MV Kulikova and DR Taylor*

A conditionally heteroskedastic time series model for certain South African stock price returns

ABSTRACT

The distributional properties of returns data have important implications for financial models and are of particular importance in risk-scenario simulation, volatility prediction and in the event of financial crisis. We present simple time-series models that capture the heteroskedasticity of financial time series and incorporate the effect of using heavy-tailed distributions. These models allow for time-varying volatility, which is an important extension of the conventional methodology. The models are an augmentation of the GARCH class of models, but allow for conditionally normal inverse Gaussian and variance gamma distributed errors. As in previous studies, this new approach permits a distinction between conditional heteroskedasticity and a conditionally leptokurtic distribution, but, compared with the well-known GARCH-t model, it allows us to capture the asymmetric behaviour observed in actual returns series. The practical applicability of the models is confirmed by implementing a fitting procedure to a carefully chosen set of South African stock price returns.

1. INTRODUCTION

Any number of studies has shown that most financial series are heteroskedastic, i.e. they exhibit changes in volatility, or variance, over time. The approach based on autoregressive conditional heteroskedasticity (ARCH) introduced by Engle (1982), and later generalized to GARCH by Bollerslev (1986), was the first attempt to take into account these changes in volatility over time. In this class of models the effect of varying volatility is captured by allowing the conditional variance of the series to be a function of past variances and of the square of previous observations.

Another interesting result of the ARCH/GARCH approach is that the *conditional* error distribution is normal, but the *unconditional* error distribution of the ARCH/GARCH model is leptokurtic. Bollerslev (1987:542) remarked that:

"It is not clear whether the GARCH(p,q) model with conditionally normal errors sufficiently accounts for the observed leptokurtosis in financial time series."

As an alternative solution, the GARCH(p,q) model with conditionally t-distributed errors was proposed in Bollerslev (1987). The results of this investigation revealed that the standardized t-distribution with constant variance fails to take account of temporal dependence in returns series, known as the volatility clustering effect. Besides this, the ARCH/GARCH models with conditionally normal errors do not seem to fully capture the leptokurtosis. Although the proposed GARCH(p,q) model with conditionally t-distributed errors was shown to be superior to the Gaussian GARCH approach, it was noted that:

"It remains an open question whether other conditional error distributions provide an even better description."

(Bollerslev, 1987:546)

Today, given the recent financial crisis, we believe that such an analysis is critical. In this paper we extend the GARCH(p,q) model to allow for conditional errors that are variance gamma (VG) or normal inverse Gaussian (NIG) distributed. As in Bollerslev (1987), this new development permits a distinction between conditional heteroskedasticity and a conditionally leptokurtic distribution, either of which could account for the observed unconditional kurtosis in the data.

Additionally, these new models allow us to capture the asymmetric behaviour observed in actual returns series, i.e. the observed unconditional skewness in the data. The considered distributions arise as either subclasses or limiting cases of the generalized hyperbolic (GH) distribution, first introduced by Barndorff-Nielsen (1977). These are a flexible, four-parameter class of distribution functions that can describe a wide range of shapes.

The variance gamma, normal inverse Gaussian and t-distribution models are frequently employed in the finance industry (Daal & Madan (2005); Madan & Seneta (1990); Madan & Milne (1991); Madan, Carr & Chang (1998); and Seneta (2004)). *However, they have been implemented with constant variance.* After careful consideration of the results of Bollerslev (1987), it is expected that these models will consequently fail to account for the volatility clustering effect (more about "stylized" facts can be found in Cont (2001); Ghysels, Harvey & Renault (2005); Harvey & Jaeger (1993); Nelson (1990))

As an alternative solution, we propose to use the GARCH approach with conditionally VG and NIG

*Respectively from CEMAT, Instituto Superior Tecnico, TU Lisbon, Portugal and Programme in Advanced Mathematics of Finance, University of the Witwatersrand - Johannesburg, Republic of South Africa. The authors wish to thank Prof. F. Lombard of the University of Johannesburg for many helpful discussions regarding the contents of this paper. The first author gratefully acknowledges the support of *Fundação para a Ciência e a Tecnologia (FCT)*, co-financed by the European community fund FEDER, under grant No. SFRH/BPD/64397/2009.
Email: David.Taylor@wits.ac.za

distributed errors. To determine their validity, the models are fitted to a set of financial time series from the South African financial market. The fitting procedure is based on the method of maximum likelihood and allows for possible dependence in the returns series.

2. THE CLASS OF GARCH-TYPE MODELS

Let $\{\varepsilon_t\}$ denote a real-valued, discrete-time stochastic process, and ψ_t the information set (σ-field) of all information through time t. Following the celebrated paper by Engle (1982) we consider $\{\varepsilon_t\}$ of the form

$$\varepsilon_t = \sigma_t z_t \quad \dots (1)$$

$$z_t \ \text{i.i.d.}\ E[z_t] = 0,\ Var[z_t] = 1 \quad \dots (2)$$

With σ_t a time-varying, positive, and measurable function of the time $t-1$ information set, ψ_{t-1}.

