Revista Colombiana de Estadística
Junio 2012, volumen 35, no. 1, pp. 15 a 38

Bayesian Analysis for Errors in Variables with Changepoint Models

Análisis bayesiano para modelos con errores en las variables con punto de cambio

Olga Cecilia Usuga[1,2,a], Freddy Hernández[2,b]

[1]Departamento de Ingeniería Industrial, Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia

[2]Departamento de Estadística, Instituto de Matemáticas y Estadística, Universidad de São Paulo, São Paulo, Brasil

Abstract

Changepoint regression models have originally been developed in connection with applications in quality control, where a change from the in-control to the out-of-control state has to be detected based on the avaliable random observations. Up to now various changepoint models have been suggested for differents applications like reliability, econometrics or medicine. In many practical situations the covariate cannot be measured precisely and an alternative model are the errors in variable regression models. In this paper we study the regression model with errors in variables with changepoint from a Bayesian approach. From the simulation study we found that the proposed procedure produces estimates suitable for the changepoint and all other model parameters.

Key words: Bayesian analysis, Changepoint models, Errors in variables models.

Resumen

Los modelos de regresión con punto de cambio han sido originalmente desarrollados en el ámbito de control de calidad, donde, basados en un conjunto de observaciones aleatorias, es detectado un cambio de estado en un proceso que se encuentra controlado para un proceso fuera de control. Hasta ahora varios modelos de punto de cambio han sido sugeridos para diferentes aplicaciones en confiabilidad, econometría y medicina. En muchas situaciones prácticas la covariable no puede ser medida de manera precisa, y un modelo alternativo es el de regresión con errores en las variables. En este trabajo estudiamos el modelo de regresión con errores en las variables con

[a]Assistant professor. E-mail: ousuga@udea.edu.co
[b]Ph.D. Student in Statistic. E-mail: fhernanb@ime.usp.br

punto de cambio desde un enfoque bayesiano. Del estudio de simulación se encontró que el procedimiento propuesto genera estimaciones adecuadas para el punto de cambio y todos los demás parámetros del modelo.

***Palabras clave*:** análisis bayesiano, modelos con errores en las variables, modelos con punto de cambio.

1. Introduction

Linear regression is one of the most widely used statistical tools to analyze the relationship between a response variable Y and a covariate x. Under the classic model of simple linear regression the relationship between Y and x is given by

$$Y_i = \alpha + \beta x_i + e_i, \quad i = 1, \dots, n \tag{1}$$

where α and β are unknown constants and $e_i \overset{\text{ind}}{\sim} N(0, \sigma_e^2)$, for $i = 1, \dots, n$, where $N(a, b^2)$ denotes the normal distribution with location parameter a and scale parameter $b > 0$. Usually it is assumed that x_i is measured without error in many practical situations this assumption is violated. Instead of observing x_i is observed

$$X_i = x_i + u_i \quad i = 1, \dots, n \tag{2}$$

where x_i is the unobservable variable and $u_i \sim N(0, \sigma_u^2)$. Measurements errors (e_i, u_i) are assumed independent and identically distribuited; see, for example, Cheng & Van Ness (1999) and Fuller (1987).

Measurement error (ME) model (also called errors-in-variables model) is a generalization of standard regression models. For the simple linear ME model, the goal is to estimate from bivariate data a straight line fit between X and Y, both of which are measured with error. Applications in which the variables are measured with error are perhaps more common than those in which the variables are measured without error. Many variables in the medical field, such as blood pressure, pulse frequency, temperature, and other blood chemical variables, are measured with error. Variables of agriculture such as rainfalls, content of nitrogen of the soil and degree of infestation of plagues can not be measured accurately. In management sciences, social sciences, and in many other sciences almost all measurable variables are measured with error.

