The Equity Index Skew, Market Crashes and Asymmetric Normal Mixture GARCH

Carol Alexander and Emese Lazar

JEL Classification Codes: C32, G13.

Keywords: GARCH process, normal mixture, equity skew, market crash, skew persistence, leverage effect

Author Details:
Professor Carol Alexander
Chair of Risk Management and Director of Research,
ISMA Centre, Business School,
The University of Reading, PO Box 242
Reading RG6 6BA, United Kingdom
Email: c.alexander@ismacentre.rdg.ac.uk
Tel: +44 (0)1183 786431 (ISMA Centre) Fax: +44 (0)1189 314741

Emese Lazar
PhD Student,
ISMA Centre, Business School,
The University of Reading, PO Box 242
Reading RG6 6BA, United Kingdom
Email: e.lazar@ismacentre.rdg.ac.uk
Tel: +44 (0)1183 786675 (ISMA Centre) Fax: +44 (0)1189 314741

The Equity Index Skew, Market Crashes and Asymmetric Normal Mixture GARCH

Abstract

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 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. In addition, single-state GARCH models naturally converge in the continuous limit to option pricing models that do not price volatility risk. By contrast, GARCH models with two volatility states can have endogenous time-varying conditional higher moments and their continuous limit is a model with uncertainty in volatility. This paper introduces additional asymmetries into the normal mixture GARCH model, leading to multi-state models with several different types of asymmetries and with time-varying higher moments. Empirical results on five major equity indices compare the fit of these models to single-state normal and *t*-GARCH models and many statistical criteria demonstrate the superiority of asymmetric normal mixture GARCH models. The estimated parameters identify agents' beliefs about the likelihood of a crash and the return and volatility behaviour in the event of a crash. Comparing the implied index skew surfaces that are generated by all the GARCH models considered we find that single-state GARCH models imply very unrealistic shapes for the skew whereas the asymmetric normal mixture GARCH models, even without a volatility risk premium, afford a sufficiently rich structure to match the empirical characteristics of implied volatility skew surfaces.

I Introduction

Out-of-the-money put options on an equity index are an attractive form of insurance for investors that fear a general market decline. 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 skew (or 'smirk') in equity index implied volatility 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.

The 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; and thus the shape of the implied volatility skew surface does not match the shape that is expected from standard time-series analysis of historical data. Consequently, 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), Jackwerth (2000) and many others. Also, if the central limit theorem applies the skewness and kurtosis of physical returns densities should diminish as the period of returns increases, yet Carr and Wu (2003) find pronounced risk neutral skews in the SP500 index that persist into options with over one year to expiry.

All this research examines the US stock market since the crash of 1987 and finds that the index skews are too pronounced, too persistent and have too much leverage to accord with standard portfolio theory and time-series analysis of the conditional densities of index returns. The risk aversion adjustments that are necessary to reconcile the physical and risk neutral distributions of equity indices 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). The conclusion they reached is that the difference between the physical and the risk neutral implied distributions can only be explained by a time-varying risk premium that also prices volatility risk. However, to quote from Bates (2003): 'To blithely attribute divergences between objective and risk neutral probability measures to the free 'risk premium' parameters within an affine model is to abdicate one's responsibilities as a financial economist.'

The large literature on risk neutral skews is matched by an even larger literature on conditional moments in the physical measure. These are almost exclusively modelled in the generalized autoregressive conditional heteroscedasticity (GARCH) framework introduced by Engle (1982) and Bollerslev (1986). As the observed non-normalities in both conditional and unconditional returns are higher than can be predicted by normal GARCH(1,1) models, Bollerslev (1987) introduced the Student's *t*-GARCH(1,1) model and Fernandez and Steel (1998) extended this to the skewed *t*-distribution. Then Nelson (1991) introduced an asymmetric volatility response in the exponential GARCH model, which is essential for modelling stock market returns. Bekaert and Wu (2000) and Wu (2001) showed that the 'leverage effect' in equities that has been documented by Black (1976) and many others determines a strong negative correlation between equity returns and volatility and this is perhaps the most important source of skewness in equity index returns. In a recent study, Christoffersen and Jacobs (2004) concluded that a simple asymmetric GARCH, or indeed any GARCH model that captures the leverage effect, performs best of all GARCH model considered.

Many studies (e.g. Christoffersen, Heston and Jacobs, 2004; Bates, 1991 and others) emphasize the connection between time-variability in the physical conditional skewness and the empirical characteristics of option implied volatility skews. However GARCH models that have only one possible state for volatility cannot capture the full extent of skewness and excess kurtosis in option data. Nor do they have time-variation in the conditional higher moments.[1] Nor are they capable of differentiating, via regime specific leverage effects, the mean reversion experienced during different market circumstances.

An option pricing model prices volatility risk only if volatility is stochastic (Fouque, Papanicolau and Sircar, 2000). However, recent research by Corradi (2000) and Alexander and Lazar (2004) proves that GARCH models with a single volatility component do *not* converge to stochastic volatility diffusions (as in Nelson, 1990) unless one makes a very specific assumption about the limiting behaviour of the parameters. That is, the assumption made by Nelson (1990) so that the GARCH limit was a stochastic volatility diffusion model is *ad hoc* and very specific (for instance, it cannot be generalised to GARCH models with more than one volatility component). Moreover, using a different approach, Wang (2002) also concludes that the continuous limit of these single component GARCH models is not a stochastic volatility model. We conclude that single component GARCH models do not price volatility risk and are therefore unable to explain the observed characteristics risk neutral skews by appealing to the existence of a time-varying volatility risk premium.

On the other hand, GARCH models with several possible states for time-varying volatility will price volatility risk, because their volatility is stochastic. They also provide a better fit to the physical conditional densities than GARCH specifications with only one volatility state (e.g. Alexander and Lazar, 2005 and Haas *et al*, 2004). Another main advantage of these models is that the conditional higher moments are time-varying and are determined endogenously. Hence the volatility skews

---

[1] Unless it is added exogenously as, for instance, in Hansen (1994) and Harvey and Siddique (1999).

implied by these models are more likely to exhibit the features of risk neutral index skews, such as a strong leverage effect and more persistence than the central limit theorem would imply.

Two variance processes are generally enough to fit the data; besides, a model with more than two variance components is likely to produce severe bias in parameter estimates because (typically) one variance component occurs with very low frequency (a consequence is that the estimated conditional skewness and excess kurtosis become very unstable over time). The continuous limit of these GARCH models with two variance components is a two factor option pricing model in which the state price density is a normal mixture, with uncertainty about which GARCH variance process governs returns (Alexander and Lazar, 2004). Since this uncertainty cannot be hedged, a volatility risk premium is necessary and time variation in this risk premium may enhance both the skew and the term structure in implied volatilities.

Until now, only symmetric variance components have been considered for normal mixture GARCH models. That is, there is only one source of skewness in the physical returns densities, i.e. that arising from the different means in the normal components of the mixture conditional density. Hence these models are not suitable for equity indices, where the leverage effect is very important. In this paper we first extend the two-state normal mixture GARCH model to include a leverage effect in each variance component. We apply this model to historical data on five major international stock market indices, showing that it provides the best fit amongst fifteen different GARCH models considered, including three symmetric and twelve asymmetric GARCH models, nine of which have only one volatility component and six of which have two volatility components.

Implementing the GARCH models with two asymmetric variance components allows us to draw some new insights about the time series behaviour of stock market returns. From these data we can recover agents' beliefs about the likelihood of a stock market crash, the risks and returns that may be experienced during a crash, the leverage effect during the crash and the persistence of volatility after a crash. Comparison of the risk neutral skews generated by the single component GARCH models with the skews that are generated by the asymmetric normal mixture GARCH models shows that, even without a risk premium, the volatility skew implied by asymmetric normal mixture GARCH models exhibits a pronounced skew that persists into long-dated options. By contrast, none of the other models can explain the observed characteristics of risk neutral skews in stock index markets.

The paper is structured as follows: Section II defines the general normal mixture GARCH model with additional asymmetries and investigates the properties (such as conditional and unconditional moments) of some specific variants. Section III describes the equity index data for five major equity markets (France, Germany, UK, Japan and US) and the estimation methodology. Section IV reports the estimation results for asymmetric and symmetric normal mixture GARCH models with two variance components and for several alternative models including symmetric and skewed *t*-GARCH with both symmetric and asymmetric variance processes. We apply several model selection criteria to identify the best model(s). Section V examines the parameter estimates from the normal mixture GARCH models and makes inferences on the likelihood of, and behaviour during, usual market circumstances and equity market crashes. Then section VI applies the models to simulate the implied equity index skews in the SP500 index and Section VII summarizes and concludes.

II The Asymmetric Normal Mixture GARCH Model

The specification of the model has one equation for the mean and *K* variance equations. For simplicity the conditional mean equation is written $y_t = \varepsilon_t$. It contains no explanatory variables as these can be estimated separately. The error term ε_t is assumed to have a conditional normal mixture density with zero mean, which is a probability weighted average of *K* normal density functions with different means and variances. We write:

$$\varepsilon_t \big| I_{t-1} \sim NM(p_1,\dots,p_K,\mu_1,\dots,\mu_K,\sigma^2_{1t},\dots,\sigma^2_{Kt}), \quad \sum_{i=1}^{K} p_i = 1, \quad \sum_{i=1}^{K} p_i\mu_i = 0$$

That is, the conditional density of the error term is

$$\eta(\varepsilon_t) = \sum_{i=1}^{K} p_i \varphi_i(\varepsilon_t)$$

where φ_i represent normal density functions with different constant means μ_i and different time-varying variances σ^2_{it} for $i = 1,\dots, K$.

The conditional variance behaviour is described by *K* variance components – and these characterize, according to one interpretation, different market circumstances. These variances can follow any GARCH process but for the purpose of this paper we assume there are three possibilities. In addition to the GARCH(1,1) processes studied in Alexander and Lazar (2005) we consider two types of asymmetric processes:

(i) NM-GARCH:

$$\sigma^2_{it} = \omega_i + \alpha_i\varepsilon^2_{t-1} + \beta_i\sigma^2_{it-1} \text{ for } i = 1,\dots, K$$

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

$$\sigma^2_{it} = \omega_i + \alpha_i(\varepsilon_{t-1} - \lambda_i)^2 + \beta_i\sigma^2_{it-1} \text{ for } i = 1,\dots, K$$

(iii) NM-GJRGARCH (based on the Glosten *et al*, 1993 model):

$$\sigma^2_{it} = \omega_i + \alpha_i\varepsilon^2_{t-1} + \lambda_i d^-_{t-1}\varepsilon^2_{t-1} + \beta_i\sigma^2_{it-1} \text{ for } i = 1,\dots, K\text{; where } d^-_t = 1 \text{ if } \varepsilon_t < 0\text{, and 0 otherwise}$$

In all cases the overall conditional variance is

$$\sigma_t^2 = \sum_{i=1}^{K} p_i \sigma_{it}^2 + \sum_{i=1}^{K} p_i \mu_i^2$$

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. For instance, Alexander and Lazar (2005) show that $\alpha_i + \beta_i < 1$ is not required and Haas *et al* (2004) found that $\alpha > 1$ can happen on the second and higher variance components.