We call $\{\varepsilon_t\}$ a GARCH(p,q) process if

$$\mathbf{E}\ [\varepsilon_t^2|\psi_{t-1}] = \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2,$$
$$\alpha_0 > 0,\ \alpha_i \geq 0, \beta_j \geq 0. \quad \dots (3)$$

From the definition given above, $\{\varepsilon_t\}$ is serially uncorrelated with zero mean, but the *conditional* variance of $\{\varepsilon_t\}$ equals σ_t^2, which is time-varying. In most applications of the model, $\{\varepsilon_t\}$ corresponds to the innovation in the mean for some other stochastic process, say $\{y_t\}$. When $p = 0$, the process (3) reduces to autoregressive conditional heteroskedasticity of order q, i.e. ARCH(q). As can be seen from (3), in the ARCH(q) process the conditional variance is specified as a linear function of past sample variances only, whereas the GARCH(p,q) process allows the inclusion of conditional variances.

Finally, if $f(z_t)$ denotes the density function of z_t, then the sample Log Likelihood Function (Log LF) for $y_T, y_{T-1} \dots y_1$ is given by the formula Bollerslev (1986):

$$\mathcal{L}(\theta; y_T, \dots, y_1) = \sum_{t=1}^{T}[\ln f(\varepsilon_t\sigma_t^{-1}) - \ln\sigma_t] \quad \dots (4)$$

where θ is an unknown parameter vector, which needs to be estimated.

Among all GARCH-type models, the GARCH$(1,1)$ model is extensively used in financial time series modelling. It provides a simple representation of the main dynamic characteristics of the returns series of a wide range of assets. It is also worth noting here that the GARCH$(1,1)$ model has proved to have a better forecasting ability when compared to traditional ARCH models; (see for example, Akgiray (1989)).

3. MODELING THE RETURN GENERATING PROCESS AS A GARCH PROCESS

3.1 The GARCH$(\mathbf{p},\mathbf{q})$ model with conditionally t-distributed errors

The GARCH-t model was first proposed by Bollerslev (1987) to describe speculative prices and their rates of return. It was shown to be very effective and out-performed the classic Gaussian GARCH(p,q) approach.

Denote by $\{y_t\}$ the de-meaned returns series, i.e. $y_t = \ln\frac{S_t}{S_{t-1}}$ where S_t is the closing price on day t. Let the conditional distribution of $\{y_t\}$ be standardized-t, with mean μ, variance σ_t^2 and degrees of freedom ν, i.e.

$$y_t = E\ [y_t|\psi_{t-1}] + \varepsilon_t = \mu + \varepsilon_t,\ \varepsilon_t = \sigma_t z_t, \quad \dots (5)$$

$$z_t \sim \mathbf{std}.f_t(x;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \nu > 2 \quad \dots (6)$$

where $\Gamma(.)$ denotes the Gamma function. In addition, assume a GARCH(p,q) model (3) for the conditional variance σ_t^2.

It is well known that, for $\frac{1}{\nu} \to 0$, the t-distribution converges to a normal distribution, but, for $\frac{1}{\nu} > 0$ it has "fatter tails" than the corresponding normal distribution. The fourth moment of the distribution only exists for $\nu > 4$. The skewness of a t-distributed random variable X is 0 when $\nu > 3$.

In this study we consider non-symmetric distributions. This new development permits a distinction between conditional heteroskedasticity and a conditionally skewed distribution, which could account for the unconditional skewness in observed returns data.

3.2. The GARCH(p,q) model with conditionally normal inverse Gaussian distributed errors.

The normal inverse Gaussian distribution (NIG) is the most extensively used distribution function in financial time series modelling. Having heavier tail dependence than the normal distribution, it is considered appropriate for modelling data sets with many extreme observations Karlis (2002).

This class of distribution functions can be considered as a mixture of the normal and the inverse Gaussian distributions. The NIG distribution is also a subclass of the GH distribution corresponding to the GH parameter value of $\lambda = -1/2$. The GARCH-NIG model can be written as follows

$$y_t = E\,[y_t|\psi_{t-1}] + \varepsilon_t = \mu + \varepsilon_t,\ \varepsilon_t = \sigma_t z_t \qquad \dots (7)$$

$$z_t \sim \mathbf{std.}\, f_{NIG}(x; a, b, \delta, m) = \frac{a}{\pi} e^{\delta\sqrt{a^2-b^2} - bm} \phi^{-\frac{1}{2}}(x) K_1\left(a\delta\phi^{\frac{1}{2}}(x)\right) e^{bx} \qquad \dots (8)$$

where $\phi(.)$ denotes $\phi(x) = 1 + \left[\frac{x-m}{\delta}\right]^2$ and $K_\lambda(.)$ is the modified Bessel function of the third kind. In addition, assume a GARCH(p, q) model (3) for the conditional variance σ_t^2.

As can be seen, the probability density function depends on 4 parameters. The distribution parameters δ and m are scaling and location parameters. The parameters a and b determine the distribution shape, since $a \pm b$ define the heaviness of the tails. If $b = 0$, then the distribution is symmetric, and the sign of b determines the kind of skewness.

