



**HAL**  
open science

# The accuracy of asymmetric GARCH model estimation

Olivier Darné, Amélie Charles

► **To cite this version:**

Olivier Darné, Amélie Charles. The accuracy of asymmetric GARCH model estimation. 2020. hal-01943883

**HAL Id: hal-01943883**

**<https://hal-audencia.archives-ouvertes.fr/hal-01943883>**

Preprint submitted on 28 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The accuracy of asymmetric GARCH model estimation

Olivier Darne, LEMNA

Amélie Charles, Audencia Business School

## Abstract

This paper reviews eight software packages when estimating asymmetric GARCH models (from their default option). We consider the numerical consistency of GJR-GARCH, TGARCH, EGARCH and APARCH estimations with Normal and Student distributions as well as out-of-sample forecasting accuracy, using the model confidence set procedure. We show that results are clearly software-dependent for both asymmetric volatility models, especially for the  $t$ -ratios. The out-of-sample forecast results show that the differences in estimating symmetric and asymmetric GARCH models imply slight differences in terms of forecast accuracy, not statistically significant, except in few cases from the QLIKE loss function. Further, the results indicated that the different specifications of the asymmetric GARCH-type models used by the different packages appear to have no significant effect on their forecast accuracy.

Keywords: EGARCH, GJR-GARCH, TARCH, APARCH, accuracy, forecasting, software.

JEL Classification: C22.

# 1 Introduction

Understanding the behavior of volatility in financial assets is important to risk management, derivative pricing and hedging, market making, market timing, portfolio selection, and many other financial activities. Autoregressive conditionally heteroscedastic (ARCH) models introduced by Engle (1982) and extended to GARCH models, independently, by Bollerslev (1986) and Taylor (1986), were developed to capture the two most important stylized facts of returns, which are heavy-tailed distributions and volatility clustering. GARCH estimation is widely available in a number of software packages (e.g., EViews, Gauss, Gretl, Matlab, Ox, R, Rats, Stata, S-Plus, TSP).

A number of reviews have investigated the accuracy, ease to use, availability of documentation and other attributes of the softwares available for the estimation of univariate (symmetric) GARCH models (see, e.g., Brooks, 1997; McCullough and Renfro, 1999; Brooks et al., 2001). McCullough and Renfro (1999) and Brooks et al. (2001, 2003) discussed numerical accuracy issues associated with maximizing the GARCH log-likelihood in several commercial softwares. They found that parameter starting values, initializations for the error and conditional variance-covariance series, optimization algorithm choice, use of analytic or numerical derivatives, and convergence criteria all influence the resulting numerical estimates of the GARCH parameters. As argued by Brooks et al. (2003) “a thorough examination of all of these issues is virtually impossible since the packages on the whole simply do not give sufficient detail on these points”.<sup>1</sup>

There is extensive empirical evidence about the existence of an asymmetric response of volatility to positive and negative past returns. In particular, increases in volatility are larger when previous returns are negative than when they have the same magnitude but are positive. The asymmetric volatility property is explained in

---

<sup>1</sup>McCullough and Renfro (1999) compared seven (anonymous) packages by using the FCP GARCH benchmark of Fiorentini et al. (1996) on the daily percentage nominal returns for the Deutschemark/British pound exchange rate of Bollerslev and Ghysels (1996). Brooks et al. (2001) used the same dataset by comparing nine software packages (EViews 3.1, Gauss 3/Fanpac, LIMDEP 8.0, Matlab 11, Microfit 4.0, Rats 4.3, SAS 6.12, Shazam 8.0, and TSP 4.5.). Laurent and Peters (2002) also compared GARCH accuracy between their Ox-GARCH 2.3 package and four econometric softwares (EViews 4, PcGive 10, S-Plus 6 and TSP 4.5).

the literature in terms of the leverage effect (Black, 1976; Christie, 1982; Schwert, 1989) and the volatility feedback effect (French et al., 1987; Campbell and Hentchel, 1992; Bollerslev et al., 1992).<sup>2</sup>

There is a large number of non-linear GARCH models, including EGARCH (Nelson, 1991), GJR-GARCH (Glosten et al., 1993), TGARCH (Rabemananjara and Zakoian, 1993) and APARCH (Ding et al., 1993) which allows for asymmetry or leverage effect, so that the conditional variance can be affected differently by positive and negative shocks of the same magnitude. Asymmetric GARCH estimation is also widely available in a number of commercial software packages (e.g., EViews, Gauss, Matlab, Ox, Stata) and there are also a few free open source implementations (e.g., R, Gretl). However, the asymmetric GARCH models can be expressed in various different forms according to the software packages. This is a severe problem that makes interpretation and diagnostic checking of the model specification extremely difficult. To the best of our knowledge, Brooks et al. (2001) and Alexander (2008) are the only studies which compare the numerical consistency of asymmetric GARCH models, especially EGARCH and/or GJR-GARCH models. Brooks et al. (2001) examine EGARCH estimation and forecasting for four econometric softwares (EViews, Microfit, Rats and SAS), and Alexander (2008) compares the estimation of GJR-GARCH and EGARCH models from EViews and Matlab using daily log returns on the FTSE 100.

The aim of this paper is to review a number of the most widely used software packages when estimating asymmetric GARCH models. In this sense, our contribution extends the work of Brooks et al. (2001) who have shown how different software packages can lead to quite different results in the standard GARCH model using updated versions of software packages and new free packages. We consider the numerical consistency of the four most widely estimated asymmetric univariate models of conditional volatility, namely the GJR-GARCH, TGARCH, EGARCH and APARCH models with Normal and Student distributions<sup>3</sup>, using a number of

---

<sup>2</sup>Hansen and Lunde (2005) compared 330 GARCH-type models and concluded that in the case of the Deutschmark-Dollar exchange rate, no model beats the GARCH(1,1) model, whereas the GARCH(1,1) model is clearly inferior to asymmetric GARCH models in the case of IBM returns.

<sup>3</sup>See Table 3 for a presentation of the GARCH-type models and distributions available in the different software packages.

different econometric software packages.<sup>4</sup> We thus evaluate eight software packages: five commercial packages, including Financial Analysis Package 3.0 (FANPAC) for GAUSS 17, Econometrics Toolbox for Matlab 13a, G@RCH 8 package in Ox 8.1, Stata 14 and EViews 9, one semi-commercial package with the free Oxford MFE Toolbox for Matlab, and two free packages with the rugarch package for R 3.4 and Gretl 2018a.

To evaluate the software accuracy in estimating a GARCH model Brooks et al. (2001) used the GARCH model proposed by Fiorentini et al. (1996) as benchmark. Brooks et al. (2001) argue that “there is a need for a benchmark and also some consensus as to what forms of the model are preferable, and the latter should be estimable by all packages” (p.54). Unfortunately, it is currently not the case as the asymmetric GARCH models can be expressed in various different forms according to the software packages. Brooks et al. (2001) argue that the establishment of benchmarks for other standard non-linear models is long overdue. Therefore, we compare the estimation of asymmetric GARCH models from financial time series by using the default option of the software packages. All the packages analyzed use conditional maximum likelihood to estimate the models. We also address the issue of whether the differences in parameter estimation between packages make a difference from a practical perspective, by evaluating the out-of-sample forecasting accuracy. For that, we extend the work of Brook et al. (2001) in several ways: (i) we use updated versions of the software packages which have been modified (improved) since their study; (ii) we analyze four asymmetric GARCH-type models with Normal and Student distributions; (iii) we compare the estimation by trying to “harmonize” the options of the packages; (iii) we estimate the models on small, medium and large samples; and (iv) we compare their forecast performance by using the model confidence set (MCS) procedure proposed by Hansen et al. (2011)<sup>5</sup> and the one-step ahead out-of-sample volatility forecasts produced automatically by the packages and by own codes.

The remainder of the paper is structured as follows. Section 2 introduces the GARCH,

---

<sup>4</sup>In this paper we do not deal with detailed descriptions of the packages, their user-friendliness, flexibility, speed ... See Brooks (1997) for these issues.

<sup>5</sup>Brooks et al. (2001) is the only study which compares the out-of-sample forecasts of the conditional variance produced by a GARCH model estimated from different packages. They only compare the forecast values from the packages for one to eight step ahead forecasts.

GJR-GARCH, TGARCH, EGARCH and APARCH specifications. Section 3 briefly discusses the maximum likelihood estimation of these models. The comparison in estimating GARCH-type models from the default option of the packages is presented in Section 4. Section 5 proposes a robustness check with a Student distribution, small and medium sample sizes, and ‘harmonized’ options for all the packages. The forecasting accuracy is displayed in Section 6. Section 7 proposes a discussion and concludes.

## 2 GARCH-type models

The GARCH model was developed independently by Bollerslev (1986) and Taylor (1986a). The GARCH model extends the ARCH model by allowing the conditional variance to be dependent upon previous own lags.

Consider the returns series  $r_t$ , defined by  $r_t = \log P_t - \log P_{t-1}$ , where  $P_t$  is the observed price at time  $t$ , follows a Normal GARCH(1,1) model defined as

$$r_t = c + \varepsilon_t \quad (1)$$

$$\varepsilon_t = \sigma_t z_t, \quad z_t \sim i.i.d.N(0, 1), \quad (2)$$

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (3)$$

where  $\varepsilon_t$  denotes the ‘market shock’ or ‘unexpected return’, and is commonly taken as the mean deviation ( $r_t - \bar{r}$ ), with  $\bar{r}$  is the sample mean.<sup>6</sup> The parameters should satisfy  $\omega > 0$ ,  $\alpha \geq 0$  and  $\beta \geq 0$  to guarantee the positivity of the conditional variance. The stationarity of the process (second-order moment condition) is achieved when the restriction  $\alpha + \beta < 1$  is satisfied. Ling and McAleer (2002a, 2002b) have derived the regularity conditions of a GARCH(1,1) model, defined as follows:  $E[\varepsilon_t^2] = \frac{\omega}{1-\alpha-\beta} < \infty$  if  $\alpha + \beta < 1$ , and  $E[\varepsilon_t^4] < \infty$  if  $k\alpha^2 + 2\alpha\beta + \beta^2 < 1$ , where  $k$  is the conditional fourth moment of  $z_t$ .<sup>7</sup> Ng and McAleer (2004) show the importance to verify these conditions. The sum of  $\alpha$  and  $\beta$  quantifies the persistence of shocks to conditional variance, meaning that the effect of a volatility shock vanishes over time at an exponential rate.

---

<sup>6</sup>Note that the OLS estimate of  $c$  is  $\bar{r}$ .

<sup>7</sup>Under the assumption of Normal distribution  $k = 3$  and thus the condition becomes  $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$ . See Ling and McAleer (2002a, 2002b) for other distributions.

## 2.1 The GJR-GARCH model

The GJR-GARCH model developed by Glosten, Jagannathan et Runkle (1993) is constructed to capture the asymmetric leverage volatility effect. Specification for the conditional variance of GJR-GARCH(1,1) model is

$$\begin{aligned}\sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \gamma I(\varepsilon_{t-1} < 0) \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ &= \omega + (\alpha + \gamma I(\varepsilon_{t-1})) \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2\end{aligned}\quad (4)$$

where  $I(\varepsilon_{t-1} < 0) = 1$  if  $\varepsilon_{t-1} < 0$ , and 0 otherwise. The volatility is positive if  $\omega > 0$ ,  $\alpha > 0$ ,  $\gamma \geq 0$ ,  $\alpha + \gamma \geq 0$  and  $\beta \geq 0$ . The process is defined as stationary if the constraint  $\alpha + \beta + (\gamma/2) < 1$  is satisfied. Ling and McAleer (2002b) have derived the regularity conditions for a GJR-GARCH(1,1), defined as follows:  $E[\varepsilon_t^2] < \infty$  if  $\alpha + \beta + \delta\gamma < 1$ , and  $E[\varepsilon_t^4] < \infty$  if  $k\alpha^2 + 2\alpha\beta + \beta^2 + \beta\gamma + k\alpha\gamma + k\delta\gamma^2 < 1$ .<sup>8</sup>

Asymmetry exists if  $\gamma > 0$ , i.e. positive and negative shocks of equal magnitude have different effects on conditional volatility. The asymmetry is observed as the impulse ( $\alpha + \gamma$ ) of negative shocks, which is larger than the impulse ( $\alpha$ ) of positive shocks. In this model, good news and bad news have different effects on the conditional variance: good news has an impact of  $\alpha$  while bad news has an impact of  $(\alpha + \gamma)$ . The GJR-GARCH model nets the GARCH model when  $\gamma = 0$ .

We can find a different parametrization for the GJR-GARCH model in other software packages. In Stata the dummy variable is defined as  $I(\varepsilon_{t-1} > 0) = 1$  if  $\varepsilon_{t-1} > 0$ , and 0 otherwise. Therefore, we have for Stata

$$\begin{aligned}\sigma_t^2 &= \omega + (\alpha_{Stata} + \gamma_{Stata}) \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad \text{if } \varepsilon_{t-1} \geq 0 \\ &= \omega - \gamma_{Stata} \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad \text{if } \varepsilon_{t-1} < 0\end{aligned}\quad (5)$$

while we have from the equation (4) for Matlab-MFE, Matlab-Ek, Gauss-Fanpac, R-rugarch and Ox-G@rch

$$\begin{aligned}\sigma_t^2 &= \omega + \gamma_{Matlab} \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad \text{if } \varepsilon_{t-1} \geq 0 \\ &= \omega - (\alpha_{Matlab} + \gamma_{Matlab}) \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad \text{if } \varepsilon_{t-1} < 0\end{aligned}\quad (6)$$

<sup>8</sup>Under a Normal distribution and a Student- $t(v)$  distribution, with  $v > 5$ ,  $\delta = \frac{1}{2}$ . See Ling and McAleer (2002a, 2002b) for other distributions.

Finally, we obtain the following correspondence

$$\begin{aligned}\alpha_{Matlab} &= \alpha_{Stata} + \gamma_{Stata} \\ \gamma_{Matlab} &= -\gamma_{Stata}\end{aligned}$$

The GJR-GARCH model is specified in Gretl without dummy variable as follows:

$$\sigma_t^2 = \omega + \alpha (|\varepsilon_{t-1}| - \gamma \varepsilon_{t-1})^2 + \beta \sigma_{t-1}^2$$

We have the following correspondence<sup>9</sup>

$$\begin{aligned}\alpha_{Matlab} &= \alpha_{Gretl}(1 - \gamma_{Gretl}) \\ \gamma_{Matlab} &= 4\alpha_{Gretl}\gamma_{Gretl}\end{aligned}$$

## 2.2 The TGARCH model

A similar way to model asymmetric effects was introduced by Zakoian (1994), and developed further in Rabemananjara and Zakoian (1993), by defining the threshold GARCH (TGARCH) model:

$$\begin{aligned}\sigma_t &= \omega + \alpha^+ \varepsilon_{t-1}^+ + \gamma^- \varepsilon_{t-1}^- + \beta \sigma_{t-1} \\ &= \omega + \alpha I(\varepsilon_{t-1} \geq 0) \varepsilon_{t-1} + \gamma I(\varepsilon_{t-1} < 0) \varepsilon_{t-1} + \beta \sigma_{t-1}\end{aligned}\quad (7)$$

where  $\varepsilon_t^+ = \varepsilon_t$  if  $\varepsilon_t > 0$ ,  $\varepsilon_t^+ \equiv 0$  otherwise,  $\varepsilon_t^- \equiv \varepsilon_t - \varepsilon_t^+$ , and  $I(\varepsilon_{t-1} < 0) = 1$  if  $\varepsilon_{t-1} < 0$ , and 0 otherwise. The TGARCH model is similar to the GJR-GARCH model but the volatility is specified in terms of  $\sigma_t$  instead of  $\sigma_t^2$ . Therefore, the TGARCH model does not impose any positivity restrictions on the volatility coefficients. However, in such a case,  $\sigma_t$  is no longer assumed to be positive which can be problematic in terms of probabilistic properties on the standardized innovations. The conditions for existence of moments can be found in He and Teräsvirta (1999a).

