A New Class of Multivariate Skew Densities, with Application to GARCH Models

Luc Bauwens

CORE and Department of Economics, Université catholique de Louvain.

E-mail: bauwens@core.ucl.ac.be

and

Sébastien Laurent

CREST (CNRS),

CORE, Université catholique de Louvain,

and Department of Quantitative Economics, Maastricht University.

E-mail: laurent@core.ucl.ac.be

April 2002, revised April 30, 2004

Abstract

We propose a practical and flexible method to introduce skewness in multivariate symmetric distributions. Applying this procedure to the multivariate Student density leads to a "multivariate skew-Student" density, in which each marginal has a specific asymmetry coefficient. Combined with a multivariate GARCH model, this new family of distributions is found to be more useful than its symmetric counterpart for modelling stock returns and especially forecasting the Value-at-Risk of portfolios.

Keywords: Multivariate skew density, Multivariate Student density, Multivariate GARCH models, Value-at-Risk.

JEL classification: C13, C32, C52.

1 INTRODUCTION

The increased importance played by risk and uncertainty considerations in modern economic theory has come with the development of new econometric time series techniques that allow for the modelling of time varying means, variances and correlations. The most widespread modelling approach to capture these properties is to specify a dynamic model for the conditional mean and the conditional variance, such as an autoregressive moving average - autoregressive conditional heteroscedasticity model (ARMA-ARCH) or one of its various extensions (see the seminal paper of Engle 1982).

Although there is a huge literature on univariate generalized ARCH (GARCH) models, much less work is concerned with their multivariate extensions. For this reason, Geweke and Amisano (2001) argue that "while univariate models are a first step, there is an urgent need to move on to multivariate modelling of the time-varying distribution of asset returns". Indeed, financial volatilities move together over time across assets and markets. Recognizing this commonality through a multivariate modelling framework can lead to obvious gains in efficiency and to more relevant financial decision making than working with separate univariate models (see Bauwens, Laurent and Rombouts 2003, for a recent survey of multivariate GARCH models).

The estimation of multivariate GARCH models is commonly done by maximizing a Gaussian likelihood function. Even if it is unrealistic in practice, the normality assumption may be justified by the fact that the Gaussian quasi-maximum likelihood (QML) estimator is consistent provided the conditional mean and the conditional variance are specified correctly. In this respect, Jeantheau (1998) has proved the strong consistency of the QML estimator of multivariate GARCH models, extending previous results of Lee and Hansen (1994) and Lumsdaine (1996).

Another well established stylized fact of financial returns, at least when they are sampled at high frequencies, is that they exhibit fat-tails (which corresponds to a kurtosis coefficient larger than three) and are often skewed (see among others Alexander 2001 and Gourieroux 1997).

As far as financial applications are concerned, and in order to gain statistical efficiency, it is of primary importance to base modelling and inference on a more suitable distribution than the multivariate normal. In particular, Engle and González-Rivera (1991) show in a univariate framework that the Gaussian QML estimator of a GARCH model is inefficient, with the degree of inefficiency increasing with the degree of departure from normality, while Peiró (1999) emphasizes the relevance of modelling the third and fourth moments in asset pricing models, portfolio selection and option pricing theories. The challenge to econometricians is to design multivariate distributions that are both easy to use for inference and compatible with the skewness and kurtosis properties of financial returns.

The main contribution of this paper is to propose a practical and flexible method to introduce skewness in multivariate symmetric distributions. Applying this procedure to the multivariate Student density leads to a "multivariate skew-Student" density, in which each marginal has a specific asymmetry parameter. We illustrate that, combined with a multivariate GARCH model, this new family of distributions is useful for modelling financial returns and forecasting the Value-at-Risk (VaR) of portfolios of assets. Indeed, throughout several examples, we find that the multivariate skew-Student density provides better, or at least not worse, out-of-sample VaR forecasts than a symmetric density.

The rest of the paper is organized as follows. In Section 2, we define the new family of multivariate skew densities. In Section 3, we investigate the usefulness of the multivariate skew-Student density in a VaR application based on a multivariate GARCH model. In Section 4, we present our conclusions and directions for future research.

2 MULTIVARIATE SKEW DENSITIES

ML estimation of GARCH models requires an assumption on the distribution of the innovations. To accommodate the leptokurtosis of financial returns, univariate GARCH models have been combined by Bollerslev (1987) with a Student distribution. Indeed, although a GARCH model generates excess kurtosis when combined with a Gaussian conditional density, it does not fully account for the excess kurtosis present in most return series. The Student density has become widely used due to its simplicity and its inherent ability of fitting excess kurtosis. It is thus natural to consider its multivariate extension as an alternative to the Gaussian when dealing with multivariate GARCH models (see Fiorentini, Sentana and Calzolari 2003).

As a reminder, the standardized multivariate Student density for the random vector $X \in \Re^k$ may be defined as:

$$g(x) = \frac{\Gamma\left(\frac{\upsilon+k}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)\left[\pi(\upsilon-2)\right]^{\frac{k}{2}}}\left[1+\frac{x'x}{\upsilon-2}\right]^{-\frac{k+\upsilon}{2}}, \qquad (1)$$

where $\Gamma(.)$ is the Gamma function. This density is denoted $ST(0, I_k, \upsilon)$. We restrict υ to be larger than 2 for ensuring the existence of the variance matrix (empirically, point estimates smaller than 2 are rarely found).

However, the main drawback of this density is that it is symmetric while the distribution of financial returns may be skewed. Consequently, using a more appropriate distribution may lead to improved empirical modelling and financial decision making.

Asymmetric densities can be defined by introducing skewness in symmetric densities by means of new parameters, such that the symmetric density results as a particular case. In Section 2.2, we propose a simple and intuitive method to introduce skewness into a multivariate symmetric unimodal density. Before that, we define the notion of symmetry that we rely on.

2.1 Multivariate symmetry

In the univariate case, symmetry corresponds to $g(x) = g(-x)$, assuming $g(x)$ is a unimodal probability density function and $E(X) = 0$. In the multivariate case, we use the following definition of symmetry of a standardized density $g(x)$:

Definition 1 (M-symmetry): *The unimodal density $g(x)$ defined on $\Re^k$, such that $E(X) = 0$, and $Var(X) = I_k$, is symmetric if and only if for any x, $g(x) = g(Qx)$ for all diagonal matrices Q whose diagonal elements are equal to +1 or to -1. If X is a random vector with a density satisfying this definition, we write*

$$X \sim M\text{-}Sym(0, I_k, g). \tag{2}$$

In the bivariate case, this definition means that $g(x_1, x_2) = g(-x_1, x_2) = g(x_1, -x_2) = g(-x_1, -x_2)$.