The mean, variance, skewness and kurtosis of a NIG distributed random variable X are given by,

$$\mathbf{E}\,[X] = m + \frac{b\delta}{\gamma},\ \mathbf{Var}\,[X] = \frac{\delta a^2}{\gamma^3},$$

where $\gamma = \sqrt{a^2 - b^2}$ … (9)

$$\text{Skewness} = \frac{3b}{a(\gamma\delta)^{1/2}},\ \text{Kurtosis} = 3\frac{1+4b^2/a^2}{\delta\gamma} \qquad \dots (10)$$

The parameter domain is restricted to $a > |b|$, $a > 0$ and $\delta > 0$.

3.3. The GARCH(p, q) model with conditionally Variance Gamma distributed errors.

Let the conditional distribution of $\{y_t\}$ be standardized variance gamma, with mean μ and variance σ_t^2, i.e.

$$y_t = E\,[y_t|\psi_{t-1}] + \varepsilon_t = \mu + \varepsilon_t,\ \varepsilon_t = \sigma_t z_t \qquad \dots (11)$$

$$z_t \sim \mathbf{std.}\, f_{VG}(x; \lambda, a, b, m) = C|x-m|^{\lambda-0.5} K_{\lambda-0.5}(a|x-m|) e^{b(x-m)} \qquad \dots (12)$$

where $C = \frac{\gamma^{2\lambda}}{\sqrt{\pi}\,\Gamma(\lambda)(2a)^{\lambda-0.5}}$, $\gamma^2 = a^2 - b^2$. … (13)

In addition, assume a GARCH(p, q) model (3) for the σ_t^2.

In Seneta (2004), the estimation of the parameters of the VG distribution via a moment matching method is presented. The first four moments, i.e. the mean, variance, skewness and kurtosis, of a VG distributed random variable X are given by,

$$\mathbf{E}\,[X] = m + 2\frac{b\lambda}{\gamma^2},\ \mathbf{Var}\,[X] = \frac{2\lambda}{\gamma^2}\left(1 + 2\frac{b^2}{\gamma^2}\right), \qquad \dots (14)$$

$$\text{Skewness} = \frac{2\theta^3 v^2 + 3\sigma^2\theta v}{(\theta^2 v + \sigma^2)^{3/2}},\quad \theta = \frac{2b\lambda}{\gamma}, \sigma^2 = \frac{2\lambda}{\gamma}, v = \frac{1}{\lambda}, \qquad \dots (15)$$

$$\text{Kurtosis} = 3 + \frac{3\sigma^4 v + 12\sigma^2\theta^2 v^2 + 6\theta^4 v^3}{(\theta^2 v + \sigma^2)^2}. \qquad \dots (16)$$

The parameter domain is restricted to $a > |b|$, $a > 0$ and $\lambda > 0$.

4. EMPIRICAL RESULTS

4.1 JSE data

In this section the method of maximum likelihood is used to estimate a GARCH model with conditionally NIG, VG and t-distributed errors. We consider the Top 40 index from the JSE and seven of its most representative components. The estimation of the GARCH model is based on daily closing prices of the JSE observed from 27 December 2005 to 27 January 2010, with a total of 1066 observations. Table 1 provides the details of the companies under study.

To validate our choice of model empirically, we first present some of the statistics of the returns data and summarize them in Table 2. The unconditional moments of the returns series are presented together with autocorrelations of the squared returns. Analyzing the obtained results, we reach two important conclusions.

Firstly, we can see that each actual returns series $\{y_t\}$ exhibits some skewness, and that all eight samples show very high kurtosis when compared with the normal distribution. This means that the actual returns distributions have heavier tails and much higher peaks than the Gaussian distribution. We can detect this graphically in Figures 1 and 2, where histograms of the observed returns series (with the best-fitting normal distribution imposed for visual reference) and QQ plots are presented. The Gaussian GARCH model is expected to be strongly rejected in favour of one of the GARCH-VG, GARCH-NIG or GARCH-t models. Additionally, the unconditional skewness for some samples is significant (see for instance BIL and MTN). Consequently, the previously proposed GARCH-t model is expected to be inappropriate for these series.

Table 1: Company profiles for seven representative components of the Top 40 index from the JSE Securities Exchange

JSE Code	Company	Sector	Market Cap
AGL	Anglo American Plc	Mining (Metals & Minerals)	GBP 22,37 bn
BIL	BHP Billiton Plc	Mining (Metals & Minerals)	GBP 74,38 bn
FSR	FirstRand Ltd	Banking	ZAR 75,64 bn
SBK	Standard Bank Group Ltd	Banking	ZAR 160,94 bn
MTN	MTN Group Ltd	Wireless Telecoms Services	ZAR 224,63 bn
NTC	Network Healthcare Holdings	Hospital Management & Long Term Care	ZAR 13,64 bn
OML	Old Mutual Plc	Life Assurance	GBP 4,21 bn

Table 2: Distributions and dynamics of the returns series for the Top 40 index of JSE and seven of its components