---

<sup>9</sup>Note that Gretl provides estimation of GJR-GARCH parameters for both parametrizations.



We can find different parametrizations for the TGARCH model among the software packages. In Gretl, Eviews<sup>10</sup> and R-rugarch the parametrization is given by

$$\sigma_t = \omega + \alpha_{Gretl} (|\varepsilon_{t-1}| - \gamma_{Gretl} \varepsilon_{t-1}) + \beta \sigma_{t-1} \quad (8)$$

which can be re-written as

$$\sigma_t = \begin{cases} \omega + \alpha_{Gretl}(1 - \gamma_{Gretl})\varepsilon_{t-1} + \beta\sigma_{t-1} & \text{if } \varepsilon_{t-1} \geq 0 \\ \omega - \alpha_{Gretl}(1 + \gamma_{Gretl})\varepsilon_{t-1} + \beta\sigma_{t-1} & \text{if } \varepsilon_{t-1} < 0 \end{cases}$$

Matlab-MFE use a dummy variable in the equation of the TGARCH(1,1) model, defined as

$$\sigma_t = \omega + \alpha_{Matlab} |\varepsilon_{t-1}| + \gamma_{Matlab} I(\varepsilon_{t-1} < 0) |\varepsilon_{t-1}| + \beta \sigma_{t-1} \quad (9)$$

where  $I(\varepsilon_{t-1} < 0) = 1$  if  $\varepsilon_{t-1} < 0$ , and 0 otherwise. This equation can be re-written as

$$\sigma_t = \begin{cases} \omega + \alpha_{Matlab} \varepsilon_{t-1} + \beta \sigma_{t-1} & \text{if } \varepsilon_{t-1} \geq 0 \\ \omega - (\alpha_{Matlab} + \gamma_{Matlab}) \varepsilon_{t-1} + \beta \sigma_{t-1} & \text{if } \varepsilon_{t-1} < 0 \end{cases}$$

We have the following correspondence between Gretl/Eviews/R-rugarch and Matlab-MFE

$$\begin{aligned} \alpha_{Matlab} &= \alpha_{Gretl}(1 - \gamma_{Gretl}) \\ \gamma_{Matlab} &= 2\alpha_{Gretl}\gamma_{Gretl} \end{aligned}$$

In Stata the dummy variable is defined as  $I(\varepsilon_{t-1} > 0) = 1$  if  $\varepsilon_{t-1} > 0$ , and 0 otherwise. Therefore, the parametrization is given by

$$\begin{aligned} \sigma_t &= \omega + (\alpha_{Stata} + \gamma_{Stata}) \varepsilon_{t-1} + \beta \sigma_{t-1} & \text{if } \varepsilon_{t-1} \geq 0 \\ &= \omega - \gamma_{Stata} \varepsilon_{t-1} + \beta \sigma_{t-1} & \text{if } \varepsilon_{t-1} < 0 \end{aligned} \quad (10)$$

---

<sup>10</sup>Note that EViews proposes the estimation of a TGARCH model but its TGARCH model is estimated on  $\sigma_t^2$  and not on  $\sigma_t$ , as defined in Zakoian (1994) and Rabemananjara and Zakoian (1993). Its TGARCH specification is rather similar to the GJR-GARCH model of Glosten et al. (1993). However, it is possible to estimate the TGARCH model in Eviews from the APARCH model by setting  $\delta = 1$ . The relationship between the parametrization of the TGARCH defined by Rabemananjara and Zakoian (1993) and that used in Gretl, Eviews and R-rugarch, which is derived from an APARCH model, is given by:  $\alpha^+ = \alpha_{Gretl}(1 - \gamma_{Gretl})$  and  $\gamma^- = -\alpha_{Gretl}(1 + \gamma_{Gretl})$ .

We obtain the following correspondence between Matlab-MFE and Stata<sup>11</sup>

$$\alpha_{Matlab} = \alpha_{Stata} + \gamma_{Stata}$$

$$\gamma_{Matlab} = -\gamma_{Stata}$$

### 2.3 The EGARCH model

Nelson (1991) proposes the exponential GARCH model (EGARCH), which can also capture volatility leverage. The EGARCH(1,1) model can be expressed in various different forms according to the software packages. This is a severe problem that makes interpretation and diagnostic checking of the model specification extremely difficult.

- Ox-G@RCH:

$$\ln(\sigma_t^2) = \omega + (1 - \beta L)^{-1}(1 + \alpha L)g(z_{t-1})$$

- Gauss-Fanpac / Matlab-MFE / Matlab-Econometrics / R-rugarch / Stata:

$$\ln(\sigma_t^2) = \omega + \theta_1 (|z_{t-1}| - E(|z_{t-1}|)) + \theta_2 z_{t-1} + \beta \ln(\sigma_{t-1}^2)$$

- SAS:

$$\ln(\sigma_t^2) = \omega + \alpha g(z_{t-1}) + \beta \ln(\sigma_{t-1}^2)$$

- EViews / Gretl:

$$\ln(\sigma_t^2) = \omega + \theta_1 (|z_{t-1}|) + \theta_2 z_{t-1} + \beta \ln(\sigma_{t-1}^2)$$

with  $g(z_t) = \theta_1 (|z_t| - E(|z_t|)) + \theta_2 z_t$ , where  $\theta_1 (|z_t| - E(|z_t|))$  denotes the magnitude effect, and  $\theta_2 z_t$  the sign effect. Thus, if  $\theta_2 < 0$  then negative innovations induce higher volatility than positive innovations of the same magnitude (asymmetry).  $E(|z_t|)$  depends on the assumption made on the unconditional density of  $\varepsilon_t$ . For the Normal distribution  $E(|z_t|) = \sqrt{2/\pi}$ . The specification of the volatility in terms of its logarithmic transformation implies that the parameters in this model are not restricted to

---

<sup>11</sup>Rodríguez and Ruiz (2012) use another parametrization for the TGARCH(1,1) model, given by  $\sigma_t = \omega + \alpha|\varepsilon_{t-1}| + \gamma\varepsilon_{t-1} + \beta\sigma_{t-1}$ . Carnero and Perez (2018) give the relationship between the parameters of this parametrization and those used in Matlab-MFE, Stata and Gretl (see Appendix).

positive values. As EGARCH model does not impose any positivity restrictions on the volatility coefficients a sufficient condition for the stationarity of the EGARCH model is  $|\beta| < 1$ . He et al. (2002) have considered the fourth moment structure of the EGARCH(1,1) model.<sup>12</sup>

The EGARCH(1,1) specification expressed in Ox-G@RCH is given by:

$$\ln(\sigma_t^2) = \omega(1 - \beta) + (1 + \alpha L)g(z_{t-1}) + \beta \ln(\sigma_{t-1}^2)$$

implying that  $\ln(\sigma_t^2)$  is a function of two lags of standardized residuals:

$$\begin{aligned} \ln(\sigma_t^2) = & \omega(1 - \beta) + \theta_1 (|z_{t-1}| - E(|z_{t-1}|)) + \theta_2 z_{t-1} \\ & + \alpha \theta_1 (|z_{t-2}| - E(|z_{t-2}|)) + \alpha \theta_2 z_{t-2} + \beta \ln(\sigma_{t-1}^2) \end{aligned}$$

To estimate the conditional variance as a function of one lag of standardized residuals the EGARCH(0,1) specification has to be chosen in Ox-G@RCH. Note there is a difference in the intercept term specification between Ox-G@RCH and Matlab-MFE, Matlab-Ek, Gauss-Fanpac, R-rugarch and Stata, with the following correspondence:  $\omega_{Matlab} = \omega_{G@RCH}(1 - \beta_{G@RCH})$ .

The EGARCH model with the Normal distribution is slightly differently estimated in EViews and Gretl, with the term  $E(|z_t|) = \sqrt{2/\pi}$  removed from the following specification:<sup>13</sup>

$$\ln(\sigma_t^2) = \omega + \theta_1 |z_{t-1}| + \theta_2 z_{t-1} + \beta \ln(\sigma_{t-1}^2)$$

Therefore, the intercept term in the conditional variance equation will differ by a factor of  $\theta_1 \sqrt{2/\pi}$  compared with that of Matlab/Gauss/R/Stata, with the following correspondence:  $\omega_{Matlab} = \omega_{EViews/Stata} + \theta_1 \sqrt{2/\pi}$ .

---

<sup>12</sup>Some authors make a distinction between asymmetry, referred to as the different impacts on conditional volatility of positive and negative shocks of equal magnitude, and leverage effect, regarded as the negative correlation between returns shocks and subsequent shocks to volatility (McAleer, 2014; Martinet and McAleer, 2018); Chang and McAleer (2017) propose the regularity conditions that an EGARCH(1,1) model obtained from a random coefficient complex nonlinear moving average process, should fulfill to capture asymmetry and/or leverage effects.

<sup>13</sup>The software packages advance some computational advantages to remove the term  $\sqrt{2/\pi}$  out of the summation operator.

## 2.4 The APARCH model

Ding, Granger et Engle (1993) propose the asymmetric power ARCH (APARCH) model which nests several ARCH-type models. The APARCH(1,1) model is defined as follows

$$\sigma_t^\delta = \omega + \alpha(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1})^\delta + \beta\sigma_{t-1}^\delta$$

where  $\omega > 0$ ,  $\alpha \geq 0$ , and  $\beta \geq 0$ . Parameter  $\delta$  ( $\delta > 0$ ) plays the role of a Box-Cox transformation of the conditional standard deviation  $\sigma_t$ .  $\gamma$ , with  $-1 < \gamma < 1$ , reflects the asymmetric effect. A positive (resp. negative) value of the  $\gamma$  means that past negative (resp. positive) shocks have a deeper impact on current conditional volatility than past positive shocks. The properties of the APARCH model have been studied by He and Teräsvirta (1999b) and He et al. (2008).

The APARCH model includes several ARCH extensions as special cases, including the GARCH(1,1) model when  $\delta = 2$  and  $\gamma = 0$ , the TGARCH(1,1) model when  $\delta = 1$ , and the GJR-GARCH(1,1) model when  $\delta = 2$ .<sup>14</sup>

This equation is the parametrization used in Eviews, Ox-G@arch, Gretl and R-rugarch. In MFE-Matlab and Stata the parametrization is slightly different with a positive sign for the asymmetric parameter  $\gamma$ , given by

$$\sigma_t^\delta = \omega + \alpha(|\varepsilon_{t-1}| + \gamma\varepsilon_{t-1})^\delta + \beta\sigma_{t-1}^\delta$$

## 3 Estimating GARCH models

### 3.1 Maximum likelihood estimation

The most commonly used method in estimating the vector of unknown parameters  $\theta$  of GARCH-type models is the method of maximum likelihood (ML) estimation. The function of the parameters set is called the log-likelihood function.<sup>15</sup>

---

<sup>14</sup>Note that some packages specify the GJR-GARCH (Gretl) and TGARCH (Eviews, Gretl and R-rugarch) models from the parametrization of the APARCH model.

<sup>15</sup>See Zivot (2009) and Xekalaki and Degiannakis (2010) for a discussion on the ML estimations for the GARCH-type models.

Given that  $\varepsilon_t = \sigma_t z_t$ , the log-likelihood function  $L$  of the standard normal distribution is

$$L = \sum_{t=1}^T l_t = -\frac{1}{2} \sum_{t=1}^T (\ln(2\pi) + \ln(\sigma_t^2) + z_t^2)$$

where  $l_t = -\frac{1}{2} (\ln(2\pi) + \ln(\sigma_t^2) + z_t^2)$ , and  $T$  is the number of observations.

As discussed in McCullough and Renfro (1999) and Brooks et al. (2001), there are several practical issues to consider in the maximization of likelihood function. Starting values for the model parameters  $\theta = (\omega, \alpha, \beta)$  need to be chosen and an initialization of  $\varepsilon_t^2$  and  $\sigma_t^2$  must be supplied.<sup>16</sup>

Once the log-likelihood is initialized, it can be maximized using numerical optimization techniques. The most popular methods of optimization make use of exact or approximated Hessian matrices, and iterative procedure that updates the current values of the estimates  $\hat{\theta}_{i+1}$  at iteration  $i + 1$ . A standard approach based on the exact Hessian matrix is the Newton-Raphson algorithm.<sup>17</sup> The iterative procedure is given by

$$\hat{\theta}_{i+1} = \hat{\theta}_i - \lambda_i \mathbf{H}(\hat{\theta}_i)^{-1} \mathbf{G}(\hat{\theta}_i) \quad (11)$$

where  $\hat{\theta}_i$  denotes the vector of estimated model parameters,  $\lambda_i$  is a scalar (called step) which may be fixed or variable,  $\mathbf{G}(\hat{\theta}_i) = \frac{\partial L(\hat{\theta}_i)}{\partial \hat{\theta}_i}$  is the gradient (or score) vector of the log-likelihood function (first derivative of the log-likelihood function with respect to the vector of unknown parameters), and  $\mathbf{H}(\hat{\theta}_i) = \frac{\partial^2 L(\hat{\theta}_i)}{\partial \hat{\theta}_i \partial \hat{\theta}_i}$  is the Hessian matrix (or a suitable approximation of it) of the log-likelihood function (second derivative of the log-likelihood function with respect to the vector of unknown parameters) all evaluated at iteration  $i$ .

An alternative procedure based on approximated Hessian matrix is the Berndt-Hall-Hausman (BHHH) algorithm (Berndt et al., 1974) which only uses first

<sup>16</sup>Almost all packages that estimate GARCH-type models use the sample variance as default for initializing the variance process ( $\sigma_t^2 = T^{-1} \sum_{t=1}^T \varepsilon_t^2$ ), except EViews which uses backcast exponential smoothing, and Matlab-MFE which uses a local average. Pelagatti and Lisi (2009) analyzed the precision of QML estimates under different choices of initialization and sample dimension.

<sup>17</sup>Analytic or numerical derivatives may be used for the Newton-Raphson algorithm. Fiorentini et al. (1996) provided algorithms for computing analytic derivatives for GARCH models.

derivatives. The BHHH algorithm is similar to the Newton-Raphson algorithm but instead of using the Hessian matrix of second derivatives  $\mathbf{H}(\hat{\theta}_i)$  it is based on an approximation formed by the sum of outer product of the gradient (OPG) vectors for the contribution of each observation to the objective function.<sup>18</sup> The Hessian matrix may be approximated by

$$-\mathbf{H}(\hat{\theta}_i) \approx \mathbf{B}(\hat{\theta}_i) = \sum_{t=1}^T \mathbf{G}_t(\hat{\theta}_i) \mathbf{G}'_t(\hat{\theta}_i) = \sum_{t=1}^T \frac{\partial l_t}{\partial \theta_i} \frac{\partial l_t}{\partial \theta'_i}$$

where  $\mathbf{G}_t(\hat{\theta})$  is the gradient of the log-likelihood function for each observation. The BHHH algorithm can be modified by employing the Marquardt correction which adds a correction matrix to the sum of the OPG vectors (Hessian approximation) as follows

$$\mathbf{B}(\hat{\theta}_i) = \sum_{t=1}^T \mathbf{G}_t(\hat{\theta}_i) \mathbf{G}'_t(\hat{\theta}_i) - aI$$

where  $I$  is the identity matrix, and  $a$  a positive number chosen by the algorithm.

