Journal of Banking & Finance 72 (2016) S216–S232

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Multivariate moments expansion density: Application of the dynamic equicorrelation model

Trino-Manuel Ñíguez [a,b], Javier Perote [c,*]

[a] Department of Economics and Quantitative Methods, Westminster Business School, University of Westminster, London, UK
[b] Bank of Spain, Madrid, Spain
[c] Department of Economics and IME, University of Salamanca, Salamanca, Spain

ARTICLE INFO

Article history:
Received 25 July 2014
Accepted 28 December 2015
Available online 18 February 2016

JEL classification:
C16
G1

Keywords:
Density forecasting
Dynamic equicorrelation
Gram–Charlier series
Multivariate GARCH
Semi-nonparametric methods

ABSTRACT

In this study, we propose a new semi-nonparametric (SNP) density model for describing the density of portfolio returns. This distribution, which we refer to as the multivariate moments expansion (MME), admits any non-Gaussian (multivariate) distribution as its basis because it is specified directly in terms of the basis density's moments. To obtain the expansion of the Gaussian density, the MME is a reformulation of the multivariate Gram–Charlier (MGC), but the MME is much simpler and tractable than the MGC when positive transformations are used to produce well-defined densities. As an empirical application, we extend the dynamic conditional equicorrelation (DECO) model to an SNP framework using the MME. The resulting model is parameterized in a feasible manner to admit two-stage consistent estimation and it represents the DECO as well as the salient non-Gaussian features of portfolio return distributions. The in- and out-of-sample performance of a MME–DECO model of a portfolio of 10 assets demonstrate that it can be a useful tool for risk management purposes.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Since the subprime crisis, financial markets have experienced significant increases in volatility and correlation. This new scenario has led to a critical need for risk models that consider the conditional second moments of portfolio returns as well as the entire portfolio distribution.

Multivariate GARCH-type (MGARCH) processes, which have been comprehensively reviewed by Bauwens et al. (2006) and Silvennoinen and Terasvirta (2009), represent a general class of models for capturing portfolio risk dynamics. Research into these models has focused mainly on the following issues: (i) the so-called "curse of dimensionality", which is an inherent feature of large portfolios; (ii) modeling the time-varying structure of correlations among assets (see Engle, 2002; Cappiello et al., 2006; Engle and Kelly, 2012; and Clements et al., 2015)[1]; (iii) the accuracy and reliability of estimation techniques,[2] and (iv) the parsimonious explanation of the dynamics in the moments of the portfolio return distribution. In the latter case, a Gaussian distribution has been assumed traditionally because it facilitates the estimation and theoretical analysis of the model. Although (Gaussian)-MGARCH models can parsimoniously capture the dynamics in the first and second moments of asset return distributions, they cannot fully consider the higher-order moments. Therefore, the Gaussian assumption may be useful for periods of economic stability, but it certainly leads to losses in forecasting accuracy during periods of higher volatility when the returns' frequency accumulates in the tails of the distribution. Thus, the ability of multivariate models to signal significant departures from normality as quickly and reliably as possible is crucial for measuring and managing risk in an efficient manner. Alternative distributions have been proposed in previous studies to address this well-known issue, as follows.

* Corresponding author at: Facultad de Economía y Empresa (Edif. FES), Campus de Miguel de Unamuno, Universidad de Salamanca, 37007 Salamanca, Spain. Tel.: +34 923 294400x3512.

E-mail address: perote@usal.es (J. Perote).

[1] In addition, see Harris and Mazibas (2013), Aslandis and Casas (2013) and Christoffersen et al. (2012) for details of semi-nonparametric, non-parametric, and copula models with time-varying correlations, respectively.

[2] Maximum likelihood techniques based on MGARCH models for asset returns yield consistent quasi-maximum likelihood estimates under normality. MGARCH with non-normally distributed errors can also be estimated consistently by maximum likelihood using the three-stage procedure described by Fan et al. (2014), (also see Fiorentini and Sentana (2007) and Sentana et al. (2008)).

http://dx.doi.org/10.1016/j.jbankfin.2015.12.012
0378-4266/© 2016 Elsevier B.V. All rights reserved.

1. Parametric multivariate non-normal distributions, e.g., skewed normal (Azzalini and Dalla Valle, 1996), Student's t (Kotz and Nadarajah, 2004), Weibull (Malevergne and Sornette, 2004), and generalized hyperbolic (Fajardo and Farias, 2010). These distributions may be good alternatives for considering skewness and kurtosis but they lack sufficient flexibility to incorporate higher-order conditional (co)-moments.
2. Copula methods for constructing an implicit multivariate probability density function (pdf) from any combination of univariate marginals; see Patton (2012) for a survey of copula models in econometrics. Although the use of copulas appears to be a highly flexible approach, its tractability for large dimensions is still a serious drawback.
3. Semi-nonparametric (SNP) densities such as the multivariate Gram–Charlier (MGC) pdfs of Del Brio et al. (2009, 2011). This approach asymptotically captures the true portfolio distribution (Phillips, 1977) but it exhibits positivity problems in practical applications, which is an issue that is usually addressed using Gallant and Nychka (1987) and Gallant and Tauchen (1989)-type (GNT) transformations. The GNT-MGC pdf is positive for all values of its parameters, but it is more difficult to implement and to handle theoretically because its moments become complex nonlinear functions of its parameters.[3]

In this study, we model the conditional[4] portfolio return distribution using a novel SNP family of multivariate densities called multivariate moments expansion (MME). Our aim is to provide a relatively simple SNP methodology, which can be applied easily to large portfolios.[5] We show that the gains in simplicity due to the use of MME significantly facilitate the statistical analysis and empirical application of the density. This simplicity is attributable to the definition of MME polynomials, which are characterized by the difference between the nth power of the random variable and the nth moment of the distribution used as the basis. These polynomials do not require orthogonality conditions to make the density integrate to one, thereby yielding tractable parameterizations when implementing GNT transformations.

In general, the MME pdf exhibits the following properties.

1. *Flexibility:* The empirical performance when fitting skewness and thick wavy tails can be improved via more accurate expansion for every dimension.[6]
2. *Simplicity:* The MME is defined in terms of very simple polynomials, which do not require orthogonality to obtain well-defined and tractable distributions.
3. *Generality:* The MME provides a natural method for expanding any multivariate distribution with finite moments up to the truncation order.
4. *Positiveness:* The MME admits GNT-type reformulations in a straightforward manner to yield more feasibly parameterized distributions than the MGC.
5. *Tractability:* When the Gaussian pdf is used as the basis, it formally admits two-stage maximum likelihood estimation (MLE) to avoid the "curse of dimensionality" when estimating the dynamic conditional correlation (DCC) (Engle, 2002) and dynamic conditional equicorrelation (DECO) (Engle and Kelly, 2012) models.
6. *Applicability:* The simple structure facilitates model implementation, thereby admitting GARCH-type processes and conditional higher-order (co)-moments. Consequently, it can be used for a wide variety of applications, which range from capturing the moment dynamics and their interactions to measuring the risk proportion that corresponds to the (co)-moment of any desired order. The former may be the basis for obtaining measures of risk contagion among markets and the latter is useful for portfolio choice theory and risk managers who are concerned with the optimal allocation of liquid wealth in risky assets.[7]

The remainder of the article is organized as follows. In Section 2, we define the MME pdf and discuss its statistical properties in relation to the MGC. In Section 3, we describe the MME–DECO model. Section 4 considers an empirical in- and out-of-sample application of the MME–DECO model to a portfolio of 10 assets. In Section 4, we summarize our conclusions. All of the proofs are provided in the Appendix A.

2. The MME distribution

In this section, we introduce the MME. Throughout the characterization of the density, we highlight theoretical advantages of the MME, such as the simplicity of the polynomial structure and the capacity to admit expansions of non-Gaussian densities. For convenience, we first define the MME density for the case of uncorrelated variables before considering the more general case of correlated variables.

2.1. MME for uncorrelated variables

The generalization of a univariate SNP density to the n-dimensional setting (Sarabia and Gómez-Déniz, 2008) may be achieved in the following two stages: (1) the multivariate pdf is defined as the product of n independent univariate SNP distributed variables, and (2) a non-diagonal correlation matrix is incorporated via a linear transformation. This appealing procedure has severe issues because the number of parameters increases dramatically with both the expansion (m) and dimension (n) orders, which affects the parameter identification process and the convergence of MLE algorithms. In this study, we tackle the curse of dimensionality in both of the aforementioned stages. In stage one, we characterize the multivariate SNP model of uncorrelated variables using a multivariate pdf of the Samarnov (1966) and Lee (1996) type, the marginals of which are univariate SNP distributions, but the number of parameters only increases linearly with n.[8] It should be noted that the family of SNP distributions constructed using this procedure are uncorrelated but not independent (with the exception of the multivariate Gaussian distribution, which is nested in our general MME specification). In stage two, we implement the two-step estimation of the DCC and DECO models: (1) the conditional variances are estimated independently and consistently, and (2) the density parameters and the DECO structure are estimated using the standardized residuals from step one. We provide more details of the

[3] See León et al. (2009) for a comprehensive analysis of GNT transformations of univariate SNP distributions. Alternatively, positivity can also be ensured by parametric constraints (Jondeau and Rockinger, 2001), where this solution does not modify the density specification but it may lead to suboptimization and underperformance.

[4] Note that although we propose the MME to model the conditional portfolio return distribution it can also be used to directly model the unconditional portfolio distribution.

[5] The financial applications in previous studies of multivariate SNP distributions (e.g., Del Brio et al., 2009, 2011) have never gone beyond the bivariate case. However, in this study, we demonstrate the applicability of our MME model to a portfolio of 10 assets.

[6] The truncation order can be set endogenously in the estimation procedure. However, if the SNP method is asymptotically valid, then a larger expansion is closer to the density of the "true" frequency function.

[7] See Eeckhoudt and Schlesinger (2006), Ebert (2013), Ebert and Wiesen (2011) and Ñíguez et al. (2015) for an overview of the literature on higher-order risk preferences.

[8] We extend Samarnov-Lee's framework to both the SNP distributions and the n-variate case.

procedures for estimating the DCC-MME and DECO-MME models in the following sections.

To facilitate comparisons, we consider two families of SNP distributions. The first family is referred to as MGC and it expands a (multivariate) Gaussian density in terms of series of its derivatives or their Hermite polynomials (HPs) counterparts.[9] By construction, the HPs constitute an orthonormal basis and this property allows us to define density functions even for truncated (finite) series. The second family called MME is defined for a general sequence of distributions used as the basis, which involve very simple polynomials that depend directly on the non-central moments of the basis distributions.

However, positivity is not held in the whole domain in any of those truncated expansions unless positivity restrictions or transformations are implemented. In particular, both the MGC and MME defined in this section incorporate GNT transformations and they are re-scaled so they can integrate up to one.[10] A remarkable characteristic of these expansions is that straightforward closed forms exist for the GNT scaling constants, which greatly simplifies the implementation of the positive transformations. To proceed with the formal definition of these two families of densities, without loss of generality, we consider the same truncation order, m, for all dimensions n.

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

$$\Pi(\mathbf{x}_t;\mathbf{\Delta}) = \frac{1}{n}\left[\prod_{i=1}^{n}\phi(x_{it})\right]\left[\sum_{i=1}^{n}\lambda_i^{-1}\left(1+\sum_{s=1}^{m}\zeta_{is}^2 H_s(x_{it})^2\right)\right], \tag{1}$$

where $\mathbf{x}_t = (x_{1t}, x_{2t}, \ldots, x_{nt})' \in \mathbb{R}^n$, $\phi(x_{it})$ denotes the $N(0,1)$ pdf, $H_s(x_{it})$ is the HP such that $d^s\phi(x_{it})/dx_{it}^s = (-1)^s\phi(x_{it})H_s(x_{it})$, $\mathbf{\Delta}$ is a $m\times n$ matrix of parameters with general elements $\{\zeta_{is}\}$, and $\{\lambda_i\}_{i=1}^{n}$ is the sequence of scaling constants that makes the density integrate up to one,

$$\lambda_i = \int\left[1+\sum_{s=1}^{m}\zeta_{is}^2 H_{is}(x_{it})^2\right]\phi(x_{it})dx_{it} = 1+\sum_{s=1}^{m}\zeta_{is}^2 s!, \quad \forall i = 1,2,\ldots,n. \tag{2}$$

The most relevant characteristics of the MGC, including the marginal pdfs, raw moments, and cumulative distribution function (cdf), were derived by Del Brio et al. (2011), who also provided a comparison of this density with other related SNP distributions.

Definition 2. Let $\{g_i(x_{it})\}_{i=1}^{n}$ be a sequence of pdfs of zero mean, unit variance, and 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 $\forall r \leqslant m$. Then, the MME pdf of $\mathbf{x}_t = (x_{1t}, x_{2t}, \ldots, x_{nt})' \in \mathbb{R}^n$ is defined as,

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

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

$$w_i = \int\left[1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right]g_i(x_{it})dx_{it} = 1+\sum_{s=1}^{m}\gamma_{is}^2\left(\mu_{i,2s}-\mu_{is}^2\right), \quad \forall i = 1,2,\ldots,n. \tag{4}$$

We consider the analysis of the statistical properties of the MME pdf in the next section and in the following, we focus on discussing the differences between the pdfs Π and F in Definitions 1 and 2, respectively.

The first difference is due to the polynomial structure of the expansions. The HPs in Π are defined in terms of the derivatives of the Gaussian pdf , whereas the polynomials in F are defined as the difference between the "empirical moment" of the distribution and its counterpart in the distribution used as the basis, $P_s(x_{it}) = x_{it}^s - \mu_{is}$. Unlike the HPs, the polynomials $P_s(x_{it})$ are not orthogonal but they are designed such that they cancel out with integration and thus the function F integrates up to one.

If $\{\varphi_i(x_{it})\}_{i=1}^{n}$ is a sequence of standard Gaussian distributions with non-central moments denoted as $\{\mu_{is}^+\}_{i=1}^{n}$, then $H_s(x_{it})$ can be re-written as a linear combination of $P_s^+(x_{it}^s) = x_{it}^s - \mu_{is}^+$, as follows:

$$\begin{aligned} H_s(x_{it}) &= s!\sum_{k=0}^{[s/2]}\frac{(-1)^k}{k!(s-2k)!2^k}x_{it}^{s-2k} = s!\sum_{k=0}^{[s/2-1]}\frac{(-1)^k}{k!(s-2k)!2^k}\left[x_{it}^{s-2k}-\mu_{s-2k}^+\right] \\ &= \sum_{k=0}^{[s/2-1]}c_k P_{s-2k}^+(x_{it}), \end{aligned} \tag{5}$$

where $c_k = (-1)^k s!/k!(s-2k)!2^k$. Therefore, if the standard Gaussian is used as the basis (we refer to F in this case as F^+, i.e., when $g_i(x_{it}) = \phi(x_{it})$ and $\mu_{is} = \mu_{is}^+$ $\forall i = 1,2,\ldots,n$) and GNT transformations are not implemented, MGC and the MME yield exactly the same expression, and thus both are asymptotic representations of a given pdf (see Cramér, 1925). Nevertheless, if GNT transformations are applied to obtain well-defined pdfs, then the MGC and MME become different pdfs, where the former has a much simpler structure. This is a very important advantage of the MME compared with the MGC from theoretical and empirical viewpoints.