Spherically symmetric (SS) densities, defined by the property that the density depends on x through $x'x$ only, i.e.

$$g(x) \propto k(x'x), \tag{3}$$

for an appropriate integrable positive function $k(.)$, are M-symmetric. The most well known examples of SS-densities are the standard normal density and the standard Student density $ST(0, I_k, \upsilon)$ (Johnson 1987 chapter 6, provides graphical illustrations of several bivariate SS-densities). However, there exists other distributions that have the desired property while not being spherically symmetric. A large class is defined by

$$g(x) = \prod_{i=1}^{k} g_i(x_i), \tag{4}$$

where $g_i(.)$, $\forall i = 1, \ldots, k$, is a univariate symmetric density (unimodal, with mean 0 and unit variance). One can show that, with respect to M-symmetry, (3) and (4) are equivalent if and only if $g(.) = N(0, I_k)$ and $g_i(.) = N(0, 1)$, $\forall i$. Nevertheless, if $g_i(.) = ST(0, 1, \upsilon)$, $\forall i$ and $g(.) = ST(0, I_k, \upsilon)$, there is a difference between (3) and (4) since the elements of (3) are not mutually independent whereas those of (4) are. Notice however that both multivariate densities have the same univariate marginal densities.

2.2 Skew densities

A brief literature review.

Although several multivariate asymmetric densities have been proposed in the literature, we did not find them simple and flexible enough to generate asymmetry and fat tails or to use in ML estimation.

Among these, we mention briefly the density of Jones (2001), where the covariances are necessarily positive, Jones (2002), and the heavily parametrized Edgeworth-Sargan density used by Mauleón and Perote (1999) for a bivariate GARCH model. Branco and Dey (2001) propose a class of skew-elliptical densities. They generalize to the complete class of elliptically contoured densities (i.e. densities obtained by linear transformation of a SS-density) the multivariate skew-normal density of Azzalini and Dalla Valle (1996). Sahu, Dey and Branco (2003) provide a further generalization and use it as the error distribution of regression models.

Asymmetric densities can also be generated by using finite mixtures of symmetric ones. In a GARCH model, Vlaar and Palm (1993) use a mixture of two normal densities. One potential drawback of mixtures is the large number of parameters. Nevertheless, this model class coupled with GARCH structures has recently gained attention; see Haas, Mittnik and Paolella

(2002) and Alexander and Lazar (2003).

Multivariate stable distributions (see Samorodnitsky and Taqqu 1994) also allow for fat tails and asymmetries with respect to the Gaussian density. When this is the case, they are incompatible with finite variance laws. Furthermore, the corresponding densities are usually not known in closed form. One has to work through the characteristic function which renders estimation by ML rather difficult.

Fernandez, Osiewalski and Steel (1995) also propose a class of continuous multivariate distributions, called *v-spherical*, that are more restrictive in the sense that the skewness is the same for every coordinate.

Ferreira and Steel (2003) introduced a novel class of skewed multivariate distributions and, more generally, a method of building such a class on the basis of univariate skewed distributions. Their method is based on a general linear transformation of a multidimensional random variable with independent components, each with a skewed distribution.

Finally, a class of multivariate densities that could be of interest is the so-called poly-t densities that contain the multivariate Student density as a particular case. Poly-t densities arise as posterior densities in Bayesian inference, see Drèze (1978), and can be heavily skewed, have fat tails and even be multimodal. However, the relations between their parameters and moments is complex (see Richard and Tompa 1980 for results on moments of poly-t densities). Moreover, ML estimation of their parameters even in the case of *i.i.d.* sampling is difficult in practice.

A new class of multivariate skew densities.

We generalize to the multivariate case the method proposed by Fernández and Steel (1998) to construct a skew density from a symmetric one.

In the univariate case, the idea of the transformation to induce symmetry is to scale X differently for negative and positive values: when $X \geq 0$, we multiply it by a positive constant ξ, and when $X < 0$, we divide it by the same constant. Assume, without loss of generality, that $\xi > 1$. The effect of this transformation on the symmetric unimodal density $g(x)$ centered at 0 is easy to understand: the new density function becomes more spread over the positive values than $g(x)$, while it is less spread over the negative part. The density is skewed to the right. It remains continuous at 0, although its first derivative is not, except in the case of symmetry. The parameter ξ clearly determines the direction and the intensity of the skewness.

We can use the same idea for each coordinate of a multivariate M-symmetric density. Let $X \sim M\text{-}Sym(0, I_k, g)$. Define k random variables W_i following independent Bernoulli distributions:

$$W_i = \begin{cases} 1 \text{ with probability } \xi_i^2/(1+\xi_i^2) \\ 0 \text{ with probability } 1/(1+\xi_i^2) \end{cases} \quad \text{for } i = 1, 2, \ldots, k. \tag{5}$$

Define also the following transformations:

$$Y_i = W_i \, |X_i| \, \xi_i - (1 - W_i) \, |X_i| \, \xi_i^{-1}, \text{ for } i = 1, 2, \ldots, k. \tag{6}$$

In the univariate case ($k = 1$), this transformation has the effect described above, since $Y_i = |X_i|\xi_i$ is positive if $W_i = 1$ and $Y_i = -|X_i|/\xi_i$ is negative if $W_i = 0$. It is easy to show that the density of Y_i is given by:

$$\begin{aligned} h(y_i|\xi_i) &= \Pr(W_i = 1|\xi_i) \, h(y_i|W_i = 1, \xi_i) \\ &+ \Pr(W_i = 0|\xi_i) \, h(y_i|W_i = 0, \xi_i) \\ &= [g(y_i/\xi_i) 1_{y_i \geq 0} + g(y_i \xi_i) 1_{y_i < 0}] \, \frac{2\xi_i}{1 + \xi_i^2} \\ &= \frac{2\xi_i}{1 + \xi_i^2} \, g(y_i^*) \end{aligned} \tag{7}$$

where

$$y_i^* = \begin{cases} y_i/\xi_i \text{ if } y_i \geq 0 \\ y_i \xi_i \text{ if } y_i < 0. \end{cases} \tag{8}$$

It should be clear (see equation (3) in Fernández and Steel 1998) that $\xi_i^2 = \Pr[Y_i \geq 0|\xi_i] / \Pr[Y_i < 0|\xi_i]$, so that if $\xi_i > 1$ (< 1), $h(y_i|\xi_i)$ is skewed to the right (resp. left). This explains the parametrization of $\Pr[W_i = 1|\xi_i]$ in (5).

