A Note on Bayesian Identification of Change Points in Data Sequences

R. H. Loschi [a], F. R. B. Cruz [a,1], R. H. C. Takahashi [b], P. L. Iglesias [c], R. B. Arellano-Valle [c], and J. MacGregor Smith [d]

[a] *Department of Statistics, Federal University of Minas Gerais, 31270-901 - Belo Horizonte - MG, Brazil*

[b] *Department of Mathematics, Federal University of Minas Gerais, 31270-901 - Belo Horizonte - MG, Brazil*

[c] *Departamento de Estadística, Facultad de Matematicas, Pontificia Universidade Catolica de Chile, Santiago, Chile*

[d] *Department of Mechanical & Industrial Engineering, University of Massachusetts, Amherst MA 01003, USA*

Abstract

Recent research in mathematical methods for finance suggests that time series for financial data should be studied with non-stationary models and with structural changes that include both jumps and heteroskedasticity (with jumps in variance). It has been recognized that discriminating between variations caused by the continuous motion of Brownian shocks and the genuine discontinuities in the path of the process constitutes a challenge for existing computational procedures. This issue is addressed here, using the Product Partition Model (PPM), for performing such discrimination and the estimation of process jump parameters. Computational implementation aspects of PPM applied to the identification of change points in data sequences are discussed. In particular, we analyze the use of a Gibbs sampling scheme to compute the estimates and show that there is no significant impact of such use on the quality of the results. The influence of the size of the data sequence on the estimates is also of interest, as well as the efficiency of the PPM to correctly identify atypical observations occurring in close instants of time. Extensive Monte Carlo simulations attest to the effectiveness of the Gibbs sampling implementation. An illustrative financial time series example is also presented.

Key words: Atypical observations, heteroskedasticity, structural changes.

[1] Corresponding author: prof. Frederico Cruz. E-mail: fcruz@est.ufmg.br. Fax: (+55 31) 3499 5929. Phone: (+55 31) 3499 5929.

Preprint submitted to Special Issue of C&OR *21 September 2005*

1 Introduction

Classical finance analysis tools were developed in the first seven decades of the XXth Century basead on the statistical quantification of asset variability. The employment of the simple assumption of stationary Gaussian stochastic processes as models for asset prices allowed the derivation of tools for composing optimal portfolios [1], for establishing asset prices [2, 3, 4] and for calculating option prices [5]. The only parameters that were needed for these calculations were the asset mean prices, their variance and their covariance matrix.

In the decade of 60's, however, empirical studies started to reveal that such assumption of Gaussian stationarity lead to large prediction errors [6, 7]. Several studies, starting in the middle of the decade of 70's, pursued better descriptions for the asset price behavior. As an example, Hamilton [8] proposed an alternative approach to nonstationarity assuming that the first differences in the observed series follow a nonlinear stationary process rather than a linear stationary process. Such nonlinearities arise if the process is subject to discrete shifts in regime.

Two main directions that were developed since then are [9]: (i) considering the variance of the models to be a variable (the stochastic volatility models), and (ii) considering discontinuous jumps in the asset price at some time instants, in addition to the usual Brownian motion model. Several recent studies, however, indicate that these approaches, when considered individually, do not lead to satisfactory results [9, 10, 11, 12]. Models that allow for joint stochastic volatility and jumps are more likely to fit the observed data.

A non-trivial task that must be performed in order to allow the usage of models with both stochastic volatility and jumps is the estimation of the parameters representing the jump arrival intensity and the distribution of jump size [11]. In particular, it has been recognized that it is empirically difficult to discriminate between variations caused by the continuous motion of Brownian shocks and genuine discontinuities in the path of the process [11]. Eraker [9] and Eraker et al. [13] proposed a model that considers jumps in the volatility, arguing that a continuous stochastic model for volatility would be structurally unable to deal with the phenomena nature.

In this paper we consider the product partition model (PPM) for estimating the mean and volatility at each instant of time. By using this approach, it is also possible to identify if the volatility and the mean experiences changes or jumps through time and when such changes take place.

In time series and data sequence analysis, a change point can be understood as the instant when a structural change is observed on the behavior of the data or simply as when an atypical observation takes place. In financial time series, for instance, a high rate of occurrence of atypical returns is common. Mainly, this phenomenon is observed in emerging markets data since these markets are more susceptible to peaks and valleys. The identification of change points plays an important role in the analysis of these markets since the rates of occurrence of atypical returns is taken into account in the evaluation of financial risks involved [14]. We should also mention that the previous identification of change points — more specifically, the identification of structural changes — plays an important role in the analysis of stationarity in time series. In general, the usual test for a unit root in a time series cannot be used if the series presents structural changes (see the paper by Perron [15], for extensions of the test for unit roots in time series with a changing mean).

Change point identification certainly is not a new problem. Indeed, many papers have been published on the subject, either Bayesian methods, as for instance the Product Partition Model (PPM) developed by Barry & Hartigan [16] back in the 90's, and non-Bayesian approaches, e.g. the papers by Dueker [17], Hawkins [18], Horváth & Kokoszka [19], and Jandhyala et al. [20]. In this paper, we choose to consider the PPM, mainly because of all flexibility it adds to change point analysis as the number of change points is modeled as a random variable. Additionally, successful applications of the PPM to change points problems have been reported recently by Crowley [21], Quintana & Iglesias [22], Loschi & Cruz [23], Loschi et al. [24], Loschi & Cruz [25], and Loschi et al. [26].

Loschi et al. [24] propose an algorithm to compute the posterior relevances based on Gibbs sampling schemes. These authors apply the algorithm to analyze two real data sets. The main contributions of this paper include a study on the PPM to show that the use of approximate schemes such as Gibbs sampling do not seem to compromise the quality of the estimates (called product estimates). Generated data sets are considered in such a study. Additionally, extensive Monte Carlo simulation experiments are performed to verify how different sizes of the sequences influence the accuracy of the product estimates and whether there is difference in this influence when the pattern is kept along the length of the sequence. We also provide Monte Carlo simulation results that address the crucial issue of correctly identifying atypical observations occurring separated by short periods of time. The ultimate goal is to verify whether the PPM can identify structural changes and atypical observations occurring in close instants. Finally, we consider the PPM to analyze the Down Jones Industrial Average (DJIA) series in the period from October, 1995, to October, 2000. It is also a goal of this paper to identify whether the DJIA series presents atypical returns as well as heteroskedasticity.

This paper is organized as follows. Section 2 reviews the parametric approach to the PPM and its application to the normal case with unknown means and variances. In Section 3, the methods to compute the product estimates are described. Section 4 presents results of the Monte Carlo simulation study performed. Section 5 presents an illustrative example. Finally, Section 6 closes the paper with final remarks and topics for future research in the area.

2 Product Partition Models

PPM is a powerful dynamic model to analyze change point problems. Consequently, it is a useful model to identify whether a particular parameter of interest changes throughout the time, or, equivalently, presents cluster structure.