The second difference between Π and F is due to the distributions used as the basis, where Π is defined for expansions of the Gaussian pdf, whereas F is defined for any sequence of distributions $\{g_i(x_{it})\}_{i=1}^{n}$ with finite moments up to the expansion order m.[11] The large family of MME distributions represent the parameter flexibility that is inherent in SNP methods, and thus they are potential instruments for capturing the conditional distributions of portfolios. Nevertheless, whether the infinite expansions consider the true distribution, as in the expansion of Gaussian densities, is still an open question.[12] Of particular interest is the expansion of the Student's t distribution, which may also be developed in terms of the derivatives of its pdf (see Mauleón and Perote, 2000, for the univariate case), but it is obtained trivially using the MME as stated in Eq. (6):

[9] HPs are based on Edgeworth and Gram–Charlier (GC) series (see Kendall and Stuart, 1977). Recently, Withers and Nadarajah (2014) found a dual relation for these expansions in a multivariate framework using Bell polynomials.

[10] We use the simplest GNT-type transformation, i.e., squaring every polynomial in the SNP expansion. Ñíguez and Perote (2012) showed that this transformation is as accurate as other types of GNT transformations. However, it yields symmetric pdfs unless the distribution employed as the basis of the expansion is skewed, and thus it is appropriate when the skewness of the data is not severe.

[11] It should be noted that it is also feasible to expand other continuous and differentiable pdfs in terms of their own derivatives. In particular, for the Poisson, Gamma, or Beta distributions, these SNP pdfs are the so-called Gram–Charlier Type B, Laguerre, and Jacobi expansions, respectively (e.g., see Abramowitz and Stegun, 1972). Nevertheless, their empirical applications and their extensions to the multivariate framework are uncommon, probably due to tractability reasons.

[12] The asymptotic properties of the MGC rely on the underlying Taylor expansion (e.g., see Blinnikov and Moessner, 1998 for the Taylor expansion of the GC characteristic function). However, the MME does not necessarily possess this property for any arbitrary combination of non-Gaussian basis distributions. Therefore, searching for an appropriate characterization of the space of the basis pdfs that retain the validity of the resulting MME as an asymptotic approximation of a given frequency function is worthwhile.

$$F^*(\mathbf{x}_t;\xi,v)=\frac{1}{n}\left[\prod_{i=1}^{n}t_v(x_{it})\right]\left[\sum_{i=1}^{n}\psi_i^{-1}\left(1+\sum_{s=1}^{m}\xi_{is}^2\left(x_{it}^s-\mu_{is}^*\right)^2\right)\right], \quad (6)$$

where $t_v(x_{it})$ represents a standard Student's t with v degrees of freedom; μ_{is}^* its sth non-central moment, i.e. $\mu_{is}^*=\mu_{is}^+(v-2)^{s/2-1}/(v-s)(v-s-2)(v-s-4)\cdots(v-4)$ $\forall s$ even (0 $\forall s$ odd); ξ is a $m\times n$ matrix of parameters with general elements $\{\xi_{is}\}$; and $\psi_i=1+\sum_{s=1is}^{m}\xi_{is}^2(\mu_{i,2s}^*-\mu_{is}^{*2})$. It should be noted that the MME expansion using the Gaussian density as the basis is a particular case of the MME of the Student's t because the latter converges to the former as v tends to infinity (since $t_v(x_{it})\underset{v\to\infty}{\to}\phi(x_{it})$ and $\mu_{is}^*\underset{v\to\infty}{\to}\mu_{is}^+$). Therefore, if GNT-type reformulations are not implemented, the MME with the Student's t as the basis generalizes the MGC, and thus it also represents a valid asymptotic method for approximating any frequency function. Furthermore, it appears that by expanding a Student's t, we may account for extreme events with shorter expansions than when using a Gaussian density as the basis. However, the MME with Student's t as the basis requires $v>m$ so the Student's t moments are well defined. These analyses, as well as those of the MME for other non-Gaussian distributions, are outside the scope of the present study, but they illustrate how the general formulation of the MME represents a very interesting avenue for future research.

These arguments demonstrate that the MME is a simple, general, flexible, and tractable SNP method for parsimoniously approximating pdfs such as those of portfolio returns.

2.1.1. MME properties

In this section, we present the main properties of the MME, i.e., up-to-one integration, marginals, (non-central) moments, cdf, and related copula density. All of the proofs are provided in the Appendix A.

1. The MME is a well defined pdf because it is positive (due to the GNT-type of transformation) and it integrates up to one (Proof 1 in the Appendix).
2. MME marginal pdfs are combinations of a Gaussian and univariate GC pdfs (Proof 2 in the Appendix).

$$f_i(x_{it})=g_i(x_{it})\left[\frac{n-1}{n}+\frac{1}{nw_i}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\right]. \quad (7)$$

3. MME non-central moments are functions of the squared density parameters (Proof 3 in the Appendix).

$$E[x_{it}^r]=\left[\frac{n-1}{n}+\frac{1}{nw_i}\right]\mu_{ir}+\frac{1}{nw_i}\sum_{s=1}^{m}\gamma_{is}^2\left[\mu_{i,2s+r}+\mu_{is}(\mu_{is}\mu_{ir}-2\mu_{i,s+r})\right],\quad \forall r\in\mathbb{Z}. \quad (8)$$

4. The MME cdf can be obtained from the univariate Gaussian and moments expansion (ME) cdfs (Proof 4 in the Appendix).

$$H(\mathbf{x}_t)=\Pr[x_{1t}\leqslant\bar{x}_1,\ldots,x_{nt}\leqslant\bar{x}_n]=\frac{1}{n}\sum_{i=1}^{n}h_i(\bar{x}_i)\left[\prod_{j=1,\,j\neq i}^{n}\int_{-\infty}^{\bar{x}_j}g_j(x_{jt})dx_{jt}\right], \quad (9)$$

where $h_i(\bar{x}_i)$ denotes the cdf of the corresponding univariate ME distribution evaluated at $\bar{x}_i$. For the case of the ME of a Gaussian pdf, this cdf is given by Eqs. (10) and (11)

$$\begin{aligned}h_i(\bar{x}_i)&=\Pr[x_i\leqslant\bar{x}_i]=\int_{-\infty}^{\bar{x}_i}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\phi(x_{it})dx_{it}\\&=\int_{-\infty}^{\bar{x}_i}\phi(x_{it})dx_{it}+w_i^{-1}\phi(\bar{x}_i)\sum_{s=1}^{n}\gamma_s^2\left[2\mu_s^+T_{x_{it}}(s)-T_{x_{it}}(2s)\right],\end{aligned} \quad (10)$$

$$T_{\bar{x}_i}(s)=\bar{x}_i^{s-1}+(s-1)\bar{x}_i^{s-3}+(s-1)(s-3)\bar{x}_i^{s-5}+\cdots+\mu_s^+\bar{x}_i^{\theta}, \quad (11)$$

where $\theta=0$ for s odd and $\theta=1$ for s even (see Ñíguez and Perote, 2014, for more details).

5. The MME copula density can be obtained as (Proof 5 in the Appendix A)[13]:

$$c(h_1(x_1),h_2(x_2)\cdots h_n(x_n))=\frac{\frac{1}{n}\left[\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\right]}{\prod_{i=1}^{n}\left[\frac{n-1}{n}+\frac{1}{nw_i}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\right]}. \quad (12)$$

2.2. MME for correlated variables

We have analyzed the case of the MME conditional distribution for uncorrelated variables, but portfolio returns exhibit time-varying correlations, which is a feature that can be incorporated directly into the MME density (Definition 3) by considering a linear transformation of the following type:

$$\mathbf{u}_t=\mathbf{\Sigma}_t^{1/2}\mathbf{x}_t=\mathbf{D}_t\mathbf{R}_t^{1/2}\mathbf{x}_t, \quad (13)$$

where the (positive definite) variance–covariance matrix, $\mathbf{\Sigma}_t=\mathbf{\Sigma}_t^{1/2}\mathbf{\Sigma}_t^{1/2}=\mathbf{D}_t\mathbf{R}_t^{1/2}\mathbf{R}_t^{1/2}\mathbf{D}_t=\mathbf{D}_t\mathbf{R}_t\mathbf{D}_t$, has been decomposed in the diagonal matrix of the conditional deviations, $\mathbf{D}_t=diag\{\sigma_{1t},\ldots,\sigma_{nt}\}$, and the correlation matrix, $\mathbf{R}_t$. Then, the MME density for time-varying correlations can be defined as follows.

Definition 3. Let $\mathbf{u}_t=(u_{1t},u_{2t},\ldots,u_{nt})'\in\mathbb{R}^n$ be a $\mathbf{0}$ mean random vector with multivariate pdf $G(\mathbf{u}_t;\mathbf{\Sigma}_t,\boldsymbol{\theta})$, where $\mathbf{\Sigma}_t$ denotes the conditional variance–covariance matrix, and $\boldsymbol{\theta}$ is a vector that contains the density parameters not included in $\mathbf{\Sigma}_t$. Let $g_i(u_{it})$ denote the ith marginal density of $G(\mathbf{u}_t;\mathbf{\Sigma}_t,\boldsymbol{\theta})$ and $\mu_{ir}<\infty$, $r\leqslant m,\forall i=1,2,\ldots,n$ is its corresponding rth order non-central moment. The MME pdf of $\mathbf{u}_t$ is defined as,

$$F(\mathbf{u}_t;\boldsymbol{\gamma},\mathbf{\Sigma}_t,\boldsymbol{\theta})=\frac{1}{n}G(\mathbf{u}_t;\mathbf{\Sigma}_t,\boldsymbol{\theta})\sum_{i=1}^{n}w_i^{-1}\left[1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right], \quad (14)$$

where $\boldsymbol{\gamma}$ is a $m\times n$ matrix of parameters with general elements $\{\gamma_{is}\}$, x_{it} is the corresponding component of the inverse transformation in Eq. (13), and w_i is the scaling constant in Eq. (4).

The MME in Eq. (14) nests the multivariate distribution used as the basis, $G(\mathbf{u}_t;\mathbf{\Sigma}_t,\boldsymbol{\theta})$, but it also encompasses a wide family of SNP expansions, depending on the basis distribution. The statistical properties of the MME family of distributions in Definition 3 may be obtained in a straightforward manner from their uncorrelated counterparts in Definition 2 and by considering the transformation

[13] An analysis of correspondence between the MME copula and known copula functions can be obtained in a Monte Carlo simulation according to the following three steps: (1) MME marginals are simulated, (2) known copula models are fitted to the simulated marginals, and (3) a Kolmogorov–Smirnov-type test is used to find the best fitted known copula for the assumed MME. This experiment is outside the scope of the present study, but we consider that it would be worthwhile investigating in further research.

in Eq. (13). In fact, the properties described in Section 2.1.1 are still valid for the inverse transformation, $\mathbf{x}_t = \mathbf{\Sigma}_t^{-1/2}\mathbf{u}_t$, and thus no theoretical discussion is necessary at this point.

More interestingly, the MME distribution that uses Gaussian pdfs as the basis (F^+) preserves the "separability" property introduced by Engle (2002) and Engle and Sheppard (2001), where the log-likelihood function can be split into the volatility part, $L_V(\mathbf{u}_t,\boldsymbol{\alpha})$ in Eq. (15) and the log-likelihood of the standardized variables $\boldsymbol{\varepsilon}_t = \mathbf{D}_t^{-1}\mathbf{u}_t$, which has the MME specification $L_{F^+}(\boldsymbol{\varepsilon}_t;\boldsymbol{\rho},\gamma)$ in Eq. (16),

$$\begin{aligned} L_V(\mathbf{u}_t;\boldsymbol{\alpha}) \quad &= -\tfrac{1}{2}\sum_{i=1}^{n}\left[T\log(2\pi) + \sum_{t=1}^{T}\left(\ln(\sigma_{it}^2) + \tfrac{u_{it}^2}{\sigma_{it}^2}\right)\right] \\ &= -\tfrac{1}{2}\sum_{i=1}^{n}\left[T\log(2\pi) + L_{V_i}(u_{it};\boldsymbol{\alpha}_i)\right], \end{aligned} \tag{15}$$

$$\begin{aligned} L_{F^+}(\boldsymbol{\varepsilon}_t;\boldsymbol{\rho},\gamma) = -\frac{1}{2}\sum_{t=1}^{T}\Bigg\{ &\ln|\mathbf{R}_t| + \boldsymbol{\varepsilon}_t'\mathbf{R}_t^{-1}\boldsymbol{\varepsilon}_t \\ &- 2\ln\left[\sum_{i=1}^{n} w_i^{-1}\left(1 + \sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s - \mu_{is}^+\right)^2\right)\right]\Bigg\}, \end{aligned} \tag{16}$$

where $\boldsymbol{\alpha}$, $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ denote the vectors that contain the parameters of the conditional variances and correlations and those associated with the MME, respectively (see Proof 6 in the Appendix).

This latter property allows the implementation of two-step MLE to MME–DCC, as follows. In the first step, the parameters of the conditional variances are estimated consistently by independent quasi-MLE (QMLE). In the second step, the dynamic correlation and the MME weighting parameters are estimated jointly by maximizing the log-likelihood of the standardized residuals obtained in the first step, i.e., $L_{F^+}(\boldsymbol{\varepsilon}_t,\widehat{\boldsymbol{\alpha}},\boldsymbol{\rho},\gamma)$, where $\widehat{\boldsymbol{\alpha}} = \arg\max\{L_V(\mathbf{u}_t,\boldsymbol{\alpha})\}$. This simplified estimation procedure was theoretically valid only under Gaussianity until Del Brio et al. (2011) extended it to the MGC by arguing that the second step is consistent under misspecification provided that the MGC is a valid asymptotic expansion. In the present study, we go a step further by extending it to densities of the MME type.

Furthermore, Engle and Kelly (2012) proved that under the conditions given by White (1994, Theorem 6.11), if DCC is consistent, then DECO will also yield consistent estimations, even under the misspecification of the correlation matrix. Therefore, the two-step MLE of the MME–DECO model is consistent and it represents a simplified specification of the correlation structure for the portfolio dynamics. In the following section, we describe the MME–DECO model in relation to the explicit transformation (13) linked to the DECO.