Series	Unconditional Moments				Autocorrelations				
	Mean	St. Dev	Skewness	Kurtosis	1st	2nd	5th	10th	20th
Top 40	0,0003	0,0175	-0,0992	2,4999	0,1698	0,2863	0,3202	0,2842	0,1244
AGL	0,0000	0,0314	-0,1406	3,4657	0,1285	0,2778	0,2480	0,3031	0,1707
BIL	0,0000	0,0286	0,2914	3,4888	0,2043	0,3093	0,2910	0,3180	0,1530
FSR	0,0000	0,0244	-0,0818	1,4552	0,1313	0,1422	0,1409	0,1058	0,0693
SBK	0,0000	0,0233	0,0785	2,0152	0,1358	0,1518	0,0810	0,0888	0,0857
MTN	0,0000	0,0269	0,3650	2,8858	0,2312	0,2730	0,0922	0,0319	0,0713
NTC	0,0000	0,0214	0,0720	1,8507	0,1450	0,1723	0,0576	0,0508	0,0320
OML	0,0000	0,0287	-0,1360	4,5465	0,3596	0,2759	0,2531	0,2124	0,1853

Our second conclusion concerns the dynamic of the daily returns and, as a consequence, the properties of a corresponding GARCH$(1,1)$ model. It can be seen from the second part of Table 2 that the autocorrelation function of the squared returns decays slowly. This indicates a relatively slow change in conditional variance and has often been observed in reality, i.e. the GARCH$(1,1)$ estimation of actual stock price and index returns usually yields $\alpha_1 + \beta_1$ very close to 1. As we will see later, this corresponds to our empirical results.

4.2 Estimation procedure

To estimate a GARCH$(1,1)$ model with VG, NIG and t-distributed innovations, we produced our own codes. All methods were implemented in MatLab. We used the build-in functions “besselk” and “gamma” to compute the Bessel and Gamma functions. The MatLab Optimization Toolbox and Symbolic Toolbox were used for optimization purposes and to satisfy condition (2).

Before discussing our empirical results, we need to mention that although the VG and NIG distributions are rich classes of distribution functions, estimation of their parameters is not easy due to the complicated quantities involved. Since the derivatives of the Log LF involve the Bessel function, direct maximization by using gradient-based optimization methods is a difficult task (Karlis (2002)). As a consequence, special numerical methods need to be used. One of the possible solutions is the development of Expectation Maximization (EM) type algorithms, which can overcome these numerical difficulties. These are now widely used in the financial industry; see Dempster (1977) for more details. However, as far as we know, the derivation of EM type methods for the considered GARCH-type models is still an open question.

In our numerical investigation we used the MatLab function “fminsearch” that is based on the Nelder-Mead simplex algorithm, a direct search method that uses function values but not its derivatives. Thus, we avoided the numerical problems arising from the LF gradient evaluation. However, in spite of these advantages, this method can be very slow to converge and may not even converge (Higham & Higham (2005)). We used two criteria for terminating the optimization method. These were based on the absolute changes of the Log LF and changes in the parameter values, θ. The algorithm was terminated if both criteria were satisfied. If we denote by $\mathcal{L}^{(k)}$ the Log LF after k iterations and $\theta^{(k)}$ as the corresponding parameter values, then the criteria used in the optimization method are $|\mathcal{L}^{(k+1)} - \mathcal{L}^{(k)}| < 10^{-7}$ and $||\theta^{(k+1)} - \theta^{(k)}|| < 10^{-5}$.

