BAYESIAN ANALYSIS OF DYNAMIC FACTOR MODELS USING MULTIVARIATE T DISTRIBUTION

Larissa Ribeiro de ANDRADE[1]
Daniel Furtado FERREIRA[2]
Thelma SÁFADI[2]
Lúcia Pereira BARROSO[3]

■ ABSTRACT: The multivariate t models are symmetric and have heavier tail than the normal distribution and produce robust inference procedures for applications. In this paper, the Bayesian estimation of a dynamic factor model is presented, where the factors follow a multivariate autoregressive model, using the multivariate t distribution. Since the multivariate t distribution is complex, it was represented in this work as a mix of the multivariate normal distribution and a square root of a chi-square distribution. This method allowed the complete define of all the posterior distributions. The inference on the parameters was made taking a sample of the posterior distribution through a Gibbs Sampler. The convergence was verified through graphical analysis and the convergence diagnostics of Geweke (1992) and Raftery and Lewis (1992).

■ KEYWORDS: Factor models; Gibbs sampler; multivariate t.

1 Introduction

A main problem in building a model for a vector time series is that the number of parameters grows with number of series. Therefore, models for reduction of dimensions are needed to model a large number or time series. Dynamic factor models are a useful tool for dimension reduction in time series.

[1]Universidade Estadual de Londrina - UEL, Departamento de Estatística, Caixa Postal 6001, CEP: 86051-990, Londrina, PR, Brazil. E-mail: *larissa.mat@gmail.com*
[2]Universidade Federal de Lavras - UFLA, Departamento de Estatística, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brazil. E-mail: *danielff@des.ufla.br; safadi@des.ufla.br*
[3]Universidade de São Paulo - USP, Instituto de Matemática e Estatística, Departamento de Estatística, CEP: 05508-090, São Paulo, SP, Brazil. E-mail: *lpbarroso@gmail.com*

Sáfadi and Peña (2007) used dynamic factor models to analyze a vector of q time series, assuming that factors follow a $VAR(p)$ model. The authors developed a full Bayesian approach, considering independent errors with normal distribution. The inference was made through a Gibbs Sampler to obtain the Markov Chain Monte Carlo (MCMC) and the convergence was verified through the convergence diagnostics of Gelman and Rubin (1992). In that work, all the posterior distributions were developed.

The statistics inference is mainly developed using the normal model. However, in some cases, this model is not appropriate; for example, when the data belongs to a distribution with heavier tail than the normal distribution, or when there is influence of outliers. Fisher (1956) pointed out that slight differences in the specification of the distribution of the model errors may play havoc on the resulting inferences.

To examine the effects on inference, Fisher (1960) analyzed Darwin's data under normal theory and later under a symmetric non-normal distribution. In the last decades, there has been an increased interest in multivariate t distribution as a robust alternative to normal distribution.

Motivated by this, Borges (2008) dealt with the model presented by Sáfadi and Peña (2007) using multivariate t errors, but the author could not evaluate the posterior distributions explicitly, because the density was highly complex.

The solution reached by Borges (2008) was to use the Griddy-Gibbs Sampler. This is a numerical method to generate random samples from a distribution, even when the posterior density is unknown.

In this study, an alternative to this numerical method is presented. This is achieved by using the multivariate t distribution as a mix of a multivariate normal distribution and a square root of a chi-square distribution. With this alternative, the posterior distributions can be obtained; therefore, the samples from these distributions can be built with no need for a numerical method.

The purpose of this study is to develop a full Bayesian approach to estimate the dynamic factor model for the main stock indices in the world. It is worth mentioning that the empirical analysis is only an illustration of the methodology. The choice of financial data is to complementing the studies of Sáfadi and Peña (2007), and Borges (2008). This methodology can also be applied to data in the biological sciences.

2 Materials and methods

In this paper, the factor model given by the following two equations is considered:

$$
\begin{aligned}
\boldsymbol{y}_t =& \boldsymbol{\beta} + \boldsymbol{C}\boldsymbol{f}_t + \boldsymbol{e}_t; \\
\boldsymbol{f}_t =& \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i} + \boldsymbol{w}_t,
\end{aligned}
\tag{1}
$$

where:

- $\boldsymbol{y}_t$ is a $q \times 1$ vector of observed time series;
- $\boldsymbol{f}_t$ is a $k \times 1$ vector which follows a multivariate autoregressive model $VAR(p)$;
- $\boldsymbol{\rho}_i$ is autoregressive matrix with $\boldsymbol{\rho}_i = diag(\rho_{i1}, ..., \rho_{ik})$ for all $i = 1, ..., p$; and $\{\rho_{1j}, \rho_{2j}, ..., \rho_{pj}\}$ satisfy the stationary condition for all $j = 1, 2, ..., k$;
- $\boldsymbol{\beta}$ is the $q \times 1$ mean vector;
- $\boldsymbol{C}$ is a $q \times k$ matrix of unknown constants;
- $\boldsymbol{e}_t$ are independent q-vectors;
- $\boldsymbol{w}_t$ are independent k-vectors;
- $\boldsymbol{e}_t$ and $\boldsymbol{w}_s$ are independent for all t and s.

As is well-known, the k-factor model must be further constrained to define only a single vector of parameters, free from identification problems. A solution adopted by Sáfadi and Peña (2007), and used here, is to constrain the matrix C, so that it is a block lower triangular matrix, assumed to be of full rank. That is,

$$\boldsymbol{C} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ c_{21} & 1 & 0 & \dots & 0 \\ c_{31} & c_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & c_{k3} & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & c_{q3} & \dots & c_{qk} \end{pmatrix}.$$

This form provides both the identification and useful interpretation of the factor model. From a Bayesian point of view, this is equivalent to assigning fixed values to these matrix C parameters with probability one, and in the analysis they are not estimated.