Denote by ρ a random partition of the set $I \cup \{0\}$, in which $I = \{1, \ldots, n\}$. In the parametric approach to the PPM presented by Barry & Hartigan [16], a sequence of unknown parameters $\theta_1, \ldots, \theta_n$ is considered, 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. Given a particular partition $\rho = \{i_0, \ldots, i_b\}$, with $b \in I$, such that $0 = i_0 < i_1 < \cdots < i_b = n$, we have that $\theta_i = \theta_{[i_{r-1}i_r]}$, for every $i_{r-1} < i \leq i_r$, for $r = 1, \ldots, b$, and that $\theta_{[i_0i_1]}, \ldots, \theta_{[i_{b-1}i_b]}$ are independent, with $\theta_{[ij]}$ having (block) prior density $\pi_{[ij]}(\theta)$, $\theta \in \Theta_{[ij]}$, where $\Theta_{[ij]}$ is the parameter space corresponding to the common parameter, say, $\theta_{[ij]} = \theta_{i+1} = \cdots = \theta_j$, which indexes the joint conditional density of $\mathbf{X}_{[ij]} = (X_{i+1}, \ldots, X_j)'$. The degree of similarity among the observations in $\mathbf{X}_{[ij]}$ is called prior cohesions and is denoted by $c_{[ij]}$. In the parametric case, two observations X_i and X_j, for $i \neq j$, are in the same block, if they are identically distributed. Thus, $(X_1, \ldots, X_n;\ \rho)$ follows the PPM if:

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

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

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) conditional in $\rho = \{i_0, \ldots, i_b\}$, the sequence $X_1, \ldots, X_n$ has the joint density given by

$$f(X_1, \ldots, X_n|\rho = \{i_0, \ldots, i_b\}) = \\ \Pi_{j=1}^{b} \int_{\Theta_{[i_{j-1}i_j]}} f_{[i_{j-1}i_j]}(\mathbf{X}_{[i_{j-1}i_j]}|\theta)\pi_{[i_{j-1}i_j]}(\theta)d\theta. \quad (1)$$

In accordance with the PPM, the posterior expectation (called the product estimate) of θ_k is given by:

$$E(\theta_k|X_1,\ldots,X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]} E(\theta_k|\mathbf{X}_{[ij]}),\ \forall\ k = 1,\ldots,n, \tag{2}$$

in which $r^*_{[ij]}$ denotes the posterior relevance for block $[ij] = \{i+1,\ldots,j\}$, that is,

$$r^*_{[ij]} = P([ij] \in \rho|X_1,\ldots,X_n) = \frac{\lambda_{[0i]} c^*_{[ij]} \lambda_{[jn]}}{\lambda_{[0n]}}. \tag{3}$$

algorithm
(1) **read** $X_1,\ldots,X_n$
    for all $i,j \in \{0,\ldots,n\}$ **such that** $i < j$ **do**
(2) $f_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta$
(3) $c^*_{[ij]} \leftarrow c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]})$
    end for
(4) *compute* $\begin{cases} \lambda_{[0j]}, \forall j = 0,\ldots,n; \\ \lambda_{[in]}, \forall i = 1,\ldots,n; \end{cases}$
    for all $i,j \in \{0,\ldots,n\}$ **such that** $i < j$ **do**
(5) $r^*_{[ij]} \leftarrow \frac{\lambda_{[0i]} c^*_{[ij]} \lambda_{[jn]}}{\lambda_{[0n]}}$
    end for
    for $k = 1$ **to** n **do**
(6) $E(\theta_k|X_1,\ldots,X_n) \leftarrow \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]} E(\theta_k|\mathbf{X}_{[ij]})$
    end for
(7) **write** $E(\theta_1),\ldots,E(\theta_n)$
end algorithm

Fig. 1. The PPM algorithm.

The posterior cohesion for block $[ij]$ is given by $c^*_{[ij]} = c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]})$ and $\lambda_{[ij]} = \sum \Pi_{k=1}^{b} c^*_{[i_{k-1} i_k]}$, in which the summation is over all partitions of $\{i+1,\ldots,j\}$ into b blocks with endpoints $i_0, i_1, \ldots, i_b$, satisfying the condition $i = i_0 < i_1 < \cdots < i_b = j$. The algorithm for the PPM presented in pseudo-language is show in Figure 1.

Remark 1 *Notice that the PPM assumes only conditional independence for $X_1,\ldots,X_n$, given the sequence of parameters $\theta_1,\ldots,\theta_n$. Consequently, the data $X_1,\ldots,X_n$ can be correlated depending on the prior distribution we assume to describe the uncertainty about $\theta_1,\ldots,\theta_n$. See, for instance, the normal case pointed out in the following.*

Remark 2 *Since the PPM is a dynamic model, which assumes that the parameters can change through the time, it is suitable to describe the behavior of data with non-constant parameters (such as, for instance, the variance).*

For the normal case, it is assumed that $\theta_1 = (\mu_1, \sigma_1^2), \ldots, \theta_n = (\mu_n, \sigma_n^2)$ are independent. It is also assumed that, given $\theta_1, \ldots, \theta_n$, the random variables $X_1, \ldots, X_n$ are independent and they are such that $X_k|\mu_k, \sigma_k^2 \sim \mathcal{N}(\mu_k, \sigma_k^2)$, for $k = 1, \ldots, n$. It is also assumed that the common parameter $\theta_{[ij]} = (\mu_{[ij]}, \sigma_{[ij]}^2)$ related to block $[ij]$ has the following conjugate normal-inverted-gamma prior distribution: $\mu_{[ij]}|\sigma_{[ij]}^2 \sim \mathcal{N}(m_{[ij]}, v_{[ij]}\sigma_{[ij]}^2)$ and $\sigma_{[ij]}^2 \sim \mathcal{IG}(a_{[ij]}/2, d_{[ij]}/2)$, in which $\mathcal{IG}(a, d)$ denotes the inverted-gamma distribution with parameters a and d, $m_{[ij]} \in \mathbb{R}$, and $a_{[ij]}$, $d_{[ij]}$ and $v_{[ij]}$ are positive values.

As a consequence of such assumptions, it follows that the random vector $\mathbf{X}_{[ij]}$ follows a $(j-i)$-dimensional Student-t distribution given by:

$$f(\mathbf{X}_{[ij]}) = \frac{K(d_{[ij]}, j-i)a_{[ij]}^{d_{[ij]}/2}}{\Big(1 + (j-i)v_{[ij]}\Big)^{1/2}\Big(a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})\Big)^{(d_{[ij]}+j-i)/2}}, \tag{4}$$

in which $K(d, k) = \Gamma((d+k)/2)/\Gamma(d/2)/\pi^{k/2}$. Thus, it is assumed that the observations within the same block are correlated and they have their behavior described by a distribution with a heavier tail than the normal distribution.

It is also possible to show that the product estimates for μ_k and σ_k^2, for $k = 1, \ldots, n$, are given, respectively, by

$$\hat{\mu}_k = E(\mu_k|X_1, \ldots, X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r_{[ij]}^* m_{[ij]}^*, \quad (\text{if } d_{[ij]}^* > 1) \tag{5}$$

and

$$\hat{\sigma}_k^2 = E(\sigma_k^2|X_1, \ldots, X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r_{[ij]}^* \frac{a_{[ij]}^*}{d_{[ij]}^* - 2}, \quad (\text{if } d_{[ij]}^* > 2), \tag{6}$$

in which

$$\begin{cases} m_{[ij]}^* &= \frac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \frac{m_{[ij]}}{(j-i)v_{[ij]}+1}, \\ v_{[ij]}^* &= \frac{v_{[ij]}}{(j-i)v_{[ij]}+1}, \\ d_{[ij]}^* &= d_{[ij]} + j - i, \\ a_{[ij]}^* &= a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]}), \end{cases}$$

with $q_{[ij]}(\mathbf{X}_{[ij]}) = \sum_{r=i+1}^{j}(X_r - \bar{X}_{[ij]})^2 + \frac{(j-i)(\bar{X}_{[ij]} - m_{[ij]})^2}{(j-i)v_{[ij]}+1}$, and $\bar{X}_{[ij]} = \frac{1}{j-i}\sum_{r=i+1}^{j} X_r$.

