Journal of Data Science **3**(2005), 101-122

Bayesian Analysis for Change Points in the Volatility of Latin American Emerging Markets

Rosangela H. Loschi, Cristiano R. Moura and Pilar L. Iglesias
Universidade Federal de Minas Gerais

Abstract: We have extended some previous works by applying the product partition model (PPM) to identify multiple change points in the variance of normal data sequence assuming mean equal to zero. This type of problem is very common in applied economics and finance. We consider the Gibbs sampling scheme proposed in the literature to obtain the posterior estimates or product estimates for the variance and the posterior distributions for the instants when changes take place and also for the number of change points in the sequence. The PPM is used to obtain the posterior behavior of the volatility (measured as the variance) in the series of returns of four important Latin American stock indexes (MERVAL-Argentina, IBOVESPA-Brazil, IPSA-Chile and IPyC-Mexico). The posterior number of change point as well as the posterior most probable partition for each index series are also obtained.

Key words: Gibbs sampling, product partition model, student-t distribution.

1. Introduction

Most methodologies used to analyse structural changes and change points problems assume that the number of change points is a known and fix value. For example, we can cite the threshold models (Chen and Lee, 1995, Geweke and Terui, 1993), the methods based on maximum likelihood estimators considered by Hawkins (2001) and many others. Other authors have studied the one change point problem using a Bayesian approach (Menzefricke, 1981, Hsu, 1984 and Smith, 1975, for example). The product partition model (PPM) developed by Hartigan (1990) introduces more flexibility into the analysis of these problems since it rather considers the number of change points as a random variable. As shown by Barry and Hartigan (1992), by applying the PPM one can easily obtain product estimates for the parameters of interest at each point the time, the posterior distribution of the random partition generated by the change points, and also the posterior distribution of the number of change points. (For multiple change point analysis using Bayesian approach see also Chernoff and Zacks (1964)).

Barry and Hartigan (1993) and Crowley (1997) applied the PPM to the identification of multiple change points in normal means only. Both papers consider Gibbs sampling approaches to obtain only the product estimates. Later, Loschi *et al.* (1999) extended the results from Barry and Hartigan (1993) and Crowley (1997) by applying the PPM to identify multiple changes in both means and variances of normal data and by proposing a Gibbs sampling scheme to compute the posterior distributions of the random partition generated by the change points and the posterior distributions of the number of change points. Quintana and Iglesias (2003) provided a decision-theoretic formulation to PPM and linked it to the Dirichlet process. Loschi *et al.* (2003a) extended even further the PPM by rather assuming a prior distribution for the parameter p that indexes Yao's (1984) cohesions, that is, the probability p of having a change point at any instants of time and by proposing a Gibbs sampling scheme to compute the posterior relevances involved in the product estimates. Despite all flexibility introduced by the PPM in the analysis of change point problems, Loschi and Cruz (2002) showed also that the product estimates may be considerably influenced by the prior specifications for p.

On the other hand, several models proposed to describe the behavior of financial time series assumes that the variance is constant throughout the time. Homoscedasticity hypothesis can be reasonable for efficient markets. However, this assumption seems to be too strong for emerging markets since they are more susceptible to shocks (Mendes (2000), Duarte and Mendes (1997)) which can produce changes in the volatility.

Hsu (1984) considers that the returns of the Dow Jones Industrial Average follows a normal distribution and "under modest efficient market assumptions" (see Hawkins (2001), p. 333) assumes that the mean (μ) is zero. Correa (1998) provides empirical evidence that the mean return in the Chilean market is also zero. Following Correa (1998) assumptions for the Chilean market, since the markets considered in this paper are also emerging markets, we suppose that, conditionally in the variance (σ^2), the returns are normally distributed with mean equal to zero. We also adopt a conjugate inverted-gamma prior distribution for σ^2. As a consequence of these assumptions, the returns are distributed according to a Student-t distribution, which discloses a structure of correlation among the returns and has heavier tails than the normal distribution.

This paper addresses the identification of multiple change points in the variance (which is understood here as a measure of volatility) of data sequences using a Bayesian approach. We apply the PPM to identify multiple change points in the normal variances. It is assumed that only contiguous blocks are possible and that the prior cohesions (Yao's cohesions, 1984) are a truncated geometric distribution with parameter p. We extended some results from Correa (1998) by assuming a beta prior distribution for p and by providing the posterior distributions of the random partition generated by the change points and for the number of change points. The Gibbs sampling scheme proposed by Loschi *et al.* (1999) is considered to obtain the estimates for these posterior distributions. Unlike Correa (1998), the product estimates for the variance are obtained using the method proposed by Loschi *et al.* (2003a). Then, the PPM will be applied to the analysis of four Latin American stock markets. To represent these markets we consider the return of their most important stock prices indexes: the MERVAL (*Índice de Mercado de Valores de Buenos Aires*) of Argentina, the IBOVESPA (*Índice da Bolsa de Valores do Estado de São Paulo*) of Brazil, the IPSA (*Índice de Precios Selectivos de Acciones*) of Chile and the IPyC (*Índice de Precios y Cotizaciones*) of Mexico.

The paper is organized as follows. In Section 2, the PPM and related results are presented following Barry and Hartigan (1992). The PPM is applied to identify multiple change points in normal variances assuming Yao's cohesions. In Section 3, we describe the Loschi *et al.*'s (2003a) computational method to obtain the product estimates of the variances, as well as a Gibbs sampling scheme to compute the posterior distributions for the random partition generated by the change points and for the number of change points. In Section 4, the methodology is applied to the four Latin American indexes aforementioned. Finally, Section 5 concludes the paper.

2. Statistical Models

2.1 The product partition model

Let $X_1, \ldots, X_n$ be a data sequence and consider the index set $I = \{1, \ldots, n\}$. Consider a random partition $\rho = \{i_0, i_1, \cdots, i_b\}$ of the set I, $0 = i_0 < i_1 < \cdots < i_b = n$, and a random variable B which denotes the number of blocks in ρ. Consider that each partition divides the sequence $X_1, \ldots, X_n$ into $B = b$ contiguous subsequences, which will be denoted here by $\mathbf{X}_{[i_{r-1}i_r]} = (X_{i_{r-1}+1}, \ldots, X_{i_r})'$, for $r = 1, \ldots, b$. Let $c_{[ij]}$ be the prior cohesion associated with the block $[ij] = \{i+1, \ldots, j\}$, for $i, j \in I \cup \{0\}$, and $j > i$, which represents the degree of similarity among the observations in $\mathbf{X}_{[ij]}$ and can be interpreted as transition probabilities in the Markov chain defined by the change points (for details, see Barry and Hartigan (1993)).

Let $\theta_1, \ldots, \theta_n$ be a sequence of unknown parameters, such that, conditionally in $\theta_1, \ldots, \theta_n$, the sequence of random variables $X_1, \ldots, X_n$ has conditional marginal densities $f_1(X_1|\theta_1), \ldots, f_n(X_n|\theta_n)$, respectively. The prior distribution of $\theta_1, \ldots, \theta_n$ is constructed as follows. Given a partition $\rho = \{i_0, \ldots, i_b\}$, for $b \in I$, one has that $\theta_i = \theta_{[i_{r-1}i_r]}$, for every $i_{r-1} < i \leq i_r$ and $r = 1, \ldots, b$, and that $\theta_{[i_0 i_1]}, \ldots, \theta_{[i_{b-1} i_b]}$ are independent, with $\theta_{[ij]}$ having prior (block) density $\pi_{[ij]}(\theta)$, $\theta \in \Theta_{[ij]}$, where $\Theta_{[ij]}$ is the parameter space corresponding to the common parameter, say, $\theta_{[ij]} = \theta_{i+1} = \ldots = \theta_j$, which indexes the conditional density of $\mathbf{X}_{[ij]}$.