This way, we only require the following set of conditions for the non-negativity of variance and the finiteness of the third moment.[2] For $i = 1, \ldots, K$ we must have:

$$0 < p_i < 1, \quad i = 1, \ldots, K-1, \quad \sum_{i=1}^{K-1} p_i < 1, \quad 0 < \alpha_i, \quad 0 \le \beta_i < 1,$$

In the NM-GARCH model (i):

$$m = \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i \omega_i}{(1-\beta_i)} > 0\,, \; n = \sum_{i=1}^{K} \frac{p_i (1-\alpha_i - \beta_i)}{(1-\beta_i)} > 0 \text{ and } \omega_i + \alpha_i \frac{m}{n} > 0$$

In the NM-AGARCH model (ii) we require that:

$$m = \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i (\omega_i + \alpha_i \lambda_i^2)}{(1-\beta_i)} > 0\,, \; n = \sum_{i=1}^{K} \frac{p_i (1-\alpha_i - \beta_i)}{(1-\beta_i)} > 0 \text{ and } \omega_i + \alpha_i \left(\frac{m}{n} + \lambda_i^2 \right) > 0\,,$$

In the NM-GJRGARCH model (iii):

$$m = \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i \omega_i}{(1-\beta_i)} > 0\,, \; n = \sum_{i=1}^{K} \frac{p_i (1-\alpha_i - 0.5\lambda_i - \beta_i)}{(1-\beta_i)} > 0 \text{ and } \omega_i + (\alpha_i + 0.5\lambda_i) \frac{m}{n} > 0$$

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. Appendix 1 shows that even the *un*conditional density will have non-zero skewness, and that this increases with the differentiation of the component means. For instance, when $K = 2$ there is negative skewness in the overall conditional returns density when the component with the higher variance has a negative mean and positive skewness in the overall conditional returns density when the component with the higher variance has a positive mean.

---

[2] There is no straightforward parameter constraint for the existence of the fourth moment. We simply require then that $0 < E(\varepsilon^4) < \infty$.

- *Dynamic Asymmetry*: This only occurs in models (ii) and (iii) and is due to the λ_i parameters in the component variance processes which capture time-varying 'short-term' asymmetries due to the leverage effect. If λ_i is positive the conditional variance in this component is higher following a negative unexpected return at time $t - 1$ than following a positive unexpected return. In equity markets, where 'bad news' normally corresponds to a negative unexpected return, we expect positive λ_i. On the other hand, negative leverage coefficients may be estimated from commodity returns.[3]

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. Not only are conditional higher moments time-varying, the unconditional skewness and excess kurtosis are both non-zero (see Appendix 1). But, since we have conditional normality for each component, the conditional skewness for each component is zero, and thus the unconditional skewness in each component is zero. Hence the unconditional skewness in the overall index returns density stems only from the 'persistent' asymmetry, i.e. from the different means in the components of the normal mixture conditional density.

Normal mixture GARCH can be viewed as a restricted form of the Markov switching GARCH model where the transition probabilities are independent of the past state. These models are 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) and Klaassen (2002). The difficulty with estimating most Markov Switching models lies in the co-dependencies of the state variances. However, normal mixture models have a very straightforward relationship between the states (the individual variances are only tied with each other through their dependence on the error term). Also, the transition probabilities are not historical state-dependent: their focus is to identify the behaviour of the returns during different states and to estimate the long-run probability of each state, based on time series data. As a result they can include quite complex volatility feedback mechanisms in each state and still be relatively easy to estimate. In the remainder of the paper we show that these models give considerably more insight to the behaviour of physical returns densities in equity index markets than has previously been recovered from time series data.

---

[3] Note that the i^{th} variance component depends on the dispersion of the unexpected return, not around its mean μ_i in the individual density, but around the overall mean 0. Hence this third effect induces skewness in each component conditional return density (if the prevailing state is not known) and not in the overall conditional return density.

III Data and Parameter Estimation

Our results will be based on the daily closing prices of five equity market indices: CAC40, DAX30, FTSE100, Nikkei225 and S&P 500 from January 1991 to May 2003. The index prices are shown in Figures 1 –5 and the following table summarises the general characteristics of the daily returns:[4]

	CAC	DAX	FTSE	Nikkei	S&P
Volatility	23.4%	24.91%	18.07%	24.36%	17.45%
Skewness	–0.0472	–0.1485***	–0.0506	0.1417***	–0.0216
Excess kurtosis	2.739***	3.964***	2.657***	2.427***	3.378***

The skewness is negative except for the Nikkei, where it is highly significant and positive due to the downward trend in Japanese equities during most of the period under study. The skewness is also highly significant (and negative) in the DAX. Moderate excess kurtosis is evident, more so in the DAX and the S&P, and the FTSE and S&P indices have been less volatile, overall, than the other markets during this period.

For each index, we estimate the conditional variance parameters separately on the residuals ε_t from AR(p) conditional means equations. All indices had significant positive autocorrelation at this daily frequency and, using information criteria we identified up to $p = 4$ lags.[5] Then, maximizing the likelihood, or equivalently, maximizing

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

gives the optimal parameter values, given the data. One major problem in any type of optimisation is the search for appropriate starting values, to ensure that the optimisation process leads to the global optimum, instead of a local one. To overcome this problem, as suggested by Doornik (2000), an initial grid search is performed. However, the difficulty of optimisation increases with the number of parameters, thus with the number of components in the mixture.

---

[4] For an (annualised) return X the first four moments of its distributions are the mean $\mu = E(X)$, variance $\sigma^2 = E[(X - \mu)^2]$, skewness, $\tau = E[(X - \mu)^3] / \sigma^3$ and excess kurtosis, $k = E[(X - \mu)^4] / \sigma^4 - 3$. The standard errors (s.e.) of the sample estimates of these parameters are as follows: s.e. sample mean = $\sigma/\sqrt{T}$, s.e. sample variance = $\sqrt{2}\,\sigma^2/T$, s.e. of the sample skewness $\approx \sqrt{6/T}$, s.e. of the sample excess kurtosis $\approx \sqrt{24/T}$, where T represents the total number of observations. In the table *** represent results significantly different from zero at the 0.1% level.

[5] We fitted the following AR(p) models:

CAC: $r_t = 0.0039 + 0.0405\, r_{t-1} - 0.0417\, r_{t-2} - 0.0494\, r_{t-3} + 0.0095\, r_{t-4} + \varepsilon_t$
DAX: $r_t = 0.0044 + 0.0066\, r_{t-1} - 0.0381\, r_{t-2} + \varepsilon_t$
FTSE: $r_t = 0.0036 + 0.0218\, r_{t-1} - 0.0506\, r_{t-2} - 0.0493\, r_{t-3} + \varepsilon_t$
NIKKEI: $r_t = -0.0056 - 0.0269\, r_{t-1} - 0.0485\, r_{t-2} + \varepsilon_t$
SP: $r_t = 0.0062 + 0.0133\, r_{t-1} + \varepsilon_t$

The updating formula has the following form, where **g** is the gradient vector, **H** the Hessian matrix and *s* represents the step-length:

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} - s[\mathbf{H}(\boldsymbol{\theta}^{(m)})]^{-1}\mathbf{g}(\boldsymbol{\theta}^{(m)})$$

To compute the Hessian matrix and the gradient vector, we can use either analytic or numerical first and second order derivatives of $L(\boldsymbol{\theta} \mid \varepsilon)$ with respect to $\boldsymbol{\theta}$ – see Appendix 2 for the numerical derivatives.[6]

IV Empirical Results

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.[7] The models are:

A. Models with normally distributed errors:
  (1) GARCH(1,1) *(symmetric)*
  (2) AGARCH(1,1)
  (3) GJR(1,1)

B. Models with symmetric Student's *t* distributed errors:
  (4) GARCH(1,1) *(symmetric)*
  (5) AGARCH(1,1)
  (6) GJR(1,1)

C. Models with skewed Student's *t* distributed errors:
  (7) GARCH(1,1)
  (8) AGARCH(1,1)
  (9) GJR(1,1)

D. Normal mixture GARCH models with zero means in the mixture component densities:
  (10) NM(2)-GARCH(1,1) *(symmetric)*
  (11) NM(2)-AGARCH(1,1)
  (12) NM(2)-GJR(1,1)

E. General normal mixture GARCH models:
  (13) NM(2)-GARCH(1,1)
  (14) NM(2)-AGARCH(1,1)
  (15) NM(2)-GJR(1,1)

---

[6] 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)

[7] 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.

The estimations are reported in Tables 1 – 5.[8] 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 highest weight of the two components in the NM(2)-GARCH models. 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.

The general formula for the time-varying state probability for the i^{th} state in an NM(*K*)-GARCH model is:

$$p_{i,t} = \frac{p_i \varphi_i(\varepsilon_t)}{\sum_{j=1}^{K} p_j \varphi_j(\varepsilon_t)}$$

Figure 1 presents the daily S&P 500 index returns for the period Jan 2001 – May 2003. It is expected that we witness a jump from the first to the second state (characterized with a larger variance) when the absolute returns increase in magnitude. 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 probability of the first state. Note that when, after a relatively normal period, the returns suddenly become extreme, the model switches to the second, more volatile state. In such a situation both conditional volatilities show an upward jump, with the jump in the second volatility being the greater. When the returns revert to the normal market regime, the system switches back to the first state. The difference between this model and the Markov switching approach is that in the latter case low values for the transition probability from the crash to the normal state may have the effect of prolonging the crash period. In the normal mixture framework we find that a switch back to the normal state is very fast. Indeed, crashes don't last long, after only a few days markets revert to normal.

Model Selection

The four model selection criteria used are

(a) *Likelihood*: The model with the largest likelihood is chosen. To account for parsimony the Akaike Information Criterion (AIC) and Schwartz's Bayesian Information Criterion (BIC) were also examined (though not reported in the Tables).

(b) *The Newey (1985) moment specification tests*: Following Newey (1985), we test for normality in the standardized residuals, checking the first four moments and for zero autocorrelations in the powers, using a Wald test. There are a total of 20 conditions and the test statistics for the

---

[8] In Tables 1 – 5 the parameters are estimated by MLE. Numbers in parenthesis represent t-ratios with * and ** signifying significance at the 5% and 1% level respectively.

moment tests have a $\chi^2(1)$ distribution. The Tables report the number of tests (out of 20) that are rejected at 1%.

(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 GARCH models to pass as it tests for the *unconditional* distributional fit. The model returns are simulated[9] 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) *The Autocorrelation Function (ACF) test*: By contrast to (iii) this test captures the *dynamic* properties of the model squared returns – namely, the fit to the empirical autocorrelations of the squared returns. Appendix 3 states the theoretical autocorrelation functions of the different models and we apply the Mean Squared Error (MSE) criterion to assess the fit of the models.