3. The MME–DECO model

Let $\mathbf{r}_t \in \mathbb{R}^n$ in Eq. (17) be a random vector of portfolio returns with a conditional distribution on the information set $\mathbf{\Omega}_{t-1}$ as an MME with a Gaussian pdf as the basis ($F = F^+$ in Definition 3, where $g_i(x_{it}) = \phi(x_{it})$ and $\mu_{is} = \mu_{is}^+ \ \forall i = 1,2,\ldots,n$) (Eq. (19)). In the following, we consider this particular case of the MME distribution. Let the portfolio's conditional mean $E_t(\mathbf{r}_t) = E(\mathbf{r}_t|\mathbf{\Omega}_{t-1}) = \boldsymbol{\mu}_t(\boldsymbol{\phi})$ and variance $V_t(\mathbf{r}_t) = E(\mathbf{u}_t\mathbf{u}_t'|\mathbf{\Omega}_{t-1}) = \mathbf{\Sigma}_t(\boldsymbol{\alpha},\boldsymbol{\rho})$ be modeled as the MGARCH process in Eqs. (18), (20), and (21), where $\boldsymbol{\phi}, \boldsymbol{\alpha}, \boldsymbol{\rho}$, and $\boldsymbol{\gamma}$ are the vector/matrices including the conditional mean, variance, correlation, and density parameters, respectively, and $\circ$ is the Hadamard product computed via element-by-element multiplication. We assume the DECO process for modeling correlations, which states a very simple positive definite variance–covariance matrix that preserves the time-varying nature of correlations. For the sake of clarity, we denote $\mathbf{R}_t^{DECO}$ and $\mathbf{R}_t^{DCC}$ as the DECO and DCC correlation matrices, respectively. Therefore, the correlation matrix for the standardized returns (Eq. (22)) is that of Eq. (23). $\mathbf{I}_n$ is the identity matrix of order n and $\mathbf{J}_n$ is an $n \times n$ matrix of ones, but imposing the DCC-type conditional correlations in Eqs. 24,25. $\widetilde{\mathbf{Q}}_t$ replaces the off-diagonal elements of $\mathbf{Q}_t$ with zeros, although it keeps its main diagonal, and $\overline{\mathbf{Q}}$ is the unconditional covariance matrix of the standardized residuals. Moreover, DECO sets ρ_t equal to the average pairwise DCC correlations as in Eq. (26), where $q_{ij,t}$ is the ith row and jth column element of $\mathbf{Q}_t$.

$$\mathbf{r}_t = \boldsymbol{\mu}_t(\boldsymbol{\phi}) + \mathbf{u}_t, \tag{17}$$
$$\boldsymbol{\mu}_t(\boldsymbol{\phi}) = \boldsymbol{\phi}_0 + \boldsymbol{\phi}_1'\mathbf{r}_{t-1}, \tag{18}$$
$$\mathbf{u}_t|\Omega_{t-1} \sim F^+(\mathbf{0},\mathbf{\Sigma}_t(\boldsymbol{\alpha},\boldsymbol{\rho}),\gamma), \tag{19}$$
$$\mathbf{\Sigma}_t(\boldsymbol{\alpha},\boldsymbol{\rho}) = \mathbf{D}_t(\boldsymbol{\alpha})\mathbf{R}_t^{DECO}(\boldsymbol{\rho})\mathbf{D}_t(\boldsymbol{\alpha}), \tag{20}$$
$$\mathbf{D}_t^2 = diag\{\alpha_{i0}\} + diag\{\alpha_{i1}\} \circ \mathbf{u}_{t-1}\mathbf{u}_{t-1}' + diag\{\alpha_{i2}\} \circ \mathbf{D}_{t-1}^2, \tag{21}$$

$$\boldsymbol{\varepsilon}_t = \mathbf{D}_t^{-1}\mathbf{u}_t, \tag{22}$$
$$\mathbf{R}_t^{DECO}(\boldsymbol{\rho}) = (1-\rho_t)\mathbf{I}_n + \rho_t\mathbf{J}_n, \tag{23}$$
$$\mathbf{R}_t^{DCC}(\boldsymbol{\rho}) = \widetilde{\mathbf{Q}}_{t-1}^{-1/2}\mathbf{Q}_t\widetilde{\mathbf{Q}}_{t-1}^{-1/2}, \tag{24}$$
$$\mathbf{Q}_t = \overline{\mathbf{Q}}(1-\delta_1-\delta_2) + \delta_1\widetilde{\mathbf{Q}}_{t-1}^{1/2}\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}'\widetilde{\mathbf{Q}}_{t-1}^{1/2} + \delta_2\mathbf{Q}_{t-1}, \tag{25}$$
$$\rho_t = \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j>i}^{n}\frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}, \tag{26}$$
$$\rho_t \in \left(\frac{-1}{n-1}, 1\right), \tag{27}$$
$$\delta_1 > 0, \delta_2 > 0, \delta_1 + \delta_2 < 1, \tag{28}$$
$$|\phi_{i1}| < 1, \alpha_{ij} > 0, \forall j = 0,1,2 \text{ and } \alpha_{i1} + \alpha_{i2} < 1, \quad \forall i = 1,\ldots,n. \tag{29}$$

The DECO-MME model in Eqs. (17)–(29) nests the DECO model for $\boldsymbol{\gamma} = \mathbf{0}$. Furthermore, the MGC–DECO model may be defined as well as the DECO-MME, but we replace Eq. (19) with $\mathbf{u}_t|\Omega_{t-1} \sim \Pi(\mathbf{0},\mathbf{\Sigma}_t(\boldsymbol{\alpha},\boldsymbol{\rho}),\gamma)$. All of these DECO models allow for time-varying correlations with a very simple dynamic structure (which depends only on two parameters, δ_1 and δ_2), but they preserve the same correlation in every period for all assets.[14] Furthermore, if $\overline{\mathbf{Q}}$ is a positive definite matrix and under the conditions in Eqs. (27) and (28), the (MGC-, MME-) DECO correlation matrix is positive definite and mean reverting.

The main problem of the (Q)MLE for the MME–DECO (as well as the MGC–DECO) model is that it needs an explicit transformation $\boldsymbol{\varepsilon}_t = \mathbf{R}_t^{1/2}\mathbf{x}_t$ (and its inverse function) such that $E_t[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \mathbf{R}_t^{1/2}\mathbf{R}_t^{1/2} = \mathbf{R}_t = (1-\rho_t)\mathbf{I}_n + \rho_t\mathbf{J}_n$. Hence, if $\mathbf{x}_t$ is distributed as described in Eq. (3), then the pdf of $\boldsymbol{\varepsilon}_t$ will be

$$\begin{aligned} F^+(\boldsymbol{\varepsilon}_t;\boldsymbol{\rho},\gamma) = \frac{(2\pi)^{-n/2}}{n}\left|\mathbf{R}_t^{-1/2}\right|\exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}_t'\mathbf{R}_t^{-1}\boldsymbol{\varepsilon}_t\right\} \\ \times\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}^+\right)^2\right). \end{aligned} \tag{30}$$

Given the constraint in Eq. (27), it is well known that the inverse of the DECO's correlation matrix and its determinant are $\mathbf{R}_t^{-1} = \theta_1\mathbf{I}_n + \theta_2\mathbf{J}_n$, where $\theta_1 = \frac{-\rho_t}{(1-\rho_t)(1-\rho_t+n\rho_t)}$ and $\theta_2 = \frac{1}{1-\rho_t}$, and $|\mathbf{R}_t| = (1-\rho_t)^{n-1}(1+(n-1)\rho_t)$, respectively. Moreover, the inverse

[14] This assumption may be loosened by considering different blocks of correlation among the portfolio assets. The block DECO (Engle and Kelly, 2012) obtained may improve the efficiency of the MLE, but at the cost of introducing more parameters and complexity into the model. The block DECO is particularly useful when stocks can be grouped according to similarity, e.g., stocks that belong to the same industry. It should also be noted that the DECO model might impose a strong restriction on the lower bound to capture negative correlations when the number of assets is excessively large.

transformation $\mathbf{x}_t = \mathbf{R}_t^{-1/2}\boldsymbol{\varepsilon}_t$ that retrieves the standardized distribution in Eq. (3) can be obtained easily through the following argument:

$$\begin{aligned}\mathbf{x}_t'\mathbf{x}_t &= \boldsymbol{\varepsilon}_t'\mathbf{R}_t^{-1/2}\mathbf{R}_t^{-1/2}\boldsymbol{\varepsilon}_t = \boldsymbol{\varepsilon}_t'\mathbf{R}_t^{-1}\boldsymbol{\varepsilon}_t = \theta_1\left(\sum_{i=1}^{n}\varepsilon_{it}\right)^2 + \theta_2\sum_{i=1}^{n}\varepsilon_{it}^2 \\ &= \theta_2\left[\sum_{i=1}^{n}\varepsilon_{it}^2 - \tfrac{n^2\rho_t}{1-\rho_t+n\rho_t}\bar{\varepsilon}_t^2\right] = \theta_2\sum_{i=1}^{n}(\varepsilon_{it} - c_t\bar{\varepsilon}_t)^2,\end{aligned} \tag{31}$$

where $\bar{\varepsilon}_t = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_{it}$ and c_t is the variable that satisfies $\frac{n^2\rho_t}{1-\rho_t+n\rho_t} = 2nc_t - nc_t^2$.

Then, Eqs. (32) and (33) allow the direct implementation of the MME–DECO and MGC–DECO models.[15]

$$x_{it} = \frac{1}{\sqrt{1-\rho_t}}(\varepsilon_{it} - c_t\bar{\varepsilon}_t), \tag{32}$$

$$c_t = 1 \pm \sqrt{\frac{1-\rho_t}{1-\rho_t+n\rho_t}}. \tag{33}$$

This transformation yields a correlation matrix $E_t[\mathbf{x}_t\mathbf{x}_t'] = \mathbf{I}_n$ provided that all of the off-diagonal elements in $\mathbf{R}_t$ are equal to ρ_t (see Proof 8 in the Appendix). Nevertheless, if a dynamic structure of the DCC type is assumed for the correlation matrix $\mathbf{R}_t$, then further assumptions on ρ_t must be considered to make both the DCC and DECO correlation structures compatible. In particular, DECO assumes that ρ_t equals the average of the DCC pairwise correlations. Unfortunately, if ρ_t is set according to DECO, then the transformed vector $\mathbf{x}_t$ does not necessarily equal $\mathbf{I}_n$, and thus the procedure may induce convergence problems in the MLE algorithms.[16] However, we argue that this is a minor shortcoming compared with the gains in simplicity and tractability of the model, especially for high-dimensional systems. Despite this, Engle and Kelly (2012) showed that DECO is not only simpler than DCC but it also outperforms it by attenuating any measurement errors and describing portfolio co-movements more accurately.

Similar to the MME–DCC (see the previous section), the MME–DECO model can be estimated in two steps, as follows. The first step estimates the conditional mean ($\boldsymbol{\phi}$) and variance ($\boldsymbol{\alpha}$) parameters by applying QML to the univariate series (see the log-likelihood in Eq. (15)), and the second step computes the conditional correlations ($\boldsymbol{\rho}$) and the remaining parameters in the MME model ($\boldsymbol{\gamma}$) by maximizing the MME–DECO log-likelihood in Eq. (34).

This two-step procedure works reasonably well for relatively large portfolios provided that the expansion order m is not excessively large. As is usual in multivariate SNP models, the key to driving the MLE algorithms to convergence is setting appropriate initial values. Furthermore, the estimation procedure may be simplified by conditioning the log-likelihood of the MME–DECO on the estimates obtained for the DECO. The following section provides an example, which illustrates this procedure for a 10-asset portfolio.

Fig. 1 illustrates the shapes available for the bivariate MME–DECO compared with the MGC–DECO and DECO, as well as its sensitivity to changes in the correlation and the weighting parameters, γ_{i4} and γ_{i6}. These figures show the high flexibility of the (bivariate) models for fitting any target distribution, even if it is different for each asset (dimension). This feature is particularly useful for modeling and forecasting portfolio risk measures because their accuracy is closely linked to the ability of the model to capture the tail thickness and possible multimodality of the portfolio return distributions. The plots highlight how the DECO (Fig. 1.1) is unable to represent heavy tails, whereas both the MME–DECO (Figs. 1.2–1.4) and MGC–DECO (Figs. 1.5 and 1.6) assign a higher probability to the distribution tails when the truncation order is higher and when the values of the parameters in the expansion are higher. It is interesting to note the flexibility of the SNP–DECO models, where in addition to the portfolio conditional correlations, they can also represent the salient features with respect to the normality of the asset return distributions. Moreover, the MME–DECO also starts by providing a higher density in the tails to lower values of the parameters than the MGC–DECO, which can be interpreted as the higher sensitivity of the density function to the weighting parameters.

4. Empirical application to portfolio returns

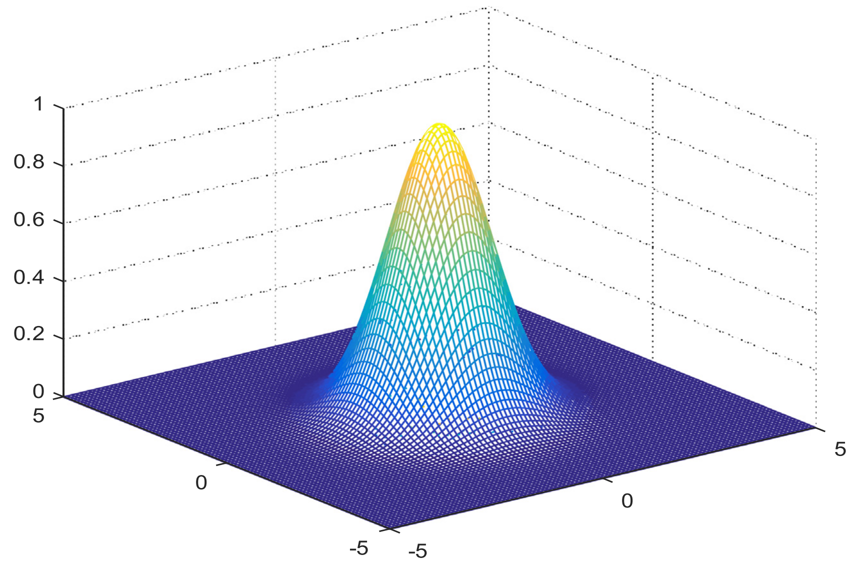
The MME represents a highly flexible and straightforward parameterization of the multivariate conditional distribution of the portfolio return, which is consistent with MGARCH models. However, the MME admits many different extensions, which depend on the focus of the study. Some of these extensions are as follows: (i) different marginals can be used to consider asymmetric correlation and tail dependence, even in an n-dimensional setting[17]; (ii) high-order conditional moments (e.g., dynamics of skewness and kurtosis) can be modeled, and the dependencies between these moments across portfolio variables can even be considered to address possible spillover effects; (iii) two- or three-step MLE (since the log-likelihood of the MME is separable) as well as the generalized method of moments can be implemented directly

$$\begin{aligned}L_{F^+}^{DECO}(\boldsymbol{\varepsilon}_t;\widehat{\boldsymbol{\alpha}},\boldsymbol{\rho},\boldsymbol{\gamma}) &= -\tfrac{1}{2}\sum_{t=1}^{T}\left\{\ln\left|\mathbf{R}_t^{DECO}\right| + \boldsymbol{\varepsilon}_t'\mathbf{R}_t^{DECO-1}\boldsymbol{\varepsilon}_t - 2\ln\left[\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2(x_{it}^s-\mu_{is}^+)^2\right)\right]\right\} \\ &= -\tfrac{1}{2}\sum_{t=1}^{T}\ln\left(\tfrac{1+(n-1\rho_t)}{(1-\rho_t)^{1-n}}\right) - \tfrac{1}{2}\sum_{t=1}^{T}\tfrac{1}{1-\rho_t}\left(\sum_{i=1}^{n}\varepsilon_{it}^2 - \tfrac{\rho_t}{1+(n-1)\rho_t}\left(\sum_{i=1}^{n}\varepsilon_{it}\right)^2\right) \\ &\quad+\sum_{t=1}^{T}\ln\left\{\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left[(1-\rho_t)^{-\frac{s}{2}}\left[\varepsilon_{it}-\bar{\varepsilon}_t-\bar{\varepsilon}_t\sqrt{\tfrac{1-\rho_t}{1-\rho_t+n\rho_t}}\right]^s-\mu_{is}^+\right]^2\right)\right\}.\end{aligned} \tag{34}$$