In the multivariate case, the transformations defined by (5) and (6) are operated on each of the 2^k orthants of $\Re^k$. On each of these, the density $g(x)$ is transformed into the density of y conditional upon y being in the relevant orthant. Each orthant corresponds to one and only one realization of the vector $W = (W_1, W_2, \ldots, W_k)$ of Bernoulli variables. The conditional densities on each orthant are then weighted by the probability that W takes the corresponding value and added up.

For example, in the bivariate case, defining $\xi = (\xi_1, \xi_2)'$, we compute the density of Y as:

$$\begin{aligned} h(y|\xi) &= \Pr(W_1 = 1, W_2 = 1|\xi)\, h(y|W_1 = 1, W_2 = 1, \xi) \\ &+ \Pr(W_1 = 1, W_2 = 0|\xi)\, h(y|W_1 = 1, W_2 = 0, \xi) \\ &+ \Pr(W_1 = 0, W_2 = 1|\xi)\, h(y|W_1 = 0, W_2 = 1, \xi) \\ &+ \Pr(W_1 = 0, W_2 = 0|\xi)\, h(y|W_1 = 0, W_2 = 0, \xi), \end{aligned}$$

and this yields:

$$\begin{aligned} h(y|\xi) &= 2^2 \frac{\xi_1}{1+\xi_1^2} \frac{\xi_2}{1+\xi_2^2} \qquad (9) \\ &\times \Big[g(y_1/\xi_1, y_2/\xi_2) 1_{(y_1 \geq 0, y_2 \geq 0)} + g(y_1 \xi_1, y_2/\xi_2) 1_{(y_1 < 0, y_2 \geq 0)} \\ &+ g(y_1 \xi_1, y_2 \xi_2) 1_{(y_1 < 0, y_2 < 0)} + g(y_1/\xi_1, y_2 \xi_2) 1_{(y_1 \geq 0, y_2 < 0)} \Big]. \end{aligned}$$

In the multivariate case, the density is expressed as follows:

$$h(y|\xi) = 2^k \left(\prod_{i=1}^{k} \frac{\xi_i}{1+\xi_i^2} \right) g(y^*), \qquad (10)$$

where

$$y^* = (y_1^*, y_2^*, \ldots, y_k^*)' \qquad (11)$$

$$\begin{aligned} y_i^* &= y_i \xi_i^{I_i} \ (i = 1, 2, \ldots, k), & (12) \\ I_i &= \begin{cases} -1 \text{ if } y_i \geq 0 \\ \\ 1 \text{ if } y_i < 0. \end{cases} & (13) \end{aligned}$$

Here is a list of properties of this density:

1) The mode of $h(y|\xi)$ is at 0.

2) Each marginal density of $h(y|\xi)$ can be obtained by transforming the corresponding marginal density of $g(x)$ using the relevant part of the transformation defined in (5) and (6). For a univariate marginal, the resulting density is like in (7), i.e. "skewed à la Fernández and Steel (1998)".

3) ξ_i^2 is a skewness measure determining the skewness of the marginal density of Y_i, see the comment after (8) and Section 2 in Fernández and Steel (1998). Right (resp. left) skewness corresponds to $\log \xi_i > 0 \, (< 0)$.

4) If the c.d.f. of X is known analytically, so is the c.d.f. of Y.

5) Following equation (5) in Fernández and Steel (1998), if X has finite moments of order r, so does Y. In particular,

$$\mathrm{E}\,(Y_i^r|\xi_i) = \frac{\xi_i^{r+1} + \frac{(-1)^r}{\xi_i^{r+1}}}{\xi_i + \frac{1}{\xi_i}} \, 2\,\mathrm{E}(X_i^r|X_i > 0). \qquad (14)$$

See Lambert and Laurent (2001) for the link between ξ_i and the first four moments of Y_i.

6) Finally, if the second-order moments exist, it is obvious that the elements of Y are uncorrelated (since those of X are uncorrelated by assumption). Hence if m is the vector of means and s is the vector of standard deviations of Y, standardization is straightforwardly achieved by the transformation $Z = (Y - m)./s$, where ./ means element by element division.

<u>Skew-Student densities.</u>

The distribution of financial returns has fat tails and is often skewed. For this reason, it seems natural to apply the skewness filter in (10) to the multivariate Student density given in (1). Note that as pointed out by a referee, our presentation could be extended to the case where $\upsilon \leq 2$, and most results hold in this case.

Doing so produces what we call the "standardized skew-Student density", which is defined as follows:

$$f(z|\xi, \upsilon) = \left(\frac{2}{\sqrt{\pi}}\right)^k \left(\prod_{i=1}^{k} \frac{\xi_i s_i}{1+\xi_i^2}\right) \frac{\Gamma(\frac{\upsilon+k}{2})}{\Gamma(\frac{\upsilon}{2})(\upsilon-2)^{\frac{k}{2}}} \left(1 + \frac{{z^*}' z^*}{\upsilon - 2}\right)^{-\frac{k+\upsilon}{2}}, \tag{15}$$

where

$$z^* = (z_1^*, \ldots, z_k^*)', \tag{16}$$

$$z_i^* = (s_i z_i + m_i)\, \xi_i^{I_i}, \tag{17}$$

$$m_i = \frac{\Gamma\left(\frac{\upsilon-1}{2}\right)\sqrt{\upsilon-2}}{\sqrt{\pi}\,\Gamma\left(\frac{\upsilon}{2}\right)} \left(\xi_i - \frac{1}{\xi_i}\right), \tag{18}$$

$$s_i^2 = \left(\xi_i^2 + \frac{1}{\xi_i^2} - 1\right) - m_i^2, \tag{19}$$

$$I_i = \begin{cases} -1 \text{ if } z_i \geq -\frac{m_i}{s_i} \\ 1 \text{ if } z_i < -\frac{m_i}{s_i}. \end{cases} \tag{20}$$

Clearly the constants m_i and s_i^2 are functions of ξ and υ and do not represent additional parameters. This density is denoted $SKST(0, I_k, \xi, \upsilon)$, and ξ is the vector of asymmetry parameters.

