SOBRAPO

COMPARISON BETWEEN THE COMPLETE BAYESIAN METHOD AND EMPIRICAL BAYESIAN METHOD FOR ARCH MODELS USING BRAZILIAN FINANCIAL TIME SERIES

Sandra C. Oliveira[1*] and Marinho G. Andrade[2]

ABSTRACT. In this work we compared the estimates of the parameters of ARCH models using a complete Bayesian method and an empirical Bayesian method in which we adopted a non-informative prior distribution and informative prior distribution, respectively. We also considered a reparameterization of those models in order to map the space of the parameters into real space. This procedure permits choosing prior normal distributions for the transformed parameters. The posterior summaries were obtained using Monte Carlo Markov chain methods (MCMC). The methodology was evaluated by considering the Telebras series from the Brazilian financial market. The results show that the two methods are able to adjust ARCH models with different numbers of parameters. The empirical Bayesian method provided a more parsimonious model to the data and better adjustment than the complete Bayesian method.

Keywords: ARCH models, Bayesian approach, MCMC methods.

1 INTRODUCTION

The dynamics of world financial markets requires increasingly sophisticated, complex and efficient models to describe the trends and characteristics of financial assets as accurately as possible.

The analysis of financial series shows they have a high rate of change of the conditional variance in time. The square root of this rate (standard deviation) is called volatility. Understanding how volatility changes over time is critical to the financial market, influencing the risk assessment of investments and asset pricing. It determines the degree of change of the price of the asset in the future, *i.e.*, a low value implies small changes (low risk) in the future, whereas a high value implies significant variations (high risk) (Enders, 2009).

*Corresponding author

[1]Curso de Administração, Campus Experimental de Tupã – CET, Universidade Estadual Paulista – UNESP, Av. Domingos da Costa Lopes, 780, 17602-496 Tupã, SP, Brazil. E-mail: sandra@tupa.unesp.br

[2]Departamento de Matemática Aplicada e Estatística – SME, Instituto de Ciências Matemáticas e de Computação – ICMC, Universidade de São Paulo – USP, Av. do Trabalhador Sãocarlense, 400, 13566-590 São Carlos, SP, Brazil. E-mail: marinho@icmc.usp.br

Let p_t be the price of a given asset at time t, normally a business day. Suppose first that no dividend was paid during the period. The compounded continuous return or simply log-return is given by $z_t = \ln(p_t/p_{t-1})$. This definition is commonly used in the literature and in this paper z_t will be called simply *return*. In practice, it is preferable to work with returns which are free of scale than with prices, because the former are easier to model using techniques of time series analysis (Morettin, 2008).

The characterization of the statistical properties of financial return series is essential for a correct application of models to the data, allowing inferences about the characteristics of this return, especially with regard to the mean and variance, which determine the expected return and volatility forecast for the coming periods. Estimates of these quantities are fundamental to investment decisions not only in assets but also on their derivatives.

There are a large number of nonlinear models for estimating the volatility of financial asset return series. The most widespread in the literature are the autoregressive conditional heteroskedasticity (ARCH) models, proposed by Engle (1982), and their extension, the generalized ARCH (GARCH) models, proposed by Bollerslev (1986). These models feature a nonlinear dependence among the returns, according to the serial dependence of the conditional variance. A comprehensive review of the properties of these models can be found in Degiannakis & Xekalaki (2004) and Bollerslev (2008).

Since it is considered that the volatility at a given instant of time depends on past values of the series, the determination of maximum likelihood estimators (MLE) of the parameters of ARCH models requires maximizing a nonlinear function. Therefore, the estimates can only be obtained numerically. Engle (1982) suggested the use of Newton's method as an iterative method to calculate maximum likelihood estimates. This relaxes the restrictions on parameters (*e.g.*, they must be positive and their sum must be less than one), which ensures stationary covariance. On the other hand, the determination of asymptotic properties of MLE with restrictions involves some difficulties and may lead to local maxima, since properties such as asymptotic normality do not allow restrictions (Barndorff-Nielsen & Cox, 1994). In addition, procedures for identification, diagnosis and adjustment of the models, as well as prediction values of econometric series, require properties of asymptotic theory. Because the models are very far from linear, the asymptotic properties of these estimators can only be verified for very long series and, in general, are more appropriate in the presence of symmetrical distribution of the errors and normal distribution of the data (Geweke, 1986).

In this context, the use of bootstrap methods can be considered to improve estimates of the parameters of models of the ARCH family. But the results of this approach can be severely affected in the MLE calculation, which is performed several times in the procedure and presents convergence difficulties, also due to the restrictions imposed on the parameters (Oliveira, 2005). An alternative to estimate these models, bypassing these difficulties, is to consider a Bayesian approach.

To the best of our knowledge, Geweke (1989b) was the first author to propose a Bayesian approach for ARCH models, in which a particular case of reparameterization allowed the use of

non-informative prior distributions. The parameters were obtained using algorithms of the Monte Carlo simulation. In a subsequent work, a Bayesian semi-nonparametric approach was proposed for ARCH models by Koop (1994), who used the advantages of Geweke's Bayesian methodology and also estimated the model in a semi-nonparametric way. A Bayesian approach was proposed for GARCH models specifically, within the class of dynamic models, in Migon & Mazucheli (1999). Nakatsuma (2000) proposed a Markov Chain Monte Carlo (MCMC) method for linear regression models with ARMA-GARCH errors. He used normal prior and Metropolis-Hastings algorithm simulation for the model parameters to generate the posterior distribution. Polasek & Kozumi (2000) and Polasek (2001) developed a Bayesian approach with hierarchical structure for VAR-VARCH (Vector-AR-Vector-ARCH) and PAR-ARCH (Periodic-AR-ARCH) models, respectively, using MCMC. In the context of unobserved components models, Giakoumatos *et al.* (2005) proposed a Bayesian approach for ARCH models by using auxiliary variables (Pitt & Walker, 2005). More recently, Ausin & Galeano (2007) proposed a Bayesian approach for GARCH models with Gaussian mixture errors, and Barreto *et al.* (2008) compared the Bayesian and maximum likelihood methods using simulated series, according to ARCH processes with different orders and under conditions of finite and infinite variances. Finally, Andrade & Oliveira (2011) presented a Bayesian approach for ARCH models using the normal prior for their parameters and compared credibility intervals with bootstrap intervals. Therefore, given the importance of the subject, in this paper we propose a complete Bayesian method and an empirical Bayesian method for estimating the parameters of ARCH processes, in order to compare the estimates obtained by those methods.

In the complete approach, we considered the prior distribution based on the non-informative proposal of Geweke (1989b) and used MCMC simulation methods to obtain posterior summaries. In the empirical approach, we partitioned the data set considering a Bayesian analysis for the first part of the data, using the non-informative prior distribution of Geweke (1989b) and MCMC simulation methods. With the information obtained from that first step, we defined an informative prior distribution, which was combined with the second part of the data, resulting in posterior summaries.

We also considered a reparameterization of those models to reduce the rejection rate of the MCMC simulation algorithm and accelerate the convergence process. The proposed methodology was used to model a series of daily prices of Telebras shares traded on the São Paulo Stock Exchange (Bovespa).

2   ARCH(q) MODELS

The regression model proposed by Engle (1982) with zero mean, expressed as a linear combination of exogenous variables, has a structure that can be summarized as:

$$z_t | \Omega_{t-1} \sim P(0, h_t) \tag{1}$$

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j z_{t-j}^2 \tag{2}$$

where z_t represents a return series, $P(\cdot)$ is a parametric distribution, usually normal or Student-t and $\Omega_{t-1} = \{z_{t-1}, z_{t-2}, \ldots\}$ represents a set of information available at $t-1$. Let z_t be a process that satisfies the model

$$z_t = h_t^{1/2} \varepsilon_t \qquad (3)$$

In this work we assumed that $\{\varepsilon_t, t \geq 0\}$ is a sequence of white noise $i.i.d.$ $N(0, 1)$.

For the model (1)-(3) to be plausible ($h_t > 0$ for all t), we must have $\alpha_0 > 0$ and $\alpha_j \geq 0$, $j = 1, \ldots, q$. Furthermore, the process z_t has finite variance, and therefore has stationary covariance if and only if all the roots of the polynomial $1 - \sum_{j=1}^{q} \alpha_j l^j$ are outside the unit circle, where l is the delay operator, such that $l^j z_t = z_{t-j}$. When this condition is satisfied, it can be shown that the unconditional variance of z_t is given by

$$\alpha_0 \Big/ \left(1 - \sum_{j=1}^{q} \alpha_j\right).$$

Therefore the sufficient condition for this process to have stationary covariance is $\sum_{j=1}^{q} \alpha_j < 1$.