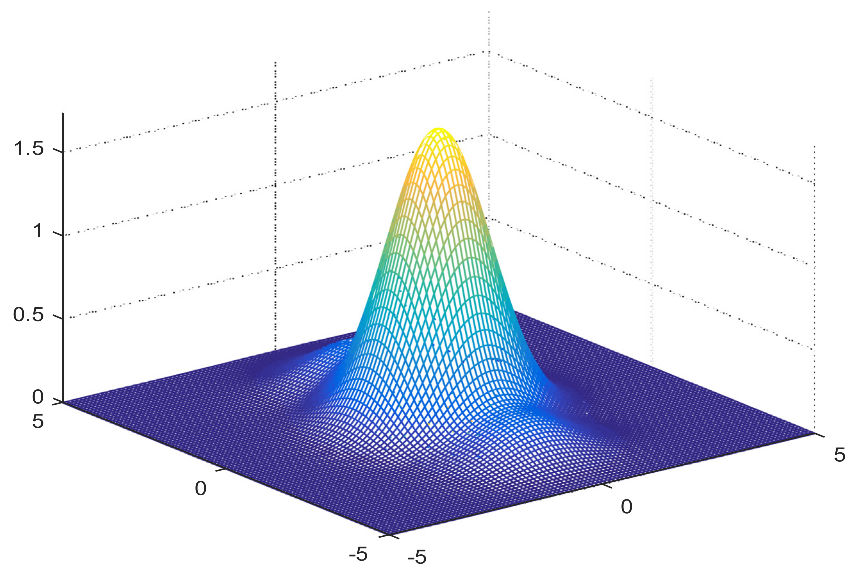
[15] This transformation allows different applications, which range from estimation to model evaluation criteria (e.g., in Section 4, we use it to evaluate multivariate cdfs using the property in Eqs. (9)–(11)). An alternative valid transformation as n tends to infinity is presented in Proof 7 in the Appendix.

[16] Note that this issue is also present in the Gaussian-DECO because this model can be obtained when the transformations in Eqs. (32) and (33) are applied to the standard Gaussian pdf $(2\pi)^{-n/2}\exp\left(-\frac{1}{2}\mathbf{x}_t'\mathbf{x}_t\right)$.

[17] The MME can be an alternative to the copula approach for modeling asymmetric dependence, particularly for large portfolios; e.g., see Ang and Chen (2002) and Patton (2006) for evidence of the asymmetric dependence of returns in stocks and exchange rates, respectively.



1.1: Gaussian DECO. $\rho_t = 0.05$.



1.2: MME-DECO. $\rho_t = 0.05$, $\gamma_{14} = 0$, $\gamma_{16} = 0.0005$, $\gamma_{24} = 0.065$, $\gamma_{16} = 0.0002$.

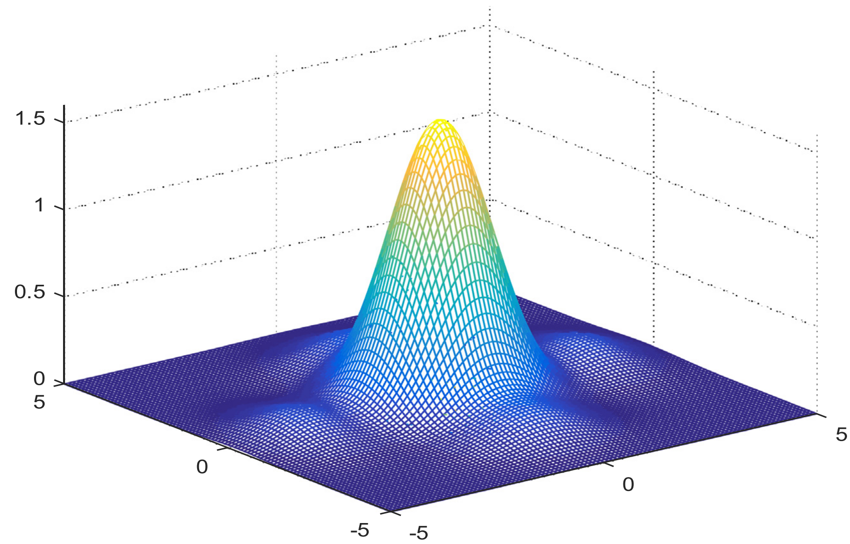
Fig. 1. Plots of bivariate Gaussian-, MGC- and MME–DECO densities.

to simplify the estimation procedure; and (iv) applications of the DCC and DECO type can also be implemented to tackle the curse of dimensionality in the variance and covariance matrix. In this study, we focused on points (iii) and (iv), and thus in this section, we investigate the in- and out-of-sample performance of MME–DECO and MGC–DECO compared with DECO based on an empirical application to portfolio returns.

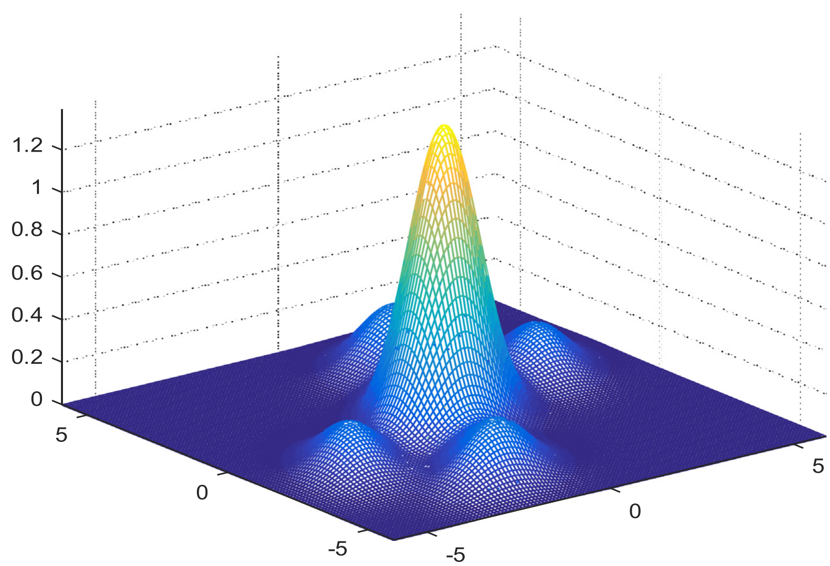
4.1. Data

The data employed in this study were (daily) percentage log returns, which were computed as $r_{it} = 100 \log (P_{it}/P_{it-1})$from series $\{P_{it}\}_{t=1}^{T}$ of daily prices for: four stock indexes (Dow Jones (DJ), Ibex 35, Nikkei 300, and Hang Seng (HS)); three individual shares (Deutsche Bank (DB), BP, and Apple); and three exchange rates (U.S. dollar to pound sterling (FX $/£), Japanese yen to pound sterling (FX ¥/£), and Swiss franc to U.S. dollar (FX Fr/$)). All of the series were sampled from February 16, 1995 to February 16, 2015 to obtain a total of 5218 observations. The data were obtained from Datastream. Table 1 reports the descriptive statistics for the series, the Jarque–Bera statistic (J–B), and its corresponding p-value (p–v) for the null of normality. The J–B test results showed that the empirical distribution of the series was not normal, where it was (mildly) skewed and highly leptokurtic. The sample correlations ranged from -0.001 to 0.64.

Fig. 2 shows plots of the return series for the full sample. The shaded areas correspond to the two periods that we used for the out-of-sample forecasting performance analysis: (1) the data per-



1.3: MME-DECO. ρ_t= 0.05, γ_{14}= 0, γ_{16}= 0.005, γ_{24}= 0.05, γ_{16}= 0.002.



1.4: MME-DECO. ρ_t= 0.4, γ_{14}= 0, γ_{16}= 0.005, γ_{18}= 0.003, γ_{24}= 0.05, γ_{26}= 0.002, γ_{28}= 0.0005

Fig. 1 (*continued*)

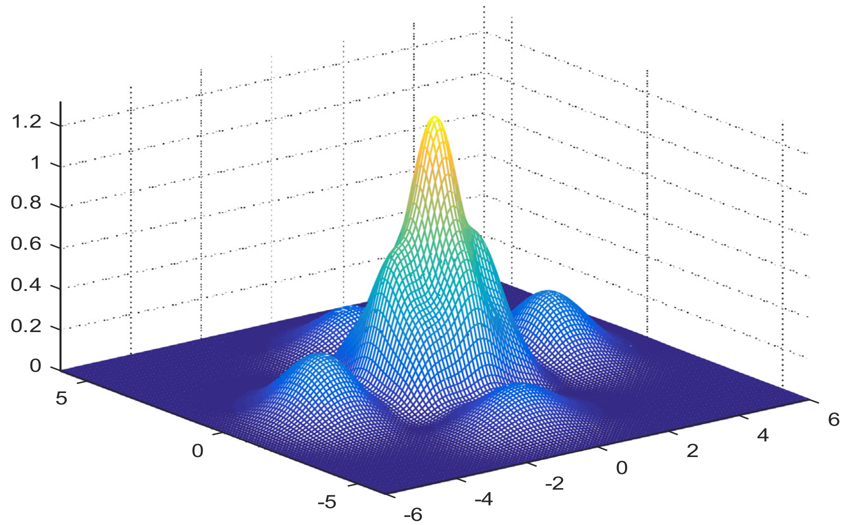
iod from March 19, 2013 to February 16, 2015 (500 observations), which exhibited relatively low volatility and it was nearer to a normal distribution; and (2) 500 observations from the recent credit crunch period (July 17, 2008 to June 16, 2010), which exhibited high volatility and a more leptokurtic distribution.

4.2. Model evaluation

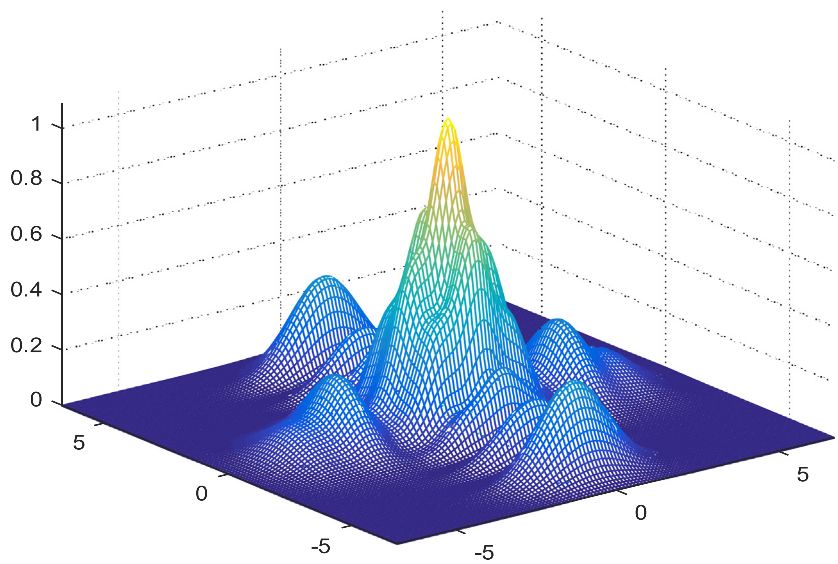
The models were estimated in two stages by MLE techniques using a rolling window with a size N equal to 4717 (3499) for the density forecast of the market calm (crisis) period. In the first stage, an AR(1) process for the conditional mean (selected according to Akaike's information criterion (AIC)) and a GARCH(1,1) process for the conditional variance were jointly estimated (under normality) independently for each asset. In the second stage, the standardized residuals from the previous stage were used to estimate the conditional equicorrelation equation in the DECO model. The DECO estimates and standardized residuals from the first and second stages were used to estimate the MGC- and MME–DECO density parameters. Finally, following Pagan (1986), a further Newton-Rapshon iteration without line search was performed based on the density and DECO process parameter estimates to ensure that the information matrix was block diagonal, thereby obtaining estimators that were asymptotically equivalent to joint QMLE.[18]

To obtain the specifications of the MGC and MME models, we proceeded by truncating the expansion at the highest-order significant parameter. The GNT-type transformation yields symmetric

[18] Given the possible limitations when estimating the MGC- and MME–DECO models, it should be mentioned that possible losses in estimation efficiency and consistency, which may be important when evaluating the in-sample performance of models, are not crucial in an out-of-sample evaluation analysis (Ruiz and Pascual, 2002). An empirical analysis of the properties of the estimator for SNP–DECO models is an area for future research.



1.5: MGC-DECO. ρ_t= 0.1, γ_{14}= 0.25, γ_{16}= 0.025, γ_{24}= 0.15, γ_{16}= 0.015.



1.6: MGC-DECO. ρ_t= 0.01, γ_{14}= 0.2, γ_{16}= 0.02, γ_{18}= 0.003, γ_{24}= 0, γ_{16}= 0.02. γ_{28}= 0.007.

Fig. 1 (*continued*)

Table 1
Descriptive statistics of daily percent returns.

	Mean	Max	Min	St. Dev.	Skew	Kurtosis	J–B stat (p-v)
Sample	2/17/1995–2/16/2015						
Observations	5217						
Dow Jones	0.03003	10.0891	−8.69495	1.11219	−0.26520	10.59539	12601.5 (0.00)
Ibex 35	0.02367	13.4836	−9.58586	1.45849	−0.01759	8.05976	5565.3 (0.00)
Nikkei 300	0.00298	12.9511	−10.2687	1.34674	−0.21636	8.79710	7345.9 (0.00)
Hang Seng	0.02131	17.2471	−14.7346	1.61924	0.09151	13.90975	25879.8 (0.00)
Deutsche Bank	−0.00203	22.3031	−18.0729	2.32904	0.18352	11.33788	15141.2 (0.00)
BP	0.01505	10.5825	−14.0368	1.67378	−0.04372	7.62108	4643.5 (0.00)
FX $/£	−0.00047	4.47445	−3.91821	0.54758	−0.03954	7.43977	4286.1 (0.00)
FX ¥/£	0.00325	6.34316	−6.94279	0.80547	−0.47496	9.74857	10096.1 (0.00)
FX Fr/$	−0.00582	8.47479	−11.4188	0.70149	−0.81471	23.92072	95717.0 (0.00)
Apple	0.08456	28.6795	−73.1231	3.00796	−2.55747	75.06267	1134522 (0.00)
Rang($\overline{\rho}$)				[-0.001,0.64]			

Notes: This table provides the descriptive statistics of the daily percent return data used in the paper. The Jarque–Bera (J–B) statistic is asymptotically distributed as a χ^2 distribution with 2 degrees of freedom under the null of normality, the *p*-value (*p*–*v*) of the test is in parenthesis next to the J–B statistic. *Rang*($\overline{\rho}$) denotes the range of the sample correlation coefficients of the ten return series.