Hence, we say that the random quantity $(X_1, \ldots, X_n; \rho)$ follows a PPM, denoted by $(X_1, \ldots, X_n; \rho) \sim PPM$, if:

i) the prior distribution of ρ is the following product distribution:

$$P(\rho = \{i_0, \ldots, i_b\}) = \frac{\Pi_{j=1}^{b} c_{[i_{j-1} i_j]}}{\sum_{\mathcal{C}} \Pi_{j=1}^{b} c_{[i_{j-1} i_j]}} \tag{2.1}$$

in which $\mathcal{C}$ is the set of all possible partitions of the set I into b contiguous blocks with endpoints $i_1, \ldots, i_b$, satisfying the condition $0 = i_0 < i_1 < \cdots < i_b = n$, for all $b \in I$;

ii) conditionally on $\rho = \{i_0, \ldots, i_b\}$, the sequence $X_1, \ldots, X_n$ has the joint density given by:

$$f(X_1, \ldots, X_n | \rho = \{i_0, \ldots, i_b\}) = \Pi_{j=1}^{b} f_{[i_{j-1} i_j]}(\mathbf{X}_{[i_{j-1} i_j]}), \tag{2.2}$$

in which $f_{[ij]}(\mathbf{X}_{[ij]}) = \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta$ is the predictive density of the random vector $\mathbf{X}_{[ij]}$, called data factor.

Consequently, if $(X_1, \ldots, X_n; \rho) \sim$ PPM, the number of blocks B in ρ has a prior distribution given by:

$$P(B = b) \propto \sum_{\mathcal{C}_1} \Pi_{j=1}^{b} c_{[i_{j-1} i_j]}, \; b \in I, \tag{2.3}$$

in which $\mathcal{C}_1$ is the set of all partitions of I into b contiguous blocks with endpoint $i_1, \ldots, i_b$, satisfying the condition $0 = i_0 < i_1 < \cdots < i_b = n$.

As shown by Barry and Hartigan (1992), the posterior distributions of ρ and B have the same functional form of the prior distribution, in which the posterior cohesion for the block $[ij]$ is given by

$$c^*_{[ij]} = c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]}). \tag{2.4}$$

Barry and Hartigan (1992) also show that the posterior distributions of θ_k is given by:

$$\pi(\theta_k | \mathbf{X}_{[0n]}) = \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \pi_{[ij]}(\theta_k | \mathbf{X}_{[ij]}), \; k = 1, \ldots, n, \tag{2.5}$$

and that the posterior expectation (or product estimate) of θ_k is given by:

$$E(\theta_k|\mathbf{X}_{[0n]}) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]} E(\theta_k|\mathbf{X}_{[ij]}),\ k = 1, \ldots, n, \tag{2.6}$$

in which $r^*_{[ij]}$ denotes the posterior relevance for the block $[ij]$, that is:

$$r^*_{[ij]} = P([ij] \in \rho|\mathbf{X}_{[0n]}) = \frac{\lambda_{[0i]} c^*_{[ij]} \lambda_{[jn]}}{\lambda_{[0n]}}, \tag{2.7}$$

with $\lambda_{[ij]} = \sum \Pi_{k=1}^{b} c^*_{[i_{k-1} i_k]}$, in which the summation is over all partitions of $\{i+1, \ldots, j\}$ in b blocks with endpoints $i_0, i_1, \ldots, i_b$, satisfying the condition $i = i_0 < i_1 < \cdots < i_b = j$.

2.2 The product partition model for normal variances

In order to specify the PPM for the normal case, assume that there is a sequence of unknown independent parameters $\theta_1 = \sigma_1^2, \ldots, \theta_n = \sigma_n^2$, such that $X_k|\sigma_k^2 \sim \mathrm{N}(0, \sigma_k^2)$, for $k = 1, \ldots, n$. It is also assumed that each common parameter $\theta_{[ij]} = \sigma^2_{[ij]}$, related to the block $[ij]$, has the conjugate inverted-gamma prior distribution denoted by $\sigma^2_{[ij]} \sim \mathrm{InvGam}(a_{[ij]}/2, d_{[ij]}/2)$, $a_{[ij]} > 0$,$d_{[ij]} > 0$, whose density function is given by:

$$f(\sigma^2_{[ij]}) = \frac{(a_{[ij]}/2)^{d_{[ij]}/2}}{\Gamma(d_{[ij]}/2)} (\sigma^2_{[ij]})^{\frac{d_{[ij]}+2}{2}} \exp\left\{\frac{a_{[ij]}}{2\sigma^2_{[ij]}}\right\},\quad \sigma^2_{[ij]} > 0. \tag{2.8}$$

The expected value, the variance and the mode of the distribution of $\sigma^2_{[ij]}$ is given, respectively, by:

$$\left.\begin{array}{rcl} E(\sigma^2_{[ij]}) & = & \frac{a_{[ij]}}{d_{[ij]}-2}, \\ Var(\sigma^2_{[ij]}) & = & \frac{2a^2_{[ij]}}{(d_{[ij]}-2)^2(d_{[ij]}-4)}, \\ Mo(\sigma^2_{[ij]}) & = & \frac{a_{[ij]}}{d_{[ij]}+2}. \end{array}\right\} \tag{2.9}$$

Hence, the conditional distribution of $\theta_{[ij]} = \sigma^2_{[ij]}$, given the observations in $\mathbf{X}_{[ij]}$, is the inverted-gamma distribution given by:

$$\sigma^2_{[ij]}|\mathbf{X}_{[ij]} \sim \mathrm{InvGam}(a^*_{[ij]}/2, d^*_{[ij]}/2), \tag{2.10}$$

where the posterior parameters are:

$$\left\{\begin{array}{rcl} a^*_{[ij]} & = & a_{[ij]} + \sum_{k=i+1}^{j} X_k^2, \quad \text{and} \\ d^*_{[ij]} & = & d_{[ij]} + j - i. \end{array}\right. \tag{2.11}$$

Consequently, it follows that

$$E(\sigma^2_{[ij]}|\mathbf{X}_{[ij]}) = \frac{a^*_{[ij]}}{d^*_{[ij]} - 2} \qquad (\text{if } d^*_{[ij]} > 2). \tag{2.12}$$

The interested reader may find details in O'Hagan (1994).

From (2.12) and (2.6), it follows that the product estimates for the parameter σ^2_k, for $k = 1, \ldots, n$, is given by:

$$\hat{\sigma}^2_k = E(\sigma^2_k|\mathbf{X}_{[0n]}) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]} \frac{a^*_{[ij]}}{d^*_{[ij]} - 2} \qquad (\text{if } d^*_{[ij]} > 2), \tag{2.13}$$

in which $a^*_{[ij]}$ and $d^*_{[ij]}$ are defined as in (2.11).

Let $\mathbf{0}_n$ be the $n \times 1$ vector of zeros and $\mathbf{I}_n$ the $n \times n$ identity matrix. The posterior relevances $r^*_{[ij]}$ can be obtained from (2.7) and (2.4) where the random vector $\mathbf{X}_{[ij]}$ follows a centred $(j-i)$-dimensional Student-t distribution denoted by $\mathbf{X}_{[ij]} \sim t_{j-i}(\mathbf{0}_{(j-i)},\, \mathbf{I}_{j-i};\, a_{[ij]},\, d_{[ij]})$ (see Arellano-Valle and Bolfarine, 1995) with density function given by

$$f(\mathbf{X}_{[ij]}) = \frac{\Gamma\left(\frac{d_{[ij]}+j-i}{2}\right)}{\Gamma\left(\frac{d_{[ij]}}{2}\right)\pi^{\frac{(j-i)}{2}}}\, a_{[ij]}^{d_{[ij]}/2}\left\{a_{[ij]} + \sum_{k=i+1}^{j} X_k^2\right\}^{-(d_{[ij]}+j-i)/2}. \tag{2.14}$$

Notice that, as a consequence of the assumptions made in this section, the observations within the same block are correlated and distributed according to a distribution with heavier tail than the normal distribution.

