MCMC Bayesian Estimation in FIEGARCH Models

Taiane S. Prass[(a)]*, Sílvia R.C. Lopes[(a)] and Jorge A. Achcar[(b)]

(a) Mathematics Institute - UFRGS, Porto Alegre, RS, Brazil

(b) Medical School - USP, Ribeirão Preto, SP, Brazil

April 11, 2014

Abstract

Bayesian inference for fractionally integrated exponential generalized autoregressive conditional heteroskedastic (FIEGARCH) models using Markov Chain Monte Carlo (MCMC) methods is described. A simulation study is presented to assess the performance of the procedure, under the presence of long-memory in the volatility. Samples from FIEGARCH processes are obtained upon considering the generalized error distribution (GED) for the innovation process. Different values for the tail-thickness parameter ν are considered covering both scenarios, innovation processes with lighter ($\nu > 2$) and heavier ($\nu < 2$) tails than the Gaussian distribution ($\nu = 2$). A comparison between the performance of quasi-maximum likelihood (QML) and MCMC procedures is also discussed. An application of the MCMC procedure to estimate the parameters of a FIEGARCH model for the daily log-returns of the S&P500 US stock market index is provided.

Key words: Bayesian inference, MCMC, FIEGARCH processes, Long-range dependence.

1 Introduction

ARCH-type (Autoregressive Conditional Heteroskedasticity) and stochastic volatility (Breidt et al., 1998) models are commonly used in financial time series modeling to represent the dynamic evolution of volatilities. By ARCH-type models we mean not only the ARCH model proposed by Engle (1982) but also several generalizations that were lately proposed.

Among the most popular generalizations of the ARCH model is the generalized ARCH (GARCH) model, introduced by Bollerslev (1986), for which the conditional variance depends not only on the p past values of the process (as in the ARCH model), but also on the q past values of the conditional variance. Although the ARCH and GARCH models are widely used in practice, they do not take into account the asymmetry in the volatility, that is, the fact that volatility tends to rise in response to "bad" news and to fall in response to "good" news. As an alternative, Nelson (1991) introduces the exponential GARCH (EGARCH) model. This model not only describes the asymmetry on the volatility, but also has the advantage that the positivity of the conditional variance is always attained since it is defined in terms of the logarithm function.

*Corresponding author: taianeprass@gmail.com

The fractionally integrated EGARCH (FIEGARCH) and fractionally integrated GARCH (FIGARCH) models proposed, respectively, by Bollerslev and Mikkelsen (1996) and Baillie et al. (1996), generalize the EGARCH (Nelson, 1991) and the GARCH (Bollerslev, 1986) models, respectively. FIEGARCH models have not only the capability of modeling clusters of volatility (as ARCH and GARCH models do) and capturing its asymmetry (as the EGARCH model does) but they also take into account the characteristic of long memory in the volatility (as the FIGARCH model does). The non-stationarity of FIGARCH models (in the weak sense) makes this class of models less attractive for practical applications. Another drawback of the FIGARCH models is that we must have $d \geq 0$ and the polynomial coefficients in its definition must satisfy some restrictions so the conditional variance will be positive. FIEGARCH models do not have this problem since the variance is defined in terms of the logarithm function, moreover, they are weak stationary whenever the long memory parameter d is smaller than 0.5 (Lopes and Prass, 2014).

A complete study on the theoretical properties of FIEGARCH processes is presented in Lopes and Prass (2014). The authors also conduct a simulation study to analyze the finite sample performance of the quasi-maximum likelihood (QML) procedure on parameter estimation. The QML procedure became popular for two main reasons. First, the expression for the quasi-likelihood function is simpler for the Gaussian case than when considering, for example, the Student's t or the generalized error distribution (GED). Second, since the parameters of the distribution function are not estimated, the dimension of the optimization problem is reduced. On the other hand, the results in Lopes and Prass (2014) indicate that, although the QML presents a relatively good performance when the sample size is 2000 and the estimation improves as the sample size increases, it does so very slowly.

In this work we propose the use of Bayesian methods considering Monte Carlo simulation techniques on the estimation of the FIEGARCH model parameters. This procedure is usually considered to analyze financial time series assuming stochastic volatility models (see, for example, Meyer and Yu, 2000), mostly because of the difficulty on applying traditional statistical techniques due to the complexity of the likelihood function. To generate samples from the joint posterior distribution for the parameters of interest we consider the Markov Chain Monte Carlo (MCMC) procedure known in the literature as *MH-whithing-Gibbs* or *Gibbs sampler with Metropolis steps*, which is a combination of Gibbs sampler (see, for example, Gelfand and Smith, 1990; Casela and George, 1992; Smith and Roberts, 1993) and Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970; Chib and Greenberg, 1995) algorithms. These samples are generated from all conditional posterior distributions for each parameter given all the other parameters, the data and a set of initial conditions.

A simulation study is conducted to assess the finite sample performance of the procedure proposed here, under the presence of long-memory in the volatility. The samples from FIEGARCH processes are obtained upon considering the GED for the innovation process. Taking into account that financial time series are usually characterized by heavy tailed distributions, different values for the tail-thickness parameter ν are considered covering both scenarios: innovation processes with lighter and heavier tails than the Gaussian distribution.

The paper is organized as follows. In Section 2 a review on the definition and main properties of FIEGARCH processes is presented. Section 3 describes the parameter estimation procedure when Bayesian inference using MCMC is considered. Section 4 describes the steps used in the simulation study, such as the data generating process, the prior selection procedure and the performance measures considered. This section also reports the simulation results for the MCMC procedure and the comparison between QML and MCMC approaches. Section 5 presents an application of the suggested MCMC algorithm to a real data set. Section 6 concludes the paper.

2 FIEGARCH Processes

Let $(1-\mathcal{B})^{-d}$ be the operator defined by its Maclaurin series expansion, namely,

$$(1-\mathcal{B})^{-d} = \sum_{k=0}^{\infty} \tau_{d,k}\,\mathcal{B}^k, \tag{1}$$

where $\tau_{d,0} := 1$, $\tau_{d,k} := \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)}$, for all $k \geq 1$, $\Gamma(\cdot)$ is the gamma function and $\mathcal{B}$ is the backward shift operator defined by $\mathcal{B}^k(X_t) = X_{t-k}$, for all $k \in \mathbb{N}$.

Assume that $\alpha(\cdot)$ and $\beta(\cdot)$ are polynomials of order p and q, respectively, defined by

$$\alpha(z) = \sum_{i=0}^{p}(-\alpha_i)z^i \quad \text{and} \quad \beta(z) = \sum_{j=0}^{q}(-\beta_j)z^j, \tag{2}$$

with $\alpha_0 = \beta_0 = -1$. If $\alpha(\cdot)$ and $\beta(\cdot)$ have no common roots and if $\beta(\cdot)$ has no roots in the closed disk $\{z : |z| \leq 1\}$, then the function $\lambda(\cdot)$, defined by

$$\lambda(z) = \frac{\alpha(z)}{\beta(z)}(1-z)^{-d} := \sum_{k=0}^{\infty}\lambda_{d,k}z^k, \qquad \text{for all } |z| < 1, \tag{3}$$

is analytic in the open disk $\{z : |z| < 1\}$, for any $d > 0$, and in the closed disk $\{z : |z| \leq 1\}$, whenever $d \leq 0$. Therefore, $\lambda(\cdot)$ is well defined and the power series representation in (3) is unique. More specifically, the coefficients $\lambda_{d,k}$, for all $k \in \mathbb{N}$, are given by (see Lopes and Prass, 2014)

$$\lambda_{d,0} = 1 \quad \text{and} \quad \lambda_{d,k} = -\alpha_k^* + \sum_{i=0}^{k-1}\lambda_{d,i}\sum_{j=0}^{k-i}\beta_j^*\delta_{d,k-i-j}, \quad \text{for all} \ \ k \geq 1, \tag{4}$$

where

$$\alpha_m^* := \begin{cases} \alpha_m, & \text{if } \ 0 \leq m \leq p; \\ 0, & \text{if } \quad m > p; \end{cases} \qquad \beta_m^* := \begin{cases} \beta_m, & \text{if } \ 0 \leq m \leq q; \\ 0, & \text{if } \quad m > q; \end{cases} \tag{5}$$

and $\delta_{d,j} := \tau_{-d,j}$, for all $j \in \mathbb{N}$, are the coefficients obtained upon replacing $-d$ by d in (1), that is

$$\sum_{k=0}^{\infty}\delta_{d,k}\mathcal{B}^k := \sum_{j=0}^{\infty}\tau_{-d,j}\mathcal{B}^j = (1-\mathcal{B})^d.$$

Let $\theta, \gamma \in \mathbb{R}$ and $\{Z_t\}_{t\in\mathbb{Z}}$ be a sequence of independent and identically distributed (i.i.d.) random variables, with zero mean and variance equal to one. Assume that θ and γ are not both equal to zero and define $\{g(Z_t)\}_{t\in\mathbb{Z}}$ by

$$g(Z_t) = \theta Z_t + \gamma[|Z_t| - \mathbb{E}(|Z_t|)], \qquad \text{for all } t \in \mathbb{Z}. \tag{6}$$

It follows that (see Lopes and Prass, 2014) $\{g(Z_t)\}_{t\in\mathbb{Z}}$ is a strictly stationary and ergodic process. Moreover, since $\mathbb{E}(Z_0^2) < \infty$, $\{g(Z_t)\}_{t\in\mathbb{Z}}$ is also weakly stationary with zero mean (hence a white noise process) and variance σ_g^2 given by

$$\sigma_g^2 = \theta^2 + \gamma^2 - [\gamma\mathbb{E}(|Z_0|)]^2 + 2\,\theta\,\gamma\,\mathbb{E}(Z_0|Z_0|). \tag{7}$$

Now, for any $d < 0.5$ and $\omega \in \mathbb{R}$, let $\{X_t\}_{t\in\mathbb{Z}}$ be the stochastic process defined by

$$X_t = \sigma_t Z_t, \tag{8}$$

$$\begin{aligned}\ln(\sigma_t^2) &= \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})}(1-\mathcal{B})^{-d} g(Z_{t-1}) \\ &= \omega + \sum_{k=0}^{\infty} \lambda_{d,k}\, g(Z_{t-1-k}), \quad \text{for all } t \in \mathbb{Z}.\end{aligned} \tag{9}$$

Then $\{X_t\}_{t\in\mathbb{Z}}$ is a *Fractionally Integrated* EGARCH *process*, denoted by FIEGARCH(p,d,q) (Bollerslev and Mikkelsen, 1996).

The properties of FIEGARCH(p,d,q) processes, with $d < 0.5$, are given below (the proofs of these properties can be found in Lopes and Prass, 2014). Henceforth *GED*(ν) denotes the generalized error distribution with tail thickness parameter ν.

Proposition 1. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a FIEGARCH(p,d,q) process. Then the following properties hold:*

1. *$\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ is a stationary (weakly and strictly) and an ergodic process and the random variable $\ln(\sigma_t^2)$ is almost surely finite, for all $t \in \mathbb{Z}$;*

2. *if $d \in (-1, 0.5)$ and $\alpha(z) \neq 0$, for $|z| \leq 1$, the process $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ is invertible;*

3. *$\{X_t\}_{t\in\mathbb{Z}}$ and $\{\sigma_t^2\}_{t\in\mathbb{Z}}$ are strictly stationary and ergodic processes;*

4. *if $\{Z_t\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. GED(ν) random variables, with $v > 1$, zero mean and variance equal to one, then $\mathbb{E}(X_t^r) < \infty$ and $\mathbb{E}(\sigma_t^{2r}) < \infty$, for all $t \in \mathbb{Z}$ and $r > 0$.*

3 Parameter Estimation: Bayesian Inference using MCMC

Let ν be the parameter (or vector of parameters) associated with the probability density function of Z_0 and denote by

- $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega, \alpha_1, \cdots, \alpha_p, \beta_1, \cdots, \beta_q) := (\eta_1, \eta_2, \cdots, \eta_{5+p+q})$ the vector of unknown parameters in (9);
- $\boldsymbol{\eta}_{(-i)}$ the vector containing all parameters in $\boldsymbol{\eta}$ except η_i, for each $i \in \{1, \cdots, 5+p+q\}$;
- $p_Z(\cdot|\nu)$ the probability density function of Z_0 given ν;
- $\mathcal{F}_t$ the σ-algebra generated by $\{Z_s\}_{s\leq t}$;
- $p_{X_t}(\cdot|\boldsymbol{\eta}, \mathcal{F}_{t-1})$ the probability density function of X_t given $\boldsymbol{\eta}$ and $\mathcal{F}_{t-1}$, for all $t \in \mathbb{Z}$.

