Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Exact Maximum Likelihood estimation for the BL-GARCH model under elliptical distributed innovations

Abdou Kâ DIONGUE, Dominique GUEGAN, Rodney C. WOLFF

2008.27

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://ces.univ-paris1.fr/cesdp/CES-docs.htm
ISSN : 1955-611X

Exact Maximum Likelihood estimation for the BL-GARCH model under elliptical distributed innovations

Abdou Kâ Diongue *, Dominique Guégan †and Rodney C. Wolff ‡

Abstract

In this paper, we discuss the class of Bilinear GARCH (BL-GARCH) models which are capable of capturing simultaneously two key properties of non-linear time series: volatility clustering and leverage effects. It has been observed often that the marginal distributions of such time series have heavy tails; thus we examine the BL-GARCH model in a general setting under some non-Normal distributions. We investigate some probabilistic properties of this model and we propose and implement a maximum likelihood estimation (MLE) methodology. To evaluate the small-sample performance of this method for the various models, a Monte Carlo study is conducted. Finally, within-sample estimation properties are studied using S&P 500 daily returns, when the features of interest manifest as volatility clustering and leverage effects.

Keywords: BL-GARCH process; Elliptical distribution; Leverage effects; Maximum likelihood; Monte Carlo method; Volatility clustering.

*Université Gaston Berger, UFR SAT, BP 234, Saint-Louis, SENEGAL, e-mail: abdou.diongue@gmail.com

†Paris School of Economics, MSE - CES, Université Paris1 Panthéon-Sorbonne, 106 boulevard de l'hopital, 75013 Paris, FRANCE, dominique.guegan@univ-paris1.fr

‡School of Mathematical Sciences, Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, AUSTRALIA, r.wolff@qut.edu.au

1 Introduction

Non-constant variance in non-linear time series is a challenging modelling exercise, considered among many others things by Tong (1990). In particular, the stylized fact that the volatility of financial time series is non-constant has been long recognized in the literature. A popular and prominent tool used to describe this phenomenon is the autoregressive conditional heteroskedasticity (ARCH) model. The ARCH model, developed by Engle (1982) and later extended to the GARCH model by Bollerslev (1986), formulates the conditional variance of a random variable as a linear function of its past squared realizations. Following the seminal work of Engle (1982), a number of applied studies surfaced to illustrate the usefulness of these models in economic and financial settings: for example, see Gouriéroux (1997) and recently Giraitis *et al.* (2007). Moreover, a large volume of the studies focuses upon the familiar observation within financial time series concerning asymmetric variation, whereby stock price changes are negatively correlated with changes in volatility; that is, negative shocks impose larger volatility relative to positive shocks of the same magnitude (Black, 1976). In the light of this empirical finding, various models which allow for asymmetry in volatility have been proposed. These models include the exponential GARCH model (EGARCH) of Nelson (1991), the APARCH model, proposed by Ding *et al.* (1993), and the BL-GARCH model recently introduced by Storti and Vitale (2003a). The lattermost model complements the Gaussian bilinear (BL) model of means, introduced by Granger and Andersen (1978), developed by Subba Rao and Gabr (1984), Guégan (1987) and Guégan and Pham (1989).

In this paper, we focus on the BL-GARCH model whose appeal is that it can take into account explosions and related volatility features of non-linear time series. Bilinear models have been shown more generally to be capable of providing an arbitrarily close second-order approximation to a general class of underlying nonlinear process that may be reasonably expressed in terms of the Volterra series expansion, as in Guégan (1988). That is, under some modifications, the BL-GARCH model can generalize various GARCH-type models including the standard GARCH, and the APARCH models, among others.

Specifically, we study the BL-GARCH$(1,1)$ process which is mainly used in applications. We substantially extend the work of Storti and Vitale (2003a) using elliptical noise. Specifically, we provide the main probabilistic properties of the model under this class of elliptical innovations (positivity of the variance, strict stationarity, $2k$-order stationarity, and fourth-order moment). However, parameter estimation of the BL-GARCH model remains problematic. Indeed, to deal with the parameter estimation for the BL-GARCH model, Storti and Vitale (2003b) use an indirect maximum likelihood procedure based on the EM algorithm. Despite the fact that the EM algorithm has become a very popular computational method in statistics, this approach presents some limitations, including slow numerical convergence; the fact that the sequence of parameter vectors converges to a maximum likelihood estimator only if a judicious choice of the starting value is made; and the underlying assumption of Normality for the data set. These limitations, coupled with the non-existence of a measure of the standard errors for the estimates, serve to limit the method's applicability severely.

In addition, in its standard form, the BL-GARCH model assumes that the conditional distribution of assets' returns is Gaussian. However, for many financial time series, this specification does not adequately take into account leptokurtosis: see Baillie and Bollerslev (1991), McAleer (2005), Zivot and Wang (2005) and references therein for some examples. Thus, it is timely to investigate non-Normal alternative distributions.

The paper is organized as follows. In Section 2, the BL-GARCH model is presented with some important properties concerning conditions for the conditional variance to be finite, as well as for strict stationarity, existence of moments and ergodic solutions. Knowlegde of these properties are essential for carrying out consistent estimator, which we address in Section 3, where we present the MLE method under elliptical distributions, such as the Normal, Student-t and GED distribution functions. Section 4 calibrates the performance of the estimation procedure through Monte Carlo simulations. Section 5 presents the data and contains the main empirical findings along with the goodness-of-fit tests, while Section 6 provides concluding remarks. In the Appendix, we provide the score functions and the Hessian matrices for these models.

2 BL-GARCH model and its specification

The BL-GARCH model has been introduced and discussed by Storti and Vitale (2003a) in a Gaussian framework. We now specify some properties of this model in a more general context, in order to apply it to financial asset price returns whose distribution function is known to be far from Gaussian.

2.1 The BL-GARCH model

Let S_t denote an asset price at time t, $y_t = log\left(S_t/S_{t-1}\right)$, and $\mu_t = E\left(y_t \mid \Psi_{t-1}\right)$ the conditional mean given an increasing sequence of σ-fields Ψ_{t-1} generated by $\left(y_{t-1}, y_{t-2}, \cdots\right)$. We assume that the series of interest, $\left(y_t\right)_t$, follows the recursive scheme, for all t,

$$y_t = \mu_t + u_t, \tag{1a}$$

$$u_t = h_t \varepsilon_t, \tag{1b}$$

$$h_t^2 = a_0 + \sum_{i=1}^{p} a_i u_{t-i}^2 + \sum_{j=1}^{q} b_j h_{t-j}^2 + \sum_{k=1}^{r} c_k h_{t-k} u_{t-k}, \tag{1c}$$

where p, q, r are non-negative integers with $r = \min\left(p,q\right)$, h_t^2 is the conditional variance of the process $\left(u_t\right)_t$ given the σ-fields Ψ_{t-1}, and ε_t is a sequence of independent identically distributed $D\left(0,1\right)$ random variables with $D\left(.\right)$ an elliptical probability density function with mean 0 and unit variance. The model (1b)-(1c) is more general than the standard GARCH model of Bollerslev (1986) in the sense that it allows innovations of different signs to have a strong impact on volatility and allows larger shocks to have a larger influence on volatility than does the standard GARCH model: see Black (1976).