2.3 The product partition model for Yao's cohesions

In this paper, the prior cohesions proposed by Yao (1984) will also be considered. Let p, for $0 \le p \le 1$, be the probability that a change occurs at any instant in the sequence. Therefore, the prior cohesion for block $[ij]$ is given by:

$$c_{[ij]} = \begin{cases} p(1-p)^{j-i-1}, & \text{if } j < n, \\ (1-p)^{j-i-1}, & \text{if } j = n, \end{cases} \tag{2.15}$$

for all $i,\ j \in I$, $i < j$, which corresponds to the probability that a new change takes place after $j - i$ instants, given that a change has taken place at instant i. These prior cohesions imply that the sequence of change points establishes a discrete renewal process, with occurence times identically distributed with geometric distribution. Such cohesions are appropriate when it is reasonable to assume the past change points are noninformative about the future change points.

In this case, the conditional prior distributions for ρ and B, given p, are obtained from (2.1) and (2.3), respectively, and assume the following form:

$$P(\rho = \{i_0, \ldots, i_b\}|p) = p^{b-1}(1-p)^{n-b}, \tag{2.16}$$

for every partition $\{i_0, \ldots, i_b\}$, satisfying $0 = i_0 < i_1 < \ldots < i_b = n, b \in I$, and

$$P(B = b|p) = C_{b-1}^{n-1} p^{b-1}(1-p)^{n-b}, \;\; b \in I, \tag{2.17}$$

where C_{b-1}^{n-1} denotes the number of distinct partitions of I into b contiguous blocks.

Let assume that p has beta prior distribution with parameters $\alpha > 0$ and $\beta > 0$, denoted by $p \sim \text{Beta}(\alpha, \beta)$. Consequently, the prior distributions for ρ and B are given, respectively, by:

$$P(\rho = \{i_0, \ldots, i_b\}) = \frac{\Gamma(\alpha+\beta)\;\Gamma(\alpha+b-1)\;\Gamma(n+\beta-b)}{\Gamma(\alpha)\;\Gamma(\beta)\;\Gamma(n+\alpha+\beta-1)}, \tag{2.18}$$

and

$$P(B = b) = C_{b-1}^{n-1}\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+b-1)\Gamma(n+\beta-b)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n+\alpha+\beta-1)}. \tag{2.19}$$

It is noticeable from (2.19) that $B \overset{D}{=} Y + 1$ where Y is a random variable which has a Binomial-Beta distribution with parameters $n-1$, α and β, $\alpha > 0$, $\beta > 0$ (see Bernardo and Smith (1994) for details). Thus, the prior mean and variance of B is given, respectively, by:

$$\left.\begin{array}{lcl} E(B) & = & \frac{(n-1)\;\alpha}{\alpha+\beta} + 1, \\ Var(B) & = & \frac{(n-1)\;\alpha\;\beta\;(\alpha+\beta+n-1)}{(\alpha+\beta)^2\;(\alpha+\beta+1)}. \end{array}\right\} \tag{2.20}$$

Assuming the 0-1 loss function (Bernardo and Smith, 1994), that is considering the following penalty function:

$$L(\delta, B) = \lim_{\epsilon \to 0} I_{|B-\epsilon|}([\epsilon, \infty)),$$

the prior Bayes estimator of B is the mode of the prior distribution of B, which is the greatest integer which does not exceed $(n+1)(\alpha-1)/(\alpha+\beta-2)+1$.

Thus, the posterior distribution of ρ is the following distribution:

$$\begin{aligned} & P(\rho = \{i_0, \ldots, i_b\}|\mathbf{X}_{[0n]}) \\ & = \prod_{j=1}^{b} f_{[i_{j-1}i_j]}(\mathbf{X}_{[i_{j-1}i_j]})\;\frac{\Gamma(\alpha+\beta)\;\Gamma(\alpha+b-1)\;\Gamma(n+\beta-b)}{\Gamma(\alpha)\;\Gamma(\beta)\;\Gamma(n+\alpha+\beta-1)}, \end{aligned} \tag{2.21}$$

and the posterior distribution of B is given by:

$$P(B=b|\mathbf{X}_{[0n]}) \\ = C_{b-1}^{n-1}\prod_{j=1}^{b} f_{[i_{j-1}i_j]}(\mathbf{X}_{[i_{j-1}i_j]})\ \frac{\Gamma(\alpha+\beta)\ \Gamma(\alpha+b-1)\ \Gamma(n+\beta-b)}{\Gamma(\alpha)\ \Gamma(\beta)\ \Gamma(n+\alpha+\beta-1)}, \quad (2.22)$$

where $f_{[ij]}(X_{[ij]})$, $i,j\in I$, $i<j$, is the Student-t distribution given in (2.14) if the normal case is under consideration.

3. Computational Methods

3.1 The Gibbs sampling

Gibbs Sampling is a Monte Carlo Markov Chain (MCMC) scheme which considers that the transition kernel is constituted by full conditional distributions (Gilks *et al.*, 1996, Gamerman, 1997). This algorithm was proposed by Geman and Geman (1984) and adapted to Bayesian statistics by Gelfand and Smith (1990). In particular, Gibbs sampling provides the posterior distributions generation scheme which is described in the following.

Let $\mathbf{X}=(\mathbf{X}_1,\ldots,\mathbf{X}_n)$ be a random quantity whose posterior distribution is $\pi(\mathbf{x})$. Suppose that all full conditional distributions $\pi_i(\mathbf{X}_i)=\pi(\mathbf{X}_i|\mathbf{X}_{-i})$ are available, where $\mathbf{X}_{-i}=(\mathbf{X}_1,\ldots,\mathbf{X}_{i-1},\mathbf{X}_{i+1},\ldots,\mathbf{X}_n)$. Given the initial values $\mathbf{X}^{(0)}=(\mathbf{X}_1^{(0)},\ldots,\mathbf{X}_n^{(0)})$, samples of the posterior distribution $\pi(\mathbf{x})$ can be generated by using the following procedure:

1. Initialize the iteration counter of the chain by making $j=0$ and obtain the $j+1$-th value $\mathbf{X}^{(j+1)}$ generating

$$\begin{array}{ccl} \mathbf{X}_1^{(j+1)} & \sim & \pi(\mathbf{x}_1|\mathbf{x}_2^{(j)},\ldots,\mathbf{x}_n^{(j)}) \\ \mathbf{X}_2^{(j+1)} & \sim & \pi(\mathbf{x}_2|\mathbf{x}_1^{(j+1)},\mathbf{x}_3^{(j)},\ldots,\mathbf{x}_n^{(j)}) \\ \vdots & & \vdots \\ \mathbf{X}_n^{(j+1)} & \sim & \pi(\mathbf{x}_n|\mathbf{x}_1^{(j+1)},\mathbf{x}_2^{(j+1)},\ldots,\mathbf{x}_{n-1}^{(j+1)}) \end{array}$$

2. Update the counter by making $j=j+1$ and return to step 1. Repeat this procedure until convergence is reached.

The sequence of values $\mathbf{X}^{(j)}$ generated by Gibbs sampling is a Markov chain. When convergence is reached, $\mathbf{X}^{(j)}$ is value from the posterior distribution $\pi(\mathbf{x})$.