From (9) it is evident that, given $\boldsymbol{\eta}$, σ_t is a $\mathcal{F}_{t-1}$-measurable random variable. Moreover, since $X_t = \sigma_t Z_t$ and $p_Z(\cdot|\nu, \mathcal{F}_{t-1}) = p_Z(\cdot|\nu)$, the following equality holds

$$p_{X_t}(x_t|\boldsymbol{\eta}, \mathcal{F}_{t-1}) = \frac{1}{\sigma_t} p_Z\big(x_t \sigma_t^{-1}|\nu\big), \quad \text{with} \quad \sigma_t = \exp\left\{\frac{1}{2}\left[\omega + \sum_{k=0}^{\infty} \lambda_{d,k}\, g(z_{t-1-k})\right]\right\}, \tag{10}$$

for all $x_t \in \mathbb{R}$ and $t \in \mathbb{Z}$. Furthermore, from (10), the conditional probability of $\boldsymbol{X} := (X_1, \cdots, X_n)'$ given $\boldsymbol{\eta}$ and $\mathcal{F}_0$ can be written as

$$
\begin{aligned}
p_{\boldsymbol{X}}\big(x_1, \cdots, x_n|\boldsymbol{\eta}, \mathcal{F}_0\big) &= p_{X_n}(x_n|\boldsymbol{\eta}, x_{n-1}, \cdots, x_1, \mathcal{F}_0) \times \cdots \times p_{X_1}(x_1|\boldsymbol{\eta}, \mathcal{F}_0) \\
&= \prod_{t=1}^{n} \frac{1}{\sigma_t} p_Z\big(x_t \sigma_t^{-1}|\nu\big). \qquad (11)
\end{aligned}
$$

Given any $I_0 \in \mathcal{F}_0$, select a prior conditional density function $p_{I_0}(\cdot|\boldsymbol{\eta})$ for I_0 given $\boldsymbol{\eta}$. Also, select a prior[1] density function $\pi_i(\cdot)$ for η_i and a prior conditional probability density function $p_{(-i)}(\cdot|\eta_i)$ for $\boldsymbol{\eta}_{(-i)}$ given η_i, for each $i \in \{1, \cdots, 5+p+q\}$.

Observe that, by applying the Bayes' rule, the conditional probability density function of η_i given $\boldsymbol{X}$, $\boldsymbol{\eta}_{(-i)}$ and any I_0, can be written as

$$
p\big(\eta_i|\boldsymbol{X}, \boldsymbol{\eta}_{(-i)}, I_0\big) \propto p_{\boldsymbol{X}}\big(\boldsymbol{X}|\boldsymbol{\eta}, I_0\big) \times p_{I_0}\big(I_0|\boldsymbol{\eta}\big) \times p_{(-i)}(\boldsymbol{\eta}_{(-i)}|\eta_i) \times \pi_i(\eta_i), \qquad (12)
$$

for each $i \in \{1, \cdots, 5+p+q\}$, where $p_{\boldsymbol{X}}\big(\cdot\,|\boldsymbol{\eta}, \mathcal{F}_0\big)$ is given in (11).

The parameter estimation is carried out by considering an MCMC procedure which is a combination of the Gibbs sampling (Geman and Geman, 1984; Gelfand and Smith, 1990) and Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970) algorithms. The main advantage of the Gibbs sampler is that it simplifies a complex high-dimensional problem by breaking it down into simple, low-dimensional problems. However, the algorithm assumes that the conditional distribution of each random variable is known and it is easy to sample from it. For the problem considered in this work, it is not possible to sample directly from $p\big(\eta_i|\boldsymbol{X}, \boldsymbol{\eta}_{(-i)}, I_0\big)$, for any $i \in \{1, \cdots, 5+p+q\}$, and hence the Metropolis-Hastings algorithm is considered instead. The procedure adopted here is known in the literature as *MH-whithing-Gibbs* or *Gibbs sampler with Metropolis steps*.

4 Simulation Study

Section 2 defines a FIEGARCH(p, d, q) process, for any $p, q \geq 0$. However, in this simulation study, the performance of the MCMC procedure is analyzed only for FIEGARCH$(0, d, 0)$ processes. Under the assumption that $p = q = 0$, one has $\lambda_{d,k} = \pi_{d,k}$, for all $k \in \mathbb{Z}$, and (9) becomes

$$
\ln(\sigma_t^2) = \omega + (1 - \mathcal{B})^{-d} g(Z_{t-1}) = \omega + \sum_{k=0}^{\infty} \pi_{d,k} g(Z_{t-1-k}), \quad \text{for all } k \in \mathbb{Z}.
$$

Therefore, the vector of unknown parameters is $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)' := (\eta_1, \cdots, \eta_5)'$.

Remark 1. Notice that, for the general case $p, q > 0$, the polynomials $\alpha(\cdot)$ and $\beta(\cdot)$, given in (9), cannot have common roots and also, $\beta(z)$ and $\alpha(z)$ must be different from zero whenever $|z| \leq 1$ (when one wishes the invertibility property for the process). These conditions must be incorporated in the expressions of the priors increasing the complexity and the computational cost of the problem. Therefore, the case $p, q > 0$ shall be discussed in a future work.

The Bayesian inference approach, using MCMC to obtain posterior density functions, is used to estimate the parameters of the model. All algorithms were implemented in FORTRAN

[1] In fact, the priors $\pi_i(\cdot)$ are not necessarily probability density functions. For instance, $\pi(x) = 1$ and $\pi(x) = 1/x$, are examples of improper priors (i.e., they do not integrate to 1) used in practice.

95 language but can be converted to other programming languages such as C, R and SPlus. The authors have not yet implemented the code for FIEGARCH models in more standard MCMC softwares such as winBUGS or openBUGS. However, the implementation should follow similar steps as in Meyer and Yu (2000), where stochastic volatility (SV) models were considered.

The authors have also considered using the so-called ARMS (Adaptive Rejection Metropolis Sampling) within Gibbs algorithm (see Gilks et al., 1995; Gelfand and Smith, 1990; Casela and George, 1992; Smith and Roberts, 1993; Chib and Greenberg, 1995) instead of MH-whithing-Gibbs with the truncated normal proposal distribution. ARMS automatically adapts to the full conditional posterior density without the need to specify tuning parameters such as μ and σ, defined in (13). However, this algorithm still depends on the specification of two values: a number n_1 of points to create the initial envelope and a number n_m of maximum points in the envelope. When applying the ARMS to sample from the posterior distribution of the parameter in the FIEGARCH model the following was observed.

- For a time series with small size, say for instance $n = 500$ observations, the values n_1 could be selected independent of the values of the parameters used to generate the FIEGARCH sample. In this case, the algorithm runs smoothly. However, for FIEGARCH processes, the parameter estimation requires higher sample sizes given the characteristic of long-range dependence in the volatility.

- When considering a larger sample size, say $n = 1000$ (or higher), the algorithm became very unstable because the log-likelihood values $\ln(p(\eta_i|\boldsymbol{X}, \boldsymbol{\eta}_{(-i)}, I_0))$ are very large, specially when one of the parameters in $\boldsymbol{\eta}_{(-i)}$ is distant from the true parameter value used to generate $\boldsymbol{X}$.

- Although one could use an artifact such as multiplying the posterior log-likelihood by a normalizing constant, the authors could not find a value that could be globally used for all posteriors, regardless the parameter η_i in question and the combination of ν and d used to generate the FIEGARCH time series. To illustrate the variation on the log-likelihood values, Figures 1 (a) and (b) show the graphs of $f_1(\cdot, \cdot)$ defined by

$$f_1(d, \omega) := \ln\big(p_{\boldsymbol{X}}\big(\boldsymbol{X}|\nu = \nu_0, d, \theta = -0.15, \gamma = 0.24, \omega, I_0 = \{g(Z_t) = 0; t \leq 0\}\big)\big)$$

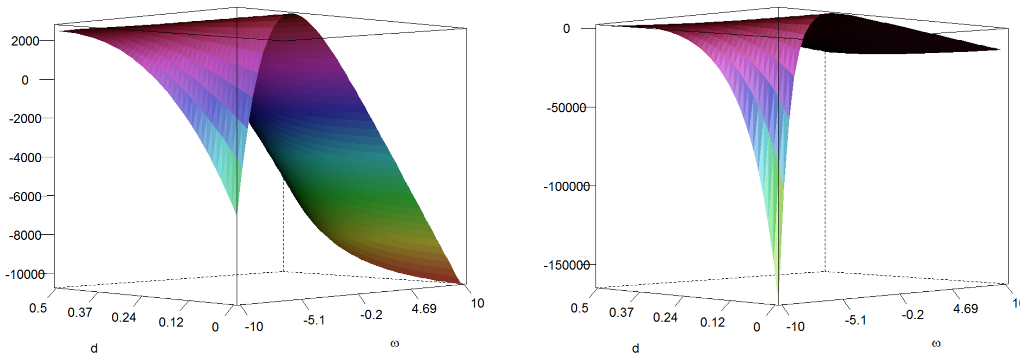
  with $\nu_0 = 1.1$ and $\nu_0 = 2.5$, respectively. The time series $\{X_t\}_{t=1}^{2000}$ were generated with $d = 0.45$, $\theta = -0.15$, $\gamma = 0.24$, $\omega = -5.4$ and $Z_0 \sim \mathrm{GED}(\nu_0)$, with $\nu_0 = 1.1$ and $\nu_0 = 2.5$, respectively, in Figures 1 (a) and (b).

- Sometimes increasing/decreasing the number of initial points in the envelope solved the instability for one model but made it worse for the others. For this simulation study we consider 4 different values of d and 5 different values of ν (totaling 20 different models) and overdispersed starting points for each model. Under these conditions a value n_1 that could be used for all models was not found. The authors concluded that if one analyzes each case separately, the algorithm could run smoothly. However this would require too much time to finish the simulations.

4.1 Data Generating Process

The samples from FIEGARCH$(0, d, 0)$ processes are obtained by setting the following:

- all time series were generated with sample sizes $n = 5000$.



(a) **(b)**

Figure 1: Graphs of $f_1(d,\omega) := \ln\big(p_{\boldsymbol{X}}\big(\boldsymbol{X}|\nu = \nu_0, d, \theta = -0.15, \gamma = 0.24, \omega, I_0 = \{g(Z_t) = 0; t \leq 0\}\big)\big)$. The time series $\{X_t\}_{t=1}^{2000}$ was generated with $d = 0.45$, $\theta = -0.15$, $\gamma = 0.24$, $\omega = -5.4$ and $Z_0 \sim \mathrm{GED}(\nu_0)$, with (a) $\nu_0 = 1.1$; (b) $\nu_0 = 2.5$.

- $Z_0 \sim \mathrm{GED}(\nu)$, with zero mean and variance equal to one. Thus,

$$p_Z(z|\nu) = \frac{\nu \exp\big\{-\frac{1}{2}|z\lambda_\nu^{-1}|^\nu\big\}}{\lambda_\nu 2^{1+1/\nu}\Gamma(1/\nu)}, \quad \lambda_\nu = \left[2^{-2/\nu}\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}\right]^{1/2}, \qquad \text{for all } z \in \mathbb{R};$$

- $d \in \{0.10,\ 0.25,\ 0.35,\ 0.45\}$ and $\nu \in \{1.1, 1.5, 1.9, 2.5, 5\}$;
- for all models, $\omega = -5.40$, $\theta = -0.15$ and $\gamma = 0.24$. These values are close to the ones already observed in practical applications (see, for instance, Nelson, 1991; Bollerslev and Mikkelsen, 1996; Ruiz and Veiga, 2008; Lopes and Prass, 2014).
- the infinite sum in (9) is truncated at $m^* = 50,000$.

For each combination of d and ν, a sample $\{z_t\}_{t=-m^*}^{n}$, of size $m^* + n + 1$, is drawn from the $\mathrm{GED}(\nu)$ distribution and then the sample $\{x_t\}_{t=1}^{n}$, from the FIEGARCH$(0,d,0)$ process, is obtained through the relation

$$\ln(\sigma_t^2) = \omega + \sum_{k=0}^{m^*}\lambda_{d,k}\, g(z_{t-1-k}) \quad \text{and} \quad x_t = \sigma_t z_t, \quad \text{for all } t = 1, \cdots, n.$$

4.2 Parameter Estimation Settings

The samples from the posterior distributions are obtained by considering the MH-within-Gibbs algorithm. The estimation was performed by considering both the entire time series $\{x_t\}_{t=1}^{n}$, with $n = 5000$, and a sub-sample of size 2000.

The transition kernel $q(\cdot|\cdot)$ considered in the Metropolis-Hastings algorithm is the function defined as

$$q(x|y) = f(x; y, \sigma, a, b),$$

where $f(\cdot;\cdot,\cdot,\cdot,\cdot)$ is the truncated normal density function, defined as

$$f(x;\mu,\sigma,a,b) = \begin{cases} \dfrac{1}{\sigma}\dfrac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}, & \text{if } a \leq x \leq b, \\ 0, & \text{otherwise,} \end{cases} \tag{13}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the probability density and cumulative distribution functions of the standard normal distribution; $a, b \in \mathbb{R}$ are, respectively, the lower and upper limits of the distribution's support; μ and σ denote, respectively, the distribution's (non-truncated version).