With regard to the computation of the relevances defined in (3), $r^*_{[ij]}$, we will use the prior cohesions for block $[ij]$ proposed by Yao [27]

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

for all $i,\ j \in I,\ i < j$, in which $0 < p < 1$ is the probability that a change occurs at any instant in the sequence. Consequently, for the normal case, the posterior cohesion of the block $[ij]$ becomes:

$$c^*_{[ij]} = \begin{cases} \frac{p(1-p)^{j-i-1}K(d_{[ij]},j-i)a_{[ij]}^{d_{[ij]}/2}}{(1+(j-i)v_{[ij]})^{1/2}\{a_{[ij]}+q_{[ij]}(\mathbf{X}_{[ij]})\}^{(d_{[ij]}+j-i)/2}}, & \text{if } j < n, \\ \\ \frac{(1-p)^{j-i-1}K(d_{[ij]},j-i)a_{[ij]}^{d_{[ij]}/2}}{(1+(j-i)v_{[ij]})^{1/2}\{a_{[ij]}+q_{[ij]}(\mathbf{X}_{[ij]})\}^{(d_{[ij]}+j-i)/2}}, & \text{if } j = n, \end{cases} \tag{7}$$

with $K(d,k) = \Gamma((d+k)/2)/\Gamma(d/2)/\pi^{k/2}$, which completes the algorithm with the computation of $r^*_{[ij]}$, from (3). For all the details, the reader is referred to the paper by Loschi & Cruz [23].

Remark 3 *It is straightforward to extend the PPM to identify change points in a sequence of data whose behavior is described by a distribution in the exponential family, for example. This extension can be found in Loschi et al. [26]. Extensions to linear models using the general PPM can be found in Quintana & Iglesias [22].*

Remark 4 *Notice that it is not necessary to consider a conjugate prior distribution for the parameters $\Theta_{[ij]}$. Assuming the Jeffreys prior, for instance, the product estimates for μ_k and σ^2_k are given, respectively, by:*

$$\hat{\mu}_k = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]}\bar{X}_{[ij]} \ \text{ and } \ \hat{\sigma^2}_k = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]}\frac{\sum_{r=i+1}^{j}(X_r - \bar{X}_{[ij]})^2}{j-i}.$$

3 Computational Methods for Product Estimates

We will use two methods to obtain the product estimates given in (5) and (6). Yao's method is based on a recursive algorithm. The second method is based on a Gibbs sampling scheme.

3.1 Yao's Method

Yao's method [27] consists of simply considering the following recursive algorithm to calculate the sums $\lambda_{[ij]}$ defined in (3):

$$\begin{cases} \lambda_{[00]} = 1, \\ \lambda_{[01]} = c^*_{[01]}, \\ \lambda_{[0j]} = c^*_{[0j]} + \sum_{t=1}^{j-1} \lambda_{[0t]} c^*_{[tj]}, \forall j = 2, \ldots, n, \\ \lambda_{[in]} = c^*_{[in]} + \sum_{t=i+1}^{n-1} \lambda_{[tn]} c^*_{[it]}, \forall i = 1, \ldots, n-2, \\ \lambda_{[(n-1)n]} = c^*_{[(n-1)n]}, \\ \lambda_{[nn]} = 1, \end{cases}$$

in such a way that the PPM algorithm of Figure 1 can be applied. Notice that the algorithm is polynomial, $\mathcal{O}(n^2)$.

3.2 Gibbs Sampling Based Method

Since Gelfand & Smith [28] showed how easily a large number of difficult problems could be approximately solved by Markov Chain Monte Carlo (MCMC) methods, many hard problems in Bayesian inference became tractable. Gibbs sampling is one of these MCMC methods and it has been used as a posterior distribution generation scheme for decades. In this paper, Gibbs sampling is used to generate samples of vector $\mathbf{U} = (U_1, \ldots, U_{n-1})$, in which the random variable U_i reflects whether or not a change point occurs at the time i, that is

$$U_i = \begin{cases} 1, & \text{if } \theta_i = \theta_{i+1}, \\ 0, & \text{if } \theta_i \neq \theta_{i+1}, \end{cases}$$

in which $i = 1, \ldots, n-1$. As noticed by Loschi & Cruz [23], any partition $\rho = \{i_0, \ldots, i_b\}$ may be associated to some value of $\mathbf{U} = (U_1, \ldots, U_{n-1})$.

The method obtains the product estimates of μ_k and σ_k^2 as follows. First we generate T samples of vector $\mathbf{U} = (U_1, \ldots, U_{n-1})$. The estimate of the posterior relevance of block $[ij]$, for $i, j = 1, \ldots, n, i < j$, is computed as follows

$$\hat{r}^*_{[ij]} = \frac{M}{T}, \tag{8}$$

in which M is the number of vectors in which it is observed that $U_i = 0, U_{i+1} = \ldots = U_{j-1} = 1$ and $U_j = 0$. Thus, the product estimates of μ_k and σ_k^2 may be obtained by substituting the estimate (8) into (5) and (6), respectively. Notice that likewise Yao's algorithm, the procedure is also polynomial $\mathcal{O}(T)$.

4 Simulation Study

All the algorithms described earlier were coded in C++, a powerful and efficient programming environment, and are available from the authors upon request or directly from the web (`ftp://ftp.est.ufmg.br/pub/loschi/pub/mtecar`). We conducted Monte Carlo simulation studies in order to analyze the influence of the size of the sequence in the accuracy of the product estimates. We also wanted to evaluate the effectiveness of the two methods presented earlier in Section 3 in presence of atypical observations occurring separated by short periods of time. The error obtained with different values of p was also evaluated. Data were generated independently from normal distributions, with common and low variance and different means. For all scenarios, 20 replications were considered and for our convenience the variance was set to 0.001.

Additionally, in all scenarios, the following prior distributions for $\mu_{[ij]}$ and $\sigma^2_{[ij]}$ were assumed

$$\mu_{[ij]}|\sigma^2_{[ij]} \sim \mathcal{N}(0.0, \sigma^2_{[ij]}) \;\text{ and }\; \sigma^2_{[ij]} \sim \mathcal{IG}\,(0.01/2, 4/2)\,.$$

Notice that the inverted gamma prior distribution above has modal value at 1.67×10^{-3}, mean at 5.0×10^{-3}, and infinite variance.

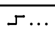
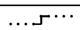
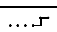
For the method based on Gibbs sampling, 1000 samples of $\mathbf{U}$'s were generated, starting from a vector of zeros. A burn-in period of 100 and a lag of one was used, since fast convergence and low autocorrelation were observed. To compare the methods, the sum of square errors per block were calculated, defined as the sum of square errors divided by the product between the sample size and the true number of blocks in the partition.

Tables 1 and 2 show the sizes of all groups tested, n, the probability p considered, and the scenario. Concerning the notation used for the scenarios, the representation $32^1\ 32^2\ 64^4$, for example, stands for a partition into three blocks of size 32, 32, and 64, generated from normal distributions with means 1, 2, and 4, respectively.