Another common approach based on approximated Hessian matrix is the Quasi-Newton methods which are to build up an estimate of the Hessian at each iteration, starting from an initial estimate. The inverse of the Hessian matrix is given by

$$\mathbf{H}(\hat{\theta}_i) = \mathbf{H}(\hat{\theta}_{i-1}) + \mathbf{C}_{i-1}$$

where  $\mathbf{C}$  is a correction matrix. The Quasi-Newton algorithms differ only in their choice of this matrix, and the most widely used are the DFP algorithm proposed by Davidson (1959) and Fletcher and Powell (1963) or the BFGS algorithm developed by Broyden (1970), Fletcher (1970), Goldfarb (1979) and Shanno (1970).<sup>19</sup>

---

<sup>18</sup>Computational speed is increased by not calculating the actual Hessian matrix at each iteration for each time step, but the approximation can be weak when the log-likelihood function is far from its maximum, thus requiring more iterations to reach the optimum.

<sup>19</sup>The Quasi-Newton BFGS and DFP methods are like Newton-Raphson method in that they use both first and second derivative information, however, the Hessian matrix is approximated, reducing the computational requirements. These methods take more iterations than the Newton-Raphson algorithm but the use of an approximation produces a gain because it can be expected to converge in less overall time.

The estimates that maximize the conditional log-likelihood  $L$  are called the maximum likelihood (ML) estimates. Under suitable regularity conditions, the ML estimates are consistent and asymptotically Normally distributed and an estimate of the asymptotic covariance matrix of the ML estimates is constructed from an estimate of the final Hessian matrix from the optimization algorithm used. However, the appropriate regularity conditions have only been done for some GARCH models.<sup>20</sup>

Typically, the GARCH models are represented as unconstrained optimization problems. However, if the likelihood maximization must be achieved subject to non-linear constraints in the vector of unknown parameters, i.e. lower and upper bounds on the parameters, like the stationary constraint  $0 < \alpha + \beta < 1$  in the GARCH(1,1) model or the constraint  $-1 < \gamma < 1$  in the GJR-GARCH(1,1) model, constrained nonlinear programming (NLP) methods can be used, such as the sequential quadratic programming (SQP) algorithms.<sup>21</sup>

### 3.2 Estimators of variance-covariance matrix

It is important to know how are computed the standard errors of the parameter estimators. The different methods for computing standard errors for GARCH coefficients can lead to different standard error estimates, and thus to important implications for test significance on the parameters.

The variance-covariance matrix of the parameter estimators when the variables are assumed to be Normally distributed is calculated as the inverse of the Fisher information matrix  $\mathbf{I}(\hat{\theta})$ , defined as

$$\mathbf{Q}(\hat{\theta}) = \mathbf{I}(\hat{\theta})^{-1}$$

---

<sup>20</sup>The asymptotic properties, in particular the asymptotic normality, of the quasi-maximum likelihood estimator (QMLE) hold under mild conditions for GARCH models (e.g., Lee and Hansen, 1994; Berkes et al., 2003; Francq and Zakoian, 2004) and for GJR-GARCH models (Hamadeh and Zakoian, 2011). However, the statistical properties for the QMLE of the EGARCH(1,1) parameters are not available under general conditions, but rather only for special cases under highly restrictive conditions (Wintenberger, 2013; Kyriakopoulou, 2015; Martinet and McAleer, 2018).

<sup>21</sup>An overview of SQP methods can be found in Nocedal and Wright (1999) and Fletcher (2013).

It is diagonal in the case of a Normal density but otherwise it can be a complicated non-diagonal matrix, and thus when MLE is applied to non-Normal distributions the computations of the covariance matrix of estimators can be difficult. In general two estimators of the Fisher information matrix can be calculated. The first estimator is the expected values of the first derivatives of the log-likelihood function. It is known as the BHHH estimator or OPG estimator:

$$\mathbf{I}(\hat{\theta}) \approx E \left[ \frac{\partial \ln L}{\partial \theta} \frac{\partial \ln L}{\partial \theta'} \right]$$

The second estimator, called Hessian-based estimator, is the expected values of the negative of the second derivatives of the log-likelihood function:<sup>22</sup>

$$\mathbf{I}(\hat{\theta}) = -E \left[ \frac{\partial^2 \ln L}{\partial \theta \partial \theta'} \right] = -E \left[ \mathbf{H}(\hat{\theta}) \right]$$

These two estimators are asymptotically equivalent, but they could give different results in finite samples. Available evidence suggests that in small or moderate sized samples, the Hessian is preferable (Greene, 2007). However, in most cases, the BHHH estimator will be the easiest to compute.

If the log-likelihood function is maximized using Hessian-based algorithms, the Fisher information matrix can be obtained as a by-product of the optimization procedure, such as Newton-Raphson and BHHH methods. On the other hand, if Quasi-Newton methods, such as BFGS method, are employed, the estimated Hessian matrix is not guaranteed to converge to its exact value when approaching the optimum. Using those methods to obtain the estimated Fisher information matrix may thus be misleading.<sup>23</sup>

The assumption of Normally distributed standardized innovations is often violated by the data, especially in finance. This has motivated the use of alternative distributional assumptions (e.g., Student- $t$  or Skewed-Student- $t$  distributions). Alternatively,

---

<sup>22</sup>Often the expected value of the second derivatives of the log likelihood have to be approximated by taking their actual values, rather than the expected values at the maximum likelihood estimates (Efron and Hinkley, 1978).

<sup>23</sup>Note that, although the Fisher information matrix is known to be block diagonal, the efficient implementation of Newton or Newton-like algorithms requires the full Hessian matrix, which is not block-diagonal in small samples (Calzolari et al., 1993).



the MLE based on the Normal density may be given by a quasi-maximum likelihood (QML) interpretation. Bollerslev and Wooldridge (1992) showed that the maximization of the Normal log-likelihood function results in QML estimates of the parameter vector  $\theta$  that are consistent and asymptotically Normally distributed even when the distribution of  $z_t$  is non-Normal. In addition, Bollerslev and Wooldridge (1992) derived an asymptotic covariance matrix for the QMLEs that is robust to conditional non-Normality.<sup>24</sup> This matrix is estimated using both matrices of second order derivatives and outer products of first derivatives:

$$\mathbf{Q}(\hat{\theta}) = \mathbf{H}(\hat{\theta}_{QML})^{-1} \mathbf{B}(\hat{\theta}_{QML}) \mathbf{H}(\hat{\theta}_{QML})^{-1}$$

where  $\hat{\theta}_{QML}$  is the QML estimator of  $\theta$ , which is also called the “sandwich” estimator, and the coefficient standard errors as the (robust) Bollerslev-Wooldridge standard errors.

Under correct specification the QML estimator is equivalent to the other covariance estimators, as it gives asymptotically the inverse of the Fisher information matrix (thus, block-diagonal).

We evaluate eight of the most widely used software packages when estimating asymmetric GARCH models: Financial Analysis Package 3.0 (Fanpac) for Gauss 17 (Gauss-Fanpac), Econometrics Toolbox for Matlab 13a (Matlab-EK), the free Oxford MFE Toolbox for Matlab (Matlab-MFE)<sup>25</sup> (Sheppard, 2009), G@RCH 8 package in Ox 8.1 (Ox-G@RCH) (Laurent and Peters, 2001; Laurent, 2009), EViews 7, Stata 14, and two free software packages with GIG (Garch In Gretl) 2.21 package for Gretl 2018a (Lucchetti and Baliotti, 2011) and rugarch 1.4-0 package for R 3.4 (R-rugarch) (Galanos, 2018).<sup>26</sup>

Table 1 summarizes the MLE in the default option of the econometric packages in terms of numerical optimization techniques and standard errors of the estimated

<sup>24</sup>The QMLE is generally close to the exact MLE for symmetric departures from Normality. However, Engle and Gonzales-Rivera (1991) showed that the loss in efficiency may be quite high for non-symmetric distributions.

<sup>25</sup>The Oxford MFE Toolbox is the follow on to the UCSD\_GARCH toolbox.

<sup>26</sup>Another package is available for GARCH modelling in R with fGarch, which is a part of the Rmetrics suite. However, this package only estimates GARCH and APARCH models.

parameters. In terms of optimization routine, Gauss-Fanpac, Matlab-MFE, Gretl and Ox-G@RCH use a version of the BFGS algorithm whereas EViews uses the BHHH algorithm with the Marquart correction. Stata switches between the BHHH and BFGS algorithms.<sup>27</sup> Finally, Matlab-Econometrics and R-rugarch apply a SQP algorithm.<sup>28</sup> In terms of standard errors for GARCH coefficients, Gauss-Fanpac, EViews, Stata and Matlab-Econometrics compute standard ML standard-error estimates (OPG or Hessian-based estimators) whereas Ox-G@RCH and Gretl displays the Bollerslev-Wooldridge robust QML standard errors. Matlab-MFE and R-rugarch provide both standard (Hessian-based estimator) ML and robust QML standard errors.

## 4 Comparison in estimating GARCH-type models

In this section, the five GARCH models previously described are fitted to represent the evolution of the volatility of two series of daily returns. The series analyzed are daily returns of the FTSE100 index observed from April 1, 2003 to December 31, 2007 (1195 observations). This series is downloaded from the webpage of the Oxford-Man Institute of Quantitative Finance (Heber et al., 2009).<sup>29</sup>

It is important when comparing econometric softwares to have an appropriate dataset, especially to avoid the misleading effects of outliers on the estimation of the volatility.<sup>30</sup> For that, we have checked that all observations are lesser than  $5\hat{\sigma}_t$ , where  $\hat{\sigma}_t$  is an estimate of the conditional standard deviation.<sup>31</sup> All packages should be able to es-

<sup>27</sup>By default, Stata performs 5 iterations with the BHHH algorithm followed by 15 iterations with the BFGS algorithm, and then switch back to BHHH for five iterations, and so on. Lombardi and Gallo (2002) show that the use of a mixed algorithm can provide some gains in terms of iterations and RMSE.

<sup>28</sup>We use the following optimization packages: the maximization package *maxlikmt2.0* for Gauss 17; the maximization package *fminunc* for Matlab-MFE; the optimization toolbox *fmincon* for Matlab-Econometrics, where the function solves a quadratic programming subproblem at each iteration, and updates an estimate of the Hessian of the Lagrangian at each iteration using the BFGS formula; the SQP solver *solnp* of Ye (1997) based on an augmented Lagrange multiplier method with a SQP interior algorithm, implemented in R by Ghalanos and Theussl (2011).

<sup>29</sup>The data are available on the following website: <http://realized.oxford-man.ox.ac.uk/data>.

<sup>30</sup>It is well known that the outliers may pose difficulties for the identification and estimation of GARCH models governing the conditional volatility of returns (e.g., Franses and Ghijssels, 1999; Carnero et al., 2007, 2012, 2016).

<sup>31</sup>Carnero et al. (2012) and Rodriguez and Ruiz (2012) advocate this approach to detect outliers in

Table 1: Maximum likelihood estimation in the default option of econometric packages.

Softwares	Ox 7.2	Eviews 9	Matlab 13a	Gauss 17	R 3.4	Stata 14	Gretl 2018a
Packages	G@RCH 7		MFE	Fanpac 3.0	rugarch 1.4-0		GIG 2.21
Initialization for	sample	backcast	local	sample	sample	expected	sample
cond. variance	uncond.	exponential	average	uncond.	uncond.	uncond.	uncond.
	variance	smoothing	variance	variance	variance	variance	variance
Optimization	BFGS	BFGS with	BFGS	BFGS	Augmented	BHHH	BFGS
algorithm		Marcquart			Lagrange	& BFGS	
		correction			based SQP		
Derivatives	numeric	numeric	numeric	numeric	numeric	numeric	numeric
		& analytic					& analytic
Variance	robust	OPG	Hessian-based	Hessian-based	Hessian-based	Hessian-based	robust
estimator	QMLE		& robust QMLE	& robust QMLE	or OPG*	QMLE	QMLE
Convergence	0.0001	0.0001	0.000001	0.000001	0.000001	0.0001	depends on
criteria							computer power
Max. iterations	1000	500	500	1000	400	16000	500

Notes: Robust QMLE denotes the Bollerslev-Wooldridge robust covariance matrix of the QML estimated parameters (sandwich estimator). OPG denotes outer product of the gradients. SQP denotes sequential quadratic programming. \* denotes that the variance estimator depends on optimization algorithm since Stata switches between the BHHH and BFGS algorithms.

timate symmetric and asymmetric GARCH models, i.e. a constant with GARCH(1,1)-type errors.

Tables 4 and 5 give default estimation results for the GARCH, GJR-GARCH, TGARCH, EGARCH and APARCH models. As in Brooks et al. (2001) we do not intend to ‘pass judgement’ on the packages that are good or not, since defaults may differ between packages in terms of initializations, starting values for parameters, convergence tolerances . . . . Rather, we only display default results for comparison with one another.

**The GARCH models.** The intercepts in the conditional variance equation ( $\omega$ ) are similar for all the packages for both returns. However, it is the GARCH parameters where the differences across packages appear. The parameters on the lagged squared errors ( $\alpha$ ) are close for Gauss-Fanpac, Matlab-Econometrics, Ox-G@RCH, Stata, Gretl and R-rugarch at around 0.71, and higher than for EViews and Matlab-MFE at around 0.61. We find similar results for the parameters on the lagged conditional variance, but the values are higher for EViews and Matlab-MFE than those for the other packages. Note that the levels of persistence ( $\alpha + \beta$ ) are broadly in agreement for all the packages, with a persistence around 0.988.

For the  $t$ -ratios associated with the coefficients (Tables 4 and 5) it is evident that the differences across packages are more marked than they were for the parameter estimates. For example, the  $t$ -ratios for the GARCH parameters given by EViews, Stata and Matlab-Econometrics (based on OPG estimator) are around two times higher than those by Ox-G@rch, Gretl and R-rugarch (based on robust QMLE estimator). However, none of these differences are important for tests of significance since all of the parameters are statistically significant at the 1% level under all packages, except the intercept for Matlab-MFE and Ox-G@rch. The differences in  $t$ -ratios are arguably unsurprising since a similar result was found by Brooks et al. (2001) for GARCH(1,1) models, who argued that differences in default calculation methods for the standard errors can explain these differences. However, these results should be take with caution because the robust QML estimator is prone to give very different results in the case where the standardized innovations are not Normal. As a consequence, it is likely that

---

GARCH models.

the standardized innovations issued from the series are not Normal because softwares that implement robust standard errors give very far away results. Inference, and thus  $t$ -ratio must be biased for the other software packages.

**The GJR-GARCH models.** As for the GARCH model the intercepts are broadly in agreement for all the packages. The parameters on the lagged conditional variance are close for Gauss-Fanpac, Matlab-Econometrics, Ox-G@RCH, Stata, Gretl and R-rugarch, and lower than for EViews and Matlab-MFE. The parameters on the lagged squared errors are also the same for Gauss-Fanpac, Matlab-Econometrics, Ox-G@RCH, Stata (with  $\alpha^*$ ), Gretl and R-rugarch, but its values are negative or null for EViews and Matlab-MFE, which do not satisfy the condition  $\alpha > 0$  to have the volatility positive.<sup>32</sup> The asymmetric parameters ( $\gamma$ ) are close for Gauss-Fanpac, Matlab-Econometrics, Ox-G@RCH, Stata (with a negative sign), Gretl and R-rugarch, and for EViews and Matlab-MFE.

