Small-sample properties of estimators in an ARCH(1) and GARCH(1,1) model with a generalized error distribution: a robustness study

Ralf Pauly and Peter Kosater

Abstract
GARCH Models have become a workhouse in volatility forecasting of financial and monetary market time series. In this article, we assess the small sample properties in estimation and the performance in volatility forecasting of four competing distribution free methods, including quasi-maximum likelihood and three regression based methods. The study is carried out by means of Monte Carlo simulations. To guarantee an utmost realistic framework, simulated time series are generated from a mixture of two symmetric generalized error distributions. This data generating process allow to reproduce the stylized facts of financial time series, in particular, peakedness and skewness. The results of the study suggest that regression based methods can be an asset in volatility forecasting, since model parameters are subject to structural change over time and the efficiency of the quasi- maximum likelihood method is confined to large sample sizes. Furthermore, the good performance of forecasts based on the historical volatility supports to use the variance targeting method for volatility forecasting.

Keywords
GARCH, volatility forecasting, Monte Carlo simulation, mixture of generalized error distributions, variance targeting.

Contents

Small-sample properties of estimators in an ARCH(1) and GARCH(1,1) model with a generalized error distribution: a robustness study

Ralf Pauly and Peter Kosater

1 Introduction

Empirical densities of financial time series such as log-returns of stock prices frequently deviate significantly from the density of the normal distribution. They exhibit a greater peakedness and heavy tails. Consequently, their kurtosis can considerably exceed the value 3 of the normal distribution. In addition of being leptokurtic, they are often skew.

The original (G)ARCH model conceived by Engle (1982) and Bollerslev (1986),which is based on normally distributed disturbances, is able to generate leptokurtic distributions. *ML*- estimators are consistent and asymptotically efficient. However, empirical results show that residuals from *ML* estimation are still leptokurtic and even skew.Thus, the distribution of the disturbances cannot be presumed to be normal.

Since we do not really know the true distribution of the disturbance, distribution free methods for estimation are crucial interest. The advantage of these methods is their robustness with respect to misspecification.

In their robustness study Fiorentini et al.(1996) have shown that the standard errors of *ML* estimators in an ARCH(1) and a GARCH(1,1) model can be strongly underestimated by covariance estimators such as the Hessian or the outer product matrix when the normal distribution is changed to a $t(5)$-distribution. Whereas the asymptotic robust quasi-maximum likelihood covariance estimator *QML* is quite reliable even in small sample sizes.

Here, we design a robustness study in order to systematically investigate the effect of peakedness and skewness on estimation. Therefore, we replace the normal distribution in Monte Carlo simulations. We use a mixture of two generalized error estimations instead. The performance of the *QML* estimation is compared to that of the *LS* and the *QGLS* estimator. Moreover, we go beyond the mere comparison of single parameter estimates in the ARCH(1) and the GARCH(1,1) model, respectively. Additionally, we particularly focus on the combination of these estimates in the volatility forecast. Volatility forecasts are of crucial interest and financial and monetary analysis. Therefore, reliable estimators for volatility are of crucial interest, too. Furthermore, forecasting based on the historical volatility can be regarded as a competing method (alternative to the (G)ARCH forecasts). Hence, we also include the historical volatility in the comparison study.

The historical volatility is of crucial interest because it allows to use simple and reliable two step procedures, such as the variance-targeting method proposed by Engle and Mezrich (1996), wich may be advantageous in forecasting conditional volatility and Value-at-Risk.

In section 2, we start with the GARCH model and the mixture of two generalized error distributions, the generalized error model. Then in section 3 we present the

QML estimation and a two step procedure which consists of the *LS* and the *QGLS* estimators. In section 4, the results of the Monte Carlo simulation for estimates of the parameters and the volatility are presented. Section 5 concludes the study and gives hints for further research.

2 The GARCH model and the generalized error model

The representation of the GARCH(p,q) model follows Fiorentini et al. (1996) and Greene (2003):

$$(2.1) \qquad y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t$$

$$(2.2) \qquad \varepsilon_t = \sqrt{h_t}v_t$$

$$(2.3) \qquad h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \delta_i h_{t-j}$$

$$(2.4) \qquad v_t \sim GEM(0,1;\gamma,\mu;g)$$

where y_t denotes the endogenous variable, x_t is a $k \times 1$ vector of explanatory variables and β is a $k \times 1$ vector of unknown coefficients. The $\varepsilon_t's$ are innovations and depend on the disturbance v_t and the conditional variance $Var[\varepsilon_t \mid \psi_{t-1}] = h_t$, conditioned on all information through time $t-1$, denoted by ψ_{t-1}. The distribution of the $v_t's$ is determined by a generalized error model which is a mixture of two symmetric generalized error distributions. Büning (1991) proposed a mixture of two normal densities to study the robustness of tests, see also Hamilton (1994) pp.685-689. Here the density of the disturbance v_t^* is

$$(2.5) \qquad f_{v*}(x) = (1-g)f_y(x) + gf_z(x) \ , \ x \in \mathbb{R}$$

where f_y and f_z are densities of the symmetric general error distribution with mean μ and unit variance,

$$(2.6) \qquad f_y(x) = \alpha \exp\left[-\frac{1}{2}|(x-\mu_y)/\lambda|^\gamma\right] \text{ and } f_z(x) = \alpha \exp\left[-\frac{1}{2}|(x-\mu_z)/\lambda|^\gamma\right]$$

with $\alpha = \gamma/[\lambda 2^{(\gamma+1)/\gamma}\Gamma(1/\gamma)]$ and $\lambda = \Gamma^{1/2}(1/\gamma)/[2^{1/\gamma}\Gamma^{1/2}(3/\gamma)]$

With $\mu_y = \mu_z = 0$ we have Nelson's generalized error distribution, normalized to have zero mean and unit variance, compare Nelson (1991). A more general version is discussed in Johnson et al. (1980) for applications to Monte Carlo studies. In order to normalize v_t^* we set $\mu_z = \mu$ and $\mu_y = -[g/(1-g)]\mu$ and we divide v_t^* by its standard deviation,

(2.7)
$$v_t = \left(1 + \frac{g\mu^2}{1-g}\right)^{-\frac{1}{2}} v_t^*$$

with $\mathrm{E}[v_t] = 0$ and $\mathrm{E}[v_t^2] = 1$. With $\mathrm{E}[v_t] = 0$ equation (2.2) yields $\mathrm{E}[\varepsilon_t] = 0$. The assumption that v_t has unit variance is not a restriction. The scaling implied by any other variance would change the parameters in (2.3).

With $\mu_z = 0$ and $\gamma = 2$ the disturbances v_t have a normal distribution and the $\varepsilon_t's$ have a conditional distribution, $\varepsilon_t|\psi_{t-1} \sim N(0, h_t)$. If $\gamma < 2$, the density has thicker tails and greater peakedness than the normal. The choice of $\mu_z \neq 0$ and $g, 0 < g < 1$, determines the degree of asymmetry and also of peakedness. In particular, the coefficient of skewness $\eta_3(v)$ is

(2.8)
$$\eta_3(v_t) = \frac{\mu^3\left(g - \frac{g^3}{(1-g)^2}\right)}{\left[1 + \frac{g\mu^2}{(1-g)}\right]^{3/2}}$$

and the coefficient to kurtosis $\eta_4(v_t)$ is

(2.9)
$$\eta_4(v_t) = \frac{\frac{\Gamma(1/\gamma)\Gamma(5/\gamma)}{\Gamma^2(3/\gamma)} + \frac{6g\mu^2}{(1-g)} + \frac{(g-3g^2+3g^3)\mu^4}{(1-g)^3}}{\left[1 + \frac{g\mu^2}{(1-g)}\right]^2}$$

To ensure positive values for the conditional variance $Var[\varepsilon_t \mid \psi_{t-1}] = h_t$ in (2.3) certain parameter restrictions have to be required. In particular, we assume for the GARCH(1,1) process that the parameters fulfill the conditions $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\delta_1 \geq 0$ and $0 < 1 - \alpha_1 + \delta_1 < 1$. Under these conditions it follows from (2.2) and (2.3) that $\mathrm{E}[\varepsilon_t^2] = \frac{\alpha_0}{1-\alpha_1-\delta_1}$. If the further condition $0 < \eta_4(v)\alpha_1^2 + \delta_1^2 + 2\alpha_1\delta_1 < 1$ is fulfilled, we find that

(2.10)
$$\eta_4[\varepsilon_t] = \eta_4(v_t)\frac{(1+\alpha_1+\delta_1)(1-\alpha_1-\delta_1)}{1-\eta_4(v_t)\alpha_1^2 - \delta_1^2 - 2\alpha_1\delta_1}.$$

Especially from (2.8), (2.9) and (2.10), we will choose values for parameters in the Monte Carlo study in such a way that the deviation from normality will be increased with regard to the peakedness and the skewness.

3 QML estimation and a two step estimation procedure

Let us further follow Engle [1982], Bollerslev [1986] and Fiorentini et al. [1996]. We define
$\mathbf{z}_{t-1} = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p})', \boldsymbol{\omega} = (\alpha_0, \alpha_1, \ldots, \alpha_q, \delta_1, \ldots, \delta_p)'$ the vector of

unknown variance parameters and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\omega}')$ the vector of all unknown parameters. Apart from some constants, the prediction error decomposition form of the log-likelihood function is

$$L_T(\boldsymbol{\theta}) = \sum_{t=1}^{T} l_t(\boldsymbol{\theta}) \text{ with } l_t(\boldsymbol{\theta}) = -\frac{1}{2}\log h_t - \frac{1}{2}\frac{\varepsilon_t^2}{h_t} \tag{3.1}$$

The first derivatives of the log-likelihood terms l_t are

$$\frac{\partial l_t}{\partial \boldsymbol{\omega}} = \frac{1}{2}\frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\left[\frac{\varepsilon_t^2}{h_t} - 1\right] \text{ and } \frac{\partial l_t}{\partial \boldsymbol{\beta}} = \frac{\varepsilon_t x_t}{h_t} + \frac{1}{2}\frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\beta}}\left[\frac{\varepsilon_t^2}{h_t} - 1\right] \tag{3.2}$$

with a further differentiation we obtain terms to build the Hessian matrix

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{12} & \mathbf{H}_{22} \end{bmatrix} \tag{3.3}$$

with

$$\begin{aligned} \mathbf{H}_{22} &= \sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}'} \\ &= \sum_{t=1}^{T} \left[\frac{\varepsilon_t^2}{h_t} - 1\right] \left[\frac{1}{2}\frac{1}{h_t}\frac{\partial^2 h_t}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'} - \frac{1}{2}\frac{1}{h_t^2}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\frac{\partial h_t}{\partial \boldsymbol{\omega}'}\right] - \frac{1}{2}\frac{1}{h_t^2}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\frac{\partial h_t}{\partial \boldsymbol{\omega}'}\frac{\varepsilon_t^2}{h_t} \end{aligned}$$

$$\begin{aligned} \mathbf{H}_{11} &= \sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \\ &= \sum_{t=1}^{T} \left[-\frac{x_t x_t'}{h_t} - \frac{1}{h_t^2}\varepsilon_t x_t \frac{\partial h_t}{\partial \boldsymbol{\beta}'} + \left[\frac{\varepsilon_t^2}{h_t} - 1\right]\frac{\partial}{\partial \boldsymbol{\beta}'}\left[\frac{1}{2}\frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\beta}}\right]\right. \\ &\quad \left. -\frac{1}{h_t^2}\varepsilon_t \frac{\partial h_t}{\partial \boldsymbol{\beta}} x_t' - \frac{1}{2}\frac{1}{h_t^2}\frac{\partial h_t}{\partial \boldsymbol{\beta}}\frac{\partial h_t}{\partial \boldsymbol{\beta}'}\frac{\varepsilon_t^2}{h_t}\right] \end{aligned}$$

$$\begin{aligned} \mathbf{H}_{12} &= \sum_{t=1}^{T} \frac{\partial^2 l_t}{\partial \boldsymbol{\beta} \partial \boldsymbol{\omega}'} \\ &= \sum_{t=1}^{T} \left[-x_t \frac{\partial h_t}{\partial \boldsymbol{\omega}'}\varepsilon_t \frac{1}{h_t^2} - \frac{1}{2}\frac{\partial h_t}{\partial \boldsymbol{\beta}}\frac{\partial h_t}{\partial \boldsymbol{\omega}'}\frac{1}{h_t^2}\left[\frac{\varepsilon_t^2}{h_t} - 1\right]\right. \\ &\quad \left. +\frac{1}{2}\frac{1}{h_t}\frac{\partial^2 h_t}{\partial \boldsymbol{\beta}\partial \boldsymbol{\omega}'}\left[\frac{\varepsilon_t^2}{h_t} - 1\right] - \frac{1}{2}\frac{\varepsilon_t^2}{h_t}\frac{\partial h_t'}{\partial \boldsymbol{\beta}}\frac{\partial h_t}{\partial \boldsymbol{\omega}'}\frac{1}{h_t^2}\right] \end{aligned}$$

The outer product matrix

$$\mathbf{OP} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}'_{12} & \mathbf{A}_{22} \end{bmatrix} \tag{3.4}$$

can be built from the first derivatives in (3.2):

$$\mathbf{A}_{11} = \sum_{t=1}^{T} \frac{\partial l_t}{\partial \beta} \frac{\partial l_t}{\partial \beta'}, \quad \mathbf{A}_{22} = \sum_{t=1}^{T} \frac{\partial l_t}{\partial \omega} \frac{\partial l_t}{\partial \omega'} \text{ and } \mathbf{A}_{12} = \sum_{t=1}^{T} \frac{\partial l_t}{\partial \beta} \frac{\partial l_t}{\partial \omega'}$$

Using the properties $E[\varepsilon_t \mid \psi_{t-1}] = 0$ and $E[\varepsilon_t^2 \mid \psi_{t-i}] = h_t$, we can construct from the expectation of the negative Hessian matrix, an estimated information matrix, which we call score matrix

$$\mathbf{S}_g = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}'_{12} & \mathbf{S}_{22} \end{bmatrix} \tag{3.5}$$

with the terms

$$\mathbf{S}_{11} = \sum_{t=1}^{T} \left[\frac{\mathbf{x}_t \mathbf{x}'_t}{h_t} + \frac{1}{2} \frac{1}{h_t^2} \frac{\partial h_t}{\partial \beta} \frac{\partial h_t}{\partial \beta'} \right], \quad \mathbf{S}_{22} = \sum_{t=1}^{T} \frac{1}{2} \frac{1}{h_t^2} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \omega'} \text{ and } \mathbf{S}_{12} = \sum_{t=1}^{T} -\frac{1}{2} \frac{\partial h_t}{\partial \beta} \frac{\partial h_t}{\partial \omega'} \frac{1}{h_t^2}$$

In the case of a symmetric distribution of ε_t we can replace the matrix $\mathbf{S}_{12}$ by a matrix of zeros and obtain the matrix $\mathbf{S}$.

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{22} \end{bmatrix} \tag{3.6}$$

With the negativ Hessian matrix $-\mathbf{H}$, we can compute the maximum likelihood estimator $\tilde{\boldsymbol{\theta}}$ by means of a gradient algorithm. The estimation $\tilde{\boldsymbol{\theta}}_k$ obtained in the k-th iteration of the gradient method with $\mathbf{H}$ computed at $\tilde{\boldsymbol{\theta}}_{k-1}$ is

$$\tilde{\boldsymbol{\theta}}_k = \tilde{\boldsymbol{\theta}}_{k-1} - \lambda \mathbf{H}^{-1} \frac{\partial L_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{3.7}$$

where λ is a scalar and the $\partial L_T(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$ is computed at $\tilde{\boldsymbol{\theta}}_{k-1}$. Here, we carry out estimation with the Scoring-Newton procedure proposed by Fiorentini et al. (1996) in Mathematica 5.0. The convergence criterion is the same as in Fiorentini et al. (1996). The parameter λ is determined by the method of squeezing, see Greene 2003, p.942.