As an illustration, Figure 1 shows the contours of a $SKST$ $(0, I_2, \xi, 6)$ density with $\xi_1 = 1$, $\xi_2 = 1.3$. The contours show the right-skewness of the density in the direction of z_2, and its symmetry in the direction of z_1. One also clearly sees from this "ovoid" that the mode is not at zero since this is a standardized version.

Obviously, the $SKST(0, I_k, \xi, \upsilon)$ density implies the same thickness of tails of each marginal density, and its components are not independent. An obvious alternative but close family is obtained by taking the product of k independent univariate $SKST(0, 1, \xi_i, \upsilon_i)$ densities. When the degrees of freedom parameters are all equal ($\upsilon_i = \upsilon, \forall i$), the marginal densities of this product are identical to those of the skew-Student density $SKST(0, I_k, \xi, \upsilon)$, although the joint densities are different. We have therefore a new copula density for asymmetric densities.

[INSERT FIGURE 3.2 HERE.]

3 VAR PREDICTION USING A MULTIVARIATE GARCH MODEL

3.1 Data, models and estimates

In the empirical applications we use two datasets. The first one consists of three major exchange rates: the Japanese yen (YEN), the EURO (EUR) and the British pound (GBP), all quoted daily against the US dollar, at 12:00 GMT+1 (Source: Reuters). For these series, we have about 12 years of data, from January 1989 to February 2001 (3066 observations). The EUR/USD series has been obtained from the Deutsche mark (DEM) USD series by using the conversion rate 1 EUR/USD = 1.95583 DEM/USD.

The second dataset is composed of three US stocks of the Dow Jones industrial index (Source: Yahoo Finance) from January 1990 to May 2002 (3113 daily observations): the Alcoa stock (AA), the Caterpillar Inc. stock (CAT) and the Walt Disney Company stock (DIS). The symbol inside the parentheses is the short notation (or ticker) for the series, which is used in the tables and comments below.

For all price series p_{it} ($\forall i = 1, \ldots, k$), daily returns are defined as $y_{it} = 100 (\ln p_{it} - \ln p_{i,t-1})$. Preliminary univariate analyses performed on the series allow us to conclude that an AR(p)-GARCH(1,1) model suffices to characterize the dynamics of the first two conditional moments of the three exchange rate series, while an additional term in the GARCH equation is needed to capture a strong leverage effect for the three stocks. In practice we have used the GJR(1,1) model (see Glosten, Jagannathan and Runkle 1993) for the stock returns, which adds to the GARCH(1,1) equation the lagged squared shock when the shock is negative. The order of the autoregressive process is set to 1 for AA, EUR and YEN, and to 0 for the other series. Note that the models have been estimated using G@RCH 3.0, an Ox package with a friendly dialog-oriented interface designed for the estimation and forecasting of various univariate ARCH-type models (see Doornik 2001 for more details about the Ox language and Laurent and Peters 2002 about the G@RCH package).

Estimation results reveal also that a univariate skew-Student distribution for the innovations (obtained by setting $k = 1$ in Definition 2) improves the quality of the models with respect to the normal and the Student distributions (likelihood ratio tests and the Schwarz information criterion clearly favor the skew-Student). Indeed, the three stocks were found to be positively skewed while the three exchange rates, i.e. EUR, YEN and GBP, where found respectively positively skewed, negatively skewed and almost symmetric.

Univariate GARCH models are extensively used to compute the Value-at-Risk (VaR) of level $\alpha\%$ of a portfolio of assets, over a given time horizon (see Jorion 2000). Indeed, the VaR of level $\alpha\%$ is directly deduced from the α-th left or right quantile of the conditional distribution of the returns (over the given horizon). The left quantile (such that the probability mass at the left tail is equal to $\alpha\%$) corresponds to the VaR of a long position, and the right quantile to the VaR of a short position. Recall that an investor holds a long position if he has bought an asset, in which case he incurs the risk of a loss of value of the asset. An investor holds a short position if he has sold an asset and detains cash, in which case he incurs a positive opportunity cost if the asset value increases.

Given an estimated parametric GARCH model, one knows the conditional distribution of the return, and one can compute the needed quantiles. For instance, Giot and Laurent (2003) have shown the supremacy of the univariate skew-Student density (compared to the normal and Student distribution) when forecasting the one-day-ahead VaR of many assets for long and short trading positions (both in- and out-of-sample).

Suppose now that one is considering a portfolio of k assets. The portfolio return is $r_t = w'y_t$, where $w' = (w_1, \ldots, w_k)$ is the vector of shares of the different assets in the portfolio, and $y_t = (y_{1t}, \ldots, y_{kt})'$ is the vector of returns. A univariate GARCH model can be fit to r_t and the VaR can be computed accordingly. However, if the weight vector changes, the model has to be specified and estimated again. On the contrary, if a *multivariate* GARCH model is fitted, the multivariate distribution of the returns can be directly used to compute the implied distribution of any portfolio. In other words, there is no need to reestimate the model for a different weight vector. Note that, since the distribution assumption is crucial to obtain accurate VaR forecast in a univariate framework (see Mittnik and Paolella 2000 or Giot and Laurent 2003 among others), one can conjecture this will also be true in the multivariate case, which calls for the use of a flexible density for the innovations.

Before presenting the results of the empirical application, let us first describe the specification we choose for the first two conditional moments. Engle (2002) and Tse and Tsui (2002) have proposed to extend the constant conditional correlation (CCC) model of Bollerslev (1990) by making the conditional correlation matrix time varying.

The dynamic conditional correlation (DCC) model of Engle (2002) is defined as follows:

$$y_t = \mu_t + \Sigma_t^{1/2} z_t, \tag{21}$$

$$\Sigma_t = D_t R_t D_t, \tag{22}$$

$$D_t = \text{diag}\,(\sigma_{11,t}^{1/2}, \ldots, \sigma_{kk,t}^{1/2}), \tag{23}$$

$$R_t = (\text{diag}\, Q_t)^{-1/2} Q_t (\text{diag}\, Q_t)^{-1/2}, \tag{24}$$

where $\mu_t = (\mu_{1t}, \ldots, \mu_{kt})'$ is the conditional mean vector, Σ_t is the conditional covariance matrix, $\sigma_{ii,t}$ is specified as a univariate GARCH-type equation, and R_t is the conditional correlation matrix.