3   LIKELIHOOD FUNCTION

Let $\mathbf{Z}_1^T = \{z_1, z_2, \ldots, z_T\}$ be an observed trajectory of the return z_t. Assuming that the first q observations $\mathbf{Z}_1^q = \{z_1, z_2, \ldots, z_q\}$ are known, the likelihood function of the observations $\mathbf{Z}_{q+1}^T = \{z_{q+1}, z_{q+2}, \ldots, z_T\}$, conditional on $\mathbf{Z}_1^q$, is given by

$$L\left(\boldsymbol{\alpha} | \mathbf{Z}_{q+1}^T, \mathbf{Z}_1^q\right) \propto \prod_{t=q+1}^{T} \left(\frac{1}{h_t}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t}\right\} \qquad (4)$$

where $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \ \alpha_1 \ \ldots \ \alpha_q \end{bmatrix}'$ is the vector of parameters of the ARCH(q) model and $h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j z_{t-j}^2$.

Thus, maximum likelihood estimators for the parameter vector α can be obtained by means of maximizing the likelihood function $L\left(\boldsymbol{\alpha} | \mathbf{Z}_{q+1}^T, \mathbf{Z}_1^q\right)$. It is worth noting that the use of the conditional likelihood function instead of the exact likelihood function can be done without great loss of precision in the estimates when the series z_t is large. This approximation is generally considered due to the great practical advantages in calculating the estimators (Box *et al.*, 1994).

4   BAYESIAN INFERENCE FOR ARCH(q) MODELS

The Bayesian approach for inferring the parameters of ARCH(q) models starts from the combination of the likelihood function, $L\left(\boldsymbol{\alpha} | \mathbf{Z}_{q+1}^T, \mathbf{Z}_1^q\right)$, with an *a prior* probability density for the parameters, $\pi(\boldsymbol{\alpha})$, by means of Bayes' rule (Gelman *et al.*, 1995):

$$\pi\left(\boldsymbol{\alpha} | \mathbf{Z}_{q+1}^T\right) \propto L\left(\boldsymbol{\alpha} | \mathbf{Z}_{q+1}^T, \mathbf{Z}_1^q\right) \pi(\boldsymbol{\alpha}) \qquad (5)$$

The expression $\pi\left(\boldsymbol{\alpha}|\mathbf{Z}_{q+1}^{T}\right)$ is called the a posterior probability density of $\boldsymbol{\alpha}$ and tells how the random variable $\boldsymbol{\alpha}$ will be distributed after the data $\mathbf{Z}_{q+1}^{T}$ have been observed.

In sections 4.1 and 4.2 we propose two different approaches for Bayesian inference for the parameters $\boldsymbol{\alpha}$ of ARCH(q) models.

4.1   Complete Bayesian model (CBM) with non-informative prior

Bayesian analysis requires the specification of an *a prior* density for $\pi(\boldsymbol{\alpha})$ to reflect the prior knowledge about the distribution of $\pi(\boldsymbol{\alpha})$. When there is little or no prior knowledge about the distribution of a model's parameters, one must adopt a non-informative prior density. In this work we adopted a non-informative prior density for the parameters $\boldsymbol{\alpha}$ of ARCH(q) models based on the proposal of Geweke (1989b), defined as:

$$\pi(\boldsymbol{\alpha}) \propto \begin{cases} \left(\dfrac{1 - \sum\limits_{j=1}^{q} \alpha_j}{\alpha_0}\right) & , \ \alpha_0 > 0, \ \alpha_j \geq 0, \ j = 1, \ldots, q \\ \\ 0 & , \ \text{otherwise} \end{cases} \tag{6}$$

Depending on $\alpha_j = 0$, $j = 1, \ldots, q$, this is the invariant prior density of Jeffreys (Box & Tiao, 1973) for the normal linear regression model. We also consider a reparameterization, which consists of:

$$\phi_j = \log\left(\frac{\alpha_j - a_j}{b_j - \alpha_j}\right) \quad j = 0, 1, 2, \ldots, q \tag{7}$$

The values of a_j and b_j can be chosen based on some prior information, for example, previous studies on the series analyzed. This reparameterization leads to the choice of a transformation that maps the intervals $(-\infty, +\infty)$ in the domain (a_j, b_j) and vice versa. Accordingly, we reduced the rejection rate of the MCMC simulation algorithm, accelerating its convergence. Using the reparametrization defined in (7), we have:

$$\pi(\boldsymbol{\varphi}) \propto \begin{cases} \left(\dfrac{1}{k(\boldsymbol{\varphi})}\right)^{\frac{1}{2}} & , \ -\infty < \varphi_j < \infty, \ j = 0, 1, \ldots, q \\ \\ 0 & , \ \text{otherwise} \end{cases} \tag{8}$$

where

$$\boldsymbol{\varphi} = \left[\phi_0 \phi_1 \ldots \phi_q\right]' \quad \text{and} \quad k(\boldsymbol{\varphi}) = \alpha_0 \left/ \left(1 - \sum_{j=1}^{q} \alpha_j\right)\right.$$

with α_j, $j = 0, 1, 2, \ldots, q$ given by (7).

Combining the expressions (4) and (8), we write the posterior joint density for the $\boldsymbol{\varphi}$ as:

$$\pi\left(\boldsymbol{\varphi}|\mathbf{Z}_{q+1}^{T}\right) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=q+1}^{T} \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\varphi})}\right\} \tag{9}$$

We note that the posterior density (9) has a form that only resembles the known probability density functions, which prevents the calculation of the quantities of interest of the parameters α_j, $j = 0, 1, \ldots, q$, such as mean, mode, median, standard deviation, among others (Kotz *et al.*, 2000).

Since the adjustment of ARCH processes to financial series generally are of high orders, the use of numerical integration methods would require large computational effort and have low accuracy. This problem can be solved by using Monte Carlo Markov chain (MCMC) simulation methods, specifically the Metropolis-Hastings algorithm. The methods based for this estimation are preferable because they can be generalized to more complex models. Other algorithms could be used (*e.g.*, acceptance/rejection, importance sampling, among others), but the Metropolis-Hastings is one of those requiring the least computational effort (Gilks *et al.*, 1996).

4.2   Empirical Bayes Model (EBM)

The empirical Bayesian methods are distinguished by using the observed data to estimate the parameters of the prior and avoid the difficulties of determining the distribution, either due to lack of knowledge about the variable of interest or some subjectivity involved in choosing it. The empirical method can be viewed as an approximation of the complete hierarchical Bayesian analysis (Gelman *et al.*, 1995).

In this work we consider a partition of the return series $\mathbf{Z}_1^T = \{z_1, z_2, \ldots, z_T\}$, such that

$$\mathbf{Z}_1^T = \left[\mathbf{Z}_1^{T_1}\ \mathbf{Z}_{T_1+1}^T\right]', \quad \text{where } \mathbf{Z}_1^{T_1} = \{z_1, z_2, \ldots, z_{T_1}\} \text{ and } \mathbf{Z}_{T_1+1}^T = \{z_{T_1+1}, z_{T_1+2}, \ldots, z_T\},$$

and we propose a Bayesian framework composed of two steps.

In the first step we must perform an analysis with the first part of the series of observed data, $\mathbf{Z}_1^{T_1}$, assuming a non-informative prior density, if possible, $\pi_1(\boldsymbol{\alpha})$, to make inferences about the parameter vector $\boldsymbol{\alpha} = \left[\alpha_0\ \alpha_1\ \ldots\ \alpha_{q_1}\right]'$. We thus propose the following Bayesian model:

$$\pi\left(\alpha|\mathbf{Z}_{q_1+1}^{T_1}\right) \propto L\left(\alpha|\mathbf{Z}_{q_1+1}^{T_1}, \mathbf{Z}_1^{q_1}\right)\pi_1(\alpha) \tag{10}$$

where $\mathbf{Z}_1^{q_1} = \{z_1, z_2, \ldots, z_{q_1}\}$ and $\mathbf{Z}_{q_1+1}^{T_1} = \{z_{q_1+1}, z_{q_1+2}, \ldots, z_{T_1}\}$. This procedure provides posterior estimates for α through the density $\pi\left(\boldsymbol{\alpha}|\mathbf{Z}_{q_1+1}^{T_1}\right)$.