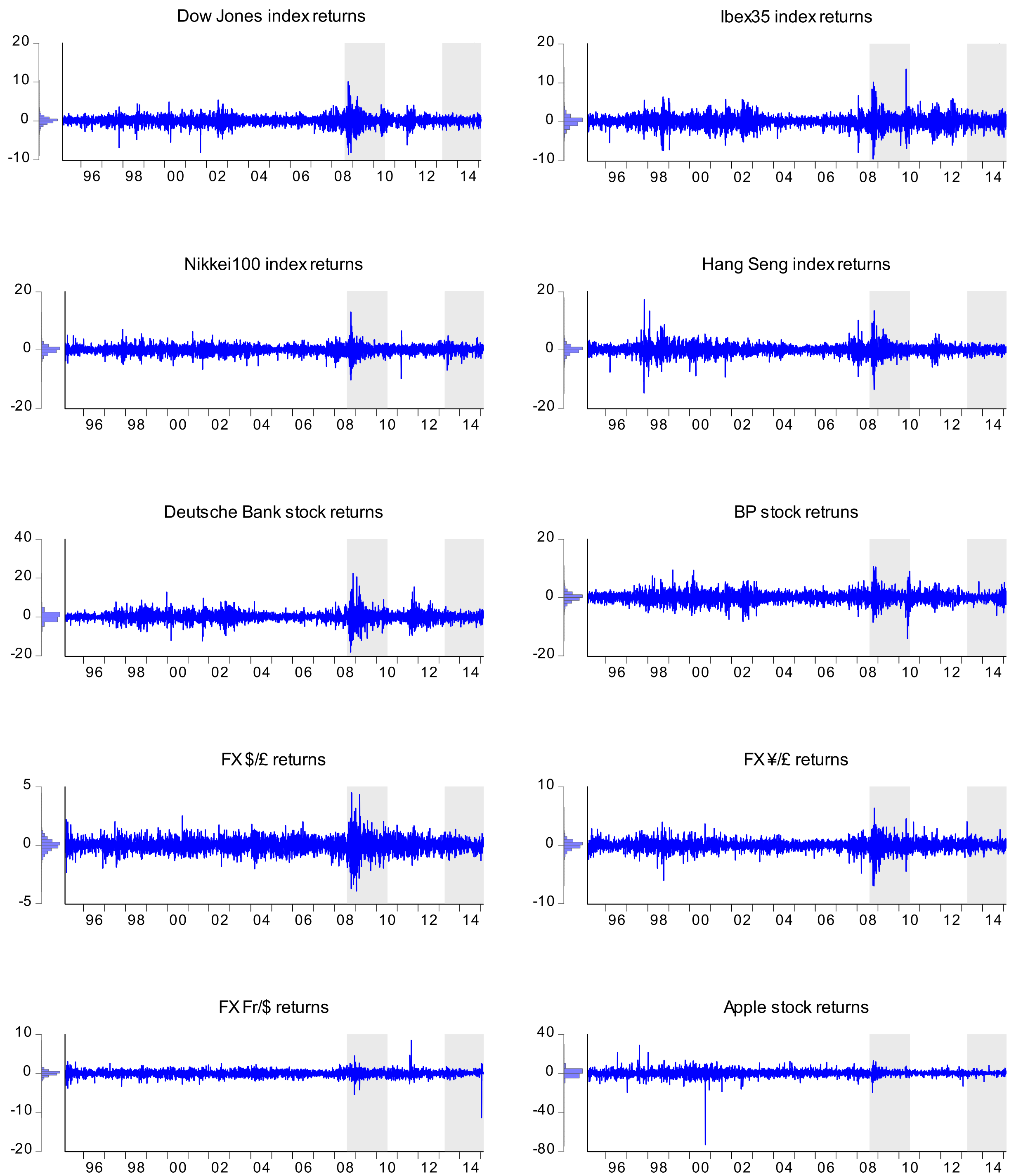


Fig. 2. Daily percent log return series. *Notes:* Plots and histograms of daily log returns of: Dow Jones, Ibex 35, Nikkei 300 and Hang Seng stock index returns; Deutsche Bank, BP and Apple stock returns; and $/£, ¥/£and Fr/$ FX returns. Full sample: 2/17/1995–2/16/2015 (5217 obs.). Out-of-sample periods (shaded areas, 500 obs.): 7/17/2008–6/16/2010 (higher volatility period) and 3/19/2013–2/16/2015 (lower volatility period).

distributions, so the odd-order parameters were omitted. The resulting SNP–DECO ι model had an sth order polynomial with weighting parameters denoted by $\gamma_{is,\iota}$ $(s = 4, 6)$; $\gamma_{i2,\iota}$ were set to zero so that the standardized SNP distributions had unit variance.[19] Estimating SNP–DECO models is not computationally demanding for lower dimensions but it may become very slow for large portfolios. In the latter case, the selection of appropriate starting values may be achieved by estimating the univariate marginals (Eq. (7)) in advance, starting with polynomials of low orders, and by considering the relations between the density parameters and their moment counterparts (Eq. (8)). The multivariate model is then estimated by starting with a low dimension, say $n = 3$. Moreover, a robustness check of the estimates is recommended, particularly when performing out-of-sample recursive estimation. We monitored the optimization by perturbing the starting values to check that the (Q)ML estimates obtained were the global optima.

Table 2 presents the estimation results obtained for the two in-sample windows considered in this study. The AR(1)-GARCH(1,1)

[19] For the sake of simplicity in notation the parameters of both the MME and MGC were denoted as $\{\gamma_{is}\}$.

Table 2
Estimation results.

	DECO	MME–DECO	MGC–DECO	DECO	MME–DECO	MGC–DECO
Sample	2/17/1995–3/18/2013 (4717 obs.)			2/17/1995–7/16/2008 (3499 obs.)		
Stage 1						
δ_1	0.0058 (6.32)			0.0102 (6.36)		
δ_2	0.9916 (680.2)			0.9825 (395.8)		
Stage 2						
$\gamma_{14,\imath}$		−0.0000 (−0.00)	0.0000 (0.00)		−0.0000 (−.00)	0.0000 (0.00)
$\gamma_{16,\imath}$		−0.0005 (−2.87)	−0.0022 (−2.83)		−0.0007 (−2.96)	−0.0027 (−2.84)
$\gamma_{24,\imath}$		−0.0000 (−0.00)	0.0000 (0.00)		−0.0000 (−.00)	−0.0000 (−0.00)
$\gamma_{26,\imath}$		0.0007 (3.80)	−0.0028 (−3.90)		0.0008 (3.91)	−0.0033 (−3.92)
$\gamma_{34,\imath}$		−0.0000 (−0.00)	0.0229 (2.09)		−0.0000 (−0.00)	−0.0000 (−0.00)
$\gamma_{36,\imath}$		0.0015 (6.06)	−0.0042 (−3.97)		0.0012 (4.66)	−0.0046 (−4.30)
$\gamma_{44,\imath}$		−0.0000 (−0.00)	0.0368 (4.71)		−0.0000 (−0.00)	0.0355 (3.70)
$\gamma_{46,\imath}$		−0.0016 (−6.85)	−0.0025 (−1.35)		−0.0017 (−6.52)	−0.0036 (−2.01)
$\gamma_{54,\imath}$		−0.0000 (−0.00)	−0.0000 (−0.00)		−0.0000 (−0.00)	−0.0000 (−0.00)
$\gamma_{56,\imath}$		−0.0011 (−4.81)	−0.0038 (−4.98)		−0.0011 (−4.51)	−0.0041 (−4.62)
$\gamma_{64,\imath}$		0.0000 (0.00)	0.0377 (4.52)		0.0000 (0.00)	0.0341 (3.28)
$\gamma_{66,\imath}$		−0.0021 (−8.91)	−0.0052 (−4.62)		−0.0017 (−5.91)	−0.0036 (−2.59)
$\gamma_{74,\imath}$		0.0684 (9.94)	0.1481 (14.1)		0.1023 (9.83)	0.2050 (11.5)
$\gamma_{76,\imath}$		0.0038 (6.95)	0.0112 (5.22)		0.0046 (5.84)	0.0168 (5.69)
$\gamma_{84,\imath}$		−0.0000 (−0.00)	0.0529 (7.08)		−0.0195 (−1.72)	−0.0681 (−8.25)
$\gamma_{86,\imath}$		−0.0026 (−11.2)	0.0063 (5.20)		0.0268 (5.47)	0.0057 (3.56)
$\gamma_{94,\imath}$		0.0852 (12.74)	0.1487 (14.1)		0.0429 (10.99)	0.0725 (9.41)
$\gamma_{96,\imath}$		−0.0032 (−5.45)	0.0127 (6.62)		0.0000 (0.00)	f−0.0000 (−0.00)
$\gamma_{10,4,\imath}$		−0.0000 (−0.00)	0.0558 (6.96)		−0.0000 (−0.00)	0.0570 (6.53)
$\gamma_{10,6,\imath}$		−0.0033 (−14.0)	0.0095 (8.42)		−0.0033 (−12.4)	0.0094 (7.10)
$AIC_\imath$	16.925	8.0853	8.1009	16.792	8.2940	8.3029
$\overline{AIC}_\imath$	16.872	8.0270	8.0457	16.563	8.0642	8.1001

Mean equation: $r_{it} = \phi_{i0} + \phi_{i1} r_{i,t-1} + u_{it},\ u_{it} = \varepsilon_{it}\sigma_{it},\quad i = 1, 2, \ldots, 10.$

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

MGC−DECO pdf: $\Pi(\mathbf{u}_t; \gamma, \Sigma_t) = \frac{1}{2}\Phi(\mathbf{u}_t; \Sigma_t)\left[\lambda_i^{-1}\left(1 + \gamma_{i4}^2 H_{i4}(x_{it})^2 + \gamma_{i6}^2 H_{i6}(x_{it})^2\right)\right]$.

MME−DECO pdf: $F^{+}(\mathbf{u}_t; \gamma, \Sigma_t) = \frac{1}{2}\Phi(\mathbf{u}_t; \Sigma_t)\left[w_i^{-1}\left(1 + \gamma_{i4}^2 P_{i4}(x_{it})^2 + \gamma_{i6}^2 P_{i6}(x_{it})^2\right)\right]$.

Notes: This table presents ML estimates of the parameters of DECO, MGC- and MME–DECO for a portfolio of 10 assets and the two in-sample periods under analysis. Heteroscedasticity-consistent standard errors are provided in parentheses next to the parameter estimates. Conditional mean and variance equations (AR(1)-GARCH(1,1)) parameter estimates are not displayed in the table to save space; these estimates take the typical values for daily stock returns. δ_1 and δ_2 denote the conditional correlation parameters, $\gamma_{is,\imath}$ $(s = 4, 6)$ denote the sth order polynomial weight parameter of the expansions in model $\imath$ (MGC- and MME–DECO). *AIC* denotes Akaike information Criterion and $\overline{AIC}$ is the average AIC for the 500 estimations performed, with a constant size rolling window, for the prediction analysis.

parameter estimates show that the data series exhibited: (1) a typical small structure in the conditional mean and (2) high persistence in the conditional variances (these estimates are not presented in the table to save space). The parameter estimates for the DECO process were both significant and they exhibited high persistence in all of the pooled asset correlations. For both the MGC- and MME–DECO models, γ_{i4} was not significant for some series, but γ_{i6} was significant for all dimensions (Bollerslev and Wooldridge (1992) robust standard errors are shown in parentheses). The estimation results were very similar for the two sample periods considered as well as for the data rolling windows used in the forecasting exercise.

In Table 2 and Fig. 3, the AIC and average AIC ($\overline{AIC}$) values indicate that the MME–DECO provided a better goodness-of-fit than the MGC–DECO, but both improved the performance of the DECO. This result was consistent across all of the rolling windows, as shown in Fig. 3. It should be noted that all of the models shared the same DECO correlation structure, so the differences among them were due only to the parameters of the expansions. Fig. 4 also provides an example of the fitted marginal densities for the FX $/£ return series based on the MME, MGC, and Gaussian pdfs compared with the data histogram. The densities yielded similar performance at the center of the distribution, but MME and MGC departed from Gaussian signaling in the presence of the significant frequency in the far left tail, thereby providing useful information when measuring the risk associated with the down-slope of the distribution.[20]

Next, we tested the performance of the models for density forecasting in the two out-of-sample periods with high and low volatility, and kurtosis (500 predictions for each period). The forecasts obtained by the models were compared according to weighted logarithmic scoring rules (Gneiting and Raftery, 2007; and Amisano and Giacomini, 2007) for the marginal and conditional portfolio return distributions (Diebold et al., 1999). A model provides better forecasting performance when its weighted average logarithmic score is lower, which is defined as:

$$\overline{\Lambda}(\widetilde{g}, x) = -N^{-1}\sum_{t=T}^{T+N-1} \varpi(x_{t+1}) \ln \widetilde{g}_t(x_{t+1}), \tag{35}$$

where $\widetilde{g}_t(x_{t+1})$ denotes the one day-ahead density forecast from model g and $\varpi(\cdot)$ is a weight function used to evaluate the performance when forecasting different regions of the return distribution.

[20] As an alternative to the SNP marginal pdfs, we checked the performance of the univariate mixture of normals (MN) as a suitable pdf to capture the frequency in the tails of the return distributions (e.g., see Alexander and Lazar, 2006). The results showed that univariate three- and two-component MN pdfs provided a good fit to the tails of the returns histogram, which were similar to the univariate ME and GC pdfs, and they were better than that of the normal. These results are available from the authors upon request.