We now specify some properties of the model (1a)-(1c), restricting to the case $p = q = r = 1$ which is the more useful for applications, as in Baillie and DeGennaro (1990), or Hansen and Lunde (2005). First of all, we give the conditions for which the conditional variance h_t^2, defined as

$$h_t^2 = a_0 + a_1 u_{t-1}^2 + b_1 h_{t-1}^2 + c_1 h_{t-1} u_{t-1}, \tag{2}$$

is nonnegative. Indeed, it is important in practice for estimation theory (using quasi-maximum likelihood methods) that a model such as that in that (1a), (1b), and (2) does not generate negative conditional variance h_t^2 in sample, since the log quasi-likelihood involves a term in $\log\left(h_t^2\right)$, which explodes to $-\infty$ as h_t^2 approaches 0 and is ill-defined for $h_t^2 \leq 0$.

The second set of properties concerns the strict stationarity and the stationarity of the process (1a), (1b), and (2), and the expression of moments.

2.1.1 Positivity of the conditional variance

Write the relationship (2) as

$$h_t^2 = [1, u_{t-1}, h_{t-1}] \begin{bmatrix} a_0 & 0 & 0 \\ 0 & a_1 & \frac{c_1}{2} \\ 0 & \frac{c_1}{2} & b_1 \end{bmatrix} \begin{bmatrix} 1 \\ u_{t-1} \\ h_{t-1} \end{bmatrix}. \tag{3}$$

Then we get the following result.

Proposition 2.1 . *Let the process $(y_t)_t$ be defined by (1a), (1b), and (2). A sufficient set* *of conditions for positivity of the conditional variance h_t^2 is*

$$a_0 > 0; a_1 > 0;\ b_1 > 0;\ \text{and } c_1^2 < 4a_1b_1. \tag{4}$$

Proof From the relationship (3), we observe that a necessary and sufficient condition for the positivity of h_t^2 is given by the positive definiteness of the matrix A, where

$$A = \begin{bmatrix} a_0 & 0 & 0 \\ 0 & a_1 & \frac{c_1}{2} \\ 0 & \frac{c_1}{2} & b_1 \end{bmatrix}.$$

Now a matrix is positive definite if all its eigenvalues are strictly positive. The set of eigenvalues is equal to

$$\left\{ a_0, \frac{a_1 + b_1 + \sqrt{(a_1 - b_1)^2 + c_1^2}}{2}, \frac{a_1 + b_1 - \sqrt{(a_1 - b_1)^2 + c_1^2}}{2} \right\}.$$

Hence, a sufficient set of conditions for positiveness of the conditional variance h_t^2 defined by the relation (2) is given by the conditions (4). ■

2.1.2 Stationarity conditions

In order to give the conditions for strict stationarity of (1a), (1b), and (2), we rewrite (2) in a formal way using random coefficients. Then, we get

$$h_t^2 = g(\varepsilon_{t-1}) + c(\varepsilon_{t-1})\, h_{t-1}^2, \tag{5}$$

with $g(\varepsilon_{t-1}) = a_0$ and $c(\varepsilon_{t-1}) = b_1 + c_1\varepsilon_{t-1} + a_1\varepsilon_{t-1}^2$.

Theorem 2.1 *Let the process $(u_t)_t$ be defined by (1b) and (5). We assume that the condi-*

tions (4) hold, then the process $(u_t)_t$ *has a strictly stationary solution and*

$$h_t^2 = g(\varepsilon_{t-1}) + \sum_{k=0}^{\infty} \prod_{j=0}^{k} c(\varepsilon_{t-1-j})\, g(\varepsilon_{t-1-k})$$

if and only if

$$E[\log\{c(\varepsilon_t)\}] < 0. \tag{6}$$

Proof First of all, we remark that the random coefficients $c(\varepsilon_t)$ and $g(\varepsilon_t)$ are nonnegative, for all t, thanks to the conditions (4). Second, applying Fubini's and Tonelli's Theorems and using the fact that the $(\varepsilon_t)_t$ are independent and identically distributed, the largest Lyapunov exponent of the model (1b) and (2) is equal to

$$\begin{aligned}\inf_{n\in\mathbb{N}}\left\{E\left[\frac{1}{n+1}\log\left(c(\varepsilon_0)\,c(\varepsilon_{-1})\cdots c(\varepsilon_{-n})\right)\right]\right\} &= \inf_{n\in\mathbb{N}}\left\{\frac{1}{n+1}\sum_{i=0}^{n} E[\log(c(\varepsilon_{-i}))]\right\}\\ &= E[\log\{c(\varepsilon_t)\}].\end{aligned}$$

Third, the process $\left(h_t^2\right)_t$ can be rewritten as a random coefficient autoregressive process, given by

$$X_t = A_t X_{t-1} + B_t,$$

where $X_t = h_t^2$, $A_t = c(\varepsilon_t)$, and $B_t = g(\varepsilon_t)$. In addition, since the variance of the process $(\varepsilon_t)_t$ is finite, we get $E[\log\{c(\varepsilon_t)\}] < \infty$. Thus, following the main lines of Theorem 1.3 of Bougerol and Picard (1992), we get our Theorem 2.1. ■

Note that the condition (6) can be difficult to verify from a data set $y_1, \cdots, y_n$. Nevertheless, this condition can be verified a posteriori after the model (1a), (1b), and (2) has been fitted. Thus, it could be useful to have a more flexible condition to verify the stationarity of a real data set. Here, we focus on an m-order stationarity condition, for any $m \in \mathbb{N}$. We get the following theorem.

Theorem 2.2 *Let the process* $(u_t)_t$ *be defined by (1b), and (5). We assume that all the absolute moments up to* $2m$ *(*$m \in \mathbb{N}$*) of the process* $(\varepsilon_t)_t$ *exist, then a necessary and sufficient* *condition for the existence of the* 2*m-order stationary solution for the process* $(u_t)_t$ *is*

$$E\{c(\varepsilon_t)\}^m < 1.$$

Proof To get this result, we follow the main lines of Theorem 2.2 in Ling and McAleer (2002). Note that, in their theorem, Ling and McAleer use the exponent km instead of m which appears to be a typographical error, since the exponent appearing in their proof is m. ■

Besides the Normal distribution, we consider, in the following, two different alternative symmetric probability density functions for the innovations $(\varepsilon_t)_t$: the standard Student-t distribution with ν degrees of freedom, following Bollerslev (1987) and the Generalized Error Distribution (GED), as in Nelson (1991). Recall that a random variable X follows a GED if its probability density function is given by

$$f(x) = \frac{\nu 2^{-\left(1+\frac{1}{\nu}\right)}}{\lambda_\nu \Gamma\left(\frac{1}{\nu}\right)} e^{-\frac{1}{2}\left|\frac{x}{\lambda_\nu}\right|^\nu}, \ -\infty < x < \infty,$$

with $\lambda_\nu = \sqrt{2^{-2/\nu}\Gamma(1/\nu)/\Gamma(3/\nu)}$, $0 < \nu < \infty$ is the tail-thickness parameter and Γ is the Euler gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$. The GED includes the Gaussian distribution ($\nu = 2$) as a special case, along with many other distributions, some more fat-tailed than the Gaussian (e.g., the double exponential distribution corresponding to $\nu = 1$) and some more thin-tailed (e.g., the Uniform distribution on the interval $\left[-\sqrt{3}, \sqrt{3}\right]$ when $\nu \to \infty$).

In the next corollary, we give conditions for the existence of the second- and fourth-order moments of the process $(u_t)_t$ under the above distributions.

Corollary 2.1 *Let the process* $(u_t)_t$ *be defined by (1b) and (2). We assume that* $(\varepsilon_t)_t$ *follows a distribution function belonging to the class of elliptical distribution functions, denoted* $D(0,1)$.