These estimates are used for inference about the density of $\boldsymbol{\alpha}^* = \left[\alpha_0\ \alpha_1\ \ldots\ \alpha_{q_2}\right]', \pi_2(\boldsymbol{\alpha}^*)$, which is used as an informative prior density for the second stage of the analysis (main analysis), where we combine the likelihood function constructed from the second part of the series of observed data, $\mathbf{Z}_{T_1+1}^T$, with the prior density $\pi(\boldsymbol{\alpha}^*)$, leading to the posterior density, given by:

$$\pi\left(\boldsymbol{\alpha}^*|\mathbf{Z}_{T_1+q_2+1}^T\right) \propto L\left(\boldsymbol{\alpha}^*|\mathbf{Z}_{T_1+q_2+1}^T, \mathbf{Z}_{T_1+1}^{T_1+q_2}\right)\pi_2(\boldsymbol{\alpha}^*) \tag{11}$$

where

$$\mathbf{Z}_{T_1+1}^{T_1+q_2} = \{z_{T_1+1}, z_{T_1+2}, \ldots, z_{T_1+q_2}\} \quad \text{and} \quad \mathbf{Z}_{T_1+q_2+1}^T = \{z_{T_1+q_2+1}, z_{T_1+q_2+2}, \ldots, z_T\}.$$

In this paper, we partition the data set so that the first half of the data series (*i.e.*, $T_1 = \frac{T}{2}$ observations) are used in the first stage of the Bayesian analysis and the second half of the data series are used in the second stage (main analysis). The choice of T_1 requires careful analysis of the sensitivity of the parameter estimates in function of the sample size. A minimum training sample is always desirable. However, the computational effort is very great for this choice. Further details and discussions about this procedure can be found in Berger & Pericchi (2004).

For the Bayesian analysis of the first step, we consider non-informative prior density of Geweke (1989b), defined in (6)-(8), assuming little or no prior knowledge about the distribution of the model's parameters. Assuming that the first q_1 observations are known, the conditional likelihood function is defined by means of the volatility h_t rewritten as a function of the parameters α_j, $j = 0, 1, \ldots, q_1$, transformed into ϕ_j, $j = 0, 1, \ldots, q_1$, by (7), *i.e.*:

$$L\left(\boldsymbol{\varphi}|\mathbf{Z}_{q_1+1}^{T_1}, \mathbf{Z}_1^{q_1}\right) \propto \prod_{t=q_1+1}^{T_1} \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\varphi})}\right\} \tag{12}$$

where $\boldsymbol{\varphi} = \begin{bmatrix} \phi_0 & \phi_1 & \ldots & \phi_{q_1} \end{bmatrix}'$. Combining expressions (12) and (8), we can write the prior joint distribution for $\boldsymbol{\varphi}$ as:

$$\pi\left(\boldsymbol{\varphi}|\mathbf{Z}_{q_1+1}^{T_1}\right) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=q_1+1}^{T_1} \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\varphi})}\right\} \tag{13}$$

where

$$k(\boldsymbol{\varphi}) = \alpha_0 \left/ \left(1 - \sum_{j=1}^{q_1} \alpha_j\right)\right. \quad \text{and} \quad \alpha_j = \frac{b_j e^{\phi_j} + a_j}{1 + e^{\phi_j}}; \quad j = 0, 1, \ldots, q_1.$$

For the Bayesian analysis of the second stage, we set an informative prior density for $\boldsymbol{\alpha}^*$, $\pi(\boldsymbol{\alpha}^*)$, which is built from the analysis of the first stage, *i.e.*, using the posterior estimates obtained by means of the density $\pi\left(\boldsymbol{\alpha}|\mathbf{Z}_{q_1+1}^{T_1}\right)$.

We consider that all α_j, $j = 0, 1, \ldots, q_2$ are independent with a prior density defined in the interval $\begin{bmatrix} a_j, b_j \end{bmatrix}$, with $a_j > 0$ and $b_j < 1$, such that $a_j \leq \alpha_j \leq b_j$, since the constraint $\sum_{j=1}^{q_2} \alpha_j < 1$ requires that $\alpha_j \in (0, 1)$, $j = 1, 2, \ldots, q_2$. Because we want to generate values for α_j close to the averages of each interval $(a_j + b_j)/2$ with greater frequency than values near the limits a_j and b_j, we adopt the proposed reparameterization in equation (7), such that ϕ_j^* is a component of the vector of parameters $\boldsymbol{\varphi}^* = \begin{bmatrix} \phi_0^* & \phi_1^* & \ldots & \phi_{q_2}^* \end{bmatrix}'$ and a_j and b_j is selected based on the upper and lower bounds of the 95% credibility intervals for $\boldsymbol{\alpha}$, obtained from the posterior density $\pi\left(\boldsymbol{\alpha}|\mathbf{Z}_{q_1+1}^{T_1}\right)$.

In this case, the conditional likelihood function is defined with volatility h_t rewritten in terms of the parameters α_j^*, $j = 0, 1, \ldots, q_2$, transformed similarly in ϕ_j^*, $j = 0, 1, \ldots, q_2$, by (7), *i.e.*:

$$L\left(\boldsymbol{\varphi}^*|\mathbf{Z}_{T_1+q_2+1}^{T}, \mathbf{Z}_{T_1+1}^{T_1+q_2}\right) \propto \prod_{t=T_1+q_2+1}^{T} \left(\frac{1}{h_t(\boldsymbol{\varphi}^*)}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\varphi}^*)}\right\} \tag{14}$$

where $\boldsymbol{\varphi}^* = \begin{bmatrix} \phi_0^* \ \phi_1^* \ \dots \ \phi_{q_2}^* \end{bmatrix}'$. Assuming that ϕ_j^*, $j = 0, 1, \dots, q_2$ are normally distributed, i.e., $\pi(\phi_j^*) \sim Normal\,(0, \sigma_j^2)$, the *a prior* joint density is given by:

$$\pi\left(\boldsymbol{\varphi}^*\right) = \prod_{j=0}^{q_2} \pi\left(\phi_j^*\right) \tag{15}$$

Combining expressions (14) and (15), we can write the posterior joint density for $\boldsymbol{\varphi}^*$ as:

$$\pi\left(\boldsymbol{\varphi}^*|Z_{T_1+q_2+1}^T\right) \propto \left(\prod_{j=0}^{q_2} \pi\left(\phi_j^*\right)\right) \times \prod_{t=T_1+q_2+1}^T \left(\frac{1}{h_t\left(\boldsymbol{\varphi}^*\right)}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t\left(\boldsymbol{\varphi}^*\right)}\right\} \tag{16}$$

Thus, the posterior conditional densities for ϕ_j^*, $j = 0, 1, \dots, q_2$, obtained by means of expression (16), are given by:

$$\pi\left(\phi_j^*|\boldsymbol{\varphi}_{-j}^*, \mathbf{Z}_{T_1+q_2+1}^T\right) \propto L\left(\mathbf{Z}_{T_1+q_2+1}^T|\mathbf{Z}_{T_1+1}^{T_1+q_2}, \boldsymbol{\varphi}^*\right) \pi\left(\phi_j^*\right) \tag{17}$$

where $\boldsymbol{\varphi}_{-j}^*$ is a vector formed by the model's parameters, except the parameter ϕ_j^*, i.e.,

$$\boldsymbol{\varphi}_{-j}^* = \begin{bmatrix} \phi_0^* \ \phi_1^* \ \dots \ \phi_{j-1}^* \phi_{j+1}^* \ \dots \ \phi_{q_2}^* \end{bmatrix}'.$$

5   ALGORITHMS TO ESTIMATE THE PARAMETERS OF ARCH(q) MODELS

To obtain representative samples from the posterior joint density (9), (13) and (16), we use MCMC simulation algorithms.

The MCMC simulation method is a form of Monte Carlo integration. The idea is to simulate an irreducible non-periodic Markov chain whose stationary density is the density of interest, *i.e.*, the posterior density. There are two methods to generate Markov chains with specified stationary density, the Metropolis-Hastings (Chib & Greenberg, 1995), which has been used for many years in statistical physics, and Gibbs sampling, which was introduced in the statistical literature by Gelfand & Smith (1990). In general, the use of these algorithms is necessary when the non-iterative generation of the density to be sampled is too complicated or costly.