The main objective of this paper is to fit the dynamic factor model to a financial vector time series which, in general, has a behavior away from the normal case; thus, the use of the multivariate t distribution is more suitable.

The data in studied are the daily values of stock indices: S&P500 (USA), Shanghai Comp Index (China), FTSE100 (UK), CAC40 (France), DAX (Germany), S&P/TSX (Canada), Bovespa (Brazil), Merval (Argentina) and Nikkei 225 (Japan); between 2008 and 2011. Consequently, there are $q = 9$ time series, with $n = 650$ observations.

The index returns were analyzed in this paper, because, according to Morettin and Toloi (2008), this process exhibits both stationarity and ergodicity.

The return is defined by $r_i = log(I_t) - log(I_{t-1})$, where I_t is the stock indices at time t.

The returns plot, the histograms and the normal Q-Q plot, were analyzed to identify a possible distance from the normal distribution, as expected, since it comes to financial data.

The number of factors k was selected through analysis of the graph of eigenvalues versus the number of factors (scree plot).

It is assumed that the errors, $\boldsymbol{e}_t$ and $\boldsymbol{w}_t$, of the models have a multivariate t distribution in the Bayesian inference.

2.1 Multivariate *t* distribution

Student's t distribution is defined as the distribution of the random variable X with

$$X = \frac{Z}{\sqrt{Q/\nu}} \tag{2}$$

where $Z \sim N(0,1)$ and $Q \sim \chi^2_\nu$ are independent, and the density function is given by

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\left(1 + \frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}. \tag{3}$$

In the multivariate case, the density function of a random vector $\boldsymbol{Y} = [Y_1, Y_2, \ldots, Y_q]^\top$ with $\boldsymbol{Y} \sim t_q(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ is, in general, given by

$$f_{\boldsymbol{Y}}(\boldsymbol{y}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{\Gamma\left(\frac{\nu+q}{2}\right)}{(\pi\nu)^{\frac{q}{2}}\Gamma\left(\frac{\nu}{2}\right)|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu}(\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right]^{-\frac{\nu+q}{2}}, \tag{4}$$

where $\boldsymbol{Y}$ has mean $\boldsymbol{\mu}$ and covariance matrix $\nu\boldsymbol{\Sigma}/(\nu-2)$ when $\nu > 2$.

Borges (2008) tried to use the density given for equation (4) to compute the posterior distribution; however, it was not possible due to its complexity. The author used a numerical method of integration to obtain the MCMC through the Griddy-Gibbs Sampler.

The alternative to the numerical method in this paper was to use the multivariate t distribution as a mix of a normal multivariate and a square root of a chi-square variable.

Consider an auxiliary variable $A \sim \sqrt{\chi^2_\nu/\nu}$, with χ^2_ν being a random variable with chi-square distribution with $\nu > 0$ degrees of freedom, and $\boldsymbol{Y} \sim t_q(\boldsymbol{\beta}, \boldsymbol{\Sigma}, \nu)$; it can be shown that the joint density function of $\boldsymbol{Y}$ and A is defined by

$$\begin{aligned} f_{\boldsymbol{Y},A}(\boldsymbol{y}, a; \nu, \Sigma) =& (2\pi)^{-q/2}a^q|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}a^2(\boldsymbol{y}-\boldsymbol{\beta})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\beta})\right\} \\ &\times \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}}a^{\nu-1}e^{-\nu a^2/2}. \end{aligned} \tag{5}$$

The concept of augmented data was used in order to apply this joint density in the Bayesian analysis.

2.2 Bayesian analysis

Once it has been diagnosed that the data do not follow the multivariate normal distribution, one can assume three sets for the distributions of the model errors, as shown in Table 1.

Table 1 - Possible sets to the distribution of errors

Set	$\boldsymbol{e}_t$	$\boldsymbol{w}_t$
1	$t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$	$t_k(\mathbf{0}, \boldsymbol{I}, \nu_w)$
2	$t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$	$N_k(\mathbf{0}, \boldsymbol{I})$
3	$N_q(\mathbf{0}, \boldsymbol{\Sigma})$	$t_k(\mathbf{0}, \boldsymbol{I}, \nu_w)$

In these three sets, for $p = 1$, $Cov(\boldsymbol{f}_t) = \boldsymbol{\Lambda}$, with $\boldsymbol{\Lambda} = \boldsymbol{\rho}_1 \boldsymbol{\Lambda} \boldsymbol{\rho}_1^\top + \boldsymbol{I}_k$.

To compute the posterior distributions of each parameter in $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho})$, the likelihood function is needed, which is given by the product of the likelihoods of $\boldsymbol{y}_t$ and $\boldsymbol{f}_t$.

When the error, $\boldsymbol{e}_t$ or $\boldsymbol{w}_t$, has a multivariate t distribution, the likelihood function must be augmented by $A_{1t} \sim \sqrt{\chi^2_{\nu_e}/\nu_e}$ or $A_{2t} \sim \sqrt{\chi^2_{\nu_w}/\nu_w}$, respectively.