1) The second order moment of the process $(u_t)_t$ *exists if and only if*

$$a_1 + b_1 < 1, \tag{7}$$

and is given by

$$E\left(u_t^2\right) = \frac{a_0}{1 - a_1 - b_1}.$$

2) The fourth order moment of the process $(u_t)_t$ *exists if and only if*

$$b_1^2 + c_1^2 + sa_1^2 + 2a_1b_1 < 1,$$

where $s = E\left(\varepsilon_t^4\right)$*, and it is equal to*

$$E\left(u_t^4\right) = \frac{sa_0^2\left(1 + a_1 + b_1\right)}{\left(1 - a_1 - b_1\right)\left(1 - b_1^2 - c_1^2 - sa_1^2 - 2a_1b_1\right)}.$$

a) If the process $(\varepsilon_t)_t$ *follows a standard Normal distribution (*$N(0,1)$*), then* $s = 3$*;*

b) If the process $(\varepsilon_t)_t$ *follows a standard Student-t distribution with* $\nu \geq 5$ *(to guarantee existence of the fourth-moment), then* $s = 3\nu^2 / \left[(\nu - 2)(\nu - 4)\right]$ *;*

c) If the process $(\varepsilon_t)_t$ *follows a standardized GED distribution with* $\nu > 0$*, then* $s = \lambda_\nu 2^{4/\nu}\Gamma(5/\nu)/\Gamma(1/\nu)$.

Proof 1) Assuming $m = 1$ in the previous theorem, we restrict the existence of a second order stationarity for the process $(u_t)_t$ defined by (1b), and (2) at the condition $a_1 + b_1 < 1$. We remark that this condition is similar to the one given by Bollerslev (1986) for a stationary solution of the GARCH$(1,1)$ process. Note that the stationarity condition is independent of c_1.

2) If we set $m = 2$ in Theorem 2.2, then we obtain

$$E\left(c(\varepsilon_t)^2\right) = b_1^2 + c_1^2 + sa_1^2 + 2a_1b_1,$$

and hence the condition for existence of the fourth order moment is obtained. In addition, with some straighforward integration, one can obtain the value of s for the different distribution functions we have considered. ■

We can remark that this model permits to model large kurtosis when compare to the standard GARCH model.

3 Maximum likelihood estimation

The estimation of conditional volatility models are typically performed by an MLE procedure, as in Bollerslev and Wooldridge (1992). Given a sample $y_1, \cdots, y_n$, the conditional likelihood function is equal to

$$\mathscr{L}(\omega) = \mathscr{L}(\omega \mid y_1, \cdots, y_n) = \prod_{t=1}^{n} g(y_t, \mu_t(\alpha), h_t(\omega)),$$

where $g(y_t, \mu_t(\alpha), h_t(\omega))$ denotes the conditional density function for the random variables y_t with mean μ_t and standard deviation h_t, and $\omega = (\alpha, \theta)$ is the parameter vector to be estimated. Here α corresponds to the set of parameters in the conditional mean assumed, in what follows, to be an ARMA(k,l) model and $\theta = (a_0, a_1, b_1, c_1)$ is the set of parameters for the BL-GARCH$(1,1)$. We assume that we work under the stationarity conditions given in the previous section. Thus, estimation proceeds by maximising $L(\omega) = \log(\mathscr{L}(\omega))$, where pre-sample values of h_t^2 are set to the unconditional sample variance.

Considering the three most typical elliptical normalized distributions that have been applied so far — the Normal, Student-t and GED distributions — we give below the expressions of the log-likelihood function, $L(\omega)$, while the score function, and the Hessian matrix for each distribution are provided in an Appendix.

3.1 Normal distribution

The Normal distribution is the most widely used when estimating GARCH models. If we assume that the innovations $(\varepsilon_t)_{t\in\mathbb{Z}}$ have a Gaussian distribution function then the conditional log-likelihood function associated with $u_t = y_t - \mu_t$ is given by

$$L(\omega) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left\{\log\left(h_t^2\right) + \frac{u_t^2}{h_t^2}\right\}, \tag{8}$$

where n is the sample size. To obtain an analytical or numerical solution for the MLE, we need to know the first-order derivative, see the Appendix for details, and to solve the equation $\partial L(\omega)/\partial\omega = 0$.

3.2 Student-t distribution

Now, if we assume that the innovations $(\varepsilon_t)_{t\in\mathbb{Z}}$ follow a Student-t distribution with ν degrees of freedom, then the MLE estimator $\hat{\omega}_n$ maximizes the log-likelihood function

$$\begin{aligned} L(\omega) = {} & n\left\{\log\Gamma\left(\frac{\nu+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log\pi(\nu-2)\right\} \\ & -\frac{1}{2}\sum_{t=1}^{n}\left[\log(h_t^2) + (\nu+1)\log\left\{1 + \frac{u_t^2}{h_t^2(\nu-2)}\right\}\right], \end{aligned} \tag{9}$$

where $2 < \nu \leq \infty$. When $\nu \to \infty$, we get the Normal distribution, so that the smaller the value of ν, the fatter the tails.

3.3 GED distribution

Knowing that skewness and kurtosis are important in financial applications, Nelson (1991) suggested to consider the family of GED distribution functions. The GED log-likelihood function is given by

$$L(\omega) = n\left\{\log\left(\frac{\nu}{\lambda_\nu}\right) - \left(1+\frac{1}{\nu}\right)\log(2) - \log\Gamma\left(\frac{1}{\nu}\right)\right\}$$
$$-\frac{1}{2}\sum_{t=1}^{n}\left\{\log\left(h_t^2\right) + h_t^{-\nu}\left|\frac{u_t}{\lambda_\nu}\right|^{\nu}\right\}. \quad (10)$$

The use and analysis of the MLE method for the estimation problem is well known. The main feature is that maximum likelihood estimators achieve optimal accuracy, in the sense that they are asymptotically consistent, and achieve the Cramér-Rao lower bound. Despite these advantages, an important obstacle to employing this method is the difficulty of computing a value $\hat{\omega}_{MLE}$ that satisfies condition (4). In the next section, the practical applicability and small sample performance of the MLE procedure for BL-GARCH processes are studied by Monte Carlo simulations.

4 Monte Carlo experiments

To our knowledge, no results exist on the properties of the MLE estimators when we observe a sample generated by (1a)-(1c) and we estimate them using the previous likelihoods. Thus, we have designed and executed a Monte Carlo experiment using the distribution functions described in the previous section as data generating processes, with the aim of analyzing the sampling properties of the exact MLE estimators of the parameter vector ω for the BL-GARCH model. Through the Monte Carlo experiments, the model considered for $u_t = y_t - \mu_t$ is a BL-GARCH$(1,1)$ given by (1a), (1b), and (2). We consider several samples size $n = 100$, 300, 1000 and 3000. Two cases are studied in the simulation experiments. In the first case, the conditional mean, μ_t, is set equal to zero, while in the second we assume that it follows an AR(1) model. The data generating processes' parameters are summarized in Table 1 with the first three lines corresponding to the case $\mu_t = 0$ and the last line to $\mu_t = \alpha_0 + \alpha_1 y_{t-1}$. Throughout the simulations, we consider a

Student-t with five degrees of freedom. Thus, the first four moments of the conditional density exist. For the GED distribution, we assume that the tail-thickness parameter is equal to three. We note particularly that we choose a small bilinear effect (c_1) in Model 4 in order to assess the capability of detecting it by maximum likelihood. Other simulation results are available upon request.