The Metropolis-Hastings algorithm

When posterior conditional densities are not easily identified as having a standard form (normal, gamma, etc.), preventing the direct generation from these densities, the Metropolis-Hastings algorithm can be used. This technique requires a nucleus, *i.e.*, a transition density $q\left(\phi, \phi'\right)$, not necessarily having equilibrium probability π, but which represents a passing rule to define a chain. Consider also the probability of acceptance $p\left(\phi^{(r-1)}, \phi'\right)$ set forth below. The algorithm follows these steps:

Step 1: Assign an initial value $\phi = \phi^{(0)}$ and start the iteration count at $r = 1$.

Step 2: Move the chain to a new value ϕ' generated from the density $q\left(\phi^{(r-1)}, \phi'\right)$.

Step 3: Generate u from the uniform density $(0, 1)$.

Step 4: Accept the value generated ϕ' if

$$u \leq p\left(\phi^{(r-1)}, \phi'\right) = \min\left(1, \frac{\pi\left(\phi'\right) q\left(\phi', \phi^{(r-1)}\right)}{\pi\left(\phi^{(r-1)}\right), q\left(\phi^{(r-1)}, \phi'\right)}\right)$$

Otherwise keep $\phi^{(r)} = \phi^{(r-1)}$.

Step 5: Update r and go to step 2.

For r sufficiently large, $\phi^{(r)}$ is a sample of the posterior distribution. For the vector case $\boldsymbol{\varphi} = \left[\phi_0 \; \phi_1 \; \ldots \; \phi_q\right]$, we have a transition density given by $q\left(\boldsymbol{\varphi}, \boldsymbol{\varphi}'\right)$ and an acceptance probability given by $p\left(\boldsymbol{\varphi}^{(r-1)}, \boldsymbol{\varphi}'\right)$, and we must do the same.

One feature that should be considered in the use of the Metropolis-Hastings algorithm is the fact that it can generate correlated samples, since if the generated candidate is rejected, the previous value is taken as a candidate. To avoid biased estimates, it is common to disregard the initial part of the chain in order to avoid the influence of initial conditions, and to select values for each k steps to eliminate the sample correlation. The evaluation of this procedure is reflected in the acceptance rate, and the convergence is checked using the criterion of Geweke (1992).

To ensure that the Metropolis-Hastings algorithm will converge to an equilibrium density, one must observe the regularity conditions of the posterior distributions, which are guaranteed in our case.

Assuming a quadratic loss function, the Bayesian estimates for $\boldsymbol{\alpha}$ (Sections 4.1 and 4.2) can be expressed as evaluating the expectancy of a function of interest $g(\boldsymbol{\alpha})$ with respect to the posterior density $\pi(\boldsymbol{\alpha}|Z)$. In general, it is given by:

$$E[g(\boldsymbol{\alpha})] = \int_{\boldsymbol{\alpha}} g(\boldsymbol{\alpha}) \pi(\boldsymbol{\alpha}|\mathbf{Z}) \, d\boldsymbol{\alpha} \tag{18}$$

The exact solution of the integral in (18) cannot be calculated analytically for the models considered in this work. However, if the expression $E[g(\boldsymbol{\alpha})]$ exists, an approximation to expression (18) can be obtained by Monte Carlo simulation (Geweke, 1989a).

6 MODEL SELECTION CRITERIA

There are several criteria that can be used for the selection of the order q of ARCH models under a Bayesian framework. In this paper, we use the following criteria: Bayesian information criterion (BIC), Akaike information criterion (AIC), deviation information criterion (DIC) and predictive ordinate criterion (POC), which are described more generally below.

6.1 Bayesian information criterion (BIC) and Akaike information criterion (AIC)

The Bayesian information criterion (BIC) is a model selection criterion proposed by Schwarz (1978) and modified by Carlin & Louis (2000) to be applied in the Bayesian context. This crite-

rion is a weighting between the maximum of the posterior expected value of the log-likelihood function and the number of model parameters. The best model is the one that has the lowest BIC value, given by:

$$BIC = -2E(l) + M \ln(T) \tag{19}$$

where $E(l)$ is the expected value taken with respect to the posterior density of the log-likelihood function, M is the size of the parameter vector and T is the size of the series.

The Akaike information criterion (AIC), proposed by Akaike (1974), was also modified to be used in the Bayesian context. The best model is the one that has the lowest AIC value, given by:

$$AIC = -2E(l) + 2M \tag{20}$$

Lower values of BIC and/or AIC indicate better models, but the BIC penalizes the number of parameters more and tends to select models with fewer parameters.

6.2   Deviation information criterion (DIC)

The deviation information criterion (DIC) is a generalization of the BIC. Proposed by Spiegel-halter *et al.* (2002), this criterion is defined as:

$$DIC = D\left(\hat{\boldsymbol{\alpha}}\right) + 2M_D \tag{21}$$

where $D\left(\hat{\boldsymbol{\alpha}}\right) = -2 \ln L\left(\mathbf{Z}|\hat{\boldsymbol{\alpha}}\right) + C$ is the deviation assessed in the posterior mean, $\ln L\left(\mathbf{Z}|\hat{\boldsymbol{\alpha}}\right)$ is the natural logarithm of the likelihood function, C is a constant that is canceled and need not be known when comparing models and M_D is the effective number of parameters of the model, given by: $M_D = \overline{D} - D\left(\hat{\boldsymbol{\alpha}}\right)$, where $\overline{D} = E\{D\left(\boldsymbol{\alpha}\right)\}$ is the posterior mean deviation, which measures the model's goodness of fit. The model with the lowest DIC value will be considered the most suitable to represent the data.

The DIC is commonly used for Bayesian models whose posterior distribution of parameters was obtained through MCMC simulation algorithms as well as proposed models for analysis of financial series.

6.3   Predictive ordinate criterion (POC)

In the criterion for selecting models based on predictive ordinate density (Carlin & Chib, 1995) we use the distribution of z_{T+m}, subject to the data $\mathbf{Z}$ and the parameters $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \ \alpha_1 \ \ldots \ \alpha_q \end{bmatrix}'$. Thus, the predictive density is given by:

$$\pi\left(z_{T+m}|\mathbf{Z}\right) = \int_{\boldsymbol{\alpha}} \pi(z_{T+m}|\mathbf{Z}, \boldsymbol{\alpha})\pi(\boldsymbol{\alpha}|\mathbf{Z})d\boldsymbol{\alpha}$$

$$c_m = \pi\left(z_{T+m}|\mathbf{Z}\right) = \int_{\boldsymbol{\alpha}} \left(\frac{1}{2\pi \ h_{T+m}}\right)^{1/2} \exp\left\{-\frac{1}{2}\frac{z_{T+m}^2}{h_{T+m}}\right\}\pi\left(\boldsymbol{\alpha}|\mathbf{Z}\right)d\boldsymbol{\alpha} \tag{22}$$

where $\pi\left(\boldsymbol{\alpha}|\mathbf{Z}\right)$ is the posterior distribution of the parameters $\boldsymbol{\alpha}$ and the integral in (22) is a multiple integral in the parameters space.

Assessing the observations $\alpha_j^{(i)}$, $j = 0, 1, \ldots, q$ obtained by MCMC simulation and considering h_{T+m} as a function of the parameters $\boldsymbol{\alpha}$, the Monte Carlo estimate of the predictive density is given by:

$$\hat{c}_m = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{\sqrt{2\pi\,h_{T+m}(\boldsymbol{\alpha}^{(i)})}} \exp\left\{ -\frac{1}{2} \frac{z_{T+m}^2}{h_{T+m}(\boldsymbol{\alpha}^{(i)})} \right\} \tag{23}$$

A low value of $\hat{c}_m$ indicates that the observation z_{T+m} is unlikely for the model in question. Thus, the POC consists of choosing the model l that provides the greatest value of

$$\hat{c}\left(l\right) = \prod_{m=T+1}^{T+K} \hat{c}_m\left(l\right).$$

7   APPLICATION

The analyzed time series consists of daily prices of shares of Telebras (Brazilian telecom holding company controlled by the federal government), traded on the Bovespa in the period from January 2, 1992 to January 5, 1996, for a total of 986 observations.

This was one of the most liquid papers (according to the number of trades and financial volume) in the Brazilian stock market during the study period, *i.e.* one of the most representative of those making up the Bovespa index (Ibovespa). The behavior of the Telebras stock has been widely studied in the literature and among the works on this issue, we can mention: Baidya & Costa (2001), Oliveira (2005) and Sáfadi & Andrade (2007).

Let p_t be the closing price of Telebras shares on a trading day, in which the return is given by $z_t = \ln\left(p_t / p_{t-1}\right)$. The share prices of Telebras and their returns are illustrated in Figure 1.

The graphs in Figure 2 also show the behavior of the series of squared returns of Telebras, z_t^2. We can see that this series is correlated. This behavior is typically associated with that of ARCH (q) models.