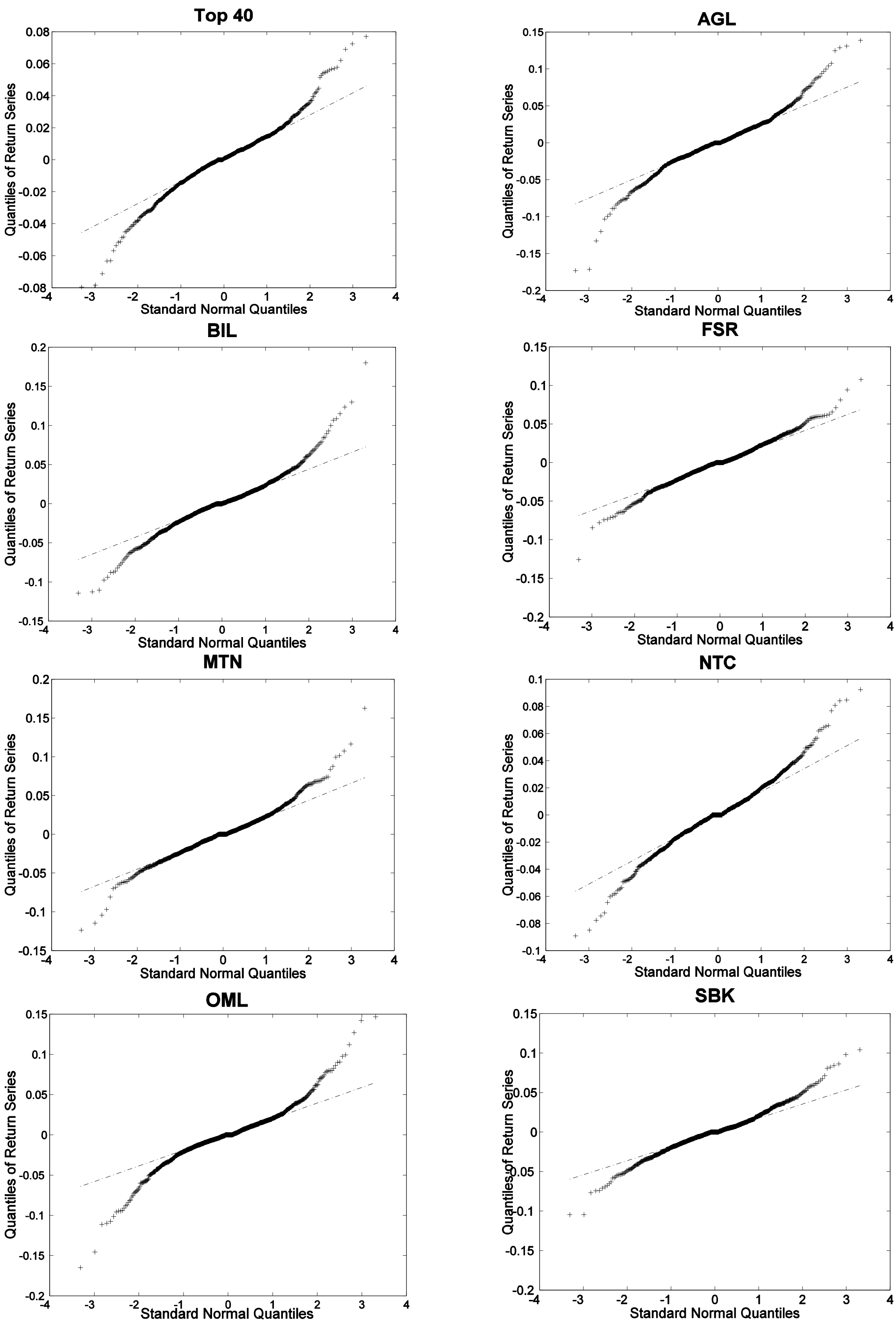

Figure1: QQ Plots of the observed return series for eight samples examined.

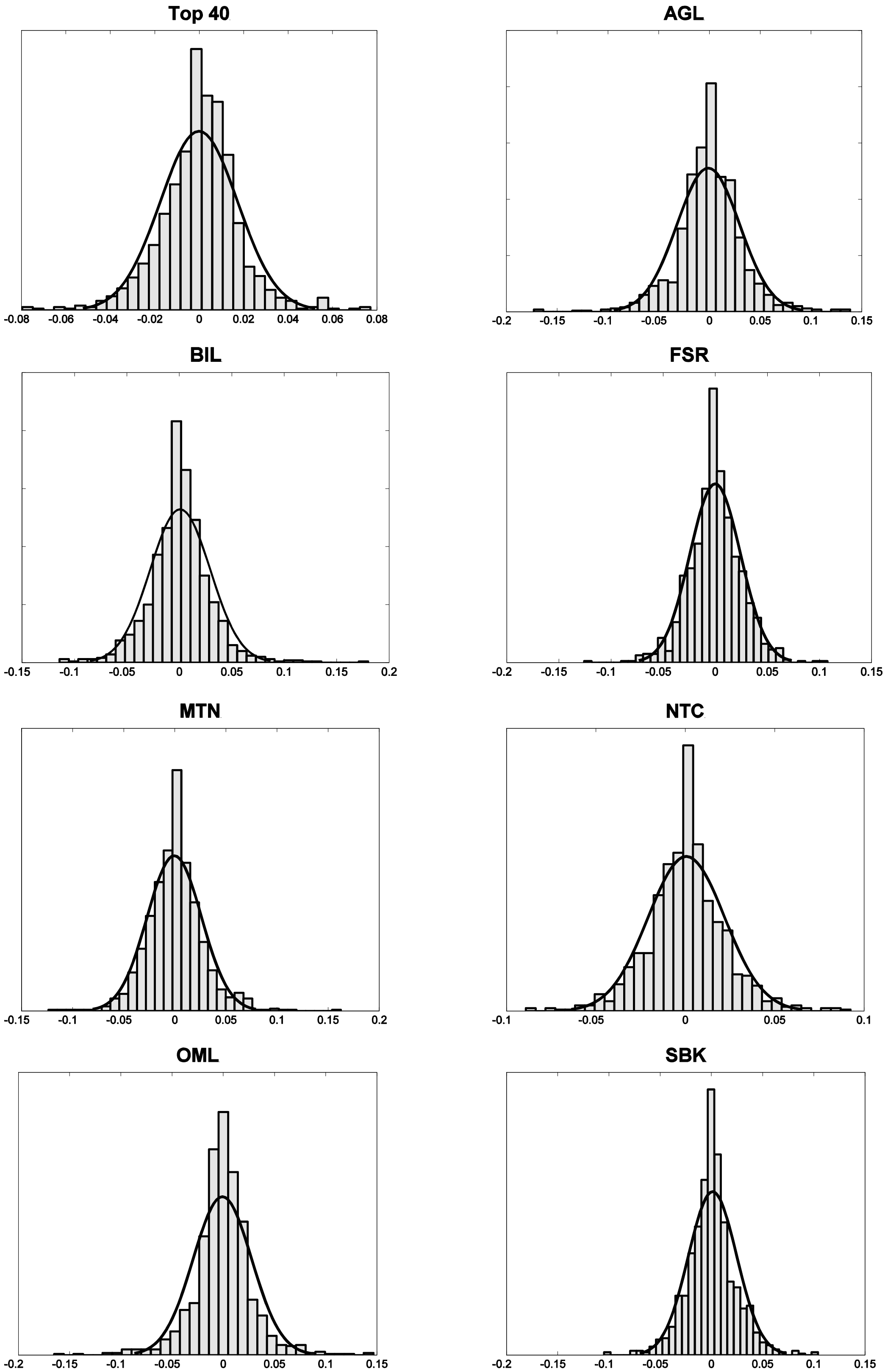

Figure 2: Histograms of the observed return series for eight samples examined, with the best-fitting normal distribution imposed for visual reference.