3.2 Gibbs sampling for PPM

As in Section 2, suppose that $p \sim \text{Beta}(\alpha, \beta)$. Assume that, given ρ, $\theta_k \in [ij]$, for $k = 1, \ldots, n$ and $i, j \in I$, $i < j$. Let $\mathbf{X}_{[0n]} = (X_1, \ldots, X_n)$ and $\theta = (\theta_1, \ldots, \theta_n)$ and denote by θ_{-k} the vector $(\theta_1, \ldots, \theta_{k-1}, \theta_{k+1}, \ldots, \theta_n)$. The full conditional distributions of p, ρ, and θ_k, for $k = 1, \ldots, n$ are given, respectively, by:

$$\begin{aligned}
\pi(p|\rho, \theta, \mathbf{X}_{[0n]}) &\propto p^{b+\alpha-2}(1-p)^{n+\beta-b-1}; \\
\pi(\rho|p, \theta, \mathbf{X}_{[0n]}) &\propto \left(\Pi_{j=1}^{b} f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]})\right) p^{b-1}(1-p)^{n-b}; \\
\pi(\theta_k|\rho, p, \theta_{-k}, \mathbf{X}_{[0n]}) &\propto f_{[ij]}(\theta_k|X_{[ij]}).
\end{aligned}$$

Notice that, since all partitions should be considered, it can be very hard to sample directly from the full conditional distribution of ρ if long sequences are considered. An easier way to sample from those distributions is described in the following.

3.3 A Gibbs sampling scheme to the PPM

Consider the auxiliary random quantity U_i, suggested by Barry and Hartigan (1993), which reflects whether or not a change point occurs at time i, that is, $U_i = 1$ if $\theta_i = \theta_{i+1}$ and $U_i = 0$ if $\theta_i \neq \theta_{i+1}$, for $i = 1, \ldots, n-1$. Notice that the random partition ρ is immediately identified by considering the vector $\mathbf{U} = (U_1, \ldots, U_{n-1})$ of these random quantities.

Each partition $\mathbf{U}^s = (U_1^s, \ldots, U_{n-1}^s)$, $s \geq 1$, is generated by using the Gibbs sampling as follows. Starting from an initial value $(U_1^0, \ldots, U_{n-1}^0)$, the rth element at step s, U_r^s, is generated from the conditional distribution $U_r|U_1^s, \ldots, U_{r-1}^s$, $U_{r+1}^{s-1}, \ldots, U_{n-1}^{s-1}$, $p^{(s-1)}$, $\theta^{(s-1)}$; $\mathbf{X}_{[0n]}$,for $r = 1, \ldots, n-1$. In order to generate the samples of $\mathbf{U}$'s above, it is sufficient to consider the following ratio:

$$R_r = \frac{P(U_r = 1|V_r^s, p^{(s-1)}, \theta^{(s-1)}; \mathbf{X}_{[0n]})}{P(U_r = 0|V_r^s, p^{(s-1)}, \theta^{(s-1)}; \mathbf{X}_{[0n]})}$$

for $r = 1, \ldots, n-1$, in which $V_r^s = \{U_1^s = u_1, \ldots, U_{r-1}^s = u_{r-1}, U_{r+1}^{s-1} = u_{r+1}, \ldots, U_{n-1}^{s-1} = u_{n-1}\}$.

Considering the Yao's prior cohesions given in (2.15) and assuming that $p \sim \text{Beta}(\alpha, \beta)$, each value U_r^s, $s \geq 1$, $r = 1, \ldots, n-1$, can be generated by using

$$R_r = \frac{f_{[xy]}(X_{[xy]})\Gamma(n+\beta-b+1)\Gamma(b+\alpha-2)}{f_{[xr]}(X_{[xr]})f_{[ry]}(X_{[ry]})\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}, \tag{3.1}$$

for $b = 1, \ldots, n$, in which:

$$x = \begin{cases} \max\{i,\ \mathbf{s.t.}: 0 < i < r, U_i^s = 0\}, & \text{if } U_i^s = 0, \text{for some} \\ & i \in \{1, \ldots, r-1\} \\ 0, & \text{otherwise,} \end{cases} \tag{3.2}$$

and

$$y = \begin{cases} \min\{i,\ \mathbf{s.t.}: r < i < n, U_i^{s-1} = 0\}, & \text{if } U_i^{s-1} = 0, \text{for some} \\ & i \in \{r+1, \ldots, n-1\} \\ n, & \text{otherwise.} \end{cases} \tag{3.3}$$

Consequently, the criterion of choosing the values U_i^s, $i = 1, \ldots, n-1$ becomes:

$$U_r^s = \begin{cases} 1, & \text{if } R_r \geq \frac{1-u}{u} \\ 0, & \text{otherwise,} \end{cases}$$

for $r = 1, \ldots, n-1$, in which $u \sim \mathcal{U}(0, 1)$. If the normal case described in Section 2 is considered, the joint density $f_{[ij]}(X_{[ij]})$ is the Student-t distribution given in (2.14). (For a degenerate prior distribution to p, see details about the generation of $\mathbf{U}$ in Loschi *et al.* (1999).)

The posterior probability for each particular partition $\rho = \{i_0, i_1, \ldots, i_b\}$ is estimated by computing the proportion of samples of $(U_2, \ldots, U_n)$ such that $U_{i_r} = 0$ for $r = i_k + 1$, $k = 1, \ldots, b-1$, and $U_r = 1$ otherwise. The posterior distribution of B is obtained by noticing that the number of blocks in $\mathbf{U}^s$ is given by $B^s = 1 + \sum_{i=1}^{n-1}(1 - U_i^s)$. Consequently, the posterior distribution of B (or the number of change points $B - 1$) is estimated as follows:

$$P(B = b | \mathbf{X}_{[0n]}) = \frac{\sum_{s=1}^{T} \mathbf{1}\{B^s = b\}}{T},$$

in which T is the net number of vectors $\mathbf{U}$ generated and $\mathbf{1}\{B\}$ denotes the indicator function of B. The method used to compute the product estimates of σ_k^2, for $k = 1, \ldots, n$, is described in the next section.

3.4 Computational procedure for the product estimates

Loschi *et al.* (2003a) obtain the product estimates as follows. Generate T vectors $\mathbf{U} = (U_1, \ldots, U_{n-1})$. The estimate of the posterior relevance of the block $[ij]$, $i, j = 1, \ldots, n, i < j$, is computed as follows:

$$\hat{r}^*_{[ij]} = \frac{M}{T}, \tag{3.4}$$

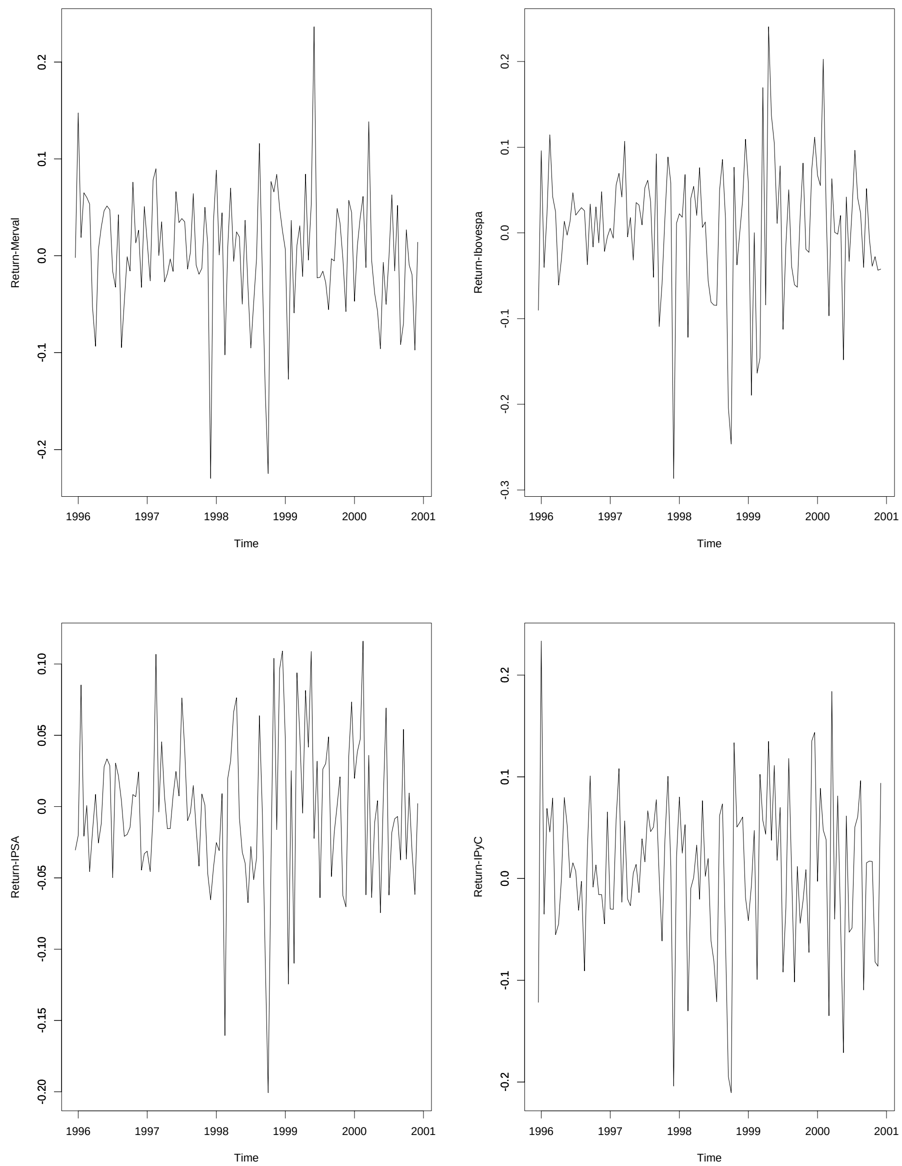

Figure 1: Return series

in which M is the number of vectors $\mathbf{U}$ for which it is observed that $U_i = 0$, $U_{i+1} = \cdots = U_{j-1} = 1$, and $U_j = 0$. The product estimates of σ_k^2, for $k = 1, \ldots, n$, are obtained substituting (3.4) in (2.13), respectively.

4. Application to Latin American Emerging Markets

In this section we focus on the identification of multiple change points in the volatility (variance) of the stock indexes of four important Latin American markets: the MERVAL (*Índice de Mercado de Valores de Buenos Aires*) of Argentina, the IBOVESPA (*Índice da Bolsa de Valores do Estado de São Paulo*) of Brazil, the IPSA (*Índice de Precios Selectivos de Acciones*) of Chile and the IPyC (*Índice de Precios y Cotizaciones*) of Mexico. We apply the methodology presented in the previous sections to analyse the behavior of these indexes within the period from October 31, 1995 to October 31, 2000, recorded fortnightly. A return series is defined by using the transformation $R_t = (P_t - P_{t-1})/P_{t-1}$, where P_t is the price in the month t. Defined in this way, the returns within each block can be considered normally distributed with mean equal to zero, given the volatility (see Correa (1998) for empirical evidence). MERVAL, IBOVESPA, IPSA and IPyC return series are plotted in Figure 1.

We notice from Figure 1 that the behavior of all these indexes suggests the existence of some changes in the variance. Our purpose here is to show that within the period from 31 October, 1995 to 31 October, 2000 the four return series present change points in their volatility - that is, we show that MERVAL, IBOVESPA, IPSA and IPyC series possess volatility clusters. We consider the PPM approach described in Section 2.

The prior distributions and cohesions related to MERVAL, IBOVESPA, IPSA and IPyC returns are formulated in accordance with the facts reported by Loschi *et al.* (1999) and Correa (1998) for the Chilean market and Duarte and Mendes (1997) for emerging Latin American markets. Brazilian, Argentinian and Mexican markets are also emerging markets and, like the Chilean market, more susceptible to the political scenario than developed markets (Mendes, 2000) justifying these specifications. As for the Chilean market, we also assume that changes in the behavior of MERVAL, IBOVESPA, IPSA and IPyC return series are a consequence of the receipt of not previously anticipated information, so that past change points are non-informative concerning future change points (see Mandelbrot, 1963). Hence, the prior cohesions presented in (2.15) are an adequate choice for the four stock market indexes we will analyse.

We suppose that returns within the same block are conditionally independent and distributed according to the normal distribution $\mathcal{N}(0, \sigma^2_{[ij]})$. We also adopt the natural conjugate prior distribution for the parameter $\sigma^2_{[ij]}$ which, in this case, is the inverted-gamma distribution given in (2.8). Consequently, we are assuming that the returns are distributed according to a Student-t PPM (Loschi *al.*, 1999), which has heavier tails than the normal distribution and discloses a structure of correlation amongst the returns within the same block similar to that one obtain

Table 1: Parameters and descriptive statistics for the prior distribution of the volatility

Index	$a_{[ij]}$	$d_{[ij]}$	Mean	Variance	Mode
MERVAL	0.001	6	2.50×10^{-4}	6.25×10^{-5}	1.25×10^{-4}
IBOVESPA	0.001	2	-	-	2.50×10^{-4}
IPSA	0.001	8	1.67×10^{-4}	1.39×10^{-5}	1.00×10^{-4}
IPyC	0.001	6	2.50×10^{-4}	6.25×10^{-5}	1.25×10^{-4}

by using ARCH model (Loschi *et al.*, 2003b).

Table 1 presents the descriptive statistics for prior distributions of $\sigma^2_{[ij]}$ for each index obtained from (2.9). These distributions are ploted in Figure 2.

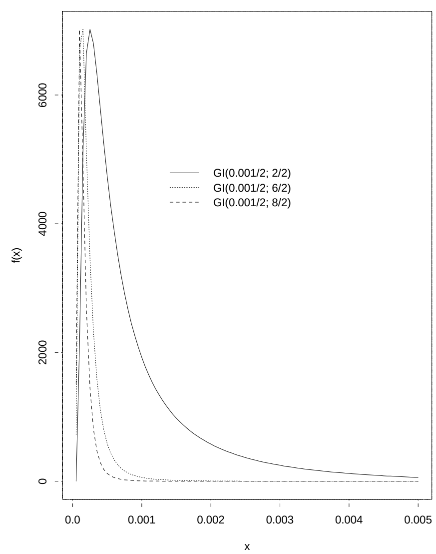

Figure 2: Prior distribution for the volatility

Notice from Figure 2 and Table 1 that we are assuming that in the prior evaluation the IBOVESPA has the highest volatility (the prior mode is 2.5×10^{-4}) and the IPSA has the smallest volatility (the prior mode is 1.0×10^{-4}). The volatility for MERVAL and IPyC are considered the same (the prior mode is 1.25×10^{-4}). Observe also that there is less certainty about the volatility of IBOVESPA since the variance for the volatility is the highest. These prior specification are in accordance with Duarte and Mendes's (1997) statements for Latin American indexes.