In implementing the Metropolis-Hastings algorithm, for each model (CBM and EBM), we simulated a chain with 50,000 iterations for each parameter, then discarded 50% of the values to decrease the effect of initial conditions and then took values spaced 5 to 5, for a total sample of 5,000 observations. The convergence of the algorithm was verified by Geweke's criterion (1992), at a significance level of 5% for the test under the null hypothesis Ho. The convergence was considered for values obtained by the diagnosis between -1.96 and 1.96. In this work, all the computational algorithms were implemented using the MATLAB$^{\copyright}$ software.

For CBM, the adjusted model for the Telebras series, according to the AIC, DIC and POC (Table 1), was an ARCH (7), described as:

$$z_t = \left(\alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \alpha_3 z_{t-3}^2 + \alpha_4 z_{t-4}^2 + \alpha_5 z_{t-5}^2 + \alpha_6 z_{t-6}^2 + \alpha_7 z_{t-7}^2 \right)^{\frac{1}{2}} \varepsilon_t \tag{24}$$
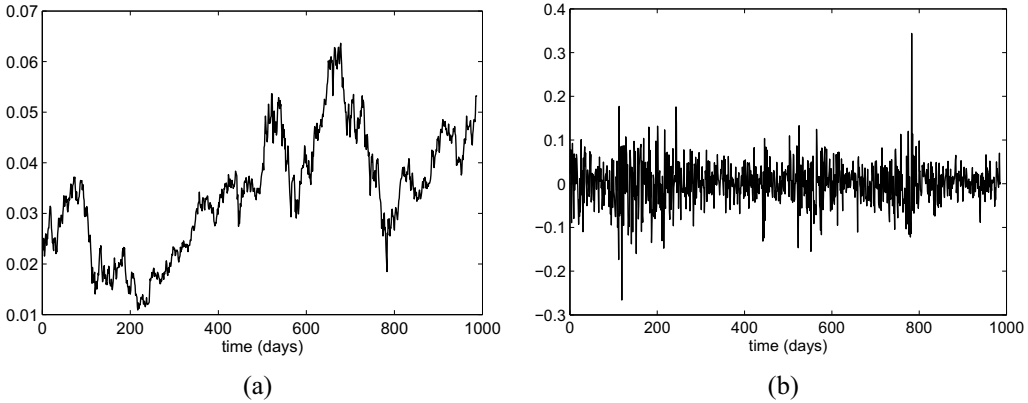
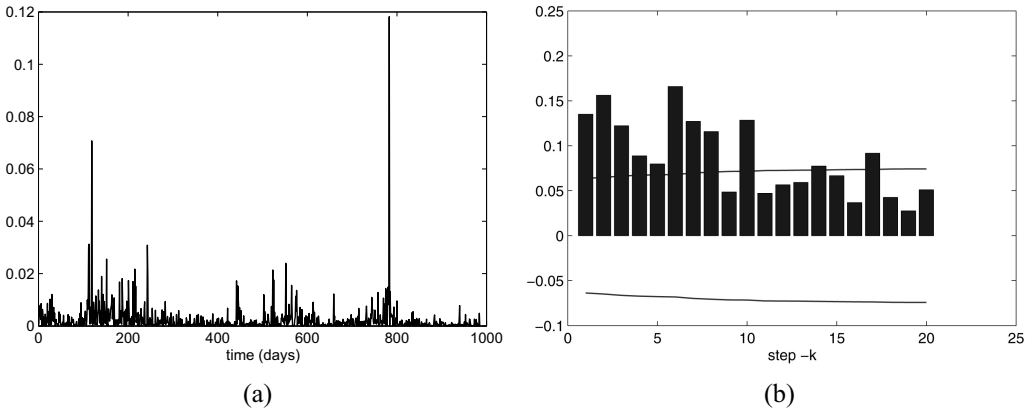**Figure 1** – Price (a) and return (b) of the Telebras series.

Figure 2 – Series of squared returns of Telebras, z_t^2 (a), and its autocorrelation function (b).

Table 1 – Model selection criteria for the ARCH (q) process, considering CBM.

Model order	AIC $(\times 10^3)$	BIC $(\times 10^3)$	DIC $(\times 10^3)$	POC $(\times 10^{27})$	$-\mathrm{E}\{\log l\}$ $(\times 10^3)$
$q = 6$	-5.13244	$\mathbf{-5.09818}$	-1.27337	1.16306	-2.00322
$q = 7$	$\mathbf{-5.13583}$	-5.09668	$\mathbf{-1.27675}$	$\mathbf{1.25362}$	$\mathbf{-2.00591}$
$q = 8$	-5.12954	-5.08550	-1.27302	0.77599	-2.00377

Table 2 shows the estimates and 95% credibility intervals (CI) for $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \ \alpha_1 \ \dots \ \alpha_q \end{bmatrix}'$, using the CBM, obtained in accordance with the procedures proposed for the inference about parameters of ARCH(q) processes, $q = 7$. The convergence of parameters was checked by Geweke's criterion (GC), which was observed for those. The last column of this table also shows the acceptance rates (AR) for the parameter values generated by the Metropolis-Hastings algorithm.

As previously mentioned, in the EBM we partitioned the data set such that the first half of the data series was used in Bayesian analysis in the first step, according to the procedure presented

Table 2 – Complete Bayesian estimate and 95% CI.

	Mean	Standard Deviation	Median	Mode	CI (95%)	GC	AR (%)
α_0	0.00069	0.00003	0.00069	0.00055	[0.00063; 0.00076]	0.90273	85.098
α_1	0.10307	0.01923	0.10219	0.11560	[0.06635; 0.13683]	0.15009	86.314
α_2	0.15264	0.01982	0.15474	0.15780	[0.11469; 0.18763]	0.11001	87.456
α_3	0.10277	0.01884	0.10270	0.10770	[0.06639; 0.13749]	1.88097	86.368
α_4	0.09878	0.01902	0.09564	0.10460	[0.06468; 0.13449]	0.00754	84.584
α_5	0.05173	0.01857	0.05338	0.05240	[0.01688; 0.08664]	-1.55891	84.318
α_6	0.10362	0.01912	0.10107	0.12090	[0.06587; 0.13771]	0.78153	85.960
α_7	0.10379	0.01871	0.10292	0.11350	[0.06727; 0.13761]	-0.79080	85.606

in Section 4.2, to obtain posterior estimates $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \ \alpha_1 \ \ldots \ \alpha_{q_1} \end{bmatrix}'$. Thus, the model fit to the data Telebras series, according to the AIC, BIC and POC (Table 3) was an ARCH (6), given by:

$$z_t = \left(\alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \alpha_3 z_{t-3}^2 + \alpha_4 z_{t-4}^2 + \alpha_5 z_{t-5}^2 + \alpha_6 z_{t-6}^2 \right)^{\frac{1}{2}} \varepsilon_t \qquad (25)$$

Table 3 – Model selection criteria for the ARCH (q_1) process, considering EBM: Bayesian analysis of the first step.

Model order	AIC $\left(\times 10^3 \right)$	BIC $\left(\times 10^3 \right)$	DIC $\left(\times 10^3 \right)$	POC $\left(\times 10^{28} \right)$	$- \mathrm{E}\{\log l\}$ $\left(\times 10^2 \right)$
$q_1 = 5$	-2.45388	-2.42452	-5.89629	2.31037	-5.82941
$q_1 = 6$	$\mathbf{-2.48262}$	$\mathbf{-2.44836}$	-6.00084	**2.66551**	$\mathbf{-5.98308}$
$q_1 = 7$	-2.47961	-2.44046	$\mathbf{-6.07358}$	2.46414	-5.97805

Table 4 shows the estimates and 95% credibility intervals (CI) for $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \ \alpha_1 \ \ldots \ \alpha_{q_1} \end{bmatrix}'$, obtained in the first stage of the EBM, according to an ARCH(q_1) process, $q_1 = 6$.

This information served to elucidate the informative prior distribution of $\boldsymbol{\alpha}^*$, $\pi(\boldsymbol{\alpha}^*)$, where their hyperparameters a_j and b_j assumed values of lower and upper bounds, respectively, in the 95% credibility intervals presented in Table 4 (in bold), and $\sigma_j^2 = 1$, for $j = 0, 1, \ldots, q_2$.