Table 1: Data generating processes (DGPs)

DGP	α_0	α_1	a_0	a_1	b_1	c_1
Model 1	0	0	0.01	0.09	0.9	0.15
Model 2	0	0	0.05	0.05	0.9	0.25
Model 3	0	0	0.2	0.05	0.75	0.35
Model 4	0.01	0.2	0.01	0.09	0.9	0.15

Tables 2-4 list the Monte Carlo mean, and the root mean square error (RMSE) for the parameter vector ω across $M = 1000$ Monte Carlo replications. The mean absolute error (MAE) are available from authors upon request. The simulation algorithm generates $n+500$ observations for each series, saving only the last n. This operation is performed in order to avoid dependence on initial values. The calculations were carried out in Matlab on a Pentium IV CPU 3.00 GHz computer. Inspection of Table 2, corresponding to the Normal case, reveals that, for all sample sizes, the averages obtained from the exact MLE are very close to the true parameter values. The corresponding RMSE are very small indicating evidence that estimators are consistent. Tables 3 and 4 present the estimated results from a non-Gaussian BL-GARCH model. We read, from these tables, that the averages of the parameter estimates are close to the true values under each of the underlying non-Normal error distributions. The RMSE are quite small and decrease when the sample sizes increase. Finally, Table 5 summarizes the results from the AR(1)-BL-GARCH$(1,1)$. Results reveal that parameter estimates are satisfactory in the sense that RMSE are small. We can also remark that, in general, the estimators of the autoregressive models seem not to be affected by the presence of the BL-GARCH errors. In addition, the method seems applicable, even if the sample size is less than 100, due to the fact that, in general, the true values are contained with ± 2 standard deviations of MLE's estimates: see Figures 1-4. But, we can note that in the simulation results, the bilinear effect probably cannot be detected for $n \leq 100$. An important finding is that our method is capable of estimating simultaneously the conditional mean and conditional variance parameters, while the proposed algorithm in Storti and Vitale (2003b) cannot.

Table 2: Estimated parameters for the centred Gaussian BL-GARCH model defined by (1b)-(2)

		$\hat{a}_0$		$\hat{a}_1$		$\hat{b}_1$		$\hat{c}_1$	
Model	n	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
1	100	0.09565	0.26478	0.09236	0.09654	0.76343	0.29031	0.15901	0.13342
	300	0.01770	0.02273	0.08398	0.04610	0.89119	0.07083	0.15306	0.04939
	1000	0.01150	0.00422	0.08943	0.01953	0.89718	0.02021	0.15252	0.03658
	3000	0.01047	0.00205	0.09022	0.01090	0.89883	0.01089	0.15129	0.01463
2	100	0.09694	0.12746	0.06368	0.08043	0.82400	0.21326	0.27502	0.12345
	300	0.05385	0.02661	0.04573	0.03235	0.89923	0.05210	0.25481	0.05712
	1000	0.05141	0.00972	0.04756	0.01637	0.90081	0.02073	0.25072	0.02904
	3000	0.05010	0.00469	0.04905	0.00884	0.90071	0.01073	0.24942	0.01540
3	100	0.22124	0.12953	0.07752	0.07592	0.70142	0.18151	0.36293	0.13423
	300	0.20409	0.06385	0.05699	0.03419	0.73837	0.08437	0.35189	0.07852
	1000	0.20094	0.03056	0.05201	0.01823	0.74715	0.03988	0.35259	0.04184
	3000	0.20013	0.01751	0.05023	0.01096	0.74947	0.02252	0.35015	0.02407

This table summarizes the estimated coefficients from the BL-GARCH$(1,1)$ model with the true value set of parameters $\{a_0, a_1, b_1, c_1\} = \{0.01, 0.09, 0.9, 0.15\}$, $\{0.05, 0.05, 0.9, 0.25\}$ and $\{0.2, 0.05, 0.75, 0.35\}$. MAE - Mean Absolute Errors, RMSE - root mean square error for the Gaussian BL-GARCH$(1,1)$. Monte Carlo simulations are computed with 1000 replications. Each replication gives a sample size $n = 100$, 300, 1000 and 3000 observations. Results for the mean absolute errors are available from authors upon request.

Table 3: Estimated parameters for the centred Student-t BL-GARCH model defined by (1b)-(2)

		$\hat{a}_0$		$\hat{a}_1$		$\hat{b}_1$		$\hat{c}_1$	
Model	n	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
1	100	0.10610	0.27096	0.10400	0.012400	0.73671	0.31462	0.14972	0.18310
	300	0.02153	0.04723	0.08986	0.05539	0.88264	0.09447	0.15937	0.06570
	1000	0.0119	0.00473	0.09018	0.0242	0.8961	0.0239	0.1522	0.0321
	3000	0.0104	0.00214	0.08951	0.0129	0.8994	0.0119	0.1493	0.0184
2	100	0.14528	0.18852	0.07323	0.09183	0.76433	0.27763	0.28221	0.18218
	300	0.07061	0.08397	0.05241	0.04431	0.87763	0.12948	0.26215	0.07734
	1000	0.05164	0.01263	0.04824	0.01606	0.90045	0.02198	0.25186	0.03909
	3000	0.05058	0.00624	0.04958	0.00906	0.90005	0.01209	0.25087	0.02241
3	100	0.24646	0.25177	0.08733	0.10274	0.67189	0.23289	0.35433	0.17808
	300	0.21268	0.09401	0.06466	0.04799	0.71045	0.13967	0.35592	0.10375
	1000	0.2018	0.04975	0.05065	0.02003	0.74725	0.05793	0.34400	0.05716
	3000	0.20206	0.02479	0.04994	0.01191	0.74910	0.02897	0.35035	0.03202

This table summarizes the estimated coefficients from the BL-GARCH$(1,1)$ model with the true value set of parameters $\{a_0, a_1, b_1, c_1\} = \{0.01, 0.09, 0.9, 0.15\}$, $\{0.05, 0.05, 0.9, 0.25\}$ and $\{0.2, 0.05, 0.75, 0.35\}$. RMSE - root mean square error for the Student-t BL-GARCH$(1,1)$. Monte Carlo simulations are computed with 1000 replications. Each replication gives a sample size $n = 100$, 300, 1000 and 3000 observations.

Table 4: Estimated parameters for the centred GED BL-GARCH model defined by (1b)-(2)

		$\hat{a}_0$		$\hat{a}_1$		$\hat{b}_1$		$\hat{c}_1$	
Model	n	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
1	100	0.07398	0.14326	0.10136	0.08697	0.77613	0.24907	0.16339	0.10712
	300	0.01609	0.01114	0.08845	0.03609	0.89029	0.04430	0.15484	0.03870
	1000	0.01105	0.00351	0.08917	0.01832	0.89884	0.01823	0.15094	0.01919
	3000	0.01025	0.00181	0.09067	0.01027	0.89866	0.01026	0.15070	0.01183
2	100	0.08216	0.09275	0.06423	0.07831	0.84761	0.16582	0.27555	0.10431
	300	0.05405	0.01935	0.04715	0.03173	0.89845	0.04357	0.25623	0.04708
	1000	0.05077	0.00738	0.04875	0.01547	0.90018	0.01819	0.25067	0.02358
	3000	0.05037	0.00393	0.04956	0.00866	0.90008	0.00976	0.25066	0.01287
3	100	0.21931	0.11359	0.07424	0.07298	0.70846	0.16225	0.36299	0.12272
	300	0.21268	0.09401	0.06466	0.04799	0.71045	0.13967	0.35592	0.10375
	1000	0.20262	0.02578	0.05130	0.01631	0.74550	0.03466	0.35097	0.03392
	3000	0.20020	0.01393	0.05082	0.01040	0.74919	0.01874	0.35114	0.01958

This table summarizes the estimated coefficients from the BL-GARCH$(1,1)$ model with the true value set of parameters $\{a_0, a_1, b_1, c_1\} = \{0.01, 0.09, 0.9, 0.15\}$, $\{0.05, 0.05, 0.9, 0.25\}$ and $\{0.2, 0.05, 0.75, 0.35\}$. RMSE - root mean square error for the GED BL-GARCH$(1,1)$. Monte Carlo simulations are computed with 1000 replications. Each replication gives a sample size $n = 100$, 300, 1000 and 3000 observations.