When one considers the  $t$ -ratios, the discrepancies between methods used by each of the packages become far more apparent. For example, the  $t$ -ratio on the asymmetric term in the conditional variance equation ( $\gamma$ ) varies from 1.44 to 5.79. These strong differences are important for test significance since the asymmetric parameters are not significant at the 10% level under Ox-G@RCH and R-rugarch, and significant for the other packages. This is strong difference can be explained by the (possible) non Normality of standardized innovations.

**The EGARCH models.** Even if the constant terms in the conditional variance equation are different across the packages due to the different forms to express the EGARCH models, once we modify the values of the coefficients ( $\omega^*$ ) in a consistent manner across the packages as defined in Section 2.2 we find that the intercepts are broadly in agreement for all the packages and returns.<sup>33</sup> The parameters on the lagged

---

<sup>32</sup>For Matlab-MFE the zero value for the lagged squared errors can be explained by the fact that its maximization package `fminunc` stopped the optimization with the following message “because the size of the current step is less than the selected value of the step size tolerance”. For Eviews by changing its default option in initializing of the conditional variance by the sample variance we found the same values of the parameters that those obtained from the other packages.

<sup>33</sup>Alexander (2008) compared the estimation of GARCH(1,1), GJR-GARCH(1,1) and EGARCH(1,1) models from EViews and Matlab using daily log returns on the FTSE 100 over the period 2003-2007.

squared errors are close for all the packages. However, the discrepancies become apparent for the two asymmetric parameters in the conditional variance equation. For Gauss-Fanpac, Matlab-Econometrics, Gretl, Stata and R-rugarch the coefficient values of the magnitude effect ( $\theta_1$ ) are close to one another, while those of EViews, Matlab-MFE and Ox-G@RCH are typically up to 30% away from the three others. We also find differences for the parameters of the sign effect ( $\theta_2$ ) for all the packages, which vary between -0.070 and -0.037, except for Gauss-Fanpac, Gretl, Stata and R-rugarch where the coefficients are very close. The differences across the packages for the EGARCH models can be explained by both the various specifications of the model and the estimation methods used by each of the packages.

As found previously the results also show differences across packages for the  $t$ -ratios. For example, the  $t$ -ratios for the lagged conditional variance displayed by EViews are around 3 times higher than those by Ox-G@rch and R-rugarch. However, these differences are not important for tests of significance since all of the parameters are statistically significant at the 1% level under all packages, except Stata with the parameter  $\theta_2$ .

**The TGARCH models.** All the coefficients are different between all the packages, except for the parameter  $\omega$ . When taking the corresponding coefficients for  $\alpha$  and  $\gamma$  we find similar values for Gretl and Stata. The differences are more important in terms of  $t$ -ratios, and the TGARCH model is rejected for R-rugarch since the asymmetric parameter is not significant at the 10% level.

**The APARCH models.** As for the TGARCH model the coefficients are different across the packages. Nevertheless, we find similarities for the values of parameters for some packages, such as  $\alpha$ ,  $\beta$  and  $\gamma$  are close for Stata and Gretl,  $\alpha$  for Matlab-MFE and R-rugarch,  $\beta$  and  $\gamma$  for Eviews and Matlab-Econometrics. Note that the power parameter  $\delta$  is close between the packages, ranging between 1.4 and 1.8, except for R-rugarch which is almost two times higher. The results also show differences for the  $t$ -ratios,

---

Her results showed that the parameter estimates are similar, except that the EViews optimizer for EGARCH model converges to a solution that is not sensible because the estimated long term volatility, given by  $\ln(\bar{\sigma}^2) = \frac{\omega}{1-\beta}$ , is too low. However, if we use the intercept term modified  $\omega^*$  we find that the long term volatility estimated by both packages is close.

with strong effects on the significance of the parameters for Matlab-MFE, R-rugarch and Gretl (based on robust QMLE estimator) which reject the APARCH model.<sup>34</sup>

Overall, for the asymmetric GARCH-type models the GJR-GARCH and APARCH models are not rejected only for Gauss-Fanpac/Matlab-Econometrics and Eviews/Stata, respectively, whereas all EGARCH and TGARCH models are not rejected, whatever the packages, except Stata and R-rugarch, respectively.<sup>35</sup> Further, the strong differences in terms of  $t$ -ratios between the packages using the OPG estimator and those using the robust QML estimator is likely due to the fact that the standardized innovations are not Normal.

## 5 Robustness

### 5.1 Student distribution

As shown in Table 2 the FTSE returns are characterized by non-Normality, with evidence of significant negative excess skewness and kurtosis. We re-estimate the five GARCH-type models with a Student distribution, which is the non-Normal distribution available for all the software packages (see Table 3). The results are given in Tables 10 and 11.

**The GARCH models.** We find similar results than those obtained with a Normal distribution, namely the parameter  $\omega$  is again similar for all the packages and the parameter  $\alpha$  and  $\beta$  are similar for Ox-G@rch, Matlab-Ek, R-rugarch, Stata and Gretl whereas they are lower and higher, respectively, for Eviews and Matlab-MFE. Opposed to the results obtained when the standardized innovations are assumed Normal, the results for the  $t$ -ratios are now slightly different when the standardized innovations follow a Student distribution since the  $t$ -ratios based on QML estimator are slightly lower than those based on OPG estimator. The parameter  $\eta$  of the Student distribution is close for

---

<sup>34</sup>Note that Ox-G@rch does not estimate this model due to problem of convergence.

<sup>35</sup>By analyzing different asymmetric models Rodriguez and Ruiz (2012) show that when the parameters satisfy the positivity, stationarity, and finite kurtosis conditions, the dynamics that the GJR-GARCH model can represent are heavily limited while those of the EGARCH and TGARCH models are less restricted.

all the packages.

**The asymmetric models.** The results are also identical to those with a Normal distribution in terms of comparison of parameters for the GJR-GARCH, EGARCH and TGARCH models. The differences in terms of  $t$ -ratios are lesser across the packages. However, some  $t$ -ratios based on the QML estimator are strongly affected. For example, Ox-G@rch and Gretl do not reject the GJR-GARCH model as the asymmetric parameter  $\gamma$  is become significant at the 10% level or Matlab-MFE rejects the TGARCH model as the parameter  $\alpha$  is become not significant at the 10% level. Finally, all the packages reject the APARCH model with a Student distribution.

Table 2: Descriptive statistics of FTSE returns (1195 observations).

mean (%)	Std. dev. (%)	Skewness	Kurtosis	JB test
0.020	0.643	-0.214*	5.278*	148.2*

Notes: \* denotes significance at the 5% level.

## 5.2 Sample size

We now evaluate the effect of the sample size on the estimations of the five GARCH-type models by using medium and small samples, namely 500 and 100 observations, respectively.

For the medium sample size the GARCH model is well estimated by all the packages with similar values for all the parameters, except for Eviews and Matlab-MFE (Tables 8 and 9). For these both packages their values of the parameter on the lagged squared errors ( $\alpha$ ) are very different (0.013 and 0.004 for Eviews and Matlab-MFE, respectively) compared to those of the other packages (0.025) and their  $t$ -ratios are not significant. However, only few asymmetric models are not rejected with a Normal distribution (the GJR-GARCH model for Matlab-Econometrics and the EGARCH models for Eviews and Ox-G@rch) because some parameter estimations are not significant and positivity or stationarity constraints are not satisfied.

When analyzing Tables 6 and 7 for the small sample size we observe that the results are worst since almost all the GARCH-type models are rejected, except the



EGARCH (Matlab-MFE and R-rugarch) and TGARCH (R-rugarch) models, whatever the distribution.<sup>36</sup> Note that in most models Gauss-Fanpac, Ox-G@rch and Stata do not succeed to estimate the parameters due to problem of convergence. This result confirms those obtained by Ng and Lam (2006) and Hwang and Pereira (2006) showing that the estimation of the GARCH-type models is dramatically affected by the sample size.

### 5.3 Identical algorithm and variance-covariance estimator

We now compare the accuracy of the packages by using the same optimization algorithm and variance-covariance estimator. This comparison can be interesting when a study estimate the volatility of asset returns using the GARCH-type models available in an econometric package from the previous options and another study compare these estimations with other GARCH-type models not available in the previous package (see Table 3), as a robustness study by reanalysis (Clements, 2017). For that we use the BFGS algorithm and the Hessian-based estimator. The choice of these options is guided by their availability in all the packages, except for Matlab-Econometrics. We also use the sample variance as option for initializing the variance process. The other options, such as parameter starting values, are not harmonized because they are impossible to modify in some packages. The results are given in Tables 12 and 13.

For the GARCH model all the packages estimate the same values for the parameters and similar values for the  $t$ -ratios. We obtain similar results for the GJR-GARCH model, with one difference in terms of  $t$ -ratios for the parameter  $\alpha$  estimated by Stata likely due to the different parametrization used by this package. The differences for the  $t$ -ratios are more important for the EGARCH models, even if the values of the parameters are similar. For example, the parameter  $\theta_2$  is not significant for R-rugarch whereas it is significant for the other packages. For the TGARCH and APARCH models the estimations are close for Eviews, Stata and Gretl whereas they are slightly different for Matlab-MFE and R-rugarch, respectively. These results show that the parametrization of the asymmetric models can be important, sometimes dramatically with the rejection of the models.<sup>37</sup>

---

<sup>36</sup>Note that all the GARCH-type models with a Student distribution are rejected. The results are given in the online Appendix.

<sup>37</sup>It would be interesting to determine a benchmark for the parametrization of the asymmetric models

## 6 Forecasting comparison

One way to produce dynamic out-of-sample volatility forecasts, is to output an out-of-sample forecast for a one-day horizon, and then assess model forecast performance by out-of-sample one-step ahead prediction errors. The models are re-estimated every trading day, thus aiming at a closer to real world forecast accuracy. The models are estimated using the first 4-year period, from April 1st 2003 to December 31st 2006 ( $T = 942$  observations), and the out-of-sample forecasts of the conditional variance are calculated for the last year, from January 2, 2007 to December 31, 2007 ( $H = 253$ ).

The one-step ahead forecast for  $T + 1$  is given for each model by

- GARCH model

$$\hat{\sigma}_{T+1}^2 = \hat{\omega} + \hat{\alpha}\hat{\varepsilon}_T^2 + \hat{\beta}\hat{\sigma}_T^2$$

- GJR-GARCH model

$$\hat{\sigma}_{T+1}^2 = \hat{\omega} + (\hat{\alpha} + \hat{\gamma}\hat{I}(\varepsilon_T))\hat{\varepsilon}_T^2 + \hat{\beta}\hat{\sigma}_T^2$$

- EGARCH model

$$\hat{\sigma}_{T+1}^2 = \exp(\hat{\omega})\exp(\hat{g}(z_T))\hat{\sigma}_T^{2\hat{\beta}}$$

- TGARCH model

$$\hat{\sigma}_{T+1} = \hat{\omega} + (\hat{\alpha} + \hat{\gamma}\hat{I}(\varepsilon_T))\hat{\varepsilon}_T + \hat{\beta}\hat{\sigma}_T$$

where  $\hat{\varepsilon}_t^2$  and  $\hat{\sigma}_t^2$  are the fitted values for the squared error and the conditional variance for observation  $t = 1, \dots, T$ , with  $T$  is the last observation in the sample. One-day out-of-sample volatility forecasts  $\hat{\sigma}_{T+1}^2$  are obtained for the forecast horizon  $H$ , and compared to the realized volatility (RV) as a proxy for actual volatility. Following Liu et al. (2015) we use the 5-min calendar-time RV as RV measure, obtained from the Oxford-Man Institute of Quantitative Finance.

---

but this issue is beyond the scope of this paper. Note that we have also estimated the GARCH-type models for the medium and small samples and obtained that the asymmetric models are rejected for both samples. The results are available in the online Appendix.

Note that the one-step ahead forecast for the asymmetric GARCH-type models can be slightly different according to the parametrization of these models (see Section 2).

Most of packages produce out-of-sample conditional variance forecasts automatically, namely EViews (*Static forecasting*)<sup>38</sup>, Gauss-Fanpac (*uforecast*), Matlab-Ek(*forecast*), Ox-G@RCH (*FORECAST*), Stata (*predict*) and R-rugarch (*ugarchforecast*). Nevertheless, we also produce conditional variance forecast from Matlab-MFE and Gretl.<sup>39</sup>

Patton (2011) shows that only two loss functions are robust to noise in the volatility proxy: the mean squared error (MSE) and the loss implied by a Gaussian likelihood (QLIKE). We thus assess volatility forecast performance by the MSE and QLIKE loss functions, defined by

$$MSE = \frac{1}{H} \sum_{t=1}^H (\sigma_t^2 - \widehat{\sigma}_t^2)^2$$

$$QLIKE = \frac{1}{H} \sum_{t=1}^H \left( \frac{\sigma_t^2}{\widehat{\sigma}_t^2} - \ln \left( \frac{\sigma_t^2}{\widehat{\sigma}_t^2} \right) - 1 \right)$$

where  $\sigma_t^2$  is 5-min calendar-time RV, and  $\widehat{\sigma}_t^2$  is the forecast volatility.

However, based only on the loss function criteria, it is difficult to conclude that the forecasting performance of one model estimated from a package dominates that of the same model estimated from another one. To draw such conclusions, we use statistical tests that can provide more reliable information, namely the model confidence set (MCS) procedure proposed by Hansen et al. (2011).

The MCS procedure proposed by Hansen et al. (2011) determine the set,  $\mathcal{M}^*$ , that consists of a subset of equivalent models in terms of superior predictive accuracy (SPA, Hansen, 2005) over the other competing collections of models,  $\mathcal{M}_0$ . The MCS procedure yields a model confidence set,  $\widehat{\mathcal{M}}^*$ , which is a set of models constructed to contain the best models with a given level of confidence. This MCS produces a number of models with the same forecasting performance, and therefore more robust

---

<sup>38</sup>In EViews we have the possibility of *Dynamic* or *Static forecasting*. *Dynamic forecasting* performs a multi-step forecast whereas *Static forecasting* performs a series of one-step ahead forecasts.

<sup>39</sup>The package GIG 2.21 for Gretl does not provide forecasts for the conditional variance for the EGARCH model.

forecasting than that using only one model. The  $t$ -statistic is defined as

$$T_{\max \mathcal{M}} = \max_{i \in \mathcal{M}} t_i \quad \text{with } t_i = \frac{\bar{d}_i}{\sqrt{\widehat{\text{var}}(\bar{d}_i)}} \quad (12)$$

where  $\widehat{\text{var}}(\bar{d}_i)$  denotes the estimate of  $\text{var}(\bar{d}_i)$ ,  $\bar{d}_i = m^{-1} \sum_{j \in \mathcal{M}} \bar{d}_{ij}$ , and  $\bar{d}_{ij} = n^{-1} \sum_{t=1}^n d_{ij,t}$ , with  $d_{ij,t} = L_{i,t} - L_{j,t}$  for all  $i, j \in \mathcal{M}_0$ , and  $L_{i,t}$  is a loss function (here MSE).<sup>40</sup> The  $t$ -statistic is associated with the null hypothesis of equal predictive ability (EPA, Hansen, 2005)  $H_{0,\mathcal{M}} : E(\bar{d}_i) = 0$  for all  $i \in \mathcal{M}$ , where  $\mathcal{M} \subset \mathcal{M}_0$ . The MCS is a sequential testing procedure, eliminating at each step the worst model from  $\mathcal{M}$ , until the null hypothesis of EPA is accepted for all the models. If the null of EPA is rejected for  $\mathcal{M} = \mathcal{M}_0$ , the worst-performing model is excluded from the set  $\mathcal{M}$ .<sup>41</sup> The iterative procedure stops when the null hypothesis of EPA of the models still included in the set cannot be rejected. If  $H_{0,\mathcal{M}}$  is accepted at level  $\alpha$  then the MCS is the set  $\widehat{\mathcal{M}}_{1-\alpha}^*$ .<sup>42</sup>

The MSE and QLIKE values as well the MCS associated with each loss function from the packages are presented in Tables 14-16, with a Normal distribution, a Student distribution and using the same optimization algorithm and variance-covariance estimator for each package, respectively. The initial model space  $\mathcal{M}_0$  consists of 8 (GARCH) to 3 (GJR-GARCH) models estimated from the different packages. We have excluded the APARCH models because the conditions are not satisfied for most of the packages.<sup>43</sup> We set the confidence level for the MCS to  $\alpha = 0.20$  and  $0.50$ .