Thus, the second half of the data series was used in the second stage (main analysis), with $\pi(\boldsymbol{\alpha}^*)$ resulting in posterior summaries of interest. The model fit to the Telebras data series in the main analysis, according to the model selection criteria (Table 5), was an ARCH (5), given by:

$$z_t = \left(\alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \alpha_3 z_{t-3}^2 + \alpha_4 z_{t-4}^2 + \alpha_5 z_{t-5}^2 \right)^{\frac{1}{2}} \varepsilon_t \qquad (26)$$

Table 6 shows the estimates and 95% credibility intervals (CI) for $\boldsymbol{\alpha}^* = \begin{bmatrix} \alpha_0 \ \alpha_1 \ \ldots \ \alpha_{q_2} \end{bmatrix}'$, using EBM, obtained from the main analysis by an ARCH(q_2) process, $q_2 = 5$.

Table 4 – Empirical Bayesian estimate and 95% CI of the first step.

	Mean	Standard Deviation	Median	Mode	CI (95%)	GC	AR (%)
α_0	0.00071	0.00003	0.00071	0.00070	**[0.00064; 0.00077]**	−0.85421	87.734
α_1	0.09549	0.03084	0.09359	0.09556	**[0.04179; 0.15933]**	0.56397	69.398
α_2	0.09081	0.03617	0.08880	0.09103	**[0.02916; 0.16344]**	1.15489	73.296
α_3	0.09826	0.03416	0.09723	0.09803	**[0.03566; 0.16578]**	−0.14777	75.122
α_4	0.17912	0.04231	0.17801	0.17971	**[0.10014; 0.26113]**	−0.01410	73.816
α_5	0.09748	0.03483	0.09578	0.09665	**[0.03348; 0.16691]**	−1.03893	76.228
α_6	0.19613	0.03854	0.19580	0.19677	**[0.12089; 0.26848]**	−1.68313	70.296

Table 5 – Model selection criteria for the ARCH(q_2) process, considering EBM: main Bayesian analysis.

Model order	AIC $\left(\times 10^3 \right)$	BIC $\left(\times 10^3 \right)$	DIC $\left(\times 10^3 \right)$	POC $\left(\times 10^{27} \right)$	$- E\{\log l\}$ $\left(\times 10^2 \right)$
$q_2 = 4$	−2.61215	−2.58768	−6.49695	1.53242	−6.61074
$q_2 = 5$	**−2.62362**	**−2.59426**	**−6.51874**	**1.83186**	**−6.67811**
$q_2 = 6$	−2.60406	−2.56980	−6.48157	0.90524	−6.59028

Table 6 – Empirical Bayesian estimate and 95% CI: Main analysis.

	Mean	Standard Deviation	Median	Mode	CI (95%)	GC	AR (%)
α_0	0.00071	0.00003	0.00071	0.00071	[0.00070; 0.00072]	−0.41660	98.400
α_1	0.12299	0.01847	0.12546	0.12881	[0.08050; 0.15075]	−1.56153	48.112
α_2	0.12926	0.01741	0.13228	0.14029	[0.08874; 0.15455]	0.97803	33.782
α_3	0.11652	0.02331	0.11922	0.13008	[0.06607; 0.15356]	−0.53695	65.944
α_4	0.15254	0.02725	0.14804	0.13122	[0.11293; 0.21611]	−0.78737	52.852
α_5	0.09784	0.02562	0.09736	0.09750	[0.05117; 0.14758]	−0.95037	89.318

The graphs in Figure 3 represent the histograms constructed from the sample selected for the parameters $\alpha_0, \alpha_1, \ldots, \alpha_7$, *i.e.*, the posterior distribution of the parameters of the ARCH (7) model, which was adjusted to the Telebras series considering a complete Bayesian model.

The graphs in Figure 4 represent the histograms constructed from the sample selected for the parameters $\alpha_0, \alpha_1, \ldots, \alpha_5$, *i.e.*, the posterior distribution of the parameters of the ARCH (5) model, which was adjusted to the Telebras series considering the main analysis of an empirical Bayesian model.

Figure 5 shows the estimated volatility *versus* the observed volatility (values of z_t^2). When the model is appropriate to the data, *i.e.*, it properly estimates the volatility of the series, we expect the graphic representation of this relationship to show those values grouped in a straight line. In this respect, this behavior can be observed in both cases (complete and empirical models).
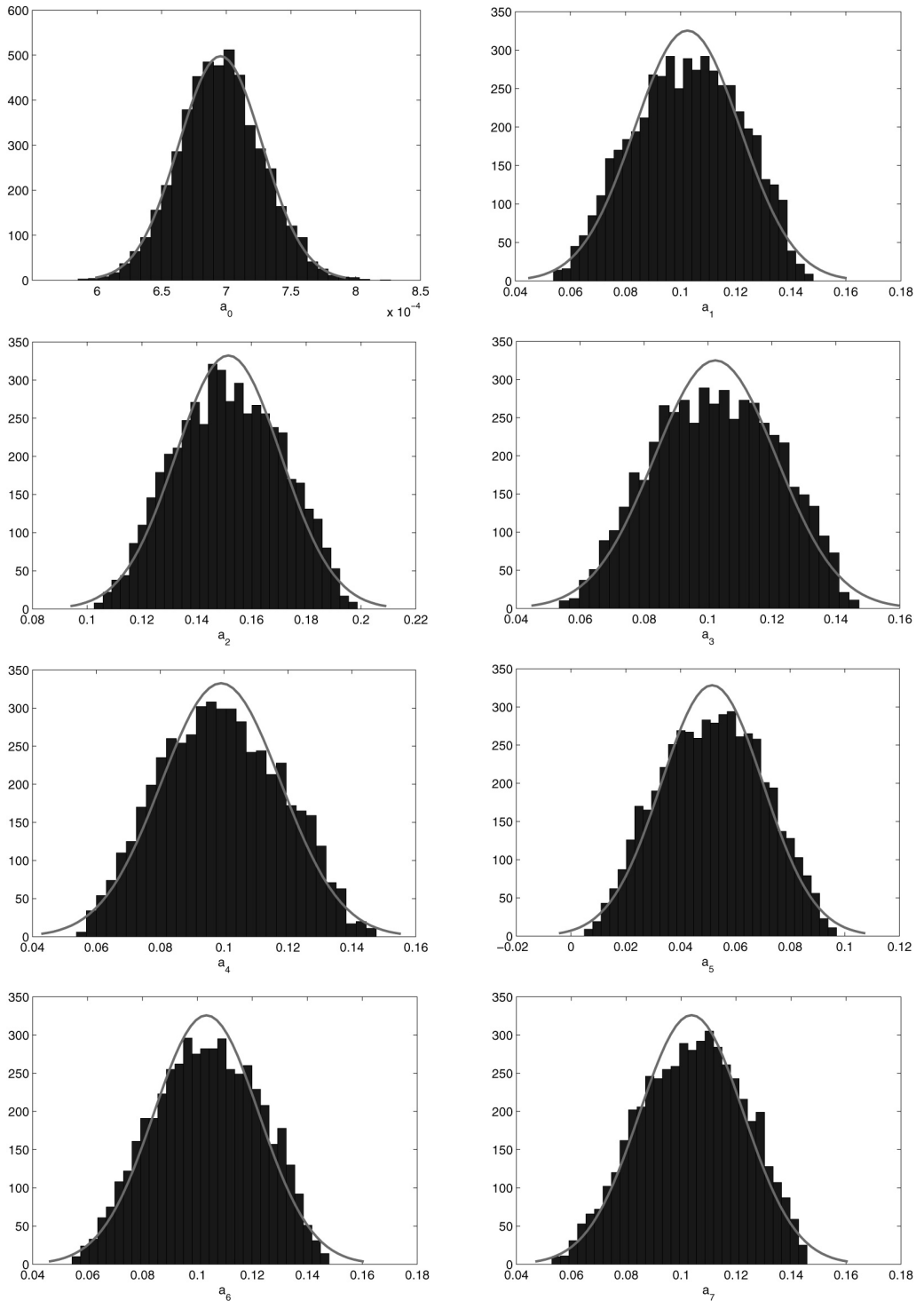
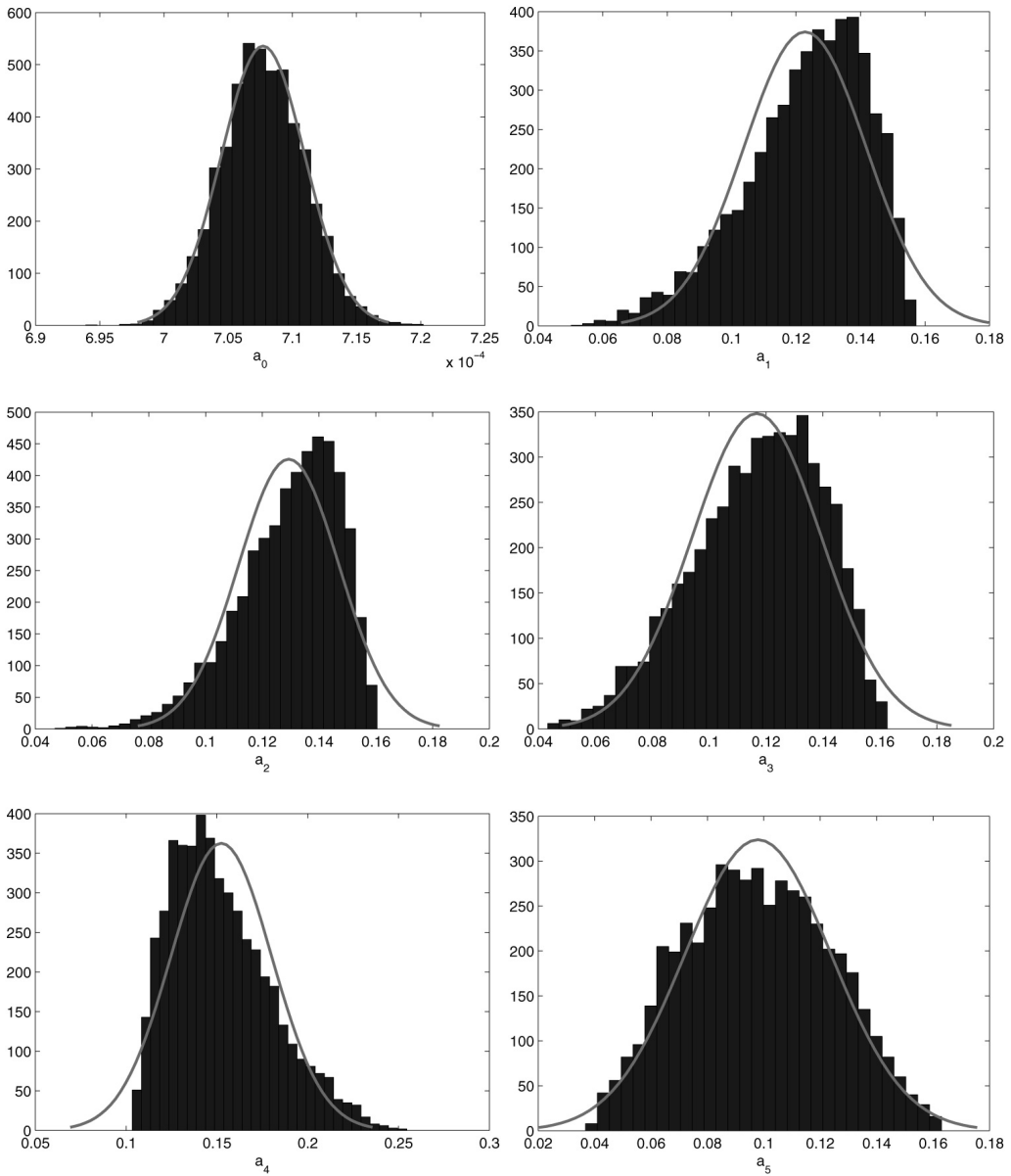**Figure 3** – Posterior distribution of the parameters – CBM.