For the evaluation of the ML estimator $\tilde{\boldsymbol{\theta}}$ we start with the assumption of normality. Thus, we assume that $\gamma = 2$ and $g = 0$ in the generalized error model $GEM(0, 1; \gamma, g, \mu)$. In this case holds $v_t \sim N(0, 1), \varepsilon_t \mid \psi_{t-1} \sim N(0, h_t)$ and the ML estimator $\tilde{\boldsymbol{\theta}}$ is consistent and asymptotically efficient. The matrices $-\mathbf{H}^{-1}, \mathbf{OP}^{-1}, \mathbf{S}_g^{-1}$ as well as $\mathbf{S}^{-1}$ computed at $\tilde{\boldsymbol{\theta}}$ are appropriate covariance estimators. The behavior of these covariance estimators will be compared to the robust quasi-maximum likelihood covariance estimators $\mathbf{QML} = \mathbf{H}^{-1}\mathbf{OPH}^{-1}$, $\mathbf{BW}_g = \mathbf{S}_g^{-1}\mathbf{OPS}_g^{-1}$ and $\mathbf{BW} = \mathbf{S}^{-1}\mathbf{OPS}^{-1}$ which from the asymptotic point view are still appropriate covariance estimators even if v_t is not normal but symmetric. However for a skew distribution, we can expect a better performance for QML and BW_g than for BW. By means of the generalized error distribution $G(0, 1; \gamma, g, \mu)$, we can systematically analyze how the deviation from normality effects the behavior of the ML estimates $\tilde{\boldsymbol{\theta}}$ by increasing the peakedness and the skewness.

The behavior of the QML estimator is compared to the performance of the LS and the $QGLS$ estimator discussed in Gouriéroux (1997) within a two step procedure, where normality is assumed, compare also Greene (2003).

Before we present the two step procedure, we will relate the $QGLS$ estimator to the ML estimator. Under the assumption of normality the method of scoring yields the block diagonal matrix $\mathbf{S}$ in (3.6). If we replace in (3.7) the Hessian matrix $\mathbf{H}$ by the score matrix $\mathbf{S}$ we obtain the estimation of the full parameter vector $\boldsymbol{\theta}$ in two parts. For $\boldsymbol{\omega}$ we find from (3.2) and (3.7)

$$\text{(3.8)} \qquad \begin{aligned} \tilde{\boldsymbol{\omega}}_k &= \tilde{\boldsymbol{\omega}}_{k-1} + \mathbf{S}_{22}^{-1} \sum_{t=1}^{T} \frac{1}{2} \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\omega}} \left[\frac{\varepsilon_t^2}{h_t} - 1 \right] \\ &= \tilde{\boldsymbol{\omega}}_{k-1} + \left[\sum_{t=1}^{T} \frac{1}{2} \frac{1}{h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\omega}} \frac{\partial h_t}{\partial \boldsymbol{\omega}'} \right]^{-1} \sum_{t=1}^{T} \frac{1}{2} \frac{1}{h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\omega}} \left[\varepsilon_t^2 - h_t \right], \end{aligned}$$

compare Greene (2003), p.242. Gouriéroux (1997) considers an ARCH(p)-model where $\tilde{\boldsymbol{\omega}}_k = \tilde{\boldsymbol{\alpha}}_k$. In this case, we can express h_t in (2.3) as

$$\text{(3.9)} \qquad h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 = \mathbf{z}_{t-1}' \boldsymbol{\alpha}$$

where $\mathbf{z}_{t-1} = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2)'$ and $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_q)'$. As h_t in (3.11) is calculated at $\tilde{\boldsymbol{\alpha}}_{k-1}$ we can replace h_t by $\mathbf{z}_{t-1}'\tilde{\boldsymbol{\alpha}}_{k-1}$ and we find with $\partial h_t / \partial \boldsymbol{\alpha} = \mathbf{z}_{t-1}$ the score estimator $\tilde{\boldsymbol{\alpha}}_s$ in form of

$$\text{(3.10)} \qquad \tilde{\boldsymbol{\alpha}}_s = \left[\sum_{t=1}^{T} \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \frac{1}{h_t^2} \right]^{-1} \sum_{t=1}^{T} \mathbf{z}_{t-1} \varepsilon_t^2 \frac{1}{h_t^2}$$

with the score matrix as estimated covariance matrix of $\tilde{\boldsymbol{\alpha}}_s$

$$\text{(3.11)} \qquad \widehat{Var}[\tilde{\boldsymbol{\alpha}}_s] = \mathbf{S}_{22}^{-1} = 2 \left[\sum_{t=1}^{T} \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \frac{1}{h_t^2} \right]^{-1}$$

The reformulation of the ARCH(p)-model in (2.2) and (2.3) as an AR(p)model for the squared innovations

$$\varepsilon_t^2 = h_t + \omega_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \omega_t = \mathbf{z}_{t-1}' \boldsymbol{\alpha} + \omega_t \tag{3.12}$$

with uncorrelated disturbances $\omega_t = h_t(v_t^2 - 1)$ having $\mathrm{E}[\omega_t] = 0$ and conditional variance $\mathrm{E}[\omega_t^2 \mid \psi_{t-1}] = h_t^2(\eta_4(v) - 1)$ shows that in the case of (3.12) the *QGLS* estimator $\hat{\hat{\boldsymbol{\alpha}}}$ is identical with the scoring estimator $\hat{\boldsymbol{\alpha}}_s$ in (3.10). The covariance estimation of the *QGLS* estimator $\hat{\hat{\boldsymbol{\alpha}}}$ is

$$\widehat{Var}[\hat{\hat{\boldsymbol{\alpha}}}] = (\eta_4(v) - 1) \left[\sum_{t=1}^{T} \mathbf{z}_{t-1} \mathbf{z}_{t-1}' \frac{1}{h_t^2} \right]^{-1} \tag{3.13}$$

with an appropriate estimation of the kurtosis $\eta_4(v)$ of v. In the case of normality $\eta_4(v) = 3$, and the *QGLS* estimator $\hat{\hat{\boldsymbol{\alpha}}}$ is asymptotically efficient, too. In small sample sizes, it is of interest whether $\hat{\hat{\boldsymbol{\alpha}}}$ is more efficient than the *QML* estimator $\tilde{\boldsymbol{\alpha}}$ and whether the covariance estimator of $\hat{\hat{\boldsymbol{\alpha}}}$ in (3.13) is more reliable than the robust covariance estimators of $\tilde{\boldsymbol{\alpha}}$.

The *QGLS* estimator $\hat{\hat{\boldsymbol{\alpha}}}$ appears in the second step of the two step estimation procedure. In the first step, the consistent *LS* estimator $\hat{\boldsymbol{\beta}}$ results form the regression of y_t on x_t in (2.1) and the unobservable variable ε_t in (3.12) can be replaced by the *LS*-residuals $\hat{\varepsilon}_t = y_t - x_t'\hat{\boldsymbol{\beta}}$. The regression of $\hat{\varepsilon}_t^2$ on 1, $\hat{\varepsilon}_{t-1}^2, \ldots, \hat{\varepsilon}_{t-q}^2$ yields the *LS*-estimator $\hat{\boldsymbol{\alpha}}$ for the coefficient $\boldsymbol{\alpha}$ in (3.12). In the second step, the *LS* estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\alpha}}$ can be improved by applying the quasi-generalized least squares to the regression y_t on x_t using the estimated conditional variance $\hat{\mathrm{E}}[\varepsilon_t^2 \mid \psi_{t-1}] = \hat{h}_t = \hat{\mathbf{z}}_{t-1}\hat{\boldsymbol{\alpha}}$. The *QGLS* estimator is

$$\hat{\hat{\boldsymbol{\beta}}} = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t' \frac{1}{\hat{h}_t} \right]^{-1} \sum_{t=1}^{T} \mathbf{x}_t y_t \frac{1}{\hat{h}_t} \tag{3.14}$$

and an estimator of its covariance matrix is

$$\widehat{\mathrm{Var}}[\hat{\hat{\boldsymbol{\beta}}}] = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t' \frac{1}{\hat{h}_t} \right]^{-1} \tag{3.15}$$

The *QGLS* estimator $\hat{\hat{\boldsymbol{\beta}}}$ is asymptotically less efficient than the *ML* estimator $\tilde{\boldsymbol{\beta}}$.

As pointed out above, the *QGLS* estimator $\hat{\hat{\boldsymbol{\alpha}}}$ is identical with the score estimator $\tilde{\boldsymbol{\alpha}}_s$ in (3.10). There as well as in the covariance estimator (3.11), we replace $\mathbf{z}_{t-1}$ by $\tilde{\mathbf{z}}_{t-1} = (1, \tilde{\varepsilon}_{t-1}^2)', \varepsilon_t$ by $\tilde{\varepsilon}_t = y_t - \mathbf{x}_t'\tilde{\boldsymbol{\beta}}$ and h_t by $\tilde{h}_t = \tilde{\mathbf{z}}_{t-1}'\tilde{\boldsymbol{\alpha}}$. In the covariance estimator in (3.13) we replace $\mathbf{z}_{t-1}$ by $\hat{\hat{\mathbf{z}}}_{t-1} = (1, \hat{\hat{\varepsilon}}_{t-1}^2)', \varepsilon_t$ by $\hat{\hat{\varepsilon}}_t = y_t - \mathbf{x}_t'\hat{\hat{\boldsymbol{\beta}}}, h_t$ by $\hat{\hat{h}}_t = \hat{\hat{\mathbf{z}}}_{t-1}'\hat{\hat{\boldsymbol{\alpha}}}$ and we estimate of $\eta_4(v)$ with standardized residuals $\hat{\hat{v}}_t = \hat{\hat{\varepsilon}}_t^4 / \sqrt{\hat{\hat{h}}_t}$ in form of

$$\eta_4(\hat{\vartheta}) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\hat{\vartheta}_t - \bar{\hat{\vartheta}}}{s_{\hat{\vartheta}}}\right)^4$$

where $\bar{\hat{\vartheta}} = \frac{1}{T}\sum_{i=1}^{T}\hat{\vartheta}$ and $s^2_{\hat{\vartheta}} = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\vartheta}_t - \bar{\hat{\vartheta}}\right)^2$.

In the case of a GARCH model the two step procedure has to be modified. Here, the GARCH(1,1) model

$$\varepsilon_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \delta_1 h_{t-1} + \omega_t \tag{3.16}$$

is of interest. In the first step, *LS* estimates for α_0, α_1 and δ_1 result by minimizing

$$Q_1(\alpha_0,\alpha_1,\delta_1) = \sum_{t=1}^{T}\hat{\omega}_t^2 = \sum_{t=1}^{T}(\hat{\varepsilon}_t^2 - \alpha_0 - \alpha_1\hat{\varepsilon}_{t-1}^2 - \delta_1 h_{t-1}(\alpha_0,\alpha_1,\delta_1))^2$$

where $\hat{\varepsilon}_t = y_t - \hat{\beta}$. We compute the estimated conditional variance $\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1\hat{\varepsilon}_{t-1}^2 + \hat{\delta}_1\hat{h}_{t-1}$ with the *LS* estimates $\hat{\alpha}_0, \hat{\alpha}_1$ and $\hat{\delta}_1$. Minimizing

$$Q_2(\alpha_0,\alpha_1,\delta_1) = \sum_{t=1}^{T}\frac{\hat{\omega}_t^2}{\hat{h}_t^2} = \sum_{t=1}^{T}\left(\frac{\hat{\varepsilon}_t^2}{\hat{h}_t} - \alpha_0\frac{1}{\hat{h}_t} - \alpha_1\frac{\hat{\varepsilon}_{t-1}^2}{\hat{h}_t} - \delta_1\frac{h_{t-1}(\alpha_0,\alpha_1,\delta_1)}{\hat{h}_t}\right)^2$$

yields the *QGLS* estimates $\hat{\hat{\alpha}}_0, \hat{\hat{\alpha}}_1$ and $\hat{\hat{\delta}}_1$. In both steps, a Marquardt algorithm based on Box and Jenkins (1976) is applied to find the minimum. A covariance estimation of the *QGLS* estimates $\hat{\hat{\alpha}}_0, \hat{\hat{\alpha}}_1$ and $\hat{\hat{\delta}}_1$ is

$$\widehat{Var}\begin{pmatrix}\hat{\hat{\alpha}}_0\\ \hat{\hat{\alpha}}_1\\ \hat{\hat{\delta}}_1\end{pmatrix} = (\eta(\hat{\vartheta}) - 1)\left[\sum\frac{1}{\hat{\hat{h}}_t^2}\hat{\hat{\mathbf{g}}}_t\hat{\hat{\mathbf{g}}}_t'\right]^{-1} \tag{3.17}$$

where $\hat{\hat{g}}_t' = \frac{\partial h_t}{\partial(\alpha_0,\alpha_1,\delta_1)}|_{\hat{\hat{\alpha}}_0,\hat{\hat{\alpha}}_1\hat{\hat{\delta}}_1}$. From $h_t = \mathbf{z}_{t-1}'\boldsymbol{\omega} = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \delta_1 h_{t-1}$ follows $\hat{\hat{\mathbf{g}}}_t' = \hat{\hat{\mathbf{z}}}_{t-1}' + \hat{\hat{\delta}}_1\hat{\hat{\mathbf{g}}}_{t-1}'$ with $\hat{\hat{\mathbf{z}}}_{t-1} = (1, \hat{\hat{\varepsilon}}_{t-1}^2, \hat{\hat{h}}_{t-1})'$, $\hat{\hat{h}}_{t-1} = \hat{\hat{\mathbf{z}}}_{t-1}'\hat{\hat{\omega}}$ and $\hat{\hat{\mathbf{g}}}_0' = \frac{1}{1-\hat{\hat{\delta}}_1}\hat{\hat{\mathbf{z}}}_0' = \frac{1}{1-\hat{\hat{\delta}}_1}(1, \hat{\varepsilon}_0^2, \hat{\varepsilon}_0^2)$ and $\hat{\varepsilon}_0^2 = \frac{1}{n}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$. For $\delta_1 = 0$ (3.17) contains (3.13) as a special case.

Even under normality the *LS* estimates $\hat{\alpha}_0, \hat{\alpha}_1$ and $\hat{\delta}_1$ are inefficient. However, as the *LS* estimates of $h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \delta h_{t-1}$ in $Q_1(\alpha_0,\alpha_1,\delta_1) = \sum(\hat{\varepsilon}_t^2 - h_t)^2$ are determined by minimizing the quadratic differences between ε_t^2 and the conditional variance $h_{1|t-1} = \mathrm{E}[\varepsilon_t^2 \mid \psi_{t-1}] = h_t$, we may expect that the *LS* forecast $\hat{h}_{1|t-1} = \hat{h}_t = \hat{\alpha}_0 = \hat{\alpha}_1\hat{\varepsilon}_{t-1}^2 + \hat{\delta}_1\hat{h}_{t-1}$ performs better than the *LS* estimates for each parameter α_0, α_1 and δ, separately. Maybe, it even outperforms the *ML* forecast $\tilde{h}_{1|t-1} = \tilde{h}_t = \tilde{\alpha}_0 + \tilde{\alpha}_1\tilde{\varepsilon}_{t-1}^2 + \tilde{\delta}\tilde{h}_{t-1}$ in relevant sample sizes.