The results of these specification tests are shown in the last four rows of Tables 1 – 5. These are discussed in turn:

(a) *Likelihood*: All series favour the NM(2)-GJRGARCH model with non-zero means in the components, except the Nikkei for which the NM(2)-AGARCH is preferable. This is not just because there are more parameters in these models, as the AIC also supports the likelihood results. Note that the BIC prefers the *t*-GARCH models (4) – (6) for some series.

(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 either most of the models pass all tests (as in the case of the FTSE index), or most have several rejections (as in the DAX). Overall, we can say that the *t*-GARCH models (4) – (9) produce marginally better results on these moment tests.

(c) *Unconditional density fit*: This shows a clear preference for the NM(2)-GJRGARCH with two sources of asymmetry (i.e. with different component means), except for the CAC (which prefers zero mean normals in the mixture).

(d) *ACF test*: This test also favours the asymmetric NM(2) models. First note that whilst one (or more) of the *t*-GARCH models may do a reasonable job to capture the dynamic properties of the squared residuals, these models perform badly for at least one index. The best *t*-GARCH model according to this criterion is model (5), AGARCH(1,1) with Student's *t* distributed

---

[9] 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.

errors. Nevertheless, this has bad MSE results for the DAX, Nikkei and S&P 500 series. By contrast, all the NM(2) models (10) – (15) perform very well according to this criteria.

In summary,

- The two most important tests (c) and (d) indicate a clear superiority of the NM(2)-GARCH models over *t*-GARCH models and the simple symmetric and asymmetric GARCH models, although slightly different specifications do better for different series.
- The 'persistent' source of asymmetry appears to be important in all indices, except the CAC. That is, the non-zero means models (14) and (15) are generally preferred.
- The 'dynamic' asymmetries (i.e. those due to leverage) are also very important – that is, both types of asymmetric components – AGARCH and GJRGARCH – greatly improve the fit.
- The NM(2)-GARCH models perform well according to criteria (a) and (b) as well.
- The normal GARCH(1,1) model is the worst fit by all criteria.
- The Students *t*-GARCH models (4) – (9) fit well according to (a) with the BIC criteria, and (albeit along with other models) they also do well on the moment tests. However, unless an AGARCH or GJR specification is used, the models yield a ridiculously high unconditional volatility.
- Interestingly, the models (7) – (9) don't perform much better than the models (4) – (6). Thus, when the variance process is either AGARCH or GJRGARCH, the additional asymmetry afforded by using the skewed *t*-distribution in place of the standard Student's *t*-distribution appears to be unnecessary.

V GARCH with 'Usual' and 'Crash' Components

This section interprets the parameter estimates for the normal mixture GARCH models. Each individual estimation reveals a high volatility component with a very low probability and a lower volatility component that occurs with a high probability. Clearly these models are capturing a 'usual market circumstances component' and a 'crash component' in equity indices. Since there are two interpretations of the mixing law – as (a) the relative frequency of each state occurring, over a very long period of time; and (b) as a representation of agents' beliefs about which of the two volatility states will govern returns within a relatively short forecast horizon – the implication is that agents' beliefs about the future are informed by the relative frequency of the states observed in the past.

The estimated parameters in models (13) – (15) differ slightly, so the following summary reports only approximate values that apply to all three models. We summarize approximate values for the parameter estimates for the 'usual' market component and the 'crash' market components in the following:[10]

Usual Component	**S&P**	**FTSE**	**CAC**	**DAX**	**Nikkei**[11]
Probability	0.93	0.92	0.96	0.96	0.95
Annualized Mean Return	11%	11%	12%	12%	-15%
Unconditional Volatility	15%	16%	21%	21%	23%
Crash Component	**S&P**	**FTSE**	**CAC**	**DAX**	**Nikkei**
Probability	0.07	0.08	0.04	0.04	0.05
Annualized Mean Return	-121%	-47%	-151%	-124%	153%
Unconditional Volatility	30%	27%	45%	47%	52%

The usual component is characterized by a high associated probability, a positive mean return (except the Nikkei that has an odd behaviour as it has been in decline for many years) and a low volatility. In normal markets the leverage effect is significant (and has the expected sign) but it is not very strong. In the crash market regime – that occurs only very rarely – all the models (13) – (15) have a relatively low beta parameter, and a larger alpha parameter than in the 'usual' component. Hence volatility is large and more reactive but less persistent in crash periods. The mean return is obviously negative and very large in absolute value (except in the Nikkei). Also, the leverage effect is much more pronounced than it is under normal circumstances.

A fact that cannot not be recovered from any single-component GARCH model, let alone simple skewness statistics, is that the crash market leverage effect is exceptionally large in the US market (especially during crash periods). Since the mean return in the US crash markets is large and negative the rapid mean-reversion in volatility during crash markets in the US is one reason why the unconditional volatility associated with the crash market is only about 30%.

VI Equity Index Implied Volatility Skews

A 'stylized fact' that has emerged from the research surveyed in the introduction is that the index skew is too pronounced and too persistent to accord with the time-series analysis of the conditional densities of index returns. But whilst standard time series model may have supported this conclusion, the asymmetric normal mixture GARCH processes considered in this paper are not standard time series models. They have time-varying conditional skewness and kurtosis, so the volatility skew will

---

[10] Note that the detailed results on Tables 1 – 5 are for variance-annualized 'unexpected' returns. That is the residuals from the conditional mean equations are pre-multiplied by √250 before estimation. Thus volatilities are quoted there in annualized terms, but to obtain the annualized mean returns we multiply the mean estimates by √250 and adjust for the expected return given by the conditional mean equation.

[11] In the Nikkei the exceptional component is actually an *upward* jump in returns component

exhibit a term structure even in the physical measure; this is not possible with most other GARCH models. Normal mixture GARCH models are also able to distinguish between two sources of asymmetry in physical returns distributions – a dynamic leverage effect and a more persistent asymmetry in the skew that arises from the difference between the mean returns under different market circumstances. We have also seen how the model quantifies differential leverage effects, which are much stronger in crash markets than in normal markets. Furthermore, we can recovers agents' beliefs about the likelihood of a market crash and learn about the behaviour of returns and volatility during each market regime. Finally, the uncertainty over two possible volatility states, each of which exhibits volatility clustering but with quite different characteristics, provides strong justification for the inclusion of a volatility risk premium whereas single state GARCH models have no such justification.

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 generated by asymmetric normal mixture GARCH models with those implied by other GARCH models. Since only the normal mixture GARCH models have a volatility risk premium we compare the physical skews from the different models. Leverage effects are clearly very important, especially in the S&P 500. In all markets we find very pronounced normal mixture GARCH skews, even in the absence of a risk premium, where the skew persistence that is captured by the difference in means of the variance components is only of secondary importance.

For illustration we use the parameter estimates of the S&P index returns given in Table 5 for the models given in the previous section.[12] We simulate volatility skew surfaces using each of these models and compare their characteristics. Since our results are based on daily returns, we also simulate daily returns. Starting with $S_0 = 100$ and using $r = 0.03$, we simulate the dynamics of the index value as:

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

For a fixed strike K and maturity T the time zero price of a European call option is computed as $exp\left(-rT\right)E\left(max\left(0, S_T - K\right)\right)$. Repeating this procedure 50,000 times and computing their average 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.

Figure 3 presents the skews based on the normal GARCH, asymmetric *t*-GARCH, asymmetric *t*-GJR, NM(2)-GARCH and NM(2)-GJR models . For comparison, we have tried to use the same vertical

---

[12] Simulated skews for the other equity indices and other models are available from the authors on request.

scale from 0 – 25% volatility for each smile.[13] Notice that the normal GARCH(1,1) skew is almost completely flat, there being nothing in the model to capture asymmetry or term structure, except a mean reversion in the deterministic variance process. The GJR skew in figure (c) is more realistic, with substantially higher volatility for ITM calls than OTM calls. Also, the additional asymmetry afforded by using the skewed *t*-distribution in place of the standard Student's *t*-distribution appears to be unnecessary. Interestingly, we reached exactly the same conclusion based on our statistical tests. But again, there is no uncertainty in the model and hence we find very little term structure in the skew.

As expected, the NM(2)-GJR model produces much the most realistic skew, even without including a volatility risk premium. Not only is the skew very pronounced and less linear than in the single component GJR model, there is much more variation of volatility over time. We again find that including non-zero means in the components of the mixture is less important than the leverage effect. On the other hand, comparing figure 3(e) with the best of the single component GJR parameterizations (figure 3(c)) the additional component in figure 3(e) allows for a richer structure in the skew, which now has a noticeable term structure.

VII Summary and Conclusions

This paper has introduced additional asymmetry into normal mixture GARCH models, which already model persistence in the skew, to capture a time-varying and regime specific leverage effect. The GJR (or AGARCH) components capture a dynamic leverage effect whilst the different component means capture a more persistent skew effect. We first ask whether this additional source of asymmetry is necessary, given that normal mixture GARCH(1,1) models with symmetric components 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 types of asymmetry. The different component means give a non-zero unconditional skewness is non-zero, but the addition of dynamic asymmetry is very highly significant and dramatically improves the time series fit of the normal mixture GARCH models with symmetric components.

To summarize our empirical findings, we have demonstrated the clear superiority of normal mixture GARCH models over single volatility component GARCH models including the GJR parameterization and the skewed and standard Student's *t*-GARCH models. Our results are also supported by a powerful behavioural interpretation for the theoretical models where traders' beliefs about the likelihood of a market crash, and the returns and volatility behaviour during the crash period

---

[13] However, for the *t*-GARCH skew (figure 3(b)) this was impossible. Recall from Table 5 that the long-term volatility estimates from this model was improbably high (at 32.43% for the asymmetric *t*-GARCH). No surprise then that the volatility skew given by this model (and the symmetric *t*-GARCH model as well) is completely unrealistic.

and during ordinary market circumstances, can be recovered from the physical data. Over the data period considered (January 1991 to May 2003) the perceived likelihood of a crash was least in the Japanese, French and German markets (about 4%); next comes the US with the crash likelihood of about 6% and finally the UK, where there is an implied probability of 8% for a crash scenario.

The UK and US markets have the lowest crash volatilities, of 30% compared with 45-50% in the Japanese, French and German markets. We expect that the index falls further during a crash market in the US than it does in the UK. There is a very pronounced leverage effect during crash markets in the US and since traders believe that the market could crash (albeit with a low probability) the leverage effect dominates the long term persistence features in the S&P500 index implied volatility skew surface.

References

Ait-Sahalia, Y. and A.W. Lo (2000): 'Nonparametric risk management and implied risk aversion', *Journal of Econometrics* 94, 9-51.