According to some expert's opinion, changes in the behavior of stock returns is mainly a consequence of crises or events that occur in other countries. Three great financial crises envolving emerging markets occurred in January, 1995 (Mexico's Crisis), August, 1997 (Asia's Crisis) and July, 1998 (Russia's Crisis). Also, in January, 1999, the Minas Gerais (Brazil) State Governor stopped paying Minas Gerais's debt with other countries. These important events are country specific. However, they can spread out across countries with a similar economy eventually producing changes in their behavior (Lopes and Migon, 2002). It is also expected that the policies adopted by the governments during and after theses crises produced changes in the behavior of the economy. In the prior evaluation we are assuming that MERVAL, IBOVESPA, IPSA and IPyC could experience changes as a consequence of (at least) these events. Thus, we assume that the probability of having a change in any instant p has a beta prior distribution with parameters $\alpha = 5$ and $\beta = 50$. This prior distribution has modal value equal to 0.091 and concentrate most of its mass in small values of p. Consequently, from equation (2.20), we can also observe that the expected number of change points in the prior evaluation for all indexes is 10.82 and the variance is high (standard deviation is 5.53) which means that we are not very sure about the number of change points. The modal value of the number of blocks is 10.0 which means that the most probable number of change points in the four indexes is 9.0. That is, for us the four indexes are equally susceptible to shocks and to the political atmosphere.

We generate 10,000 samples of 0-1 vectors with dimension 119, starting from a sequence of zeros. After convergence has been reached, we discarded the initial 4,000 interations. Since a small correlation is observed we consider lag of 10 and worked with a sample of size 600.

From Figure 3 we can notice that the product estimates for the volatility in each fortnight is smaller for IPSA followed, in general, by MERVAL and IPyC. On the whole, IBOVESPA presents the highest volatility. It is also noticeable that the product estimates for the volatility of IPSA, MERVAL and IPyC present similar behavior. The volatility of MERVAL presents an important change in the 2nd fortnight, August, 1997. Its volatility reaches its highest value in the 2nd fortnight, November, 1997 remaining at this level until the 1st fortnight, October, 1998. From the 2nd fortnight, December, 1996 to the 1st fortnight, August, 1997, MERVAL experiences its period of lowest volatility. Another change can be observed in the 2nd fortnight, June, 1999. IPSA presents an important change in its volatility in the 2nd fortnight, September, 1997 reaching its highest value in the 2nd fortnight, December, 1997. Decreases in the volatility of IPSA are observed in the 2nd fortnight, October, 1998, the 2nd fortnight, December, 1999 and in the 1st fortnight, June, 2000. The smallest volatility is observed from the

1st fortnight, February, 1997 to the 2nd fortnight, March, 1997. The volatility of IPyC increases strongly from the 1st fortnight, April, 1997, to the 2nd fortnight, September, 1997. The maximum is reached in the 2nd fortnight, November, 1997 remaining at this level until the 2nd fortnight, July, 2000. IPyC experiences its smallest volatility from the 1st fortnight, 1997 to the 2nd fortnight, March, 1997. The volatility of IBOVESPA starts increasing in the 2nd fortnight, May, 1997 and its decrease starts in the 2nd fortnight, May, 1999. IBOVESPA experiences its highest volatilities from the 2nd fortnight, September, 1997 to the 2nd fortnight, October, 1997 and from the 1st fortnight, July, 1998 to the 2nd fortnight, April, 1999. From the 2nd fortnight, March, 1996 to the 1st fortnight, November, 1996 IBOVESPA experiences its smallest volatility. In that period the volatility of IPSA and IBOVESPA were very close.

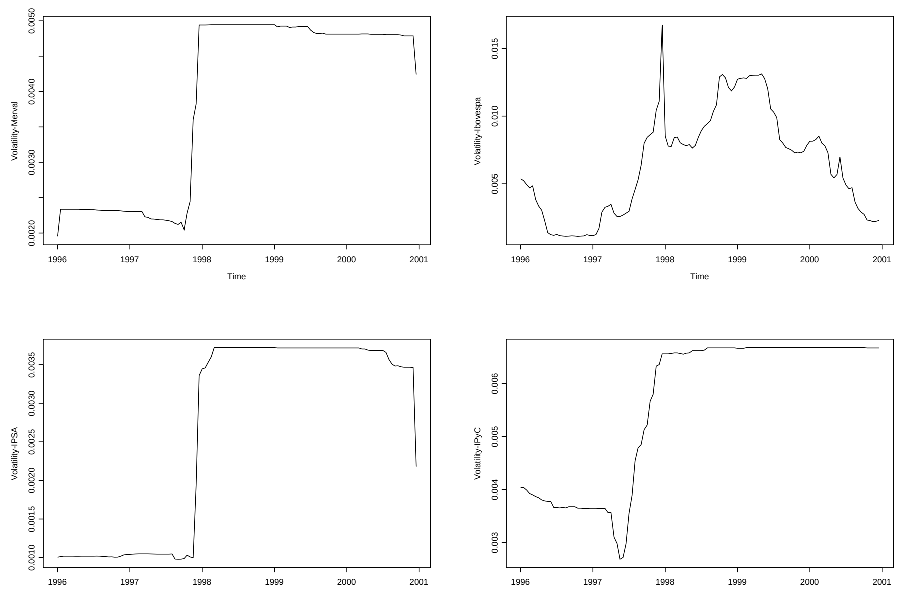

Figure 3: Product estimates for the volatility

Figure 4 presents the product estimates (solid lines) of the fortnightly volatilities for MERVAL, IBOVESPA, IPSA and IPyC return series. These estimates are contrasted with the square return series (dotted lines). We can notice that the product estimates are in agreement with the square returns behavior.

It is noticeable from Figure 5 and Table 2 that for MERVAL, IPSA and IPyC the posterior distribution of the number of blocks are very similar. All these distributions are asymetric, there are unique modes and they typically concentrate the most of their mass in small values.