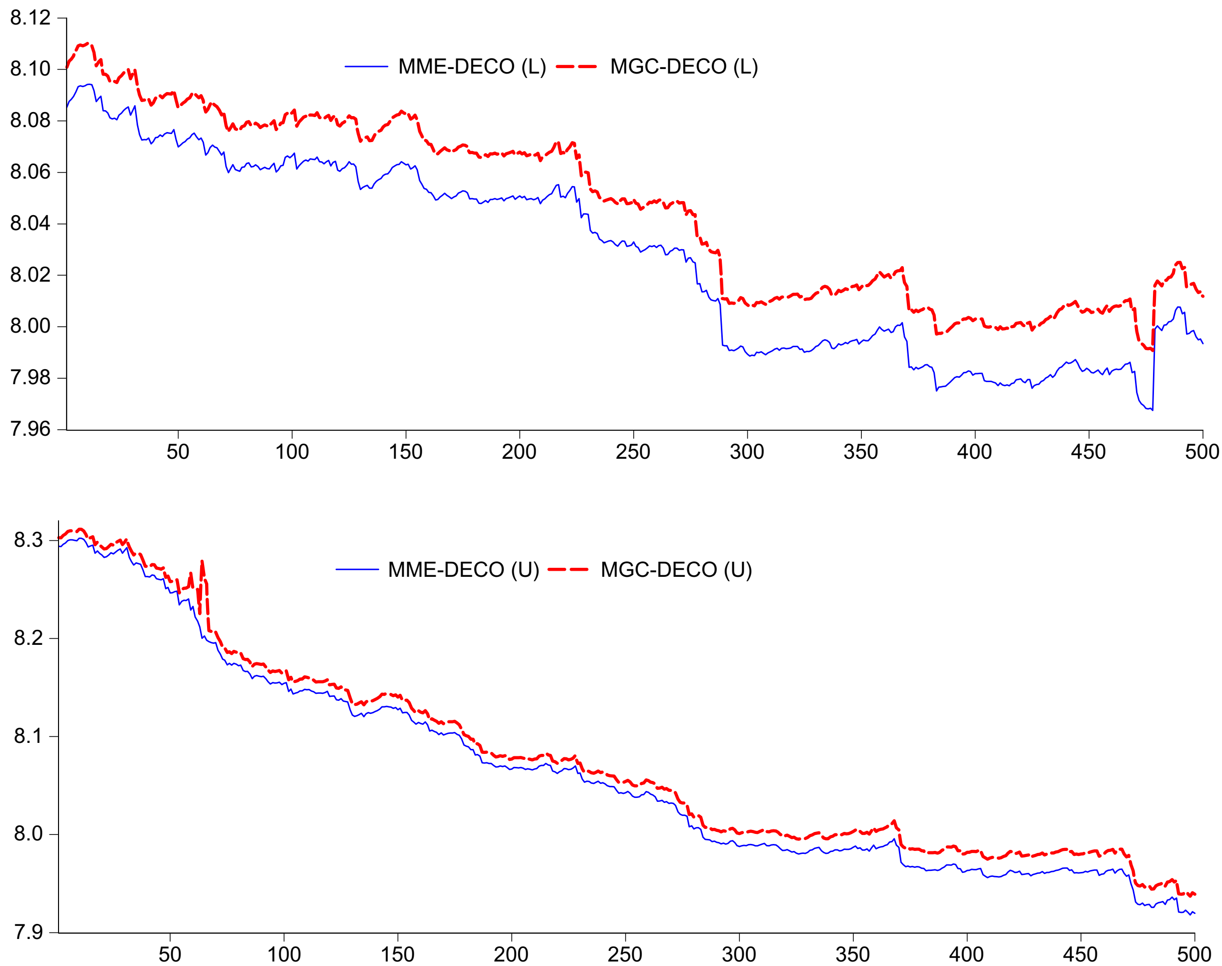


Fig. 3. Overall goodness-of-fit of MGC- and MME–DECO. *Notes:* This figure provides MGC- and MME–DECO AIC series obtained from rolling window estimation through: Low (L) volatility out-of-sample period (3/19/2013–2/16/2015), and high (U) volatility out-of-sample period (7/17/2008–6/16/2010).

The null hypothesis of equal density forecasting performance from two alternative models f and $g, H_0 : E\left(\overline{\Lambda}(\widetilde{f},x) - \overline{\Lambda}(\widetilde{g},x)\right) = 0$, was tested using the Amisano and Giacomini (2007) test. We conducted pairwise scoring rule tests using the following weights: (1) $\varpi_1 = 1$ to measure the performance for the full density, and (2) $\varpi_2 = 1 - \phi(x)/\phi(0)$ to evaluate the performance of the models when predicting the density tails.

The scoring rule test results are reported in Table 3. The main entries represent the difference between the weighted average scores using the MME–DECO with respect to the models in the columns, where the numbers in parentheses are the t-test statistics. A negative score indicates that the MME–DECO yielded more accurate forecasts than the alternative models.

The results in Table 3 Panel A showed that the MME–DECO obtained statistically significant lower scores, thereby improving the forecasts with DECO for the full, marginal, and conditional densities. These results are broadly consistent with those in Table 3 Panel B, although the differences in the scores were more pronounced when these models were compared in terms of the forecasting accuracy of the densities' tails. For the two out-of-sample periods considered in this study, we found no major differences in performance, although slightly lower differences in tail forecasting were predominant for the period of market calm compared with the period of crisis.

Compared with the MGC–DECO, the scoring rule results were better using the MME–DECO (MGC–DECO) for the period of relatively low (high) volatility and kurtosis, although the overall scoring rule differences were not significant for both of the periods and the weighting functions considered. In summary, our results show that improvements can be obtained in the accuracy of forecasting the portfolio returns density by using MME- and MGC–DECO compared with DECO.

5. Concluding remarks

Copula models and multivariate parametric leptokurtic pdfs have been used in risk management to explain the higher-order conditional (co)-moments of portfolio return series. Nevertheless, the data fits achieved using these techniques can be improved because the models are either not sufficiently flexible to consider the salient features of asset returns or, if they are, they are too complex and analytically intractable. Alternatively, the multivariate SNP approach, which is traditionally based on GC expansions, is characterized by its flexibility in fitting any target density while maintaining a reasonable analytical tractability, which is facilitated by using series based on orthogonal polynomials (e.g., HPs).

In this study, we proposed a novel SNP family of multivariate distributions (MME), which maintains the typical flexibility and generality of SNP methods, but it has both theoretical and empirical advantages due to its simplicity. The MME and the MGC differ

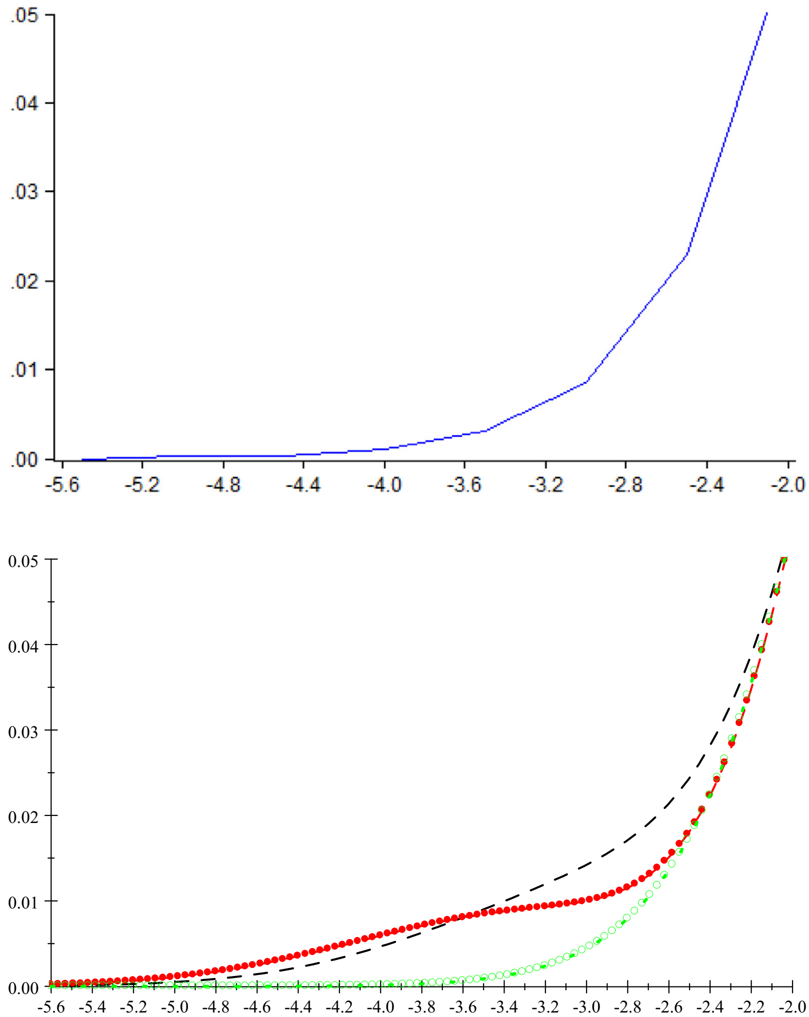


Fig. 4. Fitted and empirical marginal densities (left tail) Notes: Fitted marginal and empirical pdfs for FX $/ £standardized returns. Histogram (Solid blue), MME (Dash black), MGC (DotDash red) and Gaussian (DotDotDash green). Sample 2/17/1995–3/18/2013, 4717 obs., parameter estimates are provided in Table 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Weighted scoring rule tests.

	DECO	MGC–DECO	DECO	MGC–DECO
Out-of-sample	7/17/2008–6/16/2010		3/19/2013–2/16/2015	
Panel A: Full density (ϖ_1)				
Marginal DJ	−0.00114 (−4.19)	0.00055 (0.70)	−0.00252 (−1.74)	0.00053 (0.94)
Marginal Ibex	−0.00160 (−2.31)	0.00136 (1.48)	−0.00251 (−1.84)	0.00064 (1.10)
Marginal Nikkei	−0.00398 (−3.65)	0.00597 (1.07)	−0.01413 (−1.49)	0.00067 (0.31)
Marginal HS	−0.00466 (−2.62)	0.00000 (0.12)	−0.00692 (−3.22)	0.00044 (4.04)
Marginal DB	−0.01656 (−1.61)	0.00891 (1.92)	−0.00495 (−2.55)	−0.00072 (−3.09)
Marginal BP	−0.01940 (−1.95)	−0.00043 (0.63)	−0.02418 (−1.67)	0.00151 (0.70)
Marginal Apple	−0.08191 (−5.67)	0.00848 (0.63)	−0.04328 (−5.24)	−0.00778 (−5.88)
Marginal $/£	−0.04316 (−2.64)	0.00256 (1.46)	−0.01117 (−2.93)	−0.00013 (−1.31)
Marginal ¥/£	−0.03814 (−2.68)	−0.00244 (−1.06)	−0.07492 (−1.75)	0.00109 (−0.18)
Marginal Fr/$	−0.01637 (−1.70)	0.00140 (1.34)	−0.02954 (−2.21)	0.00086 (0.53)
Conditional DJ	−0.19857 (−6.52)	0.00555 (0.76)	−0.20000 (−4.08)	−0.00487 (−0.71)
Conditional Ibex	−0.19810 (−6.50)	0.00474 (0.68)	−0.20001 (−4.09)	−0.00498 (−0.72)
Conditional Nikkei	−0.19573 (−6.44)	0.00012 (0.02)	−0.18839 (−3.91)	−0.00500 (−0.77)
Conditional HS	−0.19505 (−6.42)	0.00607 (0.79)	−0.19560 (−4.01)	−0.00477 (−0.70)
Conditional DB	−0.18314 (−6.40)	−0.00281 (−0.52)	−0.19757 (−4.03)	−0.00361 (−0.52)
Conditional BP	−0.18031 (−6.38)	0.00576 (0.75)	−0.17834 (−3.81)	−0.00584 (−0.90)
Conditional Apple	−0.11779 (−4.99)	0.01459 (2.02)	−0.15924 (−3.30)	−0.00345 (−0.51)
Conditional $/£	−0.15655 (−6.38)	0.00354 (0.54)	−0.19134 (−3.93)	−0.00420 (−0.61)
Conditional ¥/£	−0.16156 (−5.98)	0.00854 (1.25)	−0.12760 (−5.31)	−0.00324 (−0.87)
Conditional Fr/$	−0.18333 (−6.36)	0.00470 (0.64)	−0.17298 (−3.67)	−0.00520 (−0.78)

Table 3 (*continued*)

	DECO	MGC–DECO	DECO	MGC–DECO
Panel B: Density tails (ϖ_2)				
Marginal DJ	−0.00270 (−4.14)	0.00109 (0.64)	−0.00619 (−1.71)	0.00135 (0.95)
Marginal Ibex	−0.00385 (−2.22)	0.00300 (1.48)	−0.00617 (−1.80)	0.00160 (1.10)
Marginal Nikkei	−0.00948 (−3.51)	0.01474 (1.06)	−0.03488 (−1.47)	0.00182 (0.35)
Marginal HS	−0.01082 (−2.44)	−0.00000 (−0.22)	−0.01661 (−3.10)	0.00100 (3.71)
Marginal DB	−0.04114 (−1.59)	0.02142 (1.88)	−0.01198 (−2.47)	−0.01740 (−2.98)
Marginal BP	−0.04760 (−1.91)	0.00108 (0.64)	−0.05930 (−1.63)	0.00386 (0.71)
Marginal Apple	−0.18965 (−5.33)	−0.01865 (−5.21)	−0.09846 (−4.90)	−0.01740 (−5.57)
Marginal $/£	−0.10502 (−2.56)	0.00607 (1.44)	−0.02619 (−2.76)	−0.00039 (−1.64)
Marginal ¥/£	−0.08953 (−2.52)	−0.00492 (−0.86)	−0.17853 (−1.67)	−0.00063 (0.04)
Marginal Fr/$	−0.03849 (−1.59)	0.00316 (1.26)	−0.07149 (−2.14)	−0.00227 (0.55)
Conditional DJ	−0.37188 (−6.38)	0.01046 (0.81)	−0.33953 (−4.34)	−0.01118 (−0.52)
Conditional Ibex	−0.37171 (−6.21)	0.01349 (0.90)	−0.36933 (−4.20)	−0.00814 (−0.65)
Conditional Nikkei	−0.35250 (−6.21)	0.00146 (0.15)	−0.33671 (−3.69)	−0.00812 (−0.66)
Conditional HS	−0.35685 (−6.24)	0.01653 (0.98)	−0.35579 (−3.95)	−0.00781 (−0.61)
Conditional DB	−0.34166 (−5.96)	−0.00217 (−0.20)	−0.35637 (−3.98)	−0.00563 (−0.45)
Conditional BP	−0.34509 (−6.03)	0.01754 (1.05)	−0.32799 (−3.54)	−0.00901 (−0.70)
Conditional Apple	−0.23698 (−4.90)	0.02967 (2.01)	−0.28429 (−3.46)	0.00437 (0.38)
Conditional $/£	−0.30328 (−6.22)	0.00877 (0.65)	−0.34072 (−4.22)	−0.00799 (−0.68)
Conditional ¥/£	−0.28348 (−5.92)	0.01418 (1.18)	−0.24102 (−4.85)	−0.00465 (−0.57)
Conditional Fr/$	−0.31975 (−6.08)	0.00907 (0.70)	−0.30933 (−3.20)	−0.00597 (−0.44)

Notes: This table reports the results of scoring rule tests for one-step-ahead density forecast through the two out-of-sample periods considered (density forecasts 500 each out-of-sample period: 7/17/2008–6/16/2010 and 3/19/2013–2/16/2015). The entries are pairwise differences of the weighted average logarithmic score from the MME–DECO marginal and conditional distributions with respect to the counterpart distributions of the models in the columns; a negative value means that the MME–DECO presents a lower average score than its counterpart in the column. Amisano and Giacomini (2007) test t-statistics for the significance of the pairwise difference between the average logarithmic scores are in parentheses.

in terms of two main features. First, the polynomial structure of MME is based on a very simple functional form that allows up-to-one integration without requiring orthogonality conditions, unlike the MGC. Second, the basis distribution of the MME may be any pdf with finite moments up to the expansion order, whereas MGC pdfs are defined based on a series of derivatives of the Gaussian density. Therefore, the MME is a straightforward method for obtaining well-defined pdfs, which are sufficiently flexible to incorporate the salient empirical features of the portfolio return's distribution and to asymptotically approximate its true distribution. Indeed, the MME encompasses the MGC when the Gaussian density is used as the basis and positive transformations (GNT) are not implemented, but if positivity is used via GNT, then the MME yields a much simpler pdf.