Alexander, C. and E. Lazar (2004): 'The continuous limit of normal mixture GARCH processes' *ISMA Centre Discussion Papers in Finance* DP2004-10 available from www.ismacentre.rdg.ac.uk/dp

Alexander, C. and Lazar, E. (2005): 'Normal mixture GARCH(1,1): Applications to exchange rate modelling', *ISMA Centre Discussion Paper* 2004-06, forthcoming in *Journal of Applied Econometrics*

Bakshi, G., Kapadia, N. and Madan, D. (2003): 'Stock return characteristics, skew laws, and the differential pricing for individual equity options', *The Review of Financial Studies* 16, 101-143

Bates, D.S. (1991): 'The crash of '87: Was it expected? The evidence from options markets', *Journal of Finance* 46, 1009-1044

Bates, D.S. (1997): 'The skewness premium: option pricing under asymmetric processes', *Advances in Futures and Options Research* 9, 51-82

Bates, D.S. (2000): 'Post-'87 Crash fears in the S&P 500 futures option market', *Journal of Econometrics* 94, 181-238

Bates, D.S. (2001): 'The market for crash risk' *University of Iowa and NBER,* Working paper available from http://www.biz.uiowa.edu/faculty/dbates/

Bates, D.S. (2003): 'Empricial option pricing: A retrospection', *The Journal of Econometrics 116, 387 – 404*

Bekaert, G. and Wu. G. (2000): 'Asymmetric volatility and risk in equity markets', *The Review of Financial Studies*, 13(1), 1-42

Black, F. (1976): 'Studies of stock price volatility changes', *Proceedings of the 1976 Meetings of the American Statistical Association,* 177-181

Black, F. and Scholes, M. (1973): 'The pricing of options and corporate liabilities', *Journal of Political Economy* 81, 637-659

Brigo, D. and F. Mercurio (2001): 'Displaced and mixture diffusions for analytically-tractable smile models', in: Geman, H., Madan, D.B., Pliska, S.R., Vorst, A.C.F. (Editors), *Mathematical Finance Bachelier Congress 2000*, Springer, Berlin

Chernov, M. and E. Ghysels (2000): "A Study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation" *Journal of Financial Economics* 56(3), 407-458.

Cai, J. (1994): 'A Markov model of switching-regime ARCH', *Journal of Business & Economic Statistics* 12, 309-316

Carr, P. and Wu, L. (2003): 'The finite moment log stable process and option pricing', *Journal of Finance* 58, 753-777

Christoffersen, P., Heston, S and Jacobs, K. (2004): 'Option valuation with conditional skewness', forthcoming in *The Journal of Econometrics*

Christoffersen, P. and Jacobs, K. (2004): 'Which GARCH model for option valuation?', *Management Science* 50, 1204-1221

Corradi, V. (2000): 'Reconsidering the continuous time limit of the GARCH(1,1) process', *Journal of Econometrics* 96, 145-153

Derman, E. and Kamal, M. (1997): 'The patterns of change in implied index volatilities', *Quantitative Strategies Research Notes, Goldman Sachs*

Engle, R. F. and Ng, V. K. (1993): 'Measuring and testing the impact of news on volatility', *Journal of Finance* 48, 1749 - 1778

Fernandez, C. and Steel, M. (1998): 'On Bayesian modelling of fat tails and skewness', *Journal of the American Statistical Association* 93, 359-371

Fouque, J.P., Papanicolau, G. and Sircar, K.R. (2000): 'Derivatives in financial markets with stochastic volatility', Cambridge University Press

Glosten, L. R, Jagannathan, R. and Runkle, D. E. (1993): 'On the relationship between expected value and the volatility of excess returns on stocks', *Journal of Finance* 48, 1779-1801

Gray, S. F. (1996): 'Modelling the conditional distribution of interest rates as a regime-switching process', *Journal of Financial Economics* 42, 27-62

Haas, M., Mittnik, S. and Paolella, M.S. (2004): 'Mixed normal conditional heteroskedasticity', *Journal of Financial Econometrics* 2(2), 211-250

Hamilton, J.D. and Susmel, R. (1994): 'Autoregressive conditional heteroscedasticity and changes in regime', *Journal of Econometrics* 64, 307-333

Hansen, B. E. (1994): 'Autoregressive conditional density estimation', *International Economic Review* 35, 705-730

Harvey, C.R. and Siddique, A. (1999): 'Autoregressive conditional skewness', *Journal of Financial and Quantitative Studies* 34(4), 465-487

Jackwerth, J. (2000): 'Recovering risk aversion from option prices and realized returns', *Review of Financial Studies* 13, 433-451.

Jackwerth, J.C., and Rubinstein, M. (1996): 'Recovering probabilities from option prices', *Journal of Finance* 51, 1611-1631

Khamis, H. J. (2000): 'The two-stage δ-corrected Kolmogorov-Smirnov test', *Journal of Applied Statistics* 27(4), 439-450

Klaassen, F. (2002): 'Improving GARCH volatility forecasts with regime-switching GARCH', *Empirical Economics* 27, 363-394

Kolmogoroff, A. (1933): 'Sulla determinazione empirica di una legge di distribuzione', *Giornale dell'Istituto Italiano degh Attuari* 4, 83-91

Konikov, M. and Madan, D. (2002): 'Option pricing using variance gamma markov chains', *Review of Derivatives Research* 5, 81-115

Massey, F. J. Jr. (1951): 'The Kolmogorov-Smirnov test for goodness of fit', *Journal of the American Statistical Association* 46(253), 68-78

Nelson, D.B. (1990): 'ARCH models as diffusion approximations', *Journal of Econometrics* 45, 7-38

Nelson, D. B. (1991): 'Conditional heteroscedasticity in asset returns: a new approach', *Econometrica* 59, 347-370

Newey, W.K. (1985): 'Generalized method of moments specification testing', *Journal of Econometrics* 29, 229-256

Rosenberg, J.V. and Engle, R. F. (2002): 'Empirical pricing kernels', *Journal of Financial Economics* 64, 341–372.

Rubinstein, M. (1994): 'Implied binomial trees', *Journal of Finance* 49, 771-818

Smirnov, N. (1939): 'Sur les ecarts de la courbe de distribution empirique' (Russian, French summary), *Matematicheskii Sbornik* 48(6), 3-26

Tompkins, R.G. (2001): 'Implied volatility surfaces: uncovering regularities for options on financial futures', *The European Journal of Finance* 7, 198-230

Wang, Y. (2002) 'Asymptotic nonequivalence of GARCH models and diffusions' *The Annals of Statistics* 30, 754-783

Wu, G. (2001): 'The determinants of asymmetric volatility', *The Review of Financial Studies* 14(3), 837-859

Appendix 1: Moments of the Asymmetric Normal Mixture GARCH Models

We use the following notations: $x = E(\varepsilon_t^2) = E(\sigma_t^2)$ and $y_i = E(\sigma_{it}^2)$ for $i = 1, \ldots, K$.

Taking expectations of (1) and (2) gives for NM-AGARCH

$$x = E(\varepsilon_t^2) = E(\sigma_t^2) = \frac{\sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i\left(\omega_i + \alpha_i \lambda_i^2\right)}{(1-\beta_i)}}{\left(1 - \sum_{i=1}^{K} \frac{p_i \alpha_i}{(1-\beta_i)}\right)}$$

$$y_i = \left(1-\beta_i\right)^{-1}\left(\omega_i + \alpha_i\left(x + \lambda_i^2\right)\right) \qquad i = 1, \ldots, K$$

Taking expectations of (2) and denoting α_i+0.5λ_i by δ_i gives for NM-GJRGARCH

$$x = E(\varepsilon_t^2) = E(\sigma_t^2) = \frac{\sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i \omega_i}{(1-\beta_i)}}{\left(1 - \sum_{i=1}^{K} \frac{p_i \delta_i}{(1-\beta_i)}\right)},$$

$$y_i = \left(1-\beta_i\right)^{-1}\left(\omega_i + \delta_i x\right) \qquad i = 1, \ldots, K$$

assuming (approximately) that d_t^- and ε_t^2 are independent.

Taking expectations of (3) gives:

$$x = \sum_{i=1}^{K} p_i y_i + \sum_{i=1}^{K} p_i \mu_i^2$$

The third moment is:

$$h = E(\varepsilon_t^3) = \sum_{i=1}^{K} p_i E_i(\varepsilon_t^3) = \sum_{i=1}^{K} p_i \mu_i \left(3 y_i + \mu_i^2\right)$$

and the skewness can be expressed as $s = \dfrac{h}{x^{3/2}}$.

The excess kurtosis in both models is:

$$\kappa = \frac{E(\varepsilon_t^4)}{E(\varepsilon_t^2)^2} - 3 = \frac{z}{x^2} - 3$$

where $$z = E(\varepsilon_t^4) = \frac{3\mathbf{p}'\mathbf{B}^{-1}\mathbf{f} - s}{1 - 3\mathbf{p}'\mathbf{B}^{-1}\mathbf{g}}$$

The fourth moment uses the following notation:

$$\mathbf{p} = (p_1, \dots, p_K)' \qquad s = \sum_{i=1}^{K} p_i \left(6\mu_i^2 y_i^2 + \mu_i^4\right)$$

$$\mathbf{B} = \begin{bmatrix} 1-\beta_1^2 - 2\delta_1\beta_1 e_{11} & -2\delta_1\beta_1 e_{12} & \dots & -2\delta_1\beta_1 e_{1K} \\ -2\delta_2\beta_2 e_{21} & 1-\beta_2^2 - 2\delta_2\beta_2 e_{22} & \dots & -2\delta_K\beta_K e_{2K} \\ \vdots & \vdots & \vdots & \vdots \\ -2\delta_K\beta_K e_{K1} & -2\delta_K\beta_K e_{K2} & \dots & 1-\beta_K^2 - 2\delta_K\beta_K e_{KK} \end{bmatrix}$$

where $\delta_i = \begin{cases} \alpha_i & \textit{in } (1) \\ \alpha_i + 0.5\lambda_i & \textit{in } (2) \end{cases}$, and $e_{ij} = a_{ij} p_j$

and
$$(a_{ij}) = \begin{bmatrix} 1 - \sum\limits_{\substack{k=1 \\ k\neq 1}}^{K} \dfrac{p_k\beta_1\delta_k}{1-\beta_1\beta_k} & -\dfrac{p_2\delta_1\beta_2}{1-\beta_1\beta_2} & \dots & -\dfrac{p_K\delta_1\beta_K}{1-\beta_1\beta_K} \\ -\dfrac{p_1\delta_2\beta_1}{1-\beta_2\beta_1} & 1 - \sum\limits_{\substack{k=1 \\ k\neq 2}}^{K} \dfrac{p_k\beta_2\delta_k}{1-\beta_2\beta_k} & \dots & -\dfrac{p_K\delta_2\beta_K}{1-\beta_2\beta_K} \\ \vdots & \vdots & \vdots & \vdots \\ -\dfrac{p_1\delta_K\beta_1}{1-\beta_K\beta_1} & -\dfrac{p_2\delta_K\beta_2}{1-\beta_K\beta_2} & \dots & 1 - \sum\limits_{\substack{k=1 \\ k\neq K}}^{K} \dfrac{p_k\beta_K\delta_k}{1-\beta_K\beta_k} \end{bmatrix}$$

$$\mathbf{g} = \begin{pmatrix} \gamma_1 + 2\delta_1\beta_1 d_1 \\ \vdots \\ \gamma_K + 2\delta_K\beta_K d_K \end{pmatrix}$$

where $d_i = \sum\limits_{j=1}^{K} a_{ij} \left(\sum\limits_{\substack{k=1 \\ k\neq j}}^{K} \dfrac{p_k\delta_j\delta_k}{1-\beta_j\beta_k} \right)$ and $\gamma_i = \begin{cases} \alpha_i^2 & \textit{in } (1) \\ \alpha_i^2 + \lambda_i\delta_i & \textit{in } (2) \end{cases}$.