Therefore, for $\boldsymbol{e}_t \sim N_q(\mathbf{0}, \boldsymbol{\Sigma})$, the density used to calculate the likelihood is

$$f(\boldsymbol{y}_t) = (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t)^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t)\right\}, \quad (6)$$

and for $\boldsymbol{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$, the density is augmented as follows

$$\begin{aligned} f(\boldsymbol{y}_t, a_{1t}) =& (2\pi)^{-\frac{q}{2}} a_{1t}^{(q+\nu_e-1)} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \frac{\nu_e^{\frac{\nu_e}{2}} e^{-\frac{\nu_e a_{1t}^2}{2}}}{\Gamma(\frac{\nu_e}{2}) 2^{\frac{\nu_e}{2}-1}} \\ & \times \exp\left\{-\frac{a_{1t}^2}{2}(\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t)^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t)\right\}. \end{aligned} \quad (7)$$

For $\boldsymbol{w}_t \sim N_k(\mathbf{0}, \boldsymbol{I})$ the normal density used is

$$f(\boldsymbol{f}_t) = (2\pi)^{-\frac{k}{2}} |\boldsymbol{I}_k|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i})^\top (\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i})\right\}, \quad (8)$$

and for $\boldsymbol{w}_t \sim t_k(\mathbf{0}, \boldsymbol{I}, \nu_w)$ the multivariate t distribution is given by the mix as follows

$$f(\boldsymbol{f}_t, a_{2t}) = (2\pi)^{-\frac{k}{2}} a_{2t}^{(k+\nu_w-1)} |\boldsymbol{I}_k|^{-1/2} \frac{\nu_w^{\nu_w/2} e^{-\nu_w a_{2t}^2/2}}{\Gamma(\nu_w/2) 2^{\nu_w/2-1}}$$
$$\times \exp\left\{ -\frac{a_{2t}^2}{2} (\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i})^\top (\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}) \right\}. \quad (9)$$

Thus, since the random sample, in n instants, is $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$, the vector of latent factors is $\boldsymbol{F} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$, and also, when the auxiliary variables are needed, $\boldsymbol{A}_1 = (a_{11}, \ldots, a_{1n})$ and $\boldsymbol{A}_2 = (a_{21}, \ldots, a_{2n})$, the augmented likelihoods are obtained, in each set, by the products shown in the Table 2.

Table 2 - Likelihood functions for each set

Set	**Likelihood function**
1	$f(\boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1, \boldsymbol{A}_2 \vert \boldsymbol{\theta}) = f(\boldsymbol{Y}, \boldsymbol{A}_1) f(\boldsymbol{F}, \boldsymbol{A}_2) = \prod_{t=p+1}^{n} f(\boldsymbol{y}_t, a_{1t}) f(\boldsymbol{f}_t, a_{2t})$
2	$f(\boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1 \vert \boldsymbol{\theta}) = f(\boldsymbol{Y}, \boldsymbol{A}_1) f(\boldsymbol{F}) = \prod_{t=p+1}^{n} f(\boldsymbol{y}_t, a_{1t}) f(\boldsymbol{f}_t)$
3	$f(\boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_2 \vert \boldsymbol{\theta}) = f(\boldsymbol{Y}) f(\boldsymbol{F}, \boldsymbol{A}_2) = \prod_{t=p+1}^{n} f(\boldsymbol{y}_t) f(\boldsymbol{f}_t, a_{2t})$

As Lee and Shi (2000) did, a procedure based on data augmentation was developed. The essential idea is to determine posterior distributions for all unknown parameters conditioned on the latent factor and, then, the conditional distribution of the latent factor given the observable data and the other parameters. That is, the observable data are 'augmented' by samples from the conditional distribution for the factor given the data and the parameters of the model. Specifically, the joint posterior distribution for the unknown parameters and the unobserved factors can be sampled by using a Markov Chain Monte Carlo procedure on the full set of conditional distributions. Additionally, in each set, the observable data are 'augmented' again by samples from the auxiliary variables that are needed.

Independent prior distributions given as

$$P(\boldsymbol{A}_1)P(\boldsymbol{A}_2)P(\boldsymbol{\beta})P(\boldsymbol{C})P(\boldsymbol{\rho})P(\boldsymbol{F}) \propto constant$$

and $\gamma_i = \sigma_i^{-2} \sim \Gamma(\alpha_0, \beta_0)$ are assumed, so that the distribution of σ_i^2 is an Inverse Gamma, for each component of $\boldsymbol{\Sigma} = diag(\sigma_1^2, \ldots, \sigma_q^2)$. To implement the Gibbs sampler, it is necessary to derive the full conditional posterior distribution of each parameter given all the others parameters.

Here, the full conditional distribution for $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho})$ and $\boldsymbol{F}$, and also for $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, when necessary, is presented. The posterior distributions for sets 1, 2 and 3 are presented in Tables 3, 4 and 5, respectively. The derivations for the posterior distributions are similar to the normal case, as expected, and were not shown in this paper.

The number of degrees of freedom could been considered unknown and also be estimated. According to Lange *et al.* (1989), low values of degrees of freedom have shown good performance for small samples. Thus, the determination of low values of degrees of freedom was preferred, since higher values would lead to approximately normal distributions, when, in fact, an alternative distribution is wanted.

Since the posterior distributions are known, it is possible to apply the Gibbs sampler to get the Markov chains. However, it is not known which set is more suitable to determine the right posterior distributions. Therefore, the analysis is performed in the three sets and the model that best fits is chosen according to a model comparison method, such as the Bayes factor.