An important feature of the MME is that it admits Engle (2002) decomposition of the likelihood function, which allows us to address the "curse of dimensionality" by two-step MLE. As an illustration of the applicability of the MME family of pdfs, we described an application to portfolio returns, where the correlations in the returns followed the DECO model of Engle and Kelly (2012).

We compared the in- and out-of-sample performance of the MME–DECO to that of the DECO and MGC–DECO. We found that the pdfs obtained using the MME–DECO were highly responsive to the weighting parameters, thereby providing a rich variety of shapes, which are particularly useful for modeling portfolio risk. Thus, the MME–DECO outperformed both MGC–DECO and DECO according to the AIC. The evaluation of the models in terms of the out-of-sample density forecasting showed that the MME–DECO obtained similar performance to the MGC–DECO, and both yielded significant improvements compared with the DECO. Our results highlight the flexibility of the MME pdf. In particular, it can obtain a feasibly parameterized model, which accurately captures both time-varying correlations and leptokurtosis. Thus, MME is an appealing tool for use in portfolio risk management due to its simplicity and accuracy. Its applications to other distributions, such as the Student's t, gamma, and normal inverse Gaussian, appear to be a promising avenue for further research into SNP methods.

Acknowledgements

We thank Geert Bekaert and Ike Mathur as well as two anonymous referees for their valuable comments throughout the revision process. We also thank Angel Leon, Ivan Paya, Gabriel Perez-Quiros, and seminar participants at the 35th International Symposium on Forecasting, XXII Forum of the Spanish Finance Association, XV Conference on International Economics, XV Iberian-Italian Congress of Financial and Actuarial Mathematics, 6th IFABS Conference, 10th BMRC-DEMS Conference, Bank of Spain and University of Murcia for helpful comments. Financial support from the Spanish Ministry of Economics and Competitiveness through Project ECO2013-44483-P is gratefully acknowledged.

Appendix A

This appendix includes the proofs of the properties of the MME densities presented in Sections 2.1.1. and 2.2, and the proofs related to the transformations required to implement the MME–DECO model. Proof 1 shows that MME densities integrate up to one; Proof 2–5 provide closed forms for marginal pdfs, (non-central) moments, cdf and copula density, respectively; Proof 6 shows the separability of the log-likelihood for the MME; Proof 7 includes an alternative approximation for implementing the MME–DECO model valid for very large portfolios, and Proof 8 shows that the transformation in Eqs. (32) and (33) yields uncorrelated variables in the DECO model.

Proof 1. The MME pdf integrates up to one.

$$\begin{aligned}
&F(\mathbf{x}_t)dx_{1t}\cdots dx_{nt}\\
&=\frac{1}{n}\left\{\prod_{i=1}^{n}g_i(x_{it})\right\}\left\{\sum_{i=1}^{n}w_i^{-1}\left[1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right]\right\}dx_{1t}\cdots dx_{nt}\\
&=\frac{1}{n}\sum_{i=1}^{n}\left[w_i^{-1}\int\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)g_i(x_{it})dx_{it}\prod_{j=1,\,j\neq i}^{n}\int g_j(x_{jt})dx_{jt}\right]\\
&=\frac{1}{n}n=1.
\end{aligned}\tag{36}$$

□

It must be also noted that the scaling constant w_i *is*

$$\begin{aligned} w_i &= \int \left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)g_i(x_{it})dx_{it} \\ &= \int g_i(x_{it})dx_{it} \\ &\quad +\sum_{s=1}^{n}\gamma_s^2\left(\int x_{it}^2 g_i(x_{it})dx_{it}+\mu_{is}^2\int g_i(x_{it})dx_{it}-2\mu_{is}\int x_{it}^s g_i(x_{it})dx_{it}\right) \\ &= 1+\sum_{s=1}^{n}\gamma_{is}^2(\mu_{i,2s}+\mu_{is}^2-2\mu_{is}^2)=1+\sum_{s=1}^{n}\gamma_{is}^2(\mu_{i,2s}-\mu_{is}^2). \end{aligned} \tag{37}$$

Proof 2. The MME marginals are combinations of univariate ME and Gaussian pdfs.

$$\begin{aligned} f_i(x_{it}) &= F(\mathbf{x}_t)dx_{1t}\cdots dx_{i-1,t}dx_{i+1,t}\cdots dx_{nt} \\ &= \tfrac{1}{nw_i}g_i(x_{it})\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\prod_{j=1,j\neq i}^{n}\int g_j(x_{jt})dx_{jt} \\ &\quad +\tfrac{1}{n}g_i(x_{it})\sum_{j=1,j\neq i}^{n}\left[\prod_{l=1,l\neq i}^{n}w_l^{-1}\int g_l(x_{lt})\left(1+\sum_{s=1}^{m}\gamma_{ls}^2\left(x_{lt}^s-\mu_{ls}\right)^2\right)dx_{lt}\right] \\ &= \tfrac{1}{nw_i}g_i(x_{it})\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)+\tfrac{n-1}{n}g_i(x_{it}),\quad \forall i=1,2,\ldots,n. \end{aligned} \tag{38}$$

□

Proof 3. The MME moments have a simple relation to the squared density parameters.

$$\begin{aligned} E\left[x_{it}^r\right] &= \int x_{it}^r f_i(x_{it})dx_{it} \\ &= \tfrac{n-1}{n}\int x_{it}^r g_i(x_{it})dx_{it}+\tfrac{1}{nw_i}\int x_{it}^r g_i(x_{it})dx_{it} \\ &\quad +\tfrac{1}{nw_i}\sum_{s=2}^{m}\gamma_{is}^2\int x_{it}^r\left(x_{it}^s-\mu_{is}\right)^2 g_i(x_{it})dx_{it} \\ &= \left[\tfrac{n-1}{n}+\tfrac{1}{nw_i}\right]\mu_{ir}+\tfrac{1}{nw_i}\sum_{s=1}^{m}\gamma_{is}^2\left(\int x_{it}^{2s+r}g_i(x_{it})dx_{it}+\mu_{is}^2\int x_{it}^r g_i(x_{it})dx_{it}\right. \\ &\quad \left.-2\mu_{is}\int x^{s+r}g_i(x_{it})dx_{it}\right) \\ &= \left[\tfrac{n-1}{n}+\tfrac{1}{nw_i}\right]\mu_{ir}+\tfrac{1}{nw_i}\sum_{s=1}^{m}\gamma_{is}^2\left[\mu_{i,2s+i}+\mu_{is}(\mu_{is}\mu_{ir}-2\mu_{i,s+r})\right],\quad \forall r\in\mathbb{Z}, \end{aligned} \tag{39}$$

and provided that $g_i(x_{it})$ has finite moments at least up to the $(2s+r)$th order. □

Proof 4. The cdf of the MME can be obtained through the cdfs of univariate ME and Gaussian distributions.

$$\begin{aligned} &\Pr[x_{1t}\leqslant\bar{x}_1,\ldots,x_{nt}\leqslant\bar{x}_n] \\ &= \tfrac{1}{n}\int_{-\infty}^{\bar{x}_1}\cdots\int_{-\infty}^{\bar{x}_n}\left[\prod_{i=1}^{n}g_i(x_{it})\right]\left\{\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\right\}dx_{1t}\cdots dx_{nt} \\ &= \tfrac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{\bar{x}_i}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)g_i(x_{it})dx_{it} \\ &\quad \times\prod_{j=1,j\neq i}^{n}\int_{-\infty}^{\bar{x}_j}g_j(x_{jt})dx_{jt} \\ &= \tfrac{1}{n}\sum_{i=1}^{n}h_i(\bar{x}_i)\prod_{j=1,j\neq i}^{n}\int_{-\infty}^{\bar{x}_j}g_j(x_{jt})dx_{jt}. \end{aligned} \tag{40}$$

□

Proof 5. Let $x_t\in\mathbb{R}^n$ be a random vector joint MME cdf $H(x_t)$ with margins $h_i(x_i)=z_i$, $z_i\in[0,1]$, and joint MME pdf $F(x_t)$ with margins $f_i(x_i), \forall i=1,2$. Then the Sklar's theorem (see e.g. Jondeau et al., 2007, p. 242) states that there exists a copula function C such that $H(x_t)=C(h_1(x_1),h_2(x_2),\ldots,h_n(x_n))$, the density of the copula being

$$c(h_1(x_1),h_2(x_2),\ldots,h_n(x_n))=\frac{F(\mathbf{x}_t)}{\prod_{i=1}^{n}f_i(x_i)}. \tag{41}$$

Based on this theorem and the densities in Eqs. (3) and (7) the density of the MME copula can be expressed as

$$c(h_1(x_1),h_2(x_2),\ldots,h_n(x_n))=\frac{\frac{1}{n}\left[\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\right]}{\prod_{i=1}^{n}\left[\frac{n-1}{n}+\frac{1}{nw_i}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}\right)^2\right)\right]}. \tag{42}$$

□

Proof 6. The log-likelihood of the Gaussian MME is separable thus formally admitting two-step estimation.

$$\begin{aligned} L_{F^+}(\mathbf{u}_t,\boldsymbol{\alpha},\boldsymbol{\rho},\boldsymbol{\gamma}) &= -\tfrac{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_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}^+\right)^2\right)\right]-T\ln(n) \\ &= -\tfrac{1}{2}\sum_{t=1}^{T}\left(n\ln(2\pi)+\ln|\mathbf{D}_t\mathbf{R}_t\mathbf{D}_t|+\mathbf{u}_t'\mathbf{D}_t^{-1}\mathbf{R}_t^{-1}\mathbf{D}_t^{-1}\mathbf{u}_t\right) \\ &\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}^+\right)^2\right)\right]-T\ln(n) \\ &= -\tfrac{1}{2}\sum_{t=1}^{T}\left(n\ln(2\pi)+2\ln|\mathbf{D}_t|+\ln|\mathbf{R}_t|+\boldsymbol{\varepsilon}_t'\mathbf{R}_t^{-1}\boldsymbol{\varepsilon}_t\right) \\ &\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}^+\right)^2\right)\right]-T\ln(n) \\ &= -\tfrac{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}_t|\right. \\ &\quad \left.+\boldsymbol{\varepsilon}_t'\mathbf{R}_t^{-1}\boldsymbol{\varepsilon}_t\right)+\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}^+\right)^2\right)\right] \\ &\quad -T\ln(n) \\ &= -\tfrac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{n}\left[\ln\left(2\pi\sigma_{it}^2\right)+\tfrac{u_{it}^2}{\sigma_{it}^2}\right]-\tfrac{1}{2}\sum_{t=1}^{T}\left(\ln|\mathbf{R}_t|+\boldsymbol{\varepsilon}_t'\mathbf{R}_t^{-1}\boldsymbol{\varepsilon}_t\right) \\ &\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_i^{-1}\left(1+\sum_{s=1}^{m}\gamma_{is}^2\left(x_{it}^s-\mu_{is}^+\right)^2\right)\right] \\ &\quad +\tfrac{1}{2}\sum_{t=1}^{n}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t-T\ln(n) \\ &= L_V(\mathbf{u}_t,\boldsymbol{\alpha})+L_{F^+}(\boldsymbol{\varepsilon}_t,\boldsymbol{\rho},\boldsymbol{\gamma})+\kappa. \end{aligned} \tag{43}$$

□

Proof 7. The transformation in Eq. (54) is an accurate alternative for implementing the MME–DECO model for large portfolios.

Let $R_t=(1-\rho_t)I_n+\rho_t J_n$ be the DECO correlation matrix, where I_n is the identity matrix of order n and J_n a $n\times n$ matrix of ones. Let us assume that this matrix can be decomposed as $R_t=A_tA_t$ where $A_t=\lambda_1 I_n+\lambda_2 J_n$. Then parameters λ_1 and λ_2 would be those that satisfy:

$$\begin{aligned} A_tA_t &= (\lambda_1 I_n + \lambda_2 J_n)(\lambda_1 I_n + \lambda_2 J_n) = \lambda_1^2 I_n + n\lambda_2^2 J_n + 2\lambda_1\lambda_2 J_n \\ &= \lambda_1^2 I_n + (n\lambda_2^2 + 2\lambda_1\lambda_2)J_n = (1-\rho_t)I_n + \rho_t J_n, \end{aligned} \tag{44}$$

i.e.,

$$\lambda_1^2 = 1-\rho_t \Rightarrow \lambda_1 = \sqrt{\mathbf{1}-\rho_t}, \tag{45}$$

$$\begin{aligned} n\lambda_2^2 + 2\lambda_1\lambda_2 = \rho_t \Rightarrow n\lambda_2^2 + 2\sqrt{\mathbf{1}-\rho_t}\lambda_2 - \rho_t = 0 \Rightarrow \lambda_2 \\ = \frac{-2\sqrt{\mathbf{1}-\rho_t} \pm \sqrt{4(\mathbf{1}-\rho_t) - 4n\rho_t}}{2n}. \end{aligned} \tag{46}$$