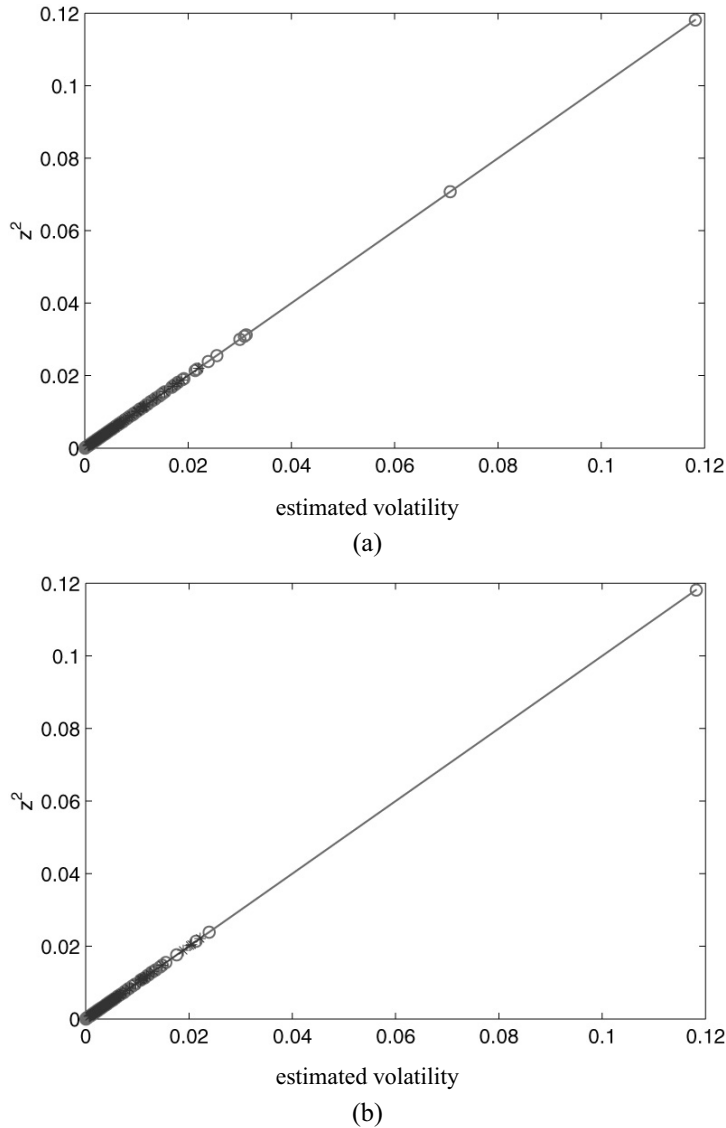**Figure 4** – Posterior distribution of the parameters – EBM.

(a)

(b)

Figure 5 – Estimated volatility *versus* z_t^2: Complete Bayesian model (a) and empirical Bayesian model (b).

As the BIC and other selection criteria similar to it are influenced by sample size, the criterion based on predictive ordinate density is the best for the final comparison between the two Bayesian procedures presented.

Thus, although the plots in Figure 5 show that both approaches presented very close results regarding fit, the graphs about predictive ordinate (and values of POC) in Figure 6 highlight that the EBM best fitted the data of daily Telebras share prices, and it is even more parsimonious than the CBM. This fact suggests that the use of *a prior* information can lead to a more representative model of that data series.
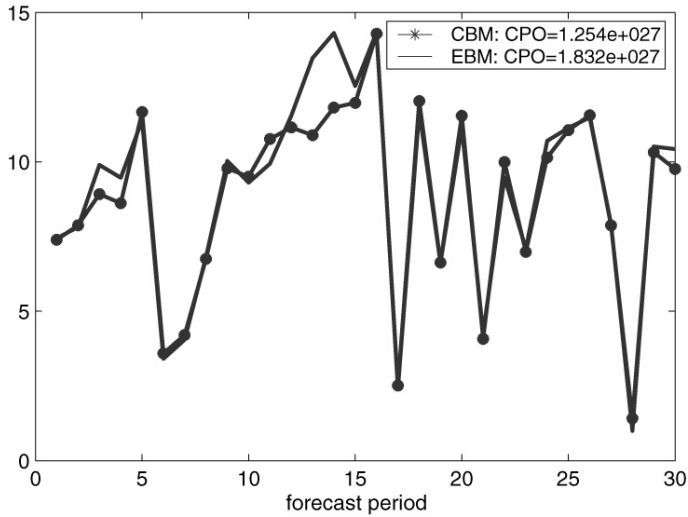
Figure 6 – Predictive ordinate criterion (c_m, $m = 30$): Complete Bayesian model and empirical Bayesian model.