We also needed to be concerned with the fact that a local maximum might be obtained. We followed the standard procedure to mitigate this whereby the algorithm was initiated at several starting points to ensure (as far as is possible) that the obtained maximum is global (see for example Karlis (2002)). For each of the eight samples we examined we used five trials, each starting from different initial values. We found that, on average, 3 out of 5 trials led to the same final estimates.

4.3 Empirical results

The results of the fitting procedure for the GARCH$(1,1)$ model with conditionally t, NIG and VG distributed innovations are summarized in Tables 3, 4 and 5, respectively. As anticipated, the GARCH$(1,1)$ estimation yields $\hat{\alpha}_1 + \hat{\beta}_1$ very close to 1 for all returns series.

Having carefully analyzed our resulting estimates, we conclude that the hypothesis of a GARCH$(1,1)$ model with conditionally normally distributed errors can be strongly rejected for all eight samples. This can readily be seen from Table 6, where the results of the likelihood ratio test are presented. For the Top 40 index return series, the maximum log LF value under the hypothesis of a Gaussian GARCH$(1,1)$ model is $2963{,}45$. This is less than the corresponding maximum log LF value of the best model among our proposed alternative GARCH models. The GARCH$(1,1)$ model with conditionally NIG distributed errors has a log LF value of $2972{,}51$; see the first row of Table 6. The number of additional system parameters in a GARCH-NIG model is 2, when compared with the Gaussian GARCH approach. As a consequence, the likelihood ratio test statistic $LR = -2(2963{,}45 - 2972{,}51) = 18{,}12$ should be compared with the $\chi^2_{2,0.99}$ value. Since $18{,}12$ is highly significant when compared with $\chi^2_{2,0.99} = 9{,}21$, we conclude that a GARCH$(1,1)$ model with conditionally normally distributed errors is too restrictive when compared with the GARCH-NIG model. The results of the likelihood ratio test for all samples are summarized in Table 6.

Table 3: MLE of the GARCH$(1,1)$ model with conditionally t-distributed errors (3), (5), (6)

Series	Estimated parameters				Log LF
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$1/\hat{v}$	max $\mathcal{L}$
Top 40	$0{,}3 \cdot 10^{-5}$	0,0906	0,8849	0,0593	2966,45
AGL	$0{,}5 \cdot 10^{-5}$	0,0590	0,9191	0,1044	2361,25
BIL	$0{,}4 \cdot 10^{-5}$	0,0551	0,9254	0,0911	2443,21
FSR	$0{,}9 \cdot 10^{-5}$	0,0793	0,8874	0,0858	2534,40
SBK	$0{,}8 \cdot 10^{-5}$	0,0687	0,8837	0,1559	2595,38
MTN	$0{,}7 \cdot 10^{-5}$	0,0473	0,9232	0,1264	2443,67
NTC	$0{,}9 \cdot 10^{-5}$	0,0520	0,8904	0,1789	2653,38
OML	$0{,}3 \cdot 10^{-5}$	0,0575	0,9170	0,1369	2526,58

Table 4: MLE of the GARCH$(1,1)$ model with conditionally NIG distributed errors (3), (7), (8)

Series	Estimated parameters							Log LF
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{a}$	$\hat{b}$	$\hat{\delta}$	$\hat{m}$	max $\mathcal{L}$
Top 40	$0{,}3 \cdot 10^{-5}$	0,0976	0,8889	3,1816	-0,9992	2,7227	0,9006	2972,51
AGL	$0{,}7 \cdot 10^{-5}$	0,0680	0,9210	1,9308	-0,1173	1,9201	0,1168	2360,15
BIL	$0{,}5 \cdot 10^{-5}$	0,0619	0,9276	2,0908	-0,1763	2,0686	0,1751	2442,85
FSR	$0{,}9 \cdot 10^{-5}$	0,0848	0,8987	2,5465	0,3286	2,4831	-0,3232	2533,23
SBK	$0{,}1 \cdot 10^{-4}$	0,0978	0,8866	1,3064	0,0895	1,3035	-0,0913	2599,27
MTN	$0{,}8 \cdot 10^{-5}$	0,0579	0,9304	1,5992	0,2442	1,5435	-0,2385	2448,22
NTC	$0{,}1 \cdot 10^{-4}$	0,0717	0,9006	1,1958	0,0072	1,1957	-0,0072	2655,30
OML	$0{,}4 \cdot 10^{-5}$	0,0693	0,9205	1,5652	-0,0982	1,5560	0,0978	2525,84