Nevertheless the solution

$$A_t = \sqrt{\mathbf{1}-\rho_t}I_n + \frac{1}{n}\left(\sqrt{\mathbf{1}-\rho_t+n\rho_t} - \sqrt{\mathbf{1}-\rho_t}\right)J_n \tag{47}$$

does not yield the exact matrix R_t but an approximation,

$$A_tA_t = R_t - 2\frac{1}{n}(1-\rho_t)J_n, \tag{48}$$

that converges to the true matrix as n tends to infinity. An interesting implication of this approach is that the inverse transformation may be straightforwardly computed:

$$A_t^{-1} = \phi_1 I_n + \phi_2 J_n, \tag{49}$$

where $A_tA_t^{-1} = I_n$. Therefore

$$\begin{aligned} A_tA_t^{-1} &= [\sqrt{\mathbf{1}-\rho_t}\mathbf{I}_n + \tfrac{1}{n}(\sqrt{\mathbf{1}-\rho_t+n\rho_t} - \sqrt{\mathbf{1}-\rho_t})\mathbf{J}_n][\phi_1\mathbf{I}_n + \phi_2\mathbf{J}_n] \\ &= \phi_1\sqrt{\mathbf{1}-\rho_t}I_n + \phi_1\tfrac{1}{n}(\sqrt{\mathbf{1}-\rho_t+n\rho_t} - \sqrt{\mathbf{1}-\rho_t})J_n \\ &\quad + \phi_2\sqrt{\mathbf{1}-\rho_t}J_n + \phi_2(\sqrt{\mathbf{1}-\rho_t+n\rho_t} - \sqrt{\mathbf{1}-\rho_t})J_n \\ &= \phi_1\sqrt{\mathbf{1}-\rho_t}I_n + [\phi_1\tfrac{1}{n}(\sqrt{\mathbf{1}-\rho_t+n\rho_t} - \sqrt{\mathbf{1}-\rho_t}) \\ &\quad + \phi_2\sqrt{\mathbf{1}-\rho_t+n\rho_t}]J_n \\ &= I_n, \end{aligned} \tag{50}$$

then,

$$\phi_1\sqrt{\mathbf{1}-\rho_t} = 1 \Rightarrow \phi_1 = \frac{1}{\sqrt{\mathbf{1}-\rho_t}}, \tag{51}$$

$$\begin{aligned} &\left[\phi_1\frac{1}{n}\left(\sqrt{\mathbf{1}-\rho_t+n\rho_t} - \sqrt{\mathbf{1}-\rho_t}\right) + \phi_2\sqrt{\mathbf{1}-\rho_t+n\rho_t}\right] = 0 \\ &\Rightarrow \phi_2 = \frac{1}{n\sqrt{\mathbf{1}-\rho_t+n\rho_t}} - \frac{1}{n\sqrt{\mathbf{1}-\rho_t}}. \end{aligned} \tag{52}$$

Therefore

$$A_t^{-1} = \frac{1}{\sqrt{\mathbf{1}-\rho_{t1}}}I_n + \left(\frac{1}{n\sqrt{\mathbf{1}-\rho_t+n\rho_t}} - \frac{1}{n\sqrt{\mathbf{1}-\rho_t}}\right)J_n. \tag{53}$$

Then, Eq. (54) represents an alternative transformation, to that in Eqs. (32) and (33), valid for large portfolios:

$$\begin{aligned} x_{it} = \frac{n-1}{n\sqrt{\mathbf{1}-\rho_t}}\varepsilon_{it} + \sum_{\substack{j=1 \\ j\neq i}}^{n}\left(\frac{1}{n\sqrt{\mathbf{1}-\rho_t+n\rho_t}} - \frac{1}{n\sqrt{\mathbf{1}-\rho_t}}\right)\varepsilon_{jt}, \\ \forall j = 1,2,\ldots,n. \end{aligned} \tag{54}$$

□

See Graybill (1983) for other related transformations and properties on the matrix$R_t = (1-\rho_t)I_n + \rho_t J_n$.

Proof 8. The MME–DECO model standardized with the transformation in Eqs. (32) and (33) has zero correlation for every pair of returns.

$$\begin{aligned} E_t(x_{it}x_{jt}) &= \tfrac{1}{1-\rho_t}E_t\left[\left(\varepsilon_{it} - \tfrac{c_t}{n}\sum_{s=1}^{n}\varepsilon_{st}\right)\left(\varepsilon_{jt} - \tfrac{c_t}{n}\sum_{h=1}^{n}\varepsilon_{ht}\right)\right] \\ &= \tfrac{1}{1-\rho_t}\left[E_t(\varepsilon_{it}\varepsilon_{jt}) - \tfrac{c_t}{n}\sum_{h=1}^{n}E_t[\varepsilon_{it}\varepsilon_{ht}] - \tfrac{c_t}{n}\sum_{s=1}^{n}E_t[\varepsilon_{jt}\varepsilon_{st}]\right. \\ &\quad \left. + \tfrac{c_t^2}{n^2}\sum_{s=1}^{n}\sum_{h=1}^{n}E_t[\varepsilon_{st}\varepsilon_{ht}]\right] \\ &= \tfrac{1}{1-\rho_t}\left[E_t(\varepsilon_{it}\varepsilon_{jt}) - \tfrac{c_t}{n}E_t(\varepsilon_{it}^2) - \tfrac{c_t}{n}\sum_{h=1,h\neq i}^{n}E_t[\varepsilon_{it}\varepsilon_{ht}] - \tfrac{c_t}{n}E_t\left(\varepsilon_{jt}^2\right)\right. \\ &\quad \left. - \tfrac{c_t}{n}\sum_{s=1,s\neq j}^{n}E_t[\varepsilon_{jt}\varepsilon_{st}] + \tfrac{c_t^2}{n^2}\sum_{s=1}^{n}E_t[\varepsilon_{st}^2] + 2\tfrac{c_t^2}{n^2}\sum_{s=1}^{n}\sum_{h=1,h>s}^{n}E_t[\varepsilon_{st}\varepsilon_{ht}]\right] \\ &= \tfrac{1}{1-\rho_t}\left[\tfrac{c_t^2}{n} - \tfrac{2c_t}{n} + 2\tfrac{c_t^2}{n^2}\sum_{s=1}^{n}\sum_{h=1,h>s}^{n}\rho_{sht} - \tfrac{c_t}{n}\sum_{h=1,h\neq i}^{n}\rho_{iht}\right. \\ &\quad \left. - \tfrac{c_t}{n}\sum_{s=1,s\neq j}^{n}\rho_{jht} + \rho_{ij}\right] \quad \forall i\neq j, \end{aligned} \tag{55}$$

and provided that $E_t(\varepsilon_{it}^2) = 1,\ \forall i = 1,2,\ldots,n$.

If $\rho_{sht} = \rho_t\ \forall s\neq h$ then

$$E_t(x_{it}x_{jt}) = \frac{1}{1-\rho_t}\left[\frac{c_t^2}{n} - \frac{2c_t}{n} + \frac{c_t^2(n-1)}{n}\rho_t - 2\frac{c_t(n-1)}{n}\rho_t + \rho_t\right] = 0. \tag{56}$$

□

Note that

$$\begin{aligned} \tfrac{c_t^2}{n} - \tfrac{2c_t}{n} &= \tfrac{1}{n}\left(1 + \tfrac{1-\rho_t}{1-\rho_t-n\rho_t} + 2\sqrt{\tfrac{1-\rho_t}{1-\rho_t-n\rho_t}}\right) - \tfrac{2}{n}\left(1 + \sqrt{\tfrac{1-\rho_t}{1-\rho_t-n\rho_t}}\right) \\ &= \tfrac{1}{n}\left(\tfrac{1-\rho_t}{1-\rho_t-n\rho_t} - 1\right) = \tfrac{-\rho_t}{1-\rho_t-n\rho_t}, \end{aligned} \tag{57}$$

and

$$\begin{aligned} (n-1)\left[\frac{c_t^2}{n} - \frac{2c_t}{n}\right]\rho_t + \rho_t = \left[\frac{-(n-1)}{1-\rho_t-n\rho_t} + 1\right]\rho_t = \frac{\rho_t}{1-\rho_t-n\rho_t}, \\ \forall i\neq j,\ i,j = 1,2,\ldots,n. \end{aligned} \tag{58}$$

References

Abramowitz, M., Stegun, I.A., 1972. Handbook of Mathematical Functions. Dover, New York.

Alexander, C., Lazar, E., 2006. Normal mixture GARCH(1,1): applications to exchange rate modelling. Journal of Applied Econometrics 21, 307–336.

Amisano, G., Giacomini, R., 2007. Comparing density forecasts via weighted likelihood ratio tests. Journal of Business and Economic Statistics 25, 177–190.

Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. Journal of Financial Economics 63, 443–494.

Aslandis, N., Casas, I., 2013. Nonparametric correlation models for portfolio allocation. Journal of Banking and Finance 37, 2268–2283.

Azzalini, A., Dalla Valle, A., 1996. The multivariate skew normal distribution. Biometrika 83, 715–726.

Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006. Multivariate GARCH models: a survey. Journal of Applied Econometrics 21, 79–109.

Blinnikov, S., Moessner, R., 1998. Expansions for nearly gaussian distributions. Astronomy & Astrophysics Supplement Series 130, 193–205.

Bollerslev, T., Wooldridge, J., 1992. Quasi maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11, 143–172.

Cappiello, L., Engle, R.F., Sheppard, K., 2006. Asymmetric dynamics in the correlations of global equity and bond returns. Journal of Financial Econometrics 4, 537–572.

Christoffersen, P., Errunza, V., Jacobs, K., Langlois, H., 2012. Is the potential for international diversification disappearing? a dynamic copula approach. Review of Financial Studies 25, 3711–3751.

Clements, A., Scott, A., Silvennoinen, A., 2015. On the benefits of equicorrelation for portfolio allocation. Journal of Forecasting 34, 507–522.
Cramér, H., 1925. On some classes of series used in mathematical statistics. Proceedings of the Sixth Scandinavian Congress on Mathematicians, 399–425.
Del Brio, E.B., Ñíguez, T.M., Perote, J., 2009. Gram–Charlier densities: a multivariate approach. Quantitative Finance 9, 855–868.
Del Brio, E.B., Ñíguez, T.M., Perote, J., 2011. Multivariate semi-nonparametric distributions with dynamic conditional correlations. International Journal of Forecasting 27, 347–364.
Diebold, F.X., Hahn, J., Tay, S.A., 1999. Multivariate density forecasts evaluation and calibration in financial risk management: high-frequency returns of foreign exchange. Review of Economics and Statistics 81, 661–673.
Ebert, S., 2013. Moment characterization of higher-order risk preferences. Theory and Decision 74, 267–284.
Ebert, S., Wiesen, D., 2011. Testing for prudence and skewness seeking. Management Science 57, 1334–1349.
Eeckhoudt, L., Schlesinger, H., 2006. Putting risk in its proper place. American Economic Review 96, 280–289.
Engle, R.F., 2002. Dynamic conditional correlation – a simple class of multivariate GARCH models. Journal of Business and Economic Statistics 20, 339–350.
Engle, R.F., Kelly, B., 2012. Dynamic equicorrelation. Journal of Business and Economic Statistics 30, 212–228.
Engle, R.F., Sheppard, K., 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Working Paper 8554, NBER.
Fajardo, J., Farias, A., 2010. Derivative pricing using multivariate affine generalized hyperbolic distributions. Journal of Banking and Finance 34, 1607–1617.
Fan, J., Qi, L., Xiu, D., 2014. Quasi-maximum likelihood estimation of GARCH models with heavy-tailed likelihoods. Journal of Business and Economic Statistics 32, 178–191.
Fiorentini, G., Sentana, E., 2007. On the efficiency and consistency of likelihood estimation in multivariate conditionally heteroskedastic dynamic regression models. Working Paper 0713, CEMFI.
Gallant, R., Nychka, D., 1987. Seminonparametric maximum likelihood estimation. Econometrica 55, 363–390.
Gallant, R., Tauchen, G., 1989. Seminonparametric estimation of conditionally constrained heterogeneous processes: asset pricing applications. Econometrica 57, 1091–1120.
Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102, 359–378.
Graybill, F.A., 1983. Matrices with Applications in Statistics. Wadsworth International Group, Belmont.
Harris, R.D.F., Mazibas, M., 2013. Dynamic hedge fund portfolio construction: a semi-parametric approach. Journal of Banking and Finance 37, 139–144.
Jondeau, E., Rockinger, M., 2001. Gram–Charlier densities. Journal of Economic Dynamics and Control 25, 1457–1483.
Jondeau, E., Poon, S.H., Rockinger, M., 2007. Financial Modeling under Non-Gaussian Distributions. Springer, London.
Kendall, M., Stuart, A., 1977. The Advanced Theory of Statistics. Griffin & Co, London.
Kotz, S., Nadarajah, S., 2004. Multivariate T-Distributions and their Applications. Cambridge University Press, Cambridge.
Lee, C.D., 1996. Properties and applications of the Samarnov family of bivariate distributions. Communications in Statistics – Theory and Methods 25, 1207–1222.
León, A., Menća, J., Sentana, E., 2009. Parametric properties of semi-nonparametric distributions, with applications to option valuation. Journal of Business and Economic Statistics 27, 176–192.
Malevergne, Y., Sornette, D., 2004. Collective origin of the coexistence of apparent random matrix theory noise and of factors in large sample correlation matrices. Physica A: Statistical Mechanics and its Applications 33, 660–668.
Mauleón, I., Perote, J., 2000. Testing densities with financial data: An empirical comparison of the Edgeworth–Sargan density to the Student's t. European Journal of Finance 6, 225–239.
Ñíguez, T.M., Perote, J., 2012. Forecasting heavy-tailed densities with positive edgeworth and Gram–Charlier expansions. Oxford Bulletin of Economics and Statistics 74, 600–627.
Ñíguez, T.M., Perote, J., 2014. The moments expansion: a semi-nonparametric method with applications for risk management. Working Paper 14-1, Westminster Business School.
Ñíguez, T.M., Paya, I., Peel, D., Perote, J., 2015. Higher-order risk preferences, constant relative risk aversion and the optimal portfolio allocation. Working Paper 1520, Bank of Spain.
Pagan, A.R., 1986. Two stage and related estimators and their applications. Review of Economic Studies 53, 517–538.
Patton, A.J., 2006. Modelling asymmetric exchange rate dependence. International Economic Review 47, 527–556.
Patton, A.J., 2012. A review of copula models for economic time series. Journal of Multivariate Analysis 110, 4–18.
Phillips, P.C.B., 1977. A general theorem in the theory of asymptotic expansions as approximations to the finite sample distributions of econometric estimators. Econometrica 45, 1517–1534.
Ruiz, E., Pascual, L., 2002. Boostraping financial time series. Journal of Economic Surveys 16, 271–300.
Samarnov, O.V., 1966. Generalized normal correlation and two-dimensional Frechet classes. Doklady (Soviet Mathematics) 168, 596–599.
Sarabia, J.M., Gómez-Déniz, E., 2008. Construction of multivariate distributions: a review of some recent results. Statistics and Operations Research Transactions 32, 3–36.
Sentana, E., Calzolari, G., Fiorentini, G., 2008. Indirect estimation of large conditionally Heteroskedastic factor models, with an application to the dow 30 stocks. Journal of Econometrics 146, 10–25.
Silvennoinen, A., Terasvirta, T., 2009. Multivariate GARCH models. In: Andersen, T. G., Davis, R.A., Krei, J.P., Mikosch, T. (Eds.), Handbook of Financial Time Series 201–229. Springer, New York.
White, H., 1994. Estimation, Inference and Specification Analysis. Cambridge University Press, New York.
Withers, C.S., Nadarajah, S., 2014. The dual multivariate Charlier and Edgeworth expansions. Statistics and Probability Letters 87, 76–85.