Therefore, initially, chains of size $N = 5000$, for each parameter, were generated for each set through a Gibbs Sampler. The Raftery and Lewis (1992) diagnostic was used to estimate the size needed (NT) for the Markov chains.

New chains were generated with size NT. The Raftery and Lewis (1992) diagnostic was used again to estimate the burn-in and thin size in order to eliminate the efect of initial values and to get a sample approximate uncorrelated sample for each parameter. The Geweke (1992) diagnostic was applied to verify the convergence of the chains.

After appropriately performing the burn-in and the thin and verifying the convergence of the chains, the parameter estimate was obtained through the mean of the respective Markov chain values. The highest probability density (HPD) intervals were computed, the chain traces were graphically analyzed and the posterior densities were graphically determined.

3 Results and discussion

The time series analyzed were the daily values of stock indices: S&P500 (USA), Shanghai Comp Index (China), FTSE100 (UK), CAC40 (France), DAX (Germany), S&P/TSX (Canada), Bovespa (Brazil), Merval (Argentina) and Nikkei 225 (Japan); between 2008 and 2011. Consequently, there were $q = 9$ time series, with $n = 650$ observations. The graphs of the returns of the indices are presented in Figure 1.

The decision of working with returns was made once, according to Morettin and Toloi (2008), they are better to work with because they are free of scale and have interesting properties, such as stationarity and ergodicity. Indeed, it can be seen in Figure 1 that the series does not present tendency or seasonality.

Through the scree plot, presented in Figure 2, the number of factors $k = 2$ was chosen, since from that value the differences in variance are not significant.

In the histograms presented in Figure 3, the estimates for the density functions of each variable can be seen. It can be seen that the shape of estimated densities is

Table 3 - Posterior distributions for the set 1

Parameter	**Posterior distribution**
A_{1t} $(t = 1, \ldots, n)$	$A_{1t}^2 \vert \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_2 \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta}{2}\right)$ $\alpha = q + \nu_e + 1;$ $\delta = (\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t)^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t) + \nu_e$
A_{2t} $(t = 1, \ldots, p)$	$A_{2t}^2 \vert \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1 \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta}{2}\right)$ $\alpha = k + \nu_w + 1$ $\delta = \boldsymbol{f}_t^\top \boldsymbol{\Lambda}^{-1} \boldsymbol{f}_t + \nu_w$
A_{2t} $(t = p+1, \ldots, n)$	$A_{2t}^2 \vert \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1 \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta}{2}\right)$ $\alpha = k + \nu_w + 1$ $\delta = \left(\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}\right)^\top \left(\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}\right) + \nu_w$
$\boldsymbol{\beta}$	$\boldsymbol{\beta} \vert \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1, \boldsymbol{A}_2 \sim N_q(\alpha, \boldsymbol{\Sigma}\delta)$ $\alpha = \delta \sum_{t=p+1}^{n} a_{1t}^2 (\boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{f}_t)$ $\delta = \left(\sum_{t=p+1}^{n} a_{1t}^2\right)^{-1}$
$\boldsymbol{C}$	$\boldsymbol{C}_i^* \vert \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1, \boldsymbol{A}_2 \sim N_k(\alpha_i; \sigma_i^2 \delta)$ $\alpha_i = \delta \sum_{t=p+1}^{n} a_{1t}^2 \boldsymbol{f}_t (y_{it} - \beta_i)$ $\delta = \left(\sum_{t=p+1}^{n} a_{1t}^2 \boldsymbol{f}_t \boldsymbol{f}_t^\top\right)^{-1}$
$\boldsymbol{\Sigma}$	$\gamma_i \vert \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1, \boldsymbol{A}_2 \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta_i}{2}\right)$ $\alpha = n - p + 2\alpha_0$ $\delta_i = 2\beta_0 + \sum_{t=p+1}^{n} a_{1t}^2 (y_{it} - \beta_i - \boldsymbol{C}_i^{*\top} \boldsymbol{f}_t)^\top (y_{it} - \beta_i - \boldsymbol{C}_i^{*\top} \boldsymbol{f}_t)$
$\boldsymbol{\rho}$	$\boldsymbol{\rho}_\nu \vert \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1, \boldsymbol{A}_2 \sim N_{kp}(\alpha, \delta)$ $\alpha = \delta \sum_{t=p+1}^{n} \boldsymbol{B}_t^\top \boldsymbol{f}_t a_{2t}^2$ $\delta = \left(\sum_{t=p+1}^{n} \boldsymbol{B}_t^\top \boldsymbol{B}_t a_{2t}^2\right)^{-1}$
$\boldsymbol{F}_t$ $(t = 1, \ldots, p)$	$\boldsymbol{f}_t \vert \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{Y}, \boldsymbol{A}_1, \boldsymbol{A}_2 \sim N_k(\alpha, \delta)$ $\alpha = \delta(\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta}) a_{1t}^2)$ $\delta = (\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1}$
$\boldsymbol{F}_t$ $(t = p+1, \ldots, n)$	$\boldsymbol{f}_t \vert \boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{C}, \boldsymbol{Y}, \boldsymbol{A}_1, \boldsymbol{A}_2 \sim N_k(\alpha, \delta)$ $\alpha = \delta\left(\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta}) a_{1t}^2 + a_{2t}^2 \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}\right)$ $\delta = (\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} a_{1t}^2 + \boldsymbol{I}_k a_{2t}^2)^{-1}$