The three parameters $\alpha_0, \alpha_1, \delta_1$ combined determine the future conditional volatility $\mathrm{E}[\varepsilon_{t+j}^2 \mid \psi_t] = h_{j|t}$ starting from given information at time t in form of

$$(3.18) \qquad h_{j|t} = \sigma_y^2 + (\alpha_1 + \delta_1)^{j-1}(h_{t+1} - \sigma_y^2)$$

where $\sigma_y^2 = \sigma_\varepsilon^2 = \alpha_0/(1 - \alpha_1 - \delta_1)$ is the unconditional variance and

$$(3.19) \qquad h_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + \delta_1 h_t$$

is the conditional variance, compare Baillie and Bollerslev(1992). In the long run, the future conditional volatility $h_{j|t}$ converges to the unconditional variance σ_y^2, $\lim_{j\to\infty} h_{j|t} = \sigma_y^2$.The conditional variance $h_{j|t}$ exhibits mean reversion with reversion level σ_y^2. If the conditional variance exceeds the long term variance $h_{t+1} - \sigma_y^2 > 0, h_{j|t}$ has a decreasing tendency, otherwise an increasing. Thus, a correct estimation of the difference $h_{t+1} - \sigma_y^2$ is of importance. An unreliable estimation of σ_y^2 may lead to a wrong direction in the forecast of the future conditional volatility $h_{j|t}$, a wrong mean reversion, and therefore to a qualitative error.

If the historical volatility s_y^2 turns out to be a relatively efficient estimator than we can replace α_0 according to $\alpha_0 = \sigma_y^2(1 - \alpha_1 - \delta_1)$ in two step procedures. For estimation, we have to incorporate the estimator s_y^2 instead of σ_y^2. Furthermore, (3.19) transform to

$$(3.20) \qquad h_t - \sigma_y^2 = \alpha_1(\varepsilon_{i-1}^2 - \sigma_y^2) + \delta(h_{i-1} - \sigma_y^2)$$

and we use (3.20) in the *QML* estimation and the two step *LS* and *QGLS* procedure, compare Engle and Merzrich (1996), where they propose this approach for the *QML* estimation as it reduces the number of parameters.

4 Results of the Monte Carlo studies for estimates of the parameters and the volatility

The Monte Carlo studies are designed to examine the effect of peakedness and skewness in the distribution of the disturbances v_t on *QML*, *LS* and *QGLS* in an ARCH(1) and in a GARCH(1,1) model with $y_t = \beta_0 + \varepsilon_t$.

As aforementioned, we do not merely assess the performance of single parameter estimators for the three methods *LS*, *QGLS* and *QML*. We particularly focus on the combination of these parameters in the *j*-step-ahead forecast of the volatility $h_{j|T}$. Here, we confine the analysis to the long run forecast $\lim_{j\to\infty} h_{j|T} = \alpha_0/(1 - \alpha_1 - \delta_1) = \sigma_\varepsilon^2 = \sigma_y^2$. The long run *ML* forecast $\widetilde{\sigma_y^2} = \tilde{\alpha}/(1 - \tilde{\alpha} - \tilde{\delta}_1)$ is compared to the *LS* forecast $\widehat{\sigma_y^2} = \hat{\alpha}/(1 - \hat{\alpha} - \hat{\delta}_1)$, to the *QGLS* forecast $\hat{\hat{\sigma}}_y^2 = \hat{\hat{\alpha}}/(1 - \hat{\hat{\alpha}} - \hat{\hat{\delta}}_1)$ as well as to the sample variance (historical volatility) $s_y^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y})^2$ which is often used in a model-free approach for the calculation of volatilities in financial applications. There, the rate of return of a financial asset is calculated as the sample mean $\bar{y}$ which is equal to the *LS* estimator of β.

In applied financial analysis the simple estimators s_y^2 and $\bar{y}$ are often restricted to a shorter sample size T, $T = 250$, as a change of σ_y^2 over time can not be excluded. Forecasts with complex GARCH models are considered as a possibility to improve the estimates of the volatility, compare Hull (2000), p.242-243 and p.368-381. Here from the asymptotic point of view, a larger sample size may be of interest. The structural change of parameters, however, raises doubts whether complex methods based on asymptotic properties should outperform the simple sample mean and sample variance.

Fiorentini et al. have confined the sample size T for the ARCH(1) model to $T \leq 400$ and for the GARCH(1,1) model to $T =\leq 800$. We augment the sample size to $T \leq 1600$. At that, we take into account that the conditional variance $\mathrm{E}[\varepsilon_t^2 - h_t \mid \psi_{t-1}] = \mathrm{E}[\omega_t^2 \mid \psi_{t-1}] = h_t^2(\eta_4(v) - 1)$ becomes greater by an increase of the kurtosis $\eta_4(v)$. A high conditional variance reduces the reliability of the estimates which we compensate by an increase of the sample size.

With regard to the deviation of normality represented by peakedness and skewness the following questions are raised:

i) How much differ the *LS* estimator $\hat{\beta}$ which is equal to the model-free sample mean $\bar{y}$, the *QGLS* estimator $\hat{\hat{\beta}}$ and the *QML* estimator $\tilde{\beta}$ from each other?

ii) Is the *QGLS* estimator for the parameters α_0, α_1 and δ_1 as robust as the *QML* estimator?

iii) Are the robust covariance estimator *QML* and BW_g closer to the Monte Carlo *MSE* than the estimator *BW*?

iv) To which extent does the approximation of robust covariance estimators to the Monte Carlo *MSE* depend on the sample size ?

v) To which extent does the approximation of the distribution of *QML* parameter estimations to the normal distribution depend on the sample size?

vi) Are the *MSE* of the *LS* estimation $\widehat{\sigma_y^2}$ and of the model-free sample variance (historical volatility HV) s_y^2 smaller than the *MSE* of the *QML* estimation $\widetilde{\sigma_y^2}$?

With regard to the last question, we can expect that for a greater sample size T the historical volatility $s_y^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{\beta})^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 = \overline{\hat{\varepsilon}^2}$ will not greatly differ from the *LS* estimation $\widehat{\sigma_y^2} = \hat{\alpha}_0/(1 - \hat{\alpha}_1 - \hat{\delta}_1)$ as the *LS* estimation of (3.19) yields a sample mean $\overline{\hat{\varepsilon}^2}$wich is approximately $\hat{\alpha}_0/(1 - \hat{\alpha}_1 - \hat{\delta}_1)$. This approximation holds for each of the 5000 replications. Thus, also the difference of the *MSE* of the historical volatility s_y^2 and of the *LS* estimation $\widehat{\sigma_y^2}$ will not be large.

Peakedness and skewness are generated by the disturbance v. The density of $v, v\ GEM(0,1;\gamma,g,\mu)$, is a mixture of two generalized error distributions. Johnson et al.(1980) have developed a random-variate generation algorithm that allows to use the generalized error distribution in Monte Carlo simulation studies.We generate Random variates with densities $f_y(x)$ and $f_z(x)$ in (2.6) as follows:

1. Generate W having a gamma distribution with shape parameter $1 + 1/\gamma$ and scale parameter 1
2. Let $V = W^{1/\gamma}$
3. Generate U having a uniform distribution on $(-1, 1)$
4. Let $Y = [\Gamma(1/\gamma)/\Gamma(3/\gamma)]^{\frac{1}{2}} VU + \mu_y$ and $Z = [\Gamma(1/\gamma)/\Gamma(3/\gamma)]^{\frac{1}{2}} VU + \mu_z$

The random variable Y has density $f_y(x)$ and Z has density $f_z(x)$. A random variate v with the mixture of both densities can be generated by a Bernoulli process.

We start the experiment with the ARCH(1) model. For the coefficients in the vector $(\beta, \alpha_0, \alpha_1)$ we assume: a)$(-0.29, 0.5.0.5)$ and b) $(0.01, 0.009, 0.22)$

For the simulation with the GARCH(1,1) model the assumptions for the coefficient vector $(\beta, \alpha_0, \alpha_1, \delta_1)$ are: a) $-0.29, 0.20, 0.35, 0.45)$ b) $0.01, 0.00015, 0.15, 0.72)$ and c) $(0.0005, 0.000005, 0.085, 0.89)$

The values in the first parameter vector are used in the Monte Carlo study in Fiorentini et al (1996). The values in the second one results from an EViews estimation using monthly data of return on S & P 500 stock index including dividend yield from Pindyck and Rubinfeld (1998). The third one contains values resulting from an EViews estimation using 2873 daily data of return on the Dax index including dividend yield from 1st January 1991 to 3rd June 2002.

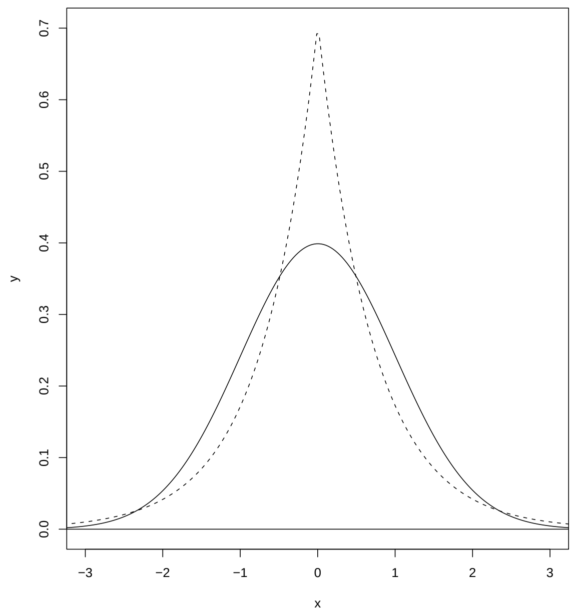


Figure 4.1

For the generalized error distribution $G(0,1;\gamma,g,\mu)$, we first exclude skewness by setting $g = 0$. Starting with the normal distribution, we increase peakedness by reducing γ from 2 to 0.5, we set $\gamma = 2, 1, 0.75, 0.6$ and 0.5. Figure 4.1 shows selected densities for $\gamma = 2$ and $\gamma = 1$

and Figure 4.2 presents the kurtosis $\eta_4(v) = \Gamma/1/\gamma)/\Gamma(5/\gamma)/\Gamma^2(3/\gamma)$ depending on γ.

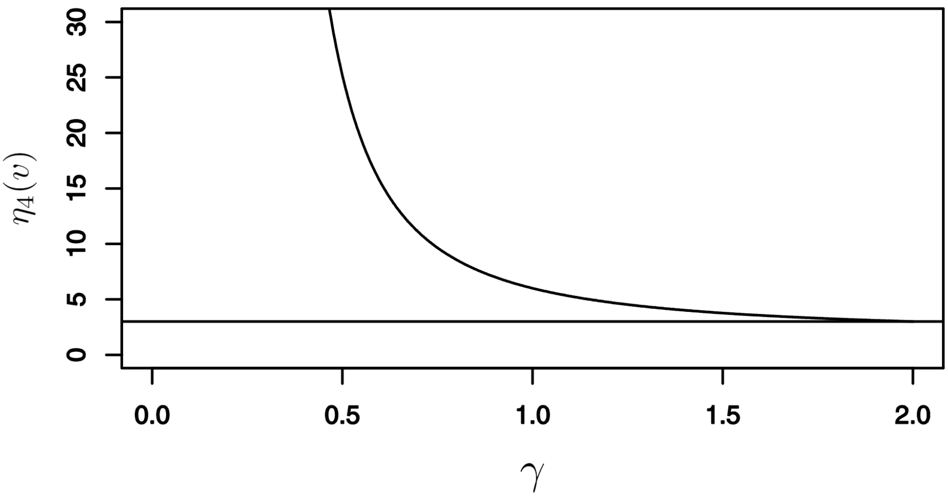


Figure 4.2: Kurtosis $\eta_4(v)$ depending on γ

Table 4.1 informs about the moment of v and ε for selected values γ.

Table 4.1: Moments of v and ε

(γ, g, μ)	$\eta_4(v)$	$\eta_4(\varepsilon)^{1)}$	$\eta_4(\varepsilon)^{2)}$	$\eta_4(\varepsilon)^{4)}$	$\eta_4(\varepsilon)^{5)}$	$\eta_4(\varepsilon)^{6)}$
$(2,0,0)$	3	9	3	9	4	4
$(1,0,0)$	6	—[3)]	8	—	11	22
$(0.75,0,0)$	10	—	17	—	48	—
$(0.6,0,0)$	16	—	60	—	—	—
$(0.5,0,0)$	25	—	—	—	—	—

(γ, g, μ)	$\eta_3(v)$	$\eta_4(v)$	$\eta_4(\varepsilon)^{1)}$	$\eta_4(\varepsilon)^{2)}$	$\eta_4(\varepsilon)^{4)}$	$\eta_4(\varepsilon)^{5)}$	$\eta_4(\varepsilon)^{6)}$
(1,0.025,6)	2	12	—	27	—	—	—
(1,0.025,8)	3	17	—	90	—	—	—

[1)] computed with (-0.29,0.5,0.5)
[2)] computed with (0.01, 0.009, 0.22)
[3)] indicates: does not exist
[4)] computed with (-0.29, 0.20, 0.35, 0.45)
[5)] computed with (0.01, 0.00015, 0.15, 0.72)
[6)] computed with (0.0005, 0.00001, 0.085, 0.89)

To obtain skewness we build with $\gamma = 2$ and $1, g = 0.025$ and 0.0125 and $\mu = 2, 4, 6$ and 8 six combinations for the vector (γ, g, μ).

Table 4.1 shows that with decreasing values of γ the kurtosis $\eta_4(v)$ considerably increases. High values of $\eta_4(v)$ yield a lower reliability of the estimation, compare (3.16), which can be compensated by a higher sample size.

Table A-1 in the appendix shows that the simulation with 5000 replications generates results which are nearly identical with those presented in Fiorentini et al.(1996) for the ARCH model with $(\beta, \alpha_0, \alpha_1) = (-0.29, 0.5, 0.5)$ and with normal disturbances v, i.e $\gamma = 2$. Moreover, Table A-2 in the appendix shows the simulation results produced by a generalized error distribution with $\gamma = 1$, *i.e.* with a kurtosis $\eta_4(v) = 6$. They are similar to those in Fiorentini et al.(1996) with a $t(5)$-distribution which has a kurtosis $\eta_4(v) = 9$. Here, their main simulation result is confirmed. Even in the smaller sample size of $T = 200$, the robust covariance estimators QML and BW perform very well. They only slightly deviate from the MSE whereas the non robust covariance estimators considerably underestimate the variances.

As mentioned in Fiorentini et.al.(1996), in very few cases and only for the shorter time series convergence was not achieved in the QML algorithm, and replications without convergence were then discarded. Even in the case of high peakedness $\eta_4(v) = 15.6$, the percentage does not exceed 1% for $T \geq 800$.