Table 1
Groups for sample size and p cases.

	group	n	p	scenario
―― …	1	32	0.01	32^0
	2		0.5	32^0
	3		0.9	32^0
	4	64	0.01	64^0
	5		0.5	64^0
	6		0.9	64^0
	7	128	0.01	128^0
	8		0.5	128^0
	9		0.9	128^0
_⊥_ …	10	32	0.01	$4^0\ 1^1\ 27^0$
	11		0.5	$4^0\ 1^1\ 27^0$
	12		0.9	$4^0\ 1^1\ 27^0$
	13	64	0.01	$4^0\ 1^1\ 59^0$
	14		0.5	$4^0\ 1^1\ 59^0$
	15		0.9	$4^0\ 1^1\ 59^0$
	16	128	0.01	$4^0\ 1^1\ 123^0$
	17		0.5	$4^0\ 1^1\ 123^0$
	18		0.9	$4^0\ 1^1\ 123^0$
… _⊥_ …	19	32	0.01	$15^0\ 1^1\ 16^0$
	20		0.5	$15^0\ 1^1\ 16^0$
	21		0.9	$15^0\ 1^1\ 16^0$
	22	64	0.01	$31^0\ 1^1\ 32^0$
	23		0.5	$31^0\ 1^1\ 32^0$
	24		0.9	$31^0\ 1^1\ 32^0$
	25	128	0.01	$63^0\ 1^1\ 64^0$
	26		0.5	$63^0\ 1^1\ 64^0$
	27		0.9	$63^0\ 1^1\ 64^0$
… _⊥_	28	32	0.01	$27^0\ 1^1\ 4^0$
	29		0.5	$27^0\ 1^1\ 4^0$
	30		0.9	$27^0\ 1^1\ 4^0$
	31	64	0.01	$59^0\ 1^1\ 4^0$
	32		0.5	$59^0\ 1^1\ 4^0$
	33		0.9	$59^0\ 1^1\ 4^0$
	34	128	0.01	$59^0\ 1^1\ 4^0$
	35		0.5	$59^0\ 1^1\ 4^0$
	36		0.9	$59^0\ 1^1\ 4^0$
_⊥⊥⊥⊥_ …	37	32	0.01	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 5^0$
	38		0.5	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 5^0$
	39		0.9	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 5^0$
	40	64	0.01	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 37^0$
	41		0.5	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 37^0$
	42		0.9	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 37^0$
	43	128	0.01	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 101^0$
	44		0.5	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 101^0$
	45		0.9	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 101^0$
_⊥⊥⊥⊥_ …	46	64	0.01	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 9^0$
	47		0.5	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 9^0$
	48		0.9	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 9^0$
	49	128	0.01	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4; 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 10^0$
	50		0.5	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4; 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 10^0$
	51		0.9	$5^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4; 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0$ $1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 6^0\ 1^4\ 10^0$

Table 1
(continued)

	group	n	p	scenario
	52	32	0.01	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	53		0.5	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	54		0.9	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	55	64	0.01	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 32^0$
	56		0.5	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 32^0$
	57		0.9	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 32^0$
	58	128	0.01	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 64^0$
	59		0.5	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 64^0$
	60		0.9	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 64^0$
	61	64	0.01	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0$ $4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	62		0.5	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0$ $4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	63		0.9	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0$ $4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	64	128	0.01	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0$ $4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$ $4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	65		0.5	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0$ $4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$ $4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	66		0.9	$4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0$ $4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$ $4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3\ 4^0\ 4^3$
	67	32	0.01	$4^0\ 28^1$
	68		0.5	$4^0\ 28^1$
	69		0.9	$4^0\ 28^1$
	70	64	0.01	$4^0\ 60^1$
	71		0.5	$4^0\ 60^1$
	72		0.9	$4^0\ 60^1$
	73	128	0.01	$4^0\ 124^1$
	74		0.5	$4^0\ 124^1$
	75		0.9	$4^0\ 124^1$
	76	32	0.01	$16^0\ 16^1$
	77		0.5	$16^0\ 16^1$
	78		0.9	$16^0\ 16^1$
	79	64	0.01	$32^0\ 32^1$
	80		0.5	$32^0\ 32^1$
	81		0.9	$32^0\ 32^1$
	82	128	0.01	$64^0\ 64^1$
	83		0.5	$64^0\ 64^1$
	84		0.9	$64^0\ 64^1$
	85	32	0.01	$28^0\ 4^1$
	86		0.5	$28^0\ 4^1$
	87		0.9	$28^0\ 4^1$
	88	64	0.01	$60^0\ 4^1$
	89		0.5	$60^0\ 4^1$
	90		0.9	$60^0\ 4^1$
	91	128	0.01	$124^0\ 4^1$
	92		0.5	$124^0\ 4^1$
	93		0.9	$124^0\ 4^1$

4.1 Sample Size and p Case

In order to analyze the influence of different sample sizes on the product estimates and to evaluate the error for different prior specifications for the parameter p, several scenarios were examined. We assumed $n = 32$, 64, and 128, and $p = 0.01$, 0.5 and 0.9. Figure 2 shows graphs for the errors for the

Table 1
(continued)