Table 4 - Posterior distributions for the set 2

Parameter	Posterior distribution
A_{1t} $(t=1,\ldots,n)$	$A_{1t}^2 \vert \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{F} \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta}{2}\right)$ $\alpha = q + \nu_e + 1;$ $\delta = (\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t)^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{f}_t) + \nu_e$
$\boldsymbol{\beta}$	$\boldsymbol{\beta} \vert \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1 \sim N_q(\alpha, \boldsymbol{\Sigma}\delta)$ $\alpha = \delta \sum_{t=p+1}^n a_{1t}^2 (\boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{f}_t)$ $\delta = \left(\sum_{t=p+1}^n a_{1t}^2\right)^{-1}$
$\boldsymbol{C}$	$\boldsymbol{C}_i^* \vert \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1 \sim N_k(\alpha_i; \sigma_i^2 \delta)$ $\alpha_i = \delta \sum_{t=p+1}^n a_{1t}^2 \boldsymbol{f}_t (y_{it} - \beta_i)$ $\delta = \left(\sum_{t=p+1}^n a_{1t}^2 \boldsymbol{f}_t \boldsymbol{f}_t^\top\right)^{-1}$
$\boldsymbol{\Sigma}$	$\gamma_i \vert \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1 \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta_i}{2}\right)$ $\alpha = n - p + 2\alpha_0$ $\delta_i = 2\beta_0 + \sum_{t=p+1}^n a_{1t}^2 (y_{it} - \beta_i - \boldsymbol{C}_i^{*\top} \boldsymbol{f}_t)^\top (y_{it} - \beta_i - \boldsymbol{C}_i^{*\top} \boldsymbol{f}_t)$
$\boldsymbol{\rho}$	$\boldsymbol{\rho}_\nu \vert \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_1 \sim N_{kp}(\alpha, \delta)$ $\alpha = \delta \sum_{t=p+1}^n \boldsymbol{B}_t^\top \boldsymbol{f}_t$ $\delta = \left(\sum_{t=p+1}^n \boldsymbol{B}_t^\top \boldsymbol{B}_t\right)^{-1}$
$\boldsymbol{F}_t$ $(t=1,\ldots,p)$	$\boldsymbol{f}_t \vert \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{Y}, \boldsymbol{A}_1 \sim N_k(\alpha, \delta)$ $\alpha = \delta(\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta}) a_{1t}^2)$ $\delta = (\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} a_{1t}^2 + \boldsymbol{\Lambda}^{-1})^{-1}$
$\boldsymbol{F}_t$ $(t=p+1,\ldots,n)$	$\boldsymbol{f}_t \vert \boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{C}, \boldsymbol{Y}, \boldsymbol{A}_1 \sim N_k(\alpha, \delta)$ $\alpha = \delta\left(\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\beta}) a_{1t}^2 + \sum_{i=1}^p \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}\right)$ $\delta = (\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} a_{1t}^2 + \boldsymbol{I}_k)^{-1}$

Table 5 - Posterior distributions for the set 3

Parameter	Posterior distribution
A_{2t} $(t = 1, \ldots, p)$	$A_{2t}^2 \| \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{F} \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta}{2}\right)$ $\alpha = k + \nu_w + 1$ $\delta = \boldsymbol{f}_t^\top \boldsymbol{\Lambda}^{-1} \boldsymbol{f}_t + \nu_w$
A_{2t} $(t = p+1, \ldots, n)$	$A_{2t}^2 \| \boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{F} \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta}{2}\right)$ $\alpha = k + \nu_w + 1$ $\delta = \left(\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}\right)^\top \left(\boldsymbol{f}_t - \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}\right) + \nu_w$
$\boldsymbol{\beta}$	$\boldsymbol{\beta} \| \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_2 \sim N_q(\alpha, \delta)$ $\alpha = (n-p)^{-1} \sum_{t=p+1}^{n} (\boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{f}_t)$ $\delta = \boldsymbol{\Sigma}(n-p)^{-1}$
$\boldsymbol{C}$	$\boldsymbol{C}_i^* \| \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_2 \sim N_k(\alpha_i; \delta_i)$ $\alpha_i = \left(\sum_{t=p+1}^{n} \boldsymbol{f}_t \boldsymbol{f}_t^\top\right)^{-1} \sum_{t=p+1}^{n} \boldsymbol{f}_t (y_{it} - \beta_i)$ $\delta_i = \sigma_i^2 \left(\sum_{t=p+1}^{n} \boldsymbol{f}_t \boldsymbol{f}_t^\top\right)^{-1}$
$\boldsymbol{\Sigma}$	$\gamma_i \| \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\rho}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_2 \sim \Gamma\left(\frac{\alpha}{2}, \frac{\delta_i}{2}\right)$ $\alpha = n - p + 2\alpha_0$ $\delta_i = 2\beta_0 + \sum_{t=p+1}^{n} (y_{it} - \beta_i - \boldsymbol{C}_i^{*\top} \boldsymbol{f}_t)^\top (y_{it} - \beta_i - \boldsymbol{C}_i^{*\top} \boldsymbol{f}_t)$
$\boldsymbol{\rho}$	$\boldsymbol{\rho}_\nu \| \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{A}_2 \sim N_{kp}(\alpha, \delta)$ $\alpha = \delta \sum_{t=p+1}^{n} \boldsymbol{B}_t^\top \boldsymbol{f}_t a_{2t}^2$ $\delta = \left(\sum_{t=p+1}^{n} \boldsymbol{B}_t^\top \boldsymbol{B}_t a_{2t}^2\right)^{-1}$
$\boldsymbol{F}_t$ $(t = 1, \ldots, p)$	$\boldsymbol{f}_t \| \boldsymbol{\beta}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{Y}, \boldsymbol{A}_2 \sim N_k(\alpha, \delta)$ $\alpha = \delta(\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}_t - \boldsymbol{\beta}))$ $\delta = (\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} + \boldsymbol{\Lambda}^{-1} a_{2t}^2)^{-1}$
$\boldsymbol{F}_t$ $(t = p+1, \ldots, n)$	$\boldsymbol{f}_t \| \boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{C}, \boldsymbol{Y}, \boldsymbol{A}_2 \sim N_k(\alpha, \delta)$ $\alpha = \delta\left(\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}_t - \boldsymbol{\beta}) + a_{2t}^2 \sum_{i=1}^{p} \boldsymbol{\rho}_i \boldsymbol{f}_{t-i}\right)$ $\delta = (\boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} + \boldsymbol{I}_k a_{2t}^2)^{-1}$