Table A-3, A-6, A-10 and A-13 present results for the ARCH model with $(\beta, \alpha_0, \alpha_1) = (0.01, 0.0009, 0.22)$, i.e. for the S&P data. They show that the approximation of the robust covariance estimator QML to the MSE considerably depends on the degree of peakedness and skewness. A higher peakedness needs a greater sample size to ensure a good approximation. Table A-10 indicates that for α_1 a good approximation is not achieved before the greater sample size of 800, i.e. only for $T \geq 800$. Here, $\gamma = 0.6$ and $\eta_4(v) = 15.58$. This result points out that asymptotic properties only hold at higher sample sizes. Table A-3, A-6, A-10 and A-13 indicate that even for higher sample sizes the BW covariance estimator systematically underestimates the MSE. The generalized version BW_g which considers the skewness does not improve the approximation. Thus, the QML estimator is to be preferable for empirical studies.

Table A-12 shows that the estimator $\eta_4(\tilde{v}) = \frac{1}{T}\sum_{t=1}^{T} [(\tilde{v}_t - \bar{\tilde{v}})/s_{\tilde{v}}]^4$ and $\eta_4(\tilde{u}) = \frac{1}{T}\sum_{t=1}^{T} [(\tilde{u}_t - \bar{\tilde{u}})/s_{\tilde{u}}]^4$ considerably underestimate the kurtosis $\eta_4(v) = 16$ and the kurtosis $\eta_4(u) = 60$, respectively. Thus, with $\eta_4(\hat{v}) = 14$ and $\eta_4(\hat{u}) = 21$ for $T = 1600$ the case of $\gamma = 0.6$ may not be regarded as an unrealistic example. Table A-9 and A-15 are further examples for the underestimation.

As expected, in the case of normality, i.e. $\gamma = 2$, the ML estimator for β performs better than the $QGLS$ estimator, compare Table A-3 and A-4. Both tables illustrate that for α_0 and α_1 the advantage of the ML estimator over the $QGLS$ estimator is only very small. With increasing peakedness and skewness, however, the QML estimator loses his dominance over the $QGLS$ and even over the LS estimator. Already for $\gamma = 0.75$, i.e. for the theoretical kurtosis $\eta_4(v) = 10$ and kurtosis $\eta_4(u) = 17$, respectively, – their estimated values are 9 and 13 for the sample size $T = 1600$ – the $QGLS$ estimator for β is more reliable than the QML estimator, compare Table A-6 and A-7. According to Table A-10 and A-13, the GLS and even the LS estimator perform better than the QML estimator. The simulation results show that with increasing peakedness and skewness, the relative efficiency of the $QGLS$ estimator for β rises.

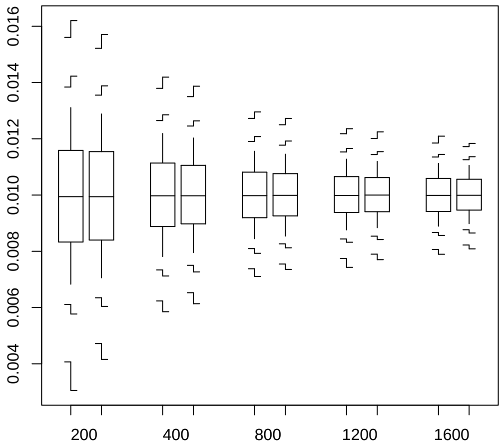


Figure 4.3: Quantile-Boxplots of $\tilde{\beta}$ (first) and $\hat{\hat{\beta}}$ (second) in the ARCH model with S&P coefficients and $\gamma = 0.6$

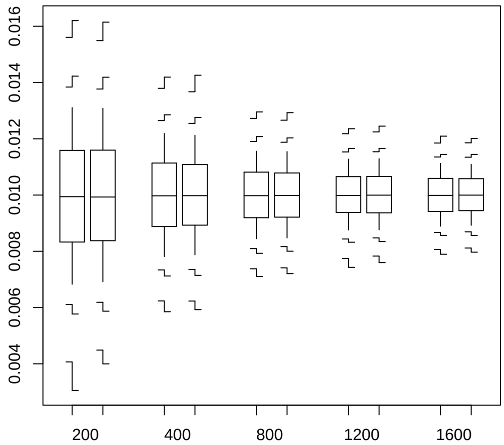


Figure 4.4: Quantile-Boxplots of $\tilde{\beta}$ (first) und $\hat{\beta}$ (second) in the ARCH model with S&P coefficients and $\gamma = 0.6$

For high peakedness, $\gamma = 0.6$, Figure 4.3 compares the distribution of the *QML* estimates $\tilde{\beta}$ with that of the distribution of the *QGLS* estimates $\hat{\hat{\beta}}$.The distributions are characterized by Quantile-Boxplots where the upper and lower quantiles are estimated, see Trenkler (2002). Figure 4.4 shows the Quantile-Boxplots of the *QML* estimates $\tilde{\beta}$ together with those of the *LS* estimates $\hat{\beta}$. Both Figures depict that the *QML* estimation $\tilde{\beta}$ has no advantage neither over the *QGLS* estimation $\hat{\hat{\beta}}$ nor over the *LS* estimates $\hat{\beta} = \bar{y}$. Thus, two step procedures look reasonable for empirical finance market analysis.

With regard to α_0 and α_1 there is a tendency that whith increasing peakedness and skewness the reliability of *QGLS* gains in relation to that of the *QML* estimator. A relative gain can be stated for the *LS* estimator, too. However, for the greater sample size $T = 1600$ the *QML* estimator for α_0 and α_1 is still the most reliable.

As the loss in reliability of the *LS* estimator for α_0 and α_1 decreases with rising peakedness and skewness, we may expect that his relative performance in forecasting the conditional volatility $h_{j|t}$ in (3.18) will increase. The simulation results for $\lim_{j\to\infty} h_{j|t} = \sigma_y^2$ are presented in Table A-5, A-8, A-11 and A-14. There, negative estimates of the variance σ_y^2 have been discarded – the column "$\%\sigma_y$" lists the percentage of positive estimated variances. In addition, the estimation results are corrected for outliers, i.e. 1% of replications due to outliers in the *QML* estimation $\widetilde{\sigma_y^2}$ are eliminated, in the comparison with the other estimates the *QML* estimates is favored. As expected, the *LS* estimation $\hat{\sigma}_y^2 = \hat{\alpha}_0/(1 - \hat{\alpha}_1)$ only slightly deviates from the sample variance $s_y^2 = \frac{1}{T}\sum(y_t - \bar{y})^2$, i.e. the historical volatility HV. In the case of higher peakedness the *LS* estimation of σ_y^2 performs in most cases better than the *QML* estimation especially in smaller sample sizes but, as we can see in Table A-11, also in higher sample sizes, see also Table 4.2.

Table 4.2: The relative efficiency of the *LS* estimate $\hat{\sigma}_y^2$ with respect to the *QML* estimate $\tilde{\sigma}_y^2$ in percent in the ARCH model with *S&P* coefficients

		Peakedness $\eta_4(v)$			
		10		16	
σ^2	T	800	1600	800	1600
LS/QML		83.9	100.0	61.1	82.1

For a comparison of the *QML* estimates $\tilde{\sigma}_y^2 = \tilde{\alpha}_0/(1-\tilde{\alpha}_1)$ with *LS* estimates $\hat{\sigma}_y^2 = \hat{\alpha}_0/(1-\hat{\alpha}_1)$, see Figure 4.5 and for a comparison of the *QML* estimates $\tilde{\sigma}_y^2$ with the historical volatility s_y^2, see Figure 4.6. Figure 4.5 shows that the *QML* estimation $\tilde{\sigma}_y^2$ has no advantage over the *LS* estimation $\hat{\sigma}_y^2$ and Figure 4.6 that the historical volatility s_y^2 performs as well as the *LS* estimation $\hat{\sigma}_y^2$. The historical volatility s_y^2 should be preferred as it yields no negative estimates for σ_y^2.

The simulation results raise doubts whether ARCH models can improve the estimation of σ_y^2. Parameters in a ARCH model may change over time. Therefor we should rather rely on the historical volatility s_y^2 than on the *QML* estimation $\widehat{\sigma_y^2} = \hat{\alpha}_0/(1-\hat{\alpha}_1)$ for empirical analyses. Quite on the contrary, the empirical results suggest to incorporate the historical volatility s_y^2 into the *QML* estimation proposed by Engle and Mezrich (1996) as variance targeting approach. In addition, the results suggest to consider the *LS* estimates in ARCH models as an alternative for forecasting the conditional volatility $h_{j|t}$.

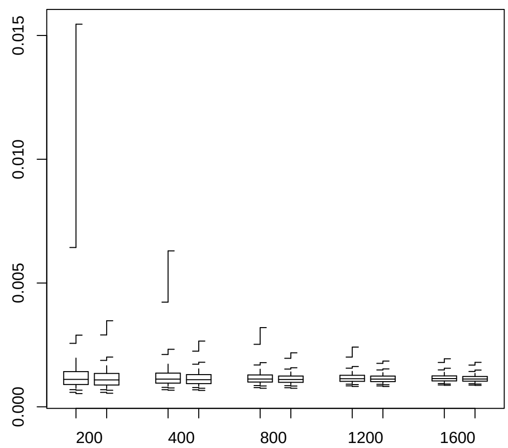


Figure 4.5: Quantile-Boxplots of $\tilde{\sigma}_y^2$ und $\hat{\sigma}_y^2$ in the ARCH model with S&P coefficients and $\gamma = 0.6$

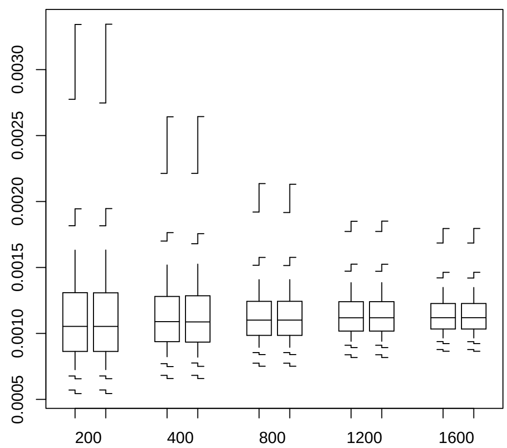


Figure 4.6: Quantile-Boxplots of $\tilde{\sigma}_y^2$ und s_y^2 in the ARCH model with S&P coefficients and $\gamma = 0.6$

Here, the simulation results favor two step procedures for empirical analysis , too. There, we can not trust in assumed distribution for the disturbance v. In a first step, β should be estimated by $\hat{\beta} = \bar{y}$ and σ^2 by the historical volatility s_y^2, i.e., by the sample mean and the sample variance. A good estimate of σ^2 is important for a reliable estimation of the mean reversion effect in forecasting the conditional volatility according to (3.18).

For the GARCH model with $(\beta, \alpha_0, \alpha_1, \delta_1) = (-0.29, 0.20, 0.35, 0.45)$ and with normal disturbances v, i.e. $\gamma = 2$, the simulation results with 5000 replications in Table T-16 in the appendix are in most cases nearly identical with those in Fiorentini et al.(1996), too. Table A-17 in the appendix shows simulation results generated by a generalized error distribution with $\gamma = 1$, i.e. with a kurtosis $\eta_4(v) = 6$. They are similar to those in Fiorentini et al.(1996) with a $t(5)$-distribution. Here, their main simulation results are confirmed, too. Even in the smaller sample size $T = 400$ the robust covariance estimator QML and BW perform very well. They only slightly deviate from the MSE whereas the non robust covariance estimators considerably underestimate the variances. As in the ARCH study, an underestimation can be noticed for the robust covariance estimator BW.

Table A-18, A-19, A-20 and A-21 show simulation results with 10000 replications for the GARCH model with $(\beta, \alpha_0, \alpha_1, \delta_1) = (0.01, 0.00015, 0.15, 0.72)$, i.e. for the S&P data. They confirm the findings in the ARCH study that approximation of the robust covariance estimator QML to the MSE considerably depends on the degree of peakedness and skewness. A higher peakedness requires a greater sample size for a good approximation. Both tables indicates that this may not be achieved before the large sample size of 2000. The asymptotic properties of the QML estimator only holds at a higher sample size. Here, the number of replications in the QML algorithm without convergence is considerably high even at a higher sample size, i.e. 5% und 3% for $T = 800$ and $T = 1200$, respectively, in the case of $\gamma = 0.75$. This technical result also indicates that longer time series are needed.

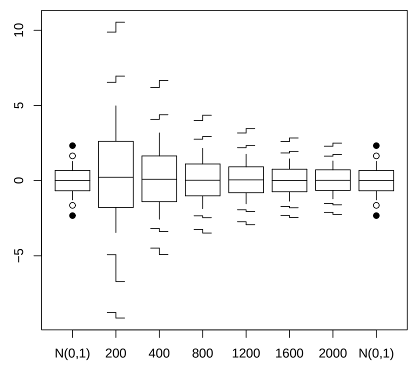


Figure 4.7: Quantile-Boxplots of $(\tilde{\alpha}_1 - \alpha)/\tilde{\sigma}_{\tilde{\alpha}_1}$ in the GARCH model with S&P coefficients and $\gamma = 2$

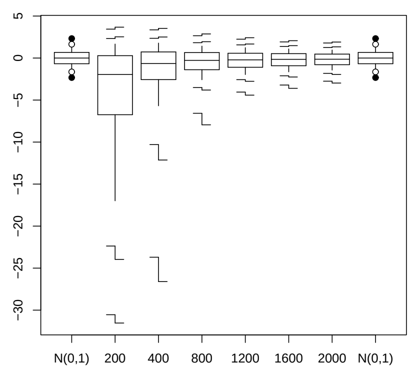


Figure 4.8: Quantile-Boxplots of $(\tilde{\delta}_1 - \delta_1)/\tilde{\sigma}_{\tilde{\delta}_1}$ in the GARCH model with S&P coefficients and $\gamma = 2$

The sequence of the Quantile Boxplots in Figure 4.7 and 4.8 show that for $\gamma = 2$, *i.e.* with normal distributed errors v, the distribution of $\tilde{\alpha}_1$ and of $\hat{\delta}_1$, respectively, approach the normal distribution, but only at the greater sample size of $T = 2000$. An increase of peakedness considerably worsens the approximation, especially with regard to the tails. Figure 4.9 and 4.10 clearly show that the estimated upper and lower quantile estimations substantially deviate from the corresponding quantiles of the normal distribution which for high sample sizes should be close to each other. Here, even for $T = 2000$ the approximation is very poor. Thus, for higher peakedness the validity of tests with regard to α_1 and δ_1 is poor. For skewness, we can observe a similar effect on testing..