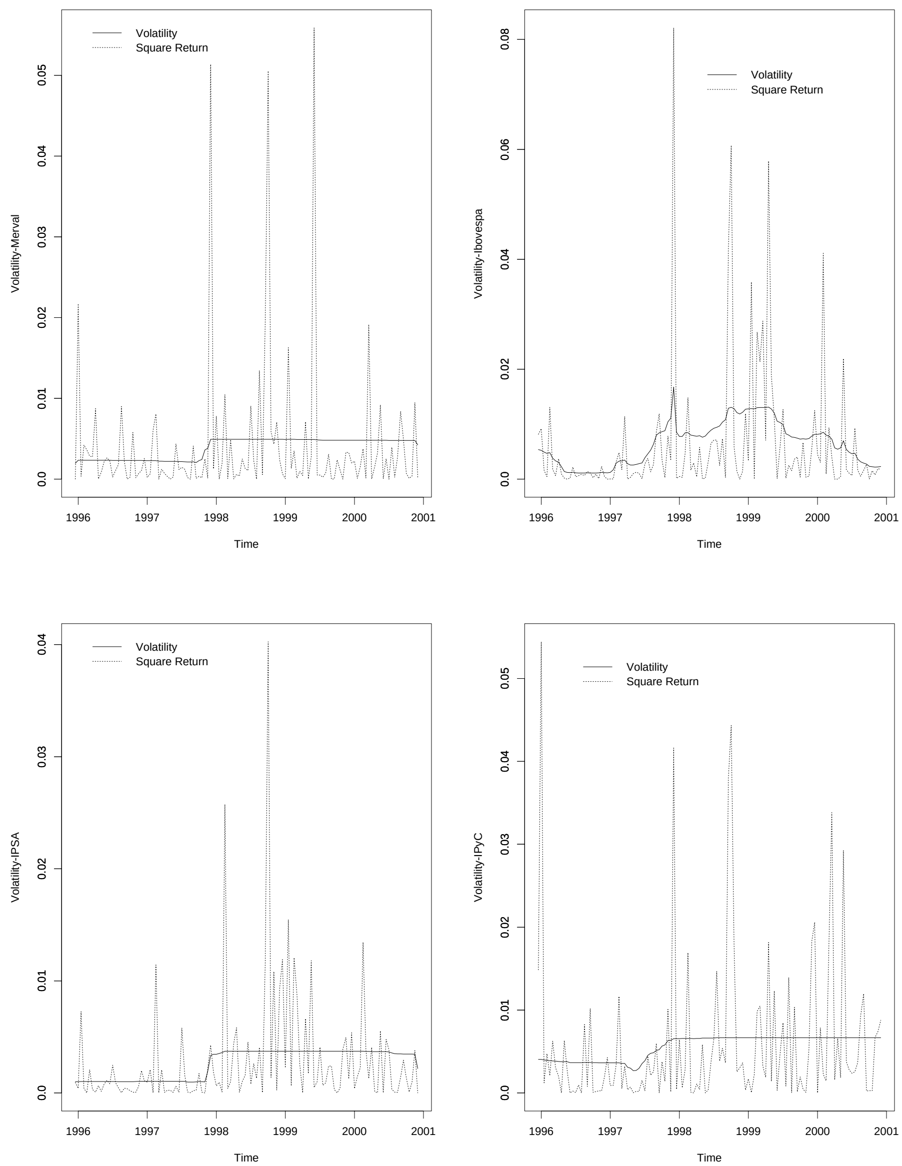

Figure 4: Product estimates for the volatility and square return series

MERVAL, IPSA and IPyC present mode equal to 2 which means that con-

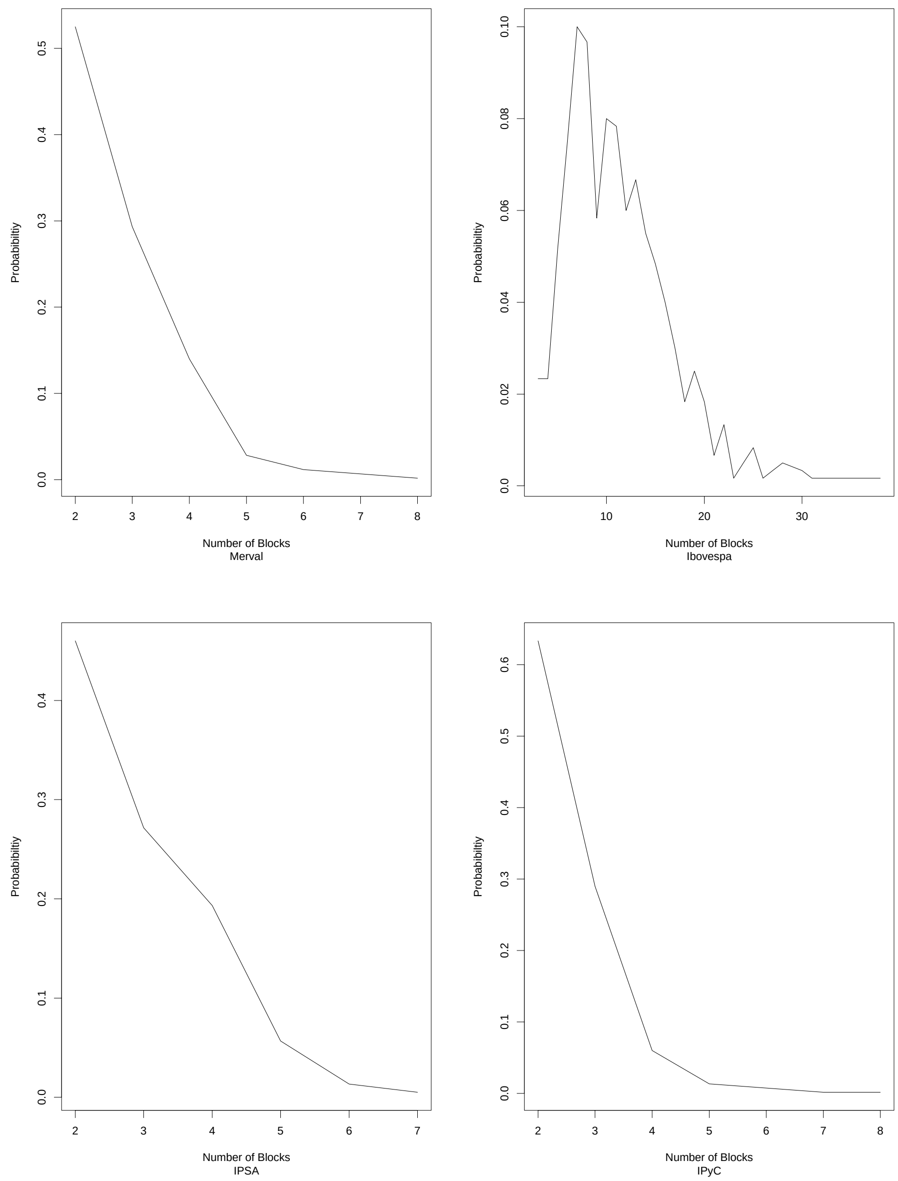

Figure 5: Product distribution for the number of Blocks

sidering the 0-1 loss function (see Section 2.3) the posterior Bayes estimates for the number of change points in such indexes are 1 - which is much smaller than 9, the prior Bayes estimate for the number of change points. It can be also observed that, since the variance of the posterior distribution of the number of change points is smallest for IPyC, there is less uncertainty about the number of changes in this index. IBOVESPA presents the highest mode (mode = 7) and mean (mean = 11.22) which means that the Brazilian stock market was more susceptible to changes in that period. Observe that, assuming the 0-1 loss function, the posterior Bayes estimates of the number of change points in the IBOVESPA is also smaller than the prior Bayes estimates.

Table 2: Descriptive statistics – Posterior distribution of the number of blocks

Index	Mean	Variance	Mode
MERVAL	2.75	0.84	2
IBOVESPA	11.22	27.50	7
IPSA	2.91	1.07	2
IPyC	2.47	0.53	2

Table 3 presents the prior and posterior probabilities for the posterior mode of the number of blocks for the four indexes. Notice, for example, that in spite of having a low prior probability for MERVAL, IPSA and IPyC series experiencing only one change point (5.5%), the posterior probability for this event increases strongly reaching 53%, 46% and 63.0%, respectively. For the IBOVESPA the posterior probability for the posterior mode also increases reaching 10% only. For IBOVESPA, we also observe that the posterior probability of the number of blocks is 7 (the posterior mode) or more is 82.66%. For the other index, the posterior probability of the number of blocks is 2 (the posterior mode) or more is 100.00%.

Table 3: Prior and posterior probabilities for the posterior mode of B

Index	Posterior mode	Prior prob.	Posterior prob.
MERVAL	2	0.055	0.53
IBOVESPA	7	0.066	0.10
IPSA	2	0.055	0.46
IPyC	2	0.055	0.63

Table 4 presents the posterior most probable partition for MERVAL, IBOVES PA, IPSA and IPyC series. We can perceive that the posterior most probable partition of each index indicates the presence of only one change point. MERVAL,IPSA and IPyC present change point in very close instants. For example, the posterior most probable partition for MERVAL is $\{0, 47, 120\}$ which means that MERVAL experiences changes in its behavior in the observation 48 (2nd fortnight, September, 1997). IBOVESPA behaves differently presenting change in the 2nd fortnight, April, 2000.