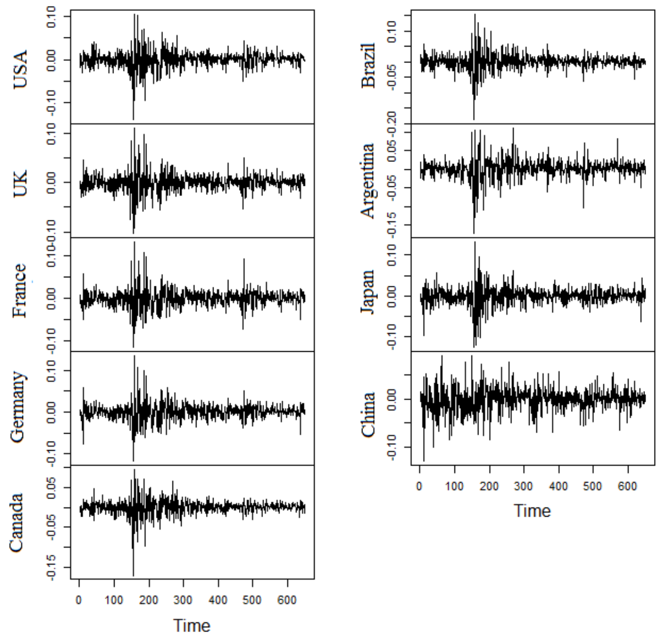

Figure 1 - Daily returns of the stock indices.

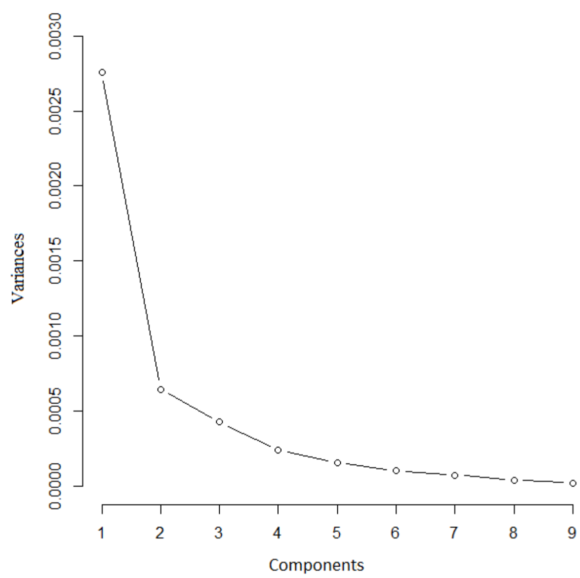

Figure 2 - Screeplot.

similar to the shape of a normal distribution or Student's t distribution.

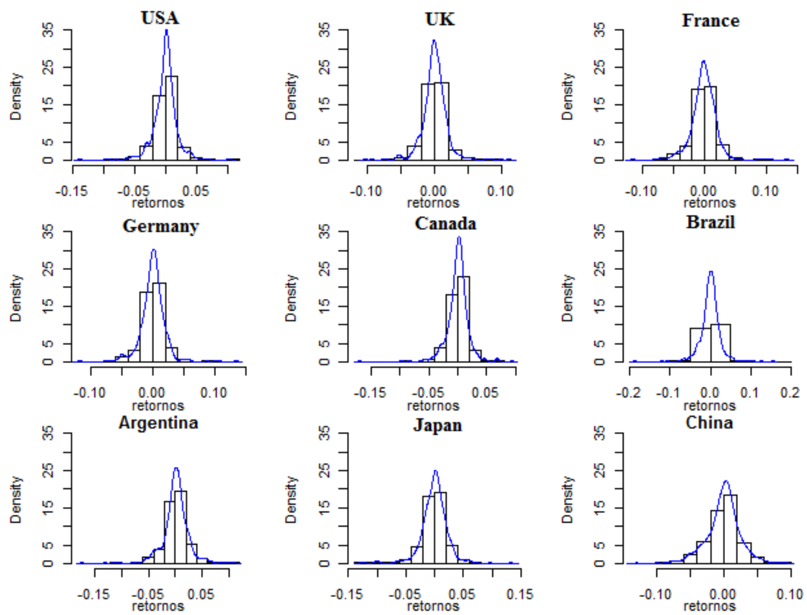

Figure 3 - Histograms of daily returns of stock indices and respective estimated densities.

On the other hand, the normal Q-Q plots, presented in Figure 4, confirm the distance from normal distribution for each variable, as expected, since financial data are being analyzed.