Table 5: Estimated parameters for the centred AR-BL-GARCH model defined by (1a),(1b), and (2)

Distribution	n	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{a}_0$	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$
Normal	100	0.0257	0.18447	0.10767	0.09233	0.78056	0.16659
		[0.09254]	[0.08737]	[0.24972]	[0.10633]	[0.28043]	[0.13219]
	300	0.01897	0.19377	0.01874	0.0838	0.88978	0.15629
		[0.04338]	[0.05949]	[0.02728]	[0.04697]	[0.07751]	[0.05055]
	1000	0.01103	0.19871	0.01142	0.08847	0.89841	0.15103
		[0.02069]	[0.03167]	[0.00441]	[0.01995]	[0.02064]	[0.02422]
	3000	0.01117	0.19905	0.01042	0.08998	0.89941	0.15086
		[0.01189]	[0.0187]	[0.00209]	[0.01134]	[0.01117]	[0.01444]
Student-t	100	0.0209	0.18876	0.08852	0.09884	0.7457	0.15474
		[0.07803]	[0.10338]	[0.23158]	[0.11957]	[0.32115]	[0.17796]
	300	0.01472	0.19351	0.02124	0.08836	0.87681	0.15205
		[0.03494]	[0.05758]	[0.03351]	[0.05389]	[0.10593]	[0.06717]
	1000	0.01123	0.19838	0.01206	0.09095	0.89505	0.15105
		[0.01737]	[0.031]	[0.00523]	[0.02215]	[0.0227]	[0.03156]
	3000	0.01053	0.19941	0.01054	0.09075	0.89861	0.15111
		[0.01013]	[0.01748]	[0.00278]	[0.01286]	[0.01323]	[0.01999]
GED	100	0.01715	0.18247	0.09296	0.09692	0.77744	0.16364
		[0.09838]	[0.10735]	[0.10735]	[0.10405]	[0.28519]	[0.11175]
	300	0.01680	0.19660	0.01619	0.08756	0.89222	0.15398
		[0.04067]	[0.05591]	[0.02660]	[0.04534]	[0.06616]	[0.04133]
	1000	0.01253	0.19973	0.01098	0.0914	0.89718	0.15228
		[0.021]	[0.03152]	[0.004]	[0.02045]	[0.02062]	[0.02077]
	3000	0.01059	0.19945	0.01026	0.09187	0.89782	0.15056
		[0.01288]	[0.01753]	[0.00232]	[0.01533]	[0.01465]	[0.01379]

This table summarizes the estimated coefficients from the AR(1)-BL-GARCH$(1,1)$ model with the true value set of parameters $\{\alpha_0, \alpha_1, a_0, a_1, b_1, c_1\} = \{0.01, 0.2, 0.01, 0.09, 0.9, 0.15\}$. RMSE - root mean square error for the AR(1)-BL-GARCH$(1,1)$ in brackets. Monte Carlo simulations are computed with 1000 replications. Each replication gives a sample size $n = 100$, 300, 1000 and 3000 observations.

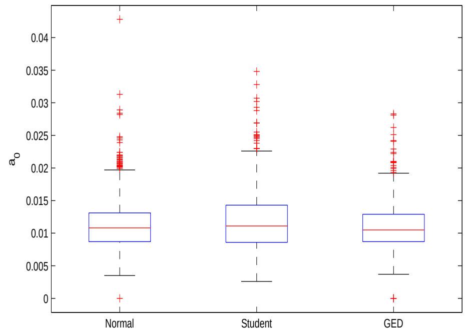

Fig. 1: Boxplot of estimates of the parameter $a_0 = 0.01$ of Model 1 under Normal, Student-t and GED distributions ($n = 1000$)

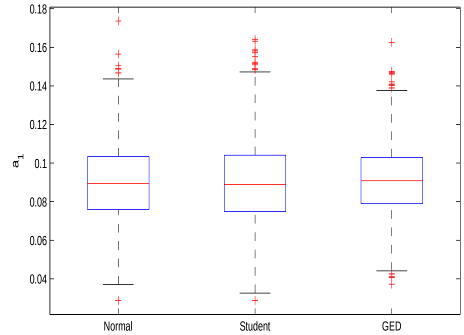

Fig. 2: Boxplot of estimates of the parameter $a_1 = 0.09$ of Model 1 under Normal, Student-t and GED distributions ($n = 1000$)

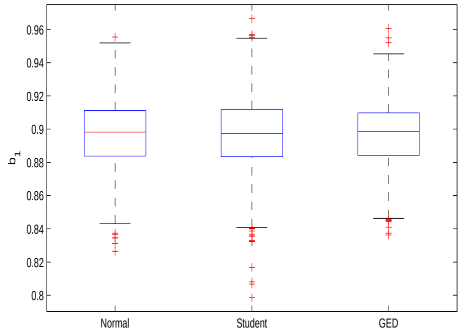

Fig. 3: Boxplot of estimates of the parameter $b_1 = 0.9$ of Model 1 under Normal, Student-t and GED distributions ($n = 1000$)

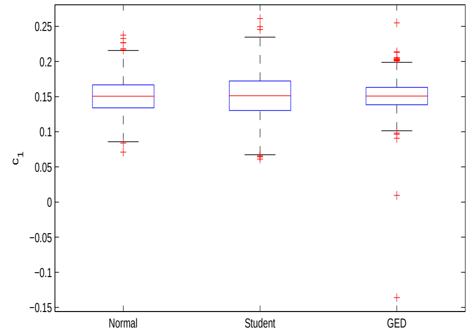

Fig. 4: Boxplot of estimates of the parameter $c_1 = 0.15$ of Model 1 under Normal, Student-t and GED distributions ($n = 1000$)

5 Empirical study

The daily continuously compounded returns of the S&P 500 stock market index are used for the empirical study in this paper to gauge the effectiveness of the BL-GARCH-type model with Normal, Student-t and GED innovations. In particular, we analyze the period from March 01, 1999, through January 31, 2001, which yields $n = 487$ daily observations, excluding public holidays. The sample closely corresponds to the data set used by Storti and Vitale (2003b). Table 6 gives the summary statistics of the S&P 500 log returns for the full sample. The mean and the standard deviation are quite small, while the estimated measure of skewness is significantly positive, indicating that the S&P 500 has non-symmetric returns. The kurtosis is a little larger than that of a Normal distribution, suggesting that fat-tailed distributions could better describe the unconditional distribution of the data. The results of the non-Normality test agree with prior literature using financial data, that is, a leptokurtic distribution is found for these S&P 500 log returns data. The Box-Pierce Q-tests of up to twenty-fourth order serial correlation for the levels and squares of the mean-corrected S&P 500 log returns were performed: $Q(24)$ and $Q^2(24)$ are significant for both the returns and squared returns series. In summary, the diagnostics suggest that a GARCH-class model would be appropriate, along with an error distribution that allows for greater kurtosis than the Gaussian distribution. Figure 5 gives the time plot of the data, and Figure 6 the returns distribution.