The $k \times k$ symmetric positive definite matrix Q_t is given by:

$$Q_t = (1 - \alpha - \beta)R + \alpha u_{t-1}u'_{t-1} + \beta Q_{t-1}, \tag{25}$$

where $u_t = (u_{1t}, \ldots, u_{kt})'$, $u_{it} = (y_{it} - \mu_{it})/\sqrt{\sigma_{ii,t}}$, R is the $k \times k$ unconditional covariance matrix of u_t, and α and β are positive scalar parameters satisfying $\alpha + \beta < 1$. We replace R by its empirical counterpart to render the estimation simpler, as suggested by Engle (2002).

Table 1 reports estimation results for the complete samples of the two trivariate models, i.e. the series AA-CAT-DIS and EUR-YEN-GBP. Based on the univariate results, we choose univariate AR(p) models ($p = 0$ or 1, according to the series) for the conditional mean equations and a GARCH(1,1) or a GJR(1,1) for the conditional variances (respectively for the exchange rates and the stocks). The models are estimated by maximum likelihood in one step, assuming a multivariate skew-Student distribution for the innovations (z_t). Note that we have conducted a Monte Carlo study about ML estimates of the parameters of the multivariate skew-Student distribution (see Bauwens and Laurent 2002). The results show that even with a sample of 1000 observations, the MLE have a close to normal sampling distribution. This holds also for mean and variance parameters, whether the volatility of the data generating process is constant or of the GARCH type. In Table 1, ML estimates using our two trivariate data sets are reported with corresponding standard errors in parentheses. We only report estimates of the parameters of the DCC part and of the skew-Student distribution.

[INSERT TABLE 1 HERE.]

Table 1 also contains two diagnostic tests for the fitted models. The first one, *LRT(CCC)*, is a likelihood ratio statistic for testing the restriction $H_0 : \alpha = \beta = 0$, asymptotically $\sim \chi^2(2)$. In both cases, the CCC model is rejected at any conventional significance level. This result suggests that the additional flexibility of the DCC model is highly justified. Notice that $\hat{\alpha} + \hat{\beta}$ is close to unity so that standard errors have to be interpreted with care (although the sample size is large). The second statistic, $LRT(ST)$, is the LR statistic, asymptotically $\sim \chi^2(3)$, for testing the null hypothesis of symmetry, i.e. $H_0 : \log \xi_1 = \log \xi_2 = \log \xi_3 = 0$. In both cases, the symmetry assumption is clearly rejected. In a previous version of this paper (Bauwens and Laurent 2002), we have estimated the multivariate GARCH model of Tse and Tsui (2002) with a normal, Student and a skew-Student likelihood on four daily indexes. Likelihood ratio tests and information criteria also clearly favored the latter distribution.

Interestingly, to avoid the dimensionality problem of most multivariate GARCH models, Engle (2002) and Engle and Sheppard (2001) show that the DCC model can be estimated consistently using a two step approach. Under the normality assumption, the log-likelihood can be decomposed as the sum of a mean-volatility term (that depends only on the parameters of the conditional means and variances), and a second term that depends on the conditional correlation parameters (holding the other parameters as fixed). In other words, it is possible to obtain consistent estimates of the DCC model by estimating firstly k univariate GARCH-type models and secondly the conditional correlation part, using u_t as given. The price to pay is that the estimates are not fully efficient (since they are limited information estimators) and that standard errors have to be corrected. This issue is not relevant to forecast the VaR. It could be for estimating the standard error of the forecast, although Ruiz and Pascual (2002) show that this does not seem to matter much.

When using a non-normal distribution, the decomposition proposed by Engle (2002) is obviously not possible and one should adopt a one step approach. However, to stay in the spirit of the DCC model, we also propose to estimate the k univariate GARCH-type model by QML (to estimate u_t), and then estimate the correlation part together with the parameters ξ and υ of the skew-Student density.

Estimates are reported between brackets in Table 1 (third rows). Both the one- and two-step methods give very similar estimates, with a small exception in the case of $\log \xi_3$ in the exchange rate model.

3.2 VaR forecasts and evaluations

In a 'real life situation', GARCH models can be used to deliver out-of-sample forecasts of VaR measures, where the model is estimated on the known returns (up to time t for example) and the VaR forecast is made for time $t+h$, where h is the horizon of the forecast. In the applications below, we use a one day horizon.

Let us denote by $\mu_{t+1|t}$ and $\Sigma_{t+1|t}$ respectively the one-step-ahead forecasts of μ_t and Σ_t, given information up to time t. Under the normality assumption, the one-step-ahead VaR computed at t for long trading positions is given by $w'\mu_{t+1|t} + z_\alpha\sqrt{w'\Sigma_{t+1|t}w}$, while for short trading positions it is equal to $w'\mu_{t+1|t} + z_{1-\alpha}\sqrt{w'\Sigma_{t+1|t}w}$, with z_α being the left quantile at $\alpha\%$ of the normal distribution and $z_{1-\alpha}$ is the right quantile at $\alpha\%$. Under the assumption of a multivariate Student distribution, the one-step-ahead VaR is obtained by replacing z_α and $z_{1-\alpha}$ by respectively $t_{\alpha,\upsilon}$ and $t_{1-\alpha,\upsilon}$, i.e. the left and right quantiles of the Student distribution with υ degrees of freedom.

Under the assumption that $z_t \sim SKST(0, I_k, \xi, \upsilon)$, there is no analytic easy-to-use formula to pass from the conditional mean and covariance matrix to the long and short VaR of the portfolio. However, the VaR can be computed using a simple Monte Carlo simulation as widely used in VaR computations. We just have to simulate a set of one-day-ahead returns of the chosen portfolio, as $r_j^* = w'\mu_{t+1|t} + z_j\sqrt{w'\Sigma_{t+1|t}w}$, for $j = 1, \ldots, j^*$, where z_j is a random draw from the $SKST(0, I_k, \xi, \upsilon)$ distribution. By definition, the one-step-ahead VaR at $\alpha\%$ is defined as the empirical quantile at $\alpha\%$ of r_j^* (over the j^* simulations), see Jorion (2000) for a discussion of Monte Carlo techniques in VaR applications. In the empirical application, we set $j^* = 100,000$.