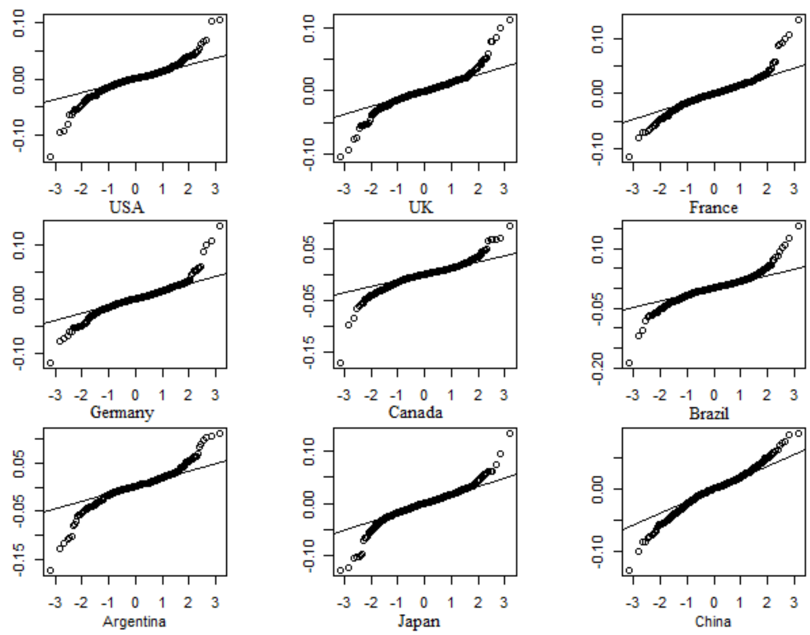

Figure 4 - Normal Q-Q plot of daily returns of stock indices.

Therefore, the methodology proposed to this data set can be applied. The assumption that the violations of normality can be treated in three ways, according to sets 1, 2 and 3 discussed in this paper, may be made in order to adjust the dynamic factor model.

The model, in the three sets, was adjusted for the actual data, and the Bayes factor was calculated in order to compare the models 2 by 2 to choose the best fit.

The results for the logarithm of the Bayes factor are presented in Table 6. The logarithm was used to avoid higher values. Besides, it is necessary for the interpretation of the Bayes factor results, in terms of information theory.

Table 6 - Logarithm of Bayes factor values to compare the models M_0 and M_1

M_0	M_1	$log_{10}B(\boldsymbol{y})$
set 1	set 2	-0.0585
set 2	set 3	167.8849

Inittialy, the models defined by the sets 1 and 2 were compared according to what was defined in Table 7 (JEFFREYS, 1961). The interpretation of the result was negative, that is, in favor of set 2.

Then, the models defined by sets 2 and 3 were compared. The interpretation in that case was decisive in favor of set 2.

Table 7 - Scale of interpretation of Bayes factor

$B(\boldsymbol{y})$	$log_{10}(B(\boldsymbol{y}))$	Interpretation in favor of H_0
< 1	< 0	Negative (in favor of H_1)
1 a 3.2	0 a 0.5	Insignificant
3.2 a 10	0.5 a 1	Significant
10 a 100	1 a 2	Strong
> 100	> 2	Decisive

Therefore, the dynamic factor model was fitted to the data, assuming that $\boldsymbol{e}_t \sim t_q(\mathbf{0}, \boldsymbol{\Sigma}, \nu_e)$ and $\boldsymbol{w}_t \sim N_k(\mathbf{0}, \boldsymbol{I}_k)$ with $\nu_e = 3$. The Bayesian analysis was performed considering the dynamic factor model with $k = 2$ factors, which follows a $VAR(1)$ model. The Markov chains were obtained through the Gibbs sampler with the posterior distributions defined by set 2.

The results from the analysis of the Markov chains are presented in Table 8, with the posterior estimates given by the mean of the chains, the standard deviation of the Markov chains, the HPD intervals and the p value of the convergence test from the Geweke (1992) diagnostic.Values in bold, for p value, indicate the chains that do not converge by the Geweke (1992) diagnostic.

The purpose of adjusting the dynamic factor model to the time series of values of stock indices was to determine a simple and parsimonious model to represent this data set.

Table 8 - Posterior means and standard deviations, HPD intervals, p value from the Geweke (1992) diagnostic, for set 2, for the series of returns of stock indices