Table 14 gives the results for the Normal distribution. For the GARCH model, the MSE values are very close for each package, which is confirmed by the MCS as all the packages are in  $\widehat{\mathcal{M}}_{50\%}^*$ , suggesting that all the packages forecast similarly in this

<sup>40</sup> $\bar{d}_{ij}$  measures the relative sample loss between the  $i$ -th and  $j$ -th models, while  $\bar{d}_i$  is the sample loss of the  $i$ -th model relative to the average across the models in  $\mathcal{M}$ .

<sup>41</sup>The choice of the worst model to be eliminated uses the following elimination rule:  $e_{\max \mathcal{M}} = \arg \max_{i \in \mathcal{M}} t_i$ .

<sup>42</sup>The MCS  $p$ -values are calculated using bootstrap implementation with 10,000 resamples (Hansen et al., 2011). The MCS test is carried out using the Ox software package MULCOM of Hansen and Lunde (2007).

<sup>43</sup>The APARCH models have been not rejected for only two packages with a Normal distribution and no package for the Student one.

case. For the QLIKE loss function the values are slightly different, where Matlab-Econometrics displays the lowest value and is also the only package to be in  $\widehat{\mathcal{M}}_{20\%}^*$ . For the asymmetric GARCH-type models the MSE and QLIKE values are close for each package. Further, most of the packages are in  $\widehat{\mathcal{M}}_{20\%}^*$  for the MSE loss function. However, it is not the case for the QLIKE loss function where only one package is in  $\widehat{\mathcal{M}}_{20\%}^*$ , such as Matlab-Econometrics for the GJR-GARCH model, Gauss-Fanpac for the EGARCH model, and Gretl for the TGARCH model.

The results obtained with a Student distribution are similar than those obtained with the Normal distribution, whatever the GARCH-type model (Table 15). The findings show that we do not obtain strong differences in terms of forecast accuracy between the packages, but only slight differences, not statistically different from the MSE loss function, and but sometimes statistically different from the QLIKE loss function. These differences can be due to the various specifications of the GARCH-type models and/or the estimation methods used by each of the packages.

Finally, Table 16 displays the results by using the same optimization algorithm (BFGS), variance-covariance estimator (Hessian-based estimator) and initialization of the variance process to examine whether the different parametrizations of the asymmetric GARCH models can have an effect on the forecast performance. The results show that the MSE values are very close for all the packages and GARCH-type models, with similar forecast accuracy as they are in  $\widehat{\mathcal{M}}_{50\%}^*$ . We obtain similar results from the QLIKE loss function, with most of the package in  $\widehat{\mathcal{M}}_{20\%}^*$ , except for the GJR-GARCH model. Overall, the different parametrizations used in the packages seem to have no impact on the forecast performance.

## 7 Discussion and conclusion

The aim of this paper was to review a number of the most widely used software packages when estimating symmetric and asymmetric GARCH models (using their default option). We considered the numerical consistency of GARCH, GJR-GARCH, TGARCH, EGARCH and APARCH estimations, with Normal and Student distributions, as well as out-of-sample forecasting accuracy, using the model confidence set procedure.

As found by Brooks et al. (2001) results are clearly software-dependent for both asymmetric volatility models, especially for the  $t$ -ratios, which makes the results from two different papers, which probably used different packages, not comparable. Further, the different specifications of the asymmetric GARCH models can have an impact on the significance of the parameters, especially for the EGARCH model.

The out-of-sample forecast results show that the differences in estimating symmetric and asymmetric GARCH models imply slight differences in terms of forecast accuracy, not statistically significant, except in few cases with the QLIKE loss function. The results indicated that the different specifications of the asymmetric GARCH-type models used by the different packages appear to have no significant effect on their forecast accuracy.

The aim of this study was not 'pass judgement' on the packages that are good or not, since defaults may differ between packages in terms of initializations, starting values for parameters, convergence tolerances, optimization routines ... but we can do some recommendations for the users. Our results showed that these choices can be important in terms of parameter values,  $t$ -ratios and, sometimes, forecasting, and thus the results can be not comparable between the packages. However, in most of times, the users do not give information on these options. In few cases the optimization algorithm and the method for computing the standard errors are given, and sometimes the package employed but not its version. The package used for the estimation is also important to know what parametrization is used for the asymmetric GARCH-type models. In this study we showed that the various specifications can lead to different values of parameters for the same model, but we provided the correspondences between the packages, making these estimations comparable. The version of the package is also important because some options can be improved or modified. For example, before the version 9 of Eviews the optimization algorithm in its default option was the BHHH algorithm, and now its is the BFGS algorithm.

Overall, we think that, based on the results obtained in our study, software users should mention the name (and version) of the package used in estimation, what the default estimation methods are, such as method to optimize the likelihood function, how the standard errors are computed, and so on. This information will be also important for reproducibility, replicability and robustness which are fundamental characteristics of

scientific studies (Burman et al., 2010; Clemens, 2017).

Our estimation results on the small, medium and large sample sizes showed that the estimation of GARCH-type models is sensitive to the sample size. As argued by Ng and Lam (2006) and Hwang and Pereira (2006) we recommend using at least 700 observations for estimating the (symmetric and asymmetric) GARCH-type models. Larger samples would be preferable but increase the presence of outliers, jumps or breaks which can lead to misleading effects on the estimation of the volatility (see, e.g., Franses and Ghijssels, 1999; Carnero et al., 2007, 2012, 2016).

There are no consensus on the choices of initial values, optimization routines and standard errors computation methods. However, it would be possible to propose a ‘benchmark’ for these options in order to make the results obtained from different packages comparable. These choices depend on the options proposed by the packages as well as their ‘flexibility’ to change the options. As we do in Section 5.3, a standard approach could be to use the BFGS optimization algorithm, the Hessian-based estimator for the standard errors, and the sample variance as option for initializing the variance process.<sup>44</sup>

As suggested by Brooks et al. (2001) there is a need for some consensus as to what forms of the asymmetric GARCH-type models are preferable, and should be estimable by all packages, as for the GARCH model. Further, a benchmark for each asymmetric GARCH model (GJR-GARCH, TGARCH, APARCH and EGARCH models) is also necessary as the benchmark for the GARCH model proposed by Fiorentini et al. (1996), namely including a specification of the parameter starting values and the initializations for the errors and conditional variances.

---

<sup>44</sup>A comparison from a Monte Carlo experiment by trying various choices of initial values, optimization routines and standard errors computation methods would be informative but it is beyond the scope of this study. Future research is encouraged to address these issues.

## References

- [1] Alexander, C. (2008). *Market Risk Analysis, Vol. II, Practical Financial Econometrics*, Wiley.
- [2] Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177-181.
- [3] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- [4] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, 69, 542-547.
- [5] Bollerslev, T., Mikkelsen, H.O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73, 151-184.
- [6] Bollerslev, T., Wooldridge, J. (1992). Quasi-maximum likelihood estimation inference in dynamic models with time-varying covariance. *Econometric Theory*, 11, 143-72.
- [7] Brooks, C. (1997). GARCH modeling in finance: A review of the software options. *Economic Journal*, 107, 1271-1276.
- [8] Brooks, C., Burke, S.P., Persaud, G. (2001). Benchmarks and the accuracy of GARCH model estimation. *International Journal of Forecasting*, 17, 45-56.
- [9] Brooks, C., Burke, S.P., Persaud, G. (2003). Multivariate GARCH models: Software choice and estimation issues. *Journal of Applied Econometrics*, 18, 725-734.
- [10] Burman, L.E., Reed, W.R., Alm, J. (2010). A call for replication studies. *Public Finance Review*, 38, 787-793.
- [11] Calzolari, G., Fiorentini, G., Panattoni, L. (1993). Alternative estimators of the covariance matrix in GARCH models. Working paper, Institute of Economics, Statistics and Territory Analysis, University of Messina.



- [12] Carnero, M.A., Peña, D., Ruiz, E. (2007). Effects of outliers on the identification and estimation of the GARCH models. *Journal of Time Series Analysis*, 28, 471-497.
- [13] Carnero, M.A., Peña, D., Ruiz, E. (2012). Estimating GARCH volatility in the presence of outliers. *Economics Letters*, 114, 86-90.
- [14] Carnero, M.A., Perez, A. (2018). Leverage effect in energy futures revisited. *Energy Economics*, in press.
- [15] Carnero, M.A., Perez, A., Ruiz, E. (2016). Identification of asymmetric conditional heteroscedasticity in the presence of outliers. *SERIEs*, 7, 179-201.
- [16] Chang, C.L., McAleer, M. (2017) The correct regularity condition and interpretation of asymmetry in EGARCH. *Economics Letters*, 161, 52-55.
- [17] Clemens, M.A. (2017). The meaning of failed replications: A review and proposal. *Journal of Economic Surveys* 31, 326-342.
- [18] Davidson, J. (2004). Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business Economic & Statistics*, 22, 16-19.
- [19] Ding, Z., Granger, C.W.J., Engle, R.F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1, 83-106.
- [20] Efron, B., Hinkley, D. (1978) Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information. *Biometrika*, 65, 457-482.
- [21] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007.
- [22] Engle, R.F., Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5, 1-50.
- [23] Engle, R.F., González-Rivera, G. (1991). Semiparametric ARCH models. *Journal of Business and Economic Statistics*, 9, 345-60.

- [24] Engle, R.F., Lee, G.G.J. (1993). A permanent and transitory component model of stock return volatility. Discussion Paper No: 92-44R, Department of Economics, UCSD.
- [25] Engle, R.F., Rangel, J.G. (2008). The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *The Review of Financial Studies*, 21, 1187-1222.
- [26] Fiorentini, G., Calzolari, G., Panattoni, L. (1996). Analytic derivatives and the computation of GARCH estimates. *Journal of Applied Econometrics*, 11, 399-417.
- [27] Fletcher, R. (2013). *Practical Methods of Optimization*, Wiley.
- [28] Francq, C., Zakoian, J-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10, 605-637.
- [29] Ghalanos, A. (2018). Introduction to the rugarch package (version 1.3-8). URL <https://cran.r-project.org/package=rugarch>.
- [30] Ghalanos, A., Theussl, S. (2015). Rsolnp: General non-linear optimization using augmented Lagrange multiplier method. URL <https://cran.r-project.org/web/packages/Rsolnp/Rsolnp.pdf>.
- [31] Glosten, L.R., Jagannathan, R., Runkle, D.E., (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48, 1779-1801.
- [32] Hamadeh, T, Zakoian, J-M. (2011). Asymptotic properties of LS and QML estimators for a class of nonlinear GARCH processes. *Journal of Statistical Planning and Inference*, 141, 488-507.
- [33] Hansen, P.R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23, 365-380.
- [34] Hansen, P.R., Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1). *Journal of Applied Econometrics*, 20, 873-89.

- [35] Hansen, P.R., Lunde, A., Nason, J.M. (2011). Model confidence sets for forecasting models. *Econometrica*, 79, 453-497.
- [36] He, C., Teräsvirta, T. (1999a). Properties of moments of a family of GARCH processes. *Journal of Econometrics*, 92, 173-192.
- [37] He, C., Teräsvirta, T. (1999b). Statistical properties of the asymmetric power ARCH process. In Engle, R.F., White, H. (eds.), *Cointegration, Causality, and Forecasting*, Chapter 19, pp. 462-474, Oxford University Press.
- [38] He, C., Malmsten, H., Teräsvirta, T. (2008). Higher-order dependence in the general power ARCH process and a special case. In Shalabh, Heumann, C. (eds.), *Recent Advances in Linear Models and Related Areas*, Chapter 12, pp. 231-251, Physica.
- [39] He, C., Teräsvirta, T., Malmsten, H. (2002). Fourth moment structure of a family of first-order exponential GARCH models. *Econometric Theory*, 18, 868-85.
- [40] Heber, G., Lunde, A., Shephard, N., Sheppard, K. (2009). Oxford-Mann Institute's realized library version 0.2. University of Oxford: Oxford-Man Institute.
- [41] Hentschel, L. (1995). All in the family nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, 39, 71-104.
- [42] Higgins, M.L., Bera, A.K. (1992). A class of nonlinear ARCH models. *International Economic Review*, 33, 137-158.
- [43] Hwang, S., Pereira, P.L. (2006). Small sample properties of GARCH estimates and persistence. *The European Journal of Finance*, 12, 473-494.
- [44] Kyriakopoulou, D. (2015). Asymptotic normality of the QML estimator of the EGARCH(1,1) model. SSRN Working Paper.
- [45] Laurent S. (2009). *G@RCH 6, Estimating and Forecasting ARCH Models*. Timberlake Consultants Press, London.
- [46] Laurent S. et Peters J.P. (2002). *G@rch 2.2: An Ox package for estimating and forecasting various ARCH models*. *Journal of Economic Surveys*, 16, 447-485.

- [47] Lee, S-W., Hansen, B.E. (1994). Asymptotic theory for the GARCH(1,1) quasi maximum likelihood estimator. *Econometric Theory*, 10, 29-52.
- [48] Ling, S., McAleer, M. (2002a). Necessary and sufficient moment conditions for the GARCH( $r,s$ ) and Asymmetric Power GARCH( $r,s$ ) models. *Econometric Theory*, 18, 722-729.
- [49] Ling, S., McAleer, M. (2002b). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, 106, 109-117.
- [50] Liu, L.Y., Patton, A.J., Sheppard, K. (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187, 293-311.
- [51] Lombardi, M., Gallo, G. (2001). Analytic Hessian matrices and the computation of FIGARCH estimates. *Statistical Methods & Applications*, 11, 247-264.
- [52] Lucchetti, K., Balietti, S. (2011). The gig package. URL [https://userpage.fu-berlin.de/sfu/gretlpkgpdf/gig\\_v221.pdf](https://userpage.fu-berlin.de/sfu/gretlpkgpdf/gig_v221.pdf)
- [53] Martinet, G.G., McAleer, M. (2018). On the invertibility of EGARCH( $p,q$ ). *Econometric Reviews*, 37, 824-849.
- [54] McAleer, M. (2014). Asymmetry and leverage in conditional volatility models. *Econometrics*, 2, 145-150.
- [55] McCullough, B.D., Renfro, C.G. (1999). Benchmarks and software standards: A case study of GARCH procedures. *Journal of Economic and Social Measurement*, 25, 59-71.
- [56] Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59, 347-370.
- [57] Ng, H.S., Lam, K.P. (2006). How does sample size affect GARCH models? *Advances in Intelligent Systems Research*. URL: <https://www.atlantispress.com/proceedings/jcis2006/139>
- [58] Nocedal, J., Wright, S.J. (1999). *Numerical Optimization*. Springer-Verlag, New-York.