We use an iterative procedure where the DCC model is estimated to predict the one-day-ahead VaR of several portfolios (we used both the one-step and two-step estimation methods described above but only report the results concerning the second approach since it is much easier to implement and the results hardly differ). The first estimation sample is the complete sample for which the data is available less the last 1000 observations (about four years). The predicted one-day-ahead VaR (both for long and short positions) is then compared with the observed return and both results are recorded for later assessment using a statistical test. At the *i-th* iteration where i goes from 2 to 1000, the estimation sample is augmented to include one more day and the VaRs are forecasted and recorded. Whenever i is a multiple of 50 (about two months), the model is re-estimated. In other words, we update the model parameters every 50 trading days. We iterate the procedure until all days except the last one have been included in the estimation sample. Corresponding failure rates are then computed by comparing the long and short forecasted VaR with the observed portfolio return for all days in the four year period.

Because the comparison of the VaR forecast and the portfolio realization defines a sequence of yes/no observations, it is possible to test $H_0 : f = \alpha$ against $H_1 : f \neq \alpha$, where f is the failure rate (estimated by $\widehat{f}$, the empirical failure rate). Under the null hypothesis, the Kupiec likelihood ratio statistic $LR = 2\ln[\hat{f}^N(1-\hat{f})^{J-N}] - 2\ln[\alpha^N(1-\alpha)^{J-N}]$ is asymptotically distributed as a $\chi^2(1)$, where N is the number of VaR violations, J is the total number of observations (1000 in our case) and α is the failure rate of the null hypothesis (see Kupiec 1995). The LR statistic is computed both for long and short trading positions.

[INSERT TABLE 2 HERE.]

In Table 2 we first present complete VaR results (i.e. P-values of the LR statistics) for the following three portfolios composed of the stocks AA, CAT and DIS:

- 1/3 of AA, 1/3 of CAT and 1/3 of DIS (equally weighted portfolio);
- 50% of AA, 20% of CAT and 30% of DIS;
- and 140% of AA, -20% of CAT and -20% of DIS.

All models are tested with a VaR level α of 5, 2.5 and 1% and their performance is then assessed by computing the failure rate for several portfolios. Estimations have been done using a preliminary version of MG@RCH 1.0, an Ox package dedicated to the estimation and forecast of multivariate GARCH models.

While the three distributions perform well in predicting the VaR for long positions (top of the table), the empirical evidence is very much in favor of the skew-Student density when we consider the short positions (bottom of the table). Indeed, if we use a 5% level for the Kupiec LR statistic, we reject the null hypothesis about the true failure rate in 5 cases (in bold in the table) for the normal, in 3 cases for the Student, while we never reject the skew-Student density. To summarize and compare the results of each model, a performance measure, called

Grade(5%), reports the percentage of P-values above the 5% critical value (by considering both the long and short trading positions). For instance, since we have 5 rejections for the normal out of 18 cases, its *Grade*(5%) is 72%, compared to 82% for the Student and 100% for the skew-Student.

Next, we do the same exercise on the exchange rate data. We assume that someone is interested in investing in the three currencies (thus we estimate as previously a trivariate model), or in pairs (then we estimate bivariate models). Doing so, we end up with four models and several weight vectors (five for the bivariate cases and three for the trivariate case).

Table 3 presents the performance measure *Grade*(5%) for various weights. In the bivariate cases (first three lines) we consider five weight vectors: (0.5,0.5), (0.2,0.8), (0.1,0.9), (0.0, 1.0) and (-0.1, 1.1). In the trivariate case (last line), the weight vectors are (0.1,0.8,0.1), (-0.1,1.2,-0.1) and (0.2,0.6,0.2).

Broadly speaking, these results are quite similar to those obtained for the stocks since the skew-Student density is never beaten by its symmetric counterparts.

[INSERT TABLE 3 HERE.]

4 CONCLUSION

It is well accepted that high-frequency financial time series are heteroscedastic and leptokurtic, and that volatilities are related over time across assets and markets. Moreover, asset returns may also be skewed, a feature that is incompatible with the choice of a symmetric density for the innovations of a multivariate GARCH model.

To improve modelling, we propose a practical and flexible method to introduce skewness in multivariate symmetric distributions. By introducing a vector of skewness parameters, the new class of distributions brings additional flexibility for modelling time series of asset returns with multivariate volatility models. Applying the procedure to the multivariate Student density leads to a "multivariate skew-Student" density, in which each marginal has a different asymmetry coefficient. An easy variant provides a multivariate skew-density that can have different tail properties on each coordinate. As a by-product, a new copula density is obtained.

The usefulness of the multivariate skew-Student density is highlighted throughout a VaR application on several portfolios of assets and currencies. We show that in several cases, this density improves the quality of out-of-sample VaR forecasts by comparison with a symmetric one, and that in no case the performance is deteriorated.

The increasing availability of databases providing the intraday prices of financial assets (stocks, stock indexes, bonds, currencies, ...) has led to new developments in applied econometrics and quantitative finance. To reduce the dimension problem inherent to most multivariate volatility models, Andersen, Bollerslev, Diebold and Labys (2003) propose to compute the daily realized covariance matrix by summing intraday cross-products of returns. Staying in the spirit of the DCC model, they propose to forecast separately each element of the realized covariance matrix (or its Choleski decomposition) using a so-called ARFIMA model (see Sowell 1992 among others). This method seems empirically justified since we have seen in our applications that the DCC was almost integrated, suggesting the presence of long-memory and/or structural change in the conditional correlation (see Diebold and Inoue 2001, Gourieroux and Jasiak 2001 and Granger and Hyung 1999 for a discussion about integration vs. long-memory and/or regime switching). Interestingly, Giot and Laurent (2004) have shown, in a univariate framework, how to compute the VaR from the daily realized volatility forecast. Furthermore they conclude that this method compares very well with the usual daily approach based on a GARCH-type model and that using an appropriate distribution is crucially important *in both cases* to obtain good VaR forecasts. We conjecture that this result will be transferred to the multivariate case, so that a high dimensional VaR problem could be solved without imposing strong restrictions on both the covariance structure and the distribution of the innovations. We leave this topic for future research.

Acknowledgments

While remaining responsible for any error in this paper, we thank the participants of the 2002 Winter Meeting of the Econometric Society in Budapest and of the ESEM meeting in Venice. We are especially grateful to Christian Gouriéroux, Gaëlle Le Fol, Bernard Lejeune, Huston McCulloch, Franz Palm, Jeroen Rombouts, Olivier Scaillet, Jean-Pierre Urbain, the editor (Alastair Hall), the anonymous associate editor and two referees for helpful comments and suggestions on the initial or revised versions.