Parameters	Mean	Standard deviation	HPD intervals *IL*	HPD intervals *SL*	Geweke (p value)
β_1	-0.0013	0.0405	-0.0814	0.0790	0.7086
β_2	-0.0010	0.0414	-0.0829	0.0802	0.7682
β_3	-0.0003	0.0024	-0.0052	0.0041	0.4459
β_4	0.0000	0.0023	-0.0049	0.0043	0.7245
β_5	0.0002	0.0023	-0.0046	0.0045	0.9433
β_6	0.0003	0.0024	-0.0047	0.0049	0.5006
β_7	0.0011	0.0025	-0.0038	0.0058	0.4357
β_8	-0.0003	0.0023	-0.0052	0.0040	0.9571
β_9	-0.0005	0.0024	-0.0052	0.0042	0.9727
C1	0.0083	0.0588	-0.1147	0.1222	0.6296
C2	0.0102	0.0127	-0.0152	0.0349	0.7401
C3	0.0098	0.0126	-0.0152	0.0341	0.0661
C4	0.0101	0.0123	-0.0140	0.0341	0.3306
C5	0.0134	0.0134	-0.0132	0.0395	0.2277
C6	0.0129	0.0135	-0.0139	0.0389	0.5857
C7	0.0047	0.0127	-0.0208	0.0291	0.3626
C8	0.0032	0.0134	-0.0240	0.0287	0.6275
C9	0.0123	0.0126	-0.0125	0.0369	**0.0378**
C10	0.0111	0.0124	-0.0132	0.0355	**0.0215**
C11	0.0083	0.0121	-0.0157	0.0319	0.4362
C12	0.0111	0.0131	-0.0144	0.0372	0.2170
C13	0.0118	0.0132	-0.0144	0.0377	0.3378
C14	0.0071	0.0126	-0.0174	0.0321	0.9438
C15	0.0044	0.0132	-0.0217	0.0305	0.2080
σ_1^2	0.1277	0.0213	0.0869	0.1687	0.4714
σ_2^2	0.1311	0.0222	0.0910	0.1758	**0.0243**
σ_3^2	0.0103	0.0006	0.0091	0.0113	**0.0032**
σ_4^2	0.0102	0.0006	0.0091	0.0113	0.7310
σ_5^2	0.0100	0.0006	0.0090	0.0111	0.9266
σ_6^2	0.0107	0.0006	0.0095	0.0118	0.7554
σ_7^2	0.0108	0.0006	0.0097	0.0120	0.6250
σ_8^2	0.0106	0.0006	0.0093	0.0116	0.2874
σ_9^2	0.0111	0.0006	0.0099	0.0123	0.4242
ρ_1	0.0332	0.1955	-0.3516	0.4180	0.3257
ρ_2	0.0312	0.1942	-0.3453	0.4151	**0.0342**

Conclusions

Full analysis of the dynamic factor model, to a vector time series, using the multivariate t distribution was developed.

In order to estimate the parameters of the model the Gibbs sampler was applied to obtain the Markov chains.

This was possible because, despite the complexity of the multivariate t distribution, this variable was used as a mix of multivariate normal distributions and a square root of a chi-square variable, and the calculations were similar to the normal case. So, the posterior distributions were calculated.

The posterior densities estimated are similar to the densities expected according to the posterior distributions calculated, indicating consistency in the results.

In a follow-up study, the authors of this paper intend to work considering unknown degrees of freedom and the order of the autoregressive model $p > 1$.

Acknowledgments

We thank reviewers and editors for their comments and suggestions.

ANDRADE, L. R.; FERREIRA, D. F.; SÁFADI, T.; BARROSO, L. P. Análise bayesiana de modelos de fatores dinâmicos usando distribuição t multivariada. *Rev. Bras. Biom.,* Lavras, v.36, n.1, p.140-156, 2018.

- *RESUMO: Os modelos t multivariados são simétricos e têm caudas mais pesadas do que a distribuição normal e produzem procedimentos de inferência robustos para aplicações. Neste artigo, apresenta-se a estimativa bayesiana de um modelo de fator dinâmico, onde os fatores seguem um modelo autoregressivo multivariado, usando a distribuição t multivariada. Uma vez que a distribuição t multivariada é complexa, ela foi representada neste trabalho como uma mistura da distribuição normal multivariada e uma raiz quadrada de uma distribuição qui-quadrado. Este método permitiu a definição completa de todas as distribuições posteriores. A inferência sobre os parâmetros foi feita tomando uma amostra da distribuição posterior através de uma amostra de Gibbs. A convergência foi verificada através de análise gráfica e os diagnósticos de convergência de Geweke (1992) e Raftery e Lewis (1992).*

- *PALAVRAS-CHAVE: Modelos fatoriais; amostrador de Gibbs; t multivariada.*

References

BORGES, L. C. *Análise Bayesiana do modelo fatorial dinâmico para um vetor de series temporais usando distribuições elípticas.* 2008. 99p. Tese (Doutorado em Estatística) - Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo.

FISHER, S. R. *Statistical Methods and Scientific Inference*. Edinburgh: Oliver & Boyd, 1956.

FISHER, S. R. On some extensions of Bayesian inference proposed by Mr Lindley. *Journal of the Royal Statist Society*, Series B, v.22, p.299-301, 1960.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. *Statistical Science*, Hayward, v.7, n.4, p.457-511, 1992.

GEWEKE, J. Evaluating the accurary of sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Ed.). *Bayesian Statistics 4*. New York: Oxford University Press, 1992. p.625-631.

JEFFREYS, H. *Theory of Probability*. London: Oxford University Press, 1961.

LANGE, K.; LITTLE, J.; TAYLOR, M. Robust statistical modeling using t-distribution. *Journal of the American Statistical Association*, v.84, p.881-896, 1989.

LEE, S.; SHI, J. Joint Bayesian analysis of factor scores and structural parameters in the factor analysis model. *Annals of the Institute of Statistical Mathematics*, v.52, p.722-736, 2000.

MORETTIN, P. A.; TOLOI, C. M. C. *Análise de séries temporais*. São Paulo: Egard Blucher, 2008.

RAFTERY, A. E.; LEWIS, S. Comment: One long run with diagnostics: implementation strategies for markov chain monte carlo. *Statistical Science*, Hayward, v.7, n.4, p.493-497, 1992.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2017.

SÁFADI, T.; PEÑA, D. Bayesian Analysis of dynamic factor models: an application to air polution and mortality in São Paulo, Brazil. *Environmetrics*, v.18, p.1-20, 2007.

Recebido em 07.07.2016.

Aprovado após revisão em 03.10.2017.