Table 4: The posterior most probable partition and their prior and posterior probabilities

Index	Partition	Prior prob.	Posterior prob.
MERVAL	{0,47,120}	4.6×10^{-4}	0.2233
IBOVESPA	{0,109,120}	4.6×10^{-4}	0.0017
IPSA	{0,47,120}	4.6×10^{-4}	0.4333
IPyC	{0,45,120}	4.6×10^{-4}	0.1117

Notice that the change points identified by the PPM in MERVAL, IPSA and IPyC indexes are close to a important international event – Asia's crisis in August. Despite of the good performance of the PPM, some other Bayes estimates could be considered in case the interest is not a retrospective analysis as it was considered in this paper (see Quintana and Iglesias, 2003).

5. Final Remarks

We applied the PPM to identify multiple change points in the variance of data sequence which are normally distributed, given the variance. We assumed an inverted-gamma prior distribution for the variance. Consequently, the data sequence is distributed according to a Student-t distribution. We considered the Yao's prior cohesions (Yao, 1984) and a beta prior distribution for the parameter p extending previous work. The Gibbs sampling scheme proposed by Loschi *al.* (1999) and Loschi *al.* (2003) was considered to obtain the posterior distributions involved in the PPM.

We applied the PPM to identify multiple change points in the volatility of four important Latin American stock markets. To represent these markets we considered the return of their most important indexes: the MERVAL (*Índice de Mercado de Valores de Buenos Aires*) of Argentina, the IBOVESPA (*Índice da Bolsa de Valores do Estado de São Paulo*) of Brazil, the IPSA (*Índice de Precios Selectivos de Acciones*) of Chile and the IPyC (*Índice de Precios y Cotizaciones*) of Mexico.

The approach presented in this paper seems to explain the behavior of the Latin American indexes - MERVAL of Argentina, IBOVESPA of Brazil, IPSA of Chile and IPyC of Mexico - satisfactorily if a change point analysis is required. We conclude that MERVAL, IPSA and IPyC have a very close behavior and could probably suffer the influences of the same non-local events. IBOVESPA seems to be the most unstable. We notice that all indexes present volatility clusters, and a small number of change points. Brazil is a developing country and has the biggest economy in South America and the Chilean economy is the more stable justifying the results we obtained.

Some open questions remain and are relevant topics for future research in this area. The posterior distribution of p could also be important for decision makers. Is it easy to implement it? How could we evaluate the probability of each instant being a change point?

Acknowledgements

The authors would you like to thank professor Wilfredo Palma (PUC-Chile) for his helpful comments and suggestions. Rosangela H. Loschi acknowledges PRPq-UFMG (grants 3893-UFMG/ RTR/ PRPq/ FUNDO/ 98 and 4801-UFMG/ RTR/ FUNDO/ PRPq/ RECEMDOUTORES/ 00), CNPq (grant 300325/2003-7) and FAPEMIG -Brasil for a partial allowance to her research. Pilar L. Iglesias's research was partially supported by FONDECYT-Chile (grant *Project Complementary Lines 800004*). Cristiano R. Moura thanks CAPES-Brazil for supporting this research.

References

Arellano-Valle, R. B. and H. Bolfarine (1995). On some characterizations of the t-distribution. *Statistics and Probability Letters* **25** 79-85.

Barry, D. and J. A. Hartigan (1992). Product partition models for change point problems. *The Annals of Statistics* **20**, 260-279.

Barry, D. and J. A.,Hartigan (1993). A Bayesian analysis for change point problem. *Journal of the American Statistical Association* **88**, 309-319.

Bernardo, J. M. and A. F. M. Smith (1994). *Bayesian Theory.* John Wiley & Sons

Chen, C. W. S. and J. C. Lee (1995). Bayesian inference of threshold autorregressive models. *Jounal of Time Series Analysis* **16**, 483-492.

Chernoff, H. and S. Zacks (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *Annals of Mathematical Statistics* **35**, 999-1018.

Correa, L. (1998). *Modelación Bayesiana de puntos de cambio en la volatilidad.* Master's thesis, Facultad de Matematicas - Pontifícia Universidad Católica de Chile, Chile. (In Spanish).

Crowley, E. M. (1997). Product partition models for normal means. *Journal of the American Statistical Association* **92**, 192-198.

Duarte Jr., A. M. and B. V. M. Mendes (1997). Robust estimation of systematic risk in emerging stock markets. *Emerging Markets Quarterly* **1**, 85-95.

Gamerman, D., (1997).*Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall.

Gelfand, A. E. and A. F. M. Smith (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398-409.

Geweke, J. and N. Terui (1993). Bayesian threshold autoregressive models for nonlinear time series. *Journal of Time Series Analysis* **14**, 441-454.

Geman, S. and D. Geman (1984). Stochastic relaxation, Gibbs distribution and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721-741.

Gilks, W. R., S. Richardson and D. J. Spiegelhalter (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall.

Hartigan, J. A. (1990). Partition models. *Communication in Statistics – Theory and Method* **19**, 2745-2756.

Hawkins, D. M. (2001). Fitting multiple change-point models to data. *Computational Statistics and Data Analysis* **37**, 323-341.

Hsu, D. A. (1984). A bayesian robust detection of shift in the risk struture of stock market returns. *Journal of the American Statistical Association* **77**, 407-416.

Loschi, R. H. and F. R. B. Cruz (2002). An analysis of the influence of some prior specifications in the identification of change points via product partition model.*Computational Statistics and Data Analysis* **39**, 477-501.

Loschi, R. H., F. R. B. Cruz, P. L. Iglesias and R. B. Arellano-Valle (2003a). A Gibbs sampling scheme to the product partition model: An application to change point problems. *Computers and Operations Research* **30**, 463-482.

Loschi, R. H., P. L. Iglesias and R. B. Arellano-Valle (2003b). Predictivistic characterization of multivariate Student-t models. *Journal of Multivariate Analysis*, forthcoming.

Loschi, R. H., P. L. Iglesias and R. B. Arellano-Valle (1999). Bayesian detection of change points in the Chilean stock market. In: *Proceedings of the Section on Bayesian Statistical Science*. Annual Meeting of American Statistical Association, Baltimore, MD, USA.

Lopes, H. F. and H. S. Migon. (2002). Comovements and Contagion in Emergin Markets: Stock Indexes Volatilities *Case Studies on Bayesian Statistics VI*, Eds. Gatsonis, C. *et al.*, Lecture Notes in Statistics Series 167.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business* **36**, 394-419.

Mendes, B. V. M (2000). Computing robust risk measures in emerging equity markets using extreme value theory. *Emerging Markets Quarterly*, Summer, 24-41.

Menzefricke, U. (1981). A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point. *Applied Statistics* **30**, 141-146.

O'Hagan, A. (1994). Bayesian Inference. In *Kendall's Advanced Theory of Statistics 2A*. John Wiley and Sons.

Quintana, F. A. and P. L. Iglesias (2003). Nonparametric Bayesian clustering and product partition models. *Journal of the Royal Statistical Society B* (forthcoming).

Smith, A. F. M. (1975). A Bayesian approach to inference about a change-point in a sequence of random variables. *Biometrika* **62**, 407-416.

Yao, Y. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches. *The Annal of Statistics* **12**, 1434-1447.

Received July 25, 2003; accepted October 20, 2003.

Rosangela H. Loschi
Departamento de Estatística
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais
31270-901 - Belo Horizonte - MG, Brazil
loschi@est.ufmg.br

Cristiano R. Moura
Departamento de Estatística
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais
31270-901 - Belo Horizonte - MG, Brazil
cmoura@est.ufmg.br

Pilar L. Iglesias
Departamento de Estadística
Pontificia Universidad Católica de Chile
Santiago - Chile
pliz@mat.puc.cl