Table 5: MLE of the GARCH$(1,1)$ model with conditionally VG distributed errors (3), (11), (12)

Series	Estimated parameters							Log LF
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{a}$	$\hat{b}$	$\hat{\lambda}$	$\hat{m}$	max $\mathcal{L}$
Top 40	$0{,}3 \cdot 10^{-5}$	0,1042	0,8817	5,4261	-1,3718	12,1247	1,2070	2972,27
AGL	$0{,}7 \cdot 10^{-5}$	0,0683	0,9204	2,8797	-0,0994	4,1315	0,0991	2359,76
BIL	$0{,}1 \cdot 10^{-5}$	0,0627	0,9262	2,7249	-0,1177	3,6920	0,1173	2448,48
FSR	$0{,}1 \cdot 10^{-5}$	0,1437	0,8352	2,7332	0,1317	3,7094	-0,1311	2532,33
SBK	$0{,}1 \cdot 10^{-5}$	0,0982	0,8867	1,7705	0,0242	1,5665	-0,0242	2602,73
MTN	$0{,}7 \cdot 10^{-5}$	0,0582	0,9302	2,4169	0,2151	2,8522	-0,2117	2448,06
NTC	$0{,}5 \cdot 10^{-5}$	0,0032	0,9852	0,8697	0,0227	0,3774	-0,0226	2654,89
OML	$0{,}5 \cdot 10^{-5}$	0,0753	0,9132	2,3788	-0,0710	2,8218	0,0709	2525,88

Table 6: Likelihood ratio (LR) test statistics

Series	Maximum Log LF value, $\mathcal{L}(\theta^*)$			LR test statistics				
	GARCH-N	Best proposed model		LR	k	$\chi^2_{k,0,99}$	p-value	
Top 40	2963,45	GARCH-NIG	2972,51	18,12	$k=2$	9,21	0,00	< 0,01
AGL	2349,46	GARCH-t	2361,25	23,58	$k=1$	6,63	0,00	< 0,01
BIL	2435,18	GARCH-VG	2448,48	26,60	$k=2$	9,21	0,00	< 0,01
FSR	2524,62	GARCH-t	2534,40	19,56	$k=1$	6,63	0,00	< 0,01
SBK	2581,26	GARCH-VG	2602,73	42,94	$k=2$	9,21	0,00	< 0,01
MTN	2430,43	GARCH-NIG	2448,22	35,57	$k=2$	9,21	0,00	< 0,01
NTC	2630,78	GARCH-NIG	2655,30	49,04	$k=2$	9,21	0,00	< 0,01
OML	2508,98	GARCH-t	2526,58	35,19	$k=1$	6,63	0,00	< 0,01

It is interesting to note that Bollerslev's GARCH-t model; see Bollerslev (1987) provides the best results only for those series with the highest unconditional kurtosis and almost zero skewness, i.e. AGL, OML and FSR; see the results in Tables 2 and 6. At the same time, for those series with significant unconditional skewness (BIL and MTN, Table 2), the GARCH-t model is too restrictive when compared with either the GARCH-NIG or GARCH-VG models. This can be seen in Table 7, where the test statistics favour the GARCH-NIG or GARCH-VG models over the GARCH-t model. As expected, for the returns data with significant skewness, the GARCH$(1,1)$ model with conditionally t-distributed errors does not seem to fully capture the skewness. Only in the case of the NTC return series is there no significant difference between the GARCH-t model and the GARCH-VG, GARCH-NIG models. This series has relatively low skewness, as can be seen in Table 2.

Now, consider the results presented in Table 8. Comparing the implied estimate of the conditional kurtosis for the proposed alternative GARCH models with the sample analogue for $\hat{\varepsilon}_t^4\hat{\sigma}_t^{-4}$, we conclude that, in most cases, our GARCH-VG and GARCH-NIG models provide the closest accordance between these values. For example, the MTN return series has an implied estimate of the conditional fourth moment from the GARCH-t model (Bollerslev (1987)) of: $3(\hat{v}-2)(\hat{v}-4)^{-1} = 4{,}534$, and a sample analogue of $5{,}153$. The best alternative model, i.e. GARCH-NIG model, gives an implied estimate of $4{,}344$. This is in close accordance with the sample analogue of $4{,}085$; see the fourth row of Table 8. A possible reason for the failure of the GARCH-t model in this regard is that the MTN return series (see Table 2) exhibits the highest unconditional skewness of all the data.

Now, let us consider the results presented at the second part of Table 8. Comparing the implied estimate of the conditional skewness for the proposed alternative GARCH models with the sample analogue for $\hat{\varepsilon}_t^3\hat{\sigma}_t^{-3}$, we conclude that, in most cases, our GARCH-VG and GARCH-NIG models provide the closest accordance between these values. For example, the MTN return series has an implied estimate of the conditional skewness from the best alternative model, i.e. GARCH-NIG, of 0.293. This is in close accordance with the sample analogue of $0{,}365$ and differs significantly from the t-value of 0.