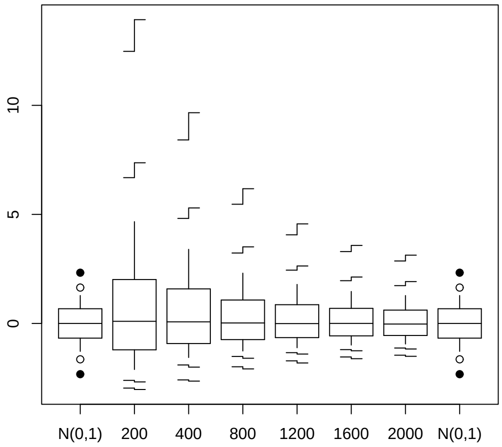


Figure 4.9: Quantile-Boxplots of $(\tilde{\alpha}_1 - \alpha)/\tilde{\sigma}_{\tilde{\alpha}_1}$ in the GARCH model with S&P coefficients and $\gamma = 0.6$

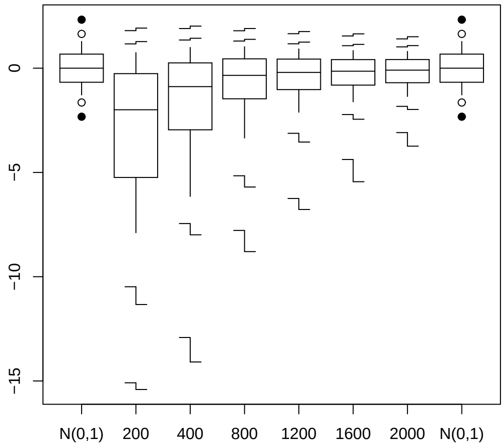


Figure 4.10: Quantile-Boxplots of $(\tilde{\delta}_1 - \delta_1)/\tilde{\sigma}_{\tilde{\delta}_1}$ in the GARCH model with S&P coefficients and $\gamma = 0.6$

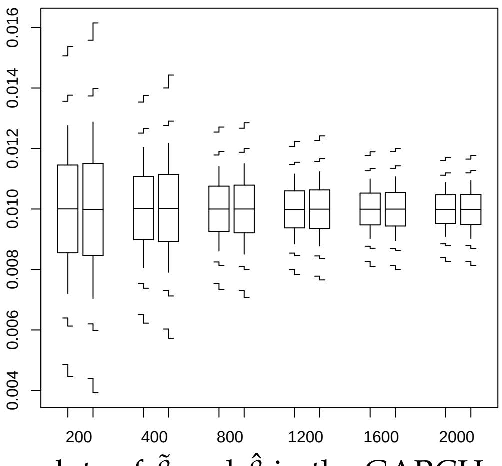


Figure 4.11: Quantile-Boxplots of $\tilde{\beta}$ and $\hat{\beta}$ in the GARCH model with S&P coefficients and $\gamma = 0.6$

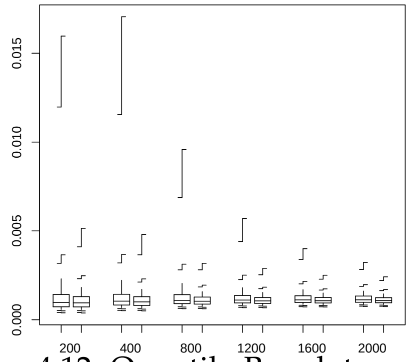


Figure 4.12: Quantile-Boxplots of $\tilde{\sigma}_y^2$ and $\hat{\sigma}_y^2$ in the GARCH model with *S&P* coefficients for $\gamma = 0.6$

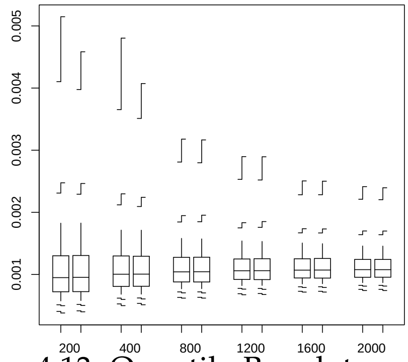


Figure 4.13: Quantile-Boxplots of $\hat{\sigma}_y^2$ and s_y^2 in the GARCH model with *S&P* coefficients for $\gamma = 0.6$

According to the estimation of the parameters $\beta, \alpha_0, \alpha_1$ and δ_1, the GARCH study can not confirm the ARCH results that *QML* estimator loses his dominance over the *QGLS* and the *LS* estimator with increasing peakedness and skewness. Figure 4.11 compares the *QML* estimation $\tilde{\beta}$ with the *LS* estimation $\hat{\beta}$. The sequence of the Boxplots shows that for higher sample sizes the *LS* estimation $\hat{\beta}$ does not substantially deviate from the *QML* estimation $\tilde{\beta}$. Thus, also in the GARCH model a two step procedure looks reasonable for empirical analysis. This holds especially with regard to the estimation of σ_y^2.

Figure 4.12 and 4.13 as well as Table A-22 in the appendix point out that for $\gamma = 0.6$ the *LS* estimation $\widehat{\sigma_y^2}$ and the historical volatility s_y^2 are relatively good estimates for the volatility σ_y^2. Here, the historical volatility HV is clearly the best one.

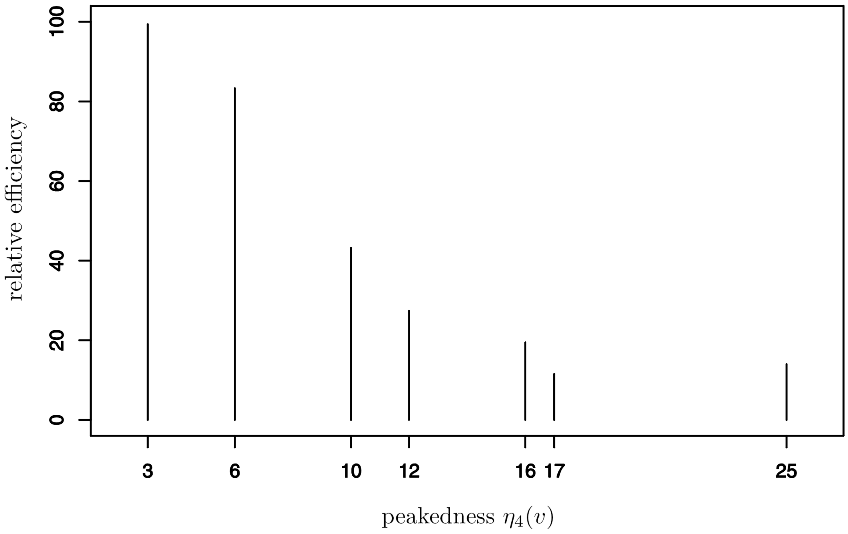
Table 4.3: The relative efficiency of the *LS* estimate $\hat{\sigma}_y^2$ with respect to the *QML* estimate $\tilde{\sigma}_y^2$ in percent depending on peakedness and skewness in the GARCH model with *S&P* coefficients for $T = 800$

	Peakedness $\eta_4(v)$						
σ^2	3	6	10	12[1)]	16	17[2)]	25
LS/*QML*	99.4	83.3	43.2	27.4	19.5	11.5	14.0

1) with skewness $\eta_3(v) = 2$
2) with skewness $\eta_3(v) = 3$

Table A-22 shows next to the *MSE* the *MSE*(cor) corrected for outliers in the *QML* estimation $\tilde{\sigma}_y^2$. This correction allows us to elaborate more clearly the dependence of the *MSE* on the peakedness and skewness as well as on the sample size. Table 4.3 as well as Figure 4.14 show for $T = 800$ the relative efficiency of the *LS* estimate $\hat{\sigma}_y^2$

with regard to the *QML* estimate $\tilde{\sigma}_y^2$ measured by the ratio of the $MSE(cor)$ of $\hat{\sigma}_y^2$ in relation to the $MSE(cor)$ of $\tilde{\sigma}_y^2$. They expose a clear gain in efficiency of the *LS* σ_y^2, and the efficiency augments with an increase of peakedness and skewness.



PSfrag replacements

$\eta_4(v)$

1: with skewness $\eta_3(v) = 2$
2: with skewness $\eta_3(v) = 3$

Figure 4.14: The relative efficiency of $\hat{\sigma}_y^2$ with respect to peakedness in the GARCH model with *S&P* coefficients for $T = 800$

The dependency on the sample size can be seen in Table 4.4. According to the barplot, the gain in efficiency can be still high for larger sample sizes.

Table 4.4: The relative efficiency of the *LS* estimate $\hat{\sigma}_y^2$ and the HV s_y^2 with respect to the *QML* estimate $\tilde{\sigma}_y^2$ in percent depending on peakedness and sample size in the GARCH model with *S&P* coefficients

	Peakedness $\eta_4(v)$								
	3			10			16		
σ^2	800	1600	2000	800	1600	2000	800	1600	2000
LS/QML	99.4	104.6	104.3	43.2	77.4	83.3	19.5	44.7	54.3

Thus, according to the estimation of the variance σ_y^2, the GARCH(1,1) analysis strengthens the result in the ARCH(1) study, that the *LS* estimator $\widehat{\sigma_y^2}$ and the historical volatility HV s_y^2 outperforms than the *QML* estimator $\widetilde{\sigma_y^2}$ and the *QGLS* estimator $\widehat{\widehat{\sigma_y^2}}$.

The doubts increase whether GARCH models can improve the estimation of σ_y^2. Furthermore, as in the ARCH analysis, the GARCH results suggest to consider the *LS* estimates in GARCH models as an alternative for forecasting the conditional volatility $h_{j|t}$, according to (3.22).

With respect to the estimation of the variance σ_y^2, the simulation study for the parameter $(\beta, \alpha_0, \alpha_1, \delta) = (0.0005, 0.00001, 0.085, 0.89)$ of the daily Dax data confirms the relative good performance of the *LS* estimator and of the historical volatility HV, compare Figure 4.15 and Figure 4.16 as well as Table A-24. Both are by far the better estimators for σ_y^2. In addition, the sample size at which a good approximation for the covariance estimator *QML* can be stated is further increased, compare the rate of replications in the *QML* algorithm without convergence in Table A-23 and Table A-25.

The performance of the *QML* estimation $\widetilde{\sigma_y^2}$ is considerably affected by outliers. This outlier effect leads to a relative weak reliability in comparison to the *LS* estimation $\widehat{\sigma_y^2}$ and the HV s_y^2. Especially with regard to the estimation of σ^2, the simulation results in the GARCH model support two step procedures, even more stronger than in the ARCH model.

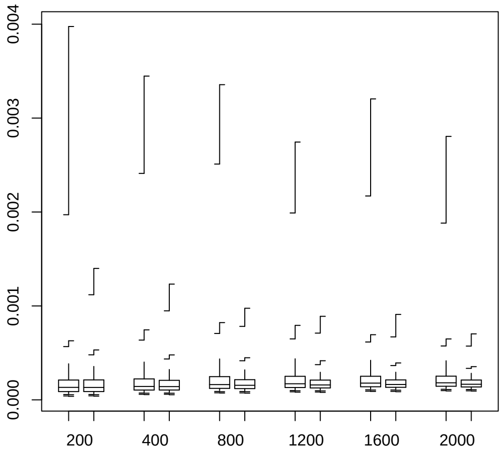


Figure 4.15: Quantile-Boxplots of $\tilde{\sigma}_y^2$ and $\hat{\sigma}_y^2$ in the GARCH model with *DAX* coefficients for $\gamma = 0.6$

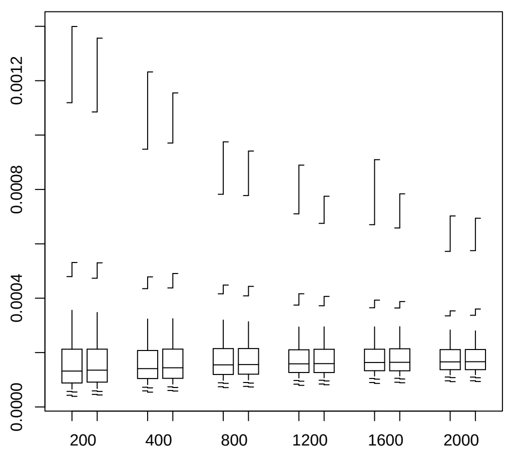


Figure 4.16: Quantile-Boxplots of $\hat{\sigma}_y^2$ and s_y^2 in the GARCH model with *DAX* coefficients for $\gamma = 0.6$

As the HV s^2 is a relative reliable and robust estimator of σ^2, it can be used as a separate input in the *QML*, *QGLS* and the *LS* method. This approach is known as variance targeting, see Engle and Mezrich (1996). The variance targeting reduces the number of estimated parameters and may improve the estimation as well as the forecast.

5 Conclusions

The main results in the simulation study of Fiorentini et al.(1996) for their ARCH and GARCH model are confirmed. The robust covariance estimators QML and BW perform very well. Even in smaller sample sizes, the asymptotic property holds, in the considered ARCH model for $T \geq 200$ and in the considered GARCH for $T \geq 400$.

In the ARCH model with parameters from monthly S&P data, the analysis shows that the approximation of the covariance estimator QML to the MSE considerably depends on the degree of peakedness and skewness. For a higher but not unrealistic degree of peakedness, the asymptotic property holds only for $T \geq 800$. As the covariance estimator BW systematically underestimates the MSE the covariance estimator QML estimator is to be preferable for empirical studies.

With higher peakedness and skewness the QML estimator for the parameter β, α_0 and α_1 loses his advantage over the $QGLS$ and even over the LS estimator. For the variance σ_y^2 the LS and the HV outperform the QML. The results raise doubts whether ARCH models can improve the HV estimation s_y^2. Quite on the contrary, the study suggest to incorporate the relative reliable and robust HV estimation in the QML, $QGLS$ and the LS method.

The GARCH simulation results with parameters from monthly S&P data as well as with daily DAX data indicate that in this important model for empirical financial analysis an even greater sample size is needed for a good approximation of the covariance estimator QML to the MSE, at least $T \geq 2000$.

Even for the large sample size $T = 2000$, the validity of test is poor when peakedness and skewness are high. In these cases, sequences of Boxplots show that upper and lower quantile estimations substantially deviate from the corresponding quantiles of the normal distribution.

Concerning σ_y^2, the GARCH analysis strengthen the ARCH results, that the LS estimator $\widehat{\sigma_y^2}$ and the HV s_y^2 perform better than the QML estimator $\widetilde{\sigma_y^2}$ and the $QGLS$ estimator$\widehat{\widehat{\sigma_y^2}}$.

The GARCH study suggests to analyse the performance of the presented methods in forecasting the conditional volatility in terms of (3.18). The high efficiency gain of the HV s_y^2 recommends to use it for estimation of the variance σ^2 in a first step. Here, it is worthwhile to scrutinize in futher research whether modified two step procedures can improve the parameter estimation as well as conditional volatility forecasts.

6 References

Baillie, R. T. and T. Bollerslev (1992), *Journal of Econometrics*, **52**, 91–113.

Berndt, E.K., B.H. Hall, R.E. Hall and J.A. Hausman (1974), Estimation and inference in nonlinear structural models, *Annals of Economic and Social Measurement*, **3**, 653–665.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, **31**, 307–327.

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

Box, G.E.P. and G.M. Jenkins (1976), *Time series analysis, forecasting and control*, Holden-Day, San Francisco et al.

Büning, H. (1991), *Robuste und adaptive Tests*, Walter de Gruyter, Berlin und New York.

Engle, R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations, *Econometrica*, **50**, 987–1007.

Engle, R.F. and T. Bollerslev (1986), Modelling the persistence of conditional variances, *Econometric Reviews*, **5**,1–50.

Engle, R.F. and J. Mezrich (1996), GARCH for groups, *RISK*, **9**, 36–40.

Fiorentini, G., G. Calzolari and L. Panattoni (1996), Analytic derivatives and the computation of GARCH estimates, *Journal of Applied Econometrics*, **11**, 399–417.

EViews 5, Quantitative Micro Software, Irvine, CA Copyright 1997-2004.

Gouriéroux, Ch. (1997), *ARCH Models and financial applications*, Springer, New York.

Greene, W. H. (2003), *Econometric Analysis*, 5.ed., Pearson Education International, Boston

Hamilton, D.J. (1994), *Time series analysis*, Princeton University Press, Princeton.

Hull, J.C. (2000), *Options, futures & other derivatives*, 4.ed., Prentice-Hall International, London.

Johnson, M.E., L.G. Tietjen and R.J. Beckman (1980), A new family probability distributions with applications to Monte Carlo studies, *Journal of the American Statistical Association*, **75**, 276–279.