$$\mathbf{f} = \begin{pmatrix} w_1 + 2\delta_1\beta_1 c_1 \\ \vdots \\ w_K + 2\delta_K\beta_K c_K \end{pmatrix}$$

Furthermore:

$$w_i = \begin{cases} \omega_i^2 + x\left(2\omega_i\alpha_i + 6\alpha_i^2\lambda_i^2\right) + \alpha_i^2\left(\lambda_i^4 - 4\lambda_i h\right) + 2\omega_i\alpha_i\lambda_i^2 + 2\left(\omega_i + \alpha_i\lambda_i^2\right)\beta_i y_i & \textit{in } (1) \\ \omega_i^2 + 2\omega_i\delta_i x + 2\omega_i\beta_i y & \textit{in } (2) \end{cases}$$

$$c_i = \sum_{j=1}^{K} a_{ij} \left[\left(\sum_{\substack{k=1 \\ k\neq j}}^{K} \frac{p_k r_{jk}}{1-\beta_j\beta_k} \right) + y_j q \right] \text{with} \quad q = \sum_{k=1}^{K} p_k \mu_k^2$$

and

$$\begin{aligned} r_{ik} = {} & \omega_i\omega_k + x\left[\left(\omega_i\alpha_k + \omega_k\alpha_i\right) + \alpha_i\alpha_k\left(\lambda_i^2 + 4\lambda_i\lambda_k + \lambda_k^2\right)\right] - 2\alpha_i\alpha_k h\left(\lambda_i + \lambda_k\right) \\ & + \beta_i y_i\left(\omega_k + \alpha_k\lambda_k^2\right) + \beta_k y_k\left(\omega_i + \alpha_i\lambda_i^2\right) + \omega_i\alpha_k\lambda_k^2 + \omega_k\alpha_i\lambda_i^2 + \alpha_i\alpha_k\lambda_i^2\lambda_k^2 \end{aligned} \quad \text{in (1)}$$

$$r_{ik} = \omega_i\omega_k + x\left(\omega_i\delta_k + \omega_k\delta_i\right) + \beta_i y_i\omega_k + \beta_k y_k\omega_i \qquad \text{in (2).}$$

The full derivation of these results is available from the authors on request.

Appendix 2: Numerical Derivatives of the Asymmetric Normal Mixture GARCH Models

The only difference from the NM(K)-GARCH model numerical derivatives (Alexander and Lazar, 2005) is the first and second order derivatives of σ_{it}^2 with respect to $\boldsymbol{\gamma}_i$ and these are as follows:

$$\frac{\partial \sigma_{it}^2}{\partial \boldsymbol{\gamma}_i} = \boldsymbol{z}_{it} + \beta_i \frac{\partial \sigma_{it-1}^2}{\partial \boldsymbol{\gamma}_i}$$

$$\frac{\partial^2 \sigma_{it}^2}{\partial \boldsymbol{\gamma}_i \partial \boldsymbol{\gamma}_i'} = w_{it} + \beta_i \frac{\partial^2 \sigma_{it-1}^2}{\partial \boldsymbol{\gamma}_i \partial \boldsymbol{\gamma}_i'} \qquad w_{it} = A_{it} + A_{it}^T$$

For NM-AGARCH:

$z_{it} = (1, (\varepsilon_{t-1} - \lambda_i)^2, -2\alpha_i(\varepsilon_{t-1} - \lambda_i), \sigma_{it-1}^2)'$. The starting values for this expression (for t=0) are:

$$\frac{\partial \sigma_{i0}^2}{\partial \boldsymbol{\gamma}_i} = (1, s^2 + \lambda_i^2, 2\alpha_i\lambda_i, s^2)', \text{ where } s^2 = \frac{\sum_{t=1}^{T} \varepsilon_t^2}{T}$$

$$A_{it} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2(\varepsilon_{t-1} - \lambda_i) & \alpha_i & 0 \\ \multicolumn{4}{c}{\left(\dfrac{\partial \sigma_{it-1}^2}{\partial \boldsymbol{\gamma}_i}\right)'} \end{bmatrix}$$

The starting values for this computation are $\dfrac{\partial^2 \sigma_{i0}^2}{\partial \boldsymbol{\gamma}_i \partial \boldsymbol{\gamma}_i'} = \dfrac{1}{(1-\beta_i)} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2\lambda_i & s^2 + \lambda_i^2 \\ 0 & 2\lambda_i & 2\alpha_i & 2\alpha_i\lambda_i \\ 1 & s^2 + \lambda_i^2 & 2\alpha_i\lambda_i & 2s^2 \end{bmatrix}$

For NM-GJRGARCH:

$\boldsymbol{z}_{it} = \left(1, \varepsilon_{t-1}^2, d_{t-1}^- \varepsilon_{t-1}^2, \sigma_{it-1}^2\right)$. The starting values for this expression (for $t = 0$) are:

$$\frac{\partial \sigma_{i0}^2}{\partial \boldsymbol{\gamma}_i} = (1, s^2, 0.5s^2, s^2)', \text{ where } s^2 = \frac{\sum_{t=1}^{T} \varepsilon_t^2}{T}$$

$$A_{it} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \multicolumn{4}{c}{\left(\dfrac{\partial \sigma_{it-1}^2}{\partial \boldsymbol{\gamma}_i}\right)'} \end{bmatrix}$$

and the starting values for this computation are $\dfrac{\partial^2 \sigma_{i0}^2}{\partial \boldsymbol{\gamma}_i \partial \boldsymbol{\gamma}_i'} = \dfrac{1}{(1-\beta_i)} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & s^2 \\ 0 & 0 & 0 & 0.5s^2 \\ 1 & s^2 & 0.5s^2 & 2s^2 \end{bmatrix}$

Appendix 3: The Autocorrelation Function of the Squared Errors in the Asymmetric Normal Mixture GARCH Models

The autocorrelations of the squared errors can be expressed as:[14]

$$\varrho_k = Corr\left(\varepsilon_t^2, \varepsilon_{t-k}^2\right) = \frac{Cov\left(\varepsilon_t^2, \varepsilon_{t-k}^2\right)}{Var\left(\varepsilon_t^2\right)} = \frac{E[\varepsilon_t^2 \varepsilon_{t-k}^2] - x^2}{E[\varepsilon_t^4] - x^2} = \frac{c_k - x^2}{z - x^2},$$

$$c_k = E\left[\varepsilon_t^2 \varepsilon_{t-k}^2\right] = x \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} p_i E\left[\sigma_{it}^2 \varepsilon_{t-k}^2\right] = x \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} p_i b_{ik}$$

For NM-AGARCH:

$$\begin{aligned} b_{ik} &= (\omega_i + \alpha_i \lambda_i^2) x + \alpha_i c_{k-1} + \beta_i b_{ik-1}, && k > 1 \\ b_{i1} &= (\omega_i + \alpha_i \lambda_i^2) x + \alpha_i c_0 + \beta_i b_{i0} + -2\alpha_i \lambda_i h, && k = 1 \end{aligned}$$

For NM-GJRGARCH:

$$b_{ik} = \omega_i x + (\alpha_i + 0.5\lambda_i) c_{k-1} + \beta_i b_{ik-1}$$

The starting values are: $c_0 = z$ and $b_{i0} = c_i + d_i z + \mathbf{e}_i{}' \mathbf{B}^{-1} (\mathbf{f} + \mathbf{g} z)$.

---

[14] 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 1. Estimation results for the CAC 40 Index