To avoid using overdispersed points so the computational time could be reduced, a reasonable $\boldsymbol{\eta}^{(0)}$ was selected by calculating $p_{\boldsymbol{X}}\big(\boldsymbol{X}|\boldsymbol{\eta},\mathcal{F}_0\big)$ for different combinations (180 in total) of ν, d, θ, γ and ω. Then $\boldsymbol{\eta}^{(0)}$ is defined as the vector $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)'$ with the highest likelihood function value. A chain of length 10000 was generated and, to eliminate any dependence on the initial $\boldsymbol{\eta}^{(0)}$, a burn-in of size $B = \max\{B_1, B_2, B_3, B_4, B_5\}$ was considered, where B_i is the burn-in size for parameter η_i, for $i \in \{1, 2, 3, 4, 5\}$, suggested by the Heidelberg and Welch (Heidelberger and Welch, 1983) diagnostic criteria.

Remark 2. The minimum length of the pilot run suggested by Raftery-Lewis diagnostic (Raftery and Lewis, 1992) test was 3746. To perform the test we consider $q = 0.025$ as quantile of interest, $r = 0.005$ as the desired margin of error of the estimate, $s = 0.95$ as the probability of obtaining an estimate in the interval $(q - r, q + r)$ and $converge.eps = 0.001$ as the precision required for estimate of time to convergence. Since the Raftery-Lewis diagnostic is usually over-conservative the burn-in and final sample size for the chain were selected taking into account the results of the Heidelberg and Welch test.

Table 1: Information available in practice for the parameter η_i in $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)' := (\eta_1, \cdots, \eta_5)'$ and the corresponding considered prior, for each $i \in \{1, \cdots, 5\}$.

Information Available	**Prior**
The generalized error distribution is well defined for any $\nu > 0$.	$\nu \sim \mathbb{I}_{(0,\infty)}(\nu)$ *
Long-memory in volatility is observed if and only if $d \in (0, 0.5)$. This characteristic can be detected, for instance, through the periodogram function of the time series $\{\ln(X_t^2)\}_{t=1}^n$ (see Lopes and Prass, 2014).	$d \sim \mathcal{U}(0, 0.5)$
Empirical evidence suggests that $\theta \in [-1, 0]$. **	$\theta \sim \mathcal{U}(-1, 0)$
Empirical evidence suggests that $\gamma \in [0, 1]$. **	$\gamma \sim \mathcal{U}(0, 1)$
$\omega = \mathbb{E}(\ln(h_t^2)) = \mathbb{E}(\ln(X_t^2)) + \mathbb{E}(\ln(Z_t^2))$. The choice of the interval for ω will depend on the magnitude of the data. The sample mean of $\{\ln(X_t^2)\}_{t=1}^n$ or $\ln(\hat{\sigma}_X^2)$, where $\hat{\sigma}_X^2$ is the sample variance of $\{X_t\}_{t=1}^n$, can be used to obtain a raw approximation for ω.	$\omega \sim \mathcal{U}(-15, 15)$

Notes: * Given $A \subset \mathbb{R}$, the symbol $\mathbb{I}_A(x)$ denotes the improper prior defined as 1, if $x \in A$, and 0, if $x \notin A$.
** See, for instance, Nelson (1991); Bollerslev and Mikkelsen (1996); Ruiz and Veiga (2008); Lopes and Prass (2014). To the best of our knowledge, a FIEGARCH model for which θ or γ are not in the intervals, respectively, $[-1, 0]$ and $[0, 1]$ has never been reported in the literature.

Table 2: Parameters of the truncated normal distribution (transition kernel) considered, at iteration m of the Gibbs sampler, to obtain the sample from the posterior distribution of the parameter η_i in $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)' := (\eta_1, \cdots, \eta_5)'$, for each $i \in \{1, \cdots, 5\}$.

Parameter	ν	d	θ	γ	ω
Mean (y)	$\nu^{(m-1)}$	$d^{(m-1)}$	$\theta^{(m-1)}$	$\gamma^{(m-1)}$	$\omega^{(m-1)}$
Standard Deviation (σ)	0.500	0.025	0.050	0.050	1.500
Lower Limit (a)	0.000	0.000	-1.000	0.000	-15.000
Upper Limit (b)	10.000	0.500	0.000	1.000	15.000

Note: $\eta_i^{(m-1)}$, for any $i \in \{1, \cdots, 5\}$, denotes the parameter value obtained in the $(m-1)$th iteration. Different combinations of standard deviation, lower and upper limits were tested for the parameter η_i in $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)'$, for each $i \in \{1, \cdots, 5\}$. The values presented here correspond to the final choice.

The prior distributions for ν, d, θ, γ and ω are selected by considering only the basic set of information usually available in practice. The information on each parameter and the corresponding prior selected are given in Table 1. Table 2 presents the mean, standard deviation, lower and upper limits for the transition kernel considered at iteration m of the Gibbs sampler with Metropolis steps, when the prior for η_i in $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)' := (\eta_1, \cdots, \eta_5)'$, for each $i \in \{1, \cdots, 5\}$, is defined as in Table 1.

Since the conditional probability density function of I_0 given $\boldsymbol{\eta}$ is difficult to obtain, in all scenarios, it is assumed that $g(Z_s) = 0$, for all $s \leq 1$, and it is fixed $p_{I_0}(\cdot|\boldsymbol{\eta}) = 1$. Moreover, since (9) is well defined regardless of the relation among the parameters of the model, it is assumed that

$$p_{(-i)}(\boldsymbol{\eta}_{(-i)}|\eta_i) \propto \prod_{j \neq i} \pi_j(\eta_j), \quad \text{for any } i \in \{1, \cdots, 5\}.$$

4.3 Estimates and Performance Measures

Let $\{\eta_i^{(k)}\}_{k=1}^{M}$ be a sample of size M from the posteriori distribution of η_i in $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)' := (\eta_1, \cdots, \eta_5)'$, for any $i \in \{1, \cdots, 5\}$. Denote by $\bar{\eta}_i$ and sd_{η_i}, respectively, the sample mean and standard deviation of $\{\eta_i^{(k)}\}_{k=1}^{M}$, namely,

$$\bar{\eta}_i = \frac{1}{M} \sum_{k=1}^{M} \eta_i^{(k)} \quad \text{and} \quad \mathrm{sd}_{\eta_i} = \sqrt{\frac{1}{M} \sum_{k=1}^{M} (\eta_i^{(k)} - \bar{\eta}_i)^2}, \quad \text{for any} \quad i \in \{1, \cdots, 5\}.$$

Then the estimate $\hat{\eta}_i$ of η_i is defined as $\hat{\eta}_i := \bar{\eta}_i$.

Moreover, let $\hat{q}_i(\alpha)$ denote the α quantile[2] for the posterior sample distribution of η_i, for any $\alpha \in [0, 1]$ and $i \in \{1, \cdots, 5\}$. Then a $100(1-\alpha)\%$ credibility interval for η_i is given by

$$CI_{1-\alpha}(\eta_i) = \left[\hat{q}_i\left(\frac{\alpha}{2}\right), \hat{q}_i\left(1 - \frac{\alpha}{2}\right)\right], \quad \text{for any} \quad i \in \{1, \cdots, 5\}.$$

Furthermore, the estimation bias is given by

$$\mathrm{bias}_{\eta_i} = \bar{\eta}_i - \eta_i, \quad \text{for any} \quad i \in \{1, \cdots, 5\}.$$

[2]In this work, the following definition is adopted (Brockwell and Davis, 1991). Given any $0 \leq \alpha \leq 1$, the number $q(\alpha)$ satisfying $\mathbb{P}(X \leq q(\alpha)) \geq \alpha$ and $\mathbb{P}(X \geq q(\alpha)) \geq 1-\alpha$, is called a quantile of order α (or α quantile) for the random variable X (or for the distribution function of X).

4.4 Simulation Results for the MCMC Bayesian Procedure

The simulation results for the procedure proposed in Section 3 are described in the sequel.

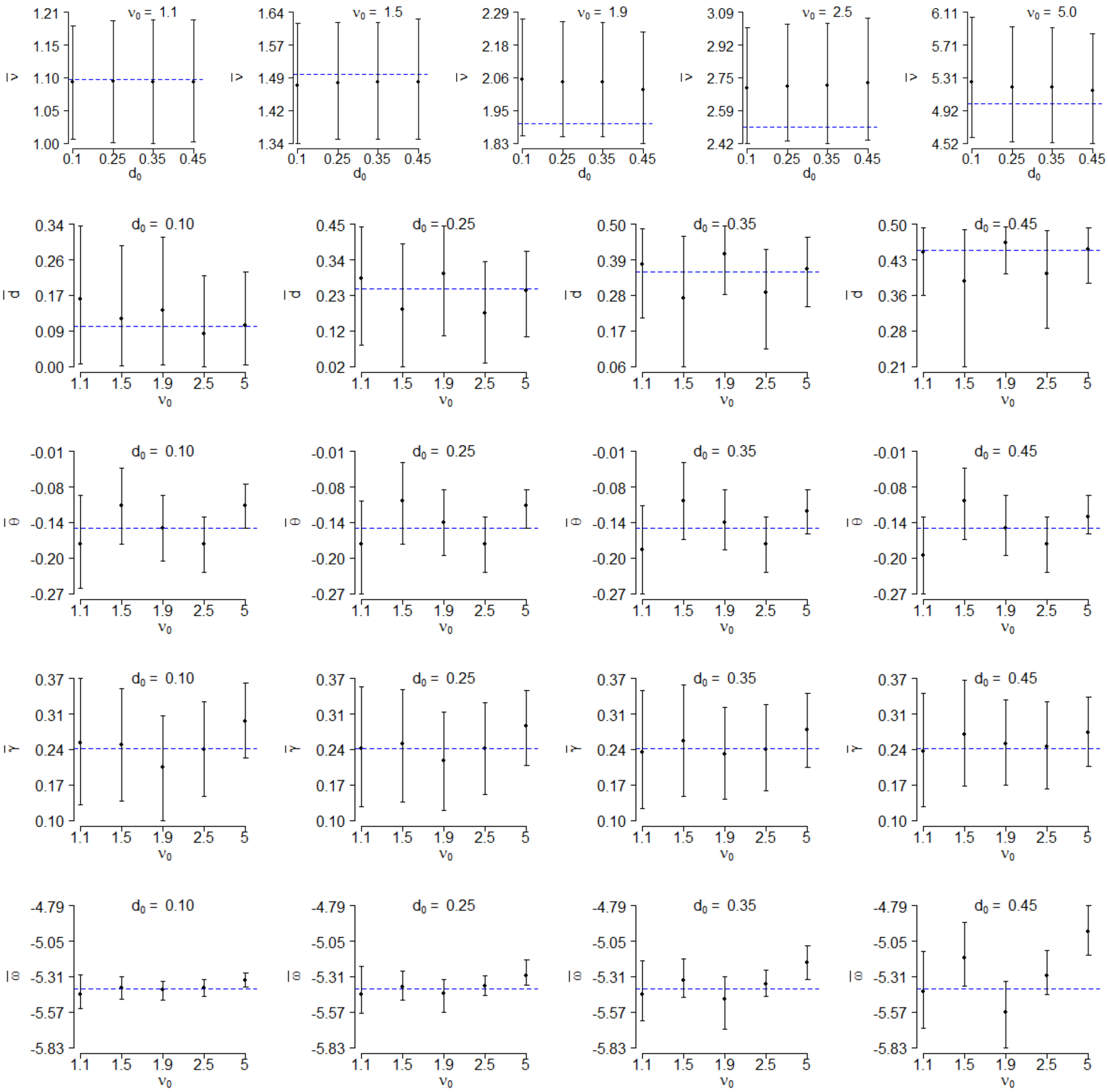


Figure 2: Posterior mean (solid circle), the true parameter value (dashed line) and the 95% credibility interval (solid line) for the parameters ν, d, θ, γ and ω (from top to bottom), for each combination of d_0 and ν_0. The posterior distributions were obtained by considering an improper prior for ν and uniform priors for d, θ, γ and ω. The size of the chains used to obtain the posterior means are given in Table 3. The sample size for the time series $\{X_t\}_{t=1}^{n}$ is $n = 2000$. The true parameters values considered in this simulation are $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$, $\nu_0 \in \{1.1, 1.5, 2.5, 5.0\}$, $\theta_0 = -0.15$, $\gamma_0 = 0.24$ and $\omega_0 = -5.4$.