	group	n	p	scenario
	94	32	0.01	$6^0\ 6^1\ 6^2\ 6^3\ 8^4$
	95		0.5	$6^0\ 6^1\ 6^2\ 6^3\ 8^4$
	96		0.9	$6^0\ 6^1\ 6^2\ 6^3\ 8^4$
	97	64	0.01	$6^0\ 6^1\ 6^2\ 6^3\ 40^4$
	98		0.5	$6^0\ 6^1\ 6^2\ 6^3\ 40^4$
	99		0.9	$6^0\ 6^1\ 6^2\ 6^3\ 40^4$
	100	128	0.01	$6^0\ 6^1\ 6^2\ 6^3\ 104^4$
	101		0.5	$6^0\ 6^1\ 6^2\ 6^3\ 104^4$
	102		0.9	$6^0\ 6^1\ 6^2\ 6^3\ 104^4$
	103	64	0.01	$6^0\ 6^1\ 6^2\ 6^3\ 6^4\ 6^5\ 6^6\ 6^7\ 6^8\ 6^9\ 4^{10}$
	104		0.5	$6^0\ 6^1\ 6^2\ 6^3\ 6^4\ 6^5\ 6^6\ 6^7\ 6^8\ 6^9\ 4^{10}$
	105		0.9	$6^0\ 6^1\ 6^2\ 6^3\ 6^4\ 6^5\ 6^6\ 6^7\ 6^8\ 6^9\ 4^{10}$
	106	128	0.01	$6^0\ 6^1\ 6^2\ 6^3\ 6^4\ 6^5\ 6^6\ 6^7\ 6^8\ 6^9\ 6^{10}\ 6^{11}\ 6^{12}\ 6^{13}\ 6^{14}\ 6^{15}\ 6^{16}\ 6^{17}\ 6^{18}\ 6^{19}\ 6^{20}\ 2^{21}$
	107		0.5	$6^0\ 6^1\ 6^2\ 6^3\ 6^4\ 6^5\ 6^6\ 6^7\ 6^8\ 6^9\ 6^{10}\ 6^{11}\ 6^{12}\ 6^{13}\ 6^{14}\ 6^{15}\ 6^{16}\ 6^{17}\ 6^{18}\ 6^{19}\ 6^{20}\ 2^{21}$
	108		0.9	$6^0\ 6^1\ 6^2\ 6^3\ 6^4\ 6^5\ 6^6\ 6^7\ 6^8\ 6^9\ 6^{10}\ 6^{11}\ 6^{12}\ 6^{13}\ 6^{14}\ 6^{15}\ 6^{16}\ 6^{17}\ 6^{18}\ 6^{19}\ 6^{20}\ 2^{21}$
	109	32	0.01	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7$
	110		0.5	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7$
	111		0.9	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7$
	112	64	0.01	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 32^{10}$
	113		0.5	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 32^{10}$
	114		0.9	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 32^{10}$
	115	128	0.01	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 96^{10}$
	116		0.5	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 96^{10}$
	117		0.9	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 96^{10}$
	118	64	0.01	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 12^8\ 6^9\ 4^{11}\ 1^{12}\ 2^{14}\ 7^{16}$
	119		0.5	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 12^8\ 6^9\ 4^{11}\ 1^{12}\ 2^{14}\ 7^{16}$
	120		0.9	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 12^8\ 6^9\ 4^{11}\ 1^{12}\ 2^{14}\ 7^{16}$
	121	128	0.01	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 12^8\ 13^9\ 4^{11}\ 1^{12}\ 2^{14}\ 7^{15}\ 10^{16}\ 25^{17}\ 1^{19}\ 13^{20}\ 8^{23}$
	122		0.5	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 12^8\ 13^9\ 4^{11}\ 1^{12}\ 2^{14}\ 7^{15}\ 10^{16}\ 25^{17}\ 1^{19}\ 13^{20}\ 8^{23}$
	123		0.9	$12^0\ 6^2\ 4^3\ 1^{4.5}\ 2^5\ 7^7\ 12^8\ 13^9\ 4^{11}\ 1^{12}\ 2^{14}\ 7^{15}\ 10^{16}\ 25^{17}\ 1^{19}\ 13^{20}\ 8^{23}$

product estimates of the variances as a function of the errors for the estimates of the means. Arcs are connecting the results for the same scenarios. We notice that both methods performed similarly in sequences in which changes did not occur or occurred only once (scenarios 1 to 36), such a rare event in practice. Additionally, in sequences in which we had several atypical observations or structural changes (scenarios 37 to 93), the Gibbs sampling based method had a superior performance, with smaller square errors for both the estimates of variances and means. Finally, in staircase style sequences (scenarios 94 to 123), Yao's method presented lower errors for most of the scenarios.

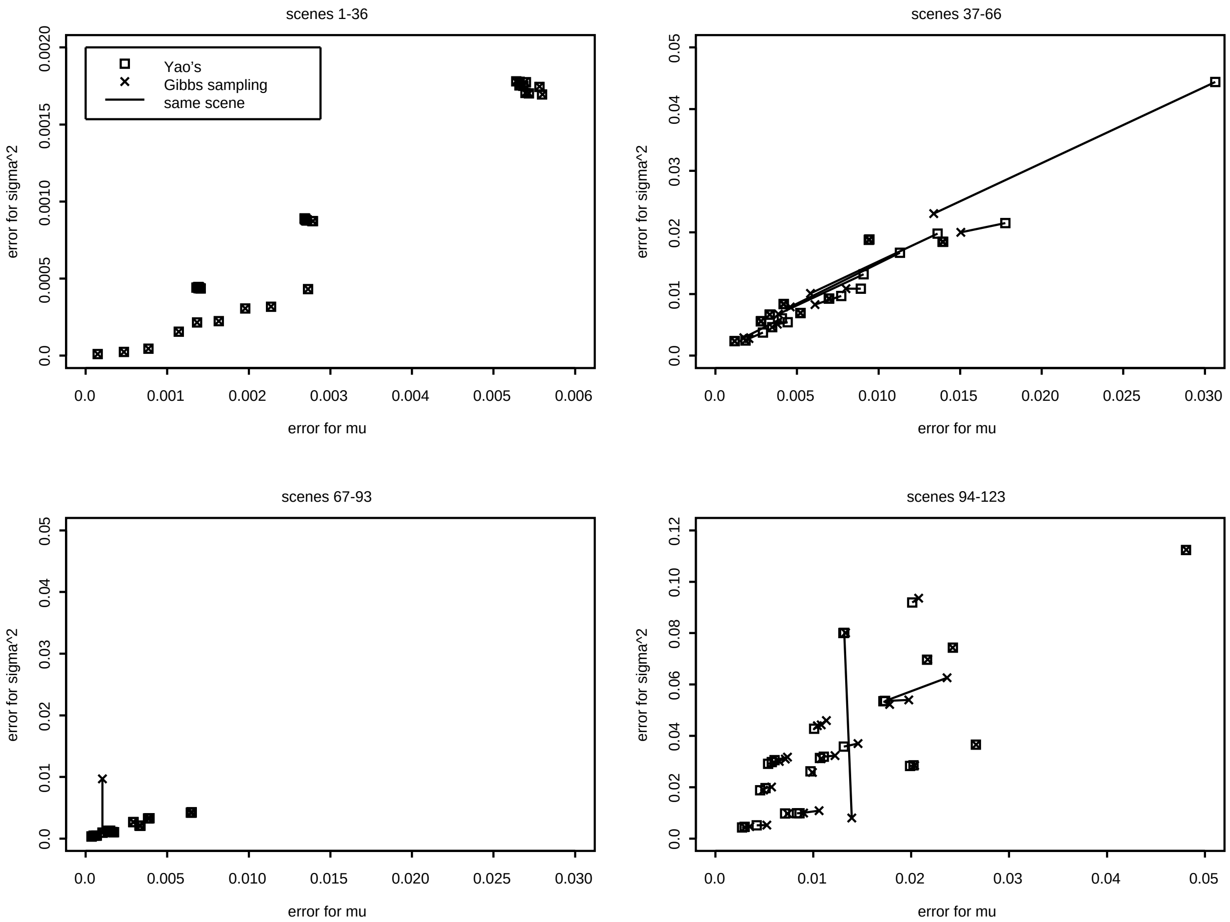


Fig. 2. Errors for the mean *versus* variance – sample size and p cases.