Sébastien Laurent acknowledges the financial support provided through the European Community's Human Potential Programme under contract HPRN-CT-2002-00232, "Microstructure of Financial Markets in Europe."

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

References

Alexander, C.O. (2001) *A Practitioners Guide to Financial Data Analysis,* J. Wiley & Sons.

Alexander, C.O., and Lazar, E. (2003), "Symmetric Normal Mixture GARCH," ISMA Center Discussion Papers in Finance, 2003-09.

Andersen, T.G., Bollerslev, T., Diebold, F.X., and Labys, P. (2003), "Modeling and Forecasting Realized Volatility," *Econometrica,* 71, 579–626.

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

Bauwens, L., and Laurent, S. (2002), "A New Class of Multivariate Skew Densities, with Application to GARCH Models," CORE DP 2002/20.

Bauwens, L., Laurent, S., and Rombouts, J. (2003), "Multivariate GARCH Models: a Survey," CORE DP 2003/31.

Bollerslev, T. (1987), "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics,* 69, 542–547.

—— (1990), "Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *Review of Economics and Statistics,* 72, 498–505.

Branco, M.D., and Dey, D.K. (2001), "A Class of Multivariate Skew-Elliptical Distributions," *Journal of Multivariate Analysis,* 79, 99–113.

Diebold, F.X., and Inoue, A. (2001), "Long Memory and Regime Switching," *Journal of Econometrics,* 105, 131–159.

Doornik, J. A. (2001) *Ox: Object Oriented Matrix Programming, 3.0,* London: Timberlake Consultant Ltd.

Drèze, J.H. (1978), "Bayesian Regression Analysis Using Poly-t Densities," *Journal of Econometrics,* 6, 329–354.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica,* 50, 987–1007.

—— (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business and Economics Statistics,* 20, 339–350.

Engle, R.F., and González-Rivera, G. (1991), "Semiparametric ARCH Model," *Journal of Business and Economic Statistics,* 9, 345–360.

Engle, R.F., and Sheppard, K. (2001), "Theorical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," Mimeo, UCSD.

Fernández, C., and Steel, M.F.J. (1998), "On Bayesian Modelling of Fat Tails and Skewness," *Journal of the American Statistical Association,* 93, 359–371.

Fernandez, C., Osiewalski, J., and Steel, M.F.J. (1995), "Modeling and Inference with v-Spherical Distributions," *Journal of the American Statistical Association,* 90, 1331–1340.

Ferreira, J.T.A.S., and Steel, M.F.J. (2003), "Bayesian Multivariate Regression Analysis with a New Class of Skewed Distributions," Statistics Research Report 419, University of Warwick.

Fiorentini, G., Sentana, E., and Calzolari, G. (2003), "Maximum Likelihood Estimation and Inference in Multivariate Conditionally Heteroskedastic Dynamic Regression Models with Student t Innovations," *Journal of Business and Economic Statistics,* 21, 532–546.

Geweke, J., and Amisano, G. (2001), "Compound Markov Mixture Models with Application in Finance," Mimeo, University of Iowa.

Giot, P., and Laurent, S. (2003), "Value-at-Risk for Long and Short Positions," *Journal of Applied Econometrics,* 18, 641–664.

—— (2004), "Modelling Daily Value-at-Risk Using Realized Volatility and ARCH Type Models," *Journal of Empirical Finance,* 11, 379–398.

Glosten, L.R., Jagannathan, R., and Runkle, D.E. (1993), "On the Relation Between Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance,* 48, 1779–1801.

Gourieroux, C. (1997) *ARCH Models and Financial Applications,* Springer Series in Statistics.

Gourieroux, C., and Jasiak, J. (2001), "Memory and Infrequent Breaks," *Economics Letters,* 70, 2941.

Granger, C.W.J., and Hyung, N. (1999), "Occasional Structural Breaks and Long Memory," UCSD Discussion Paper 99-14, June 1999.

Haas, M., Mittnik, S., and Paolella, M. (2002), "Mixed Normal Conditional Heteroskedasticity," Center for Financial Studies Working Paper 2002/10.

Jeantheau, T. (1998), "Strong Consistency of Estimators for Multivariate ARCH Models," *Econometric Theory,* 14, 70–86.

Johnson, M.E. (1987) *Multivariate Statistical Simulation,* Wiley.

Jones, M.C. (2001), "Multivariate T and Beta Distributions Associated with the Multivariate F Distribution," *Metrika,* 54, 215–231.

—— (2002), "Marginal Replacement in Multivariate Densities, with Application to Skewing Spherically Symmetric Distributions," *Journal of Multivariate Analysis,* 81, 85–99.

Jorion, P. (2000) *Value-at-Risk: The New Benchmark for Managing Financial Risk,* McGraw-Hill.

Kupiec, P. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives,* 2, 173–84.

Lambert, P., and Laurent, S. (2001), "Modelling Financial Time Series Using GARCH-Type Models and a Skewed Student Density," Mimeo, Université de Liège.

Laurent, S., and Peters, J.-P. (2002), "G@RCH 2.2 : An Ox Package for Estimating and Forecasting Various ARCH Models," *Journal of Economic Surveys,* 16, 447–485.

Lee, S.W., and Hansen, B.E. (1994), "Asymptotic Properties of the Maximum Likelihood Estimator and Test of the Stability of Parameters of the GARCH and IGARCH Models," *Econometric Theory,* 10, 29–52.

Lumsdaine, R.L. (1996), "Asymptotic Properties of the Quasi Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models," *Econometrica,* 64, 575–596.

Mauleón, I., and Perote, J. (1999), "Estimation of Multivariate Densities with Financial Data: the Performance of the Multivariate Edgeworth-Sargan Density," Proceedings of the 12th Australian Finance and Banking Conference, Sidney.

Mittnik, S., and Paolella, M.S. (2000), "Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates," *Journal of Forecasting,* 19, 313–333.

Peiró, A. (1999), "Skewness in Financial Returns," *Journal of Banking and Finance,* 23, 847–862.

Richard, J.-F., and Tompa, H. (1980), "On the Evaluation of Poly-t Density Functions," *Journal of Econometrics,* 12, 335–351.

Ruiz, E., and Pascual, L. (2002), "Bootstrapping Financial Time Series," *Journal of Economic Surveys,* 16, 271–300.