Figures 2 and 3 show the sample mean (solid circle) and the 95% credibility interval (solid line) for the sample obtained from the posterior distribution of ν, d, θ, γ and ω (respectively, from top to bottom), for each combination of d_0 and ν_0. The true parameter values $\nu_0, d_0, \theta_0, \gamma_0$ and ω_0 are represented in the corresponding row by the dashed line. The graphs related to θ, γ and ω (respectively, the third, fourth and fifth rows, from top to bottom) consider the same

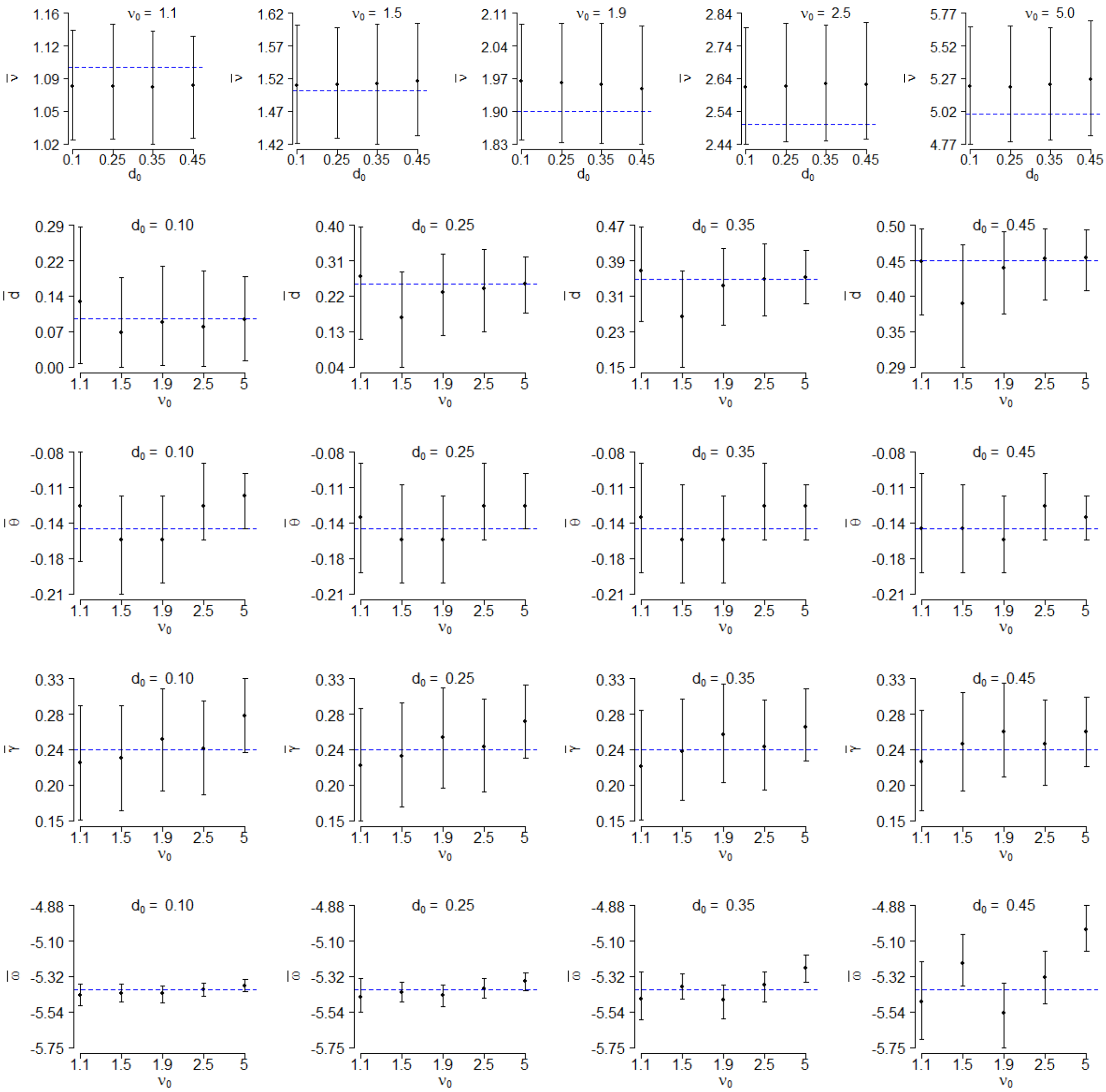


Figure 3: Posterior mean (solid circle), the true parameter value (dashed line) and the 95% credibility interval (solid line) for the parameters ν, d, θ, γ and ω (from top to bottom), for each combination of d_0 and ν_0. The posterior distributions were obtained by considering an improper prior for $\nu$and uniform priors for d, θ, γ and ω. The size of the chains used to obtain the posterior means are given in Table 3. The sample size for the time series $\{X_t\}_{t=1}^{n}$ is $n = 5000$. The true parameters values considered in this simulation are $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$, $\nu_0 \in \{1.1, 1.5, 2.5, 5.0\}$, $\theta_0 = -0.15$, $\gamma_0 = 0.24$ and $\omega_0 = -5.4$.

scale for all $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$. Also, for the parameters θ, γ and ω, there is one graph for each d_0 and, for each one of these graphs, the true value of ν_0 is indicated in the x-axis. The size of the chain used to obtain the posterior means showed in Figures 2 and 3 are given in Table 3. The sample size for the time series $\{X_t\}_{t=1}^{n}$ considered in Figures 2 and 3 are, respectively, $n = 2000$ and 5000. The same results shown in Figures 2 and 3 are reported, respectively, in Tables 11 and 12, given in the Appendix.

Table 3: Burn-in size $B = \max\{B_1, B_2, B_3, B_4, B_5\}$ considered to obtain the posterior mean showed in Figures 2 and 3. The value B_i is the burn-in size for parameter η_i, for $i \in \{1, 2, 3, 4, 5\}$, suggested by the Heidelberg and Welch diagnostic criteria. The corresponding chain size used to obtain the posterior means is $M = 10000 - B$.

ν_0	1.1		1.5		1.9		2.5		5.0	
n	2000	5000	2000	5000	2000	5000	2000	5000	2000	5000
$d_0 = 0.10$	4000	0	0	0	0	0	0	0	0	0
$d_0 = 0.25$	0	0	0	0	0	1000	2000	0	4000	0
$d_0 = 0.35$	0	0	0	4000	0	0	4000	0	0	0
$d_0 = 0.45$	0	2000	0	0	0	0	4000	0	0	0

From Figures 2 and 3 one observes that, as the sample size n increases, parameter estimation improves significantly and the width of the credibility interval decreases. The estimation bias for ν is always negative when $\nu_0 < 1.9$ ($\nu_0 = 1.1$, if $n = 5000$) and positive when $\nu_0 \geq 1.9$ ($\nu_0 > 1.1$, if $n = 5000$), when $n = 2000$. For $(\eta_2, \cdots, \eta_5) = (d, \theta, \gamma, \omega)$ the following pattern is observed (with a few exceptions): for each ν_0 and $i \in \{2, \cdots, 5\}$ fixed, either $\bar{\eta}_i \leq \eta_i$ or $\bar{\eta}_i \geq \eta_i$ for all $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$. For $n = 2000$, the exceptions are $\bar{d}$ when $(d_0, \nu_0) \in \{(0.10, 1.5), (0.45, 1.1)\}$ and $\bar{\gamma}$ when $(d_0, \nu_0) \in \{(0.35, 1.1), (0.45, 1.9)\}$. For $n = 5000$, the exceptions are $\bar{d}$ when $(d_0, \nu_0) \in \{(0.35, 2.5), (0.45, 2.5)\}$ and $\bar{\gamma}$ when $(d_0, \nu_0) \in \{(0.45, 1.5)\}$, $\bar{\omega}$ when $(d_0, \nu_0) \in \{(0.35, 1.5), (0.45, 1.5)\}$. Moreover, let $\zeta(d_0, \nu_0, \eta_i)$ be the value of $\bar{\eta}_i$ when the true parameter values are $d = d_0$ and $\nu = \nu_0$, for any $i \in \{1, \cdots, 5\}$. Then, the following pattern also follows: for any $i \in \{1, \cdots, 4\}$, if $\zeta(0.1, \nu_{(i)}, \eta_i) \leq \zeta(0.1, \nu_{(i+1)}, \eta_i)$, then $\zeta(d_0, \nu_{(i)}, \eta_i) \leq \zeta(d_0, \nu_{(i+1)})$, for all $d_0 \in \{0.25, 0.35, 0.45\}$, where $\nu_{(1)} < \nu_{(2)} < \nu_{(3)} < \nu_{(4)} < \nu_{(5)}$ are the values of ν_0 in ascending order. The same follows when $\zeta(0.1, \nu_{(i)}, \eta_i) \geq \zeta(0.1, \nu_{(i+1)}, \eta_i)$, for any $i \in \{1, \cdots, 4\}$,

When $n = 2000$ (see Figure 2), the credibility intervals $CI_{0.95}(\nu)$, $CI_{0.95}(d)$ and $CI_{0.95}(\gamma)$ contain the true parameter values (respectively, ν_0, d_0 and γ_0) in all cases. The true parameter value $\theta_0 = -0.15$ is contained in all credibility intervals $CI_{0.95}(\theta)$, except when $\nu_0 = 5.0$ and $d_0 = 0.10$. For the parameter ω, the true parameter value $\omega_0 = -5.4$ is not contained in $CI_{0.95}(\omega)$ for the following combinations of ν_0 and d_0: $(1.5, 0.45)$ and $(5.0, d_0)$, for all $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$. When $n = 5000$ all parameters all contained in their respective credibility intervals, except ω_0 when $(\nu_0, d_0) \in \{(5.0, 0.35), (1.5, 0.45), (5.0, 0.45)\}$.

It is our believe that some credibility intervals failed to include the true parameter values due to the fact that the variation on the log-likelihood values in the region around the true parameter value (θ_0 or ω_0) is too small. To illustrate the behavior of the log-likelihood function, Figures 4 (a) - (d) show the graphs of $f_2(\cdot, \cdot)$ defined by

$$f_2(\theta, \omega) := \ln\big(p_{\boldsymbol{X}}\big(\boldsymbol{X} | \nu = 5.0, d = d_0, \theta, \gamma = 0.24, \omega, I_0 = \{g(Z_t) = 0; t \leq 0\}\big)\big)$$

with $d_0 = 0.10, 0.25, 0.35$ and 0.45, respectively. The time series $\{X_t\}_{t=1}^{2000}$ were generated with $\theta = -0.15$, $\gamma = 0.24$, $\omega = -5.4$, $Z_0 \sim \text{GED}(5.0)$, and $d_0 = 0.10, 0.25, 0.35$ and 0.45, respectively, in Figures 4 (a) - (d).

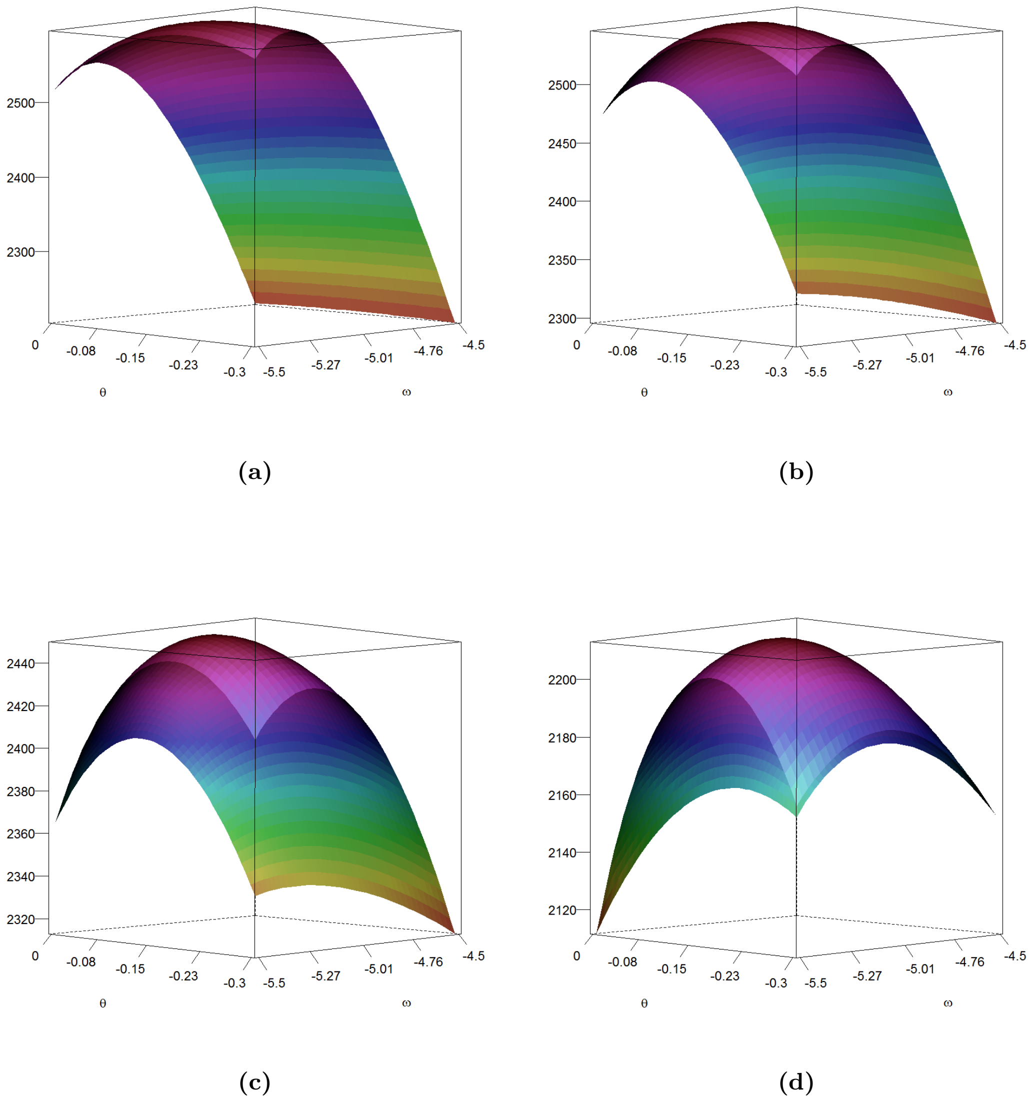


Figure 4: Graphs of $f_2(\theta,\omega) := \ln\big(p_{\boldsymbol{X}}\big(\boldsymbol{X}|\nu = 5.0, d = d_0, \theta, \gamma = 0.24, \omega, I_0 = \{g(Z_t) = 0; t \leq 0\}\big)\big)$. The time series $\{X_t\}_{t=1}^{2000}$ was generated with $d = 0.45$, $\theta = -0.15$, $\gamma = 0.24$, $\omega = -5.4$ and $Z_0 \sim \text{GED}(5.0)$, with (a) $d_0 = 0.10$; (b) $d_0 = 0.25$; (c) $d_0 = 0.35$; (d) $d_0 = 0.45$.