There are three ME models depending on the assumptions about x_i. If the $x_i^{,}s$ are unknown constan, then the model is known as a functional ME model; whereas, if the $x_i^{,}s$ are independent identically distribuited random variables and independent of the errors, the model is known as a structural ME model. A third model, the ultrastructural ME model, assumes that the $x_i^{,}s$ are independent random variables but not identically distributed, instead of having possibly different means, μ_i, and common variance σ^2. The ultrastructural model is a generalization of the functional and structural models: if $\mu_1 = \cdots = \mu_n$, then the ultrastructural model reduces to the structural model; whereas if $\sigma^2 = 0$, then the ultrastructural model reduces to the functional model (Cheng & Van Ness 1999).

It is common to assume that all the random variables in the ME model are jointly normal in this case the structural ME model, is not identifiable. This means that different sets of parameters can lead to the same joint distribution of X and Y. For this reason, the statistical literature have considered six assumptions about the parameters which lead to an identifiable structural ME model. The six assumptions have been studied extensively in econometrics; see for example Reiersol (1950), Bowden (1973), Deistler & Seifert (1978) and Aigner, Hsiao, Kapteyn & Wansbeek (1984). They make identifiable the structural ME model.

1. The ratio of the error variances, $\lambda = \sigma_e^2/\sigma_u^2$, is known
2. The ratio $k_x = \sigma_x^2/(\sigma_x^2 + \sigma_u^2)$ is known
3. σ_u^2 is known
4. σ_e^2 is known
5. The error variances, σ_u^2 and σ_e^2, are known
6. The intercept, α, is known and $E(X) \neq 0$

Assumption 1 is the most popular of these assumptions and is the one with the most published theoretical results; the assumption 2 is commonly found in the social science and psychology literatures; the assumption 3 is a popular assumption when working with nonlinears models; the assumption 4 is less useful and cannot be used to make the equation error model or the measurement error model with more than one explanatory variable identifiable; the assumption 5 frequently leads to the same estimates as those for assumption 1 and also leads to an overidentified model, and finally the assumption 6 does not make the normal model, with more than one identifiable explanatory variable.

In the structural ME model, usually it is assumed that $x_i \sim N(\mu_x, \sigma_x^2)$, $e_i \sim N(0, \sigma_e^2)$ and $u_i \sim N(0, \sigma_u^2)$ with x_i, e_i and u_i independent. A variation of the structural ME model proposed by Chang & Huang (1997) consists in relaxing the assumption of $x_i \sim N(\mu_x, \sigma_x^2)$, so that the $x_i^{`}s$ are not identically distribuited. Consider an example that can be stated as follows. Let x_i denote some family's true income at time i, let X_i denote the family's measured income, let Y_i denote its measured comsumption. During the observations (X_i, Y_i), some new impact on the financial system in the society may occur, for instance, a new economic policy may be announced. The family's true income structure may start to change some time after the announcement; however, the relation between income and consumption remains unchanged. Under this situation Chang & Huang (1997) considered the structural ME model defined by (1) and (2), where the covariate x_i has a change in its distribution given by:

$$
\begin{aligned}
x_i &\sim N(\mu_1, \sigma_x^2) \qquad i = 1, \dots, k \\
x_i &\sim N(\mu_2, \sigma_x^2) \qquad i = k+1, \dots, n
\end{aligned}
$$

This model with change in the mean of x_i at time k is called structural ME model with changepoint.

The problems with changepoint have been extensively studied. Hinkley (1970) developed a frequentist approach to the changepoint problems and Smith (1975) developed a Bayesian approach. The two works were limited to the inference about the point in a sequence of random variables at which the underlying distribution changes. Carlin, Gelfand & Smith (1992) extended the Smith approach using Markov chain Monte Carlo (MCMC) methods for changepoint with continuos time. Lange, Carlin & Gelfand (1994) and Kiuchi, Hartigan, Holford, Rubinstein & Stevens (1995) used MCMC methods for longitudinal data analysis in AIDS studies. Although there are works in the literature on changepoint problems with Bayesian approach, the Bayesian approach for ME models has not been studied. Hernandez & Usuga (2011) proposed a Bayesian approach for reliability models. The goal of this paper is to propose a Bayesian approach to make inferences in structural ME model with changepoint.