McCullough, B.D. and C.G. Renfro (1998), Benchmarks and Software standards: A case study of GARCH procedures, *Journal of Economic and Social Measurement*, **25**, 59–71. IOS Press.

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

Pindyck, R.S. and D.L. Rubinfeld (1998), *Econometric models and economic forecasts*, 4th ed., McGraw-Hill, Boston.

Trenkler, D. (2002), Quantile Boxplots, *Communication in Statistics – Simulation and Computation*, **31**, 1–12.

White, H. (1982), Maximum likelihood estimation of misspecified models, *Econometrica*, **50**, 1–25.

White, H. (1983), Corrigendum, *Econometrica*, **51**, 513.

Wolfram, Stephen (2003), *The Mathematica Book*, 5.ed., Cambridge University Press.

7 Appendix

Table A-1: *ML* estimates in the Fiorentini ARCH model with $\beta = 0.29, \alpha_0 = 0.5, \alpha_1 = 0.5$ and $\gamma = 2$

QML	*T*	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*
$\tilde{\beta}$	100	-0.29018	64.006	60.033	66.777	61.885	62.400	59.400
	200	-0.28928	31.718	30.157	31.943	30.530	30.507	30.066
	400	-0.28946	15.115	15.024	15.514	15.133	15.108	14.972
	800	-0.28996	7.573	7.511	7.630	7.533	7.530	7.506
	1200	-0.29037	5.096	5.010	5.063	5.018	5.015	5.009
	1600	-0.28988	3.846	3.760	3.794	3.767	3.764	3.756
$\tilde{\alpha}_0$	100	0.50517	140.948	128.595	167.205	140.535	137.449	115.698
	200	0.50400	63.074	62.982	72.645	65.745	63.686	59.553
	400	0.50104	31.300	30.952	33.387	31.624	31.089	30.146
	800	0.50045	15.798	15.384	15.924	15.527	15.435	15.240
	1200	0.50074	10.345	10.249	10.531	10.320	10.251	10.153
	1600	0.50103	7.515	7.696	7.840	7.735	7.709	7.654
$\tilde{\alpha}_1$	100	0.47303	448.254	377.418	496.841	427.889	436.281	341.369
	200	0.48433	207.765	191.978	224.950	204.253	202.650	182.671
	400	0.49249	99.987	97.155	106.054	100.202	99.368	94.603
	800	0.49768	49.496	48.921	51.120	49.592	49.426	48.414
	1200	0.49755	33.477	32.606	33.671	32.928	32.819	32.334
	1600	0.49883	24.882	24.511	25.081	24.688	24.647	24.379

MSE and variance estimates multiplied by 10000

Table A-2: *QML* estimates in the Fiorentini ARCH model with $\beta = 0.29, \alpha_0 = 0.5, \alpha_1 = 0.5$ and $\gamma = 1$

QML	*T*	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*
$\tilde{\beta}$	100	-0.29113	68.280	55.886	62.026	58.562	69.230	63.042
	200	-0.28784	35.439	28.611	28.927	29.299	34.176	33.060
	400	-0.28996	16.757	14.389	13.526	14.538	17.074	16.952
	800	-0.28945	8.413	7.197	6.418	7.226	8.611	8.621
	1200	-0.29043	5.639	4.800	4.207	4.818	5.750	5.746
	1600	-0.28990	4.276	3.601	3.118	3.609	4.320	4.324
$\tilde{\alpha}_0$	100	0.49663	255.848	99.210	67.446	109.658	236.949	203.439
	200	0.50187	121.419	49.115	27.566	52.359	118.853	109.722
	400	0.50194	61.417	24.309	11.977	25.116	58.850	57.034
	800	0.50071	29.660	12.054	5.466	12.285	29.717	29.148
	1200	0.50049	19.442	8.007	3.514	8.101	19.683	19.497
	1600	0.50061	15.451	6.005	2.577	6.059	14.884	14.783
$\tilde{\alpha}_1$	100	0.48856	1149.967	449.815	339.591	542.336	1254.569	883.868
	200	0.48492	581.226	222.339	139.245	249.980	574.500	472.095
	400	0.49037	296.767	112.257	60.210	119.149	279.366	257.193
	800	0.49641	141.143	56.637	27.278	58.626	141.817	135.480
	1200	0.49605	93.658	37.652	17.382	38.487	93.084	90.623
	1600	0.49776	73.806	28.359	12.699	28.850	70.639	69.267

MSE and variance estimates multiplied by 10000

Table A-3: *ML* estimates in the ARCH model with S&P coefficients $\beta = 0.01, \alpha_0 = 0.0009, \alpha_1 = 0.22$ and $\gamma = 2 (\eta_3(v) = 0, \eta_4(v) = 3.00, \eta_4(\varepsilon) = 3)$

QML	*T*	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*	*BWg*
$\tilde{\beta}$	100	0.01000	0.10835	0.09964	0.10803	0.10368	0.10927	0.09934	0.10003
	200	0.01001	0.05138	0.05053	0.05259	0.05124	0.05202	0.05056	0.05067
	400	0.01002	0.02574	0.02538	0.02591	0.02554	0.02570	0.02538	0.02540
	800	0.01004	0.01307	0.01275	0.01290	0.01279	0.01282	0.01274	0.01274
	1200	0.00999	0.00892	0.00850	0.00856	0.00851	0.00853	0.00850	0.00850
	1600	0.00999	0.00648	0.00638	0.00642	0.00639	0.00639	0.00637	0.00638
$\tilde{\alpha}_0$	100	0.00091	0.00034	0.00033	0.00042	0.00036	0.00036	0.00030	0.00030
	200	0.00090	0.00017	0.00016	0.00018	0.00017	0.00017	0.00015	0.00015
	400	0.00090	0.00008	0.00008	0.00009	0.00008	0.00008	0.00008	0.00008
	800	0.00090	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004
	1200	0.00090	0.00003	0.00003	0.00003	0.00003	0.00003	0.00003	0.00003
	1600	0.00090	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002
$\tilde{\alpha}_1$	100	0.20644	268.091	232.683	329.032	273.442	294.460	201.111	201.497
	200	0.20852	126.961	116.377	143.770	126.650	125.650	106.559	106.626
	400	0.21476	62.172	59.085	66.719	61.771	60.850	56.334	56.346
	800	0.21932	30.439	29.872	31.950	30.505	30.168	29.171	29.174
	1200	0.21880	20.489	19.889	20.859	20.224	20.093	19.550	19.551
	1600	0.21777	15.825	14.876	15.414	15.039	14.959	14.705	14.706

MSE and variance estimates multiplied by 10000

Table A-4: *QGLS* estimates in the ARCH model with S&P coefficients $\beta = 0.01, \alpha_0 = 0.0009, \alpha_1 = 0.22$ and $\gamma = 2$

QGLS	*T*	*Mean*	*MSE*	$\hat{\hat{\sigma}}^2_{\hat{\hat{\theta}}}$
$\hat{\hat{\beta}}$	100	0.00999	0.11192	0.10831
	200	0.01001	0.05391	0.05397
	400	0.01002	0.02715	0.02701
	800	0.01003	0.01366	0.01354
	1200	0.01000	0.00940	0.00901
	1600	0.00999	0.00686	0.00676
$\hat{\hat{\alpha}}_0$	100	0.00093	0.00033	0.00034
	200	0.00092	0.00017	0.00016
	400	0.00091	0.00008	0.00008
	800	0.00091	0.00004	0.00004
	1200	0.00090	0.00003	0.00003
	1600	0.00090	0.00002	0.00002
$\hat{\hat{\alpha}}_1$	100	0.17834	221.734	214.755
	200	0.19205	120.998	111.003
	400	0.20569	61.052	57.601
	800	0.21416	30.226	29.448
	1200	0.21528	20.613	19.697
	1600	0.21499	15.898	14.762

MSE and variance estimates multiplied by 10000

$\hat{\hat{\sigma}}^2_{\hat{\hat{\theta}}}$ according to (3.16) and (3.18)

Table A-5: Estimation of the variance $\sigma_y^2 = \alpha_0/(1-\alpha_1) = 0.0009/(1-0.22) = 0.00115$

	T	% σ_y^2	*Mean*	*MSE*
QML	100	0.902	0.00119	0.00069
	200	0.972	0.00116	0.00025
	400	0.998	0.00115	0.00012
	800	1	0.00116	0.00006
	1200	1	0.00115	0.00004
	1600	1	0.00115	0.00003
QGLS	100	0.902	0.00117	0.00050
	200	0.972	0.00115	0.00024
	400	0.998	0.00115	0.00011
	800	1	0.00115	0.00006
	1200	1	0.00115	0.00004
	1600	1	0.00115	0.00003
LS	100	0.903	0.00117	0.00049
	200	0.972	0.00115	0.00024
	400	0.998	0.00115	0.00012
	800	1	0.00116	0.00006
	1200	1	0.00115	0.00004
	1600	1	0.00115	0.00003
HV	100	1	0.00115	0.00048
	200	1	0.00115	0.00024
	400	1	0.00115	0.00012
	800	1	0.00116	0.00006
	1200	1	0.00115	0.00004
	1600	1	0.00115	0.00003

MSE multiplied by 10000, %σ_y^2 percentage of positive estimated variances

Table A-6: *QGLS* estimates in the ARCH model with S&P coefficients $\beta = 0.01, \alpha_0 = 0.0009, \alpha_1 = 0.22$ and $\gamma = 0.75 (\eta_3(v) = 0, \eta_4(v) = 10, \eta_4(\varepsilon) = 17)$

QML	T	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*	*BWg*
$\tilde{\beta}$	100	0.01003	0.10615	0.09168	0.10890	0.09719	0.11541	0.10095	0.10507
	200	0.01001	0.05572	0.04741	0.05114	0.04879	0.05783	0.05308	0.05427
	400	0.01000	0.02829	0.02419	0.02453	0.02450	0.02823	0.02709	0.02741
	800	0.01000	0.01389	0.01228	0.01193	0.01236	0.01400	0.01373	0.01381
	1200	0.01000	0.00933	0.00820	0.00780	0.00823	0.00934	0.00922	0.00926
	1600	0.00998	0.00715	0.00616	0.00579	0.00618	0.00700	0.00693	0.00695
$\tilde{\alpha}_0$	100	0.00088	0.00098	0.00024	0.00011	0.00026	0.00087	0.00079	0.00079
	200	0.00089	0.00049	0.00012	0.00004	0.00012	0.00046	0.00043	0.00043
	400	0.00089	0.00024	0.00006	0.00002	0.00006	0.00024	0.00023	0.00023
	800	0.00090	0.00013	0.00003	0.00001	0.00003	0.00012	0.00012	0.00012
	1200	0.00090	0.00009	0.00002	0.00001	0.00002	0.00008	0.00008	0.00008
	1600	0.00090	0.00006	0.00001	0	0.00001	0.00006	0.00006	0.00006
$\tilde{\alpha}_1$	100	0.24662	1142.865	278.124	179.470	370.560	1384.500	770.578	790.326
	200	0.22887	500.722	124.105	62.973	152.721	622.340	397.222	400.648
	400	0.22238	260.764	59.764	24.195	68.559	280.292	215.501	216.266
	800	0.21890	127.135	29.080	10.093	31.374	129.674	111.661	111.801
	1200	0.21894	84.584	19.339	6.210	20.356	82.905	77.981	78.030
	1600	0.21922	60.980	14.488	4.434	15.085	61.498	57.923	57.948

MSE and variance estimates multiplied by 10000
0 means a number smaller than 0.00005

Table A-7: $QGLS$ estimates in the ARCH model with S&P coefficients $\beta = 0.01, \alpha_0 = 0.0009, \alpha_1 = 0.22$ and $\gamma = 0.75$

$QGLS$	T	$Mean$	MSE	$\hat{\sigma}^2_{\hat{\theta}}$
$\hat{\beta}$	100	0.01002	0.10273	0.10588
	200	0.00999	0.05279	0.05282
	400	0.01001	0.02662	0.02652
	800	0.01000	0.01297	0.01325
	1200	0.01000	0.00862	0.00883
	1600	0.00999	0.00654	0.00661
$\hat{\alpha}_0$	100	0.00092	0.00102	0.00118
	200	0.00092	0.00052	0.00066
	400	0.00091	0.00027	0.00036
	800	0.00091	0.00014	0.00019
	1200	0.00091	0.00009	0.00013
	1600	0.00091	0.00007	0.00010
$\hat{\alpha}_1$	100	0.16692	517.235	760.310
	200	0.18000	363.107	485.378
	400	0.18953	232.179	292.948
	800	0.20009	129.101	163.705
	1200	0.20643	92.864	118.505
	1600	0.20982	72.457	91.824

MSE and variance estimates multiplied by 10000

Table A-8: Estimation of the variance $\sigma_y^2 = \alpha_0/(1-\alpha_1) = 0.00115$

	T	% σ_y^2	$Mean$	MSE	$Mean(cor)$	$MSE(cor)$
QML	100	0.768	0.00168	0.71205	0.00133	0.00786
	200	0.896	0.00150	0.85154	0.00123	0.00233
	400	0.969	0.00123	0.01065	0.00118	0.00079
	800	0.996	0.00118	0.00059	0.00116	0.00031
	1200	0.999	0.00116	0.00027	0.00116	0.00020
	1600	1	0.00116	0.00017	0.00115	0.00014
$QGLS$	100	0.777	0.00137	0.11275	0.00120	0.00288
	200	0.891	0.00129	0.28304	0.00116	0.00117
	400	0.963	0.00119	0.00612	0.00115	0.00057
	800	0.995	0.00116	0.00045	0.00115	0.00028
	1200	0.998	0.00116	0.00032	0.00115	0.00019
	1600	0.999	0.00116	0.00026	0.00115	0.00014
LS	100	0.801	0.00119	0.00299	0.00116	0.00186
	200	0.904	0.00117	0.00132	0.00115	0.00093
	400	0.970	0.00115	0.00070	0.00114	0.00048
	800	0.996	0.00115	0.00035	0.00115	0.00026
	1200	0.999	0.00115	0.00031	0.00114	0.00018
	1600	1	0.00115	0.00025	0.00115	0.00014
HV	100	1	0.00115	0.00265	0.00112	0.00121
	200	1	0.00115	0.00128	0.00113	0.00065
	400	1	0.00115	0.00070	0.00114	0.00036
	800	1	0.00115	0.00035	0.00114	0.00019
	1200	1	0.00115	0.00031	0.00114	0.00013
	1600	1	0.00115	0.00025	0.00115	0.00010

MSE multiplied by 10000, $\%\sigma_y^2$ percentage of positive estimated variances, Mean (cor) and $MSE(cor)$: after elimination of 1% of replications due to outliers in the QML estimation $\widetilde{\sigma_y^2}$.

