# Multivariate Generalized Distributions with an Application to Pricing Time-Varying Risk

G. C. Lim, Thomas H. McCurdy and Vance L. Martin*

September 1999

**Abstract**

A new class of multivariate distributions is proposed which extends the univariate class of generalized exponential distributions. This family of distributions is used to develop new classes of multivariate GARCH models and applied to a single index conditional capital asset pricing model over the period 1988 to 1995. Special attention is given to extending the standard multivariate conditional distributions of normality to higher order moments which can admit both skewness and kurtosis. This new family of distributions can also exhibit multimodality, which is a property that is found to be important in the empirical application as it provides an alternative explanation of periods when there is an increase in risk. A further advantage of this class of models is that a natural parameterization of the distribution to achieve positive definiteness of the conditional covariance matrices is possible which is simpler than that used in existing multivariate GARCH models.

*The authors are respectively at the University of Melbourne, University of Toronto, and the University of Melbourne. All correspondence should be directed to Martin at the University of Melbourne, Economics Department, Parkville, Victoria, Australia, 3052.

# 1 Introduction

A new class of multivariate distributions is proposed which extends the univariate class of generalized exponential distributions investigated by Cobb, Koppstein and Chen (1983) and Lye and Martin (1993). Special attention is given to investigating the properties of generalizations of the multivariate normal and Student t distributions.

Important features of this new family of distributions are its ability to model higher order moment behaviour for multivariate processes, including multivariate skewness, kurtosis and multimodality. These properties are exploited in the empirical application where a flexible framework is developed for pricing risk in international equity markets. The skewness and kurtosis properties are used to model the non-normal behaviour of financial returns whilst the multimodality. property provides a new interpretation of volatility in equity markets.

In formulating the model of risk, a new framework for modelling multivariate volatility processes is proposed. This model is referred to as transcendental GARCH as the moments are computed implicitly from the underlying distribution specified. An important advantage of this model is that it circumvents the need to impose constraints on the variance-covariance matrix to achieve positive-definiteness as is the case with the model suggested by Engle and Kroner (1995).

The rest of the paper proceeds as follows. The multivariate generalized distributions are derived in Section 2. Two special subordinate distributions are discussed which are referred to as the multivariate generalized normal and generalized Student t distributions. Section 3 discusses the transcendental GARCH model which is a multivariate extension of GARCH models based on the proposed generalized distributions for modelling time-varying moments. The multivariate transcendental GARCH model is illustrated in Section 4 where it is applied to pricing risk based on a single index conditional CAPM between the NYSE excess returns and the excess returns of a world portfolio. The results are compared to the estimates based on the BEKK framework assuming conditional normality. The key empirical result is that the transcendental GARCH model identifies periods of increased risk when the underlying distribution exhibits bimodality. The BEKK model, in general, correctly identifies these periods of increased risk, but because it is based on conditional normality it fails to identify that the underlying cause of changes in risk is the result of changes in the third and fourth moments. This empirical result has implications for the CAPM as it highlights a nonlinear relationship between the market portfolio and a particular asset during periods of high volatility. It also shows that increases in risk can arise from periods of bimodality when returns switch between the stable modes of the distribution. Thus, while in tranquil periods the use of the second moment to model risk may suffice, the CAPM will be misspecified at least, in more turbulent periods when higher order moments are needed to model risk. Section 5 contains some concluding comments.

# 2 Generalized Multivariate Distributions

In this section, new families of multivariate distributions are derived from multivariate analogues of the generalized exponential family. The univariate generalizations are initially reviewed to provide the motivation for the solution of the multivariate distributions. Much of the discussion is devoted to analyzing bivariate distributions as this is the appropriate framework that is needed for the empirical analysis of the single index CAPM investigated below. Higher dimensional generalizations are not discussed, although in most cases the forms of these generalizations will be fairly obvious from the two dimensional distributions.

## 2.1 Univariate Generalizations

The univariate generalizations discussed here are based on the work of Cobb, Koppstein and Chen (1983) and Lye and Martin (1993), which represent extensions of the Pearson family. The Pearson family is derived from the differential equation; see Kendall and Stuart (1969, p.148)

$$\frac{d\ln f}{dy} = \frac{-\left(\alpha_0 + \alpha_1 y\right)}{\left(\beta_0 + \beta_1 y + \beta_2 y^2\right)}, \tag{1}$$

where $f(y)$ represents the distribution, and $\{\alpha_0, \alpha_1, \beta_0, \beta_1, \beta_2\}$ are the parameters which control the characteristics of the distribution. For example, the numerator controls the number of modes, which for this class of densities is one or less. The denominator controls the type of density. Two density types which are important in empirical ARCH models is the normal distribution, $\beta_1 = \beta_2 = 0$, and the Student t distribution, $\beta_1 = 0$ and $\beta_2 = 1$. The general solution to (1) is

$$f(y) = \exp\left[-\int\left(\frac{\alpha_0 + \alpha_1 s}{\beta_0 + \beta_1 s + \beta_2 s^2}\right) ds - \eta\right], \qquad y \in D, \tag{2}$$

where the normalizing constant is given by

$$\eta = \ln \int \exp\left[-\int\left(\frac{\alpha_0 + \alpha_1 s}{\beta_0 + \beta_1 s + \beta_2 s^2}\right) ds\right] dy. \tag{3}$$

The domain $D$, of $f(y)$ in (2) is the open interval where $\beta_0 + \beta_1 s + \beta_2 s^2 > 0$.

An obvious way to generalize (1) is to extend the polynomial in the numerator to higher order terms

$$\frac{df}{dy} = \frac{-\sum_{i=0}^{M-1} \alpha_i y^i f(y)}{\left(\beta_0 + \beta_1 y + \beta_2 y^2\right)}. \tag{4}$$

Analogous to the normal distribution, the generalized normal is given by setting $\beta_1 = \beta_2 = 0$. Solving the differential equation in (4) gives the generalized normal distribution; see Cobb, Koppstein and Chen (1983)

$$f(y) = \exp\left[\sum_{i=1}^{M} \Theta_i y^i - \eta^{GN}\right], \tag{5}$$

where the normalizing constant is given by

$$\eta^{GN} = \ln \int_{y \in D} \exp\left[\sum_{i=1}^{M} \Theta_i y^i\right] dy, \tag{6}$$

and the parameters $\Theta_i, i = 1, 2, ..., M$, are functions of the parameters $\{\alpha_i, i = 0, 1, ..., M-1; \beta_0\}$.[1] This distribution allows for skewness, kurtosis and other characteristics by the inclusion of the higher order terms. All moments of the generalized normal distribution exist provided that $\Theta_M < 0$, where $M$ is a positive and even integer. For the normal distribution $M = 2$, and the distribution is summarized by the first two moments. For the generalized normal given by (5), the distribution is summarized by its first $M$ moments.

A subordinate class of the generalized exponential family of distributions that is found to be useful in modelling financial data is the generalized Student t distribution; see Lye and Martin (1993), Lim, Lye, Martin and Martin (1998), as well as Creedy, Lye and Martin (1996). This distribution is derived by setting $\beta_1 = 0$ and $\beta_2 = 1$, in (4). The general form of the distribution is

$$f(y) = \exp\left[\Theta_1 \tan^{-1}\left(\frac{y}{\beta_0}\right) + \Theta_2 \ln\left(\beta_0^2 + y^2\right) + \sum_{i=3}^{M} \Theta_i y^i - \eta^{GS}\right], \tag{7}$$

where $\Theta_i, i = 1, 2, ..., M$, are functions of the parameters $\{\alpha_i, i = 0, 1, ..., M-1; \beta_0\}$ and $\eta^{GS}$ is the normalizing constant. A special case of the generalized Student t distribution is the Student t distribution, which is obtained by setting $\Theta_1 = \Theta_3 = ... = \Theta_M = 0$ and $\Theta_2 = -0.5\left(1 + \beta_0^2\right)$. For this model, $\beta_0^2$ is commonly referred to as the degrees of freedom parameter.

## 2.2 Multivariate Generalizations

The univariate families discussed above all derive from the Pearson differential equation (1), or its generalization (4). The multivariate analogue of these differential equations are partial differential equations. To highlight the key features of the derivations of the multivariate distributions, the bivariate generalized normal distribution followed by the bivariate generalized Student t distribution are discussed in turn.

[1] It is not possible to identify all of the parameters $\{\alpha_i, i = 0, 1, ..., M-1; \beta_0\}$, from the distribution parameters $\Theta_i, i = 1, 2, ..., M$, unless some identifying restriction is imposed. Given that the focus of attention with the present paper is with the distribution parameters $\Theta_i, i = 1, 2, ..., M$, the form of the identifying restriction is not of concern here.

### 2.2.1 Generalized Normal

Consider the following bivariate normal distribution

$$f(x, y) = \exp\left[\Theta_1 xy + \Theta_2 x^2 + \Theta_3 y^2 - \eta^{BN}\right], \tag{8}$$

where the normalizing constant is

$$\eta^{BN} = \ln \int\int \exp\left[\Theta_1 xy + \Theta_2 x^2 + \Theta_3 y^2\right] dxdy. \tag{9}$$

This implies that the corresponding pair of partial differential equations is

$$\frac{\partial \ln f}{\partial x} = \Theta_1 y + 2\Theta_2 x \tag{10}$$

$$\frac{\partial \ln f}{\partial y} = \Theta_1 x + 2\Theta_3 y. \tag{11}$$

Notice that this bivariate system of equations contains a cross-equation restriction as the parameter $\Theta_1$, appears in both equations. It is this restriction which enables a closed form solution to be derived. The parameters in (8) have the following interpretations: $\Theta_1$ controls the covariance between $x$ and $y$, and $\Theta_2$ and $\Theta_3$ control the respective variances of $x$ and $y$.

Motivated by the generalizations of the Pearson univariate system discussed above, a natural generalization of the bivariate normal distribution in (8) and hence in equations (10) to (11), is to include higher order terms in $x$ and $y$. Of course, while an infinite number of parameterizations can be entertained, it is necessary to choose a parameterization which not only achieves flexibility, but is also parsimonious. One generalization that is implemented in the empirical section below which achieves both properties is the following multivariate generalized normal distribution

$$f(x, y) = \exp\left[\Theta_1 xy + \Theta_2 x^2 + \Theta_3 y^2 + \Theta_4 x^4 + \Theta_5 y^4 - \eta^{MGN}\right], \tag{12}$$

where $\eta^{MGN}$ is the normalizing constant of the multivariate generalized normal distribution. All moments of this density exist provided that $\Theta_4, \Theta_5 < 0$. It is worth emphasizing at this stage, that these moment conditions are independent of the numerical values of the remaining parameters.[2] The implied pair of partial differential equations is

$$\frac{\partial \ln f}{\partial x} = \Theta_1 y + 2\Theta_2 x + 4\Theta_4 x^3 \tag{13}$$

$$\frac{\partial \ln f}{\partial y} = \Theta_1 x + 2\Theta_3 y + 4\Theta_5 y^3. \tag{14}$$

[2]Notice that (12) excludes the third order powers as well as intermediate cross-product terms with the exception of $xy$. These additional terms could be included, but they would not change the qualitative characteristics of this distribution. More formally, there exists a diffeomorphic transformation from this more general form of the distribution to the form given by (12).