The plan of the paper is as follows. Section 2 presents the Bayesian formulation of the model, Section 3 presents the simulation study and Section 4 presented an application with a real dataset and finally some concluding remarks are presents in Section 5.

2. Structural Errors in Variables Models with Changepoint

The structural ME model with one changepoint that will be studied in this paper is defined by the following equations:

$$\left.\begin{array}{ll} Y_i = \alpha_1 + \beta_1 x_i + e_i & i = 1, \dots, k \\ Y_i = \alpha_2 + \beta_2 x_i + e_i & i = k+1, \dots, n \end{array}\right\} \tag{3}$$

and

$$X_i = x_i + u_i \qquad i = 1, \dots, n\} \tag{4}$$

where X_i and Y_i are observable random variables, x_i is an unobservable random variable, e_i and u_i are random errors with the assumption that $(e_i, u_i, x_i)^T$ are independents for $i = 1, \dots, n$ with distribution given by:

$$\begin{pmatrix} e_i \\ u_i \\ x_i \end{pmatrix} \sim N_3\left(\begin{pmatrix} 0 \\ 0 \\ \mu_1 \end{pmatrix}, \begin{pmatrix} \sigma^2_{e_1} & 0 & 0 \\ 0 & \sigma^2_{u_1} & 0 \\ 0 & 0 & \sigma^2_{x_1} \end{pmatrix}\right), \qquad i = 1, \dots, k$$

$$\begin{pmatrix} e_i \\ u_i \\ x_i \end{pmatrix} \sim N_3\left(\begin{pmatrix} 0 \\ 0 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma^2_{e_2} & 0 & 0 \\ 0 & \sigma^2_{u_2} & 0 \\ 0 & 0 & \sigma^2_{x_2} \end{pmatrix}\right), \qquad i = k+1, \dots, n$$

The observed data (Y_i, X_i) have the following joint distribution for $i = 1, \dots, n$.

$$\begin{pmatrix} Y_i \\ X_i \end{pmatrix} \overset{\text{i.i.d}}{\sim} N_2\left(\begin{pmatrix} \alpha_1 + \beta_1\mu_1 \\ \mu_1 \end{pmatrix}, \begin{pmatrix} \beta_1^2\sigma_{x_1}^2 + \sigma_{e_1}^2 & \beta_1\sigma_{x_1}^2 \\ \beta_1\sigma_{x_1}^2 & \sigma_{x_1}^2 + \sigma_{u_1}^2 \end{pmatrix}\right), \; i = 1, \dots, k$$
$$\begin{pmatrix} Y_i \\ X_i \end{pmatrix} \overset{\text{i.i.d}}{\sim} N_2\left(\begin{pmatrix} \alpha_2 + \beta_2\mu_2 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \beta_2^2\sigma_{x_2}^2 + \sigma_{e_2}^2 & \beta_2\sigma_{x_2}^2 \\ \beta_2\sigma_{x_2}^2 & \sigma_{x_2}^2 + \sigma_{u_2}^2 \end{pmatrix}\right), \; i = k+1, \dots, n$$

The likelihood function $L(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{Y})$, where $\boldsymbol{\theta} = (k, \alpha_1, \beta_1, \mu_1, \sigma_{x_1}^2, \sigma_{e_1}^2, \sigma_{u_1}^2, \alpha_2, \beta_2, \mu_2, \sigma_{x_2}^2, \sigma_{e_2}^2, \sigma_{u_2}^2)^T$, $\boldsymbol{X} = (X_1, \dots, X_n)^T$ and $\boldsymbol{Y} = (Y_1, \dots, Y_n)^T$ can be written as:

$$\begin{aligned} L(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{Y}) \propto\ & (\beta_1^2\sigma_{u_1}^2\sigma_{x_1}^2 + \sigma_{e_1}^2\sigma_{x_1}^2 + \sigma_{u_1}^2\sigma_{e_1}^2)^{-k/2} \exp\left\{-\frac{A}{C}\right\} \\ & \times (\beta_2^2\sigma_{u_2}^2\sigma_{x_2}^2 + \sigma_{e_2}^2\sigma_{x_2}^2 + \sigma_{u_2}^2\sigma_{e_2}^2)^{-(n-k)/2} \exp\left\{-\frac{B}{D}\right\} \end{aligned} \tag{5}$$

where

$$\begin{aligned} A =& (\beta_1^2\sigma_{x_1}^2 + \sigma_{e_1}^2)\sum_{i=1}^{k}(X_i - \mu_1)^2 - 2\beta_1\sigma_{x_1}^2\sum_{i=1}^{k}(Y_i - \alpha_1 - \beta_1\mu_1)(X_i - \mu_1) \\ & + (\sigma_{x_1}^2 + \sigma_{u_1}^2)\sum_{i=1}^{k}(Y_i - \alpha_1 - \beta_1\mu_1)^2 \\ B =& (\beta_2^2\sigma_{x_2}^2 + \sigma_{e_2}^2)\sum_{i=k+1}^{n}(X_i - \mu_2)^2 - 2\beta_2\sigma_{x_2}^2\sum_{i=k+1}^{n}(Y_i - \alpha_2 - \beta_2\mu_2)(X_i - \mu_2) \\ & + (\sigma_{x_2}^2 + \sigma_{u_2}^2)\sum_{i=k+1}^{n}(Y_i - \alpha_2 - \beta_2\mu_2)^2 \\ C =& 2(\beta_1^2\sigma_{u_1}^2\sigma_{x_1}^2 + \sigma_{e_1}^2\sigma_{x_1}^2 + \sigma_{u_1}^2\sigma_{e_1}^2) \\ D =& 2(\beta_2^2\sigma_{u_2}^2\sigma_{x_2}^2 + \sigma_{e_2}^2\sigma_{x_2}^2 + \sigma_{u_2}^2\sigma_{e_2}^2) \end{aligned}$$

2.1. Prior and Posterior Distributions

It was considered the discrete uniform distribution for k in the range $1, \dots, n$ allowing values of $k = 1$ or $k = n$, which would indicate the absence of changepoint. Also, it was considered inverse Gamma distribution for each of the variances and normal distributions for the remaining parameters to obtain posterior distributions. The above distributions with their hyperparameters are given below.

$$p(k) = \begin{cases} P(K = k) = \frac{1}{n}, & k = 1, \dots, n, \\ 0, & \text{otherwise}, \end{cases}$$

$$\begin{aligned} \sigma_{e_1}^2 &\sim GI(a_{e_1}, b_{e_1}) \quad & \sigma_{e_2}^2 &\sim GI(a_{e_2}, b_{e_2}) \\ \sigma_{u_1}^2 &\sim GI(a_{u_1}, b_{u_1}) \quad & \sigma_{u_2}^2 &\sim GI(a_{u_2}, b_{u_2}) \end{aligned}$$

$$\sigma^2_{x_1} \sim GI(a_{x_1}, b_{x_1}) \quad \sigma^2_{x_2} \sim GI(a_{x_2}, b_{x_2})$$
$$\alpha_1 \sim N(\alpha_{01}, \sigma^2_{\alpha_1}) \quad \alpha_2 \sim N(\alpha_{02}, \sigma^2_{\alpha_2})$$
$$\beta_1 \sim N(\beta_{01}, \sigma^2_{\beta_1}) \quad \beta_2 \sim N(\beta_{02}, \sigma^2_{\beta_2})$$
$$\mu_1 \sim N(\mu_{01}, \sigma^2_{\mu_1}) \quad \mu_2 \sim N(\mu_{02}, \sigma^2_{\mu_2})$$