Table A-9: Moments of QML residuals $\tilde{v}$ and $\tilde{\varepsilon}$

T	$\tilde{v}$	$s^2_{\tilde{v}}$	$\eta_3(\tilde{v})$	$\eta_4(\tilde{v})$	$\tilde{\varepsilon}$	$s^2_{\tilde{\varepsilon}}$	$\eta_3(\tilde{\varepsilon})$	$\eta_4(\tilde{\varepsilon})$
100	-0.00037	0.99841	-0.00928	6.67219	-0.00002	0.00115	-0.00405	7.90015
200	-0.00053	0.99917	-0.01261	7.69700	-0.00002	0.00115	0.00064	9.35890
400	0.00000	0.99957	0.00210	8.42688	-0.00001	0.00115	-0.00786	10.81308
800	0.00020	0.99978	0.00622	8.88707	0.00001	0.00115	0.00866	11.88406
1200	0.00013	0.99986	0.00415	9.19369	0.00001	0.00115	0.00896	12.76738
1600	0.00018	0.99989	-0.00085	9.30884	0.00001	0.00115	-0.01419	13.21444

Table A-10: QML, $QGLS$ and LS estimates in the ARCH model with S&P coefficients $\beta = 0.01, \alpha_0 = 0.0009, \alpha_1 = 0.22$ and $\gamma = 0.6 (\eta_3(v) = 0, \eta_4(v) = 16, \eta_4(\varepsilon) = 60)$

		QML				QGLS			LS	
	T	*Mean*	*MSE*	*QML*	*BW*	*Mean*	*MSE*	$\hat{\sigma}^2_{\hat{\theta}}$	*Mean*	*MSE*
β	100	0.00994	0.11529	0.13941	0.09990	0.00996	0.10983	0.10579	0.00995	0.11834
	200	0.00996	0.06229	0.06058	0.05484	0.00996	0.05402	0.05338	0.00998	0.06131
	400	0.01000	0.02881	0.03083	0.02877	0.01000	0.02470	0.02669	0.01000	0.02737
	800	0.01000	0.01468	0.01501	0.01445	0.01000	0.01253	0.01319	0.01000	0.01405
	1200	0.01001	0.00967	0.01015	0.00975	0.01001	0.00843	0.00883	0.01001	0.00957
	1600	0.01000	0.00743	0.00743	0.00725	0.01001	0.00639	0.00659	0.01001	0.00712
α_0	100	0.00086	0.00148	0.00137	0.00120	0.00091	0.00158	0.00167	0.00101	0.00564
	200	0.00088	0.00073	0.00066	0.00062	0.00092	0.00081	0.00095	0.00101	0.00261
	400	0.00089	0.00038	0.00037	0.00035	0.00092	0.00044	0.00054	0.00100	0.00101
	800	0.00089	0.00020	0.00019	0.00018	0.00091	0.00022	0.00029	0.00098	0.00046
	1200	0.00090	0.00013	0.00013	0.00013	0.00091	0.00015	0.00020	0.00098	0.00035
	1600	0.00090	0.00010	0.00010	0.00009	0.00091	0.00011	0.00015	0.00097	0.00027
α_1	100	0.27395	1822.118	2869.610	1227.906	0.171	662.739	1068.457	0.107	323.463
	200	0.25475	1047.638	1458.254	787.812	0.175	506.937	757.544	0.113	257.064
	400	0.23181	483.193	609.709	405.458	0.182	356.931	458.849	0.123	233.704
	800	0.21895	227.141	252.587	195.045	0.190	225.484	261.883	0.133	180.134
	1200	0.22036	148.256	166.613	134.962	0.200	155.227	187.316	0.142	148.575
	1600	0.22066	109.919	113.274	99.215	0.204	119.514	146.458	0.146	140.890

MSE and variance estimates multiplied by 10000

Table A-11: Estimation of the variance $\sigma_y^2 = \alpha_0/(1-\alpha_1) = 0.00115$

	T	$\% \sigma_y^2$	*Mean*	*MSE*	*Mean(cor)*	*MSE(cor)*
QML	100	0.719	0.00179	0.78087	0.00133	0.00994
	200	0.834	0.00165	1.00551	0.00128	0.00512
	400	0.937	0.00141	0.17572	0.00123	0.00233
	800	0.990	0.00123	0.02551	0.00117	0.00072
	1200	0.998	0.00119	0.00126	0.00117	0.00042
	1600	0.999	0.00117	0.00126	0.00116	0.00028
QGLS	100	0.744	0.00268	39.00784	0.00123	0.00515
	200	0 .848	0.00123	0.00899	0.00117	0.00188
	400	0.931	0.00129	0.30060	0.00117	0.00120
	800	0.983	0.00117	0.00475	0.00114	0.00051
	1200	0.994	0.00118	0.00725	0.00115	0.00036
	1600	0.998	0.00117	0.00399	0.00115	0.00026
LS	100	0.775	0.00122	0.01393	0.00116	0.00275
	200	0.870	0.00119	0.00650	0.00115	0.00149
	400	0.945	0.00116	0.00152	0.00114	0.00081
	800	0.992	0.00114	0.00065	0.00113	0.00044
	1200	0.998	0.00115	0.00046	0.00114	0.00031
	1600	0.999	0.00115	0.00034	0.00114	0.00023
HV	100	1	0.00116	0.01136	0.00111	0.00178
	200	1	0.00116	0.00582	0.00112	0.00106
	400	1	0.00116	0.00148	0.00113	0.00059
	800	1	0.00114	0.00065	0.00113	0.00032
	1200	1	0.00115	0.00046	0.00114	0.00023
	1600	1	0.00115	0.00034	0.00114	0.00017

MSE multiplied by 10000, $\%\sigma_y^2$ percentage of positive estimated variances

Table A-12: Moments of *QML* residuals $\tilde{v}$ and $\tilde{\varepsilon}$

T	$\tilde{v}$	$s^2_{\tilde{v}}$	$\eta_3(\tilde{v})$	$\eta_4(\tilde{v})$	$\tilde{\varepsilon}$	$s^2_{\tilde{\varepsilon}}$	$\eta_3(\tilde{\varepsilon})$	$\eta_4(\tilde{\varepsilon})$	rate
100	0.00058	0.99983	-0.00784	8.74001	0.00000	0.00116	-0.02076	10.20088	0.81
200	0.00005	0.99963	-0.02117	10.37299	0.00002	0.00116	0.00301	12.89472	0.90
400	0.00002	0.99940	-0.00859	12.20757	-0.00000	0.00116	-0.02526	15.73314	0.96
800	0.00013	0.99971	-0.01173	13.44752	0.00000	0.00114	0.01299	18.39020	0.99
1200	0.00005	0.99980	-0.00328	13.94305	0.00000	0.00115	0.00080	20.05828	1.00
1600	0.00013	0.99985	0.00500	14.22863	0.00000	0.00115	-0.00210	20.71042	1.00

Table A-13: *QML*, *QGLS* and *LS* estimates in the ARCH model with S&P coefficients $\beta = 0.01, \alpha_0 = 0.0009, \alpha_1 = 0.22$ and $\gamma = 1, \mu_2 = 6, g = 0.025 (\eta_3(v) = 2.02, \eta_4(v) = 12, \eta_4(\varepsilon) = 27)$

		QML					QGLS			LS	
	n	*Mean*	*MSE*	*QML*	*BW*	*BWg*	*Mean*	*MSE*	$\hat{\sigma}^2_{\hat{\theta}}$	*Mean*	*MSE*
β	100	0.00973	0.12140	0.14029	0.10770	0.12064	0.00999	0.11464	0.10911	0.01011	0.12507
	200	0.00983	0.05990	0.06576	0.05738	0.06268	0.01001	0.05314	0.05394	0.01006	0.05904
	400	0.00983	0.02915	0.03022	0.02909	0.03132	0.00993	0.02613	0.02672	0.00996	0.02885
	800	0.00998	0.01380	0.01451	0.01477	0.01585	0.01000	0.01233	0.01333	0.01003	0.01403
	1200	0.00992	0.00981	0.00950	0.00980	0.01050	0.00995	0.00891	0.00879	0.00997	0.01001
	1600	0.00995	0.00688	0.00708	0.00735	0.00786	0.00997	0.00618	0.00660	0.00998	0.00699
α_1	100	0.00088	0.00130	0.00150	0.00104	0.00107	0.00094	0.00128	0.00151	0.00105	0.00260
	200	0.00089	0.00065	0.00064	0.00055	0.00056	0.00093	0.00064	0.00081	0.00102	0.00116
	400	0.00089	0.00032	0.00031	0.00029	0.00029	0.00091	0.00032	0.00042	0.00100	0.00058
	800	0.00090	0.00016	0.00016	0.00015	0.00015	0.00091	0.00016	0.00021	0.00099	0.00034
	1200	0.00090	0.00010	0.00010	0.00010	0.00010	0.00090	0.00010	0.00014	0.00097	0.00023
	1600	0.00090	0.00008	0.00008	0.00007	0.00008	0.00091	0.00008	0.00011	0.00097	0.00019
α_2	100	0.28212	2138.688	7169.387	1319.890	1400.020	0.166	806.363	1325.168	0.086	329.033
	200	0.25325	1023.443	1415.210	646.492	659.531	0.175	536.194	790.352	0.103	257.920
	400	0.23020	413.825	478.891	313.199	314.922	0.190	334.182	438.181	0.120	196.801
	800	0.21991	197.229	201.516	157.837	158.091	0.197	189.458	228.826	0.135	154.197
	1200	0.21920	124.693	127.543	108.690	108.790	0.203	126.240	157.478	0.143	129.649
	1600	0.21666	93.411	90.486	82.057	82.105	0.205	96.306	118.537	0.148	110.754

MSE and variance estimates multiplied by 10000

Table A-14: Estimation of the variance $\sigma_y^2 = \alpha_0/(1-\alpha_1) = 0.00115$

	T	% σ_y^2	*Mean*	*MSE*	*Mean*(*cor*)	*MSE*(*cor*)
QML	100	0.649	0.00338	74.68363	0.00154	0.02049
	200	0.828	0.00182	3.29567	0.00137	0.00695
	400	0.945	0.00136	0.13826	0.00122	0.00140
	800	0.991	0.00120	0.00190	0.00118	0.00049
	1200	0.998	0.00117	0.00035	0.00116	0.00029
	1600	0.999	0.00116	0.00024	0.00116	0.00021
QGLS	100	0.675	0.00218	9.50793	0.00137	0.00941
	200	0.833	0.00155	0.80065	0.00124	0.00277
	400	0.942	0.00121	0.00333	0.00118	0.00096
	800	0.990	0.00117	0.00060	0.00116	0.00044
	1200	0.998	0.00116	0.00034	0.00115	0.00027
	1600	0.999	0.00115	0.00024	0.00115	0.00020
LS	100	0.720	0.00127	0.00517	0.00123	0.00328
	200	0.859	0.00119	0.00190	0.00117	0.00135
	400	0.951	0.00116	0.00102	0.00115	0.00069
	800	0.991	0.00116	0.00052	0.00115	0.00039
	1200	0.998	0.00115	0.00034	0.00114	0.00026
	1600	0.999	0.00115	0.00030	0.00114	0.00020
HV	100	1	0.00118	0.00415	0.00114	0.00193
	200	1	0.00116	0.00177	0.00114	0.00096
	400	1	0.00115	0.00100	0.00114	0.00051
	800	1	0.00116	0.00052	0.00115	0.00028
	1200	1	0.00115	0.00034	0.00114	0.00019
	1600	1	0.00115	0.00030	0.00114	0.00014

MSE multiplied by 10000, %σ_y^2 percentage of positive estimated variances

Table A-15: Moments of QML residuals $\tilde{v}$ and $\tilde{u}$

T	$\tilde{v}$	$s^2_{\tilde{v}}$	$\eta_3(\tilde{v})$	$\eta_4(\tilde{v})$	$\tilde{\varepsilon}$	$s^2_{\tilde{\varepsilon}}$	$\eta_3(\tilde{\varepsilon})$	$\eta_4(\tilde{\varepsilon})$	rate
100	0.00365	0.99755	1.55140	9.53594	0.00037	0.00118	1.71353	11.35415	0.84
200	0.00288	0.99879	1.78779	10.74138	0.00023	0.00116	1.93298	13.15655	0.94
400	0.00162	0.99946	1.89501	11.32281	0.00013	0.00115	2.06507	14.99902	0.99
800	-0.00002	0.99976	1.95841	11.60277	0.00005	0.00116	2.18984	16.90020	1.00
1200	0.00040	0.99984	1.98001	11.72125	0.00004	0.00115	2.22699	17.67808	1.00
1600	0.00036	0.99989	1.98989	11.74605	0.00003	0.00115	2.25114	18.23502	1.00

Table A-16: ML estimates in the Fiorentini GARCH model with $\beta = -0.29, \alpha_0 = 0.20, \alpha_1 = 0.35, \delta = 0.45$ and $\gamma = 2$

ML	T	$Mean$	MSE	S	OP	H	QML	BW
$\tilde{\beta}$	200	-0.2890	31.990	30.507	32.865	31.216	31.662	30.419
	400	-0.2908	15.399	15.266	15.887	15.434	15.491	15.246
	800	-0.2903	7.887	7.623	7.788	7.664	7.666	7.615
	1200	-0.2897	5.190	5.076	5.151	5.089	5.083	5.073
$\tilde{\alpha}_1$	200	0.2433	212.156	186.437	257.152	145.895	158.544	158.389
	400	0.2204	61.022	54.039	63.200	57.941	76.771	51.064
	800	0.2107	25.064	23.234	25.591	23.958	24.907	22.413
	1200	0.2057	15.071	14.675	15.651	14.947	15.187	14.369
$\tilde{\alpha}_2$	200	0.3500	150.642	143.836	179.446	152.103	157.370	132.771
	400	0.3484	73.974	71.463	81.098	73.997	76.170	68.337
	800	0.3506	35.500	35.674	38.135	36.239	36.645	35.013
	1200	0.3488	24.186	23.549	24.793	23.795	23.771	23.146
$\tilde{\delta}$	200	0.3907	439.644	432.716	623.532	324.014	353.089	359.521
	400	0.4228	156.962	139.555	164.283	146.288	186.461	131.803
	800	0.4355	66.907	62.441	68.926	63.823	67.196	60.342
	1200	0.4429	42.658	40.314	43.158	40.814	41.604	39.453

MSE and variance estimates multiplied by 10000

Table A-17: QML estimates in the Fiorentini GARCH model with $\beta = -0.29, \alpha_0 = 0.20, \alpha_1 = 0.35, \delta = 0.45$ and $\gamma = 1$

ML	T	$Mean$	MSE	S	OP	H	QML	BW
$\tilde{\beta}$	200	-0.2895	31.873	27.128	28.594	28.026	32.253	30.864
	400	-0.2902	16.336	13.681	13.402	13.918	15.997	15.697
	800	-0.2899	8.021	6.872	6.386	6.926	8.008	7.972
	1200	-0.2900	5.394	4.588	4.173	4.611	5.332	5.327
$\tilde{\alpha}_1$	200	0.2495	281.074	619.570	760.139	131.449	253.938	666.938
	400	0.2226	105.530	39.444	26.079	39.732	117.869	77.230
	800	0.2108	40.613	16.066	8.813	16.754	40.430	34.852
	1200	0.2080	25.473	10.259	5.205	10.647	26.289	23.185
$\tilde{\alpha}_2$	200	0.3619	383.262	154.835	110.119	172.505	407.011	308.626
	400	0.3576	188.332	75.603	43.864	80.405	195.058	168.439
	800	0.3558	92.356	37.229	18.942	38.730	96.064	87.602
	1200	0.3516	61.736	24.436	11.752	25.059	61.506	58.005
$\tilde{\delta}$	200	0.3684	672.491	1788.821	2257.209	359.949	637.495	1855.250
	400	0.4119	307.014	116.944	83.327	115.271	308.479	211.644
	800	0.4314	127.094	50.315	28.902	52.080	124.486	104.719
	1200	0.4371	81.446	32.772	17.255	33.823	83.095	71.781

MSE and variance estimates multiplied by 10000

Table A-18: *ML* estimates in the GARCH model with *S&P* coefficients $\beta = 0.01, \alpha_0 = 0.00015, \alpha_1 = 0.15, \delta = 0.72$ and $\gamma = 2$