Table 6: Statistics of daily log returns of the S&P 500 stock market index.

Number of observations	487	Skewness	0.03708
Mean	0.0002055	Kurtosis	4.4643
Standard deviation	0.01283	Jarque-Bera test	43.3524
Minimum	-0.06004	$Q(24)$	39.2169
Maximum	0.04888	$Q^2(24)$	45.7208

The Jarque-Bera test critical value at significance level of 5% is 5.85423. Q and Q^2 are the Box Pierce statistics for the levels and square of the S&P 500 log returns, respectively, using 24 lags. The critical value at level 5% is 36.4150.

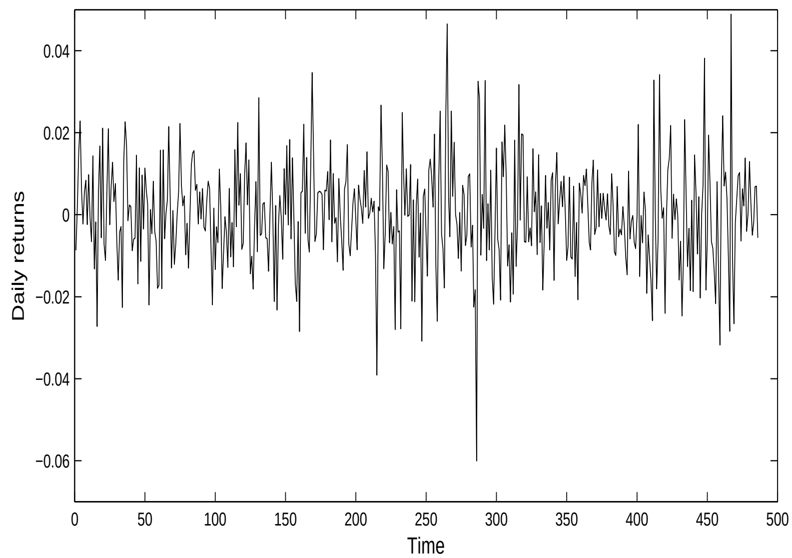

Fig. 5: S&P 500 daily returns 03/01/99-01/31/01

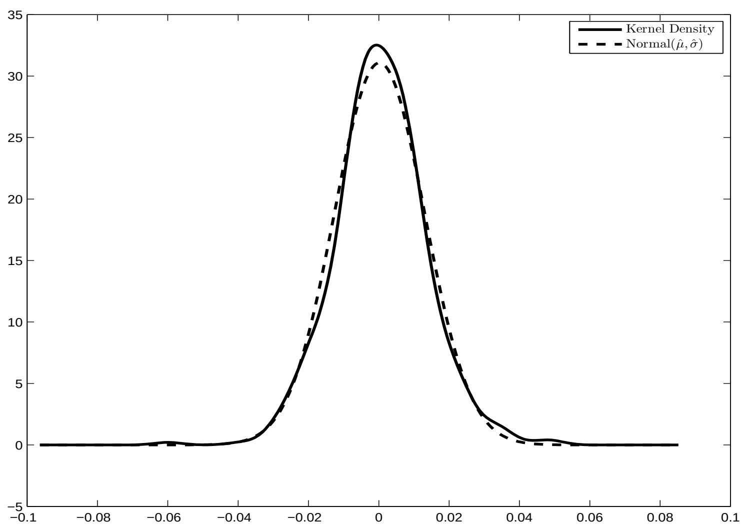

Fig. 6: Non-parametric density of S&P 500 daily returns (solid line) and probability density function of the Normal distribution (dotted line)

The parameter estimates from the GARCH model with Normal, Student-t and GED errors are also provided in Table 7, with the robust standard errors due to Bollerslev and Wooldridge (1992) shown in parentheses. While Table 8 summerizes the maximum likelihood estimation results from the BL-GARCH model. In order to compare the different models, various goodness-of-fit statistics are used. The diagnostics, summarized in Tables 7 and 8, are the log-likelihood function at its maximum, and the Akaike Information Criterion (AIC). We further report the values of the Box-Pierce (Q) statistics for the standardized and squared standardized residuals with, in parentheses, the corresponding p-values as a check of the empirical validity of the models. It appears from Table 7 that the estimates $\hat{a}_1$ and $\hat{b}_1$ in the GARCH$(1,1)$ are significant at the 5% level with the volatility coefficient greater in magnitude. Hence the hypothesis of constant variance can be rejected, at least within sample. Furthermore, the stationarity condition is satisfied for the three distributions because $\hat{a}_1 + \hat{b}_1 < 1$ at the maximum of the respective log-likelihood functions.

The fit of the BL-GARCH$(1,1)$ model, in Table 8, shows that the innovation and volatility spillovers are significantly different from zero. Further, the estimated asymmetric volatility response ($\hat{c}_1$) is negative and significant for all models, confirming the usual expectation in stock markets where downward movements (falling returns) are followed by higher volatility than upward movements (increasing returns). In all cases, the tail parameter estimates are strongly significant: far from two for the GED distribution, ν large through insignificant for the Student-t distribution but $1/\nu$ different to zero.

We note that, in the Gaussian case, our estimate's results in the fifth column of Table 8 are significantly similar to those given in Table 3 in Storti and Vitale (2003b). The latter results are recalled in the footnote of Table 8. In addition, we can remark that the standard errors in our study are more robust than those given in Storti and Vitale.

The results for the Q-statistics given in Tables 7 and 8 are not significant up to order 12 and also order 24, which indicates that the GARCH$(1,1)$ as well as the BL-GARCH$(1,1)$ process are appropriate to model the conditional variance of the S&P 500 log-returns. However, the goodness-of-fit statistics as well as the residuals diagnostics indicate that the BL-GARCH performs better, compare to the GARCH model, in describing the conditional variance of the S&P 500 returns. Moreover, the possible usefulness of using fat-tailed innovations for the BL-GARCH model seems to be confirmed by the log-likelihood and the AIC values. This result confirms the general finding in the financial literature when comparing GARCH-family models with Normal and fat-tailed distributions.

Table 7: GARCH model estimates for the S&P 500 return

Parameters	Normal	Student-t	GED
$\hat{a}_0$	0.00000715	0.000006361	0.000006108
	($0.28104\ 10^{-10}$)	($0.20586\ 10^{-10}$)	($0.20144\ 10^{-10}$)
$\hat{a}_1$	0.0568184	0.0504452	0.0499355
	(0.0007625)	(0.0006168)	(0.0006111)
$\hat{b}_1$	0.900428	0.911250	0.913097
	(0.0026979)	(0.0017480)	(0.0018101)
$\hat{v}$	-	8.676471	1.514841
	(-)	(9.929491)	(0.017712)
Goodness-of-fit statistics			
Log-lik	1435.91706	1441.94452	1441.04650
AIC	-2865.83441	-2875.88904	-2874.09300
Diagnostics			
$Q(12)$	15.892571	16.018537	16.010225
	(0.196206)	(0.190388)	(0.190768)
$Q(24)$	32.082112	32.206641	32.159427
	(0.124969)	(0.121951)	(0.123088)
$Q^2(12)$	4.682633	4.589348	4.566436
	(0.967748)	(0.970309)	(0.970918)
$Q^2(24)$	19.326186	19.483667	19.377018
	(0.734378)	(0.725728)	(0.731595)

This table provides the estimated coefficients, standard errors for the GARCH equation for the S&P 500 log returns index market. $\hat{a}_0$ is the constant in the conditional standard deviation equation, $\hat{a}_1$ is the ARCH coefficient, $\hat{b}_1$ is the GARCH coefficient, $\hat{c}_1$ is the leverage effect, $\hat{v}$ is the degrees of freedom. Log-lik is the maximized log likelihood. AIC is the Akaike Information Citerion and BIC the Bayesian Information Criterion. Q and Q^2 are the Box Pierce statistics for the standardized and squared standardized residuals respectively, using 12 and 24 lags with p-values in square brackets. The critical values at significant level of 5% are 21.026069 and 36.415028 respectively.