Table 7: Likelihood Ratio (LR) test statistics

Series	Maximum Log LF value, $\mathcal{L}(\theta^*)$			LR test statistics			
	GARCH-t	Best examined model		LR	$\chi^2_{1,0,99}$	p-value	
Top 40	2966,45	GARCH-NIG	2972,51	12,12	6,63	0,00	< 0,01
BIL	2443,21	GARCH-VG	2448,48	10,54	6,63	0,00	< 0,01
SBK	2595,38	GARCH-VG	2602,73	14,69	6,63	0,00	< 0,01
MTN	2443,67	GARCH-NIG	2448,22	9,09	6,63	0,00	< 0,01
NTC	2653,38	GARCH-NIG	2655,30	3,84	6,63	0,02	> 0,01

Table 8: Implied estimates of the conditional fourth moment and conditional skewness and their sample analogues for $\hat{\varepsilon}_t^4\hat{\sigma}_t^{-4}$ and $\hat{\varepsilon}_t^3\hat{\sigma}_t^{-3}$, respectively

Stock Price	Conditional fourth moment					Conditional skewness		
	GARCH-t		Best alternative			Best alternative		
Series	Implied	Sample	Model	Implied	Sample	Model	Implied	Sample
Top 40	3,432	3,612	GARCH-NIG	3,508	3,568	GARCH-NIG	-0,328	-0,326
BIL	3,860	3,870	GARCH-VG	3,829	3,869	GARCH-VG	-0,122	-0,169
SBK	5,485	5,761	GARCH-VG	4,917	4,694	GARCH-VG	0,158	0,188
MTN	4,534	5,153	GARCH-NIG	4,344	4,085	GARCH-NIG	0,293	0,365
NTC	5,590	6,771	GARCH-NIG	5,098	4,190	GARCH-NIG	0,015	0,011

Our final remark addresses the comparison of our proposed alternative models with the Gaussian GARCH$(1,1)$ model. Having compared the results summarized in Table 8, we can conclude that the GARCH$(1,1)$ model with conditionally normally distributed errors does not capture the leptokurtosis and skewness observed in returns series data. The implied estimates of the conditional fourth and third moments obtained from the GARCH-t, GARCH-VG and GARCH-NIG models differ significantly from the normal Gaussian of three and zero. We note that the same conclusion holds for all eight samples examined.

5. CONCLUSION

This paper examined two new GARCH-type models with conditionally variance gamma or normally inverse Gaussian distributed errors. The new developments allow for modelling the volatility clustering effect and permits conditionally leptokurtic and skewed distributions, which accounts for the observed unconditional kurtosis and skewness in the data.

The models were fitted to a set of financial time series from the South African financial market. The empirical results showed that a GARCH model with conditionally normal errors is strongly rejected in favour of the GARCH-t, GARCH-NIG and GARCH-VG models. The previously proposed GARCH-t model does not fully capture the skewness observed in actual returns series.

Our alternative GARCH models with conditionally NIG and VG distributed errors overcome these difficulties. In order to accommodate returns series that exhibit the volatility clustering effect, these alternative models should be implemented in preference to the standard GARCH models. Since most financial time series exhibit this effect, it would make these models preferable in almost all circumstances.

REFERENCES

Akgiray V. 1989. Conditional heteroskedasticity in time series of stock returns: evidence and forecast. *Journal of Business*, 62:55-80.

Barndorff-Nielsen OE. 1977. Exponentially decreasing distributions for the logarithm of a particle size. *Proc. Roy. Soc. London Ser A*, 353:401-419.

Barndorff-Nielsen OE. and Prause K. 2001. Apparent scaling. *Finance and Stochastics*, 5:103-113.

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307-327.

Bollerslev T. 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Econometrics and Statistics*, 542-547.

Bollerslev T., Chou R.Y. and Kroner K.F. 1992. ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, 52:5-59.

Cont R. 2001. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223-236.

Daal E and Madan D. 2005. An empirical examination of the variance-gamma model for foreign currency options. *Journal of Business*, 78:79-105.

Dempster AP, Laird NM and Rubin D. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. B*, 39:1-38.

Engle RF. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50:987-1008.

Ghysels E, Harvey A and Renault E. 2005. *Stochastic volatility*. Montreal: Scientific Series.

Harvey AC and Jaeger A. 1993. Detrending stylized facts and the business cycle. *Journal of Applied Econometrics*, 8:231-247.

Higham DJ and Higham NJ. 2005. *MATLAB Guide*. Philadelphia: Society for Industrial and Applied Mathematics.

Karlis D. 2002. An EM type algorithm for maximum likelihood estimation of the normal-inverse Gaussian distribution. *Statistic & Probability Letters*, 57:43-52.

Madan DB. and Seneta E. 1990. The variance gamma model for share market returns. *Journal of Business*, 63:511-524.

Madan DB. and Milne F. 1991. Option pricing with VG martingale components. *Mathematical Finance*, 4:39-55.

Madan DB., Carr PP. and Chang EC. 1998. The variance gamma process and option pricing. *European Finance Review*, 2:79-105.

Nelson DB. 1990. Stationarity and persistence in the GARCH$(1,1)$ model. *Econom. Theory.*, 6:318-344.

Seneta E. 2004. Fitting the Variance-Gamma model to financial data. *Journal of Applied Probability*, 41:177-187.