ML	*T*	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*
$\tilde{\beta}$	400	0.009985	0.023979	0.023827	0.024559	0.024040	0.024322	0.023892
	800	0.009989	0.011886	0.011873	0.012072	0.011924	0.011963	0.011877
	1200	0.009999	0.007805	0.007918	0.008003	0.007938	0.007955	0.007926
	1600	0.010002	0.005810	0.005940	0.005987	0.005951	0.005962	0.005945
	2000	0.009997	0.004696	0.004746	0.004778	0.004754	0.004760	0.004748
$\tilde{\alpha}_1$	400	0.000240	0.000696	0.001438	0.001842	0.000256	0.000313	0.001276
	800	0.000176	0.000092	0.000064	0.000071	0.000058	0.000075	0.000061
	1200	0.000166	0.000034	0.000029	0.000031	0.000030	0.000035	0.000029
	1600	0.000162	0.000025	0.000020	0.000021	0.000021	0.000023	0.000020
	2000	0.000160	0.000016	0.000015	0.000016	0.000016	0.000016	0.000015
$\tilde{\alpha}_2$	400	0.155208	34.665029	34.163404	40.185192	34.896348	36.910025	31.930475
	800	0.152280	15.912496	15.921534	17.459049	16.230682	17.070181	15.326092
	1200	0.151985	10.510070	10.391240	11.054135	10.574944	10.958017	10.164867
	1600	0.150728	7.721586	7.672395	8.036404	7.765652	7.968565	7.555073
	2000	0.151051	6.093172	6.111097	6.349549	6.167958	6.276940	6.030999
$\tilde{\delta}$	400	0.627255	698.287546	*	*	299.498608	341.602924	*
	800	0.692165	118.837997	89.495139	99.215006	83.104477	105.417232	85.446554
	1200	0.703180	51.512609	44.774565	47.774876	46.202658	52.829229	43.919473
	1600	0.707649	35.997877	31.972537	33.585732	32.488778	35.456071	31.539443
	2000	0.709801	25.846380	24.507326	25.491539	24.717805	25.980988	24.250893

MSE and variance estimates multiplied by 10000
* unreliable estimate (exceeds more than two times the *MSE* value)

Table A-19: *QML* estimates in the GARCH model with *S&P* coefficients $\beta = 0.01, \alpha_0 = 0.00015, \alpha_1 = 0.15, \delta = 0.72$ and $\gamma = 0.75$

ML	*T*	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*
$\tilde{\beta}$	400	0.009999	0.024359	0.020815	0.021376	0.021156	0.024757	0.024064
	800	0.010013	0.012024	0.010449	0.010129	0.010530	0.012175	0.012040
	1200	0.009986	0.008128	0.006972	0.006589	0.007008	0.008077	0.008027
	1600	0.010002	0.006102	0.005244	0.004879	0.005264	0.006049	0.006044
	2000	0.009995	0.004803	0.004191	0.003861	0.004202	0.004827	0.004827
$\tilde{\alpha}_1$	400	0.000288	0.001070	21.102062	27.435340	0.000214	0.000480	16.431927
	800	0.000205	0.000296	0.000080	0.000051	0.000043	0.000180	0.000180
	1200	0.000178	0.000121	0.000018	0.000007	0.000019	0.000101	0.000060
	1600	0.000170	0.000074	0.000012	0.000004	0.000013	0.000056	0.000041
	2000	0.000163	0.000039	0.000008	0.000003	0.000009	0.000040	0.000029
$\tilde{\alpha}_2$	400	0.175327	136.370913	33.918759	15.413689	37.926862	202.469482	114.767566
	800	0.163658	60.270965	14.472905	5.353295	15.483723	70.544890	54.347606
	1200	0.158194	38.086409	8.875458	2.986825	9.333275	42.947335	34.950818
	1600	0.155250	26.695053	6.369299	2.019288	6.635194	28.831341	25.367866
	2000	0.154527	20.374321	4.967593	1.516026	5.137731	22.039439	19.856429
$\tilde{\delta}$	400	0.552410	1299.479532	*	*	287.328817	637.022832	*
	800	0.651764	408.379443	119.474057	75.912416	66.548917	271.855435	262.067319
	1200	0.684285	180.415137	30.935418	12.584935	32.503040	165.140469	96.566709
	1600	0.695203	117.464964	20.888199	7.782649	21.825726	95.911272	68.954021
	2000	0.702380	68.255820	14.850344	5.079317	15.627472	69.827289	51.975478

MSE and variance estimates multiplied by 10000
* unreliable estimate (exceeds more than two times the *MSE* value)

Table A-20: *QML* estimates in the GARCH model with *S&P* coefficients $\beta = 0.01, \alpha_0 = 0.00015, \alpha_1 = 0.15, \delta = 0.72$ and $\gamma = 0.6$

QML	*T*	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*
	400	0.010032	0.024558	0.019603	0.020471	0.020031	0.025635	0.024638
	800	0.010009	0.012320	0.009904	0.009556	0.010007	0.012560	0.012431
$\tilde{\beta}$	1200	0.009991	0.008241	0.006643	0.006198	0.006691	0.008609	0.008286
	1600	0.010002	0.006035	0.004996	0.004553	0.005021	0.006250	0.006227
	2000	0.009989	0.004987	0.003995	0.003597	0.004012	0.004995	0.004984
	400	0.000299	0.001107	68.630125	134.766125	0.000204	0.000410	41.037056
	800	0.000219	0.000398	11.041167	10.735009	0.000046	0.000241	11.683868
$\tilde{\alpha}_1$	1200	0.000191	0.000203	0.000031	0.000014	0.000021	0.000144	0.000115
	1600	0.000176	0.000107	0.000011	0.000003	0.000012	0.000093	0.000057
	2000	0.000168	0.000063	0.000008	0.000002	0.000008	0.000062	0.000040
	400	0.191365	266.633762	40.394923	14.436780	48.255705	365.488836	207.079463
	800	0.172752	115.636597	16.336691	4.568930	18.403758	136.557269	95.156130
$\tilde{\alpha}_2$	1200	0.164340	70.853354	9.748340	2.394210	10.608007	85.716600	61.553544
	1600	0.160063	47.142065	6.814800	1.553504	7.326165	53.544349	43.835734
	2000	0.156763	38.088251	5.204430	1.123961	5.556717	42.812362	34.367995
	400	0.525659	1453.589507	*	*	288.958667	524.347817	*
	800	0.627315	603.342167	*	*	72.895044	348.637479	*
$\tilde{\delta}$	1200	0.665402	314.018853	49.847216	22.281513	35.120852	253.938018	177.717625
	1600	0.684532	179.957457	20.313980	6.010881	21.949546	157.927518	93.669209
	2000	0.695936	115.599856	14.413514	3.834310	15.687110	116.839738	71.229445

MSE and variance estimates multiplied by 10000
* unreliable estimate (exceeds more than two times the *MSE* value)

Table A-21: *QML* estimates in the GARCH model with *S&P* coefficients $\beta = 0.01, \alpha_0 = 0.00015, \alpha_1 = 0.15, \delta = 0.72, \gamma = 1, g = 0.025$ and $\mu_2 = 6$

QML	*T*	*Mean*	*MSE*	*S*	*OP*	*H*	*QML*	*BW*
	400	0.009795	0.024865	0.019603	0.026872	0.020132	0.025455	0.025999
	800	0.009893	0.012354	0.009897	0.013005	0.010046	0.012257	0.013180
$\tilde{\beta}$	1200	0.009915	0.007937	0.006619	0.008556	0.006694	0.008048	0.008803
	1600	0.009958	0.005928	0.004974	0.006378	0.005021	0.005977	0.006639
	2000	0.009968	0.004808	0.003990	0.005109	0.004023	0.004762	0.005315
	400	0.000330	0.001427	1.782753	3.793751	0.000243	0.000664	0.947438
	800	0.000228	0.000466	0.000366	0.000302	0.000051	0.000179	0.000671
$\tilde{\alpha}_1$	1200	0.000192	0.000203	0.000035	0.000015	0.000020	0.000120	0.000119
	1600	0.000175	0.000100	0.000014	0.000005	0.000011	0.000067	0.000054
	2000	0.000168	0.000059	0.000007	0.000002	0.000008	0.000056	0.000034
	400	0.190733	253.783096	43.054958	15.084097	52.004212	291.732950	176.480349
	800	0.172263	99.064311	17.089455	4.682566	19.340010	107.735843	77.187966
$\tilde{\alpha}_2$	1200	0.163438	55.960507	10.005918	2.432833	11.004563	63.815196	47.878205
	1600	0.159643	39.383601	6.976699	1.586842	7.476095	42.765826	34.476318
	2000	0.156063	28.923702	5.295695	1.160328	5.609834	31.842292	26.638507
	400	0.491464	1844.647502	*	*	373.245558	917.276856	*
	800	0.617817	690.204659	568.578447	510.158376	79.696092	281.706670	969.426044
$\tilde{\delta}$	1200	0.665059	322.151505	55.304576	23.512562	33.991914	198.930003	187.902412
	1600	0.686715	173.976539	25.193799	8.887737	20.547883	125.119564	98.003932
	2000	0.696106	111.164791	13.816354	3.376656	15.093561	105.440458	66.007392

MSE and variance estimates multiplied by 10000
* unreliable estimate (exceeds more than two times the *MSE* value)

Table A-22: Estimation of the variance $\sigma_y^2 = \alpha_0/(1-\alpha_1-\delta) = 0.00115$ in the GARCH model with $S\&P$ coefficients and $\gamma = 0.6$

	T	$\%\ \hat{\sigma}_y^2$	$Mean$	MSE	Mittel$(cor.)$	$MSE(cor.)$
QML	400	0.851200	0.001970	0.663215	0.001430	0.018869
	800	0.953400	0.001569	0.230604	0.001316	0.006293
	1200	0.981800	0.001430	0.192501	0.001244	0.002726
	1600	0.992800	0.001315	0.048208	0.001215	0.001553
	2000	0.996200	0.001467	5.343096	0.001195	0.001088
$QGLS$	400	0.843900	0.001830	1.004775	0.001349	0.012523
	800	0.945700	0.001516	0.336445	0.001270	0.005054
	1200	0.977600	0.001390	0.105158	0.001218	0.002547
	1600	0.989900	0.001318	0.060702	0.001195	0.001562
	2000	0.993900	0.001237	0.024938	0.001175	0.001071
LS	400	0.816700	0.001206	0.015127	0.001135	0.002308
	800	0.893300	0.001158	0.004976	0.001117	0.001230
	1200	0.916500	0.001146	0.001971	0.001118	0.000901
	1600	0.935100	0.001143	0.001480	0.001119	0.000694
	2000	0.938800	0.001142	0.001860	0.001117	0.000591
HV	400	1	0.001175	0.010829	0.001115	0.000755
	800	1	0.001157	0.004991	0.001117	0.000434
	1200	1	0.001153	0.002172	0.001123	0.000312
	1600	1	0.001149	0.001543	0.001125	0.000242
	2000	1	0.001150	0.001997	0.001123	0.000198

MSE multiplied by 10000, $\%\sigma_y^2$ percentage of positive estimated variances, $MSE(cor)$: elimination of 1% of replications due to outliers in the QML estimation $\tilde{\sigma}_y^2$.

Table A-23: Moments of QML residuals $\tilde{v}$ and $\tilde{\varepsilon}$ in the GARCH model with $S\&P$ coefficients and $\gamma = 0.6$

T	$\bar{\tilde{v}}$	$s^2_{\tilde{v}}$	$\eta_3(\hat{v})$	$\eta_4(\hat{v})$	$\bar{\tilde{\varepsilon}}$	$s^2_{\tilde{\varepsilon}}$	$\eta_3(\hat{\varepsilon})$	$\eta_4(\hat{\varepsilon})$	$rate$
400	-0.00009	0.999	0.00585	11.3	-0.00000	0.00118	0.01674	15.9	0.82
800	-0.00009	1.003	-0.00168	12.7	-0.00001	0.00116	-0.01195	19.9	0.90
1200	0.00024	0.999	0.01594	13.4	0.00001	0.00115	0.00435	22.8	0.94
1600	0.00001	0.999	0.00194	13.7	-0.00000	0.00115	0.00542	24.7	0.96
2000	-0.00007	0.999	0.00201	14.0	-0.00001	0.00115	0.00044	26.2	0.97

rate of replications in the QML algorithm without convergence

Table A-24: Estimation of the variance $\sigma_y^2 = \alpha_0/(1-\alpha_1-\delta) = 0.00020$ in the GARCH model with DAX coefficients and $\gamma = 0.6$

	T	% $\hat{\sigma}_y^2$	*Mean*	*MSE*	*Mean(cor.)*	*MSE(cor.)*
$ML_{(99p)}$	400	0.726100	0.003212	583.287018	0.000241	0.000985
	800	0.848100	0.000438	0.357907	0.000247	0.000775
	1200	0.914100	0.000390	0.215043	0.000242	0.000588
	1600	0.938400	0.000348	0.026509	0.000242	0.000520
	2000	0.962600	0.000355	0.060930	0.000240	0.000444
$TSLS_{(99p)}$	400	0.735300	0.000376	0.097119	0.000225	0.000691
	800	0.859700	0.000375	0.166630	0.000235	0.000680
	1200	0.917800	0.000356	0.091866	0.000224	0.000413
	1600	0.939300	0.000315	0.034853	0.000227	0.000412
	2000	0.959700	0.000303	0.018330	0.000225	0.000336
$LS_{(99p)}$	400	0.779400	0.000215	0.001076	0.000193	0.000216
	800	0.903500	0.000210	0.001412	0.000189	0.000121
	1200	0.950600	0.000197	0.000396	0.000185	0.000088
	1600	0.966800	0.000200	0.000894	0.000186	0.000072
	2000	0.976200	0.000201	0.004372	0.000184	0.000057
$MM_{(99p)}$	400	1	0.000198	0.000642	0.000181	0.000026
	800	1	0.000203	0.000799	0.000185	0.000016
	1200	1	0.000195	0.000359	0.000183	0.000011
	1600	1	0.000199	0.000820	0.000186	0.000010
	2000	1	0.000195	0.000212	0.000185	0.000008

MSE multiplied by 10000, $\%\sigma_y^2$ percentage of positive estimated variances, $MSE(cor)$: elimination of 1% of replications due to outliers in the QML estimation $\tilde{\sigma}_y^2$.

Table A-25: Moments of QML residuals $\tilde{v}$ and $\tilde{\varepsilon}$ in the GARCH model with DAX coefficients and $\gamma = 0.6$

	$\bar{\hat{v}}$	$s^2_{\hat{v}}$	$\eta_3(\hat{v})$	$\eta_4(\hat{v})$	$\bar{\hat{\varepsilon}}$	$s^2_{\hat{\varepsilon}}$	$\eta_3(\hat{\varepsilon})$	$\eta_4(\hat{\varepsilon})$	*rate*
400	-0.00014	1.001	0.00850	11.4	0.00000	0.00020	0.01272	14.6	0.74
800	-0.00002	1.001	0.02234	12.7	0.00000	0.00020	0.02367	19.2	0.85
1200	-0.00019	1.000	0.01340	13.4	-0.00000	0.00020	0.00602	22.0	0.90
1600	-0.00010	1.000	-0.00245	13.8	-0.00000	0.00020	-0.01819	25.3	0.93
2000	-0.00017	1.000	-0.00384	14.1	0	0.00020	0.01988	26.9	0.95

rate of replications in the QML algorithm without convergence