Table 8: BL-GARCH model estimates for the S&P 500 return

Parameters	Normal	Student-t	GED
$\hat{a}_0$	0.000011394	0.000009243	0.00001057
	(0.14775 10^{-10})	(0.13496 10^{-10})	(0.15424 10^{-10})
$\hat{a}_1$	0.060119	0.0513906	0.0559965
	(0.0006226)	(0.0005249)	(0.0006235)
$\hat{b}_1$	0.880531	0.900687	0.888784
	(0.0013771)	(0.0011259)	(0.0013873)
$\hat{c}_1$	-0.271323	-0.249673	-0.261943
	(0.0027800)	(0.0030085)	(0.0031051)
$\hat{v}$	-	14.943269	1.741412
	(-)	(67.446372)	(0.024844)
Goodness-of-fit statistics			
Log-lik	1456.47965	1458.63396	1457.65676
AIC	-2904.95930	-2907.26792	-2905.31353
Diagnostics			
$Q(12)$	15.876516	15.739415	15.833319
	(0.196958)	(0.203465)	(0.198990)
$Q(24)$	32.314595	32.501633	32.417392
	(0.119381)	(0.115033)	(0.116975)
$Q^2(12)$	4.477202	4.455119	4.485078
	(0.973208)	(0.973756)	(0.973011)
$Q^2(24)$	22.177684	22.542585	22.475412
	(0.568662)	(0.546916)	(0.550914)

This table provides the estimated coefficients, standard errors for the conditional standard deviation equation for the S&P 500 log returns index market. $\hat{a}_0$ is the constant in the conditional standard deviation equation, $\hat{a}_1$ is the ARCH coefficient, $\hat{b}_1$ is the GARCH coefficient, $\hat{c}_1$ is the leverage effect, $\hat{v}$ is the degrees of freedom. Log-lik is the maximized log likelihood. AIC is the Akaike Information Citerion and BIC the Bayesian Information Criterion. Q and Q^2 are the Box Pierce statistics for the standardized and squared standardized residuals respectively, using 12 and 24 lags with p-values in square brackets. The critical values at significant level of 5% are 21.026069 and 36.415028 respectively. The estimated parameters and standard errors (in parentheses) in Storti and Vitale (2003b) are: $a_0 = 2.09 \times 10^{-5}$ (7.07×10^{-6}), $a_1 = 0.071$ (0.031), $b_1 = 0.801$ (0.061) and $c_1 = -0.266$ (0.056).

6 Conclusion

This work extends the study of the BL-GARCH model proposed by Storti and Vitale (2003a) considering elliptical distributions and it gives new probabilistic results concerning the stationarity of the process and the moments. We obtain exact maximum likelihood estimates for this class of processes under conditional elliptical distributions. The small-sample properties indicate that the approach can yield asymptotically efficient estimates. In addition, these results strongly suggest that the maximum likelihood estimation inference procedure can be used to estimate the parameters of the BL-GARCH model, even in small samples (100 observations). Further, one advantage of the maximum likelihood estimator procedure proposed in this paper, compared to the method used by Storti and Vitale (2003b), is that it could simultaneously estimate the parameters of the BL-GARCH when the conditional mean is assumed non-constant. Further, the empirical results reveal that the BL-GARCH-t$(1,1)$, i.e., a BL-GARCH model with conditional errors that are t-distributed, fits the choosen data set best. This result points out the interest of using fat-tailed distributions combined with non-linear variance when modelling financial time series.

Acknowledgement

This research was supported by an Australian Research Council Discovery Project grant (DP0666266).

Appendix

Score functions and Hessian matrix for the Normal distribution

In the expression (8), taking the differential of $L(\omega)$ with respect to the ω yields

$$\frac{\partial L(\omega)}{\partial \omega} = \sum_{t=1}^{n} \frac{u_t}{h_t^2} \frac{\partial \mu_t}{\partial \omega} + \frac{1}{2} \sum_{t=1}^{n} \frac{1}{h_t^2} \left(\frac{u_t^2}{h_t^2} - 1 \right) \frac{\partial h_t^2}{\partial \omega}. \tag{11}$$

The Hessian matrix is given by

$$\begin{aligned} \frac{\partial^2 L(\omega)}{\partial \omega \partial \omega'} &= -\sum_{t=1}^{n} \frac{1}{h_t^2} \frac{\partial \mu_t}{\partial \omega} \frac{\partial \mu_t}{\partial \omega'} - \sum_{t=1}^{n} \frac{u_t}{h_t^4} \frac{\partial \mu_t}{\partial \omega} \frac{\partial h_t^2}{\partial \omega'} \\ &\quad - \sum_{t=1}^{n} \frac{u_t}{h_t^4} \frac{\partial \mu_t}{\partial \omega} \frac{\partial h_t^2}{\partial \omega'} + \sum_{t=1}^{n} \frac{1}{h_t^4} \left(\frac{1}{2} - \frac{u_t^2}{h_t^2} \right) \frac{\partial h_t^2}{\partial \omega} \frac{\partial h_t^2}{\partial \omega'}. \end{aligned} \tag{12}$$

Score functions and Hessian matrix for the Student-t distribution

The score function is given by

$$\frac{\partial L(\omega)}{\partial\omega} = \sum_{t=1}^{n}\left[\frac{\nu+1}{\nu-2}\frac{u_t}{h_t^2}\left(1+\frac{u_t^2}{h_t^2(\nu-2)}\right)^{-1}\right]\frac{\partial\mu_t}{\partial\omega} + \frac{1}{2}\sum_{t=1}^{n}\left[\frac{\nu+1}{\nu-2}\frac{u_t^2}{h_t^2}\left(1+\frac{u_t^2}{h_t^2(\nu-2)}\right)^{-1}-1\right]\frac{1}{h_t^2}\frac{\partial h_t^2}{\partial\omega}, \tag{13}$$

and the Hessian matrix is equal to

$$\begin{aligned}\frac{\partial^2 L(\omega)}{\partial\omega\partial\omega'} &= \frac{\nu+1}{\nu-2}\sum_{t=1}^{n}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}\frac{u_t}{h_t^2}\left[\frac{2}{\nu-2}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}\frac{u_t}{h_t^2}-1\right]\frac{\partial\mu_t}{\partial\omega}\frac{\partial\mu_t}{\partial\omega'}\\
&+\frac{\nu+1}{\nu-2}\sum_{t=1}^{n}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}\frac{u_t}{h_t^4}\left[\frac{1}{\nu-2}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}\frac{u_t^2}{h_t^2}-1\right]\frac{\partial\mu_t}{\partial\omega}\frac{\partial h_t^2}{\partial\omega'}\\
&+\frac{\nu+1}{\nu-2}\sum_{t=1}^{n}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}\frac{u_t}{h_t^4}\left[\frac{1}{\nu-2}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}\frac{u_t^2}{h_t^2}-1\right]\frac{\partial h_t^2}{\partial\omega}\frac{\partial\mu_t}{\partial\omega'}\\
&+\frac{1}{2}\sum_{t=1}^{n}\frac{1}{h_t^4}\left[1+\frac{(\nu+1)u_t^2}{(\nu-2)h_t^2}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}\right]\\
&\times\left[\frac{u_t^2}{(\nu-2)h_t^2}\left(1+\frac{u_t^2}{(\nu-2)h_t^2}\right)^{-1}-2\right]\frac{\partial h_t^2}{\partial\omega}\frac{\partial h_t^2}{\partial\omega'}.\end{aligned} \tag{14}$$