- [59] Pelagatti, M., Lisi, F. (2009). Variance initialisation in GARCH estimation. In Paganoni, A.M., Sangalli, L.M., Secchi, P., Vantini, S. (eds.), S.Co. 2009 Sixth Conference Complex Data Modeling and Computationally Intensive Statistical Methods for Estimation and Prediction, Maggioli Editore, Milan.
- [60] Rabemananjara, R., Zakoian, J.M. (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, 8, 31-49.
- [61] Rodriguez, M.J., Ruiz, E. (2012). Revisiting several popular GARCH models with leverage effect: Differences and similarities. *Journal of Financial Econometrics*, 10, 637-668.
- [62] Sheppard, K. (2009). MFE MATLAB Function Reference Financial Econometrics. URL [https://www.kevinsheppard.com/images/9/95/MFE\\_Toolbox\\_Documentation.pdf](https://www.kevinsheppard.com/images/9/95/MFE_Toolbox_Documentation.pdf)
- [63] Taylor, S.J. (1986a). Forecasting the volatility of currency exchange rates. *International Journal of Forecasting*, 3, 159-170.
- [64] Taylor, S.J. (1986b). *Modelling Financial Time Series*. Wiley.
- [65] Tse, Y.K. (1998). The conditional heteroscedasticity of the yen-dollar exchange rate. *Journal of Applied Econometrics*, 13, 49-55.
- [66] Wintenberger, O. (2013). Continuous invertibility and stable QML estimation of the EGARCH(1,1) model. *Scandinavian Journal of Statistics*, 40, 846-867.
- [67] Xekalaki, E., Degiannakis, S. (2010). *ARCH Models for Financial Applications*. Wiley.
- [68] Ye, Y. (1997). *Interior point algorithms: Theory and analysis*. Wiley-Interscience.
- [69] Zakoian, J.M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18, 931-955.
- [70] Zivot, E. (2009). Practical issues in the analysis of univariate GARCH models. In Andersen, T.G. et al. (eds) *Handbook of Financial Time Series*, Springer: Berlin.

Table 3: GARCH-type models and the econometric softwares.

Softwares	Ox 8.1	Eviews 9	Matlab 13a	Matlab 13a	Gauss 17	R 3.4	Stata 14	Gretl 2018a
Packages	G@RCH 8		MFE	Ek	Fanpac 3.0	rugarch 1.4-0		GIG 2.21
<i>Models</i>								
GARCH	✓	✓	✓	✓	✓	✓	✓	✓
IGARCH	✓		✓		✓	✓		
GJR-GARCH	✓	✓	✓	✓	✓	✓	✓	✓
TGARCH			✓			✓	✓	✓
EGARCH	✓	✓	✓	✓	✓	✓	✓	✓
APARCH	✓	✓	✓			✓	✓	✓
NGARCH			✓			✓	✓	✓
AGARCH			✓					
AVGARCH			✓			✓		✓
CGARCH		✓	✓			✓		
ACGARCH		✓						
fGARCH						✓		
FIGARCH	✓		✓		✓	✓		
HYGARCH	✓							
FIEGARCH	✓							
FIAPARCH	✓							
Spline-GARCH	✓							
<i>Distrib.</i>								
Normal	✓	✓	✓	✓	✓	✓	✓	✓
Student	✓	✓	✓	✓	✓	✓	✓	✓
Skew-Student	✓		✓		✓	✓		✓
GED	✓	✓	✓			✓	✓	✓
Skew-GED						✓		✓
Others*						✓		

*Notes:* IGARCH: Integrated GARCH (Engle and Bollerslev, 1986); NGARCH: Nonlinear GARCH (Higgins et Bera, 1992); AVGARCH: Absolute Value GARCH (Taylor, 1986b); CGARCH: Component GARCH (Engle and Lee, 1993); ACGARCH: Asymmetric Component GARCH (Engle and Lee, 1993); fGARCH: family GARCH (Hentchel, 1995); HYGARCH: HYperbolic GARCH (Davidson, 2004); FIEGARCH: Fractionally Integrated EGARCH (Bollerslev and Mikkelsen, 1996); FIAPARCH: Fractionally Integrated APARCH (Tse, 1998); Spline-GARCH (Engle and Rangel, 2008). \*: the rugarch package propose four other conditional distributions (Generalized Hyperbolic, GH; Normal Inverse Gaussian; GH Skew-Student; Johnson's reparametrized SU).

Table 4: Default estimation for volatility models with Normal distribution on large sample (1195 observations).

Parameters	Packages								
	Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl	
<i>GARCH</i>									
$\omega$	coef.	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
	std.	0.003	0.002	0.003	0.002	0.002	0.003	0.002	0.003
	t-stat	1.60*	2.65	1.81	2.41	1.99	1.54*	2.37	1.60*
$\alpha$	coef.	0.071	0.063	0.061	0.072	0.071	0.071	0.071	0.071
	std.	0.024	0.011	0.020	0.011	0.016	0.024	0.012	0.024
	t-stat	3.02	6.00	3.01	6.26	4.37	2.92	6.21	3.01
$\beta$	coef.	0.917	0.923	0.926	0.917	0.917	0.917	0.917	0.917
	std.	0.027	0.014	0.023	0.015	0.020	0.029	0.015	0.028
	t-stat	33.4	68.6	39.8	62.7	46.6	31.8	62.6	33.3
LL		-1045.0	-1039.2	-1039.3	-1045.1	-1041.5	-1045.0	-1045.0	-1045.0
<i>GJR-GARCH</i>									
$\omega$	coef.	0.005	0.003	0.003	0.005	0.005	0.005	0.005	0.005
	std.	0.003	0.001	0.002	0.002	0.002	0.003	0.002	0.003
	t-stat	1.48*	2.99	1.73	2.44	1.96	1.42*	2.43	1.47*
$\alpha$	coef.	0.046	-0.001 <sup>b</sup>	0.000	0.048	0.046	0.046	0.088	0.046
$\alpha^*$	coef.	0.046	-0.001	0.000	0.048	0.046	0.046	0.047	0.046
	std.	0.021	0.013	0.000	0.014	0.016	0.025	0.014	0.021
	t-stat	2.23	-0.09*	0.84*	3.49	2.83	1.85**	6.20	2.24
$\beta$	coef.	0.921	0.952	0.955	0.921	0.921	0.921	0.921	0.921
	std.	0.028	0.012	0.013	0.015	0.020	0.031	0.014	0.029
	t-stat	32.5	81.4	74.5	63.7	47.0	30.0	64.1	32.4
$\gamma$	coef.	0.043	0.075	0.068	0.039	0.043	0.043	-0.041	0.043
	std.	0.028	0.013	0.020	0.015	0.020	0.030	0.015	0.028
	t-stat	1.52*	5.79	3.46	2.64	2.12	1.44*	-2.76	1.51*
LL		-1042.7	-1031.1	-1030.3	-1043.1	-1039.1	-1042.7	-1042.9	-1042.7
<i>EGARCH</i>									
$\omega$	coef.	-0.939	-0.091	-0.012	-0.013	-0.013	-0.013	-0.012	-0.135
$\omega^*$	coef.	-0.009	-0.011	-0.012	-0.013	-0.013	-0.013	-0.012	-0.012
	std.	0.333	0.021	0.006	0.007	0.008	0.009	0.007	0.034
	t-stat	-2.83	-4.36	-2.10	-1.89	-1.66**	-1.47*	-1.76**	-4.02
$\beta$	coef.	0.990	0.988	0.990	0.987	0.986	0.986	0.987	0.987
	std.	0.006	0.004	0.005	0.006	0.006	0.008	0.006	0.007
	t-stat	176.4	260.7	191.7	174.5	155.7	130.0	174.2	134.0
$\theta_2$	coef.	-0.056	-0.070	-0.061	-0.037	-0.042	-0.042	-0.042	-0.042
	std.	0.019	0.012	0.019	0.011	0.016	0.025	0.012	0.021
	t-stat	-2.95	-5.76	-3.29	-3.29	-2.67	-1.78	-3.57	-1.97
$\theta_1$	coef.	0.115	0.100	0.101	0.154	0.154	0.154	0.154	0.154
	std.	0.029	0.024	0.027	0.022	0.028	0.040	0.022	0.037
	t-stat	3.92	4.24	3.79	6.87	5.42	3.82	6.89	4.15
LL		-1034.8	-1030.8	-1030.7	-1043.6	-1039.5	-1043.1	-1043.1	-1043.1

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.

Table 5: Default estimation for volatility models with Normal distribution on large sample (1195 observations).

Parameters		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl
<i>TGARCH</i>									
$\omega$	coef.	–	0.007	0.007	–	–	0.010	0.008	0.005
	std.	–	0.002	0.004	–	–	0.006	0.003	0.003
	t-stat	–	3.16	1.60*	–	–	1.71**	2.38	1.69**
$\alpha$	coef.	–	0.055	0.024	–	–	0.099	0.105	0.082
$\alpha^*$	coef.	–	0.016	0.024	–	–	0.080	0.060	0.057
	std.	–	0.013	0.015	–	–	0.028	0.015	0.021
	t-stat	–	4.39	1.65**	–	–	3.48	7.22	3.87
$\beta$	coef.	–	0.945	0.944	–	–	0.906	0.923	0.923
	std.	–	0.012	0.018	–	–	0.030	0.013	0.022
	t-stat	–	79.1	51.7	–	–	30.5	71.7	42.3
$\gamma$	coef.	–	0.705	0.064	–	–	0.196	-0.045	0.303
$\gamma^*$	coef.	–	0.078	0.064	–	–	0.039	0.045	0.050
	std.	–	0.218	0.021	–	–	0.167	0.013	0.136
	t-stat	–	3.24	3.09	–	–	1.18*	-3.57	2.22
LL		–	-1031.0	-1032.0	–	–	-1051.3	-1043.5	-1042.9
<i>APARCH</i>									
$\omega$	coef.	– <sup>a</sup>	0.005	0.003	–	–	0.001	0.006	0.005
	std.	–	0.002	0.002	–	–	0.003	0.003	0.003
	t-stat	–	2.33	1.35*	–	–	0.45*	1.93**	1.55*
$\alpha$	coef.	–	0.041	0.023	–	–	0.020	0.074	0.077
	std.	–	0.022	0.034	–	–	0.047	0.017	0.026
	t-stat	–	1.82**	0.82*	–	–	0.43*	4.50	2.95
$\beta$	coef.	–	0.949	0.954	–	–	0.939	0.922	0.922
	std.	–	0.013	0.017	–	–	0.038	0.014	0.025
	t-stat	–	75.2	55.5	–	–	24.7	66.2	36.8
$\gamma$	coef.	–	0.769	-0.740	–	–	0.152	-0.193	0.223
	std.	–	0.449	0.753	–	–	0.091	0.074	0.170
	t-stat	–	1.72**	-0.98*	–	–	1.67**	-2.60	1.31*
$\delta$	coef.	–	1.412	1.777	–	–	3.500	1.612	1.476
	std.	–	0.368	0.558	–	–	2.400	0.497	0.630
	t-stat	–	3.83	3.18	–	–	1.46*	3.25	2.35
LL		–	-1030.4	-1030.2	–	–	-1040.3	-1042.8	-1042.3

Notes:  $\alpha^*$  and  $\gamma^*$  denote the parameters modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively.<sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.



Table 6: Default estimation for volatility models with Normal distribution on small sample (100 observations).

Parameters	Packages								
	Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl	
<i>GARCH</i>									
$\omega$	coef.	– <sup>a</sup>	– <sup>a</sup>	0.020	0.000	0.616	0.000	– <sup>a</sup>	0.000
	std.	–	–	0.018	0.054	0.788	0.001	–	0.062
	t-stat	–	–	1.11*	0.00*	0.78*	0.00*	–	0.00*
$\alpha$	coef.	–	–	0.000	0.017	-0.037 <sup>b</sup>	0.012	–	0.015
	std.	–	–	0.000	0.000	0.072	0.023	–	0.052
	t-stat	–	–	0.34*	0.37*	-0.51*	0.52*	–	0.29*
$\beta$	coef.	–	–	0.961	0.975	0.367	0.981	–	0.978
	std.	–	–	0.020	0.105	0.810	0.024	–	0.124
	t-stat	–	–	48.3	9.26	0.45*	40.2	–	7.90
LL	–	–	-135.1	-138.3	-136.9	-137.4	–	-137.5	
<i>GJR-GARCH</i>									
$\omega$	coef.	– <sup>a</sup>	0.033	0.021	0.000	– <sup>a</sup>	0.014	– <sup>a</sup>	0.000
	std.	–	0.270	0.000	0.040	–	0.008	–	0.037
	t-stat	–	1.23*	1.11*	0.00*	–	1.74**	–	0.00*
$\alpha$	coef.	–	-0.098 <sup>b</sup>	0.000	0.024	–	0.022	–	0.028
$\alpha^*$	coef.	–	-0.098	0.000	0.024	–	0.022	–	0.028
	std.	–	0.014	0.000	0.052	–	0.007	–	0.056
	t-stat	–	-6.89	0.00*	0.47*	–	3.02	–	0.49*
$\beta$	coef.	–	1.049 <sup>b</sup>	0.959	0.978	–	1.000 <sup>b</sup>	–	0.980
	std.	–	0.000	0.021	0.097	–	0.000	–	0.097
	t-stat	–	9537.2	45.3	10.1	–	28384.0	–	10.1
$\gamma$	coef.	–	-0.002	0.000	-0.024	–	-0.079	–	-0.027
	std.	–	0.072	0.000	0.070	–	0.015	–	0.069
	t-stat	–	-0.03*	0.00*	-0.35*	–	-5.33	–	-0.39*
LL	–	-130.4	-133.5	-137.9	–	-135.7	–	-136.9	
<i>EGARCH</i>									
$\omega$	coef.	-0.091	0.172	-0.005	-0.019	– <sup>a</sup>	0.028	-0.051	0.087
$\omega^*$	coef.	-0.051	-0.008	-0.005	-0.019	–	0.028	-0.051	-0.057
	std.	0.125	0.026	0.001	0.021	–	0.000	0.101	0.166
	t-stat	-0.73*	6.63	-6.87	-0.90*	–	365.2	-0.51*	0.52*
$\beta$	coef.	0.442	0.985	0.960	1.000 <sup>b</sup>	–	0.911	0.429	0.432
	std.	0.247	0.001	0.000	0.037	–	0.001	0.650	0.238
	t-stat	1.79**	1662.3	123338.0	27.0	–	948.2	0.66*	1.81**
$\theta_2$	coef.	0.223	-0.017	-0.030	0.157	–	0.165	0.226	0.222
	std.	0.114	0.040	0.002	0.106	–	0.000	0.185	0.113
	t-stat	1.95**	-0.42*	-17.7	1.49*	–	365.8	1.22*	1.97
$\theta_1$	coef.	-0.181	-0.225	-0.353	-0.232	–	-0.452	-0.182	-0.180
	std.	0.181	0.015	0.002	0.154	–	0.001	0.012	0.179
	t-stat	-0.10*	-14.8	-149.7	-1.50*	–	-359.7	-3.57	-1.01*
LL	–	-136.9	-130.3	-127.6	-135.3	–	-130.8	-137.0	-136.9

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.

Table 7: Default estimation for volatility models with Normal distribution on small sample (100 observations).