Table 2
Groups for atypical observation cases.

group	p	scenario
1	0.01	$20^0\ 1^1\ 1^0\ 1^1\ 27^0$
2		$20^0\ 1^2\ 1^0\ 1^2\ 27^0$
3		$20^0\ 1^4\ 1^0\ 1^4\ 27^0$
4		$20^0\ 1^1\ 2^0\ 1^1\ 26^0$
5		$20^0\ 1^2\ 2^0\ 1^2\ 26^0$
6		$20^0\ 1^4\ 2^0\ 1^4\ 26^0$
7		$20^0\ 1^1\ 4^0\ 1^1\ 24^0$
8		$20^0\ 1^2\ 4^0\ 1^2\ 24^0$
9		$20^0\ 1^4\ 4^0\ 1^4\ 24^0$
10		$20^0\ 1^1\ 8^0\ 1^1\ 20^0$
11		$20^0\ 1^2\ 8^0\ 1^2\ 20^0$
12		$20^0\ 1^4\ 8^0\ 1^4\ 20^0$
13		$20^0\ 1^1\ 14^0\ 1^1\ 14^0$
14		$20^0\ 1^2\ 14^0\ 1^2\ 14^0$
15		$20^0\ 1^4\ 14^0\ 1^4\ 14^0$
16		$20^0\ 1^4\ 1^0\ 1^2\ 27^0$
17		$20^0\ 1^4\ 2^0\ 1^2\ 26^0$
18		$20^0\ 1^4\ 4^0\ 1^2\ 24^0$
19		$20^0\ 1^4\ 8^0\ 1^2\ 20^0$
20		$20^0\ 1^4\ 14^0\ 1^2\ 14^0$
21		$20^0\ 1^3\ 1^0\ 1^2\ 27^0$
22		$20^0\ 1^3\ 2^0\ 1^2\ 26^0$
23		$20^0\ 1^3\ 4^0\ 1^2\ 24^0$
24		$20^0\ 1^3\ 8^0\ 1^2\ 20^0$
25		$20^0\ 1^3\ 14^0\ 1^2\ 14^0$
26		$20^0\ 1^1\ 1^{-1}\ 28^0$
27		$20^0\ 1^1\ 2^0\ 1^{-1}\ 26^0$
28		$20^0\ 1^1\ 6^0\ 1^{-1}\ 22^0$
29		$20^0\ 1^1\ 12^0\ 1^{-1}\ 16^0$
30		$20^0\ 1^4\ 1^{-4}\ 28^0$
31		$20^0\ 1^4\ 2^0\ 1^{-4}\ 26^0$
32		$20^0\ 1^4\ 6^0\ 1^{-4}\ 22^0$
33		$20^0\ 1^4\ 12^0\ 1^{-4}\ 16^0$
34		$20^0\ 1^4\ 1^{-1}\ 28^0$
35		$20^0\ 1^4\ 2^0\ 1^{-1}\ 26^0$
36		$20^0\ 1^4\ 6^0\ 1^{-1}\ 22^0$
37		$20^0\ 1^4\ 12^0\ 1^{-1}\ 16^0$
38		$20^0\ 1^{-1}\ 1^4\ 28^0$
39		$20^0\ 1^{-1}\ 2^0\ 1^4\ 26^0$
40		$20^0\ 1^{-1}\ 6^0\ 1^4\ 22^0$
41		$20^0\ 1^{-1}\ 12^0\ 1^4\ 16^0$
42		$20^0\ 1^2\ 1^4\ 28^0$
43		$20^0\ 1^1\ 1^6\ 28^0$
44		$20^0\ 1^1\ 1^4\ 28^0$
45	0.1	$10^0\ 1^2\ 4^0\ 1^2\ 4^0\ 1^2\ 4^0\ 1^2\ 24^0$
46		$10^0\ 1^4\ 4^0\ 1^4\ 4^0\ 1^4\ 4^0\ 1^4\ 24^0$
47		$10^0\ 1^2\ 8^0\ 1^2\ 8^0\ 1^2\ 8^0\ 1^2\ 12^0$
48		$10^0\ 1^4\ 8^0\ 1^4\ 8^0\ 1^4\ 8^0\ 1^4\ 12^0$
49		$10^0\ 1^4\ 4^0\ 1^1\ 4^0\ 1^4\ 4^0\ 1^2\ 24^0$
50		$10^0\ 1^4\ 8^0\ 1^1\ 8^0\ 1^4\ 8^0\ 1^2\ 12^0$
51		$10^0\ 1^4\ 4^0\ 1^{-4}\ 4^0\ 1^4\ 4^0\ 1^{-4}\ 24^0$
52		$10^0\ 1^4\ 8^0\ 1^{-4}\ 8^0\ 1^4\ 8^0\ 1^{-4}\ 12^0$
53		$10^0\ 1^4\ 4^0\ 1^{-1}\ 4^0\ 1^1\ 4^0\ 1^{-4}\ 24^0$
54		$10^0\ 1^4\ 8^0\ 1^{-1}\ 8^0\ 1^1\ 8^0\ 1^{-4}\ 12^0$
55		$20^0\ 1^4\ 4^0\ 1^3\ 4^0\ 1^1\ 4^0\ 1^5\ 14^0$
56		$20^0\ 1^4\ 8^0\ 1^3\ 8^0\ 1^1\ 8^0\ 1^5\ 2^0$
57		$20^0\ 1^4\ 4^0\ 1^3\ 8^0\ 1^1\ 6^0\ 1^5\ 2^0\ 1^2\ 5^0$
58		$20^0\ 1^4\ 4^0\ 1^4\ 8^0\ 1^4\ 6^0\ 1^4\ 2^0\ 1^4\ 5^0$

4.2 Atypical Observation Case

In this section, the goal is to provide a Monte Carlo simulation study to analyze the performance of the PPM when atypical observations are observed

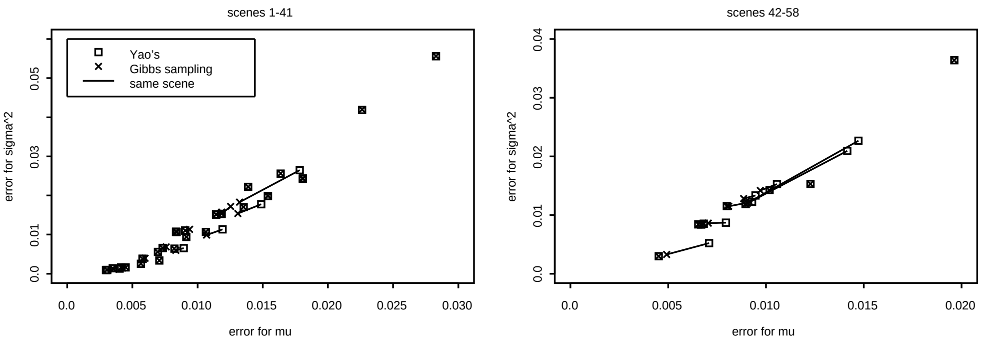


Fig. 3. Errors for the mean *versus* variance – atypical observation cases.

in close instants. Sequences of size $n = 50$ and different prior values for p were considered (see Table 2). The results are shown in Figure 3. In the presence of atypical observations, both methods performed similarly and the errors for the estimates for the mean were significantly lower than for the variance. We notice that the distance between the atypical observations — as well as the difference between the means of the normal distributions that generated the atypical data and the remaining observations — had influence in the error. In general, for all scenarios, smaller errors were observed if the number of observations among the atypical data increased. The errors were also smaller for scenarios in which the typical and atypical observations were from normal distributions with similar mean values. Figure 3 shows graphs for the errors for the product estimates of the variances as a function of the errors for the estimates of the means.

As a final remark, when atypical observations occur in close instants, an error that may happen is to find nonexistent structural change in the data sequence. For instance, Figure 4 shows the product estimates corresponding to scenarios 3, 6, 9, and 12 (seen in Table 2), and two other similar scenarios (not shown), by the Gibbs sampling method. By using the PPM, one would identify atypical observations only when a considerable number of typical observations occurred between them. In other words, one might assume a real structural change in the sequence, even if all that had happened was two atypical observations too close in time. It is also noticeable that all methods identify changes in the variance that do not occur.