Features of the bivariate generalized normal distribution are highlighted in Figure 1. The bivariate normal distribution is given in (a) which serves as the benchmark distribution. Setting $\Theta_4 = \Theta_5 = -0.25$, in (b) has the effect of making the distribution more peaked and less dispersed. Now changing the signs of $\Theta_2$ and $\Theta_3$, in (c) causes the distribution to become multimodal. Figure 1 (d) shows that increasing $\Theta_2$ and $\Theta_3$ from $\Theta_2 = \Theta_3 = 0.5$ to $\Theta_2 = \Theta_3 = 1.5$, increases the peakedness of the modes. Asymmetries are introduced by changing $\Theta_1$: (e) and (f) show that $\Theta_1$ controls the relative heights of the modes.

### 2.2.2 Generalized Student t

The main characteristic of the generalized Pearson differential equation in (4) that identifies the univariate generalized normal and generalized Student t distributions, is the form of the denominator: for the generalized normal it is a constant, whereas for the generalized Student t it is a quadratic. This suggests that multivariate generalized Student t distributions can be formed by writing down partial differential equations as in (13) and (14), but with the addition of polynomial terms in the denominator.

In choosing a particular parameterization of the distribution, there is no reason why the choice of the polynomials in the denominators have to be the same. As an example, consider the following pair of partial differential equations

$$\frac{\partial \ln f}{\partial x} = \frac{\Theta_1 y}{1+x^2y^2} + \frac{2\Theta_2 x}{1+x^2} + 2\Theta_4 x \tag{15}$$

$$\frac{\partial \ln f}{\partial y} = \frac{\Theta_1 x}{1+x^2y^2} + \frac{2\Theta_3 y}{1+y^2} + 2\Theta_5 y. \tag{16}$$

The solution is

$$\begin{aligned} f(x,y) &= \exp\Big[\Theta_1 \tan^{-1}(xy) + \Theta_2 \ln\left(1+x^2\right) + \Theta_3 \ln\left(1+y^2\right) \\ &\quad +\Theta_4 x^2 + \Theta_5 y^2 - \eta^{MGS}\Big], \end{aligned} \tag{17}$$

where $\eta^{MGS}$ is the normalizing constant of the multivariate generalized Student t distribution.[3] All moments of this density exist provided that $\Theta_4, \Theta_5 < 0$. This particular parameterization of the multivariate generalized Student t distribution is just one possibility out of a very large set of parameterizations; however, it proves to be useful in the CAPM empirical application.

To help understand the properties of the generalized Student t distribution, consider the special case $\Theta_1 = 0$ in (17). The joint distribution is now simply

[3] Another possible parameterization is to choose the ratio of the polynomials in (15) to (16) to yield a bivariate generalized Student t distribution which has as a special case the univariate generalized Student t distribution in (7) when $x = y$. This choice would be appropriate only for $x, y > 0$, as the distribution would need to include a term of the form $\tan^{-1}\left(\sqrt{xy}\right)$ in (17).

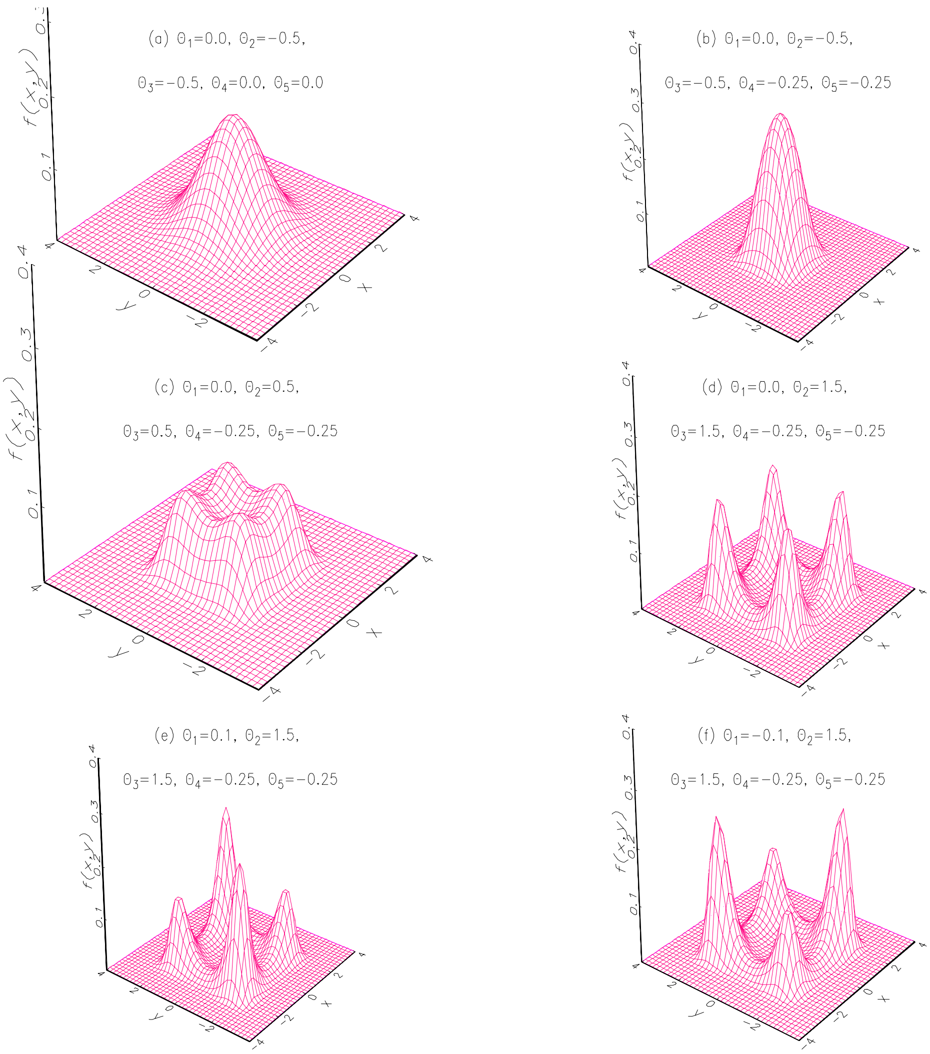


Figure 1: $f(x,y) = \exp\left[\Theta_1 xy + \Theta_2 x^2 + \Theta_3 y^2 + \Theta_4 x^4 + \Theta_5 y^4 - \eta^{MGN}\right]$

a product of two marginals with each marginal distribution being part of the univariate generalized Student t family. In particular, these marginals can be considered as a mixture of a normal distribution and a Student t distribution. This implies that $\Theta_1$ controls the strength of the dependence between the two random variables. Another way of understanding the properties of this form of the distribution is to note that by setting $\Theta_1 = 0$ in equations (15) and (16), this uncouples the partial differential equations thereby resulting in a system of ordinary differential equations which can be solved separately.

Further distributional characteristics of the bivariate generalized Student t in (17) are given in Figure 2. The bivariate normal distribution is given in Figure 2 (a). Setting $\Theta_2 = \Theta_3 = -1.0$, in (b) results in the distribution becoming more peaked and less dispersed. Changing the signs of $\Theta_2$ and $\Theta_3$, in (c) causes the distribution to become multimodal with equal heights associated with all modes. The relative heights of the modes change in (d) by setting $\Theta_1 = 0.5$. The distribution becomes bimodal in (e) for the parameterization $\Theta_2 = 0.0$, $\Theta_2 = 3.0$ and $\Theta_3 = -3.0$, with the modes centred on $x = 1$. Increasing $\Theta_1 = 0.0$ to $\Theta_1 = 5.0$ in (f), pushes the modes away from the $x = 0$ line.

## 2.3 Hypothesis Testing

An important feature of the generalized multivariate distributions discussed so far, is that they nest many of the standard multivariate distributions used in statistics. This has the advantage that hypothesis testing methods based on Lagrange multiplier tests for example, have standard asymptotic distributions under the null hypothesis. To highlight this feature a multivariate test of normality is derived based on the generalized bivariate normal distribution and is shown to be related to existing tests of multivariate normality proposed in the literature.

Consider the following generalized bivariate normal distribution

$$f(x, y) = \exp\left[\Theta_1 x^2 + \Theta_2 y^2 + \Theta_3 x^2 y^2 - \eta\right], \tag{18}$$

where $\eta$ is the normalizing constant. The random variables are assumed to be standardized to have zero means and unit variance-covariance matrix. The degree of nonnormality is controlled by $\Theta_3$, thus the null hypothesis is $H_0 : \Theta_3 = 0$.

For a sample of $t = 1, 2, ..., T$, the log-likelihood is given by

$$\ln L = \sum_t \left[\Theta_1 x_t^2 + \Theta_2 y_t^2 + \Theta_3 x_t^2 y_t^2 - \eta_t\right]. \tag{19}$$

This yields the following scores

$$\begin{aligned} \frac{\partial \ln L}{\partial \Theta_1} &= \sum_t x_t^2 - T\frac{\partial \eta_t}{\partial \Theta_1} \\ &= \sum_t x_t^2 - TE\left[x_t^2\right] \end{aligned}$$

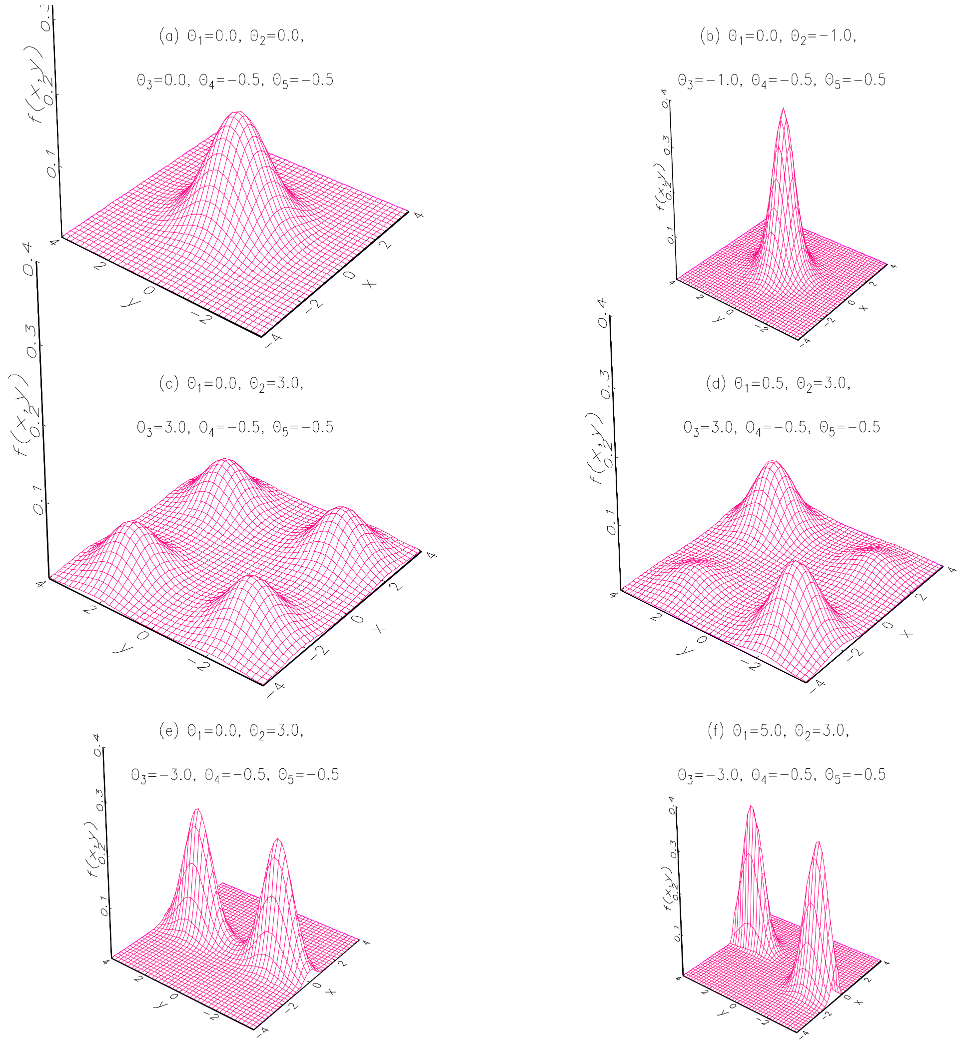


Figure 2: See equation (17) for the form of the generalized Student t distribution.

$$
\begin{aligned}
\frac{\partial \ln L}{\partial \Theta_2} &= \sum_t y_t^2 - T\frac{\partial \eta_t}{\partial \Theta_2} \\
&= \sum_t y_t^2 - TE\left[y_t^2\right] \\
\frac{\partial \ln L}{\partial \Theta_3} &= \sum_t x_t^2 y_t^2 - T\frac{\partial \eta_t}{\partial \Theta_3} \\
&= \sum_t x_t^2 y_t^2 - TE\left[x_t^2 y_t^2\right],
\end{aligned}
$$

The second derivatives are given by

$$
\begin{aligned}
\frac{\partial^2 \ln L}{\partial \Theta_1^2} &= -T\frac{\partial^2 \eta}{\partial \Theta_1^2} = -T\left(E\left[x_t^4\right] - E\left[x_t^2\right]E\left[x_t^2\right]\right) \\
\frac{\partial^2 \ln L}{\partial \Theta_2^2} &= -T\frac{\partial^2 \eta}{\partial \Theta_2^2} = -T\left(E\left[y_t^4\right] - E\left[y_t^2\right]E\left[y_t^2\right]\right) \\
\frac{\partial^2 \ln L}{\partial \Theta_3^2} &= -T\frac{\partial^2 \eta}{\partial \Theta_3^2} = -T\left(E\left[x_t^4 y_t^4\right] - E\left[x_t^2 y_t^2\right]E\left[x_t^2 y_t^2\right]\right) \\
\frac{\partial^2 \ln L}{\partial \Theta_1 \partial \Theta_2} &= -T\frac{\partial^2 \eta}{\partial \Theta_1 \partial \Theta_2} = -T\left(E\left[x_t^2 y_t^2\right] - E\left[x_t^2\right]E\left[y_t^2\right]\right) \\
\frac{\partial^2 \ln L}{\partial \Theta_1 \partial \Theta_3} &= -T\frac{\partial^2 \eta}{\partial \Theta_1 \partial \Theta_3} = -T\left(E\left[x_t^4 y_t^2\right] - E\left[x_t^2\right]E\left[x_t^2 y_t^2\right]\right) \\
\frac{\partial^2 \ln L}{\partial \Theta_2 \partial \Theta_3} &= -T\frac{\partial^2 \eta}{\partial \Theta_2 \partial \Theta_3} = -T\left(E\left[x_t^2 y_t^4\right] - E\left[y_t^2\right]E\left[x_t^2 y_t^2\right]\right).
\end{aligned}
$$

The moment generating function for a bivariate normal distribution with zero means and unit variance-covariance matrix is (see Kendall and Stuart, 1969, p.82)

$$
\begin{aligned}
M(t_1, t_2) &= \exp\left[\frac{1}{2}\left(t_1^2 + t_2^2\right)\right] \\
&= 1 + \frac{1}{2^1 \times 1!}\left(t_1^2 + t_2^2\right) + \frac{1}{2^2 \times 2!}\left(t_1^4 + t_2^4 + 2t_1^2 t_2^2\right) \\
&\quad + \frac{1}{2^3 \times 3!}\left(t_1^6 + 3t_1^4 t_2^2 + 3t_1^2 t_2^4 + t_2^6\right) \\
&\quad + \frac{1}{2^4 \times 4!}\left(t_1^8 + 4t_1^6 t_2^2 + 6t_1^4 t_2^4 + 4t_1^2 t_2^6 + t_2^8\right) \\
&\quad + \ldots
\end{aligned}
$$

Now

$$E\left[x_t^2 y_t^2\right] = 1,$$

as this is the coefficient of $\frac{t_1^2 t_2^2}{2!2!}$. Similarly, the coefficient of $\frac{t_1^2 t_2^4}{2!4!}$ is

$$E\left[x_t^2 y_t^4\right] = 3.$$

Finally, the coefficient of $\frac{t_1^4 t_2^4}{4!4!}$ is

$$E\left[x_t^4 y_t^4\right] = 9.$$

Under the null hypothesis of bivariate normality, the gradient vector is

$$G = \left[0, 0, \frac{\Sigma x_t^2 y_t^2}{T} - 1\right],$$

whilst the information matrix is

$$\begin{aligned} I &= -E\left[\frac{\partial^2 \ln L}{\partial \Theta_i \partial \Theta_j}\right]_{\Theta_3=0} \\ &= T\begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 2 \\ 2 & 2 & 8 \end{bmatrix}. \end{aligned}$$

Thus the Lagrange multiplier test reduces to

$$\begin{aligned} LM &= GI^{-1}G \\ &= \frac{T}{4}\left[\frac{\Sigma x_t^2 y_t^2}{T} - 1\right]^2, \end{aligned}$$

which is asymptotically distributed as $\chi_1^2$ under the null hypothesis.. An alternative form is

$$\frac{\frac{\Sigma x_t^2 y_t^2}{T} - 1}{4/T} \xrightarrow{a} N(0,1).$$

This is the same form of the multivariate normality test statistic proposed by Bera and John (1983).

# 3 Transcendental Multivariate GARCH

Multivariate GARCH models were initially studied by Engle, Granger and Kraft (1984) and Bollerslev, Engle and Wooldridge (1988). These earlier models displayed a diagonal representation whereby conditional second moments were expressed as functions of own lags and own residuals. To overcome the problem of achieving positive definite conditional covariance matrices while at the same time adopting a less restrictive parameterization than previous multivariate GARCH models, Engle and Kroner (1995) suggested an alternative parameterization which is characterized by a number of across and within equation restrictions on the parameters in the conditional variance and covariance equations. Letting $H_t$, be the conditional variance covariance matrix at time $t$, and $\varepsilon_t$, be the vector of residuals, this alternative representation is given for a GARCH(1,1) model as

$$H_t = C'C + A'\varepsilon_{t-1}\varepsilon_{t-1}' A + B' H_{t-1} B, \tag{20}$$

where $C$ is an upper triangular matrix, and $A$ and $B$, may or may not be symmetric matrices. The specification of the multivariate GARCH model is completed by specifying the conditional distribution as normal

$$\varepsilon_t | \, I_{t-1} \sim N(0, H_t), \tag{21}$$

where $I_{t-1}$, represents the information matrix at time $t-1$.

To begin the analysis, consider the following bivariate generalized normal distribution, which is a special case of (12)

$$f(\varepsilon_{1t}, \varepsilon_{2,t}) = \exp\left[\Theta_{1,t}\varepsilon_{1,t}\varepsilon_{2,t} + \Theta_{2,t}\varepsilon_{1,t}^2 + \Theta_{3,t}\varepsilon_{2,t}^2 - 0.25\varepsilon_{1,t}^4 - 0.25\varepsilon_{2,t}^4 - \eta_t^{MGN}\right], \tag{22}$$

where, as before, $\eta_t^{MGN}$ is the normalizing constant which is given by

$$\eta_t^{MGN} = \ln \int\int \exp\left[\Theta_{1,t}\varepsilon_{1,t}\varepsilon_{2,t} + \Theta_{2,t}\varepsilon_{1,t}^2 + \Theta_{3,t}\varepsilon_{2,t}^2 - 0.25\varepsilon_{1,t}^4 - 0.25\varepsilon_{2,t}^4\right] d\varepsilon_{1,t}\varepsilon_{2,t}. \tag{23}$$

From the above discussion of the properties of this distribution, by allowing the parameters $\Theta_{i,t}$, $i = 1, 2, 3$, to be time-varying, this distribution can allow for time-varying moments. In the context of GARCH modelling, one natural parameterization is

$$\begin{aligned}
\Theta_{1,t} &= \theta_{1,0} + \theta_{1,1}\varepsilon_{1,t-1}^2 + \theta_{1,2}\varepsilon_{2,t-1}^2 + \theta_{1,3}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \theta_{1,4}\Theta_{1,t-1} + \theta_{1,5}\Theta_{2,t-1} + \theta_{1,6}\Theta_{3,t-1} && (24) \\
\Theta_{2,t} &= \theta_{2,0} + \theta_{2,1}\varepsilon_{1,t-1}^2 + \theta_{2,2}\varepsilon_{2,t-1}^2 + \theta_{2,3}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \theta_{2,4}\Theta_{1,t-1} + \theta_{2,5}\Theta_{2,t-1} + \theta_{2,6}\Theta_{3,t-1} && (25) \\
\Theta_{3,t} &= \theta_{3,0} + \theta_{3,1}\varepsilon_{1,t-1}^2 + \theta_{3,2}\varepsilon_{2,t-1}^2 + \theta_{3,3}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\
&\quad + \theta_{3,4}\Theta_{1,t-1}. + \theta_{2,5}\Theta_{2,t-1} + \theta_{2,6}\Theta_{3,t-1}. && (26)
\end{aligned}$$

For each $\Theta_{i,t}$, the information is the same; namely it includes a constant, lagged squared residuals and lagged cross-products of residuals. The lags of the squares and cross-products of the residuals represents the ARCH part, whereas the lags of the $\Theta_{j,t}, j = 1, 2, 3$, captures the GARCH part. Another possible parameterization of the GARCH model which is a more parsimonious representation than the specifications in (24) to (26), is

$$\begin{aligned}
\Theta_{1,t} &= \theta_{1,0} + \theta_{1,1}\varepsilon_{1,t-1} + \theta_{1,2}\varepsilon_{2,t-1} + \theta_{1,3}\Theta_{1,t-1} && (27) \\
\Theta_{2,t} &= \theta_{2,0} + \theta_{2,1}\varepsilon_{1,t-1} + \theta_{2,2}\varepsilon_{2,t-1} + \theta_{2,3}\Theta_{2,t-1} && (28) \\
\Theta_{3,t} &= \theta_{3,0} + \theta_{3,1}\varepsilon_{1,t-1} + \theta_{3,2}\varepsilon_{2,t-1} + \theta_{3,3}\Theta_{3,t-1}. && (29)
\end{aligned}$$

In the empirical application, this form of the specification is adopted.

The expressions in (24) to (26) and (27) to (29) are not the conditional covariances, however they do affect the shape of the conditional distribution at each point in time and hence affect the conditional variances and covariances. The conditional second moments are given by

$$h_{11,t} = \int\int \varepsilon_{1,t}^2 \exp\left[\psi\left(\varepsilon_{1,t},\varepsilon_{2,t}\right)\right] d\varepsilon_{1,t} d\varepsilon_{2,t} \tag{30}$$

$$h_{22,t} = \int\int \varepsilon_{2,t}^2 \exp\left[\psi\left(\varepsilon_{1,t},\varepsilon_{2,t}\right)\right] d\varepsilon_{1,t} d\varepsilon_{2,t} \tag{31}$$

$$h_{12,t} = \int\int \varepsilon_{1,t}\varepsilon_{2,t} \exp\left[\psi\left(\varepsilon_{1,t},\varepsilon_{2,t}\right)\right] d\varepsilon_{1,t} d\varepsilon_{2,t}, \tag{32}$$

where

$$\psi\left(\varepsilon_{1,t},\varepsilon_{2,t}\right) = \Theta_{1,t}\varepsilon_{1,t}\varepsilon_{2,t} + \Theta_{2,t}\varepsilon_{1,t}^2 + \Theta_{3,t}\varepsilon_{2,t}^2 - 0.25\varepsilon_{1,t}^4 - 0.25\varepsilon_{2,t}^4 - \eta_t^{MGN}.$$

The form of the conditional distribution discussed so far has been couched in terms of the generalized normal distribution. Another choice of the generalized distribution is, for example, the multivariate generalized Student t distribution based on (17)

$$\begin{aligned} f(\varepsilon_{1t},\varepsilon_{2,t}) = {} & \exp\Big[\Theta_{1,t}\tan^{-1}\left(\varepsilon_{1,t}\varepsilon_{2,t}\right) + \Theta_{2,t}\ln\left(1+\varepsilon_{1,t}^2\right) + \Theta_{3,t}\ln\left(1+\varepsilon_{2,t}^2\right) \\ & -0.5\varepsilon_{1,t}^2 - 0.5\varepsilon_{2,t}^2 - \eta_t^{MGS}\Big], \end{aligned} \tag{33}$$

where, as before, $\eta_t^{MGS}$ is the normalizing constant which is given by

$$\begin{aligned} \eta_t^{MGS} = {} & \ln\int\int \exp\Big[\Theta_{1,t}\tan^{-1}\left(\varepsilon_{1,t}\varepsilon_{2,t}\right) + \Theta_{2,t}\ln\left(1+\varepsilon_{1,t}^2\right) + \Theta_{3,t}\ln\left(1+\varepsilon_{2,t}^2\right) \\ & -0.5\varepsilon_{1,t}^2 - 0.5\varepsilon_{2,t}^2 d\varepsilon_{1,t}\varepsilon_{2,t}\Big]. \end{aligned} \tag{34}$$

The expressions for computing the conditional variances and covariances in (30) to (32) remain the same except the bivariate generalized normal distribution is replaced by the bivariate generalized Student t distribution in (33).

The conditional covariances in (30) to (32) do not have a closed form expression which is in contrast with the existing class of multivariate GARCH models as represented by (20) and (21). For this reason the class of GARCH models introduced in this paper is referred to as transcendental GARCH.

An important advantage of the transcendental GARCH model is that the conditions to ensure that all moments of the underlying distribution exist are easily imposed and do not depend on time. In particular, by a simple application of the Cauchy-Schwarz inequality, the conditional variance-covariance matrix as given by (30) to (32) are positive definite at each point in time. The problem of estimating multivariate GARCH models under the assumption of normality while ensuring that the moments of the distribution exist, is highlighted by inspecting the bivariate normal distribution in (8). The functions $\Theta_{1,t}, \Theta_{2,t}$ and $\Theta_{3,t}$ serve two roles: the first is to represent the time-varying conditional variance and covariance specifications as given by (20), and the second, in the case of $\Theta_{2,t}$ and $\Theta_{3,t}$ at least, is to control the moments of the distribution. In particular, for moment existence the restrictions $\Theta_{2,t} < 0$ and $\Theta_{3,t} < 0$, are needed. The problem is that because these functions have to perform a dual role, conflicts can arise at particular points in time whereby for certain values of the GARCH parameters $\Theta_{2,t} > 0$ and $\Theta_{3,t} > 0$, resulting in the moments not existing. The adoption of the bivariate generalized distributions in (22) and (33) circumvent this problem by allocating the two roles to different $\Theta_{i,t}$ functions. For example, the restriction that the coefficients attached to $\varepsilon_{1,t}^2$ and $\varepsilon_{2,t}^2$, in (33) ensure moment existence, whilst the role of time-varying second moments is assigned to the remaining functions, $\Theta_{1,t}, \Theta_{2,t}, \Theta_{3,t}$.

Finally, it is not appropriate to compare the parameters in (27) to (29) with the parameters in the GARCH specifications of the BEKK estimator in (20) as they are not commensurate. One solution is to differentiate numerically (30) to (32) with respect to $\theta_{i,j}$.

# 4 Conditional CAPM and Estimates of Time-Varying Risk Premia

The multivariate transcendental GARCH model set out above is now applied to estimating time-varying risk premia in a single index international conditional CAPM. For comparative purposes, the estimates from using the BEKK model assuming conditional normality, are also reported.

## 4.1 Data

The data are excess returns on the NYSE ($R_{n,t}$) and MSCI ($R_{w,t}$) which represents the excess returns on a world index. The sample period is from the 3rd of February, 1988 to the 29th of December, 1995, a total of 2000 observations. The data are presented in Figure 3. The lag length in (35) and (36) is set at one, which makes the effective sample size equal to $T = 1999$ observations.

Some descriptive statistics on the two excess returns are given in Table 1. Inspection of the estimates of the first four moments shows that both excess returns exhibit very similar moments with the exception of the third moment: the world excess returns show no significant skewness whereas the NYSE excess returns do. Both excess returns show significant autocorrelation of various orders based on the Ljung-Box statistic ($Q_x(j)$).[4] The Ljung-Box test statistic applied

[4]The p-values of this test statistic should be interpreted as being suggestive of autocorrelation as it is not strictly correct in the presence of heterogeneity.

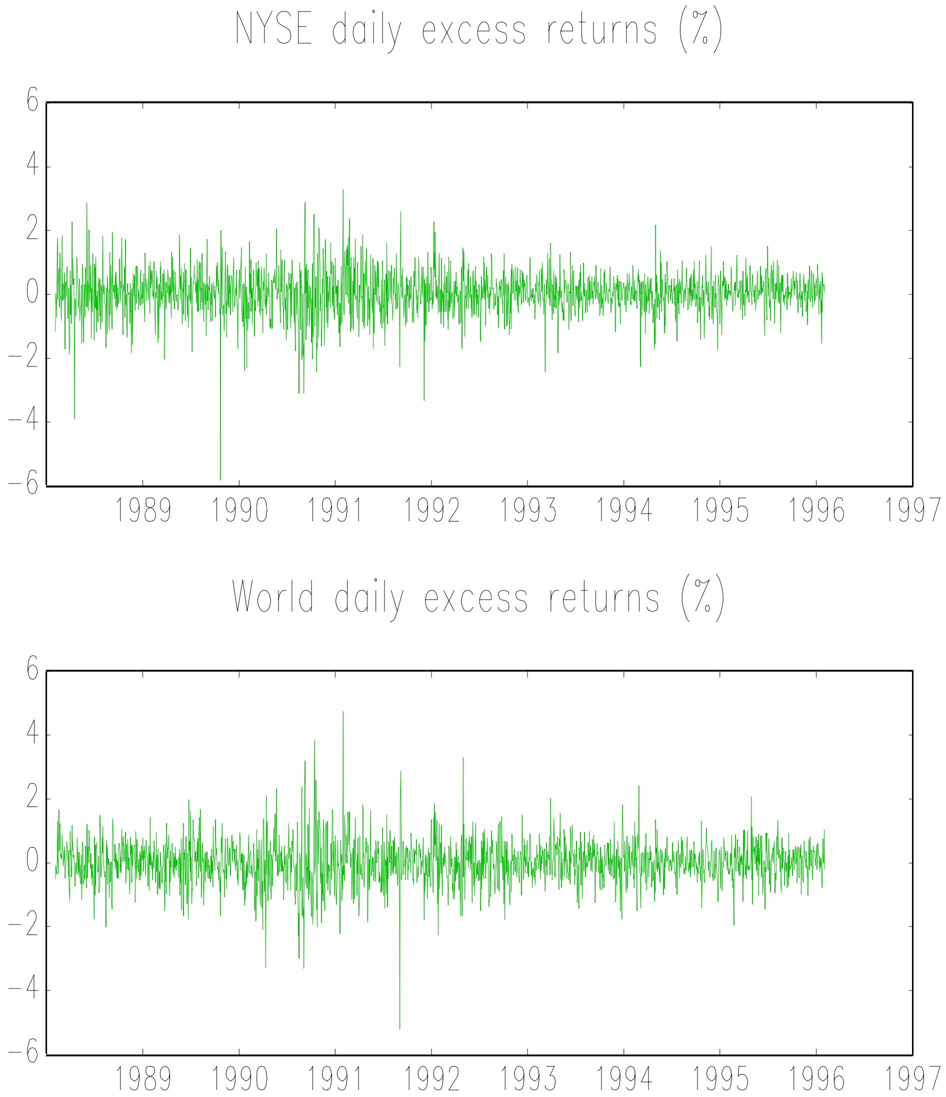


Figure 3: Daily excess returns (%): 1988 to 1995.

to the squared excess returns ($Q_{xx}(j)$) shows evidence of significant ARCH; see also Figure 4 which gives plots of the squares of the two excess returns.

## 4.2 Theory and Estimating Equations

Following McCurdy and Morgan (1993), the key equations of the single index international CAPM are

$$\begin{aligned} R_{w,t} &= \delta_0 + \delta_1 R_{w,t-1} + \varepsilon_{w,t} && (35) \\ R_{n,t} &= \gamma_0 + h_{nw,t}\left(\frac{\delta_0 + \delta_1 R_{w,t-1}}{h_{w,t}}\right) + \varepsilon_{n,t}, && (36) \end{aligned}$$

where $R_{w,t}$ and $R_{n,t}$, are the excess returns for the world portfolio and the NYSE index respectively, $\{\gamma_0, \delta_0, \delta_1\}$ are parameters, $h_{w,t}$ is the conditional variance of the world portfolio, and $h_{nw,t}$ is the conditional covariance between the NYSE and world portfolios excess returns. The information set is constrained to include just one lag. Other lags were tried but were found to be statistically insignificant.

For the transcendental GARCH model, the error vector $\varepsilon_t^{'} \equiv [\varepsilon_{w,t}, \varepsilon_{n,t}]$ is assumed to be either the conditional bivariate generalized normal distribution

$$f(\varepsilon_{w,t}, \varepsilon_{n,t}) = \exp\left[\Theta_{1,t}\varepsilon_{w,t}\varepsilon_{n,t} + \Theta_{2,t}\varepsilon_{w,t}^2 + \Theta_{3,t}\varepsilon_{n,t}^2 - 0.25\varepsilon_{w,t}^4 - 0.25\varepsilon_{n,t}^4 - \eta_t^{MGN}\right], \quad (37)$$

or the conditional bivariate Student t distribution

$$\begin{aligned} f(\varepsilon_{w,t}, \varepsilon_{n,t}) &= \exp\Big[\Theta_1 \tan^{-1}(\varepsilon_{w,t}\varepsilon_{n,t}) + \Theta_2 \ln\left(1 + \varepsilon_{w,t}^2\right) + \Theta_3 \ln\left(1 + \varepsilon_{n,t}^2\right) \\ &\quad -0.5\varepsilon_{w,t}^2 - 0.5\varepsilon_{n,t}^2 - \eta^{MGS}\Big]. && (38) \end{aligned}$$

The $\Theta_i$ functions of the joint distribution are assumed to have the following representations

$$\begin{aligned} \Theta_{1,t} &= \theta_{1,0} + \theta_{1,1}R_{w,t-1} + \theta_{1,2}R_{n,t-1} + \theta_{1,3}\Theta_{1,t-1} && (39) \\ \Theta_{2,t} &= \theta_{2,0} + \theta_{2,1}R_{w,t-1} + \theta_{2,2}R_{n,t-1} + \theta_{2,3}\Theta_{2,t-1} && (40) \\ \Theta_{3,t} &= \theta_{3,0} + \theta_{3,1}R_{w,t-1} + \theta_{3,2}R_{n,t-1} + \theta_{3,3}\Theta_{3,t-1}. && (41) \end{aligned}$$

The conditional variances and covariances are computed numerically at each $t$, as

$$\begin{aligned} h_{n,t} &= \int\int \varepsilon_{n,t}^2 f(\varepsilon_{w,t}, \varepsilon_{n,t}) d\varepsilon_{n,t} d\varepsilon_{w,t} && (42) \\ h_{w,t} &= \int\int \varepsilon_{w,t}^2 f(\varepsilon_{w,t}, \varepsilon_{n,t}) d\varepsilon_{n,t} d\varepsilon_{w,t} && (43) \\ h_{nw,t} &= \int\int \varepsilon_{n,t}\varepsilon_{w,t} f(\varepsilon_{w,t}, \varepsilon_{n,t}) d\varepsilon_{n,t} d\varepsilon_{w,t}. && (44) \end{aligned}$$

As noted above, the choice of the lag distributions is based on some preliminary analysis which found that lags greater than one were statistically insignificant.

Table 1:

Descriptive statistics of daily excess returns (%): 1988 to 1996.

| Statistic | NYSE | | World | |
|---|---|---|---|---|
| | Estimate | pv | Estimate | pv |
| Mean | 0.037 | | 0.011 | |
| Standard deviation | 0.676 | | 0.684 | |
| Maximum | 3.282 | | 4.749 | |
| Minimum | -5.841 | | -5.225 | |
| Skewness$^{(a)}$ | -0.538 | 0.000 | -0.039 | 0.482 |
| Kurtosis$^{(b)}$ | 8.288 | 0.000 | 7.719 | 0.000 |
| BJ$^{(c)}$ | 2426.473 | 0.000 | 1856.115 | 0.000 |
| $Q_x(1)^{(d)}$ | 14.541 | 0.000 | 59.901 | 0.000 |
| $Q_x(5)$ | 16.207 | 0.006 | 65.741 | 0.000 |
| $Q_x(10)$ | 31.348 | 0.001 | 75.374 | 0.000 |
| $Q_x(15)$ | 43.482 | 0.000 | 83.843 | 0.000 |
| $Q_x(20)$ | 50.363 | 0.000 | 87.639 | 0.000 |
| $Q_{xx}(1)^{(e)}$ | 12.427 | 0.000 | 52.903 | 0.000 |
| $Q_{xx}(5)$ | 31.852 | 0.000 | 155.704 | 0.000 |
| $Q_{xx}(10)$ | 53.012 | 0.000 | 199.139 | 0.000 |
| $Q_{xx}(15)$ | 78.462 | 0.125 | 232.236 | 0.000 |
| $Q_{xx}(20)$ | 89.496 | 0.235 | 249.812 | 0.000 |

(a) The skewness coefficient is computed as $\sum z_t^3/T$, where $z_t$ is the standardized excess return. The p-value (pv) is for a test of skewness with normality as the null.

(b) The kurtosis coefficient is computed as $\sum z_t^4/T$, where $z_t$ is the standardized excess return. The p-value (pv) is for a test of excess kurtosis with normality as the null.

(c) Bera-Jarque test of normality.

(d) $Q_x(j)$ is the Ljung-Box statistic for the first j lags of the autocorrelation function of excess returns.

(e) $Q_{xx}(j)$ is the Ljung-Box statistic for the first j lags of the autorcorrelation function of squared excess returns.

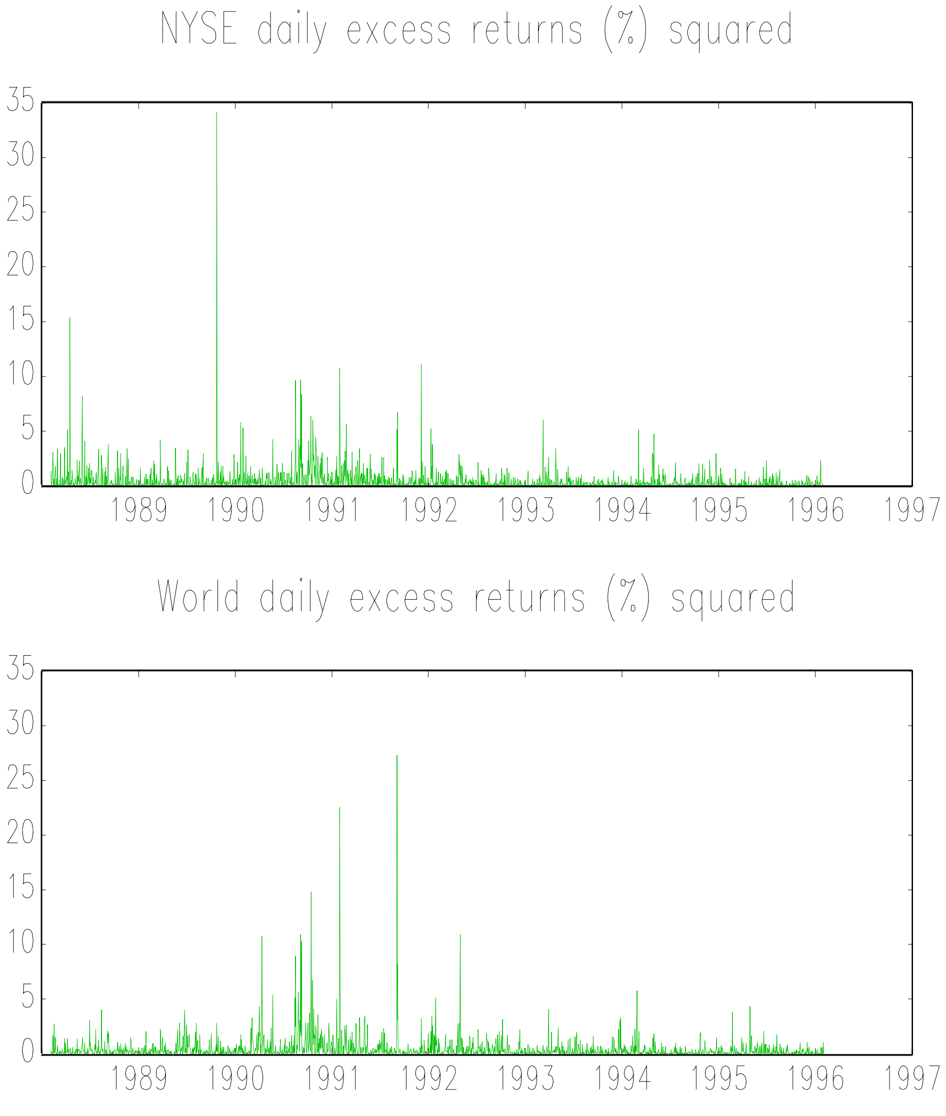


Figure 4: Daily excess returns (%) squared: 1988 to 1995.

### 4.3 Maximum Likelihood Estimation

As the class of multivariate generalized distributions considered here falls within the exponential family, maximum likelihood represents a natural estimation framework. For the bivariate generalized normal distribution, the logarithm of the likelihood at the $t^{th}$ observation is

$$\ln L_t = \Theta_{1,t}\varepsilon_{w,t}, \varepsilon_{n,t} + \Theta_{2,t}\varepsilon_{w,t}^2 + \Theta_{3,t}\varepsilon_{n,t}^2 - 0.25\varepsilon_{w,t}^4 - 0.25\varepsilon_{n,t}^4 - \eta_t^{MGN}, \qquad (45)$$

whereas for the bivariate generalized Student t distribution given in (38), it is

$$\begin{aligned} \ln L_t &= \Theta_{1,t}\tan^{-1}(\varepsilon_{w,t}\varepsilon_{n,t}) + \Theta_{2,t}\ln\left(1+\varepsilon_{w,t}^2\right) + \Theta_{3,t}\ln\left(1+\varepsilon_{n,t}^2\right) \\ &\quad -0.5\varepsilon_{w,t}^2 - 0.5\varepsilon_{n,t}^2 - \eta_t^{MGS}. \end{aligned} \qquad (46)$$

For a sample of $t = 1, 2, ..., T$, observations, the log of the likelihood is then

$$\ln L = \sum_{t=1}^{T} \ln L_t. \qquad (47)$$

This expression can be maximized with respect to the distribution parameters in (39) to (41), $\{\theta_{i,j}\}$, and the mean parameters $\{\gamma_0, \delta_0, \delta_1\}$ in (35) and (36).

The likelihood function (47), can be maximized using standard gradient optimization algorithms. As a result of the nonlinearities in the model, it is convenient to use numerical derivatives. In evaluating both the likelihood function and derivatives of the log likelihood, it is necessary to compute the normalizing constant, as well as the conditional covariances as given by (30) to (32). These bivariate integrals can be computed using standard Gaussian quadrature procedures. All computations are performed using the GAUSS computer language. The optimization algorithm chosen is the BFGS option in MAXLIK, and the integrations are calculated using the procedure INTQUAD2.

### 4.4 Empirical Results

The results of the transcendental GARCH model are given in Table 2 using both the generalized normal and Student t as the conditional distributions. Robust asymptotic t-statistics as based on the Hessian and the outer product of the gradients, are reported. Although the chosen parameterizations of the generalized normal and Student t distributions do not provide a suitable nesting of the two distributions, a comparison of the values of the log-likelihoods suggests that the generalized Student t distribution does a better job than the generalized normal distribution in capturing the characteristics of the data.[5] Given this result, the rest of the analysis concerning the transcendental GARCH model is based on

[5] More formally, by calculating information statistics based either on the AIC or the BIC, shows that these statistics are minimized for the bivariate generalized Student t distribution.

the generalized Student t distribution. Table 2 also shows that $\theta_{1,0}$ and $\theta_{1,1}$ are statistically significant for the generalized Student t distribution, showing that there is significant dependence between the two error terms.

For comparison, the error vector $\varepsilon_t' \equiv [\varepsilon_{w,t}, \varepsilon_{n,t}]$ in (35) and (36) is also modelled as a bivariate normally distributed process with a BEKK multivariate GARCH format given by (20), where the parameter matrices are specified as

$$C = \begin{bmatrix} c_{11} & \\ c_{21} & c_{22} \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}. \tag{48}$$

Parameter estimates of the BEKK model are given in Table 3. Two versions of this model are estimated. The first is the asymmetric model and the second is the symmetric model which is obtained by imposing in (48) the restrictions

$$a_{21} = a_{12}, \; b_{21} = b_{12}.$$

A likelihood ratio test of these restrictions yields a value of 10.5547, which has a p-value of 0.005.

Conditional second moments of the transcendental GARCH and the asymmetric BEKK models are presented in Figures 5 and 6 respectively. The main difference between the estimates is with the NYSE conditional variance estimate: the transcendental GARCH model picks up the large movements in the NYSE volatility as highlighted in Figure 4, better than the BEKK model which tends to produce relatively smoother conditional second moment estimates.[6]

Estimates of the time-varying beta coefficient $h_{nw,t}/h_{w,t}$, from the two models are also given in Figures 5 and 6. The transcendental GARCH model yields beta estimates which tend to hover around the mean value of 0.444. The largest deviation occurs on the 24th of August, 1990, where the estimated value of the beta coefficient jumps to 0.927. The mean of the time-varying beta coefficient estimates from the BEKK model is relatively higher with a value of 0.510. The BEKK model time-varying beta estimates also tend to show a downward trend starting in 1991. For comparison, the estimate of the fixed beta coefficient model is 0.527, which is obtained from regressing $R_{n,t}$ on a constant and $R_{w,t}$.

Estimates of the time-varying risk premium, as defined by

$$Risk_t = h_{nw,t}\left(\frac{\delta_0 + \delta_1 R_{w,t-1}}{h_{w,t}}\right), \tag{49}$$

for the transcendental GARCH and BEKK models are given in Figure 7. The two estimates of the time-varying risk premium tend to exhibit similar overall patterns with a correlation of 0.889. The main difference between the two estimates are the few, large negative estimates of the risk premium estimates obtained using the transcendental GARCH model.

[6] An extension of the BEKK model suggested by these results is to follow Glosten, Jagannathan and Runkle (1993) and include dummy variables to allow for additional asymmetries in the conditional second moments.

Table 2:

Transcendental Multivariate GARCH Model Estimates.

| Parameter | Gen. Normal | | Gen. Student t | |
|---|---|---|---|---|
| | Estimate | s.e. | Estimate | s.e. |
| $\gamma_0$ | -0.011 | 0.036 | 0.034 | 0.012 |
| $\delta_0$ | -0.028 | 0.039 | 0.003 | 0.014 |
| $\delta_1$ | 0.235 | 0.060 | 0.174 | 0.027 |
| $\theta_{1,0}$ | 1.604 | 0.102 | 1.666 | 0.670 |
| $\theta_{1,1}$ | -0.263 | 0.185 | -0.436 | 0.172 |
| $\theta_{1,2}$ | 0.373 | 0.082 | 0.325 | 0.271 |
| $\theta_{1,3}$ | 0.000 | | 0.159 | 0.336 |
| $\theta_{2,0}$ | -0.012 | 0.003 | -0.017 | 0.005 |
| $\theta_{2,1}$ | -0.187 | 0.029 | -0.178 | 0.031 |
| $\theta_{2,2}$ | 0.056 | 0.024 | 0.033 | 0.023 |
| $\theta_{2,3}$ | 0.991 | 0.002 | 0.992 | 0.002 |
| $\theta_{3,0}$ | -0.358 | 0.114 | -0.190 | 0.387 |
| $\theta_{3,1}$ | -0.192 | 0.166 | -0.043 | 0.126 |
| $\theta_{3,2}$ | -0.319 | 0.090 | -0.282 | 0.245 |
| $\theta_{3,3}$ | 0.647 | 0.125 | 0.906 | 0.187 |
| $\ln L_t/T$ | -2.26075 | | -1.76882 | |

Table 3:

BEKK Model Estimates.

| Parameter | Asymmetric | | Symmetric | |
|---|---|---|---|---|
| | Estimate | s.e. | Estimate | s.e. |
| $\gamma_0$ | 0.033 | 0.013 | 0.037 | 0.012 |
| $\delta_0$ | 0.022 | 0.014 | 0.020 | 0.013 |
| $\delta_1$ | 0.178 | 0.027 | 0.166 | 0.026 |
| $c_{11}$ | 0.014 | 0.026 | -0.011 | 0.054 |
| $c_{21}$ | -0.046 | 0.032 | 0.066 | 0.013 |
| $c_{22}$ | 0.153 | 0.083 | -0.075 | 0.033 |
| $a_{11}$ | 0.105 | 0.018 | 0.140 | 0.025 |
| $a_{12}$ | -0.084 | 0.042 | -0.041 | 0.028 |
| $a_{21}$ | 0.026 | 0.073 | -0.041 | 0.028 |
| $a_{22}$ | 0.376 | 0.110 | 0.260 | 0.040 |
| $b_{11}$ | 0.987 | 0.016 | 0.974 | 0.008 |
| $b_{12}$ | 0.061 | 0.031 | 0.026 | 0.011 |
| $b_{21}$ | 0.004 | 0.033 | 0.026 | 0.011 |
| $b_{22}$ | 0.878 | 0.081 | 0.945 | 0.017 |
| $\ln L/T$ | -1.77386 | | -1.77650 | |

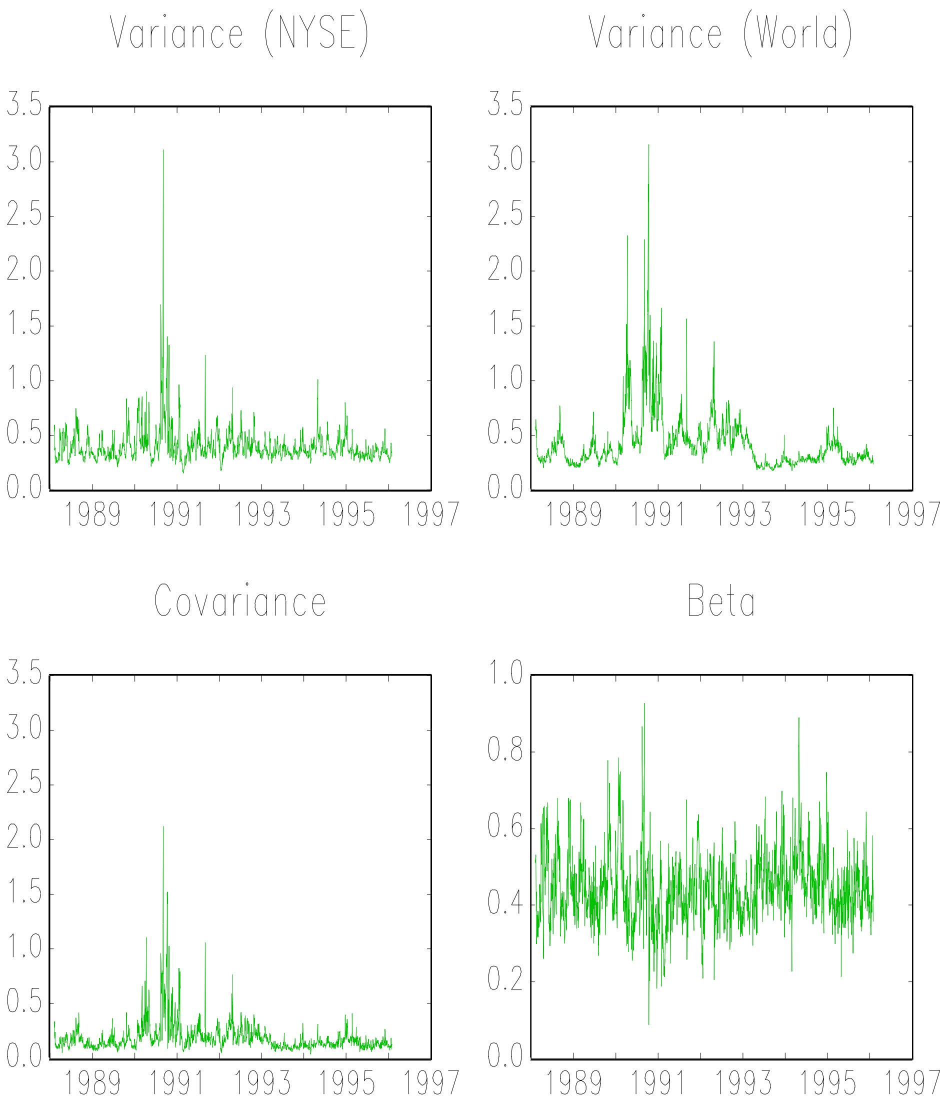


Figure 5: Conditional moments of the transcendental GARCH model: 1988 to 1995.

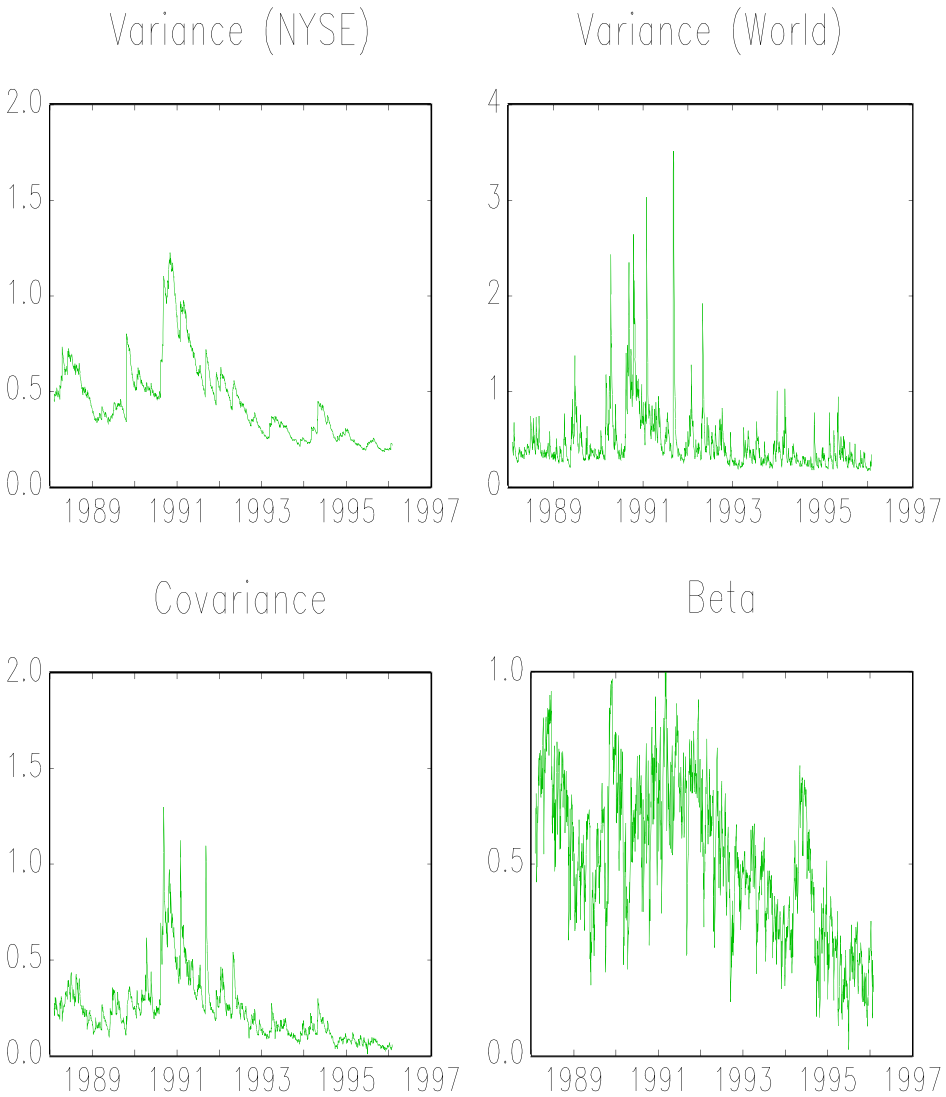


Figure 6: Conditional moments of the BEKK model: 1988 to 1995.

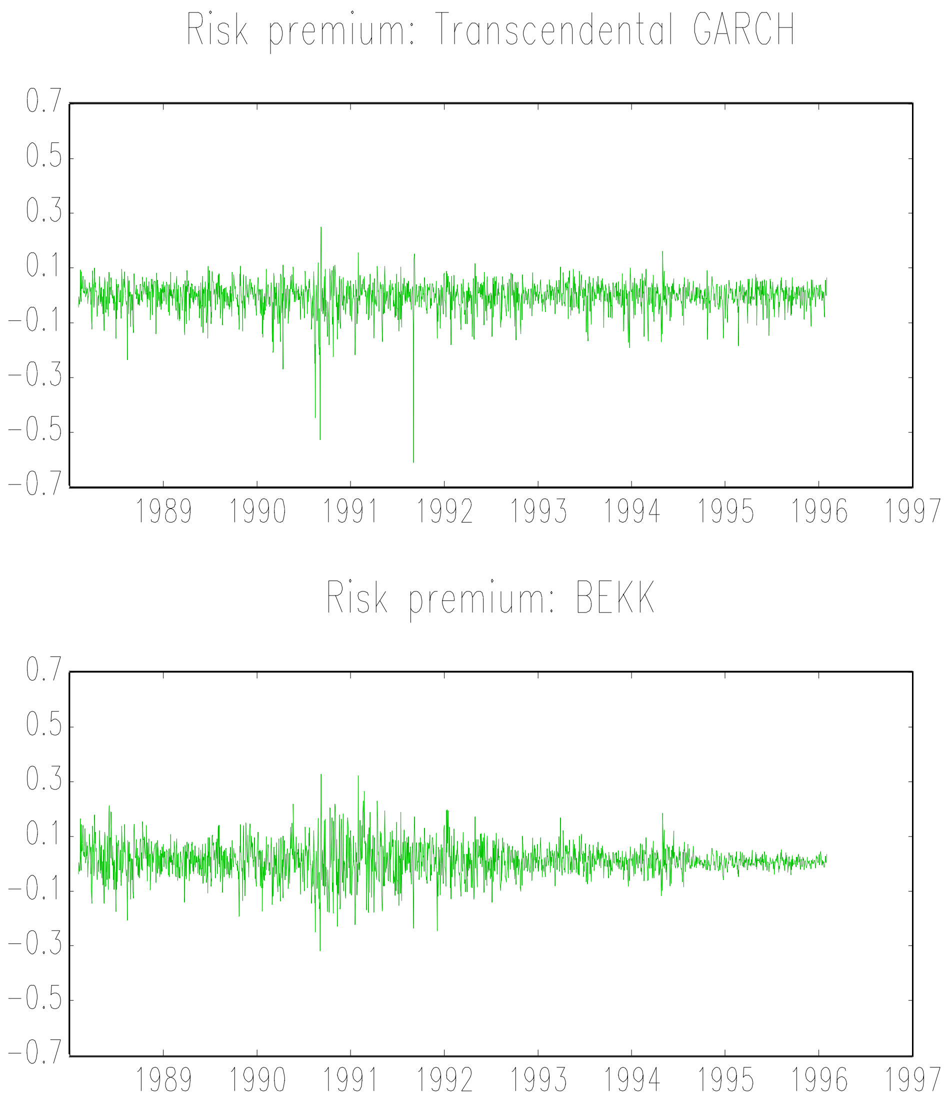


Figure 7: Alternative estimates of the risk premium: 1992 to 1995.

The time series plots of the transcendental GARCH time-varying beta coefficient in Figure 5 and the risk premium in Figure 7, suggest a change in risk around the 24th of August, 1990. To investigate this change, snapshots of the bivariate generalized Student t distributions are given in Figure 8 for six consecutive trading days starting Monday, the 20th of August. The key property to note here is that at the start of the week, the distribution is just slightly bimodal, but that the strength of bimodality increases over the week.[7] The distribution still exhibits bimodality on the following Monday, the 27th of August.

Evidence of bimodality in the excess returns bivariate distribution provides further information for understanding the nature of the increase in risk around the 24th of August, 1990: during these trading days the increase in volatility is associated with the excess returns switching between the regions of the stable modes. Uncovering evidence of bimodality of course is a feature of the generalized Student t distribution which is undetectable within the BEKK framework based on the assumption of conditional normality. Whilst Figure 6 for example, shows that the BEKK model is able to detect changes in volatility during this period through the second conditional moment, the transcendental GARCH model shows that it is changes in the higher order moments which underlie these movements in risk.

An alternative interpretation of the occurrence of bimodality during periods of increased volatility is that one of the modes is spurious. To examine this proposition, Table 4 gives the actual excess returns from the 20th to the 27th of August, 1990. Also given in the table are the areas of the bivariate distribution for various quadrants corresponding to whether the NYSE and world excess returns are increasing or decreasing. The probabilities show that the largest probability regions coincide with the two excess returns having the same signs. This is consistent with the positive estimate obtained for the time-varying beta coefficient during this period. Concentrating on the excess returns for the 22nd and 23rd of August, Table 4 shows that both excess returns are negative. However, on these days the probability estimates also show that there is a probability of around 0.44 of positive excess returns. Inspection of Figure 8 shows that the mode in the positive quadrant is around 1.5%. It is interesting to observe that on the 24th of August, excess returns did in fact become positive, with the NYSE and world excess returns being 1.279% and 1.301% respectively. This result suggests that the additional mode, whilst dormant on the 22nd and 23rd of August, became active on the 24th.

A further implication of bimodal distributions during turbulent periods is that the relationship between $R_{n,t}$ and $R_{w,t}$, is nonlinear which is not captured by the CAPM. In contrast, during the relatively more tranquil periods the bivariate

[7] At this stage, no formal statistical test of bimodality is offered. One approach would be to extend the bimodality test for univariate processes of Lim, Martin and Teo (1998), to bivariate processes.

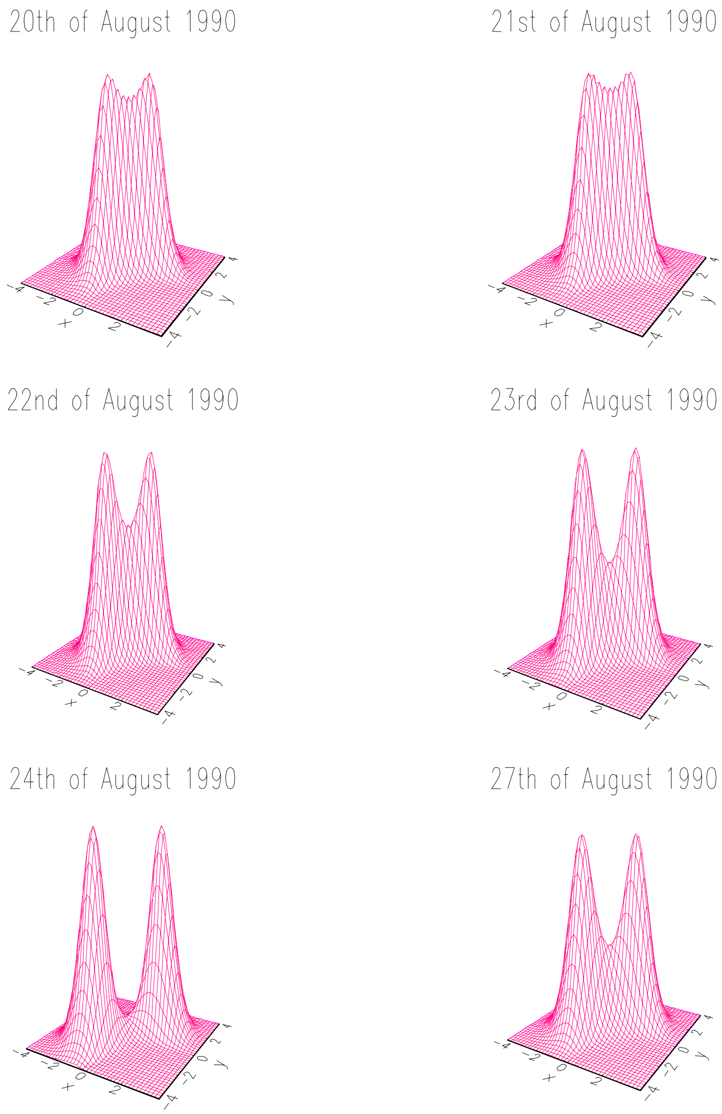


Figure 8: Daily 2-dimensional distributions of NYSE excess returns ($y$) and world excess returns ($x$): 20/8/1990 to 27/8/1990.

Table 4:

Statistics for the period the 20th to the 27th of August, 1990: excess NYSE returns ($y$), excess world returns ($x$).

| Date | y | x | Area of probability regions | | | |
|---|---|---|---|---|---|---|
| | | | y<0,x<0 | y>0,x<0 | y<0,x>0 | y>0,x>0 |
| 20/8/1990 | 0.036 | -0.457 | 0.417 | 0.083 | 0.083 | 0.417 |
| 21/8/1990 | -1.908 | 1.752 | 0.414 | 0.086 | 0.086 | 0.414 |
| 22/8/1990 | -1.536 | -1.437 | 0.430 | 0.070 | 0.070 | 0.430 |
| 23/8/1990 | -3.110 | -3.301 | 0.443 | 0.058 | 0.058 | 0.443 |
| 24/8/1990 | 1.279 | 1.301 | 0.479 | 0.021 | 0.021 | 0.479 |
| 27/8/1990 | 2.891 | 3.202 | 0.438 | 0.062 | 0.062 | 0.438 |

distributions are found to be unimodal (not shown here). For these periods, risk is adequately modelled using the conditional second moments.

# 5 CONCLUSIONS

This paper has introduced a new class of distributions which represent multivariate analogues of the univariate generalized distributions proposed by Cobb, Koppstein and Chen (1983) and Lye and Martin (1993). An important feature of this family of distributions is that it is flexible enough to exhibit various distributional characteristics such as leptokurtosis, asymmetry and multimodality.

This family of distributions was applied to formulating a new class of conditional variance models known as transcendental GARCH. A computational advantage of the transcendental GARCH model over the existing conditional variance models based on the BEKK specification is that a more natural, simpler set of restrictions are needed to achieve positive definite conditional variance-covariance matrices. A further advantage of the transcendental GARCH model over the BEKK model was that it could provide further insight into changes in risk over time. This was indeed the case with the empirical application concerning an international CAPM between excess NYSE returns and excess world returns over the period 1988 to 1995, using daily data. The empirical results showed that increases in the time-varying beta coefficient were associated with periods of bimodality in the distribution of returns. Whilst the BEKK model was able to identify changes in risk over time via the conditional second moments, it was not able to identify the cause of these changes; namely, changes in the higher order conditional moments. During periods of relatively high volatility the occurrence of bimodality meant that a nonlinear relationship existed between the NYSE and world excess returns with the implication that the CAPM was misspecified. In contrast, for tranquil periods, the distributions were found to be unimodal giving support to the CAPM and the use of the second conditional moments as a measure of risk.

The empirical results presented in this paper are preliminary. Possible extensions of this framework would include more in depth testing of alternative generalized Student t distributions. It would also be of interest to extend the BEKK model to allow for additional asymmetries. Another line of research would be to maintain the BEKK conditional variance specification and replace the conditional normality assumption by a multivariate generalized Student t distribution. This would have the advantage of making comparisons between the BEKK model and the multivariate GARCH model introduced here, more straightforward. The disadvantage would be the convenience of imposing positivity on the second moments would no longer exist.

# References

[1] Bera, A. and John, S. (1983), "Tests for Multivariate Normality with Pearson Alternatives", *Communications in Statistics- Theory and Methods*, 12(1), 103-117.

[2] Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

[3] Bollerslev, T. (1987), "A Conditionally Heteroskedastic Time Series Model fir Speculative Prices and rates of Return", *Review of Economics and Statistics*, 69, 542-547.

[4] Bollerslev, T. and Wooldridge, J.M. (1992), "Quasi-Maximum Likelihood Estimation and inference in Dynamic Models with Time-varying Covariances", *Econometric Reviews*, 11, 143-172.

[5] Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence", *Journal of Econometrics*, 52, 5-59.

[6] Bollerslev, T., Engle, R.F. and Nelson, D. (1994), "ARCH Models", *The Handbook of Econometrics*, Vol.4, North-Holland, forthcoming.

[7] Bollerslev, T., Engle, R.F. and Wooldridge (1988), "A Capital Asset Pricing Model with Time-Varying Covariances", *Journal of Political Economy*, 96, 116-131.

[8] Cobb, L., P. Koppstein, and N.H. Chen (1983), "Estimation and Moment Recursion Relations for Multimodal Distributions of the Exponential Family", *Journal of the American Statistical Association*, 78, 124-130.

[9] Creedy, J, Lye, J.N. and Martin, V.L. (1996), "A Nonlinear Model of the Real US/UK Exchange Rate", *Journal of Applied Econometrics*, 11, 669-686.

[10] Engle, C. and Rodrigues, (1989), "Tests of International CAPM with Time-Varying Covariances", *Journal of Applied Econometrics*, 4, 119-138.

[11] Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica,* 50, 987-1008.

[12] Engle, R.F., Granger, C.W.J. and Kraft, D. (1984), "Combining Competing Forecasts of Inflation Using a Bivariate ARCH Model", *Journal of Economic Dynamics and Control*, 151-165.

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

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

[15] Friedman, D. and Vandersteel, S. (1982), "Short-Run Fluctuations in Foreign Exchange Rates: Evidence from the data 1973-79", *Journal of International Economics*, 13, 171-186.

[16] Higgins, M.L. and Bera, A.K. (1992), "A Class of Nonlinear ARCH Models", *International Economic Review*, 33, 137-158.

[17] Gallant, R.A., Hsieh, D.A. and Tauchen, G. (1991), "On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974-83", in W.A. Barnett, J. Powell and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press, Chapter 8.

[18] Glosten, L.R., Jagannathan, R. and Runkle, D. (1993), "On the relation between the expected value and the volatility of the nominal excess return on stocks", *Journal of Finance,* 48, 1779-1802.

[19] Harvey, A., Ruiz, E. and Sentana, E. (1992), "Unobserved Component Time Series Models with ARCH Disturbances", *Journal of Econometrics*, 52, 129-157.

[20] Hsieh, D.A. (1989), "Modelling Heteroskedasticity in Daily Foreign Exchange Rates", *Journal of Business and Economic Statistics*, 7, 307-317.

[21] Kendall, M.G. and A. Stuart (1969), *The Advanced Theory of Statistics*, Vol. 1, 3rd edition, London: Charles Griffen and Co.

[22] Kroner, K.F. and Ng, V.K. (1996), "Multivariate GARCH Modelling of Asset Returns", *Papers and Proceedings of the American Statistical Association, Business and Economic Sections*, forthcoming.

[23] Lim, G.C., Martin, V.L. and Teo, L. (1998), "Endogenous Jumping and Asset Price Dynamics", *Macroeconomic Dynamics*, 2, 1-25.

[24] Lim, G.C., Lye, J.N., Martin, G.M. and Martin, V.L. (1998), "The Distribution of Exchange Rate returns and the Pricing of Currency Options", *Journal of International Economics*, 45, 351-368.

[25] Lye, J.N. and V.L. Martin (1993), "Robust Estimation, Non-Normalities and Generalized Exponential Distributions", *Journal of the American Statistical Association*, 88, 253-259.

[26] McCurdy, T.H. and Morgan, I.G. (1993), "Intertemporal Risk in Foreign Currency Markets", in *The Exchange Rate and the Economy*, Bank of Canada, 325-355.

[27] McCurdy, T.H. and Morgan, I.G. (1991), "Tests for a Systematic Risk Component in Deviations from Uncovered Interest Parity", *Review of Economic Studies*, 58, 587-602.

[28] McCurdy, T.H. and Stengos, T. (1992), "A Comparison of Risk Premium Forecasts Implied by Parametric Versus Nonparametric Conditional Mean Estimators", *Journal of Econometrics*, 52, 225-244.

[29] Pagan A.R. and Hong, Y.S. (1991), Non-Parametric Estimation and the Risk Premium", in W.A. Barnett, J. Powell and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press.

[30] Pagan, A.R. and Schwert, G.W. (1990), "Alternative Models for Conditional Stock Volatility", *Journal of Econometrics*, 45, 267-290.