4.5 A Comparison Between the QML and MCMC Bayesian Procedures

Following the same steps as in Lopes and Prass (2014), we have also conduct a simulation study to analyze the finite sample performance of the quasi-maximum likelihood (QML) procedure on parameter estimation. All samples were generated as described in Section 4.1. For each model 1000 replications were considered. For each replication, the parameter estimation was carried out by considering both the entire time series $\{x_t\}_{t=1}^{n}$, with $n = 5000$, and a sub-sample of size 2000.

The simulation results considering QML procedure are showed in Tables 4-8. Statistics reported, based on 1000 replications, are, respectively, the mean ($\bar{\eta}$), the standard deviation (sd), the bias ($bias$), the mean absolute error (mae) and the mean square error (mse) of the estimates. In these tables *Neval* denotes the mean for the number of function evaluations required to achieve the convergence in the algorithm.

Table 4: Estimation results for the simulated FIEGARCH models considering the QML procedure. The FIEGARCH time series were obtained by considering $\nu = 1.1$, $d \in \{0.10, 0.25, 0.35, 0.45\}$, $\theta = -0.15$, $\gamma = 0.24$ and $\omega = -5.4$. The statistics reported are based on 1000 replications. *Neval* denotes the mean for the number of function evaluations required to achieve the algorithm convergence.

	$n = 2000$						$n = 5000$					
η	$\bar{\eta}$	*sd*	*bias*	*mae*	*mse*	*Neval*	$\bar{\eta}$	*sd*	*bias*	*mae*	*mse*	*Neval*
d = 0.1000	0.0449	0.1907	-0.0551	0.1403	0.0394	101.6070	0.0841	0.0966	-0.0159	0.0762	0.0096	105.7680
θ = -0.1500	-0.1519	0.0483	-0.0019	0.0384	0.0023	101.6070	-0.1509	0.0302	-0.0009	0.0240	0.0009	105.7680
γ = 0.2400	0.2307	0.0684	-0.0093	0.0544	0.0048	101.6070	0.2358	0.0419	-0.0042	0.0334	0.0018	105.7680
ω = -5.4000	-5.3647	0.0684	0.0353	0.0601	0.0059	101.6070	-5.3570	0.0472	0.0430	0.0510	0.0041	105.7680
d = 0.2500	0.1959	0.1357	-0.0541	0.1047	0.0213	108.8400	0.2213	0.0685	-0.0287	0.0566	0.0055	118.6380
θ = -0.1500	-0.1538	0.0468	-0.0038	0.0376	0.0022	108.8400	-0.1523	0.0294	-0.0023	0.0234	0.0009	118.6380
γ = 0.2400	0.2308	0.0675	-0.0092	0.0537	0.0046	108.8400	0.2360	0.0412	-0.0040	0.0327	0.0017	118.6380
ω = -5.4000	-5.3162	0.0997	0.0838	0.1036	0.0170	108.8400	-5.2861	0.0762	0.1139	0.1171	0.0188	118.6380
d = 0.3500	0.2944	0.0996	-0.0556	0.0868	0.0130	118.6400	0.3054	0.0534	-0.0446	0.0549	0.0048	128.6600
θ = -0.1500	-0.1563	0.0449	-0.0063	0.0364	0.0021	118.6400	-0.1552	0.0283	-0.0052	0.0229	0.0008	128.6600
γ = 0.2400	0.2310	0.0667	-0.0090	0.0528	0.0045	118.6400	0.2371	0.0403	-0.0029	0.0321	0.0016	128.6600
ω = -5.4000	-5.2534	0.1441	0.1466	0.1669	0.0423	118.6400	-5.1909	0.1187	0.2091	0.2113	0.0578	128.6600
d = 0.4500	0.3864	0.0793	-0.0636	0.0810	0.0103	127.5360	0.3885	0.0446	-0.0615	0.0642	0.0058	137.8490
θ = -0.1500	-0.1602	0.0427	-0.0102	0.0351	0.0019	127.5360	-0.1597	0.0267	-0.0097	0.0228	0.0008	137.8490
γ = 0.2400	0.2321	0.0651	-0.0079	0.0512	0.0043	127.5360	0.2404	0.0391	0.0004	0.0313	0.0015	137.8490
ω = -5.4000	-5.1460	0.2516	0.2540	0.2912	0.1278	127.5360	-5.0217	0.2240	0.3783	0.3847	0.1933	137.8490

Table 5: Estimation results for the simulated FIEGARCH models considering the QML procedure. The FIEGARCH time series were obtained by considering $\nu = 1.5$, $d \in \{0.10, 0.25, 0.35, 0.45\}$, $\theta = -0.15$, $\gamma = 0.24$ and $\omega = -5.4$. The statistics reported are based on 1000 replications. *Neval* denotes the mean for the number of function evaluations required to achieve the algorithm convergence.

	$n = 2000$						$n = 5000$					
η	$\bar{\eta}$	*sd*	*bias*	*mae*	*mse*	*Neval*	$\bar{\eta}$	*sd*	*bias*	*mae*	*mse*	*Neval*
d = 0.1000	0.0730	0.1378	-0.0270	0.1067	0.0197	96.3460	0.0892	0.0775	-0.0108	0.0620	0.0061	99.4470
θ = -0.1500	-0.1530	0.0377	-0.0030	0.0300	0.0014	96.3460	-0.1510	0.0240	-0.0010	0.0194	0.0006	99.4470
γ = 0.2400	0.2346	0.0599	-0.0054	0.0479	0.0036	96.3460	0.2384	0.0374	-0.0016	0.0298	0.0014	99.4470
ω = -5.4000	-5.3860	0.0486	0.0140	0.0395	0.0026	96.3460	-5.3827	0.0317	0.0173	0.0289	0.0013	99.4470
d = 0.2500	0.2206	0.1094	-0.0294	0.0846	0.0128	104.0570	0.2353	0.0578	-0.0147	0.0474	0.0036	112.2100
θ = -0.1500	-0.1543	0.0371	-0.0043	0.0297	0.0014	104.0570	-0.1517	0.0237	-0.0017	0.0193	0.0006	112.2100
γ = 0.2400	0.2344	0.0591	-0.0056	0.0474	0.0035	104.0570	0.2384	0.0366	-0.0016	0.0291	0.0013	112.2100
ω = -5.4000	-5.3622	0.0695	0.0378	0.0616	0.0063	104.0570	-5.3489	0.0494	0.0511	0.0585	0.0051	112.2100
d = 0.3500	0.3200	0.0912	-0.0300	0.0715	0.0092	111.8830	0.3319	0.0478	-0.0181	0.0407	0.0026	121.2610
θ = -0.1500	-0.1557	0.0360	-0.0057	0.0290	0.0013	111.8830	-0.1528	0.0230	-0.0028	0.0188	0.0005	121.2610
γ = 0.2400	0.2347	0.0579	-0.0053	0.0466	0.0034	111.8830	0.2388	0.0355	-0.0012	0.0284	0.0013	121.2610
ω = -5.4000	-5.3290	0.1071	0.0710	0.1010	0.0165	111.8830	-5.2970	0.0835	0.1030	0.1121	0.0176	121.2610
d = 0.4500	0.4213	0.0732	-0.0287	0.0608	0.0062	122.4590	0.4295	0.0410	-0.0205	0.0368	0.0021	132.9790
θ = -0.1500	-0.1571	0.0343	-0.0071	0.0278	0.0012	122.4590	-0.1538	0.0218	-0.0038	0.0179	0.0005	132.9790
γ = 0.2400	0.2354	0.0557	-0.0046	0.0447	0.0031	122.4590	0.2394	0.0339	-0.0006	0.0273	0.0012	132.9790
ω = -5.4000	-5.2698	0.2090	0.1302	0.1953	0.0606	122.4590	-5.2025	0.1793	0.1975	0.2230	0.0711	132.9790

Table 6: Estimation results for the simulated FIEGARCH models considering the QML procedure. The FIEGARCH time series were obtained by considering $\nu = 1.9$, $d \in \{0.10, 0.25, 0.35, 0.45\}$, $\theta = -0.15$, $\gamma = 0.24$ and $\omega = -5.4$. The statistics reported are based on 1000 replications. $Neval$ denotes the mean for the number of function evaluations required to achieve the algorithm convergence.

	$n = 2000$						$n = 5000$					
η	$\bar{\eta}$	sd	$bias$	mae	mse	$Neval$	$\bar{\eta}$	sd	$bias$	mae	mse	$Neval$
d = 0.1000	0.0689	0.1244	-0.0311	0.0965	0.0165	93.3030	0.0881	0.0724	-0.0119	0.0569	0.0054	97.6140
θ = -0.1500	-0.1498	0.0317	0.0002	0.0254	0.0010	93.3030	-0.1502	0.0198	-0.0002	0.0158	0.0004	97.6140
γ = 0.2400	0.2364	0.0521	-0.0036	0.0415	0.0027	93.3030	0.2389	0.0337	-0.0011	0.0268	0.0011	97.6140
ω = -5.4000	-5.3982	0.0382	0.0018	0.0308	0.0015	93.3030	-5.3972	0.0257	0.0028	0.0206	0.0007	97.6140
d = 0.2500	0.2203	0.0980	-0.0297	0.0774	0.0105	100.7190	0.2368	0.0564	-0.0132	0.0454	0.0034	109.3460
θ = -0.1500	-0.1506	0.0311	-0.0006	0.0250	0.0010	100.7190	-0.1506	0.0194	-0.0006	0.0155	0.0004	109.3460
γ = 0.2400	0.2366	0.0515	-0.0034	0.0409	0.0027	100.7190	0.2391	0.0328	-0.0009	0.0260	0.0011	109.3460
ω = -5.4000	-5.3945	0.0525	0.0055	0.0424	0.0028	100.7190	-5.3918	0.0388	0.0082	0.0315	0.0016	109.3460
d = 0.3500	0.3237	0.0839	-0.0263	0.0662	0.0077	107.5320	0.3384	0.0468	-0.0116	0.0381	0.0023	117.9370
θ = -0.1500	-0.1514	0.0302	-0.0014	0.0241	0.0009	107.5320	-0.1510	0.0188	-0.0010	0.0151	0.0004	117.9370
γ = 0.2400	0.2367	0.0503	-0.0033	0.0400	0.0025	107.5320	0.2393	0.0316	-0.0007	0.0250	0.0010	117.9370
ω = -5.4000	-5.3891	0.0840	0.0109	0.0673	0.0072	107.5320	-5.3831	0.0696	0.0169	0.0567	0.0051	117.9370
d = 0.4500	0.4302	0.0684	-0.0198	0.0554	0.0051	117.1190	0.4416	0.0391	-0.0084	0.0318	0.0016	128.6340
θ = -0.1500	-0.1520	0.0287	-0.0020	0.0230	0.0008	117.1190	-0.1511	0.0179	-0.0011	0.0144	0.0003	128.6340
γ = 0.2400	0.2369	0.0482	-0.0031	0.0383	0.0023	117.1190	0.2395	0.0298	-0.0005	0.0236	0.0009	128.6340
ω = -5.4000	-5.3798	0.1795	0.0202	0.1426	0.0326	117.1190	-5.3674	0.1640	0.0326	0.1322	0.0280	128.6340

Table 7: Estimation results for the simulated FIEGARCH models considering the QML procedure. The FIEGARCH time series were obtained by considering $\nu = 2.5$, $d \in \{0.10, 0.25, 0.35, 0.45\}$, $\theta = -0.15$, $\gamma = 0.24$ and $\omega = -5.4$. The statistics reported are based on 1000 replications. $Neval$ denotes the mean for the number of function evaluations required to achieve the algorithm convergence.