where $GI(a, b)$ denotes the inverse Gamma distribution with shape parameter $a > 0$ and scale parameter $b > 0$. The hyperparameters a_{e_1}, b_{e_1}, a_{e_2}, b_{e_2}, a_{u_1}, b_{u_1}, a_{u_2}, b_{u_2}, a_{x_1}, b_{x_1}, a_{x_2} y b_{x_2}, α_{01}, $\sigma^2_{\alpha_1}$, α_{02}, $\sigma^2_{\alpha_2}$, β_{01}, $\sigma^2_{\beta_1}$, β_{02}, $\sigma^2_{\beta_2}$, μ_{01}, $\sigma^2_{\mu_1}$, μ_{02} and $\sigma^2_{\mu_2}$ are considered as known. The prior distribution for the vector $\boldsymbol{x}$ is denoted by $\pi(\boldsymbol{x})$ and it is based on the assumption of independence and normality of the model.

The likelihood function based on complete data $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{x} = (x_1, \dots, x_n)^T$ is denoted by $L^*(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{Y})$ and can be expressed as

$$L^*(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{Y}) \propto \prod_{i=1}^{k} (\sigma^2_{e_1} \sigma^2_{u_1} \sigma^2_{x_1})^{-\frac{1}{2}} e^E \prod_{i=k+1}^{n} (\sigma^2_{e_2} \sigma^2_{u_2} \sigma^2_{x_2})^{-\frac{1}{2}} e^F \tag{6}$$

where

$$E = -\frac{(Y_i - \alpha_1 - \beta_1 x_i)^2}{2\sigma^2_{e_1}} - \frac{(X_i - x_i)^2}{2\sigma^2_{u_1}} - \frac{(x_i - \mu_1)^2}{2\sigma^2_{x_1}}$$
$$F = -\frac{(Y_i - \alpha_2 - \beta_2 x_i)^2}{2\sigma^2_{e_2}} - \frac{(X_i - x_i)^2}{2\sigma^2_{u_2}} - \frac{(x_i - \mu_2)^2}{2\sigma^2_{x_2}}$$

Based on the prior distributions for each parameter the posterior distribution for $\boldsymbol{\theta}$ can be written as

$$\begin{aligned}\pi(\boldsymbol{\theta}, \boldsymbol{x} \mid \boldsymbol{X}, \boldsymbol{Y}) \propto & \prod_{i=1}^{k} (\sigma^2_{e_1} \sigma^2_{u_1})^{-\frac{1}{2}} e^G \prod_{i=k+1}^{n} (\sigma^2_{e_2} \sigma^2_{u_2})^{-\frac{1}{2}} e^H p(k) \\ & \times \pi(\alpha_1)\, \pi(\alpha_2)\, \pi(\beta_1)\, \pi(\beta_2)\, \pi(\mu_1)\, \pi(\mu_2) \\ & \times \pi(\sigma^2_{e_1})\, \pi(\sigma^2_{e_2})\, \pi(\sigma^2_{u_1})\, \pi(\sigma^2_{u_2})\, \pi(\sigma^2_{x_1})\, \pi(\sigma^2_{x_2})\, \pi(\boldsymbol{x})\end{aligned} \tag{7}$$

where

$$G = -\frac{(Y_i - \alpha_1 - \beta_1 x_i)^2}{2\sigma^2_{e_1}} - \frac{(X_i - x_i)^2}{2\sigma^2_{u_1}}$$
$$H = -\frac{(Y_i - \alpha_2 - \beta_2 x_i)^2}{2\sigma^2_{e_2}} - \frac{(X_i - x_i)^2}{2\sigma^2_{u_2}}$$

The conditional posterior distributions for the parameters obtained from the previous posterior distribution are given in Appendix. For almost all the parameters posterior distributions with pdf known were obtained, except for the parameter k. The conditional posterior distribution of the changepoint k in the model has not pdf known, making it necessary to use the Gibbs sampler, introduced by

Geman & