Modelling Regime Specific Stock Volatility Behaviour

Abstract

Any GARCH model with a single volatility state identifies only one mechanism governing the dynamic response of volatility to market shocks, and the conditional standardized higher moments are constant, unless modelled explicitly. Therefore such models cannot capture state dependent behaviour of volatility and neither can they explain why the equity index skew persists into long dated options. However the family of Markov switching (MS) GARCH models, which specify several deterministic mean-reverting volatility states, have endogenous time-varying conditional variance, skewness and kurtosis. Within this model class the normal mixture (NM) GARCH model, a MS GARCH with constant state probabilities, is simplest to estimate. This paper extends previous theoretical and empirical work on the NM GARCH model by introducing a state dependent leverage effect., which is essential for modelling equity market volatility. With this construction we can explain the typical empirical characteristics of equity index returns and implied volatility skews, even with a zero or constant volatility risk premium. An empirical study of fifteen GARCH models of four European equity indices between 1991 and 2005 identifies two-state asymmetric NM GARCH as the best fitting model. The models show that the volatility behaviour is broadly similar in all indices during usual market circumstances. But there is a small probability (between 2% and 5%) of the market crashing, when the volatility behaviour of different indices becomes quite diverse, and the mean-reversion and leverage effects are quite different from those in the usual regime.

JEL Classification Codes: C32, G13.

Keywords: GARCH, normal mixture, leverage effect, mean-reverting volatility, equity skew, stock market regimes, market crash
I Introduction

Modelling regimes in market volatility is of interest to most finance practitioners. For their assessment of risk capital and for setting traders’ limits, risk managers need to understand how volatility may respond differently to market shocks during a predominantly stable period, compared with a crash or recovery period. When computing economic capital, and to enable traders’ limits to be raised when markets are considered to be exceptionally volatile, risk models are subjected to ‘stress tests’. These tests are also required by most banking regulators but they are often performed using ad hoc adjustments. However, a more precise approach to stress testing would specify an intensity for the shock to volatility and the probability associated with the stressful scenario, and this requires a regime dependent volatility model. Managing volatility following large price falls in an index is essential for governments and policy makers as well, who are concerned about systemic risk, where insolvency in one sector of the economy may lead to an economic crisis as in the global crash of 1987. Regulators need to understand how markets behave under stressful conditions if their aim is to avert insolvency in a particular sector of the economy. Finally, an understanding of best discrete time volatility model should help option prices to select from the plethora of possible stochastic volatility models available, many of which will fit the market implied volatility smile equally well.

A main characteristic of volatility is ‘mean-reversion’, whereby volatility persists after a market shock but in the absence of any further shocks it will eventually return to a long-run volatility level. Mean-reversion is captured in the standard approach to modelling volatility behaviour, i.e. the generalized autoregressive conditional heteroscedasticity (GARCH) framework introduced by Engle (1982) and Bollerslev (1986). However, normal GARCH models do not explain why both conditional and unconditional returns on financial assets have such high skewness and kurtosis. For this reason Bollerslev (1987) extended the GARCH model to include the Student’s $t$ conditional distribution and Fernandez and Steel (1998) enhanced it further using the skewed $t$-distribution.

Bekaert and Wu (2000), Wu (2001), Hansen, Lunde and Nason (2003) and many others argue that it is important to also characterise the ‘leverage effect’ in stock markets, whereby volatility responds more to negative shocks than to positive shocks of the same magnitude. Nelson (1991) introduced the first GARCH process with a leverage effect. In this exponential model the returns are conditionally symmetric (assuming volatility is known) but unconditional asymmetry is obtained from the dynamics of the variance process. Amongst other GARCH processes with leverage effects are the absolute value GARCH based on Taylor (1986, pp. 78-79) and Schwert (1989), the asymmetric and nonlinear asymmetric GARCH introduced by Engle (1990) and Engle and Ng (1993), the GJR model of Glosten, Jagannathan and

---

1 As kindly noted by one of the referees, the asymmetric GARCH model was initially suggested by Engle (1990) and subsequently discussed in Engle and Ng (1993).
Runkle (1993), the asymmetric power GARCH model of Ding, Granger and Engle (1993), the threshold GARCH of Zakoïan (1994), the quadratic GARCH of Sentana (1995), the Box-Cox transformed GARCH of Hentschel (1995) and Hwang and Basawa (2004), which nests several different models, the model of Brännäs and de Gooijer (2004) that extends QGARCH and, more recently, the model of Lanne and Saikkonen (2007) that generalizes the AGARCH model.

All the above models assume that the GARCH volatility process has regime-invariant parameters and they do not allow different types of shocks (of the same size) to precipitate different types of responses. Yet a rumour may have an enormous effect but die out very quickly whereas an announcement of important changes to economic policy (if represented by a shock of the same size as the rumour) may produce a less volatile but more persistent response, after which the general level of volatility is raised. If the mean-reversion and leverage effects are state dependent, the parameters estimated in a single state GARCH process can represent only an average of these. Most single-state GARCH models also have constant conditional skewness and conditional kurtosis. As a result, the option implied volatility smiles and skews that are generated from these models are often unrealistic. Bates (1991), Christoffersen, Heston and Jacobs (2006) and others argue that, for a model to capture the empirical characteristics of option implied volatility skews, it is essential to account for time-variability in the physical conditional skewness and kurtosis. So another reason for using more than one volatility state in a GARCH model is that both conditional skewness and kurtosis can be endogenously time-varying.

More recent research examines Markov Switching (MS) GARCH and its more tractable relation: normal mixture (NM) GARCH. Switching ARCH models were first introduced by Cai (1994) and Hamilton and Susmel (1994). The straightforward generalization of this model to GARCH processes has estimation problems due to the dependence of the conditional variance on the entire history of states. A solution was suggested by Gray (1996) by replacing the lagged variance in the GARCH equation with its conditional expectation; also the error term is allowed to follow a Student’s t distribution with switching degrees of freedom parameter. Later on an enhancement was suggested by Klaassen (2002) making use of the full information set available when computing the conditional probability of a state; see also Marcucci (2005) for a model that allows for regime-specific means. However, these models are not very tractable.

One type of MS-GARCH was introduced by Francq, Roussignol and Zakoïan (2001); here there is only one variance process but the parameters are state-dependent. The existence of the moments and the $L^2$ structures of this model were described in Francq and Zakoïan (2005), whilst the autocovariance structure of the squared residuals were discussed in Francq and Zakoïan (2007); this paper also proposes a GMM procedure for the estimation of the parameters. Henneke, Rachev and Fabozzi (2006) and Bauwens,
Preminger and Rombouts (2006, 2007) use Bayesian techniques to estimate this type of MS-GARCH processes, whilst Dueker (1997) discusses an estimation method using a collapsing procedure to approximate the likelihood function for a simplified version of MS-GARCH.

An improved MS-GARCH model was introduced and its properties presented by Haas, Mittnik and Paolella (2004b) assuming that each state is characterized by its own variance and that the variance process of each state depends only on its lagged values (as opposed to lagged values of the other variance processes) and the squared residuals; for this model ML estimation is feasible. Weak stationarity properties of the (1,1) model and conditions for the existence of the higher moments, relaxing the assumption of initial finite variance, were discussed in Liu (2006); Abramson and Cohen (2007) also develop stationarity properties for the models of Klaassen (2002) and Haas, Mittnik and Paolella (2004b) – for models of order $(p,q)$ – using a technique based on backward recursion.

Liu (2007) introduces a general model that allows for asymmetries in the variance equations; the focus of this paper is the discussion of the stationarity properties and existence of higher moments of a generalization of the model of Haas, Mittnik and Paolella (2004b) – a MS Box-Cox transformed threshold GARCH(1,1), to be more precise.

All the above models are rather complex to estimate; a simpler model that ignores the autocorrelation in the state variable (but, as opposed to some MS models, allows for different means in the states) is normal mixture GARCH. Here the error term follows a mixture distribution. Early versions of NM-GARCH models were discussed in Vlaar and Palm (1993), Bauwens, Bos and van Dijk (1999), Bai, Russell and Tiao (2001, 2003) and Ding and Granger (1996). Concentrating on the individual means in the mixture density, Wong and Li (2001) introduced a NM-ARCH model with individual AR processes in the mean equations. This model was later extended to GARCH processes by Lanne and Saikkonen (2003); they also assume that the mixing weights are functions of past observations. The general format of this model was introduced by Haas, Mittnik and Paolella (2004a) and Alexander and Lazar (2006), providing strong evidence that these models provide a better fit to exchange rate data than single-state GARCH models. Bayesian estimation of the NM-GARCH was discussed in Bauwens and Rombouts (2007). Multivariate extensions were considered by Haas, Mittnik and Paolella (2006), Fong, Li and An (2006) and Bauwens, Hafner and Rombouts (2007) – the latter uses an estimation technique based on the EM algorithm.

Yet none of these models, with the exception of Henneke, Rachev and Fabozzi (2006) and Liu (2007), have leverage effects in the state dependent GARCH volatility processes. The only source of skewness in the physical returns densities is the different means in the normal components of the mixture conditional
density. Hence such models are not suitable when the leverage effect is known to be very important, as in individual stock and stock index markets.

This paper examines the theoretical and empirical behaviour of state-dependent asymmetric volatility processes in a normal mixture GARCH model. A main strength of the model is that it allows the asymmetric volatility response to stock price shocks to be quantified in both usual and stressful market circumstances. When two-state normal mixture models are applied to financial markets, often two states can be differentiated: a ‘normal’ state that occurs most of the time and a ‘stressful’ state that occurs only rarely. At every point of time the state is selected randomly, independently of the previous state. However, as we will see later, following this random and unknown state-selection and after a return is realized from the given distribution, it is possible to estimate the probability that a state was selected, and this will be the ex-post inference about the probability of the state at time \( t \), and it will be closely related to the actual state of the world. Also, it allows the conditional higher moments to change in time. Implementing such a model would allow risk capital assessments to be more refined and help risk managers to derive regime specific limits for traders. It would also bring new insights, relevant to investors and policy makers, about the likelihood of a market crash and the returns/volatility behaviour during a crash period. Two asymmetric models are considered, namely the AGARCH and GJR extension of NM-GARCH, based on the papers of Engle (1990), Engle and Ng (1993) and Glosten et al. (1993). One of the models, the NM-GJR is a special case of the model of Liu (2007), with \( \delta = 1 \) and with no autocorrelation for the state variable; thus the theoretical results therein can be applied for the NM-GJR model in this paper. However, this is not a restricted model of Henneke, Rachev and Fabozzi (2006), even though both consider GJR variance processes. The other model discussed in this paper, NM-AGARCH, is not nested within the family of models considered by Liu (2007).

Our empirical study on four major European equity market indices demonstrates the superiority of a two-state asymmetric normal mixture GARCH model to capture the empirical characteristics of stock index returns. We restrict our study to asymmetric NM-GARCH models as MS models don’t necessarily perform better than mixture models and our focus is on the regime-specific asymmetric behaviour of the variance process. We also compare the risk neutral skews generated by single component GARCH models with the skews that are generated by the two-state asymmetric GARCH models. Even without a risk premium the volatility smile implied by asymmetric normal mixture GARCH models exhibits a pronounced skew that persists into long-dated options.

\[ \text{(4) This is due to that NM-GARCH is a special case of the MS-GARCH of Haas, Mittnik and Paolella (2004b), which is a different formulation than the model of Henneke, Rachev and Fabozzi (2006).} \]

\[ \text{(5) A thorough theoretical and empirical comparison of MS and NM-GARCH models is done by Haas, Mittnik and Paolella (2004b) by evaluating the fit of these models for three currencies. Using the BIC criterion, NM models are preferred in 2 out of 3 cases. When the forecasting performance of these models is studied, then again in most cases the NM models proved to dominate. The only criterion according to which MS models are (marginally) preferred is the fit of the conditional higher moments.} \]
The rest of the paper is organized as follows: Section II defines the general asymmetric normal mixture GARCH model and describes two sources of asymmetry in these models; Section III characterizes the equity index data for four major equity markets and presents the estimation methodology; Section IV reports our empirical results. We start with our estimations for fifteen different GARCH models, including two-state asymmetric and symmetric normal mixture GARCH models, and symmetric and skewed t-GARCH with both symmetric and asymmetric variance processes. Then several model selection criteria are applied to identify the best model. Choosing the two-state normal mixture AGARCH and GJR models as best overall, the estimated parameters are used to infer various aspects of volatility behaviour in European stock markets. Section V simulates prices of European calls and puts of various maturities and compares the implied equity index skews of the FTSE index simulated by symmetric and asymmetric GARCH models with one and two states. Section VI summarizes and concludes.

II The Asymmetric Normal Mixture GARCH Model

The model extends the GARCH processes studied in Haas, Mittnik and Paoella (2004a) and Alexander and Lazar (2006) to include a leverage effect in the GARCH variance equations. Specifically, the model has one equation for the mean and K variance equations. For simplicity the conditional mean equation is written \( y_t = \epsilon_t \), but this can easily be extended to allow for explanatory variables because their coefficients can be estimated separately. The error term \( \epsilon_t \) is assumed to have a conditional normal mixture density with zero mean, which is a weighted average of K normal density functions with different means and variances. We write:

\[
\epsilon_t \mid \mathbf{1}_{t-1} \sim \text{NM}\left(p_1, \ldots, p_K, \mu_1, \ldots, \mu_K, \sigma^2_1, \ldots, \sigma^2_K\right), \quad \sum_{i=1}^{K} p_i = 1, \quad \sum_{i=1}^{K} p_i \mu_i = 0
\]

and the conditional density of the error term is:

\[
\eta_\epsilon(\epsilon_t) = \sum_{i=1}^{K} p_i \varphi_i(\epsilon_t)
\]

where \( \varphi_i \) represent normal density functions at time \( t \) with different means \( \mu_i \) and different time-varying variances \( \sigma^2_i \) for \( i = 1, \ldots, K \). The mixing law \( \mathbf{p} = (p_1, \ldots, p_K) \) has elements that may be interpreted as the relative frequency of each state occurring over a long period of time; for reasons of identifiability we assume that \( 1 \geq p_1 \geq p_2 \geq \ldots \geq p_K \geq 0 \). There are \( K \) conditional variance components and these can follow any GARCH process, but for the purpose of this paper we assume that there are three possibilities:

(i) **NM-GARCH(1,1):**

\[
\sigma^2_i = \omega_i + \alpha_i \sigma^2_{i-1} \quad \text{for} \quad i = 1, \ldots, K
\]

(ii) **NM-AGARCH(1,1) (based on the Engle, 1990 and Engle and Ng, 1993 model):**

\[
\sigma^2_i = \omega_i + \alpha_i \left(\epsilon_{i-1} - \lambda_i \right)^2 + \beta_i \sigma^2_{i-1} \quad \text{for} \quad i = 1, \ldots, K
\]
(iii) \textbf{NM-GJR(1,1) (based on the Glosten et al., 1993 model):}

\begin{equation}
\sigma^2_t = \omega_i + \alpha_i \varepsilon^2_{t-1} + \lambda_i d_{t-1}^i \varepsilon^2_{t-1} + \beta_i \sigma^2_{t-1} \quad \text{for } i = 1, \ldots, K; \tag{5}
\end{equation}

where \( d_t^i = 1 \) if \( \varepsilon_t < 0 \), and 0 otherwise.

For the normal densities to be well identified, we also assume that:

\[ |\mu_i - \mu_j| + |\omega_i - \omega_j| + |\alpha_i - \alpha_j| + |\beta_i - \beta_j| + |\lambda_i - \lambda_j| > 0 \quad \text{for } i \neq j. \]

The main difference between NM-GARCH and the MS-GARCH model of Haas, Mittnik and Paolella (2004b) is that the latter allows for autocorrelation in the state variable; in other words, for the NM-GARCH model the transition matrix is of rank 1, so it can be expressed as

\[ P = (p_{ij}) = P[\Delta_t = i|\Delta_{t-1} = j] - \mathbf{1} \]

where \( \mathbf{1} = (1, \ldots, 1)' \).

Also, the NM-GARCH adds asymmetry by allowing for non-zero means for the different components.

The overall conditional variance can be expressed as

\begin{equation}
\sigma^2_t = \sum_{i=1}^{K} p_i \sigma^2_{t-i} + \sum_{i=1}^{K} p_i |\mu_i|^2 \tag{6}
\end{equation}

and, given the data, it is possible to compute an ex-post estimate of the probability of the \( i \)th state at an arbitrary time \( t \), expressed by:

\begin{equation}
p_{i,\delta} = \frac{p_i p_i (e_i)}{\sum_{j=1}^{K} p_j p_j (e_i)} \tag{7}
\end{equation}

This is a time-varying probability, giving a realistic ex-post inference about the probability of state \( i \) occurring at time \( t \). Thus at any point in time we have an ex-ante probability of the states, given by the mixing law, and an ex-post estimate of the state-probabilities, given by the above formula.

The weak stationarity conditions for the two models are stated in the Appendix. For \( K > 1 \), the existence of second, third and fourth moments are assured by imposing less stringent conditions than in the single component (\( K = 1 \)) models. Namely, it is not necessary to assume that each component is stationary in the conventional sense. For instance, it may happen that the more volatile component would be considered non-stationary according to standard interpretation (i.e. \( \alpha_i + \beta_i > 1 \) for that component). However, the model has to be considered as a whole, and interpreting its components in isolation would be misleading. \(^5\)

\(^4\) Also note that we do not require that \( \omega_i > 0 \); actually it often happens that the \( \omega \) parameter of one of the components is slightly negative. However, as long as it is small in absolute value, it does not affect the positivity of the variance series (also, see Alexander and Lazar, 2006).

\(^5\) When the models are restricted to NM-GARCH(1,1), then the stationarity conditions are equivalent to those given by Alexander and Lazar (2006). Also, they are equivalent to those given by Haas, Mittnik and Paolella (2004a) and Liu (2006) when the \( \omega_i > 0 \).
There are two distinct sources of asymmetry in the model:

- **Persistent Asymmetry:** This arises in all three models when the conditional returns density is a mixture of normal density components having different means; it is generated by the difference between the expected returns under different market circumstances. The Appendix shows that even the unconditional density will have non-zero skewness, and that this increases with the differentiation of the component means. For instance, when $K = 2$, there is positive skewness in the overall conditional returns density if the component with the higher probability has low volatility and a negative mean (as it generally happens), and negative skewness in the overall conditional returns density if the component with the higher probability has low volatility and a positive mean.6

- **Dynamic Asymmetry:** This only occurs in models (ii) and (iii) and is due to the leverage parameters in the component variance processes. If the leverage parameter $\lambda_i$ is positive, the conditional variance in this component is higher following a negative unexpected return than following a positive unexpected return. In equity markets, where a negative unexpected return increases the leverage of a firm we expect positive values of the leverage parameter. However, some studies (e.g. Koutmos et al., 1998 and Chortareas et al., 2000) report negative leverage coefficients in stock indices of developing economies.7

Taken together, these two sources of skewness in the physical conditional returns density offer a much richer structure for capturing the shape of equity index skews than is given by traditional GARCH models. The unconditional skewness and excess kurtosis are both non-zero and the conditional higher moments are also non-zero and time-varying. See the Appendix for the derivation of the unconditional higher moments.

The conditional normal mixture densities can be interpreted as merely a device to increase the flexibility of the returns density, because the model will capture the behaviour of returns during different regimes and identify the long-run probability of each state, based on time series data. The model is considerably easier to estimate than the class of Markov switching GARCH models introduced by Hamilton and Susmel (1994) even with the restrictions and improvements introduced by Cai (1994), Gray (1996), Klaassen (1994) even with the restrictions and improvements introduced by Cai (1994), Gray (1996), Klaassen

---

6 As suggested by one of the referees, this can be seen by expressing the third moment as

\[
m_3 = \rho_p^2 \left[ (\rho_2 - \rho_1)(\mu_1 - \mu_2)^3 + 3(\mu_1 - \mu_2)(\sigma_1^2 - \sigma_2^2) \right].
\]

7 Note that the $i$th variance component depends on the dispersion of the unexpected return, not around its mean $\mu_i$ in the individual density, but around the overall mean 0. Hence there is a third effect that induces skewness in each component conditional return density but not in the overall conditional return density.
(2002) and Haas, Mittnik and Paoella (2004b). The difficulty with estimating most of these models (except the last one) lies in the co-dependencies of the state variances. However, the normal mixture GARCH models considered here have a very straightforward relationship between the individual variances, because they are tied to each other only through their dependence on the error term. Hence an advantage of normal mixture GARCH is that quite complex volatility feedback mechanisms can be included and the model remains easy to estimate.

III Data and Parameter Estimation

Our results are based on the daily closing prices of four major European equity market indices: CAC40 (total number of observations is 3733), DAX30 (3730), FTSE100 (3739) and DJ Eurostoxx 50 (3810) from 1 January 1991 to 21 October 2005. This period encompasses a number of stock market crises, including the Asian crisis in 1997, the Russian crisis in 1998, the technology market crash in 2000 and the terrorist attack on the US in 2001. Table 1 summarises the general characteristics of the daily returns. The skewness is negative and the excess kurtosis is positive and in most indices these are highly significant. At first, it can be noticed that during this period the FTSE index was less volatile in general than the other markets.

Table 1: Summary Statistics of the Stock Market Indices

<table>
<thead>
<tr>
<th></th>
<th>CAC</th>
<th>DAX</th>
<th>FTSE</th>
<th>Eurostoxx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>21.24%</td>
<td>22.82%</td>
<td>16.41%</td>
<td>20.13%</td>
</tr>
<tr>
<td>Skewness</td>
<td>–0.099*</td>
<td>–0.271***</td>
<td>–0.113**</td>
<td>–0.166***</td>
</tr>
<tr>
<td>Excess kurtosis</td>
<td>2.86***</td>
<td>4.03***</td>
<td>3.19***</td>
<td>4.25***</td>
</tr>
<tr>
<td>Maximum moment exponent</td>
<td>5.58**</td>
<td>4.59***</td>
<td>6.70*</td>
<td>5.66***</td>
</tr>
</tbody>
</table>

For each index, we estimate the conditional variance parameters on the residuals $\varepsilon_t$ from constant conditional mean equations. Then, maximizing the likelihood function, or equivalently, maximizing

$$L(\theta | \varepsilon) = \sum_{t=1}^{T} \ln \left[ \eta(\varepsilon_t) \right]$$

the optimal parameter values are found, given the data. The updating formula has the following form, where $g$ is the gradient vector, $H$ the Hessian matrix and $s$ represents the step-length:

$$\theta_{m+1} = \theta_m - s \left[ H(\theta_m) \right]^{-1} g(\theta_m)$$

---

8 The standard errors (s.e.) of the sample estimates of the mean, variance, skewness and excess kurtosis parameters are as follows:

- s.e. of the sample mean $= \sigma / \sqrt{T}$,
- s.e. of the sample variance $= \sigma / \sqrt{T}$,
- s.e. of the sample skewness $= \sqrt{6 / T}$,
- s.e. of the sample excess kurtosis $= \sqrt{24 / T}$,

where $T$ represents the total number of observations. In Table 1 *** represent results significantly different from zero at the 0.1% level, ** at 1% and * at 5%. The last row reports the HKKP estimate for the maximum moment exponent; here ***, ** and * denote results that are significantly higher than 4 at 0.1%, 1% and 5% significance level, respectively.

9 We fitted ARMA models to the series; the best fitting model for all series was an ARMA(0,0) model. Thus we only removed the means of the series and continued with the demeaned series.
To compute the Hessian matrix and the gradient vector we can use either analytic or numerical first and second order derivatives of the likelihood – see the Appendix for the numerical derivatives.\textsuperscript{10}

IV Empirical Results

IV.1 Overview of Models

We fitted three symmetric and twelve asymmetric GARCH models to the equity index data. The first nine models have a single variance component and the last six models have two variance components.\textsuperscript{11} The models are:\textsuperscript{12}

A. Models with normally distributed errors:
   \begin{enumerate}
   \item GARCH
   \item AGARCH
   \item GJR
   \end{enumerate}

B. Models with symmetric Student’s $t$ distributed errors:
   \begin{enumerate}
   \item GARCH
   \item AGARCH
   \item GJR
   \end{enumerate}

C. Models with skewed Student’s $t$ distributed errors:
   \begin{enumerate}
   \item GARCH
   \item AGARCH
   \item GJR
   \end{enumerate}

D. Normal mixture GARCH models with zero means in the mixture component densities:
   \begin{enumerate}
   \item NM-GARCH
   \item NM-AGARCH
   \item NM-GJR
   \end{enumerate}

E. General normal mixture GARCH models:
   \begin{enumerate}
   \item NM-GARCH
   \item NM-AGARCH
   \item NM-GJR
   \end{enumerate}

\textsuperscript{10} The results were generated using C++ and Ox version 3.30 (Doornik, 2002) and the G@rch package version 3.0 (Laurent, S. and Peters, J.-P., 2002)

\textsuperscript{11} Several restricted versions of the normal mixture GARCH model were also fitted to the data (assuming a constant variance component or assuming constant difference between the two variance processes) – but these performed quite badly according to some of the selection criteria and thus these models are not discussed here.

\textsuperscript{12} All models are GARCH(1,1) specifications. To simplify notation, this will be implied in the following. Models (2), (3), (5), (6), (7), (11), (12) and (13) have one source of asymmetry, whilst models (8), (9), (14) and (15) have two sources. Also, all normal mixture models will be based on two normal distributions in the mixture, as important conclusions can be drawn from such a setup – see also Marcucci (2005) and Bauwens, Preminger and Rombouts (2007). We do not consider models with more than two components as those would increase the number of parameters from 10 to at least 16 and, as the results of Haas, Mittnik and Paolella (2004a,b) and Alexander and Lazar (2006) clearly indicate, in the case of three-component mixture models convergence is harder to achieve and also these models prove to be over-parameterized when it comes to out-of-sample forecasting.
The estimation results are reported in Tables 2 – 5. The upper figure in each cell reports the parameter estimate and the lower figure is the t-ratio. Note that in these tables the first row reports the degrees of freedom for the t-GARCH models (4) – (9) and the weight of the first component in the NM-GARCH models for models (10) – (15). Similarly, the third row reports the skewness parameter for models (7) – (9) and the mean of the first normal density for the normal mixture GARCH models.

IV.2 Ex-post State Probabilities

Figure 1 presents the daily Eurostoxx index returns for the period Jan 2004 – October 2005. It is expected that the ex-post state probabilities will indicate a jump from the first state (characterized by a lower variance) to the second state (characterized by a larger variance) when the absolute returns increase in magnitude – or, from the second state to the first one, when absolute returns decrease. Figure 2 shows that this is indeed the case – the graph presents the estimated conditional volatilities of the two components plotted against the time-varying ex-post probability of the first state. Note that when the returns increase in absolute value after a relatively tranquil period, the ex-post probability of the first state decreases, suggesting a switch from the first, less volatile state to the second state. In such a situation both conditional volatilities show an upward jump. The volatilities exhibit a downward jump when the switch is from the second to the first state.

IV.3 Model Selection:

The following selection criteria are used:

(a) **Bayesian Information Criterion**: The model with the lowest BIC is chosen.

(b) **Moment specification tests**: Following Newey (1985), we test for normality in the standardized residuals, checking the first four moments and for zero autocorrelations in their first four powers, using a Wald test. Although this test has the disadvantage that it tends to reject the null too often in small samples (Godfrey, 1988), this test was used for GARCH-type processes by Nelson (1991), Harvey and Siddique (1999) and Brooks, Burke and Persand (2005) and a multivariate test was developed by Ding and Engle (2001). We proceed with the transformation:

\[ u_t = \Phi^{-1} \left( \frac{1}{K} \sum_{i=1}^{K} \Phi \left( z_i \right) \right) \]

where \( \Phi \) is the cumulative normal density function; if the model is the true DGP, then this transformation should lead to i.i.d. standard normal series. We test a total of 20 conditions (the

---

13 Note that the results in Tables 2 – 5 are for variance-annualized unexpected returns. That is daily returns are pre-multiplied by \( \sqrt{250} \) before estimation. Thus, volatilities are quoted in annualized terms.

14 Interestingly, the author have shown using Monte Carlo simulations that the small sample properties of this test’s statistics are good.

15 A similar transformation was used by Haas, Mittnik and Paolella (2004a). This transformation is different from the one of Lanne and Saikkonen (2003), where the resulting series have unit variance and preserve autocorrelation, but are not i.i.d. normal.
first four moments of $u_t$ and the first four autocorrelations in its first four powers) and the test statistics for the moment tests have a $\chi^2(1)$ distribution. The tables report the number of tests (out of 20) that are rejected at 1\% significance level.

(c) **Unconditional density fit:** The density test is on the histogram fit between the model simulated data and the original data. This is one of the most difficult tests for models to pass as it tests for the unconditional distributional fit. The model returns are simulated\(^\text{16}\) and their histogram is estimated using a nonparametric kernel approach. Several alternatives are available for the kernel, our chosen function being that of Epanechnikov (1969). Then the model selection criterion is based on the modified Kolmogorov-Smirnov (KS) statistic (Kolmogoroff, 1933, Smirnov, 1939, Massey, 1951 and Khamis, 2000).

(d) **ACF analysis:** In contrast to (c) this test captures the dynamic properties of the model squared returns – namely, the fit to the empirical autocorrelations of the squared returns. The Appendix states the theoretical autocorrelation functions (ACF) of the different models and we apply the Mean Squared Error (MSE) criterion to assess the fit of the models.

(e) **Value at Risk (VaR) analysis:** The VaR tests developed by Christoffersen (1998) are performed. Using the estimated parameters, the in-sample VaR estimates are computed. Based on these values, the occurrences of exceeding the VaR are counted. This is achieved by building several indicator functions. The first one equals zero except when a return lower than minus the VaR occurs, when it equals one:

$$I_{t;\alpha} = \begin{cases} 1, & r_t < -\text{VaR}_{1-t;\alpha} \\ 0, & \text{otherwise} \end{cases}$$

where $\alpha$ is the level of significance

Then there are a set of four indicator functions:

$$J_{j,i;\alpha} = \begin{cases} 1, & I_{t;\alpha} = i \text{ and } I_{t;\alpha} = j \\ 0, & \text{otherwise} \end{cases}$$

for $i, j = 0, 1$

The three test statistics, for unconditional coverage, independence and conditional coverage are as follows:

$$LR_{uc;\alpha} = -2\log \left( \frac{1 - \alpha}{1 - \pi_{1;\alpha}} \right)^{n_{1;\alpha}} \left( \frac{\pi_{1;\alpha}}{\pi_{1;\alpha}} \right)^{n_{0;\alpha}} \sim \chi^2(1)$$

$$LR_{ind;\alpha} = -2\log \left( \frac{1 - \pi_{2;\alpha}}{1 - \pi_{01;\alpha}} \right)^{n_{01;\alpha}} \left( \frac{\pi_{2;\alpha}}{\pi_{01;\alpha}} \right)^{n_{10;\alpha}} \left( \frac{1 - \pi_{2;\alpha}}{1 - \pi_{11;\alpha}} \right)^{n_{02;\alpha}} \left( \frac{\pi_{2;\alpha}}{\pi_{11;\alpha}} \right)^{n_{12;\alpha}} \sim \chi^2(1)$$

$$LR_{cc;\alpha} = LR_{uc;\alpha} + LR_{ind;\alpha} \sim \chi^2(2)$$

\(^\text{16}\) We simulate returns, based on the estimated parameters, and to ensure that the simulated density is not affected by small sample size we use 50,000 replications. Also, to avoid any influence of the starting values, each simulation has 1000 steps ahead in time but we only use the last simulated return.
where $n_{i}\alpha = \sum_{t=1}^{T} I_{t\alpha}$; $n_{0}\alpha = T - n_{i}\alpha$; $\pi_{i} = n_{i}\alpha / T$; $n_{i}\pi = \sum_{t=1}^{T} J_{i\alpha t}$; $\pi_{01}\alpha = n_{01}\alpha / (n_{01}\alpha + n_{00}\alpha)$; $\pi_{11}\alpha = n_{11}\alpha / (n_{11}\alpha + n_{10}\alpha)$; $\pi_{2}\alpha = (n_{01}\alpha + n_{11}\alpha) / T$

For simplicity, only the first two test statistics are reported and analyzed.

The results of these specification tests are shown in the last four rows of Tables 2 – 5 and in Table 6; these can be interpreted as follows:

(a) **Bayesian Information Criterion:** Most series favour the skewed $t$-AGARCH and skewed $t$-GJR models, except the FTSE for which the symmetric $t$-GJR model is preferable. According to the BIC, normal mixture models are outperformed by $t$-GARCH models. Asymmetric NM-GARCH models fare better than symmetric models, and also models with non-zero means have a lower BIC than normal mixture models that assume zero means for the normal components. However, this is a very simple criterion and it ignores the models’ ability to capture the characteristics of the data.

(b) **Moment specification tests:** These tests show that the most basic models, i.e. (1) – (3) do not capture the higher moments. But beyond this observation, the moment tests do not distinguish well between the models. We find that most models have several rejections for these tests. None of the models performs consistently well based on these moment tests.

(c) **Unconditional density fit:** This shows a clear preference for the NM-GARCH models (13), (14) and (15) (i.e. with different component means). We conclude that it is necessary to model both sources of asymmetry. It is important to notice that the $t$-GARCH models perform quite badly according to this criterion. Often they have very unrealistic (negative) unconditional kurtosis estimates, suggesting that for these models the fourth moment does not exist. To test for the existence of the fourth moment, we performed the HKKP procedure derived by Huisman et al. (2001, 2002) which estimates the maximum moment exponent based on the method introduced by Hill (1975). The estimates for the maximum exponent are found in Table 1.17 As the results illustrate, we can reject the hypothesis that the 4th moment does not exist for all series.

(d) **ACF analysis:** This test also favours the asymmetric NM-GARCH models. Again we see that models based on the $t$ distribution perform badly in many situations and quite often the ACF estimates of these models are negative. This is probably because these models do not capture the fourth moment. By contrast, all the normal mixture models perform very well according to this criterion.

---

17 We have used OLS for the estimation of the maximum exponent instead of WLS because the WLS procedure suggested by Huisman et al. (2001, 2002) actually increased the level of heteroscedasticity of the residuals. To correct for heteroscedasticity we used White’s heteroscedasticity consistent covariance matrix. We used the 500 most extreme observations in the HKKP regression.
(e) **VaR analysis**: The independence and conditional coverage test results in Table 6 show that normal GARCH models are in general very unsuitable for VaR estimations. Additionally, restricted NM-GARCH models fail some of the VaR tests, again proving that they are not suitable for VaR calculations. Symmetric and skewed $t$-GARCH models fail only few of the tests but the NM-AGARCH and NM-GJR models achieve the best results, the latter passing both tests for all indices.

In summary, the normal GARCH model is the worst fit by all criteria and the Student’s $t$-GARCH models fit well according to the BIC criteria but often don’t capture the fourth moment of the returns, performing badly according to the density fit and ACF criteria. The normal mixture models with different component means fit better than those with identical means. We conclude that the in-sample fit of a two-state normal mixture GARCH model with two sources of asymmetry, both persistent and dynamic, is superior to the other models considered. However, for the dynamic asymmetry (i.e. the leverage effect) it is not possible to decide which asymmetric specification for the two volatility components is preferred – this depends on the specific index considered.

**IV.4 Regime Specific Volatility Behaviour**

Each of the two-state normal mixture AGARCH models reveals a lower volatility component that occurs with a high probability (the ‘usual’ component) and a high volatility component with a very low probability (the ‘crash’ component). Table 7 summarizes the characteristics of these volatility components in each of the four European stock markets during the sample period considered. The results in this table are based on the NM-AGARCH model (14). To obtain the annualized mean returns we multiply the mean estimates by $\sqrt{250}$.

In each market the ‘usual’ component is characterized by a high associated probability, a positive mean return and a low volatility. The ‘usual’ mean return is lowest in the FTSE, at only 4.6% per annum but the unconditional volatility is also low: at 15% it is the lowest of all. Notably the Eurostoxx index has the highest return (8.5%) and the second lowest volatility of the four markets (17%). On the other hand, the CAC and DAX indices have higher volatility, around 21%, but their mean return is less than that of Eurostoxx during their usual state. However, they have been in their ‘usual’ regime more often than the FTSE and Eurostoxx: for the latter a ‘usual’ regime occurred only about 95% of the time between 1991 and 2005. In the ‘usual’ regime the least reactive and most persistent volatility is that of the FTSE series but the Eurostoxx and DAX display slightly more reaction and less persistence than the other indices. The leverage effect is most pronounced in the CAC index.

The ‘crash’ market regime occurred between about 2% and 5% of the time, depending on the index. The CAC and the DAX have the highest unconditional volatility (over 40%) and the lowest return during a
crash (it is over −350% in annual terms for the CAC). This is probably because the CAC and DAX indices are less liquid than the FTSE and Eurostoxx. By comparison, during the 1987 crisis the FTSE lost over 400% in annual terms. Yet in our sample the ‘crash’ mean returns are around −100% for the FTSE and −160% for the Eurostoxx. All indices, and the FTSE and CAC in particular, are highly reactive to market shocks in the ‘crash’ regime, yet the effect dies out soon because the persistence parameters are all low, especially on the FTSE. The leverage effects are similar to the ‘usual’ leverage effects for the DAX and Eurostoxx, but the highly reactive FTSE and CAC indices have an inverse leverage effect in a crash regime. This indicates that a positive return will lead to higher volatility than a negative return, but only during the crash regime. A possible explanation of this is that investors anticipate further falls after a modest recovery.

<table>
<thead>
<tr>
<th>Table 7: Summary of Regime Specific Behaviour in Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>‘Usual’ Component</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>‘Crash’ Component</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

V Equity Index Implied Volatility Skews

Out-of-the-money put options on an equity index are an attractive form of insurance for investors that fear a general market decline; and with index option market makers in relatively short supply, the market prices of these options are often far higher than the Black-Scholes (1973) model prices based on the at-the-money volatility. Consequently the implied volatility of these options is commonly found to be higher than the implied volatility of at-the-money call and put options and out-of-the-money calls. This leads to the skew (or ‘smirk’) in equity index implied volatility that has been very pronounced since the global stock market crash in 1987, as shown by Bates (1991), Rubinstein (1994), Jackwerth and Rubinstein (1996), Derman and Kamal (1997), Tompkins (2001) and many others. This equity index implied volatility skew is associated with a negatively skewed implied risk neutral returns density.

Since the global crash of 1987 the skewness in risk neutral index densities has, in general, been much greater than the skewness estimated from historical data on stock index returns. The difference between the physical and risk neutral skews has been the subject of extensive academic research, as in Bakshi et. al (2003), Bates (1997, 2000) and many others. These papers argue that a volatility risk aversion adjustment
is necessary to reconcile the physical and risk neutral distributions of equity indices. The volatility risk aversion can be recovered using empirical pricing kernels based either on unconditional historical returns, as in Ait-Sahalia and Lo (2000) and Jackwerth (2000) or on conditional returns densities, as in Rosenberg and Engle (2002). However, Bates (2003) argued that the difference between the risk-neutral and real-world distributions cannot be explained, except perhaps by the existence of a time-varying volatility risk premium.

One conclusion of this research was that the index skew is too pronounced and too persistent to verify the standard approach and that more sophisticated time series models of the conditional densities of index returns were required. The asymmetric normal mixture GARCH process considered here provides just such a model: since time-varying conditional skewness and kurtosis are endogenous the model volatility skew exhibits a term structure even in the physical measure. With single state GARCH models this is impossible. The uncertainty over two possible volatility states, each of which exhibits volatility clustering but with quite different characteristics, provides a strong justification for the inclusion of a time-varying volatility risk premium.

A natural question to ask, therefore, is how these properties are reflected in the equity skews implied by these models. In this section we compare the implied volatility skews in the physical measure generated by asymmetric normal mixture GARCH models with those implied by other GARCH models. As we shall see, leverage effects are clearly very important. In all markets we find very pronounced normal mixture GARCH skews, even in the absence of a risk premium, and the skew persistence that is captured by the difference in the means of the variance components is only of secondary importance.

For illustration we use the parameter estimates of the FTSE index returns given in Table 5 for a representative selection of five of the fifteen models: normal GARCH, asymmetric t-GARCH, asymmetric t-GJR, NM-GARCH and NM-GJR models. We simulate volatility skew surfaces using each of these five models and compare their characteristics. Starting with $S_0 = 100$ and using $r = 0.03$, we simulate the dynamics of the index value as:

$$S_t = S_{t-1} \exp \left( \left( r - \sigma_t^2 / 2 \right) \Delta t + \varepsilon_t \sqrt{\Delta t} \right)$$

Then, for a fixed strike $K$ and maturity $T$ the time zero price of a European call option is computed as $\exp(-rT)E\left( \max(0, S_T - K) \right)$. Simulating 50,000 times and computing the average call value gives the estimate of the option price. Then, applying the inverse Black-Scholes formula gives the simulated implied volatility at $(K, T)$. We take a range of strikes between 80 and 130 and a range of maturities from 3 to 18 months. The results are shown in Figure 3.
We have used the same vertical scale from 10% – 25% volatility for each smile, as this makes the comparison easier. The first two skews, from the normal and skewed $t$-GARCH models are unrealistic. The volatility level is too low and there is no evidence of a negative skew even for the skewed $t$-GARCH model. The GJR skew in figure (c) is more realistic, with substantially higher volatility for ITM calls than OTM calls. Still, there is little evidence of a volatility term structure as, for a given moneyness, volatility is almost constant as maturity increases. The NM-GARCH without leverage effect has more of a term structure and the short-term skew is not exactly linear with respect the moneyness, a feature that is quite common in index markets. Finally, the NM-GJR model produces the most realistic skew: it is more pronounced and less linear than in the single component GJR model, and there is more variation of volatility over time. We conclude that including non-zero means in the components of the mixture provides a skew that persists into longer dated options but it is the leverage effect that primarily determines the slope of the skew for short dated options.

VII Summary and Conclusions

The majority of generalised autoregressive conditional heteroscedasticity (GARCH) models specify a single time-varying volatility state and thus offer only one scenario for market behaviour. Also their standardized higher moment specification is not realistic: time-variation in conditional skewness and kurtosis is ignored except in a few instances where it is specified exogenously to the GARCH model. This paper considers the normal mixture GARCH model with two volatility states and endogenous time-varying conditional higher moments and introduces additional skewness to model the asymmetry due to leverage effects in equity markets.

We first ask whether this additional source of asymmetry is necessary, given that normal mixture GARCH$(1,1)$ models with different mean returns already exhibit time-varying conditional skewness and kurtosis, in contrast to single-state GARCH models. The answer to this question is undoubtedly yes. Both the statistical criteria and the simulations of the index skew justify the addition of both unconditional and dynamic types of asymmetry. The different component means give a non-zero unconditional skewness, but the addition of dynamic asymmetry is very highly significant and dramatically improves the time series fit of the normal mixture GARCH models.

Empirical results on four European indices compare the fit of these models to single-state normal and symmetric and skewed $t$-GARCH specifications. The overall conclusion from applying five statistical criteria to fifteen fitted models is a clear superiority of asymmetric normal mixture GARCH models. These models are then used to provide considerable insight to the regime specific behaviour of the European equity indices. Between 1991 and 2005 the relative frequency of a ‘crash’ regime varied from about 2% for the CAC to about 5% for the Eurostoxx50. In the ‘usual’ regime the indices have broadly similar behaviour (except that leverage effects are strongest in the CAC and weakest in the Eurostoxx) but
the crash regime behaviour is more diverse. The CAC index displays the most extreme crash behaviour, returning −350% in annual terms with an unconditional volatility of 44%; by contrast the FTSE returned only −100% and the unconditional volatility is far lower. All indices are highly sensitive to market shocks in the crash regime, particularly the CAC, but persistence in volatility is very low indicating that volatility quickly returns to more normal levels. The leverage effect in the CAC and the FTSE is positive (but small) during a crash period, indicating that investors can anticipate further falls after only a modest recovery. Although the Eurostoxx has the highest crash probability, in the ‘usual’ regime it has the highest mean return (8.5%) and a low unconditional volatility (17%), although this is not as low as the FTSE volatility. Since the crash behaviour of the Eurostoxx is also the least extreme, apart from the FTSE, it may be the best investment alternative of the indices considered.

If agents’ beliefs about the future can be informed by past behaviour, the rich structure that is revealed from an estimation of a two-state asymmetric normal mixture GARCH model could provide invaluable information not only to investors, but also to risk managers and policy makers. Stress tests are commonly applied using ad hoc rules, such as changing volatilities to 100% without associating any probability to this event. The NM-GARCH model that we have implemented in this paper has the potential to provide detailed, objective stress tests of equity risk factor volatility and this is clearly an interesting subject for further research. Finally we have demonstrated that single-state GARCH models imply unrealistic shapes for the implied index skew surface. However, asymmetric normal mixture GARCH models, even without a volatility risk premium, provide a rich behavioural structure that can match the empirical characteristics of implied volatility skew surfaces. We conclude that option pricing based on two-state asymmetric GARCH models is another area suitable for further research.
References


Derman, E. and M. Kamal (1997): ‘The Patterns of Change in Implied Index Volatilities’, *Quantitative Strategies Research Notes, Goldman Sachs*


Appendix: Theoretical Properties of Asymmetric NM-GARCH Models

I: Moments

We use the following notations: \( x = E(\varepsilon_i^2) = E(\sigma_i^2) \) and \( y_i = E(\sigma_i^2) \) for \( i = 1, \ldots, K \).

Taking expectations of (3) and (6) yields the result for NM-AGARCH and taking expectations of (4) and (6) yields the result for the NM-GJR model. We get:

\[
x = E(\varepsilon_i^2) = E(\sigma_i^2) = \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i \omega_i^*}{1 - \sum_{i=1}^{K} p_i (1 - \beta_i)}.
\]

\[
y_i = (1 - \beta_i)^{-1}(\omega_i^* + \delta_i x) \quad i = 1, \ldots, K
\]

where

\[
\omega_i^* = \begin{cases} 
\omega_i + \alpha_i \lambda_i^2 & \text{for NM-AGARCH} \\
\omega_i & \text{for NM-GJR} 
\end{cases}
\]

\[
\delta_i = \begin{cases} 
\alpha_i & \text{for NM-AGARCH} \\
\alpha_i + \lambda_i E(d_i^-)(1 + \rho) & \text{for NM-GJR} 
\end{cases}
\]

and \( \rho \) is the correlation between \( d_i^- \) and \( \varepsilon_i^2 \).

The third moment is \( h = E(\varepsilon_i^3) = \sum_{i=1}^{K} p_i E(\varepsilon_i^3) = \sum_{i=1}^{K} p_i \mu_i^3 \) and the skewness can be expressed as:

\[
s = \frac{h}{x^{3/2}}.
\]

The excess kurtosis in both models is:

\[
x = \frac{E(\varepsilon_i^4)}{E(\varepsilon_i^2)^2} - 3 = \frac{z}{x^2} - 3 \quad \text{where} \quad z = E(\varepsilon_i^3) = \frac{3p'^{B^{-1}}f - s}{1 - 3p'^{B^{-1}}g}
\]

The fourth moment uses the following notation:

\[
p = (p_1, \ldots, p_K)^t \quad s = \sum_{i=1}^{K} p_i \left( 6\mu_i^2 y_i^2 + \mu_i^4 \right)
\]

\[
B = \begin{bmatrix} 
1 - \beta_1^2 & -2\beta_1 \beta_2 e_{11} & \ldots & -2\beta_1 \beta_K e_{1K} \\
-2\beta_2 \beta_1 e_{21} & 1 - \beta_2^2 & \ldots & -2\beta_2 \beta_K e_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
-2\beta_K \beta_1 e_{K1} & -2\beta_K \beta_2 e_{K2} & \ldots & 1 - \beta_K^2 - 2\beta_K \beta_K e_{KK} 
\end{bmatrix}
\]

where and \( e_{ij} = a_{ij} p_j \)

---

\(^{18}\) In our empirical application we have approximated \( E(d_i^-) \) by 0.5 and \( \rho \) by 0.
and

\[
A = (a_{ij}) = \begin{bmatrix}
1 - \sum_{k=1}^{K} \frac{p_k \beta_k}{1 - \beta_k} & - \frac{p_2 \beta_2}{1 - \beta_2} & \ldots & - \frac{p_K \beta_K}{1 - \beta_K} \\
- \frac{p_1 \beta_1}{1 - \beta_1} & 1 - \sum_{k=1}^{K} \frac{p_k \beta_k}{1 - \beta_k} & \ldots & - \frac{p_K \beta_K}{1 - \beta_K} \\
\vdots & \vdots & \ddots & \vdots \\
- \frac{p_1 \beta_1}{1 - \beta_1} & - \frac{p_2 \beta_2}{1 - \beta_2} & \ldots & 1 - \sum_{k=1}^{K} \frac{p_k \beta_k}{1 - \beta_k}
\end{bmatrix}
\]

Also:

\[
g = \left( \gamma_1 + 2 \delta_i \beta_i d_i \right) \quad \text{and} \quad f = \left( w_i + 2 \delta_i \beta_i c_i \right)
\]

Furthermore:

\[
w_i = \begin{cases}
\omega_i^2 + x \left( 2 \omega_i x_i + 6 \omega_i \lambda_i^2 \right) + \alpha_i^2 \left( \lambda_i^4 - 4 \lambda_i h \right) + 2 \omega_i \alpha_i \lambda_i^2 + 2 \left( \omega_i + \alpha_i \lambda_i^2 \right) \beta_i y_i, & \text{for NM - AGARCH} \\
\omega_i^2 + 2 \omega_i \delta_i x_i + 2 \omega_i \beta_i y_i, & \text{for NM - GJR}
\end{cases}
\]

\[
c_i = \sum_{j=1}^{K} a_{ij} \left( \sum_{k=1}^{K} \frac{p_k r_{jk}}{1 - \beta_k^2} \right) + y_i q \quad \text{with} \quad q = \sum_{k=1}^{K} p_k r_{kk}^2
\]

\[
r_n = \begin{cases}
\omega_i \omega_k + x \left[ \left( \omega_i \alpha_k + \omega_k \alpha_i \right) + \alpha_i \alpha_k \left( \lambda_i^2 + 4 \lambda_i h + \lambda_k^2 \right) \right] - 2 \alpha_i \alpha_k h \left( \lambda_i + \lambda_k \right) + \beta_i y_i \left( \omega_i + \alpha_i \lambda_i^2 \right) + \omega_i \alpha_i \lambda_i^2 + \omega_k \alpha_i \lambda_i^2 + \alpha_i \lambda_i^2 \lambda_k^2 + \alpha_i \lambda_i^2 \lambda_k^2 + \omega_i \omega_k + x \left( \omega_i \delta_k + \omega_k \delta_i \right) + \beta_i y_i \omega_i + \beta_k y_i \omega_k, & \text{for NM - AGARCH} \\
\omega_i \omega_k + x \left( \omega_i \delta_k + \omega_k \delta_i \right) + \beta_i y_i \omega_i + \beta_k y_i \omega_k, & \text{for NM - GJR}
\end{cases}
\]

II: Parameter Constraints

We require the following set of conditions for the existence and non-negativity of variance and the finiteness of the third moment. These conditions are required for the existence and positivity of the overall variance and individual variances as well. However, as kindly noted by one of the referees, these conditions don’t guarantee the existence of the third moment. These conditions only assure that, if the third moment exists, then it will be finite. Nelson 1990 gave sufficient and necessary conditions for the existence of the moments of the GARCH 1,1 model, proving that for the normal GARCH model given by the equation

\[
\sigma_t^2 = \omega + \left( \alpha_{t-1}^2 + \beta \right) \sigma_{t-1}^2
\]

the \( \rho \)th moment of the error term exists iff

---

19
\[ E\left[ (\alpha_{n-1}^2 + \beta)\right]^{p/2} < 1. \]

For example, when we assume that \( \alpha = 0.09875 \) and \( \beta = 0.9 \) then the second moment exists, but the third and higher moments do not. Unfortunately, no such sufficient and necessary conditions exist for the NM-GARCH model. Also, no straightforward parameter constraint exists for the existence and finiteness of the fourth moment; we simply put \( 0 < E\left( \varepsilon_i^4 \right) < \infty \).

In both the NM-AGARCH and NM-GJR models we require that:
\[
m = \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i \omega_i^*}{\omega_i} > 0, \quad n = \sum_{i=1}^{K} \frac{p_i \left(1 - \delta_i - \beta_i \right)}{\delta_i} > 0 \quad \text{and} \quad \omega_i^* + \delta_i \frac{m}{n} > 0
\]

Also, we must have that \( 0 < p_i < 1, \quad i = 1, \ldots, K-1, \quad \sum_{i=1}^{K-1} p_i < 1, \quad 0 < \alpha_i, \quad 0 \leq \beta_i < 1. \)

**III: Numerical Derivatives of the Asymmetric NM-GARCH Models**

The only difference from the NM-GARCH model numerical derivatives (Alexander and Lazar, 2006) is the first and second order derivatives of \( \sigma_i^2 \) with respect to \( \gamma \) and these are as follows:

\[
\frac{\partial \sigma_i^2}{\partial \gamma} = z_u + \beta_i \frac{\partial \sigma_{i-1}^2}{\partial \gamma}; \quad \frac{\partial^2 \sigma_i^2}{\partial \gamma \partial \gamma} = w_u + \beta_i \frac{\partial^2 \sigma_{i-1}^2}{\partial \gamma \partial \gamma}, \quad \text{with} \quad w_u = A_u + A_u^T.
\]

**NM-AGARCH:** \( z_u = \left(1, \left(\varepsilon_{i-1} - \lambda_i\right)^2, -2\alpha_i \left(\varepsilon_{i-1} - \lambda_i\right), \sigma_{i-1}^2 \right)' \). The starting values \((t = 0)\) are:

\[
\frac{\partial \sigma_{i0}^2}{\partial \gamma_i} = \left(1, s^2 + \lambda_i^2, 2\alpha_i \lambda_i, s^2 \right)' \text{, where } s^2 = \frac{\sum_{i=1}^{T} \varepsilon_i^2}{T},
\]

\[A_u = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \left(\frac{\partial \sigma_{i-1}^2}{\partial \gamma_i}\right)' \end{bmatrix}, \quad \text{with} \quad \frac{\partial^2 \sigma_{i0}^2}{\partial \gamma_i \partial \gamma_i} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 2\lambda_i & s^2 + \lambda_i^2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 2\lambda_i & s^2 + \lambda_i^2 \\ 2\lambda_i & 2\alpha_i & 2\alpha_i \lambda_i & 2s^2 \\ s^2 + \lambda_i^2 & 2\alpha_i \lambda_i & 2s^2 \end{bmatrix}.
\]

**NM-GJR:** \( z_u = \left(1, \varepsilon_{i-1}^2, d_{i-1}^2 \varepsilon_{i-1}^2, \sigma_{i-1}^2 \right)' \). The starting values for this expression \((t = 0)\) are:

\[
\frac{\partial \sigma_{i0}^2}{\partial \gamma_i} = \left(1, s^2, 0.5s^2, s^2 \right)' \text{, where } s^2 = \frac{\sum_{i=1}^{T} \varepsilon_i^2}{T}
\]
\[ A_t = \begin{bmatrix} \sigma_{it}^2 \\ \sigma_{it-1}^2 + \sigma_{it-2}^2 \\ \sigma_{it-2}^2 + \sigma_{it-4}^2 \\ \sigma_{it-4}^2 + \sigma_{it-6}^2 \end{bmatrix} \] with \[ \frac{\partial^2 \sigma_{it}^2}{\partial \gamma_i \partial \gamma_j} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & s^2 \\ 0 & 0 & 0.5s^2 \\ 0 & 1 & 2s^2 \end{bmatrix} \]

IV: ACF of the Squared Errors in the Asymmetric NM-GARCH Models

\[ \varrho_k = \text{Corr}(\varepsilon_i^2, \varepsilon_{i-k}^2) = \frac{\text{Cov}(\varepsilon_i^2, \varepsilon_{i-k}^2)}{\text{Var}(\varepsilon_i^2)} = \frac{E[\varepsilon_i^2 \varepsilon_{i-k}^2] - x^2}{E[\varepsilon_i^2]} - x^2 = \frac{c_k - x^2}{z - x^2}, \text{where} \]

\[ c_k = E[\varepsilon_i^2 \varepsilon_{i-k}^2] = x \sum_{i=1}^K p_i \mu_i^2 + \sum_{i=1}^K p_i[E[\sigma_{it}^2 \varepsilon_{i-k}^2]] = x \sum_{i=1}^K p_i \mu_i^2 + \sum_{i=1}^K p_i b_k^i \]

NM-AGARCH:
\[ b_{ik} = (\omega_i + \alpha_i \lambda_i^2) x + \alpha_i c_{k-1} + \beta_i b_{ik-1}, \quad k > 1 \]

NM-GJR:
\[ b_{ik} = (\omega_i + \alpha_i \lambda_i^2) x + \alpha_i c_0 + \beta_i b_{i0} + -2\alpha_i \lambda_i h, \quad k = 1 \]

The starting values are: \[ c_0 = z \] and \[ b_{i0} = c_i + \delta_i z + \varepsilon_i' B^{-1}(f + g\varepsilon) \].

\(^{20}\) Since the variance of the NM(K)-GARCH(1,1) model can be expressed as a GARCH(K,K) variance, according to Bollerslev (1986) the autocorrelations can also be written as an AR(K) process.
<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pl(d.f. for (4)-(9))</td>
<td>11.5408</td>
<td>12.2000</td>
<td>12.1752</td>
<td>12.1359</td>
<td>12.6697</td>
<td>12.7269</td>
<td>0.9356</td>
<td>0.9790</td>
<td>0.9765</td>
<td>0.9141</td>
<td>0.9780</td>
<td>(30.24)</td>
<td>(79.76)</td>
<td></td>
</tr>
<tr>
<td>μ(Sk..for (7)-(9))</td>
<td>-0.0724</td>
<td>-0.0728</td>
<td>-0.0751</td>
<td>(2.96)</td>
<td>(2.99)</td>
<td>(3.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1</td>
<td>6.4E-4</td>
<td>9.6E-5</td>
<td>5.9E-4</td>
<td>3.9E-4</td>
<td>8.2E-5</td>
<td>4.6E-4</td>
<td>3.9E-4</td>
<td>5.8E-5</td>
<td>4.7E-4</td>
<td>2.2E-4</td>
<td>1.2E-4</td>
<td>4.0E-4</td>
<td>1.7E-4</td>
<td>-1.1E-4</td>
</tr>
<tr>
<td>z1</td>
<td>0.0663</td>
<td>0.0495</td>
<td>0.0166</td>
<td>0.0600</td>
<td>0.0545</td>
<td>0.0166</td>
<td>0.0606</td>
<td>0.0553</td>
<td>0.0173</td>
<td>0.0486</td>
<td>0.0517</td>
<td>0.0157</td>
<td>0.0446</td>
<td>0.0524</td>
</tr>
<tr>
<td>λ1</td>
<td>0.1060</td>
<td>0.0664</td>
<td>0.0302</td>
<td>0.1082</td>
<td>0.0732</td>
<td>0.0302</td>
<td>0.1067</td>
<td>0.0732</td>
<td>0.0302</td>
<td>0.1094</td>
<td>0.0692</td>
<td>0.1058</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β1</td>
<td>0.9186</td>
<td>0.9345</td>
<td>0.9348</td>
<td>0.9310</td>
<td>0.9319</td>
<td>0.9346</td>
<td>0.9303</td>
<td>0.9312</td>
<td>0.9340</td>
<td>0.9386</td>
<td>0.9313</td>
<td>0.9344</td>
<td>0.9432</td>
<td>0.9311</td>
</tr>
<tr>
<td>p2</td>
<td>0.0644</td>
<td>0.0210</td>
<td>0.0235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p2</td>
<td>-0.0890</td>
<td>-0.2235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2</td>
<td>0.0148</td>
<td>0.0087</td>
<td>0.0174</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z2</td>
<td>0.5600</td>
<td>1.6171</td>
<td>1.5287</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λ2</td>
<td>-0.0375</td>
<td>-0.4967</td>
<td>-0.0311</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β2</td>
<td>0.6794</td>
<td>0.6336</td>
<td>0.6352</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional σ</td>
<td>20.35%</td>
<td>20.18%</td>
<td>19.65%</td>
<td>20.86%</td>
<td>20.22%</td>
<td>19.47%</td>
<td>20.69%</td>
<td>20.55%</td>
<td>19.79%</td>
<td>20.80%</td>
<td>21.30%</td>
<td>20.80%</td>
<td>21.77%</td>
<td></td>
</tr>
<tr>
<td>Unconditional σ1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional σ2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional τ</td>
<td>-0.0881</td>
<td>-0.0861</td>
<td>-0.0883</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional k</td>
<td>1.2503</td>
<td>0.8485</td>
<td>1.1207</td>
<td>3.3426</td>
<td>2.6990</td>
<td>4.8650</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>-0.4634</td>
<td>-0.4728</td>
<td>-0.4722</td>
<td>-0.4842</td>
<td>-0.4941</td>
<td>-0.4931</td>
<td>-0.4844</td>
<td>-0.4943</td>
<td>-0.4934</td>
<td>-0.4768</td>
<td>-0.4851</td>
<td>-0.4838</td>
<td>-0.4794</td>
<td>-0.4875</td>
</tr>
<tr>
<td>Moment tests I%</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Density</td>
<td>1.3101</td>
<td>1.4771</td>
<td>0.9944</td>
<td>0.9228</td>
<td>1.1713</td>
<td>1.5195</td>
<td>0.8417</td>
<td>0.9824</td>
<td>0.8340</td>
<td>0.7382</td>
<td>0.8662</td>
<td>0.9965</td>
<td>0.5757</td>
<td>1.1570</td>
</tr>
<tr>
<td>ACF</td>
<td>0.0500</td>
<td>0.1848</td>
<td>0.1033</td>
<td>0.3976</td>
<td>0.0677</td>
<td>0.3012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Estimation results for the CAC 40 Index
<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1(d.f. for (4)-(9))</td>
<td>9.002572</td>
<td>9.537122</td>
<td>9.596825</td>
<td>(11.72)</td>
<td>(11.94)</td>
<td>(11.80)</td>
<td>9.392559</td>
<td>9.8487</td>
<td>9.9290</td>
<td>(100.65)</td>
<td>(118.45)</td>
<td>(98.95)</td>
<td>0.9613</td>
<td>0.9701</td>
<td>0.9604</td>
</tr>
<tr>
<td>μl(Sk..for (7)-(9))</td>
<td>-0.1056</td>
<td>-0.1064</td>
<td>-0.1069</td>
<td>(4.41)</td>
<td>(4.43)</td>
<td>(4.46)</td>
<td>1.5E-4</td>
<td>3.2E-5</td>
<td>2.5E-4</td>
<td>(2.29)</td>
<td>(2.12)</td>
<td>(2.19)</td>
<td>0.0044</td>
<td>0.0041</td>
<td>0.0043</td>
</tr>
<tr>
<td>ω1</td>
<td>8.8E-4</td>
<td>4.7E-4</td>
<td>8.9E-4</td>
<td>(8.02)</td>
<td>(3.80)</td>
<td>(7.83)</td>
<td>3.4E-4</td>
<td>1.4E-4</td>
<td>4.4E-4</td>
<td>(8.48)</td>
<td>(2.87)</td>
<td>(3.83)</td>
<td>0.0787</td>
<td>0.0776</td>
<td>0.0414</td>
</tr>
<tr>
<td>α1</td>
<td>0.0825</td>
<td>0.0668</td>
<td>0.0320</td>
<td>(11.59)</td>
<td>(13.59)</td>
<td>(4.93)</td>
<td>0.0789</td>
<td>0.0772</td>
<td>0.0413</td>
<td>(8.53)</td>
<td>(8.33)</td>
<td>(3.92)</td>
<td>0.0608</td>
<td>0.0664</td>
<td>0.0363</td>
</tr>
<tr>
<td>λ1</td>
<td>0.0764</td>
<td>0.0768</td>
<td>0.0629</td>
<td>(7.82)</td>
<td>(7.53)</td>
<td>(5.01)</td>
<td>0.0628</td>
<td>0.0731</td>
<td>0.0628</td>
<td>(4.69)</td>
<td>(5.00)</td>
<td>0.0633</td>
<td>0.0549</td>
<td>0.0639</td>
<td>0.0538</td>
</tr>
<tr>
<td>β1</td>
<td>0.0938</td>
<td>0.1135</td>
<td>0.09080</td>
<td>(100.01)</td>
<td>(134.74)</td>
<td>(115.60)</td>
<td>0.9155</td>
<td>0.9142</td>
<td>0.9121</td>
<td>(97.01)</td>
<td>(96.17)</td>
<td>(93.55)</td>
<td>0.9282</td>
<td>0.9200</td>
<td>0.9216</td>
</tr>
<tr>
<td>p2</td>
<td>0.0340</td>
<td>0.0238</td>
<td>0.0320</td>
<td>0.0387</td>
<td>0.0320</td>
<td>0.0387</td>
<td>0.0320</td>
<td>0.0387</td>
<td>0.0320</td>
<td>0.0387</td>
<td>0.0320</td>
<td>0.0387</td>
<td>0.0320</td>
<td>0.0387</td>
<td>0.0320</td>
</tr>
<tr>
<td>μ2</td>
<td>-0.1096</td>
<td>-0.1328</td>
<td>-0.1039</td>
<td>0.0354</td>
<td>-0.0150</td>
<td>0.0394</td>
<td>0.0354</td>
<td>-0.0150</td>
<td>0.0394</td>
<td>0.0354</td>
<td>-0.0150</td>
<td>0.0394</td>
<td>0.0354</td>
<td>-0.0150</td>
<td>0.0394</td>
</tr>
<tr>
<td>ω2</td>
<td></td>
<td></td>
<td></td>
<td>0.4442</td>
<td>0.2226</td>
<td>-0.1441</td>
<td>0.4442</td>
<td>0.2226</td>
<td>-0.1441</td>
<td>0.4442</td>
<td>0.2226</td>
<td>-0.1441</td>
<td>0.4442</td>
<td>0.2226</td>
<td>-0.1441</td>
</tr>
<tr>
<td>α2</td>
<td></td>
<td></td>
<td></td>
<td>0.58</td>
<td>0.31</td>
<td>-0.31</td>
<td>0.58</td>
<td>0.31</td>
<td>-0.31</td>
<td>0.58</td>
<td>0.31</td>
<td>-0.31</td>
<td>0.58</td>
<td>0.31</td>
<td>-0.31</td>
</tr>
<tr>
<td>λ2</td>
<td></td>
<td></td>
<td></td>
<td>0.7468</td>
<td>0.8197</td>
<td></td>
<td>0.7468</td>
<td>0.8197</td>
<td></td>
<td>0.7468</td>
<td>0.8197</td>
<td></td>
<td>0.7468</td>
<td>0.8197</td>
<td></td>
</tr>
<tr>
<td>β2</td>
<td></td>
<td></td>
<td></td>
<td>0.7120</td>
<td>0.4451</td>
<td>0.7110</td>
<td>0.7120</td>
<td>0.4451</td>
<td>0.7110</td>
<td>0.7120</td>
<td>0.4451</td>
<td>0.7110</td>
<td>0.7120</td>
<td>0.4451</td>
<td>0.7110</td>
</tr>
<tr>
<td>Unconditional σ</td>
<td>21.31%</td>
<td>20.86%</td>
<td>20.22%</td>
<td>24.85%</td>
<td>22.75%</td>
<td>20.97%</td>
<td>24.10%</td>
<td>23.31%</td>
<td>21.45%</td>
<td>21.83%</td>
<td>21.52%</td>
<td>20.11%</td>
<td>21.40%</td>
<td>21.71%</td>
<td>20.22%</td>
</tr>
<tr>
<td>Unconditional σ1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional σ2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional τ</td>
<td>-0.1517</td>
<td>-0.1467</td>
<td>-0.1464</td>
<td>-7.97E-6</td>
<td>-8.40E-6</td>
<td>-1.01E-5</td>
<td>-0.1517</td>
<td>-0.1467</td>
<td>-0.1464</td>
<td>-7.97E-6</td>
<td>-8.40E-6</td>
<td>-1.01E-5</td>
<td>-0.1517</td>
<td>-0.1467</td>
<td>-0.1464</td>
</tr>
<tr>
<td>BIC</td>
<td>-0.4339</td>
<td>-0.4393</td>
<td>-0.4410</td>
<td>-0.4884</td>
<td>-0.4925</td>
<td>-0.4927</td>
<td>-0.4912</td>
<td>-0.4955</td>
<td>-0.4957</td>
<td>-0.4853</td>
<td>-0.4878</td>
<td>-0.4875</td>
<td>-0.4870</td>
<td>-0.4891</td>
<td>-0.4893</td>
</tr>
<tr>
<td>Moment tests I%</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Density</td>
<td>2.2739</td>
<td>2.3948</td>
<td>1.9538</td>
<td>1.4085</td>
<td>1.8253</td>
<td>2.4724</td>
<td>1.2864</td>
<td>1.4249</td>
<td>1.3516</td>
<td>1.0503</td>
<td>1.2268</td>
<td>2.1285</td>
<td>0.9681</td>
<td>1.0154</td>
<td>1.5470</td>
</tr>
<tr>
<td>ACF</td>
<td>0.1889</td>
<td>0.3591</td>
<td>0.3116</td>
<td>–</td>
<td>0.9764</td>
<td>–</td>
<td>0.0570</td>
<td>–</td>
<td>0.0584</td>
<td>0.1345</td>
<td>0.1800</td>
<td>0.0443</td>
<td>0.0620</td>
<td>0.2201</td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
<td>(10)</td>
<td>(11)</td>
<td>(12)</td>
<td>(13)</td>
<td>(14)</td>
<td>(15)</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>p1(d.f. for (4)-(9))</td>
<td></td>
<td></td>
<td></td>
<td>13.7316</td>
<td>14.1768</td>
<td>14.0745</td>
<td>14.4429</td>
<td>14.6183</td>
<td>14.6046</td>
<td>0.9462</td>
<td>0.9598</td>
<td>0.9587</td>
<td>0.9288</td>
<td>0.9583</td>
<td>0.9564</td>
</tr>
<tr>
<td>p1(Sk., for (7)-(9))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.0679</td>
<td>-0.0657</td>
<td>-0.0669</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.87)</td>
<td>(2.80)</td>
<td>(2.85)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω1</td>
<td>2.5E-4</td>
<td>-4.7E-5</td>
<td>2.4E-4</td>
<td>(4.40)</td>
<td>(-0.54)</td>
<td>(5.39)</td>
<td>2.1E-4</td>
<td>-1.2E-4</td>
<td>2.1E-4</td>
<td>2.0E-4</td>
<td>-1.1E-4</td>
<td>2.1E-4</td>
<td>1.3E-4</td>
<td>-1.7E-4</td>
<td>1.6E-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(3.34)</td>
<td>(-1.02)</td>
<td>(4.26)</td>
<td>(2.52)</td>
<td>(-1.74)</td>
<td>(3.72)</td>
</tr>
<tr>
<td>α1</td>
<td>0.0734</td>
<td>0.0558</td>
<td>0.0155</td>
<td>(10.94)</td>
<td>(8.54)</td>
<td>(2.42)</td>
<td>0.0674</td>
<td>0.0547</td>
<td>0.0114</td>
<td>0.0674</td>
<td>0.0547</td>
<td>0.0122</td>
<td>0.0570</td>
<td>0.0493</td>
<td>0.0074</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(8.19)</td>
<td>(7.32)</td>
<td>(1.75)</td>
<td>(7.47)</td>
<td>(7.25)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>λ1</td>
<td>0.0785</td>
<td>0.0781</td>
<td></td>
<td>(6.39)</td>
<td>(8.96)</td>
<td></td>
<td>0.0862</td>
<td>0.0820</td>
<td></td>
<td>0.0849</td>
<td>0.0810</td>
<td></td>
<td>0.0902</td>
<td>0.0756</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(6.07)</td>
<td>(7.72)</td>
<td></td>
</tr>
<tr>
<td>β1</td>
<td>0.9171</td>
<td>0.9316</td>
<td>0.9350</td>
<td>(118.21)</td>
<td>(129.34)</td>
<td>(137.35)</td>
<td>0.9247</td>
<td>0.9331</td>
<td>0.9379</td>
<td>0.9247</td>
<td>0.9333</td>
<td>0.9379</td>
<td>0.9305</td>
<td>0.9348</td>
<td>0.9416</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(101.89)</td>
<td>(112.72)</td>
<td>(121.93)</td>
<td>(103.05)</td>
<td>(114.24)</td>
<td>(122.91)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(109.31)</td>
<td>(118.63)</td>
<td>(133.69)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional σ</td>
<td>16.24%</td>
<td>15.38%</td>
<td>15.05%</td>
<td>16.13%</td>
<td>15.17%</td>
<td>14.79%</td>
<td>15.95%</td>
<td>15.37%</td>
<td>15.01%</td>
<td>15.86%</td>
<td>15.21%</td>
<td>14.81%</td>
<td>15.72%</td>
<td>15.31%</td>
<td>15.24%</td>
</tr>
<tr>
<td>Unconditional σ1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional σ2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional τ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.0742</td>
<td>-0.0713</td>
<td>-0.0726</td>
<td>-7.26E-6</td>
<td>-8.67E-6</td>
<td>-1.13E-5</td>
</tr>
<tr>
<td>Unconditional k</td>
<td>3.9916</td>
<td>1.5162</td>
<td>3.0582</td>
<td>5.5489</td>
<td>2.7292</td>
<td>8.2826</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.8609</td>
<td>3.8048</td>
<td>3.6864</td>
</tr>
<tr>
<td>BIC</td>
<td>-1.0583</td>
<td>-1.0689</td>
<td>-1.0704</td>
<td>-1.0687</td>
<td>-1.0794</td>
<td>-1.0812</td>
<td>-1.0685</td>
<td>-1.0791</td>
<td>-1.0810</td>
<td>-1.0638</td>
<td>-1.0738</td>
<td>-1.0760</td>
<td>-1.0657</td>
<td>-1.0750</td>
<td>-1.0777</td>
</tr>
<tr>
<td>Moment tests 1%</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Density</td>
<td>60.7739</td>
<td>1.2624</td>
<td>0.6579</td>
<td>0.7213</td>
<td>1.2385</td>
<td>1.7367</td>
<td>1.0483</td>
<td>1.1607</td>
<td>0.9953</td>
<td>0.6626</td>
<td>1.0954</td>
<td>1.6242</td>
<td>0.8146</td>
<td>0.8484</td>
<td>0.8520</td>
</tr>
<tr>
<td>ACF</td>
<td>0.4478</td>
<td>0.0988</td>
<td>0.1832</td>
<td>0.8115</td>
<td>0.0976</td>
<td>0.6602</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.2547</td>
<td>0.1275</td>
<td>0.1529</td>
<td>0.2636</td>
<td>0.1087</td>
<td>0.1239</td>
</tr>
</tbody>
</table>
Table 5. Estimation results for the Eurostoxx 50 Index

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pl(d.f. for (4)-(9))</td>
<td>8.1884</td>
<td>8.5490</td>
<td>8.5869</td>
<td>8.5988</td>
<td>8.8517</td>
<td>8.9319</td>
<td>0.9506</td>
<td>0.9567</td>
<td>0.9571</td>
<td>0.9444</td>
<td>0.9493</td>
<td>0.9490</td>
<td>0.9457</td>
<td>0.9054</td>
<td>0.9055</td>
</tr>
<tr>
<td>μl(Sk..for (7)-(9))</td>
<td>0.0057</td>
<td>0.0084</td>
<td>0.0055</td>
<td>0.0057</td>
<td>0.0084</td>
<td>0.0055</td>
<td>0.0057</td>
<td>0.0084</td>
<td>0.0055</td>
<td>0.0057</td>
<td>0.0084</td>
<td>0.0055</td>
<td>0.0057</td>
<td>0.0084</td>
<td>0.0055</td>
</tr>
<tr>
<td>ω1</td>
<td>4.5E-4</td>
<td>4.5E-4</td>
<td>1.6E-4</td>
<td>4.5E-4</td>
<td>8.30</td>
<td>8.34</td>
<td>2.7E-4</td>
<td>3.1E-4</td>
<td>4.5E-4</td>
<td>2.7E-4</td>
<td>3.1E-4</td>
<td>4.5E-4</td>
<td>2.7E-4</td>
<td>3.1E-4</td>
<td>4.5E-4</td>
</tr>
<tr>
<td>z1</td>
<td>0.0691</td>
<td>0.0586</td>
<td>0.0299</td>
<td>0.0743</td>
<td>0.0720</td>
<td>0.0410</td>
<td>0.0752</td>
<td>0.0734</td>
<td>0.0419</td>
<td>0.0624</td>
<td>0.0661</td>
<td>0.0394</td>
<td>0.0605</td>
<td>0.0642</td>
<td>0.0384</td>
</tr>
<tr>
<td>λ1</td>
<td>0.0709</td>
<td>0.0630</td>
<td>0.0247</td>
<td>0.0508</td>
<td>0.0623</td>
<td>0.0504</td>
<td>0.0635</td>
<td>0.0428</td>
<td>0.0510</td>
<td>0.0447</td>
<td>0.0457</td>
<td>0.0466</td>
<td>0.0432</td>
<td>0.0450</td>
<td>0.0452</td>
</tr>
<tr>
<td>β1</td>
<td>0.9209</td>
<td>0.9208</td>
<td>0.9198</td>
<td>0.9198</td>
<td>0.9194</td>
<td>0.9185</td>
<td>0.9249</td>
<td>0.9208</td>
<td>0.9203</td>
<td>0.9261</td>
<td>0.9206</td>
<td>0.9196</td>
<td>0.9261</td>
<td>0.9206</td>
<td>0.9196</td>
</tr>
<tr>
<td>p2</td>
<td>0.0494</td>
<td>0.0433</td>
<td>0.0429</td>
<td>0.0556</td>
<td>0.0507</td>
<td>0.0510</td>
<td>0.0096</td>
<td>0.1018</td>
<td>0.1030</td>
<td>0.0098</td>
<td>0.0884</td>
<td>0.0089</td>
<td>0.0901</td>
<td>0.0820</td>
<td>0.0839</td>
</tr>
<tr>
<td>a2</td>
<td>0.0140</td>
<td>0.0055</td>
<td>0.0253</td>
<td>0.0139</td>
<td>0.0055</td>
<td>0.0253</td>
<td>0.0139</td>
<td>0.0055</td>
<td>0.0253</td>
<td>0.0139</td>
<td>0.0055</td>
<td>0.0253</td>
<td>0.0139</td>
<td>0.0055</td>
<td>0.0253</td>
</tr>
<tr>
<td>λ2</td>
<td>0.3060</td>
<td>0.2710</td>
<td>0.0625</td>
<td>0.3155</td>
<td>0.2288</td>
<td>0.1122</td>
<td>0.3155</td>
<td>0.2288</td>
<td>0.1122</td>
<td>0.3155</td>
<td>0.2288</td>
<td>0.1122</td>
<td>0.3155</td>
<td>0.2288</td>
<td>0.1122</td>
</tr>
<tr>
<td>β2</td>
<td>0.7860</td>
<td>0.4909</td>
<td>0.7008</td>
<td>0.7996</td>
<td>0.8332</td>
<td>0.8416</td>
<td>0.7996</td>
<td>0.8332</td>
<td>0.8416</td>
<td>0.7996</td>
<td>0.8332</td>
<td>0.8416</td>
<td>0.7996</td>
<td>0.8332</td>
<td>0.8416</td>
</tr>
</tbody>
</table>