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
p1(d.f. for (4)-(9))				9.7319	10.2452	10.2712	9.8402	10.3367	10.3913	0.9423	0.9640	0.9603	0.9428	0.9660	0.9590
				(7.92)	*(7.88)*	*(7.79)*	*(7.65)*	*(7.63)*	*(7.50)*	*(37.82)*	*(54.62)*	*(45.63)*	*(37.02)*	*(62.74)*	*(44.14)*
$\mu1$(Sk..for (7)-(9))							-0.0330	-0.0245	-0.0292	0	0	0	0.0046	0.0039	0.0037
							(-1.22)	*(-0.90)*	*(-1.07)*				*(1.16)*	*(1.10)*	*(1.17)*
$\omega1$	1.2E-3	3.0E-4	9.9E-4	6.8E-4	1.1E-4	7.5E-4	6.8E-4	1.2E-4	7.4E-4	4.3E-4	-1.2E-5	5.6E-4	4.2E-4	-2.6E-5	5.4E-4
	(6.30)	*(1.36)*	*(5.46)*	*(2.99)*	*(0.38)*	*(3.63)*	*(2.99)*	*(0.42)*	*(3.63)*	*(2.56)*	*(-0.05)*	*(3.48)*	*(2.55)*	*(-0.11)*	*(3.41)*
$\alpha1$	0.0653	0.0458	0.0120	0.0583	0.0522	0.0140	0.0587	0.0523	0.0143	0.0488	0.0470	0.0116	0.0487	0.0470	0.0115
	(8.34)	*(9.22)*	*(2.48)*	*(6.45)*	*(6.07)*	*(1.68)*	*(6.50)*	*(6.08)*	*(1.71)*	*(6.66)*	*(6.35)*	*(1.51)*	*(6.68)*	*(6.33)*	*(1.48)*
$\lambda1$		0.1268	0.0706		0.1191	0.0743		0.1178	0.0737		0.1195	0.0662		0.1210	0.0661
		(7.11)	*(7.91)*		*(4.94)*	*(5.53)*		*(4.88)*	*(5.51)*		*(5.05)*	*(5.68)*		*(5.09)*	*(5.69)*
$\beta1$	0.9116	0.9324	0.9320	0.9286	0.9307	0.9338	0.9283	0.9306	0.9337	0.9343	0.9328	0.9371	0.9347	0.9328	0.9377
	(84.92)	*(113.46)*	*(111.88)*	*(81.57)*	*(87.17)*	*(91.19)*	*(81.83)*	*(87.22)*	*(91.35)*	*(96.81)*	*(97.82)*	*(101.39)*	*(98.41)*	*(97.99)*	*(102.03)*
p2										0.0577	0.0360	0.0397	0.0572	0.0340	0.0410
$\mu2$										0	0	0	-0.0752	-0.1115	-0.0869
$\omega2$										0.0289	0.0398	0.0410	0.0244	0.0238	0.0362
										(1.46)	*(0.51)*	*(0.69)*	*(1.55)*	*(0.37)*	*(0.77)*
$\alpha2$										0.7699	0.7734	0.3160	0.8945	1.3708	0.4438
										(0.81)	*(0.53)*	*(0.38)*	*(0.84)*	*(0.68)*	*(0.36)*
$\lambda2$											0.0262	0.6255		-0.0126	0.5815
											(0.19)	*(0.43)*		*(-0.17)*	*(0.43)*
$\beta2$										0.5865	0.5868	0.5961	0.5705	0.5518	0.5736
										(1.90)	*(0.92)*	*(1.08)*	*(1.90)*	*(1.15)*	*(1.13)*
Unconditional σ	22.34%	21.83%	21.83%	22.75%	22.27%	22.27%	22.75%	22.26%	22.24%	23.01%	22.36%	22.20%	23.29%	23.10%	22.30%
Unconditional $\sigma1$										21.42%	21.14%	20.97%	21.65%	21.71%	21.03%
Unconditional $\sigma2$										41.05%	43.73%	42.23%	41.22%	46.56%	41.30%
Unconditional τ							-0.0480	-0.0344	-0.0407	0	0	0	-0.1275	-0.1604	-0.1242
Unconditional *k*	0.6908	0.5428	0.7521	2.4794	2.3639	4.0811	-1.0446	-1.3430	-1.5027	3.6878	2.5241	1.9386	4.8295	6.1933	2.0312
Loglikelihood	355.4	374.4	375.2	396.9	414.5	414.1	397.7	414.9	414.7	396.9	415.3	415.1	398.3	416.7	416.4
Moment tests 1%	2	1	1	0	0	0	0	0	0	1	1	2	1	0	0
Density	1.5008	1.5799	1.3769	0.7360	1.1287	1.1680	0.6706	1.1383	0.9778	0.9175	0.7898	0.6122	0.8579	0.9032	0.7717
ACF	0.2260	0.3422	0.2353	0.1598	0.0654	0.2161	7.5529	17.5556	37.1285	0.0794	0.2051	0.2672	0.0677	0.1061	0.2334

Table 2. Estimation results for the DAX 30 Index

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
p1(d.f. for (4)-(9))				7.1878 *(10.85)*	7.5348 *(10.92)*	7.5359 *(10.83)*	7.2108 *(10.69)*	7.5438 *(10.77)*	7.5476 *(10.67)*	0.9578 *(86.41)*	0.9647 *(91.62)*	0.9581 *(81.68)*	0.9551 *(83.91)*	0.9629 *(90.89)*	0.9558 *(80.85)*
μ1(Sk..for (7)-(9))							-0.0596 *(-2.35)*	-0.0483 *(-1.89)*	-0.0507 *(-1.99)*	0	0	0	0.0034 *(1.58)*	0.0029 *(1.36)*	0.0034 *(1.53)*
ω1	1.0E-3 *(8.04)*	4.9E-4 *(3.43)*	1.0E-3 *(7.41)*	2.9E-4 *(2.53)*	5.5E-5 *(0.31)*	4.1E-4 *(3.27)*	2.9E-4 *(2.54)*	7.1E-5 *(0.41)*	4.0E-4 *(3.26)*	8.2E-5 *(1.30)*	-6.8E-5 *(-0.55)*	1.6E-4 *(2.21)*	7.8E-5 *(1.26)*	-7.0E-5 *(-0.58)*	1.5E-4 *(2.16)*
α1	0.0844 *(10.13)*	0.0630 *(11.67)*	0.0355 *(6.00)*	0.0734 *(7.32)*	0.0734 *(7.18)*	0.0398 *(3.72)*	0.0726 *(7.35)*	0.0726 *(7.17)*	0.0399 *(3.73)*	0.0542 *(8.23)*	0.0577 *(7.90)*	0.0301 *(3.70)*	0.0534 *(8.22)*	0.0568 *(7.93)*	0.0296 *(3.69)*
λ1		0.0839 *(9.38)*	0.0735 *(7.74)*		0.0723 *(4.41)*	0.0727 *(4.66)*		0.0701 *(4.27)*	0.0696 *(4.53)*		0.0698 *(4.31)*	0.0506 *(4.44)*		0.0694 *(4.27)*	0.0496 *(4.40)*
β1	0.8980 *(91.15)*	0.9196 *(128.18)*	0.9090 *(107.41)*	0.9245 *(94.20)*	0.9217 *(93.03)*	0.9192 *(91.69)*	0.9254 *(95.87)*	0.9227 *(94.09)*	0.9207 *(93.22)*	0.9361 *(127.27)*	0.9307 *(117.53)*	0.9331 *(121.73)*	0.9367 *(128.68)*	0.9317 *(119.57)*	0.9341 *(123.82)*
p2										0.0422	0.0353	0.0419	0.0449	0.0371	0.0442
μ2										0	0	0	-0.0727	-0.0744	-0.0735
ω2										0.0554 *(0.67)*	0.0621 *(0.48)*	0.0627 *(0.56)*	0.0516 *(0.77)*	0.0600 *(0.55)*	0.0598 *(0.64)*
α2										0.7095 *(0.67)*	0.4451 *(0.52)*	0.1364 *(0.31)*	0.6915 *(0.80)*	0.4253 *(0.58)*	0.0982 *(0.25)*
λ2											0.1595 *(0.43)*	0.8868 *(0.49)*		0.1432 *(0.40)*	0.8641 *(0.58)*
β2										0.5984 *(1.04)*	0.5871 *(0.73)*	0.5722 *(0.78)*	0.5999 *(1.22)*	0.5940 *(0.85)*	0.5764 *(0.91)*
Unconditional σ	24.03%	23.11%	23.09%	37.12%	29.92%	29.75%	37.26%	30.01%	29.54%	24.87%	24.12%	23.60%	24.82%	24.11%	23.54%
Unconditional σ1										23.17%	22.69%	22.00%	23.06%	22.64%	21.90%
Unconditional σ2										49.72%	49.05%	47.12%	48.52%	47.97%	45.89%
Unconditional τ							-0.1153	-0.0894	-0.0937	0	0	0	-0.1179	-0.1070	-0.1228
Unconditional *k*	2.0706	1.1340	1.9110	*NA*	*NA*	*NA*	*NA*	*NA*	11.4562	7.6984	3.8355	4.2003	7.0208	3.6691	3.6944
Loglikelihood	373.6	387.5	389.3	478.6	490.1	490.3	481.1	491.7	492.0	490.7	502.0	502.6	492.0	503.0	504.0
Moment tests 1%	2	2	3	2	1	1	2	1	2	4	3	4	3	3	4
Density	2.4236	2.3173	2.0420	0.9949	0.9186	1.3234	0.9354	0.7601	0.7884	0.8168	1.2321	0.9142	0.6848	1.0770	0.6858
ACF	0.1887	0.3718	0.2512	11.8904	3.1054	10.1459	19.8213	3.7574	1.8564	0.1660	0.1180	0.1073	0.1374	0.1201	0.1304

Table 3. Estimation results for the FTSE 100 Index

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
p1(d.f. for (4)-(9))				11.4353	11.5482	11.7353	11.5386	11.5999	11.8102	0.8820	0.9127	0.9207	0.8142	0.9135	0.9226
				(5.59)	*(5.75)*	*(5.84)*	*(5.63)*	*(5.75)*	*(5.84)*	*(16.24)*	*(25.35)*	*(23.88)*	*(11.34)*	*(25.10)*	*(23.41)*
µ1(Sk..for (7)-(9))							-0.0375	-0.0189	-0.0211	0	0	0	0.0063	0.0019	0.0026
							(-1.40)	*(-0.71)*	*(-0.80)*				*(2.01)*	*(0.88)*	*(1.22)*
ω1	3.1E-4	-4.4E-5	2.9E-4	2.6E-4	-1.5E-4	2.8E-4	2.6E-4	-1.4E-4	2.8E-4	1.1E-4	-3.9E-4	2.0E-4	6.0E-5	-3.8E-4	1.9E-4
	(3.96)	*(-0.44)*	*(4.79)*	*(2.92)*	*(-1.04)*	*(3.80)*	*(2.90)*	*(-0.99)*	*(3.77)*	*(1.46)*	*(-2.25)*	*(3.11)*	*(0.90)*	*(-2.21)*	*(3.03)*
α1	0.0738	0.0537	0.0151	0.0703	0.0545	0.0115	0.0700	0.0544	0.0119	0.0533	0.0468	0.0051	0.0443	0.0466	0.0054
	(10.19)	*(8.87)*	*(2.55)*	*(7.38)*	*(6.46)*	*(1.53)*	*(7.45)*	*(6.47)*	*(1.58)*	*(5.61)*	*(5.74)*	*(0.77)*	*(4.79)*	*(5.73)*	*(0.81)*
λ1		0.0876	0.0779		0.0980	0.0852		0.0969	0.0842		0.1224	0.0812		0.1212	0.0797
		(6.79)	*(8.22)*		*(5.54)*	*(6.63)*		*(5.49)*	*(6.60)*		*(5.57)*	*(6.65)*		*(5.51)*	*(6.66)*
β1	0.9171	0.9335	0.9363	0.9224	0.9327	0.9370	0.9227	0.9329	0.9371	0.9330	0.9328	0.9397	0.9423	0.9332	0.9405
	(112.77)	*(138.57)*	*(139.99)*	*(90.08)*	*(100.46)*	*(108.25)*	*(91.32)*	*(100.85)*	*(108.59)*	*(87.12)*	*(93.91)*	*(105.37)*	*(89.22)*	*(93.91)*	*(106.62)*
p2										0.1180	0.0873	0.0793	0.1858	0.0865	0.0774
µ2										0	0	0	-0.0275	-0.0202	-0.0312
ω2										0.0074	0.0042	0.0060	0.0037	0.0035	0.0058
										(1.35)	*(0.86)*	*(1.26)*	*(1.67)*	*(0.74)*	*(1.25)*
α2										0.4591	0.3713	0.4638	0.3234	0.3695	0.5275
										(2.36)	*(2.29)*	*(1.93)*	*(3.14)*	*(2.49)*	*(2.03)*
λ2											-0.0267	-0.1438		-0.0354	-0.2101
											(-0.41)	*(-0.34)*		*(-0.54)*	*(-0.47)*
β2										0.6789	0.7903	0.7654	0.7633	0.7993	0.7565
										(4.68)	*(7.72)*	*(7.31)*	*(9.51)*	*(8.75)*	*(7.19)*
Unconditional σ	18.53%	16.97%	17.40%	18.91%	17.15%	17.73%	18.90%	17.14%	17.68%	17.95%	17.00%	17.19%	17.99%	17.07%	17.27%
Unconditional σ1										16.52%	15.73%	16.00%	16.09%	15.78%	16.09%
Unconditional σ2										26.31%	26.91%	27.39%	24.48%	27.10%	27.52%
Unconditional *τ*							-0.0483	-0.0244	-0.0269	0	0	0	-0.0903	-0.0513	-0.0706
Unconditional *k*	4.5022	1.3501	3.5253	8.5581	3.0208	22.5555	0.7667	-1.3738	-1.5777	6.1528	3.1377	2.8508	6.3090	3.3577	3.0345
Loglikelihood	1192.0	1213.9	1217.2	1212.9	1235.8	1239.1	1213.9	1236.0	1239.4	1214.6	1242.5	1244.7	1216.4	1242.9	1245.4
Moment tests 1%	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Density	1.5140	1.2135	0.9711	0.6399	0.8276	0.8717	0.6906	0.8302	0.6942	0.7519	0.7869	0.7800	0.6732	0.7569	0.5355
ACF	0.7160	0.1117	0.3850	1.4849	0.1577	1.4929	0.1840	25.1860	78.2128	0.2608	0.1364	0.1232	0.3295	0.1183	0.1187