5 Illustrative Example

A simple and yet interesting example will be shown next. Let us consider the Dow Jones Industrial Average (DJIA) return series recorded fortnightly from October, 1995, to October, 2000. The ultimate goal of this section is to identify

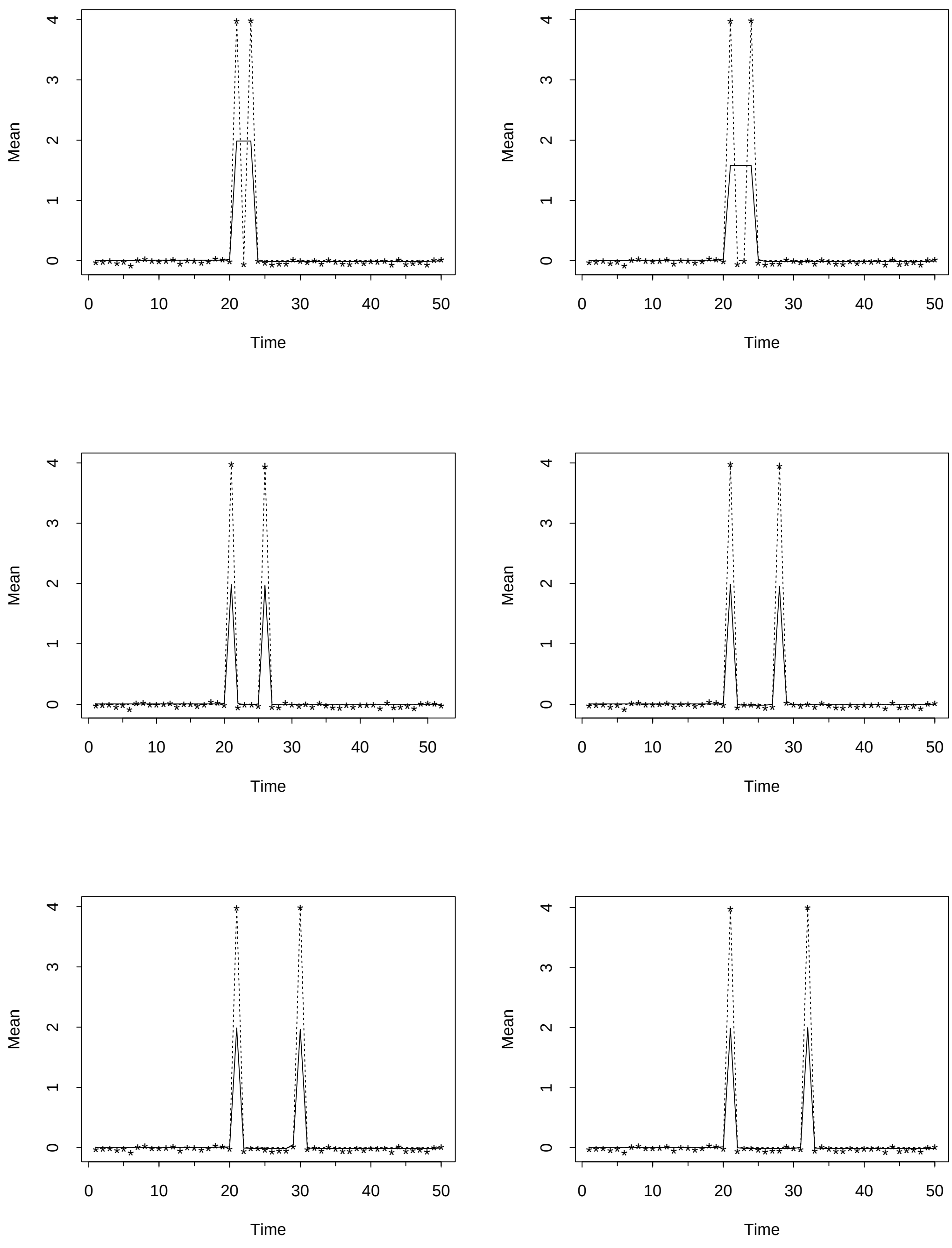


Fig. 4. Product estimates for the means by the Gibbs sampling method (* → data; traced line → real mean; full line → product estimates).

whether the DJIA series presents volatility (measured as variance) clusters as well as expected return clusters. That is, the main interest is to verify if the DJIA series presents heteroskedasticity and atypical returns. The advantage of using such a methodology (if compared with McLeod-Li test, for example) is that, besides identifying if the Down Jones series presents heteroskedasticity, we can also identify whether the change is structural or else it is an atypical value, as well as we can acertain when the changes occurred and in what

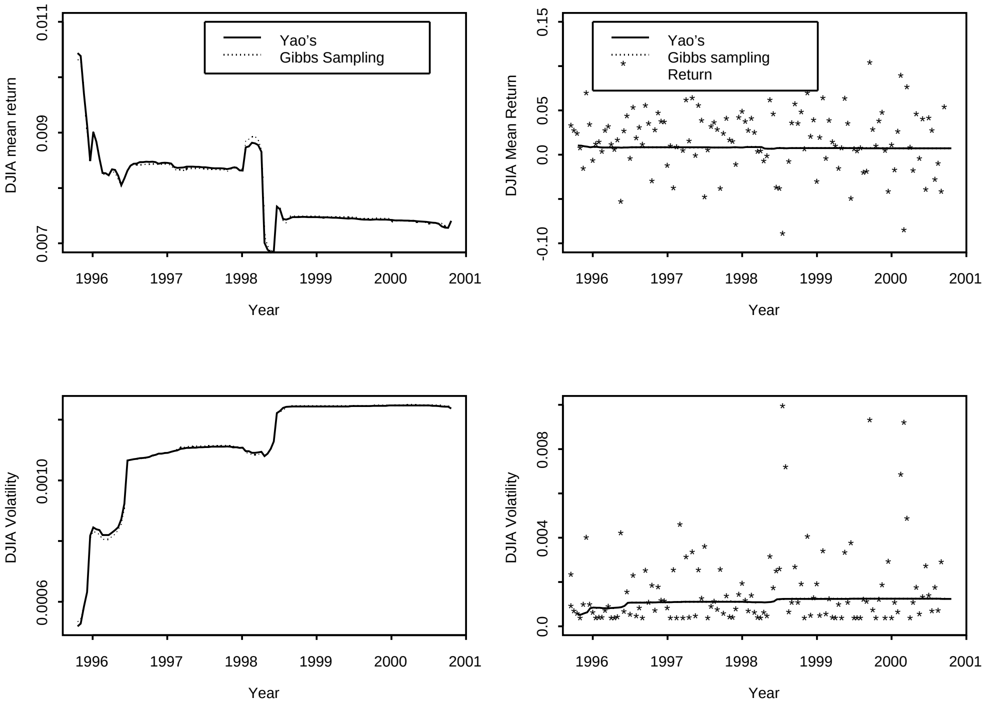


Fig. 5. Product estimates for the mean return and Volatility - DJIA

directions they were.

For modeling the returns, conditionally in the mean return (μ) and in the volatility (σ^2), assume that the returns are normally distributed. Since the USA market is efficient, few changes are expected in its behavior and the mean returns can be considered close to zero [18]. Then, let us assume that the prior distribution for p is the degenerate distribution $p = 0.1$ and that the prior distribution for the common parameter $\Theta_{[ij]} = (\mu_{[ij]}, \sigma^2_{[ij]})$ is:

$$\mu_{[ij]}|\sigma^2_{[ij]} \sim \mathcal{N}(0, \sigma^2_{[ij]}) \text{ and } \sigma^2_{[ij]} \sim \mathcal{IG}(0.001/2, 8/2).$$

Notice that under the square loss function the prior Bayes' estimates for the mean return and for the volatility are 0 and 1.67×10^{-4}, respectively. For the Gibbs sampling scheme, 10,000 samples were generated. The initial 1,000 samples were discarded and a lag of 10 was considered.