Sahu, S.K., Dey, D.K., and Branco, D. (2003), "A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models," *Canadian Journal of Statistics,* 31, 129–150.

Samorodnitsky, G., and Taqqu, M.S. (1994) *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance,* London: Chapman and Hall.

Sowell, F. (1992), "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models," *Journal of Econometrics,* 53, 165–188.

Tse, Y.K., and Tsui, A.K.C (2002), "A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations," *Journal of Business and Economics Statistics,* 20, 351–362.

Vlaar, P.J.G., and Palm, F.C. (1993), "The Message in Weekly Exchange Rates in the European Monetary System: Mean

Reversion, Conditional Heteroskedasticity and Jumps,"
Journal of Business and Economic Statistics, 11, 351–360.

Tables

Table 1: Estimation results for trivariate models

	AA-CAT-DIS	EUR-YEN-GBP
β	0.9846	0.9684
	(0.0047)	(0.0037)
	[0.9837]	[0.9689]
α	0.0088	0.0303
	(0.0021)	(0.0033)
	[0.0095]	[0.0294]
$\log \xi_1$	0.1050	-0.0875
	(0.0257)	(0.0242)
	[0.0977]	[-0.0724]
$\log \xi_2$	0.0786	0.0987
	(0.0263)	(0.0253)
	[0.0698]	[0.0983]
$\log \xi_3$	0.0667	-0.0677
	(0.0276)	(0.0238)
	[0.0591]	[-0.0353]
υ	7.2858	6.1928
	(0.5335)	(0.3960)
	[7.4020]	[6.4896]
T	3113	3066
$LRT(CCC)$	58.41	895.62
$LRT(ST)$	34.58	33.45

Note: For each parameter, we report the one step ML estimate and below (in parentheses) it the corresponding standard error. The estimate of the two-step approach is reported between brackets (third rows). $LRT(CCC)$ and $LRT(ST)$ are likelihood ratio statistics respectively for the assumption of constant correlations and symmetry with respect to the Student density.

Table 2: VaR results for AA, CAT and DIS

w	α	Normal	Student	SKST
VaR for long positions				
(1/3,1/3,1/3)	5.0%	0.475	0.203	0.122
	2.5%	0.241	0.690	0.174
	1.0%	0.362	0.538	0.362
(0.5,0.2,0.3)	5.0%	0.770	1.000	0.159
	2.5%	0.689	0.689	0.429
	1.0%	0.232	0.538	0.362
(1.4,-0.2,-0.2)	5.0%	0.885	0.475	0.097
	2.5%	1.000	0.838	0.326
	1.0%	0.362	0.314	0.139
VaR for short positions				
(1/3,1/3,1/3)	5.0%	0.666	0.474	0.773
	2.5%	**0.014**	**0.036**	0.550
	1.0%	**0.005**	0.139	0.362
(0.5,0.2,0.3)	5.0%	0.257	0.093	0.320
	2.5%	0.122	0.241	0.326
	1.0%	**0.005**	0.139	0.362
(1.4,-0.2,-0.2)	5.0%	0.159	0.069	0.257
	2.5%	**0.008**	**0.023**	0.326
	1.0%	**0.002**	**0.043**	0.139
$Grade(5\%)$		72%	83%	100%

Note: Entries are P-values for the null hypothesis $f_l = \alpha$ (i.e. failure rate for the long trading positions is equal to α, top part of the table) and $f_s = \alpha$ (i.e. failure rate for the short trading positions is equal to α, bottom of the table). w is the weight vector of the portfolio. $Grade(5\%)$ summarizes the performance of each model and is defined as the percentage of P-values above the 5% critical value (both for long and short trading positions).

Table 3: $Grade(5\%)$ for the exchange rates

Series	Normal	Student	SKST
EUR/YEN	90%	97%	97%
EUR/GBP	70%	67%	87%
GBP/YEN	67%	67%	87%
EUR/YEN/GBP	67%	67%	89%

Note: In the bivariate cases (first three lines) we consider five weight vectors: (0.5,0.5), (0.2,0.8), (0.1,0.9), (0.0, 1.0) and (-0.1, 1.1). In the trivariate case (last line), the weight vectors are (0.1,0.8,0.1), (-0.1,1.2,-0.1) and (0.2,0.6,0.2).

Figure Titles and Legends

Figure 1: Contour of the bivariate $SKST\left(0, I_2, (1, 1.3), 6\right)$ density

Figure Artwork

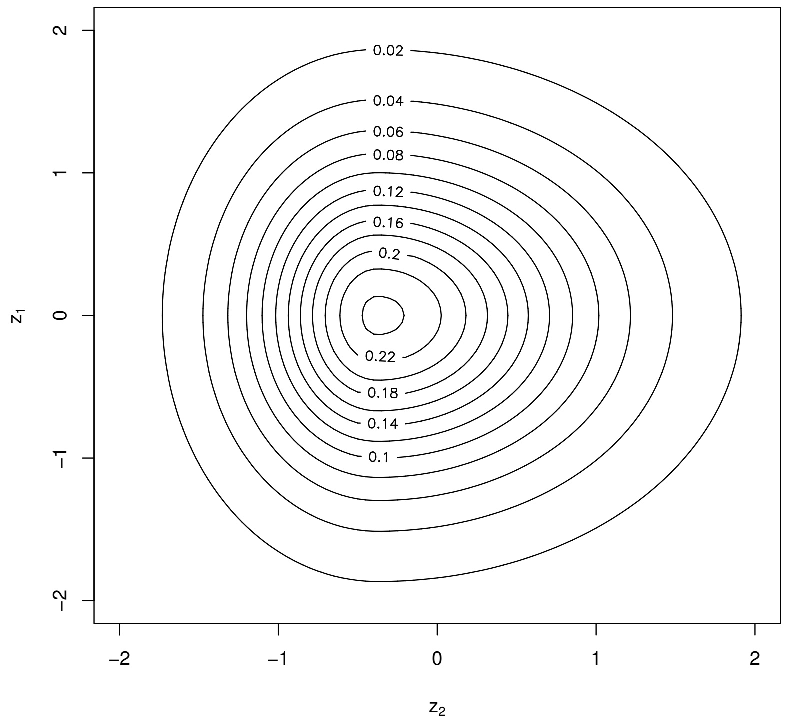

Figure 1: Contour of the bivariate $SKST\,(0, I_2, (1, 1.3), 6)$ density