Table 4. Estimation results for the NIKKEI 225 Index

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
p1(d.f. for (4)-(9))				7.0281 *(9.15)*	7.7837 *(9.03)*	7.5387 *(9.24)*	7.0218 *(9.14)*	7.8057 *(9.01)*	7.5417 *(9.23)*	0.9294 *(38.61)*	0.9556 *(52.26)*	0.9538 *(50.41)*	0.9408 *(43.36)*	0.9576 *(53.82)*	0.9570 *(53.18)*
μ1(Sk..for (7)-(9))							-0.0056 *(-0.22)*	0.0111 *(0.43)*	0.0020 *(0.08)*	0	0	0	-0.0044 *(-1.42)*	0.0028 *(1.00)*	-0.0043 *(-1.52)*
ω1	1.9E-3 *(7.49)*	-1.6E-4 *(-0.53)*	1.5E-3 *(7.05)*	1.1E-3 *(3.29)*	-6.6E-4 *(-1.53)*	8.9E-4 *(3.38)*	1.1E-3 *(3.28)*	-6.7E-4 *(-1.55)*	9.0E-4 *(3.37)*	6.8E-4 *(2.95)*	-8.0E-4 *(-2.23)*	6.1E-4 *(3.20)*	7.4E-4 *(3.10)*	-7.7E-4 *(-2.15)*	6.3E-4 *(3.27)*
α1	0.0807 *(10.62)*	0.0681 *(9.96)*	0.0264 *(3.96)*	0.0733 *(6.75)*	0.0590 *(6.39)*	0.0215 *(2.56)*	0.0734 *(6.75)*	0.0590 *(6.39)*	0.0215 *(2.55)*	0.0571 *(6.99)*	0.0482 *(6.85)*	0.0179 *(2.77)*	0.0597 *(7.21)*	0.0490 *(6.87)*	0.0183 *(2.82)*
λ1		0.1587 *(10.34)*	0.0981 *(8.27)*		0.1730 *(6.43)*	0.0965 *(5.89)*		0.1738 *(6.44)*	0.0966 *(5.89)*		0.1857 *(6.77)*	0.0839 *(6.57)*		0.1839 *(6.84)*	0.0844 *(6.57)*
β1	0.8887 *(89.33)*	0.9066 *(111.89)*	0.9026 *(94.21)*	0.9100 *(68.25)*	0.9228 *(82.86)*	0.9184 *(77.70)*	0.9101 *(68.26)*	0.9228 *(82.74)*	0.9184 *(77.65)*	0.9154 *(79.62)*	0.9253 *(94.36)*	0.9195 *(89.65)*	0.9126 *(78.60)*	0.9239 *(92.58)*	0.9187 *(89.14)*
p2										0.0706	0.0444	0.0462	0.0592	0.0424	0.0430
μ2										0	0	0	0.0695	-0.0635	0.0960
ω2										0.0346 *(1.19)*	0.0329 *(0.63)*	0.0432 *(0.72)*	0.0468 *(1.09)*	0.0317 *(0.50)*	0.0518 *(0.73)*
α2										0.4924 *(1.26)*	0.5178 *(0.76)*	0.3315 *(0.59)*	0.5665 *(1.10)*	0.5716 *(0.68)*	0.3388 *(0.52)*
λ2											0.0930 *(0.56)*	0.2382 *(0.26)*		0.1240 *(0.75)*	0.3338 *(0.30)*
β2										0.7102 *(3.59)*	0.7423 *(2.49)*	0.7245 *(2.20)*	0.6603 *(2.59)*	0.7207 *(2.03)*	0.6836 *(1.80)*
Unconditional σ	24.99%	24.80%	25.96%	25.85%	24.66%	27.52%	25.87%	24.66%	27.52%	25.06%	24.32%	25.92%	25.19%	24.46%	26.03%
Unconditional σ1										22.44%	22.28%	23.98%	22.75%	22.39%	24.11%
Unconditional σ2										47.55%	51.38%	51.63%	49.36%	51.71%	52.17%
Unconditional *τ*							-0.0114	0.0200	0.0037	0	0	0	0.1495	0.2010	0.1525
Unconditional *k*	0.8242	1.2460	2.2529	4.3800	4.6234	*NA*	0.3059	-0.4480	0.3453	3.8474	3.7111	3.3675	3.7088	3.5502	3.2459
Loglikelihood	120.0	158.5	146.2	189.7	221.0	213.2	189.7	221.1	213.2	191.3	227.2	217.8	192.5	229.1	219.3
Moment tests 1%	2	2	3	0	0	0	0	0	0	0	0	0	1	0	1
Density	1.9647	1.5412	1.7282	0.8046	0.9122	1.0593	0.8718	0.8086	1.0423	0.7429	0.7305	0.9732	0.7339	0.7198	0.7777
ACF	0.1451	0.2707	0.6610	1.4882	1.2539	4.9502	0.1357	0.2768	0.2251	0.1837	0.1210	0.3343	0.1684	0.0975	0.3114

Table 5. Estimation results for the S&P 500 Index

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
$p1$(d.f. for (4)-(9))				6.1960 *(9.47)*	6.7604 *(9.24)*	6.9690 *(9.32)*	6.2970 *(9.15)*	6.9093 *(8.87)*	7.1501 *(8.84)*	0.8783 *(33.40)*	0.9213 *(37.26)*	0.9448 *(51.31)*	0.8899 *(36.05)*	0.9293 *(42.25)*	0.9430 *(51.02)*
$\mu 1$(Sk..for (7)-(9))							-0.0439 *(-1.73)*	-0.0488 *(-1.88)*	-0.0524 *(-1.99)*	0	0	0	0.0069 *(3.16)*	0.0061 *(3.29)*	0.0057 *(3.31)*
$\omega 1$	1.3E-4 *(4.64)*	-2.9E-4 *(-2.31)*	2.2E-4 *(7.35)*	6.8E-5 *(1.77)*	-2.3E-4 *(-1.59)*	1.7E-4 *(3.63)*	7.6E-5 *(1.91)*	-2.3E-4 *(-1.55)*	1.8E-4 *(3.73)*	1.0E-5 *(0.45)*	-1.8E-4 *(-1.74)*	1.3E-4 *(3.42)*	1.6E-5 *(0.69)*	-1.8E-4 *(-1.74)*	1.3E-4 *(3.37)*
$\alpha 1$	0.0471 *(10.40)*	0.0475 *(10.10)*	0.0018 *(0.35)*	0.0474 *(6.64)*	0.0535 *(6.40)*	0.0079 *(0.91)*	0.0486 *(6.70)*	0.0544 *(6.47)*	0.0082 *(0.95)*	0.0288 *(6.21)*	0.0433 *(6.76)*	0.0063 *(0.92)*	0.0304 *(6.28)*	0.0459 *(6.99)*	0.0069 *(1.01)*
$\lambda 1$		0.1119 *(7.18)*	0.0852 *(10.77)*		0.0937 *(4.81)*	0.0893 *(6.44)*		0.0933 *(4.84)*	0.0907 *(6.55)*		0.0887 *(4.87)*	0.0829 *(7.40)*		0.0876 *(5.10)*	0.0837 *(7.31)*
$\beta 1$	0.9495 *(201.78)*	0.9416 *(194.11)*	0.9472 *(219.31)*	0.9522 *(139.55)*	0.9391 *(114.36)*	0.9421 *(118.16)*	0.9507 *(135.04)*	0.9378 *(112.28)*	0.9406 *(115.93)*	0.9607 *(163.43)*	0.9381 *(116.94)*	0.9375 *(120.65)*	0.9587 *(156.71)*	0.9349 *(113.75)*	0.9359 *(116.98)*
$p2$										0.1217	0.0787	0.0552	0.1101	0.0707	0.0570
$\mu 2$										0	0	0	-0.0556	-0.0804	-0.0942
$\omega 2$										0.0080 *(1.71)*	-0.0032 *(-0.50)*	0.0317 *(0.91)*	0.0084 *(1.78)*	-0.0039 *(-0.63)*	0.0240 *(0.83)*
$\alpha 2$										0.5245 *(2.71)*	0.0820 *(1.45)*	-0.0383 *(-0.09)*	0.6001 *(2.56)*	0.0855 *(1.45)*	0.0684 *(0.14)*
$\lambda 2$											0.2395 *(1.09)*	1.3872 *(1.02)*		0.2504 *(1.20)*	1.0860 *(0.90)*
$\beta 2$										0.7271 *(6.24)*	0.9600 *(39.36)*	0.5113 *(1.03)*	0.7026 *(5.78)*	0.9575 *(40.98)*	0.5377 *(1.05)*
Unconditional σ	19.25%	16.63%	16.10%	41.21%	17.93%	17.77%	32.43%	17.72%	17.47%	17.56%	16.52%	16.41%	17.64%	16.31%	16.13%
Unconditional $\sigma 1$										15.11%	14.72%	15.03%	15.26%	14.64%	14.77%
Unconditional $\sigma 2$										29.78%	30.49%	31.77%	30.15%	29.60%	29.38%
Unconditional τ							-0.1010	-0.0997	-0.1026	0		0	-0.2297	-0.2683	-0.2588
Unconditional k	5.3756	1.5033	3.8951	*NA*	9.2975	*NA*	*NA*	1.1879	3.5948	8.7107	3.8027	3.3077	9.7287	3.5860	2.6070
Loglikelihood	1292.3	1317.7	1324.3	1375.5	1391.4	1396.8	1377.0	1393.2	1398.9	1376.2	1391.7	1401.1	1381.9	1397.5	1406.5
Moment tests 1%	1	1	3	0	0	0	1	0	1	0	0	1	1	0	2
Density	2.8043	1.8920	1.3301	3.0903	1.1384	1.9504	2.3198	1.1729	1.2590	0.9592	0.8875	1.2573	0.9444	1.0892	0.7582
ACF	2.2051	0.1166	0.7790	29.8164	1.8548	7.8304	29.7108	0.1499	1.1739	0.2005	0.0568	0.0667	0.2246	0.0632	0.0697