	$n = 2000$						$n = 5000$					
η	$\bar{\eta}$	sd	$bias$	mae	mse	$Neval$	$\bar{\eta}$	sd	$bias$	mae	mse	$Neval$
d = 0.1000	0.0774	0.1059	-0.0226	0.0836	0.0117	91.7590	0.0896	0.0669	-0.0104	0.0539	0.0046	96.9160
θ = -0.1500	-0.1515	0.0280	-0.0015	0.0221	0.0008	91.7590	-0.1504	0.0183	-0.0004	0.0147	0.0003	96.9160
γ = 0.2400	0.2349	0.0496	-0.0051	0.0398	0.0025	91.7590	0.2378	0.0324	-0.0022	0.0257	0.0011	96.9160
ω = -5.4000	-5.4109	0.0362	-0.0109	0.0301	0.0014	91.7590	-5.4106	0.0230	-0.0106	0.0202	0.0006	96.9160
d = 0.2500	0.2257	0.0830	-0.0243	0.0658	0.0075	99.0640	0.2385	0.0509	-0.0115	0.0412	0.0027	108.1180
θ = -0.1500	-0.1526	0.0277	-0.0026	0.0219	0.0008	99.0640	-0.1510	0.0180	-0.0010	0.0144	0.0003	108.1180
γ = 0.2400	0.2345	0.0493	-0.0055	0.0394	0.0025	99.0640	0.2377	0.0319	-0.0023	0.0252	0.0010	108.1180
ω = -5.4000	-5.4251	0.0517	-0.0251	0.0454	0.0033	99.0640	-5.4312	0.0359	-0.0312	0.0389	0.0023	108.1180
d = 0.3500	0.3265	0.0700	-0.0235	0.0559	0.0055	106.1320	0.3385	0.0417	-0.0115	0.0342	0.0019	114.9920
θ = -0.1500	-0.1536	0.0271	-0.0036	0.0215	0.0007	106.1320	-0.1517	0.0175	-0.0017	0.0140	0.0003	114.9920
γ = 0.2400	0.2347	0.0481	-0.0053	0.0384	0.0023	106.1320	0.2381	0.0310	-0.0019	0.0245	0.0010	114.9920
ω = -5.4000	-5.4455	0.0841	-0.0455	0.0758	0.0092	106.1320	-5.4641	0.0649	-0.0641	0.0758	0.0083	114.9920
d = 0.4500	0.4293	0.0591	-0.0207	0.0481	0.0039	114.8900	0.4395	0.0349	-0.0105	0.0290	0.0013	125.2540
θ = -0.1500	-0.1547	0.0261	-0.0047	0.0209	0.0007	114.8900	-0.1526	0.0167	-0.0026	0.0135	0.0003	125.2540
γ = 0.2400	0.2357	0.0459	-0.0043	0.0367	0.0021	114.8900	0.2395	0.0295	-0.0005	0.0233	0.0009	125.2540
ω = -5.4000	-5.4806	0.1797	-0.0806	0.1576	0.0388	114.8900	-5.5241	0.1539	-0.1241	0.1633	0.0391	125.2540

Table 8: Estimation results for the simulated FIEGARCH models considering the QML procedure. The FIEGARCH time series were obtained by considering $\nu = 5.0$, $d \in \{0.10, 0.25, 0.35, 0.45\}$, $\theta = -0.15$, $\gamma = 0.24$ and $\omega = -5.4$. The statistics reported are based on 1000 replications. $Neval$ denotes the mean for the number of function evaluations required to achieve the algorithm convergence.

	$n = 2000$						$n = 5000$					
η	$\bar{\eta}$	sd	$bias$	mae	mse	$Neval$	$\bar{\eta}$	sd	$bias$	mae	mse	$Neval$
d = 0.1000	0.0784	0.0939	-0.0216	0.0721	0.0093	90.1900	0.0937	0.0519	-0.0063	0.0408	0.0027	94.4360
θ = -0.1500	-0.1495	0.0238	0.0005	0.0188	0.0006	90.1900	-0.1498	0.0154	0.0002	0.0123	0.0002	94.4360
γ = 0.2400	0.2386	0.0432	-0.0014	0.0346	0.0019	90.1900	0.2390	0.0278	-0.0010	0.0223	0.0008	94.4360
ω = -5.4000	-5.4280	0.0324	-0.0280	0.0345	0.0018	90.1900	-5.4291	0.0211	-0.0291	0.0308	0.0013	94.4360
d = 0.2500	0.2242	0.0691	-0.0258	0.0557	0.0054	98.0300	0.2347	0.0378	-0.0153	0.0321	0.0017	106.5890
θ = -0.1500	-0.1508	0.0233	-0.0008	0.0184	0.0005	98.0300	-0.1509	0.0151	-0.0009	0.0121	0.0002	106.5890
γ = 0.2400	0.2373	0.0430	-0.0027	0.0344	0.0019	98.0300	0.2385	0.0273	-0.0015	0.0218	0.0007	106.5890
ω = -5.4000	-5.4677	0.0501	-0.0677	0.0712	0.0071	98.0300	-5.4830	0.0367	-0.0830	0.0833	0.0082	106.5890
d = 0.3500	0.3184	0.0564	-0.0316	0.0498	0.0042	105.2960	0.3241	0.0312	-0.0259	0.0324	0.0016	114.6520
θ = -0.1500	-0.1529	0.0227	-0.0029	0.0180	0.0005	105.2960	-0.1533	0.0146	-0.0033	0.0119	0.0002	114.6520
γ = 0.2400	0.2366	0.0427	-0.0034	0.0342	0.0018	105.2960	0.2390	0.0269	-0.0010	0.0215	0.0007	114.6520
ω = -5.4000	-5.5223	0.0840	-0.1223	0.1273	0.0220	105.2960	-5.5614	0.0674	-0.1614	0.1619	0.0306	114.6520
d = 0.4500	0.4115	0.0485	-0.0385	0.0495	0.0038	112.9310	0.4134	0.0288	-0.0366	0.0391	0.0022	121.4870
θ = -0.1500	-0.1564	0.0218	-0.0064	0.0178	0.0005	112.9310	-0.1575	0.0140	-0.0075	0.0125	0.0003	121.4870
γ = 0.2400	0.2374	0.0426	-0.0026	0.0337	0.0018	112.9310	0.2421	0.0267	0.0021	0.0213	0.0007	121.4870
ω = -5.4000	-5.6166	0.1777	-0.2166	0.2349	0.0785	112.9310	-5.7034	0.1551	-0.3034	0.3060	0.1161	121.4870

From Tables 4-8 one observes that, as the sample size increases, the estimation improves. For all models, θ and γ are better estimated than the other parameters. In all cases the bias for the parameter d is negative. For ω, the bias is positive when $\nu \leq 1.9$ and negative when $\nu > 1.9$. Except for a few cases, the biases for θ and γ are always negative. In summary, the same fact reported in Lopes and Prass (2014) can be observed here, that is, although the QML presents a relatively good performance when the sample size is 2000 and the estimation improves as the sample size increases, it does so very slowly.

Table 9 presents the values for the $RMSE$ measure, defined by

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sigma_t^2 - \hat{\sigma}_t^2)^2},$$

where $n \in \{2000, 5000\}$ are the sample sizes of the considered FIEGARCH$(0, d, 0)$ time series; $\{\sigma_t^2\}_{t=1}^{n}$ are the volatility values generated by considering the true parameter values $\boldsymbol{\eta} = (\nu, d, \theta, \gamma, \omega)$ and $\{\hat{\sigma}_t^2\}_{t=1}^{n}$ are the smoothed volatility values obtained by considering the estimates $\hat{\boldsymbol{\eta}} = (\hat{\nu}, \hat{d}, \hat{\theta}, \hat{\gamma}, \hat{\omega})$. For the MCMC approach $\hat{\nu}, \hat{d}, \hat{\theta}, \hat{\gamma}, \hat{\omega}$ are the posterior means, obtained by considering a sample of size $M = 10000 - B$, with B given in Table 3. Therefore, there is only one corresponding $RMSE$ value for each model. For the QML method we consider $re = 1000$ replications so that 1000 $RMSE$ values are obtained for each model. In this case, only the minimum, maximum and the mean $RMSE$ values are reported (the mean $RMSE$ value corresponds to the mean over 1000 replications for this statistic).

From Table 9 one observes that the $RMSE$ values for the QML and MCMC procedures are very close to each other, for all models. Moreover, the $RMSE$ values not always decrease as n increases. This result indicates that the improvement in parameter estimation when n increases does not significantly improves the volatility estimation, in terms of $RMSE$.

Table 9: $RMSE$ values for the QML and MCMC procedures, for $n \in \{2000, 5000\}$. For the QML procedure only the minimum (Min), maximum (Max) and the mean $RMSE$ values are reported (the mean is taken over 1000 replications).

		$n = 2000$				$n = 5000$			
ν_0	d_0	QML			MCMC	QML			MCMC
		Min	mean	Max		Min	mean	Max	
1.1	0.10	1.0647	1.0654	1.0663	1.0656	1.0649	1.0654	1.0662	1.0652
	0.25	1.0629	1.0655	1.0690	1.0651	1.0634	1.0655	1.0689	1.0643
	0.35	1.0594	1.0656	1.0752	1.0640	1.0603	1.0657	1.0746	1.0630
	0.45	1.0508	1.0660	1.0931	1.0617	1.0519	1.0662	1.0907	1.0607
1.5	0.10	1.0646	1.0654	1.0663	1.0654	1.0648	1.0654	1.0660	1.0654
	0.25	1.0628	1.0654	1.0687	1.0656	1.0632	1.0655	1.0682	1.0657
	0.35	1.0588	1.0655	1.0740	1.0663	1.0590	1.0656	1.0725	1.0667
	0.45	1.0488	1.0660	1.0880	1.0690	1.0492	1.0661	1.0834	1.0704
1.9	0.10	1.0645	1.0654	1.0661	1.0653	1.0649	1.0654	1.0659	1.0655
	0.25	1.0626	1.0655	1.0686	1.0653	1.0631	1.0654	1.0679	1.0657
	0.35	1.0584	1.0656	1.0728	1.0655	1.0596	1.0656	1.0724	1.0661
	0.45	1.0484	1.0661	1.0865	1.0664	1.0510	1.0660	1.0847	1.0672
2.5	0.10	1.0646	1.0653	1.0662	1.0654	1.0648	1.0654	1.0660	1.0655
	0.25	1.0625	1.0654	1.0689	1.0661	1.0631	1.0654	1.0682	1.0659
	0.35	1.0587	1.0655	1.0742	1.0677	1.0600	1.0655	1.0730	1.0667
	0.45	1.0506	1.0660	1.0887	1.0718	1.0531	1.066	1.0869	1.0691
5.0	0.10	1.0644	1.0653	1.0662	1.0650	1.0647	1.0653	1.0660	1.0654
	0.25	1.0623	1.0653	1.0691	1.0654	1.0630	1.0653	1.0682	1.0661
	0.35	1.0590	1.0653	1.0751	1.0668	1.0593	1.0653	1.0728	1.0681
	0.45	1.0508	1.0655	1.0916	1.0714	1.0502	1.0655	1.0856	1.0743

It is a fact that for the QML procedure the expression for the quasi-likelihood function is simpler since the Gaussian distribution is used instead of the true density (in our case, the GED). Moreover, the results shown in Figures 2 and 3 (see also Tables 11 and 12), in Tables 4-8 and in Table 9 indicate that the MCMC and QML approaches are very competitive. Also, the number of iterations for the QML procedure ($Neval/5$) is really small if compared with the number of iterations (size of the chain) in the MCMC algorithm.

To generate a chain of length 10000, the MCMC procedure takes around 10 minutes, when $n = 2000$, and 40 minutes, if $n = 5000$. For the QML procedure, 1000 replications can be performed in around 5 minutes, if $n = 2000$, and 35 minutes, if $n = 5000$. Although the computational time for the MCMC procedure is much higher than for the QML, it has some advantages such as:

- the possibility of selecting the prior distribution based on a known information from the literature or provided by a specialist;
- standard errors for the parameters tend to be smaller when the Bayesian approach is considered;
- the posterior predictive distribution makes use of the entire posterior distribution of the

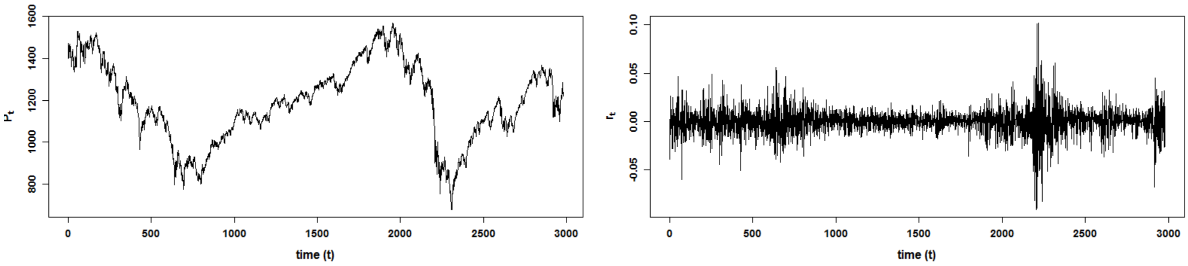
parameter(s) given the observed data to yield a probability distribution over an interval rather than simply a point estimate (as it would do in the frequentist approach);

- when Bayesian approach is considered, a sample from the posterior distribution of each parameter is obtained, rather than a single point, which makes it possible to construct confidence intervals and performing hypothesis test without making further assumptions on the estimates distribution.