Notice from Figure 5 that the product estimates for the mean returns are close to zero and reach their minimum at May 2nd fortnight, 1998. It is noticeable that the mean return is approximately constant from June 1st fortnight, 1996, to January 1st fortnight, 1998, and from July 2nd fortnight, 1998, to October 2nd fortnight, 2000. The volatility for the DJIA return series is also small and experiences few changes. We perceive some important increases in June 2nd fortnight, 1996, and in June 2nd fortnight, 1999. We also observe that changes in the volatility and in the mean return occur in similar instants and in opposite directions, that is, the volatility increases when the mean return decreases. In a descriptive analysis (not shown), we have found three occasions with atypical returns, which are August 2nd fortnight, 1998 (-0.0098), October 2nd fortnight, 1999 (0.0095), and April 1st fortnight, 2000 (-0.094). These atypical observations were not detected by the PPM. That is, we noticed the presence of heteroskedasticity in the DJIA return series, which is in agreement with Hsu [29] and Hawkins [18]. Contrary to what we expected, for the period analyzed, we also noticed that there are not atypical observations in the DJIA return series.

6 Conclusions

Yao's method and a Gibbs sampling based idea were applied to implement an extension of the Product Partition Model (PPM). The goal was to deal with the change-point identification problem both in means and variances of normal data sequences.

An extensive Monte Carlo simulation study was performed with these two computational methods. In conclusion, for the scenarios considered in this paper, the results obtained by the approximate Gibbs sampling based algorithm were not significantly worse than those obtained Yao's method. Indeed, both algorithms presented similar errors. However, Yao's method may be preferable for staircase-style scenarios. The length of the sequence influenced the accuracy of the product estimates and the errors were smaller for longer sequences. Additionally, in scenarios without changes, smaller errors were observed for $p = 0.01$.

We also noticed that by using the PPM, one is likely to misinterpret atypical observations as structural changes. The PPM usually cannot properly identify atypical observations that occur separated by short periods of time. In spite of this, the PPM is a valuable tool in identifying change points in data sequences as we can see in numerous applications published in the literature.

As an illustrative example, we verified this by means of the results for the Dow Jones Industrial Average (DJIA) return series. The results indicated that both methods identified change points in the series and produced similar product estimates for the means and variances of the returns. Finally, a question that remains open is how the PPM could be adapted for forecasting. This is only one of the many possible interesting directions for future research in the area.

Acknowledgements

Rosangela H. Loschi acknowledges the CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) of the Ministry for Science and Technology of Brazil (grant 300325/2003-7 and 472066/2004-8) and PRPq-UFMG (grant 3893-UFMG/RTR/FUNDO/PRPq/RECEMDOUTORES/00), for a partial allowance to her research. The research of Frederico R. B. Cruz has been partially funded by the CNPq, grants 201046/1994-6, 301809/1996-8, 307702/2004-9, and 472066/2004-8, the FAPEMIG (*Fundação de Amparo à Pesquisa do Estado de Minas Gerais*), grants CEX-289/98 and CEX-855/98, and PRPq-UFMG, grant 4081-UFMG/RTR/FUNDO/PRPq/99. Ricardo Takahashi acknowledges the CNPq, grant 550722/2002-5.

References

[1] Markowitz HM. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.
[2] Sharpe WF. Capital asset prices: a theory of market equilibrium under conditions risk. *Journal of Finance*, 19:425–442, 1964.
[3] Lintner J. The valuation of risk assets and the selectionof risky investment in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47:13–37, 1965.
[4] Mossin J. Equilibrium in a capital asset market. *Econometrica*, 34(4): 768–783, 1966.
[5] Black F & Scholes M. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.
[6] Mandelbrot BB. The variation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.
[7] Fama E. The behavior of stock-market prices. *Journal of Business*, 38 (1):34–105, 1965.
[8] Hamilton JD. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.
[9] Eraker B. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *The Journal of Finance*, 59(3):1367–1403, 2004.
[10] Chacko G & Viceira LM. Spectral GMM estimation of continuous-time processes. *Journal of Econometrics*, 116:259–292, 2003.
[11] Bandi FM & Nguyen TH. On the functional estimation of jump-diffusion models. *Journal of Econometrics*, 116:293–328, 2003.
[12] Johannes M. The statistical and economic role of jumps in continuous-time interest rate models. *The Journal of Finance*, 59(1):227–260, 2004.
[13] Eraker B, Johannes M & Polson N. The impact of jumps in volatility and returns. *The Journal of Finance*, 58(3):1269–1300, 2003.

[14] Mendes BV. Computing robust risk measures in emerging markets using extreme value theory. *Emerging Markets Quarterly*, 4(2):25–42, 2000.

[15] Perron P. Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics*, 5(2):153–162, 1990.

[16] Barry D & Hartigan JA. Product partition models for change point problems. *The Annals of Statistics*, 20(1):260–279, 1992.

[17] Dueker MJ. Markov switching in GARCH processes and mean-reverting stock-market volatility. *Journal of Business & Economic Statistics*, 15 (1):26–34, 1997.

[18] Hawkins DM. Fitting multiple change-point models to data. *Computational Statistics & Data Analysis*, 37(3):323–341, 2001.

[19] Horváth L & Kokoszka P. Change-point detection with non-parametric regression. *Statistics*, 36(1):9–31, 2002.

[20] Jandhyala VK, Fotopoulos SB & Hawkins DM. Detection and estimation of abrupt changes in the variability of a process. *Computational Statistics & Data Analysis*, 40(1):1–19, 2002.

[21] Crowley EM. Product partition models for normal means. *Journal of the American Statistical Association*, 92(437):192–198, 1997.

[22] Quintana FA & Iglesias PL. Bayesian clustering and product partition models. *Journal of the Royal Statistical Society B*, 65(2):557–574, 2003.

[23] Loschi RH & Cruz FRB. An analysis of the influence of some prior specifications in the identification of change points via product partition model. *Computational Statistics & Data Analysis*, 39(4):477–501, 2002.

[24] Loschi RH, Cruz FRB, Iglesias PL & Arellano-Valle RB. A Gibbs sampling scheme to the product partition model: An application to change point problems. *Computers & Operations Research*, 30(3):463–482, 2003.

[25] Loschi RH & Cruz FRB. Extension to the product partition model: Computing the probability of a change. *Computational Statistics & Data Analysis*, 48(2):255–268, 2005.

[26] Loschi RH, Cruz FRB & Arellano-Valle RB. Multiple change point analysis for the regular exponential family using the product partition model. *Journal of Data Science*, 3(3):305–330, 2005.

[27] Yao Y. Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches. *The Annals of Statistics*, 12(4):1434–1447, 1984.

[28] Gelfand AE & Smith AFM. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85: 398–409, 1990.

[29] Hsu DA. A Bayesian robust detection of shift in the risk structure of stock market returns. *Journal of the American Statistical Association*, 77 (2):29–39, 1982.