Finally, Figure 7 illustrates the estimated volatility from the best model, EBM. It can be seen that events were detected generating significant behavioral changes in the series of returns during the period studied, often irregular, of which we can highlight the crisis in Mexico in February and March 1995.
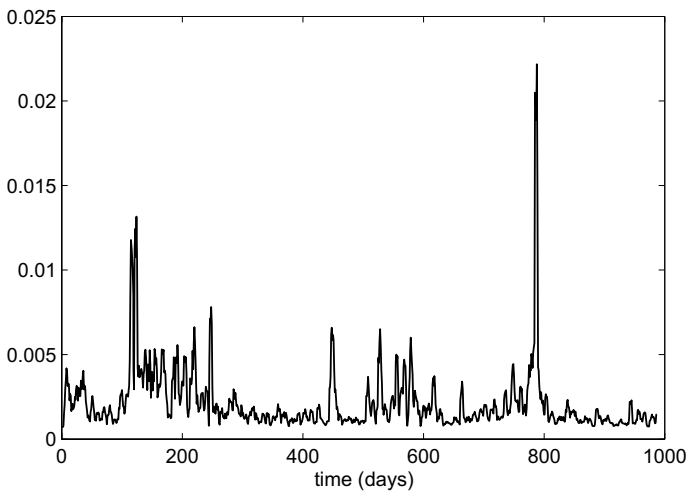

Figure 7 – Estimated volatility – Telebras series.

8   CONCLUDING REMARKS

In general, it is important to emphasize the feasibility of the Bayesian approach in the inference of the parameters of ARCH(q) processes, since it allows the possibility of incorporating the

experience of experts in finance, which usually is a relevant issue in the analysis of economic and financial series.

In this context, studies show that empirical Bayesian models present similar results to those presented by complete Bayesian models, but the former have the advantage of being easily integrated for various cases.

In this paper, we performed an empirical Bayesian analysis, through a detailed study. However, due to the complexity of this analysis, posterior estimates could only be found numerically and, so we used MCMC simulation algorithms.

The results were compared to those obtained by using the complete Bayesian model and showed that the two methods have adjusted ARCH processes with different orders. We noted that the empirical Bayesian model provides a more parsimonious adjustment to the Telebras series, *i.e.*, with a smaller number of parameters for the ARCH process, compared to the complete Bayesian model. Moreover, the selection criterion of models based on the ordinate predictive density showed that the empirical Bayesian model provided better adjustment than the complete Bayesian model.

It is worth noting that the use of empirical prior distributions and of reparametrization contributed to faster convergence of the process of inference of the parameters of ARCH(q) models using MCMC simulation algorithms.

9   ACKNOWLEDGMENTS

REFERENCES

[1] AKAIKE H. 1974. A new look at the statistical identification model. *IEEE Trans on Automatic Control*, **19**: 716–723.

[2] ANDRADE MG & OLIVEIRA SC. 2011. A comparative study of Bayesian and Maximum Likelihood approaches for ARCH models with evidence from Brazilian financial series. *New Mathematics and Natural Computation*, **7**: 347–361.

[3] AUSÍN MC & GALEANO P. 2007. Bayesian estimation of the Gaussian mixture GARCH model. *Computational Statistics and Data Analysis*, **51**: 2636–2652.

[4] BARNDORFF-NIELSEN OE & COX DR. 1994. *Inference and Asymptotics.* Chapman-Hall, London.

[5] BARRETO GA, OLIVEIRA SC & ANDRADE MG. 2008. Estimação de Parâmetros de Modelos ARCH(p): Abordagem Bayes-MCMC *versus* Máxima Verossimilhança. *Revista Brasileira de Estatística*, **69**: 7–24.

[6] BERGER JO & PERICCHI LR. 2004. Training samples in objective Bayesian model selection. *The Annals of Statistics*, **32**: 841–869.

[7] BOLLERSLEV T. 2008. Glossary to ARCH (GARCH). *CREATES Research Paper 2008-49.* Available at SSRN: <http://ssrn.com/abstract=1263250>.

[8]   BOLLERSLEV T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**: 307–327.

[9]   BOLLERSLEV T, CHOU RY & KRONER KF. 1992. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, **52**: 5–59.

[10] BOX GE, JENKINS GM & REINSEL G. 1994. *Time Series Analysis: Forecasting and Control.* Third Edition. Englewood Cliffs: Prentice Hall, San Francisco.

[11] BOX GE & TIAO GC. 1973. *Bayesian Inference in Statistical Analysis.* Addison-Wesley, New York.

[12] CARLIN B & CHIB S. 1995. Bayesian Model Choice via Markov Chain Monte Carlo Methods. *Journal of the Royal Statistical Society*, **57**: 473–484.

[13] CARLIN B & LOUIS T. 2000. *Bayes and Empirical Bayes Methods for Data Analysis.* Texts in Statistical Science Series, Chapman and Hall, London.

[14] CHIB S & GREENBERG E. 1995. Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, **49**: 327–335.

[15] COSTA PHS & BAIDYA TKN. 2001. Propriedades estatísticas das séries de retornos das principais ações brasileiras. *Pesquisa Operacional*, **21**: 61–87.

[16] DEGIANNAKIS S & XEKALAKI E. 2004. Autoregressive Conditional Heteroscedasticity (ARCH) Models: A review. *Quality Technology and Quantitative Management*, **1**: 271–324.

[17] ENDERS W. 2009. *Applied Econometric Times Series.* 3rd edition, John Wiley & Sons, New York.

[18] ENGLE R. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation. *Econometrica*, **50**: 987–1007.

[19] GELFAND AE & SMITH AF. 1990. Sampling – Based Approaches to Calculating Marginal Densities. *Journal of the Statistical Association*, **5**: 398–409.

[20] GELMAN A, CARLIN J, STERN H & RUBIN D. 1995. *Bayesian Data Analysis.* London: Chapman & Hall.

[21] GEWEKE J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (with discussion). In: BERGER J, BERNARDO J, DAWID A & SMITH A. (editors), *Bayesian Statistics*, **4**: 164–193.

[22] GEWEKE J. 1989a. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica*, **57**: 1317–1339.

[23] GEWEKE J. 1989b. Exact Predictive Densities for Linear Models with ARCH Disturbances. *Journal of Econometrics*, **40**: 63–86.

[24] GEWEKE J. 1986. Exact Inference in the Inequality Constrained Normal Linear Regression Model. *Journal of Applied Econometrics*, **1**: 127–141.

[25] GIAKOUMATOS S, DELLAPORTAS P & POLITIS D. 2005. Bayesian Analysis of the Unobserved ARCH Model using Auxiliary Variable Sampling. *Statistics and Computing*, **15**: 103–112.

[26] GILKS WR, RICHARDSON S & SPIEGELHALTER DJ. 1996. *Markov Chain Monte Carlo in Practice.* London: Chapman and Hall.

[27] KOOP G. 1994. Bayesian Semi-nonparametric ARCH Models. *Review of Econometrics and Statistics*, **76**: 176–181.

[28] KOTZ S, BALAKRISHNAN N & JOHNSON NL. 2000. Continuous Multivariate Distributions, Models and Applications, 2nd edition, *Wiley Series in Probability and Statistics: Applied Probability and Statistics.* Wiley-Interscience, **1**: 1–722.

[29] MIGON HS & MAZUCHELI J. 1999. Modelos GARCH Bayesianos: Métodos Aproximados e Aplicações. *The Brazilian Review of Econometrics*, **19**: 111–138.

[30] MORETTIN PA. 2008. *Econometria Financeira.* Edgard Blucher Ltda, São Paulo.

[31] NAKATSUMA T. 2000. Bayesian Analysis of ARMA-GARCH Models: A Markov Chain Sampling Approach. *Journal of Econometrics*, **95**: 57–69.

[32] NAKATSUMA T & TSURUMI H. 1996. ARMA-GARCH Models: Bayes Estimation *versus* MLE, and Bayes Non-Stationarity Test. Technical report, Department of Economics, New Brunswick.

[33] OLIVEIRA SC. 2005. Modelos Estocásticos com Heterocedasticidade para Séries Temporais em Finanças. Doctoral Thesis, Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação.

[34] PITT M & WALKER S. 2005. Constructing stationary time series models using auxiliary variables with applications. *Journal of the American Statistical Association*, **100**: 554–564.

[35] POLASEK W. 2001. MCMC Methods for Periodic AR-ARCH Models. Technical report, University of Basel.

[36] POLASEK W & KOZUMI H. 2000. The VAR-VARCH Model: A Bayesian Approach. Technical report, University of Basel.

[37] SÁFADI T & ANDRADE MG. 2007. Abordagem Bayesiana de Modelos de Séries Temporais. Mini-curso, Universidade Federal do Rio Grande do Sul, 12ª Escola de Séries Temporais e Econometria.

[38] SCHWARZ G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, **6**: 461–464.

[39] SPIEGELHALTER D, BEST N, CARLIN B & VAN DER LINDE A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, **64**: 583–639.