5 An Application

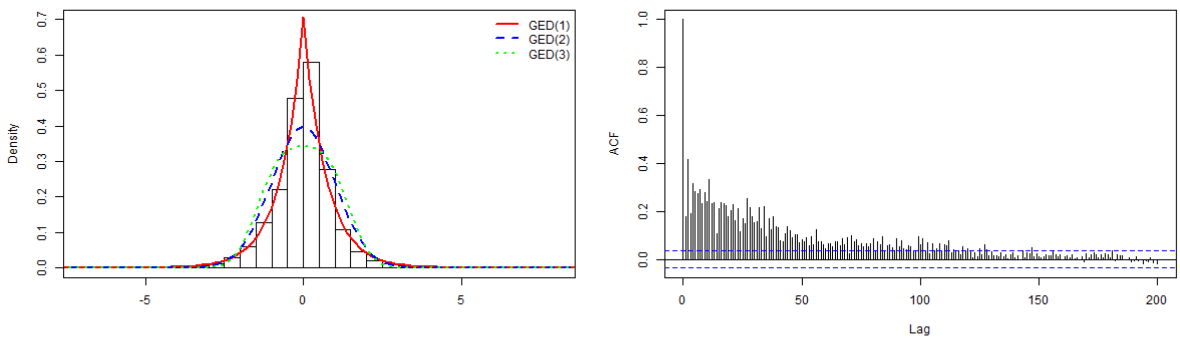
In this section we present an application of the MCMC procedure previously discussed to the daily log-returns of the S&P500 US stock market index in the period from January 03, 2000 to November 03, 2011. This time series was already considered in Pumi and Lopes (2013) were a different methodology was applied to estimate the long-range parameter associated to the absolute log-returns.

The sample size of the index time series is $n = 2980$, which leads to $n = 2979$ log-returns. Figure 5 presents the S&P500 time series and the corresponding log-returns $\{r_t\}_{t=1}^{2979}$.



(a) (b)

Figure 5: S&P500 index time series: (a) original time series; (b) log-return time series.



(a) (b)

Figure 6: (a) Histogram of the log-return time series and the curves of the GED(ν) density, for $\nu = 1, 2$ and 3; (b) autocorrelation function of the squared log-returns time series.

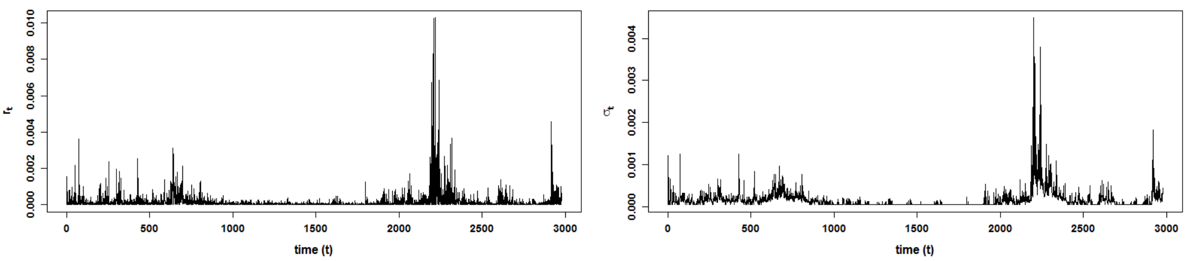
It is well known that log-returns of financial time are usually uncorrelated while the absolute and squared returns are correlated and present slow decay of covariance, characteristic associated to long-range dependence in volatility. Moreover, the distribution of the log-return time series presents heavier tails than the Gaussian one. Figure 6 presents the histogram of the standardized log-returns series and the autocorrelation function of the squared log-return time series. The histogram clearly indicates that the distribution of the log-returns, although symmetric, is not Gaussian ($\nu = 2$). The slow decay of the sample autocorrelation function suggests long-range dependence on the absolute return time series, therefore, in the volatility.

A FIEGARCH$(0, d, 0)$ model, with $Z_0 \sim \text{GED}(\nu)$ was assumed and the MCMC approach was applied to estimate the parameters of the model. The priors and kernel parameters are the same as in Section 3. A chain of size 10000 was generated and the Heidelberg and Welch diagnostic criteria was used to define the burn-in size. The estimation results are shown in Table 10. For this table a burn-in of size 1000 was considered (Heidelberg and Welch diagnostic suggested $B_2 = 0$ and $B_i = 1000$, for $i = 1, 2, 4, 5$) so that the posterior means were obtained by considering a chain of size 9000. The time series $\{\hat{\sigma}_t^2\}_{t=1}^{2979}$, obtained by considering the estimated parameter values $\hat{\boldsymbol{\eta}} = (\hat{\nu}, \hat{d}, \hat{\theta}, \hat{\gamma}, \hat{\omega})$, and $\{r_t^2\}_{t=1}^{2979}$ are presented in Figure 7.

Table 10 shows that the estimate for parameter d is extremely close to the non-stationarity region ($d \geq 0.5$). This results is coherent with the ones reported in the literature. Moreover, $\hat{\nu} = 1.3222$ confirms the hypothesis that the distribution for the log-returns is not Gaussian. Figure 7 shows that the fitted model was able to capture very well the behavior of the volatility, which in practical applications is roughly approximated by the squared log-returns time series.

Table 10: Statistics for the sample obtained from the posterior distribution of the parameters corresponding to the FIEGARCH$(0, d, 0)$ model for the S&P500 log-returns time series. The the lower and upper limits for the 95% credibility interval, the mean and the standard deviation were obtained by considering a chain of size 9000.

Parameter	ν	d	θ	γ	ω
lower limit	1.2415	0.4917	-0.3505	0.2263	-9.9547
mean	1.3222	0.4977	-0.3007	0.2694	-9.8515
upper limit	1.4014	0.4999	-0.2655	0.3178	-9.7530
standard deviation	0.0414	0.0022	0.0218	0.0238	0.0497



(a) (b)

Figure 7: S&P500 index time series: (a) squared log-return time series; (b) smoothed volatility obtained by considering the MCMC estimates for the parameters of the FIEGARCH model.

6 Conclusions

The Bayesian inference approach for parameter estimation on FIEGARCH models was described and a Monte Carlo simulation study was conducted to analyze the performance of the method under the presence of long-range dependence in volatility. The samples from FIEGARCH processes were obtained by considering the infinite sum representation for the logarithm of the volatility. A recurrence formula was used to obtain the coefficients for this representation. The generalized error distribution, with different tail-thickness parameters was considered so both innovation processes with lighter and heavier tails than the Gaussian distribution, were covered.

Markov Chain Monte Carlo (MCMC) methods were used to obtain samples from the posterior distribution of the parameters. A sensitivity analysis was performed by considering the following steps. An improper prior for ν and uniform priors d, θ, γ and ω were selected so only the basic set of information usually available in practice was considered. Two different sample sizes for the FIEGARCH time series were considered: $n = 2000$ and 5000 and the performance of MCMC procedure was compared for these two cases.

The simulation study showed that as the sample size n increases the parameter estimation improves. The true parameters values ν_0, d_0 and γ_o were contained in the 95% credibility interval, for any combination of $\nu_0 \in \{1.1, 1.5, 1.9, 2.5, 5.0\}$ and $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$ considered. The true parameters values γ_0 and ω_0 were not contained in any credibility intervals for some combinations of d_0 and ν_0. The number of combinations for which the credibility interval fails to contain the true parameter values decreases as n increases. It is our believe that the estimation failed in those cases due to the slow variation of the log-likelihood value in the neighborhood of θ_0 and ν_0.

We have also conducted a simulation study to analyze the finite sample performance of the quasi-maximum likelihood (QML) procedure on parameter estimation. The smoothed volatility $\{\hat{\sigma}_t^2\}_{t=1}^n$ calculated upon considering $\hat{\boldsymbol{\eta}} = (\hat{\nu}, \hat{d}, \hat{\theta}, \hat{\gamma}, \hat{\omega})'$ obtained from the MCMC and QML procedures was compared with the simulated volatility $\{\sigma_t^2\}_{t=1}^n$ in order to analyze which method provides better estimates for it. We have concluded that the MCMC and QML approaches are very competitive in terms of parameter estimation. The number of iterations for the QML procedure ($Neval/5$) is really small if compared with the number of iterations (size of the chain) in the MCMC algorithm. However, when Bayesian approach is considered, a sample from the posterior distribution of each parameter is obtained, rather than a single point, which makes it possible to construct confidence intervals and performing hypothesis test without making further assumptions on the estimates distribution.

To illustrate the use of the methodology an application to a real data set was considered. The daily log-returns of the S&P500 US stock market index in the period from January 03, 2000 to November 03, 2011 were analyzed. A FIEGARCH$(0, d, 0)$ model with $Z_0 \sim \text{GEd}(\nu)$ was considered based on the histogram of the log-returns and on the decay of the autocorrelation function of the squared log-returns. The estimated values for d and ν were, respectively, $\hat{d} = 0.4977$ and $\hat{\nu} = 1.3222$, yielding results coherent with the ones found in the literature for the S&P500 log-returns index time series.

Acknowledgments

T.S. Prass was supported by CNPq-Brazil. S.R.C. Lopes was partially supported by CNPq-Brazil, by CAPES-Brazil, by INCT em *Matemática* and by Pronex *Probabilidade e Processos Estocásticos* - E-26/170.008/2008 -APQ1. The authors are grateful to the (Brazilian) National

Center of Super Computing (CESUP-UFRGS) for the computational resources. The authors would like to thank the Editor and two anonymous referees for helpful and constructive comments, which helped to improve the manuscript.

References

Baillie, R.; T. Bollerslev and H.O. Mikkelsen (1996). "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity". *Journal of Econometrics*, vol. 74, 3-30.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity". *Journal of Econometrics*, vol. 31, 307-327.

Bollerslev, T. and H.O. Mikkelsen (1996). "Modeling and Pricing Long Memory in Stock Market Volatility". *Journal of Econometrics*, vol. 73, 151-184.

Breidt, F.; N. Crato and P.J.F. de Lima (1998). "On the Detection and Estimation of Long Memory in Stochastic Volatility". *Journal of Econometrics*, vol. 83, 325-348.

Brockwell, P.J. and R.A. Davis (1991). *Time Series: Theory and Methods.* Second Edition. New York: Springer-Verlag.

Casella, G. and E.I. George (1992). "Explaining the Gibbs Sampler", *The American Statistician*, vol 46(3), 167-174.

Chib, S. and E. Greenberg (1995). "Understanding the Metropolis-Hastings Algorithm", *The American Statistician*, vol. 49(4), 327-335.

Engle, R.F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation". *Econometrica*, vol. 50, 987-1008.

Gelfand, A.E. and A.F.M Smith (1990). "Sampling-Based Approaches to Calculating Marginal Densities". *Journal of the American Statistical Association*, vol. 85(410), 398-409.

Geman, S. and D. Geman (1984). "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, 609-628.

Gelman, A. and D. Rubin (1992). "Inference from Iterative Simulation Using Multiple Sequences". *Statistical Science*, vol. 7, 457-511.

W. R. Gilks; N.G. Best and K.K.C. Tan (1995). "Adaptive Rejection Metropolis Sampling within Gibbs Sampling". *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 44(4), 455-472.

Hastings, W.K. (1970). "Monte Carlo Sampling Methods Using Markov Chains and Their Applications". *Biometrika*, vol. 57(1), 97-109.

Heidelberger, P. and P.D. Welch (1983). "Simulation run length control in the presence of an initial transient". *Operations Research* , vol. 31, 1109-1144.

Lopes, S.R.C and T.S. Prass (2014). "Theoretical Results on Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroskedastic Processes." *Physica A*, vol. 401, 278-307

Metropolis, N.; A.W. Rosenbluth; M.N. Rosenbluth; A.H. Teller and E. Teller (1953). "Equations of State Calculations by Fast Computing Machines". *Journal of Chemical Physics*, vol. 21(6), 1087-1092.

Meyer, R. and T. Yu (2000). "Bugs for a Bayesian Analysis of Stochastic Volatility Models". *Econometrics Journal*, vol. 3, 198-215.

Nelson, D.B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach". *Econometrica*, vol. 59, 347-370.

Pumi, G. and Lopes, S.R.C. (2013a). "Copulas Related to Piecewise Monotone Functions of the Interval and Associated Processes". arXiv:1305.5232

Raftery, A.E. and S.M. Lewis (1992). "One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo". *Statistical Science*, vol. 7, 493-497.

Ruiz E. and H. Veiga (2008). "Modelling long-memory volatilities with leverage effect: A-LMSV versus FIEGARCH". *Computational Statistics and Data Analysis*, vol. 52(6), 2846-2862.

Smith, A.F.M. and G.O. Roberts (1993). "Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods". *Journal of the Royal Statistical Society*, Ser. B, vol. 55, 3-23.

Appendix

This appendix contains extra tables which help to illustrate the results discussed in the text.