Unconditional σ
| Unconditional σ1 | 17.89% | 17.78% | 17.02% | 21.60% | 19.57% | 18.06% | 21.17% | 20.37% | 18.80% | 18.65% | 18.26% | 17.32% | 18.21% | 18.47% | 17.64% |
| Unconditional σ2 | 17.43% | 17.27% | 16.25% | 16.91% | 17.28% | 16.43% |

Unconditional τ
| Unconditional τ1 | -0.159 | -0.156 | -0.157 | -1.05 | -8.26 | -1.20 | -0.20 | -0.1882 | -0.2101 |
| Unconditional τ2 | 34.54% | 33.29% | 33.09% | 31.77% | 32.09% | 31.09% |

Unconditional k
| Unconditional k1 | 1.5502 | 1.2574 | 1.5431 | 1.5870 | 1.3780 | 1.2102 | 3.2070 | 2.8952 | 4.7779 | 3.8641 | 3.3103 |
| Unconditional k2 | 0.7481 | 0.7553 | 0.7547 | 0.7597 | 0.8014 | 0.8015 | 0.8008 | 0.8045 | 0.8047 | 0.7895 | 0.7910 | 0.7912 | 0.7927 | 0.7936 | 0.7943 |

BIC
| BIC1 | 2.1384 | 2.230 | 1.8409 | 1.1173 | 1.5988 | 2.2568 | 1.0173 | 1.1380 | 1.0852 | 1.2707 | 1.2732 | 1.4062 | 0.7634 | 0.9386 | 0.7297 |
| BIC2 | 0.2479 | 0.4044 | 0.3427 | 0.8580 | 0.2339 | 0.0508 | 0.3303 | 0.3867 | 0.0720 | 0.1501 | 0.2778 |
Notes for tables 2, 3, 4 and 5: The models are: (1) GARCH, (2) AGARCH and (3) GJR, all three with normally distributed errors; (4) GARCH, (5) AGARCH and (6) GJR, all three with symmetric Student’s t distributed errors; (7) GARCH, (8) AGARCH and (9) GJR, all three with skewed Student’s t distributed errors; (10) NM-GARCH, (11) NM-AGARCH and (12) NM-GJR, all three normal mixture GARCH models with zero means in the mixture component densities; (13) NM-GARCH, (14) NM-AGARCH and (15) NM-GJR, all three general normal mixture GARCH models. Numbers in parenthesis represent t-values. A ‘-‘ indicates negative kurtosis, or an unreasonably high value for the ACF statistic; meaning that in that model the fourth moment is not defined.
Table 6. In-sample VaR results

<table>
<thead>
<tr>
<th>Test</th>
<th>Significance</th>
<th>Data</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>1%</td>
<td>CAC</td>
<td>7.4***</td>
<td>6.6**</td>
<td>7.4***</td>
<td>2.3</td>
<td>1.9</td>
<td>1.5</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>2.8*</td>
<td>5.2**</td>
<td>5.2**</td>
<td>0.0</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAX</td>
<td>5.9**</td>
<td>5.2**</td>
<td>3.9**</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>2.4</td>
<td>4.6**</td>
<td>4.6**</td>
<td>0.0</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FTSE</td>
<td>2.8*</td>
<td>2.3</td>
<td>1.5</td>
<td>1.1</td>
<td>0.0</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
<td>0.3</td>
<td>1.5</td>
<td>1.1</td>
<td>0.6</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EUROs</td>
<td>14***</td>
<td>9.1***</td>
<td>9.1***</td>
<td>2.9*</td>
<td>0.6</td>
<td>0.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>6.7***</td>
<td>2.9*</td>
<td>4.6**</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>CAC</td>
<td>9.0***</td>
<td>11***</td>
<td>17***</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>1.3</td>
<td>1.9</td>
<td>3.7*</td>
<td>0.9</td>
<td>5.8**</td>
<td>5.8**</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAX</td>
<td>14***</td>
<td>11***</td>
<td>13***</td>
<td>2.0</td>
<td>0.9</td>
<td>1.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.9</td>
<td>2.6</td>
<td>1.9</td>
<td>2.7*</td>
<td>1.9</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FTSE</td>
<td>7.8***</td>
<td>4.0**</td>
<td>5.8**</td>
<td>1.4</td>
<td>0.3</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.6</td>
<td>0.9</td>
<td>1.9</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EUROs</td>
<td>19***</td>
<td>13***</td>
<td>15***</td>
<td>3.7*</td>
<td>3.0</td>
<td>3.0</td>
<td>1.7</td>
<td>0.2</td>
<td>0.8</td>
<td>4.5**</td>
<td>0.8</td>
<td>1.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>CAC</td>
<td>7.2***</td>
<td>5.3**</td>
<td>7.2***</td>
<td>3.7*</td>
<td>2.3</td>
<td>2.3</td>
<td>0.2</td>
<td>1.2</td>
<td>0.4</td>
<td>1.0</td>
<td>0.2</td>
<td>0.2</td>
<td>1.0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAX</td>
<td>14***</td>
<td>9.3***</td>
<td>9.3***</td>
<td>2.3</td>
<td>0.9</td>
<td>1.4</td>
<td>1.2</td>
<td>0.4</td>
<td>0.0</td>
<td>0.2</td>
<td>1.2</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FTSE</td>
<td>14***</td>
<td>11***</td>
<td>14***</td>
<td>1.2</td>
<td>3.7*</td>
<td>5.3**</td>
<td>0.4</td>
<td>1.2</td>
<td>3.7*</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
<td>0.4</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EUROs</td>
<td>41***</td>
<td>24***</td>
<td>31***</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>0.2</td>
<td>0.0</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Conditional</td>
<td>1%</td>
<td>CAC</td>
<td>0.0</td>
<td>0.1</td>
<td>1.6</td>
<td>0.2</td>
<td>1.1</td>
<td>1.1</td>
<td>0.8</td>
<td>0.8</td>
<td>0.9</td>
<td>1.3</td>
<td>1.5</td>
<td>1.5</td>
<td>0.7</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Coverage</td>
<td></td>
<td>DAX</td>
<td>4.0**</td>
<td>4.2**</td>
<td>1.8</td>
<td>3.0*</td>
<td>2.8*</td>
<td>2.7</td>
<td>3.7*</td>
<td>4.0**</td>
<td>0.8</td>
<td>2.1</td>
<td>4.4**</td>
<td>4.4**</td>
<td>3.5*</td>
<td>2.8*</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FTSE</td>
<td>1.2</td>
<td>1.2</td>
<td>1.1</td>
<td>1.0</td>
<td>0.7</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EUROs</td>
<td>0.7</td>
<td>1.1</td>
<td>1.1</td>
<td>2.0</td>
<td>2.7*</td>
<td>2.7*</td>
<td>3.6*</td>
<td>0.7</td>
<td>0.7</td>
<td>1.3</td>
<td>1.3</td>
<td>0.1</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>CAC</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.5</td>
<td>0.2</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAX</td>
<td>3.7*</td>
<td>0.9</td>
<td>0.8</td>
<td>6.5**</td>
<td>2.3</td>
<td>2.1</td>
<td>3.4*</td>
<td>3.4*</td>
<td>3.0*</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FTSE</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EUROs</td>
<td>3.0*</td>
<td>0.7</td>
<td>0.7</td>
<td>1.6</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>CAC</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAX</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FTSE</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EUROs</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Note: The models are: (1) GARCH, (2) AGARCH and (3) GJR, all three with normally distributed errors; (4) GARCH, (5) AGARCH and (6) GJR, all three with symmetric Student’s t distributed errors; (7) GARCH, (8) AGARCH and (9) GJR, all three with skewed Student’s t distributed errors; (10) NM-GARCH, (11) NM-AGARCH and (12) NM-GJR, all three normal mixture GARCH models with zero means in the mixture component densities; (13) NM-GARCH, (14) NM-AGARCH and (15) NM-GJR, all three general normal mixture GARCH models. The reported test statistics follow a χ²(1) distribution. *, ** and *** indicate values significant at 10%, 5% and 1% significance level, respectively.
Fig. 1. The Returns on the Eurostoxx Index

[Graph showing time series data with x-axis from Jan-2004 to Jan-2005 and y-axis ranging from -0.5 to 0.5]

Fig. 2. The Conditional Volatilities and the Ex-post Time-varying Probability of the First State for the Eurostoxx index

[Graph showing time series data with x-axis from Jan-2004 to Jan-2005 and y-axis ranging from -1 to 100%]
Fig. 3. Simulated FTSE Index Skews up to 3 Months

(a) GARCH

(b) GARCH with skewed $t$ distribution

(c) GJR with skewed $t$ distribution
Note: Implied volatility skews for call options are simulated using 50,000 runs. Zero interest rate is assumed. The parameters for the simulation are taken from Table 3.