Fig. 1. The Returns on the S&P Index

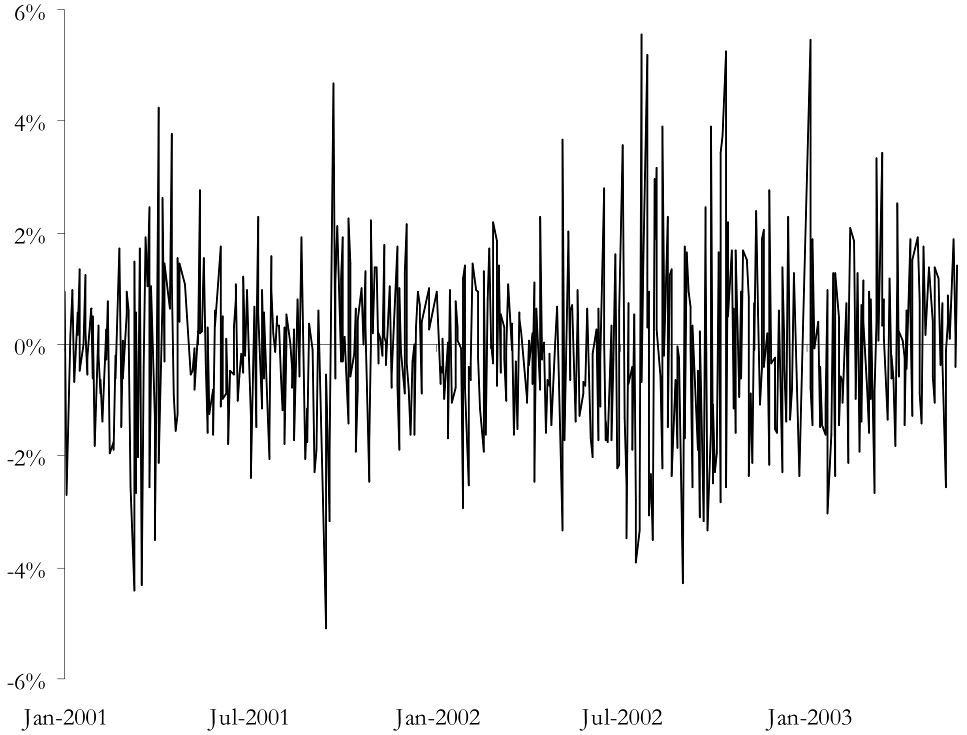


Fig. 2. The Conditional Volatilities and the Time-varying Probability of the First State

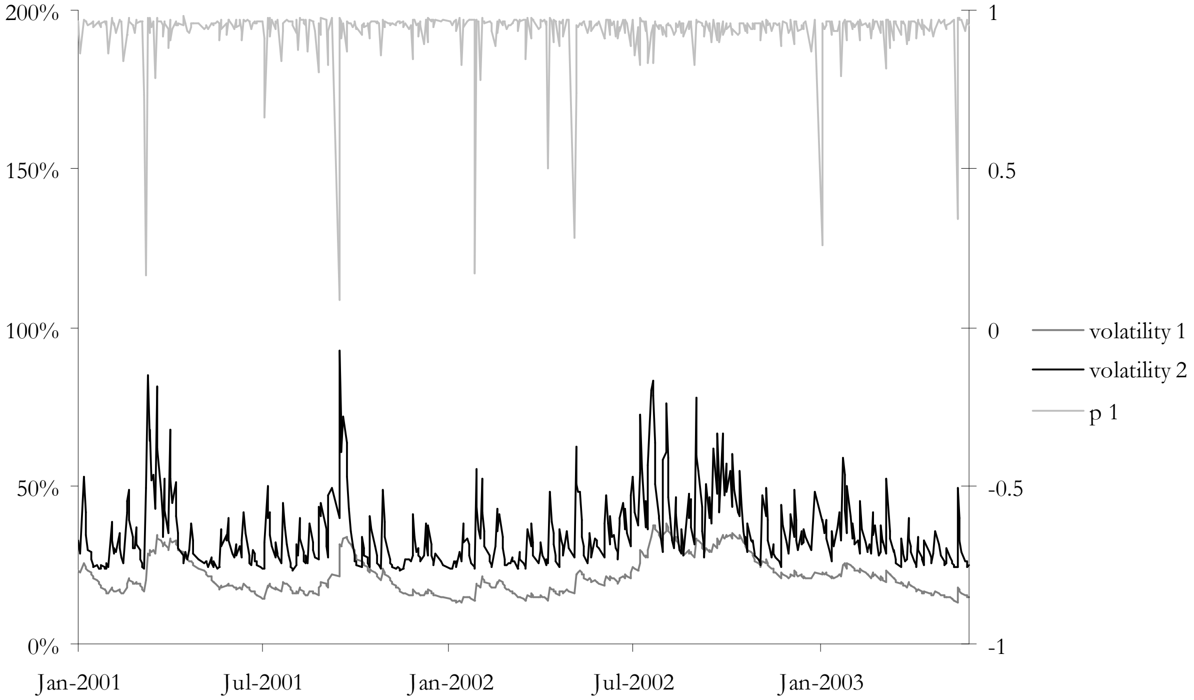
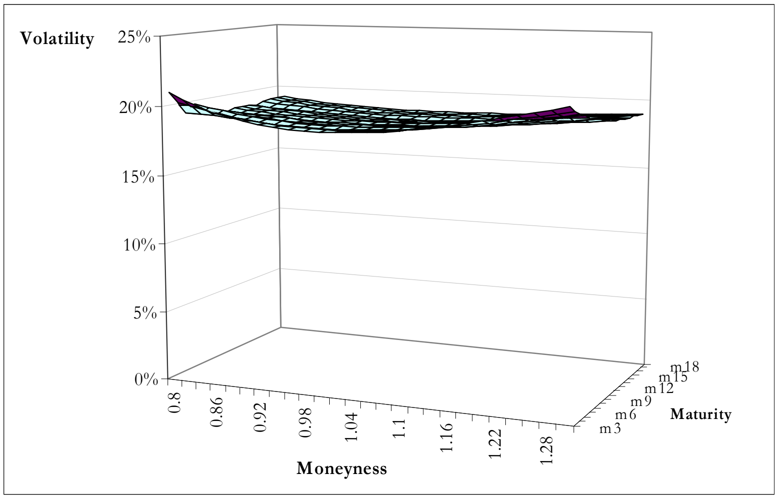
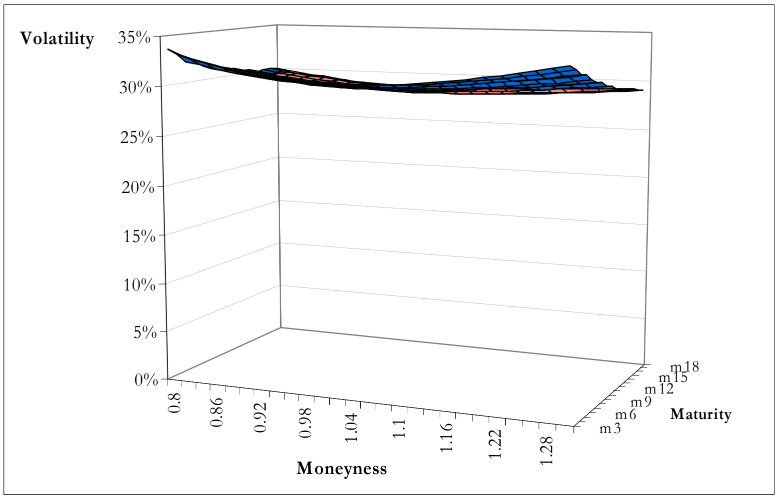


Fig. 3. Simulated Equity Index Skews

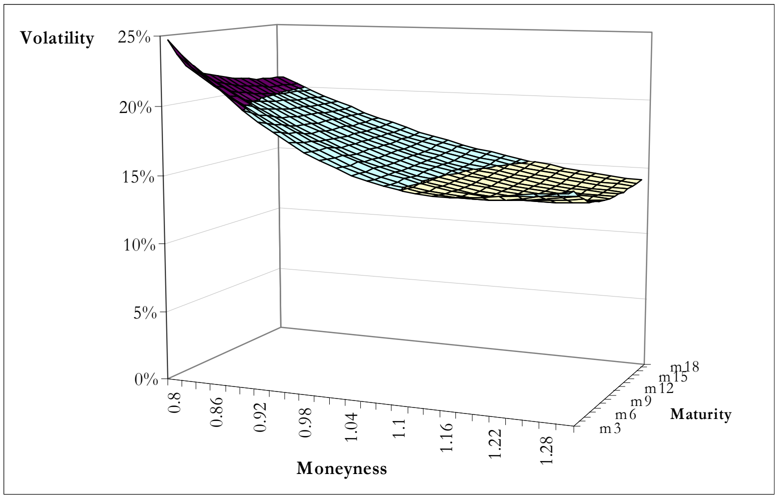
(a) Normal GARCH(1,1)



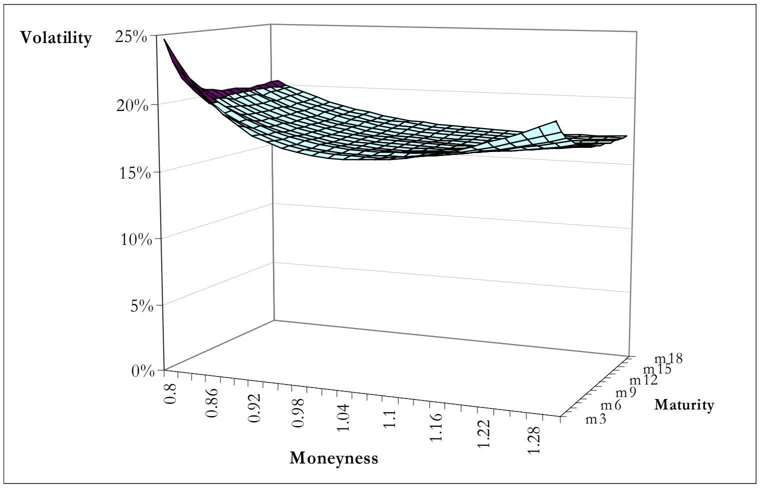
(b) Asymmetric *t*-GARCH(1,1)



(c) Asymmetric *t*-GJR



(d) NM(2)-GARCH(1,1)



(e) NM(2)-GJR