Table 11: Summary for the sample obtained from posterior distributions considering an improper prior for ν and uniform priors for the remaining parameters. The statistics are based on a chain of size $10000 - B$, were $B = \max\{B_1, B_2, B_3, B_4, B_5\}$ and B_i is the burn-in size for parameter η_i, for $i \in \{1,2,3,4,5\}$, suggested by the Heidelberg and Welch diagnostic criteria. The true parameter values considered in this simulation are $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$, $\nu_0 \in \{1.1, 1.5, 2.5, 5.0\}$, $\theta_0 = -0.15$, $\gamma_0 = -0.24$ and $\omega_0 = -5.4$. The sample size for the FIEGARCH time series is $n = 2000$.

d_0	ν_0	B	$\bar{\nu}$ (sd$_\nu$) $CI_{0.95}(\nu)$	$\bar{d}$ (sd$_d$) $CI_{0.95}(d)$	$\bar{\theta}$ (sd$_\theta$) $CI_{0.95}(\theta)$	$\bar{\gamma}$ (sd$_\gamma$) $CI_{0.95}(\gamma)$	$\bar{\omega}$ (sd$_\omega$) $CI_{0.95}(\omega)$
0.10	1.1	4000	1.096 (0.048) [1.007; 1.185]	0.164 (0.094) [0.010; 0.340]	-0.178 (0.042) [-0.260; -0.094]	0.251 (0.061) [0.134; 0.373]	-5.446 (0.067) [-5.544; -5.299]
	1.5	0	1.475 (0.069) [1.345; 1.615]	0.117 (0.081) [0.005; 0.291]	-0.106 (0.037) [-0.178; -0.035]	0.248 (0.054) [0.141; 0.354]	-5.399 (0.045) [-5.481; -5.314]
	1.9	0	2.055 (0.106) [1.859; 2.270]	0.137 (0.084) [0.008; 0.312]	-0.146 (0.031) [-0.207; -0.087]	0.205 (0.050) [0.104; 0.303]	-5.413 (0.037) [-5.485; -5.345]
	2.5	0	2.701 (0.150) [2.420; 3.005]	0.081 (0.059) [0.003; 0.220]	-0.181 (0.027) [-0.234; -0.130]	0.239 (0.045) [0.150; 0.328]	-5.396 (0.032) [-5.456; -5.335]
	5.0	0	5.260 (0.380) [4.587; 6.049]	0.101 (0.061) [0.007; 0.229]	-0.110 (0.020) **[-0.148; -0.072]**	0.292 (0.036) [0.222; 0.364]	-5.341 (0.029) **[-5.391; -5.285]**
0.25	1.1	0	1.098 (0.049) [1.001; 1.193]	0.283 (0.090) [0.082; 0.437]	-0.182 (0.041) [-0.267; -0.102]	0.241 (0.059) [0.131; 0.358]	-5.447 (0.084) [-5.581; -5.240]
	1.5	0	1.481 (0.068) [1.355; 1.617]	0.187 (0.102) [0.016; 0.387]	-0.103 (0.036) [-0.175; -0.034]	0.249 (0.054) [0.139; 0.352]	-5.388 (0.054) [-5.486; -5.276]
	1.9	0	2.048 (0.105) [1.855; 2.261]	0.295 (0.084) [0.110; 0.441]	-0.140 (0.032) [-0.205; -0.080]	0.218 (0.048) [0.124; 0.310]	-5.440 (0.061) [-5.578; -5.332]
	2.5	2000	2.710 (0.151) [2.430; 3.025]	0.177 (0.080) [0.028; 0.332]	-0.180 (0.027) [-0.233; -0.128]	0.241 (0.043) [0.154; 0.327]	-5.386 (0.038) [-5.454; -5.307]
	5.0	4000	5.199 (0.360) [4.537; 5.937]	0.245 (0.064) [0.105; 0.363]	-0.113 (0.019) [-0.152; -0.076]	0.283 (0.036) [0.209; 0.351]	-5.306 (0.042) **[-5.379; -5.193]**
0.35	1.1	0	1.096 (0.048) [1.000; 1.194]	0.374 (0.069) [0.211; 0.484]	-0.186 (0.039) [-0.267; -0.109]	0.234 (0.057) [0.126; 0.351]	-5.441 (0.112) [-5.638; -5.199]
	1.5	0	1.482 (0.068) [1.355; 1.618]	0.272 (0.106) [0.062; 0.459]	-0.102 (0.036) [-0.173; -0.034]	0.254 (0.054) [0.150; 0.361]	-5.344 (0.071) [-5.467; -5.182]
	1.9	0	2.047 (0.104) [1.854; 2.257]	0.405 (0.056) [0.283; 0.492]	-0.137 (0.029) [-0.195; -0.081]	0.229 (0.044) [0.144; 0.318]	-5.480 (0.096) [-5.695; -5.311]
	2.5	4000	2.713 (0.151) [2.419; 3.028]	0.287 (0.075) [0.117; 0.419]	-0.177 (0.026) [-0.229; -0.127]	0.239 (0.042) [0.161; 0.323]	-5.371 (0.051) [-5.457; -5.268]
	5.0	0	5.194 (0.357) [4.525; 5.920]	0.359 (0.053) [0.246; 0.456]	-0.119 (0.018) [-0.155; -0.084]	0.275 (0.036) [0.205; 0.345]	-5.211 (0.061) **[-5.333; -5.087]**
0.45	1.1	0	1.096 (0.049) [1.003; 1.194]	0.447 (0.037) [0.358; 0.497]	-0.196 (0.036) [-0.268; -0.125]	0.235 (0.055) [0.131; 0.345]	-5.424 (0.140) [-5.692; -5.132]
	1.5	0	1.482 (0.068) [1.354; 1.624]	0.387 (0.075) [0.211; 0.493]	-0.102 (0.034) [-0.170; -0.038]	0.266 (0.050) [0.169; 0.370]	-5.180 (0.111) **[-5.382; -4.917]**
	1.9	0	2.020 (0.099) [1.831; 2.223]	0.466 (0.026) [0.401; 0.499]	-0.146 (0.028) [-0.200; -0.094]	0.249 (0.041) [0.172; 0.333]	-5.572 (0.122) [-5.835; -5.346]
	2.5	4000	2.727 (0.156) [2.436; 3.055]	0.402 (0.053) [0.289; 0.491]	-0.178 (0.026) [-0.229; -0.128]	0.244 (0.042) [0.164; 0.328]	-5.305 (0.081) [-5.442; -5.125]
	5.0	0	5.159 (0.342) [4.517; 5.848]	0.452 (0.032) [0.382; 0.497]	-0.126 (0.016) [-0.158; -0.094]	0.271 (0.034) [0.206; 0.338]	-4.985 (0.089) **[-5.155; -4.793]**

Note: The bold-face font for the **credibility interval** indicates that the interval does not contain the true parameter value.

Table 12: Summary for the sample obtained from posterior distributions considering an improper prior for ν and uniform priors for the remaining parameters. The statistics are based on a chain of size $10000 - B$, were $B = \max\{B_1, B_2, B_3, B_4, B_5\}$ and B_i is the burn-in size for parameter η_i, for $i \in \{1, 2, 3, 4, 5\}$, suggested by the Heidelberg and Welch diagnostic criteria. The true parameter values considered in this simulation are $d_0 \in \{0.10, 0.25, 0.35, 0.45\}$, $\nu_0 \in \{1.1, 1.5, 2.5, 5.0\}$, $\theta_0 = -0.15$, $\gamma_0 = -0.24$ and $\omega_0 = -5.4$. The sample size for the FIEGARCH time series is $n = 5000$.

d_0	ν_0	B	$\bar{\nu}$ (sd$_\nu$) $CI_{0.95}(\nu)$	$\bar{d}$ (sd$_d$) $CI_{0.95}(d)$	$\bar{\theta}$ (sd$_\theta$) $CI_{0.95}(\theta)$	$\bar{\gamma}$ (sd$_\gamma$) $CI_{0.95}(\gamma)$	$\bar{\omega}$ (sd$_\omega$) $CI_{0.95}(\omega)$
0.10	1.1	0	1.080 (0.029) [1.025; 1.138]	0.134 (0.075) [0.010; 0.284]	-0.134 (0.026) [-0.185; -0.083]	0.223 (0.037) [0.152; 0.294]	-5.435 (0.035) [-5.498; -5.363]
	1.5	0	1.508 (0.047) [1.418; 1.601]	0.072 (0.049) [0.003; 0.183]	-0.161 (0.023) [-0.207; -0.115]	0.229 (0.034) [0.164; 0.294]	-5.423 (0.028) [-5.474; -5.366]
	1.9	0	1.964 (0.061) [1.840; 2.082]	0.093 (0.054) [0.007; 0.205]	-0.162 (0.020) [-0.202; -0.123]	0.252 (0.032) [0.188; 0.315]	-5.421 (0.026) [-5.482; -5.377]
	2.5	0	2.613 (0.093) [2.438; 2.796]	0.084 (0.052) [0.005; 0.196]	-0.126 (0.017) [-0.160; -0.093]	0.241 (0.030) [0.184; 0.300]	-5.401 (0.021) [-5.440; -5.361]
	5.0	0	5.215 (0.230) [4.770; 5.671]	0.099 (0.044) [0.015; 0.184]	-0.125 (0.012) [-0.150; -0.101]	0.281 (0.023) [0.236; 0.328]	-5.376 (0.020) [-5.411; -5.337]
0.25	1.1	0	1.080 (0.030) [1.026; 1.144]	0.268 (0.072) [0.109; 0.394]	-0.137 (0.026) [-0.188; -0.088]	0.220 (0.036) [0.151; 0.291]	-5.443 (0.049) [-5.535; -5.329]
	1.5	0	1.509 (0.046) [1.426; 1.598]	0.164 (0.063) [0.039; 0.281]	-0.159 (0.023) [-0.204; -0.115]	0.231 (0.033) [0.168; 0.298]	-5.417 (0.030) [-5.474; -5.353]
	1.9	1000	1.960 (0.064) [1.836; 2.084]	0.227 (0.053) [0.119; 0.325]	-0.161 (0.020) [-0.200; -0.123]	0.255 (0.032) [0.192; 0.316]	-5.432 (0.035) [-5.503; -5.372]
	2.5	0	2.615 (0.091) [2.446; 2.810]	0.238 (0.054) [0.129; 0.337]	-0.125 (0.018) [-0.161; -0.091]	0.243 (0.030) [0.187; 0.303]	-5.393 (0.031) [-5.453; -5.330]
	5.0	0	5.208 (0.228) [4.787; 5.675]	0.250 (0.037) [0.176; 0.318]	-0.128 (0.012) [-0.152; -0.105]	0.275 (0.023) [0.229; 0.320]	-5.347 (0.028) [-5.405; -5.298]
0.35	1.1	0	1.079 (0.029) [1.021; 1.137]	0.368 (0.055) [0.254; 0.468]	-0.141 (0.025) [-0.189; -0.093]	0.218 (0.035) [0.152; 0.289]	-5.457 (0.075) [-5.586; -5.291]
	1.5	4000	1.511 (0.044) [1.417; 1.603]	0.265 (0.056) [0.151; 0.368]	-0.157 (0.023) [-0.201; -0.113]	0.237 (0.032) [0.177; 0.302]	-5.384 (0.040) [-5.457; -5.301]
	1.9	0	1.957 (0.062) [1.834; 2.085]	0.335 (0.044) [0.247; 0.420]	-0.160 (0.020) [-0.198; -0.122]	0.258 (0.031) [0.199; 0.321]	-5.463 (0.055) [-5.580; -5.370]
	2.5	0	2.624 (0.092) [2.448; 2.805]	0.351 (0.042) [0.268; 0.431]	-0.125 (0.017) [-0.159; -0.093]	0.243 (0.029) [0.189; 0.301]	-5.373 (0.047) [-5.473; -5.291]
	5.0	0	5.227 (0.222) [4.802; 5.665]	0.354 (0.030) [0.294; 0.415]	-0.132 (0.012) [-0.155; -0.110]	0.268 (0.023) [0.225; 0.315]	-5.269 (0.042) **[-5.351; -5.186]**
0.45	1.1	2000	1.081 (0.028) [1.027; 1.132]	0.449 (0.034) [0.371; 0.497]	-0.149 (0.023) [-0.193; -0.104]	0.224 (0.032) [0.164; 0.289]	-5.473 (0.117) [-5.705; -5.225]
	1.5	0	1.515 (0.045) [1.430; 1.604]	0.387 (0.047) [0.295; 0.474]	-0.154 (0.021) [-0.195; -0.112]	0.246 (0.031) [0.188; 0.311]	-5.240 (0.079) **[-5.379; -5.062]**
	1.9	0	1.948 (0.064) [1.832; 2.079]	0.440 (0.032) [0.373; 0.493]	-0.158 (0.018) [-0.194; -0.124]	0.262 (0.029) [0.206; 0.322]	-5.541 (0.100) [-5.755; -5.362]
	2.5	0	2.622 (0.093) [2.454; 2.814]	0.453 (0.028) [0.393; 0.497]	-0.128 (0.016) [-0.160; -0.098]	0.246 (0.027) [0.195; 0.301]	-5.323 (0.086) [-5.488; -5.162]
	5.0	0	5.264 (0.219) [4.835; 5.717]	0.455 (0.023) [0.407; 0.495]	-0.137 (0.011) [-0.159; -0.115]	0.262 (0.022) [0.219; 0.305]	-5.033 (0.074) **[-5.166; -4.883]**

Note: The bold-face font for the **credibility interval** indicates that the interval does not contain the true parameter value.