Parameters		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl
<i>TGARCH</i>									
$\omega$	coef.	–	0.023	0.009	–	–	0.021	– <sup>a</sup>	– <sup>a</sup>
	std.	–	0.035	0.000	–	–	0.041	–	–
	t-stat	–	0.65*	0.57*	–	–	0.52*	–	–
$\alpha$	coef.	–	-0.105	0.000	–	–	0.034	–	–
	std.	–	0.044	0.000	–	–	0.014	–	–
	t-stat	–	-2.38	0.00*	–	–	2.42	–	–
$\alpha^*$	coef.	–	-0.105	0.000	–	–	0.068	–	–
	std.	–	0.044	0.000	–	–	0.014	–	–
	t-stat	–	-2.38	0.00*	–	–	2.42	–	–
$\beta$	coef.	–	1.056 <sup>b</sup>	0.984	–	–	0.948	–	–
	std.	–	0.000	0.017	–	–	0.051	–	–
	t-stat	–	166331.9	59.2	–	–	18.7	–	–
$\gamma$	coef.	–	-0.003	0.000	–	–	-1.000	–	–
$\gamma^*$	coef.	–	0.001	0.000	–	–	-0.068	–	–
	std.	–	0.327	0.000	–	–	0.447	–	–
	t-stat	–	-0.01*	0.01*	–	–	-2.24	–	–
LL		–	-130.6	-133.4	–	–	-138.4	–	–
<i>APARCH</i>									
$\omega$	coef.	– <sup>a</sup>	0.031	0.019	–	–	0.002*	– <sup>a</sup>	– <sup>a</sup>
	std.	–	0.039	0.023	–	–	0.005	–	–
	t-stat	–	0.79*	0.85*	–	–	0.30	–	–
$\alpha$	coef.	–	-0.092 <sup>b</sup>	0.000	–	–	0.000	–	–
	std.	–	0.079	0.000	–	–	0.001	–	–
	t-stat	–	-1.17*	0.24*	–	–	0.00*	–	–
$\beta$	coef.	–	1.048 <sup>b</sup>	0.964	–	–	0.978	–	–
	std.	–	0.000	0.036	–	–	0.010	–	–
	t-stat	–	178703.7	26.7	–	–	102.0	–	–
$\gamma$	coef.	–	0.057	-0.000	–	–	-0.999	–	–
	std.	–	0.461	0.000	–	–	0.000	–	–
	t-stat	–	0.13*	-0.62*	–	–	-2612.8	–	–
$\delta$	coef.	–	0.336	1.790	–	–	3.500	–	–
	std.	–	1.669	1.177	–	–	0.382	–	–
	t-stat	–	0.20*	1.52*	–	–	9.17	–	–
LL		–	-131.0	-1033.2	–	–	-135.0	–	–

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.

Table 8: Default estimation for volatility models with Normal distribution on medium sample (500 observations).

Parameters	Packages								
	Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl	
<i>GARCH</i>									
$\omega$	coef.	0.001	0.002	0.002	0.001	0.001	0.001	0.001	0.001
	std.	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.002
	t-stat	0.34*	2.64	2.40	0.50*	0.41*	0.53*	0.50*	0.32*
$\alpha$	coef.	0.025	0.013	0.004	0.025	0.025	0.025	0.025	0.025
	std.	0.015	0.009	0.021	0.009	0.011	0.005	0.009	0.015
	t-stat	1.66**	1.45*	0.17*	2.89	2.21	5.28	2.89	1.64**
$\beta$	coef.	0.971	0.974	0.985	0.971	0.971	0.971	0.971	0.971
	std.	0.017	0.009	0.021	0.010	0.013	0.005	0.010	0.018
	t-stat	56.5	114.6	48.1	97.0	74.7	186.2	97.0	55.1
LL		-446.9	-438.6	-436.7	-447.0	-443.5	-446.9	-447.0	-447.0
<i>GJR-GARCH</i>									
$\omega$	coef.	— <sup>a</sup>	0.002	0.001	0.001	0.001	0.001	0.001	0.001
	std.	—	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	t-stat	—	3.33	1.57*	0.69*	0.59*	0.73*	0.75*	0.45*
$\alpha$	coef.	—	-0.047 <sup>b</sup>	0.000	0.026	0.026	0.026	-0.000	0.026
$\alpha^*$	coef.	—	-0.047	0.000	0.026	0.026	0.026	0.025	0.026
	std.	—	0.005	0.000	0.008	0.011	0.009	0.012	0.015
	t-stat	—	-9.90	1.17*	3.14	2.40	2.96	-0.03*	1.73**
$\beta$	coef.	—	1.005 <sup>b</sup>	0.982	0.981	0.981	0.982	0.982	0.980
	std.	—	0.000	0.008	0.008	0.013	0.002	0.008	0.017
	t-stat	—	11680.0	128.6	119.1	77.8	636.9	123.6	57.1
$\gamma$	coef.	—	0.053	0.017	-0.024	-0.024	-0.025	0.025	-0.023
	std.	—	0.014	0.015	0.014	0.016	0.018	0.013	0.015
	t-stat	—	3.79	1.13*	-1.72**	-1.53*	-1.36*	1.96	-1.52*
LL		—	-433.5	-433.5	-446.1	-442.6	-446.0	-445.9	-446.1
<i>EGARCH</i>									
$\omega$	coef.	— <sup>a</sup>	0.030	-0.006	-0.007	-0.007	-0.007	-0.006	-0.041
$\omega^*$	coef.	—	-0.006	-0.006	-0.007	-0.007	-0.007	-0.006	-0.006
	std.	—	0.000	0.001	0.007	0.004	0.002	0.004	0.048
	t-stat	—	248.9	-105.4	-1.63*	-1.59*	-4.24	-1.69**	-0.87*
$\beta$	coef.	—	0.997	0.999	0.997	0.997	0.997	0.997	0.997
	std.	—	0.002	0.000	0.003	0.004	0.001	0.003	0.005
	t-stat	—	484.5	24131.4	307.0	269.8	5742.0	302.9	220.8
$\theta_2$	coef.	—	-0.062	-0.053	0.020	0.021	0.021	0.021	0.022
	std.	—	0.006	0.001	0.015	0.019	0.016	0.016	0.025
	t-stat	—	-10.1	-105.9	1.29*	1.10*	1.35*	1.38*	0.88*
$\theta_1$	coef.	—	-0.046	-0.043	0.047	0.045	0.044	0.045	0.044
	std.	—	0.000	0.001	0.022	0.035	0.002	0.022	0.057
	t-stat	—	-1118.6	-35.8	2.17	1.28*	23.5	2.05	0.77*
LL		—	-425.7	-423.7	-446.9	-443.4	-446.8	-446.8	-446.8

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. \* means no convergence. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.

Table 9: Default estimation for volatility models with Normal distribution on medium sample (500 observations).

Parameters		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl
<i>TGARCH</i>									
$\omega$	coef.	–	0.003	0.002	–	–	0.005	0.542	0.001
	std.	–	0.002	0.001	–	–	0.004	0.014	0.002
	t-stat	–	1.50*	1.15*	–	–	1.23*	39.2	0.59*
$\alpha$	coef.	–	0.018	0.000	–	–	0.054	-0.344	0.024
	$\alpha^*$ coef.	–	0.000	0.000	–	–	0.070	-0.635	0.034
	std.	–	0.017	0.000	–	–	0.019	0.009	0.030
$\beta$	t-stat	–	1.04*	1.53*	–	–	2.79	-36.8	0.81*
	coef.	–	0.979	0.991	–	–	0.921	0.712	0.977
	std.	–	0.014	0.007	–	–	0.031	0.008	0.026
$\gamma$	t-stat	–	67.9	142.3	–	–	30.0	94.6	37.3
	coef.	–	0.999	0.009	–	–	-0.294	-0.291	-0.405
	$\gamma^*$ coef.	–	0.036	0.009	–	–	0.070	0.291	-0.019
LL	std.	–	1.303	0.022	–	–	0.305	0.009	0.913
	t-stat	–	0.77*	0.42*	–	–	-0.96*	-32.7	-0.44*
	LL	–	-435.4	-431.4	–	–	-455.2	-760.5	-446.8
<i>APARCH</i>									
$\omega$	coef.	– <sup>a</sup>	0.004	0.002	–	–	– <sup>a</sup>	– <sup>a</sup>	– <sup>a</sup>
	std.	–	0.003	0.001	–	–	–	–	–
	t-stat	–	1.24*	1.55*	–	–	–	–	–
$\beta$	coef.	–	0.984	0.996	–	–	–	–	–
	std.	–	0.014	0.003	–	–	–	–	–
	t-stat	–	68.4	361.5	–	–	–	–	–
$\gamma$	coef.	–	0.903	-0.000	–	–	–	–	–
	std.	–	0.456	0.000	–	–	–	–	–
	t-stat	–	1.98	-0.52*	–	–	–	–	–
$\delta$	coef.	–	0.272	0.626	–	–	–	–	–
	std.	–	0.563	0.399	–	–	–	–	–
	t-stat	–	0.48*	1.57*	–	–	–	–	–
LL	–	-434.9	-431.2	–	–	–	–	–	

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.

Table 10: Default estimation for volatility models with Student distribution on large sample (1195 observations).

Parameters	Packages								
	Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl	
<i>GARCH</i>									
$\omega$	coef.	0.004	0.004	0.004	0.004	– <sup>a</sup>	0.004	0.004	0.004
	std.	0.002	0.002	0.002	0.002	–	0.003	0.002	0.003
	t-stat	1.45*	2.12	1.76**	1.79**	–	1.34*	1.79**	1.44*
$\alpha$	coef.	0.059	0.052	0.050	0.059	–	0.059	0.059	0.059
	std.	0.019	0.013	0.015	0.014	–	0.021	0.014	0.019
	t-stat	3.09	4.12	3.30	4.24	–	2.76	4.24	3.08
$\beta$	coef.	0.932	0.936	0.940	0.932	–	0.932	0.932	0.932
	std.	0.023	0.015	0.018	0.017	–	0.026	0.017	0.023
	t-stat	40.3	62.9	52.4	56.4	–	36.1	56.4	40.2
$\nu$	coef.	11.0	11.2	11.3	11.0	–	11.0	11.0	11.0
	std.	2.89	3.20	3.21	3.17	–	2.70	3.17	2.89
	t-stat	3.80	3.49	3.53	3.45	–	4.06	3.45	3.79
LL		-1035.9	-1030.5	-1030.5	-1036.0	–	-1035.9	-1035.9	-1035.9
<i>GJR-GARCH</i>									
$\omega$	coef.	0.003	0.002	0.002	0.003	0.003	0.003	0.003	0.003
	std.	0.002	0.001	0.001	0.002	0.002	0.003	0.002	0.002
	t-stat	1.33*	2.15	1.79**	1.71**	1.51*	1.26*	1.71**	1.32*
$\alpha$	coef.	0.036	-0.010 <sup>b</sup>	0.000	0.036	0.036	0.036	0.073	0.037
$\alpha^*$	coef.	0.036	-0.010	0.000	0.036	0.036	0.036	0.037	0.037
	std.	0.016	0.013	0.000	0.017	0.016	0.021	0.017	0.016
	t-stat	2.29	-0.09*	0.51*	2.19	2.27**	1.72**	4.33	2.29
$\beta$	coef.	0.936	0.963	0.960	0.936	0.936	0.936	0.936	0.936
	std.	0.023	0.011	0.010	0.016	0.019	0.025	0.016	0.023
	t-stat	41.5	88.2	97.7	57.8	48.8	37.5	57.7	41.3
$\gamma$	coef.	0.038	0.074	0.062	0.038	0.038	0.038	-0.036	0.038
	std.	0.022	0.017	0.015	0.020	0.021	0.025	0.020	0.022
	t-stat	1.71**	4.50	4.01	1.91**	1.83**	1.50*	-1.85**	1.70**
$\nu$	coef.	11.3	11.7	11.8	11.3	11.3	11.3	11.2	11.3
	std.	2.98	3.50	3.29	3.41	3.19	2.79	3.39	2.97
	t-stat	3.80	3.34	3.57	3.30	3.54	4.05	3.31	3.79
LL		-1034.1	-1023.2	-1021.9	-1034.1	-1030.7	-1034.1	-1034.3	-1034.2
<i>EGARCH</i>									
$\omega$	coef.	-1.281	-0.078	-0.010	-0.013	-0.003	-0.013	-0.010	– <sup>a</sup>
$\omega^*$	coef.	-0.010	-0.009	-0.010	-0.013	-0.003	-0.013	-0.010	–
	std.	0.370	0.023	0.005	0.008	0.008	0.009	0.008	–
	t-stat	-3.46	-3.38	-2.12	-1.70**	-0.37*	-1.46*	-1.34*	–
$\beta$	coef.	0.992	0.991	0.992	0.989	0.989	0.989	0.989	–
	std.	0.005	0.004	0.004	0.006	0.006	0.008	0.006	–
	t-stat	212.6	239.9	254.5	160.3	156.3	127.2	160.2	–
$\theta_2$	coef.	-0.052	-0.068	-0.062	-0.038	-0.039	-0.038	-0.038	–
	std.	0.017	0.016	0.016	0.016	0.018	0.022	0.016	–
	t-stat	-3.12	-4.24	-3.83	-2.35	-2.20	-1.72**	-2.35	–
$\theta_1$	coef.	0.103	0.086	0.086	0.137	0.137	0.137	0.137	–
	std.	0.025	0.026	0.021	0.029	0.030	0.039	0.029	–
	t-stat	4.13	3.28	4.11	4.80	4.58	3.48	4.80	–
$\nu$	coef.	13.2	12.9	12.5	11.9	11.9	11.9	11.8	–
	std.	4.06	4.39	3.66	3.75	3.61	3.05	3.74	–
	t-stat	3.25	2.95	3.42	3.16	3.29	3.89	3.16	–
LL		-1028.7	-1024.6	-1023.6	-1035.4	-1031.9	-1035.4	-1035.4	–

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.

Table 11: Default estimation for volatility models with Student distribution on large sample (1195 observations).

Parameters		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl
<i>TGARCH</i>									
$\omega$	coef.	–	0.005	0.006	–	–	0.008	0.007	0.004
	std.	–	0.003	0.004	–	–	0.006	0.004	0.003
	t-stat	–	2.18	1.55*	–	–	1.37*	1.79**	1.52*
$\alpha$	coef.	–	0.047	0.018	–	–	0.085	0.093	0.072
$\alpha^*$	coef.	–	0.010	0.018	–	–	0.067	0.053	0.050
	std.	–	0.014	0.013	–	–	0.027	0.018	0.019
	t-stat	–	3.43	1.39*	–	–	3.14	5.12	3.92
$\beta$	coef.	–	0.954	0.952	–	–	0.920	0.932	0.933
	std.	–	0.013	0.015	–	–	0.029	0.015	0.020
	t-stat	–	72.3	61.8	–	–	31.3	60.2	47.8
$\gamma$	coef.	–	0.783	0.061	–	–	0.211	-0.040	0.303
$\gamma^*$	coef.	–	0.074	0.061	–	–	0.036	0.040	0.044
	std.	–	0.320	0.018	–	–	0.177	0.017	0.132
	t-stat	–	2.45	3.43	–	–	1.20*	-2.30	2.29
$\nu$	coef.	–	13.0	12.9	–	–	10.5	11.8	11.9
	std.	–	4.42	3.80	–	–	2.61	3.71	3.34
	t-stat	–	2.93	3.39	–	–	4.01	3.19	3.57
LL	—	–	-1025.0	-1025.1	–	–	-1040.8	-1035.7	-1035.4
<i>APARCH</i>									
$\omega$	coef.	– <sup>a</sup>	0.003	0.002	–	–	0.001	0.003	0.003
	std.	–	0.019	0.001	–	–	0.003	0.002	0.003
	t-stat	–	1.61*	1.55*	–	–	0.25*	1.24*	1.18*
$\alpha$	coef.	–	0.023	0.014	–	–	0.014	0.051	0.055
	std.	–	2.065	0.011	–	–	0.067	0.023	0.035
	t-stat	–	0.01*	1.29*	–	–	0.21*	2.21	1.57*
$\beta$	coef.	–	0.956	0.960	–	–	0.954	0.937	0.935
	std.	–	0.014	0.010	–	–	0.034	0.016	0.024
	t-stat	–	70.6	94.8	–	–	27.8	57.7	38.3
$\gamma$	coef.	–	0.999	-0.992	–	–	0.171	-0.162	0.182
	std.	–	105.8	0.051	–	–	0.109	0.109	0.158
	t-stat	–	0.01	-19.5	–	–	1.57*	-1.48*	1.15*
$\delta$	coef.	–	1.730	2.098	–	–	3.500	2.090	1.932
	std.	–	0.561	0.567	–	–	4.555	0.805	1.032
	t-stat	–	3.08	3.70	–	–	0.77*	2.60	1.87**
$\nu$	coef.	–	12.2	11.7	–	–	11.0	11.2	11.3
	std.	–	3.95	3.30	–	–	2.92	3.38	3.09
	t-stat	–	3.09	3.53	–	–	3.78	3.31	3.67
LL	–	–	-1023.4	-1021.9	–	–	-1030.8	-1034.4	-1034.1

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied.