Score functions and Hessian matrix for the GED distribution

In that case, the score function and the Hessian matrix, respectively, are given by

$$\frac{\partial L(\omega)}{\partial\omega} = \frac{1}{2}\sum_{t=1}^{n}\frac{\nu}{|\lambda_\nu|}\left(\frac{u_t}{h_t}\right)^{\nu}\frac{1}{u_t}\frac{\partial\mu_t}{\partial\omega} + \frac{1}{2}\sum_{t=1}^{n}\frac{1}{h_t^2}\left[\frac{1}{2}\frac{l}{|\lambda_\nu|^{\nu}}\left(\frac{u_t}{h_t}\right)^{\nu}-1\right]\frac{\partial h_t^2}{\partial\omega}, \tag{15}$$

and

$$\begin{aligned}
\frac{\partial^2 L(\omega)}{\partial\omega\partial\omega'} &= -\frac{\nu}{|\lambda_\nu|^\nu}\left(\frac{\nu-3}{2}\right)\sum_{t=1}^{n}\frac{1}{h_t^2}\left(\frac{u_t^2}{h_t^2}\right)^{\frac{\nu}{2}-1}\frac{\partial\mu_t}{\partial\omega}\frac{\partial\mu_t}{\partial\omega'} \\
&\quad -\frac{1}{2}\frac{\nu}{|\lambda_\nu|^\nu}\sum_{t=1}^{n}\frac{u_t}{h_t^4}\left(\frac{u_t^2}{h_t^2}\right)^{\frac{\nu}{2}-1}\left[1+\left(\frac{\nu}{2}-1\right)\frac{u_t^2}{h_t^2}\right] \\
&\quad -\frac{1}{2}\frac{\nu}{|\lambda_\nu|^\nu}\sum_{t=1}^{n}\frac{u_t}{h_t^4}\left(\frac{u_t^2}{h_t^2}\right)^{\frac{\nu}{2}-1}\left[1+\left(\frac{\nu}{2}-1\right)\frac{u_t^2}{h_t^2}\right] \\
&\quad -\frac{1}{2}\sum_{t=1}^{n}\left[\frac{1}{4}\frac{\nu(\nu+2)}{|\lambda_\nu|^\nu}\left(\frac{u_t}{h_t}\right)^{\nu}-1\right]\frac{1}{h_t^4}\frac{\partial h_t^2}{\partial\omega}\frac{\partial h_t^2}{\partial\omega'}. \qquad (16)
\end{aligned}$$

Equations (11) through (16) require the computation of $\partial h_t^2/\partial\omega$. We provide it:

$$\begin{aligned}
\frac{\partial h_t^2}{\partial\omega} &= \left(1,u_{t-1}^2,h_{t-1}^2,h_{t-1}u_{t-1}\right)+2a_1u_{t-1}\frac{\partial u_{t-1}}{\partial\omega}+b_1\frac{\partial h_{t-1}^2}{\partial\omega} \\
&\quad +c_1\left(h_{t-1}\frac{\partial u_{t-1}}{\partial\omega}+\frac{1}{2}\frac{u_{t-1}}{h_{t-1}}\frac{\partial h_{t-1}^2}{\partial\omega}\right).
\end{aligned}$$

References

[1] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t student distribution. *Journal of the Royal Statistical Society Series B* **65**, 579–602.

[2] Baillie, R.T. and Bollerslev, T. (1991). Intra-day and inter-market volatility in foreign exchange rates. *Review of Economic Studies* **58**, 565–585.

[3] Baillie, R. and DeGennaro, R.P. (1990). Stock Returns and Volatility. *Journal of Financial & Quantitative Analysis* **25**, 203–214.

[4] Bingham, N.H. and Kiesel, R. (2002). Semi-parametric modelling in finance: theoretical foundations. *Quantitative Finance* **2**, 368–385.

[5] Black, F. (1976). Studies in stock price volatility changes. *in Proceedings of 1976 Meeting of Business and Economics Statistics Section, American Statistical Association*, 177–181.

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

[7] Bollerslev, T. (1987). A conditional heteroscedastic time series model for speculation prices and rates of return. *Review of Economics and Statistics* **69**, 542–547.

[8] Bollerslev, T., and Wooldridge, J. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* **11**, 143–172.

[9] Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* **52**, 115–127.

[10] Brockett, R.W. (1976). Volterra series and geometric ccontrol theory. *Automatica* 12, 167–176.

[11] Ding, Z., Granger, C.W., and Engle, R.F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1**, 83–106.

[12] Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the U.K. inflation. *Econometrica* **50**, 987–1008.

[13] Guégan, D. (1987) Different Representations of Bilinear Models , *Journal of Time series Analysis* **8**, 38-58.

[14] Guégan, D. (1988) Modèles Bilinéaires et Polynômiaux de Séries Chronologiques : Etude Probabiliste et Analyse Statistique, Thèse d'état, Université de Grenoble I.

[15] Guégan, D. and T.D. Pham (1989) A Note on the Estimation of the Parameters of the Diagonal Bilinear Models by the Least Squares Method, *Scand. Journ. of Stat. Th. and Appl.* **16**. 129-136.

[16] Giraitis, L., Leipus, R. and Surgailis, D. (2007). *Recent advances in ARCH modelling*. in G. Teyssière and A.P. Kirman (Eds.), Long Memory in Economics, Berlin: Springer.

[17] Gouriéroux, C. (1997). *ARCH model and financial application*. New York: Springer.

[18] Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review* **35**, 705–730.

[19] Hansen, P.R. and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?. *Journal of Applied Econometrics* **20**, 873–889.

[20] He, C. and Terasvirta, T. (1999). Properties of moments of a family of GARCH processes. *Journal of Econometrics* **92**, 173–192.

[21] Jones, M.B., and Faddy, M.J. (2003). A skew extension of the t distribution with applications. *Journal of the Royal Statistical Society series B* **65**, 159–174.

[22] Ling, S., and McAleer, M. (2002). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics* **106**, 109–117.

[23] McAleer, M. (2005). Automated inference and learning in modeling financial volatility. *Econometric Theory* **21**, 232–261.

[24] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* **59**, 347–370.

[25] Patton, A. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics* **2**, 130–168.

[26] Peiró, A. (1999). Skewness in financial returns. *Journal of Banking and Finance* **23**, 847–862.

[27] Storti, G., Vitale, C. (2003a). BL-GARCH models and asymmetries in volatility. *Statistical Methods and Applications* **12**, 19–40.

[28] Storti, G., Vitale, C. (2003b). Likelihood inference in BL-GARCH models. *Computational Statistics* **18**, 387–400.

[29] Tong, H. (1990). *Non-linear time series: a dynamical system approach*. Oxford Publications: Oxford University Press.

[30] Tsay, R.S. (2002). *Analysis of financial time series*. New York: Wiley.

[31] Zivot, E., and Wang, J. (2005). *Modeling financial time series with SPLUS*. New York: Springer Berlin Heidelberg.

[32] Subba Rao, T., Gabr,M.M. (1984). *An introduction to bispectral analysis and bilinear time series models*. Lectures Notes in Statistics, 24. Springer Verlag.