Table 12: Estimation for volatility models with Normal distribution on large sample (1195 observations) from BFGS algorithm and Hessian-based estimator.

Parameters		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl
<i>GARCH</i>									
$\omega$	coef.	0.005	0.005	0.005	–	0.005	0.005	0.005	0.005
	std.	0.002	0.002	0.002	–	0.002	0.002	0.002	0.002
	t-stat	1.99	1.98	2.01	–	1.99	2.12	1.98	1.98
$\alpha$	coef.	0.071	0.071	0.072	–	0.071	0.072	0.071	0.071
	std.	0.016	0.016	0.016	–	0.016	0.016	0.016	0.016
	t-stat	4.37	4.36	4.37	–	4.37	4.50	4.36	4.36
$\beta$	coef.	0.917	0.917	0.917	–	0.917	0.917	0.917	0.917
	std.	0.020	0.020	0.020	–	0.020	0.019	0.020	0.020
	t-stat	46.5	46.5	46.5	–	46.6	48.2	46.5	46.5
LL		-1045.0	-1045.0	-1045.1	–	-1041.5	-1045.2	-1045.0	-1045.0
<i>GJR-GARCH</i>									
$\omega$	coef.	0.005	0.005	0.004	–	0.005	– <sup>a</sup>	0.005	0.005
	std.	0.002	0.002	0.002	–	0.002	–	0.002	0.002
	t-stat	1.95**	1.94**	1.89**	–	1.96	–	1.94**	1.94**
$\alpha$	coef.	0.046	0.047	0.043	–	0.046	–	0.088	0.046
$\alpha^*$	coef.	0.046	0.047	0.043	–	0.046	–	0.047	0.046
	std.	0.016	0.016	0.016	–	0.016	–	0.021	0.016
	t-stat	2.82	2.87	2.62	–	2.83	–	4.10	2.83
$\beta$	coef.	0.921	0.921	0.925	–	0.921	–	0.921	0.921
	std.	0.020	0.020	0.020	–	0.020	–	0.020	0.020
	t-stat	46.6	46.6	47.4	–	47.0	–	46.5	46.5
$\gamma$	coef.	0.043	0.041	0.042	–	0.043	–	-0.041	0.043
	std.	0.020	0.020	0.020	–	0.020	–	0.020	0.020
	t-stat	2.12	2.04	2.15	–	2.12	–	-2.03	2.10
LL		-1042.7	-1042.9	-10341.4	–	-1039.1	–	-1042.9	-1042.7
<i>EGARCH</i>									
$\omega$	coef.	-0.939	-0.135	-0.013	–	-0.013	-0.014	-0.012	-0.135
$\omega^*$	coef.	-0.009	-0.012	-0.013	–	-0.013	-0.014	-0.012	-0.012
	std.	0.275	0.027	0.007	–	0.008	0.008	0.008	0.027
	t-stat	-3.42	-5.11	-1.72**	–	-1.66**	-1.77**	-1.64**	-5.11
$\beta$	coef.	0.990	0.987	0.987	–	0.986	0.986	0.986	0.986
	std.	0.005	0.006	0.006	–	0.006	0.006	0.006	0.006
	t-stat	203.4	155.1	160.8	–	155.7	153.2	155.0	154.5
$\theta_2$	coef.	-0.056	-0.042	-0.038	–	-0.042	-0.037	-0.042	-0.042
	std.	0.015	0.016	0.015	–	0.016	0.025	0.016	0.016
	t-stat	-3.84	-2.66	-2.53	–	-2.67	-1.49*	-2.65	-2.66
$\theta_1$	coef.	0.115	0.154	0.146	–	0.154	0.154	0.154	0.154
	std.	0.026	0.029	0.028	–	0.028	0.039	0.029	0.029
	t-stat	4.50	5.39	5.26	–	5.42	3.96	5.39	5.39
LL		-1034.8	-1043.1	-1043.3	–	-1039.5	-1043.7	-1043.1	-1043.1

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied. <sup>c</sup> means failure to improve likelihood (non-zero gradients).

Table 13: Estimation for volatility models with Normal distribution on large sample (1195 observations) from BFGS algorithm and Hessian-based estimator.

Parameters		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Matlab-Ek	Gauss-Fanpac	R-rugarch	Stata	Gretl
<i>TGARCH</i>									
$\omega$	coef.	–	0.008	0.013	–	–	– <sup>a</sup>	0.008	0.005
	std.	–	0.004	0.006	–	–	–	0.004	0.003
	t-stat	–	2.04	2.20	–	–	–	2.03	2.05
$\alpha$	coef.	–	0.082	0.076	–	–	–	0.105	0.082
$\alpha^*$	coef.	–	0.058	0.076	–	–	–	0.060	0.057
	std.	–	0.016	0.017	–	–	–	0.020	0.016
	t-stat	–	5.19	4.50	–	–	–	5.22	5.19
$\beta$	coef.	–	0.923	0.903	–	–	–	0.923	0.923
	std.	–	0.017	0.021	–	–	–	0.017	0.017
	t-stat	–	55.9	44.0	–	–	–	55.5	55.7
$\gamma$	coef.	–	0.298	0.040	–	–	–	-0.045	0.303
$\gamma^*$	coef.	–	0.049	0.040	–	–	–	0.045	0.050
	std.	–	0.103	0.019	–	–	–	0.017	0.107
	t-stat	–	2.88	2.11	–	–	–	-2.61	2.83
LL		–	-1043.3	-1050.4	–	–	–	-1043.5	-1042.9
<i>APARCH</i>									
$\omega$	coef.	– <sup>a</sup>	0.006	–	–	–	0.002	0.006	0.005
	std.	–	0.003	–	–	–	0.001	0.003	0.003
	t-stat	–	1.68**	–	–	–	3.14	1.65**	1.96
$\alpha$	coef.	–	0.075	–	–	–	0.022	0.074	0.077
	std.	–	0.020	–	–	–	0.010	0.020	0.019
	t-stat	–	3.77	–	–	–	2.23	3.65	4.03
$\beta$	coef.	–	0.922	–	–	–	0.935	0.922	0.922
	std.	–	0.019	–	–	–	0.016	0.019	0.018
	t-stat	–	49.5	–	–	–	57.5	49.4	50.3
$\gamma$	coef.	–	0.209	–	–	–	0.134	-0.193	0.223
	std.	–	0.113	–	–	–	0.081	0.107	0.118
	t-stat	–	1.85**	–	–	–	1.66**	-1.81**	1.90**
$\delta$	coef.	–	1.580	–	–	–	3.485	1.613	1.476
	std.	–	0.542	–	–	–	0.593	0.561	0.516
	t-stat	–	2.91	–	–	–	5.88	2.87	2.86
LL		–	-1030.8	–	–	–	-1041.0	-1042.8	-1042.3

Notes:  $\omega^*$  denotes the intercept term modified for comparison. \* and \*\* mean significant at the 5% and 10% level, respectively. <sup>a</sup> means no convergence. <sup>b</sup> means that the positivity or stationarity constraint is not satisfied. <sup>c</sup> means failure to improve likelihood (non-zero gradients).



Table 14: Out-of-sample results of GARCH-type models with a Normal distribution.

Models		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Gauss-Fanpac	R-rugarch	Stata	MatlabEk	Gretl
<i>Panel A: Normal</i>									
GARCH	MSE	0.3969	0.4053	0.4001	0.3958	0.3957	0.3968	<b>0.3946</b>	0.3957
	MCS-MSE	0.6567**	0.6567**	0.6567**	0.6567**	0.6567**	0.6567**	1.000**	0.6567**
	QLIKE	0.2243	0.2407	0.2342	0.2216	0.2214	0.2240	<b>0.2195</b>	0.2212
	MCS-QLIKE	0.0676	0.0496	0.0380	0.0676	0.0676	0.0676	1.000**	0.0676
GJR	MSE	–	–	–	<b>0.3833</b>	–	0.3900	0.3839	–
	MCS-MSE	–	–	–	1.000**	–	0.4209*	0.4209*	–
	QLIKE	–	–	–	0.2108	–	0.2161	<b>0.2107</b>	–
	MCS-QLIKE	–	–	–	0.7555**	–	0.0995	1.000**	–
EGARCH	MSE	0.4012	0.4125	0.3904	<b>0.3812</b>	0.3813	0.3903	0.3822	–
	MCS-MSE	0.4300*	0.4229*	0.4300*	1.000**	0.4300*	0.4300*	0.4300*	–
	QLIKE	0.2336	0.2477	0.2284	<b>0.2130</b>	0.2132	0.2200	0.2139	–
	MCS-QLIKE	0.0583	0.0508	0.0583	1.000**	0.1826	0.0583	0.1826	–
TGARCH	MSE	–	0.4130	0.4160	–	–	0.3957	–	<b>0.3852</b>
	MCS-MSE	–	0.3010*	0.1601	–	–	0.3597*	–	1.000**
	QLIKE	–	0.2506	0.2727	–	–	0.2246	–	<b>0.2180</b>
	MCS-QLIKE	–	0.0509	0.0048	–	–	0.1073	–	1.000**

Notes: MCS-MSE and MCS-MAD denote the  $p$ -value of the statistic  $T_{\max \mathcal{M}}$  of Hansen et al. (2011) based on the MSE and QLIKE loss functions respectively. \* and \*\* mean that the forecasts are in  $\widehat{\mathcal{M}}_{20\%}^*$  and  $\widehat{\mathcal{M}}_{50\%}^*$ , respectively. The lowest MSE and QLIKE values appear in bold.

Table 15: Out-of-sample results of GARCH-type models with a Student distribution.

Models		Packages							
		Ox-G@RCH	Eviews	Matlab-MFE	Gauss-Fanpac	R-rugarch	Stata	MatlabEk	Gretl
GARCH	MSE	0.4022	0.4143	0.4070	0.4016	0.4004	0.4021	<b>0.3996</b>	0.4004
	MCS-MSE	0.7266**	0.6445**	0.6445**	0.7195**	0.7266**	0.7266**	1.000**	0.7266**
	QLIKE	0.2324	0.2549	0.2457	0.2286	0.2282	0.2321	<b>0.2270</b>	0.2280
	MCS-QLIKE	0.1801	0.0730	0.0479	0.1801	0.1801	0.1801	1.000**	0.1801
GJR	MSE	0.3943	–	–	<b>0.3875</b>	–	0.3949	0.3879	0.3877
	MCS-MSE	0.4697*	–	–	1.000**	–	0.4697*	0.4697*	0.4697*
	QLIKE	0.2222	–	–	0.2162	–	0.2228	0.2163	<b>0.2159</b>
	MCS-QLIKE	0.1158	–	–	0.2416*	–	0.1158	0.3754*	1.000**
EGARCH	MSE	0.4074	0.4239	0.3961	<b>0.3860</b>	0.3861	–	0.3863	–
	MCS-MSE	0.3530*	0.3091*	0.3751*	1.000**	0.3751*	–	0.3751*	–
	QLIKE	0.2424	0.2660	0.2370	<b>0.2171</b>	<b>0.2171</b>	–	0.2177	–
	MCS-QLIKE	0.0307	0.0224	0.0307	1.000**	0.3998*	–	0.1891	–
TGARCH	MSE	–	0.4226	–	–	–	0.3985	–	<b>0.3898</b>
	MCS-MSE	–	0.0748	–	–	–	0.4091*	–	1.000**
	QLIKE	–	0.2660	–	–	–	0.2299	–	<b>0.2218</b>
	MCS-QLIKE	–	0.0246	–	–	–	0.0920	–	1.000**

Notes: MCS-MSE and MCS-MAD denote the  $p$ -value of the statistic  $T_{\max \mathcal{M}}$  of Hansen et al. (2011) based on the MSE and QLIKE loss functions respectively. \* and \*\* mean that the forecasts are in  $\widehat{\mathcal{M}}_{20\%}^*$  and  $\widehat{\mathcal{M}}_{50\%}^*$ , respectively. The lowest MSE and QLIKE values appear in bold.

Table 16: Out-of-sample results of GARCH-type models with BFGS algorithm.

Models		<i>Packages</i>							
		Ox-G@RCH	Eviews	Matlab-MFE	Gauss-Fanpac	R-rugarch	Stata	MatlabEk	Gretl
GARCH	MSE	0.3969	0.3968	0.4001	0.3958	–	0.3969	–	<b>0.3957</b>
	MCS-MSE	0.8840**	0.8840**	0.8740**	0.8840**	–	0.8840**	–	1.000**
	QLIKE	0.2249	0.2238	0.2342	0.2216	–	0.2240	–	<b>0.2212</b>
	MCS-QLIKE	0.3254*	0.3375*	0.1076	0.3375*	0.3375*	1.000**		
GJR	MSE	0.3895	0.3900	<b>0.3832</b>	0.3833	0.3833	0.3900	–	0.3834
	MCS-MSE	0.8404**	0.7516**	1.000**	0.9931**	0.9931**	0.7011**	–	0.9924**
	QLIKE	0.2161	0.2159	<b>0.2106</b>	0.2108	0.2108	0.2161	–	<b>0.2106</b>
	MCS-QLIKE	0.1355	0.1355	0.1355	0.1355	0.1355	0.1355	–	1.000**
EGARCH	MSE	0.4013	0.3902	0.3837	<b>0.3812</b>	–	0.3903	–	–
	MCS-MSE	0.3593*	0.3593*	0.3593*	1.000**	–	0.3593*	–	–
	QLIKE	0.2337	0.2197	0.2152	<b>0.2130</b>	–	0.2200	–	–
	MCS-QLIKE	0.1997	0.3297*	0.5920**	1.000**	–	0.3297*	–	–
TGARCH	MSE	–	0.3931	0.3885	–	–	0.3942	–	<b>0.3852</b>
	MCS-MSE	–	0.5739**	0.5739**	–	–	0.5030**	–	1.000**
	QLIKE	–	0.2229	0.2201	–	–	0.2246	–	<b>0.2180</b>
	MCS-QLIKE	–	0.3468*	0.5617**	–	–	0.2419	–	1.000**

Notes: MCS-MSE and MCS-MAD denote the  $p$ -value of the statistic  $T_{\max \mathcal{M}}$  of Hansen et al. (2011) based on the MSE and QLIKE loss functions respectively. \* and \*\* mean that the forecasts are in  $\widehat{\mathcal{M}}_{20\%}^*$  and  $\widehat{\mathcal{M}}_{50\%}^*$ , respectively. The lowest MSE and QLIKE values appear in bold.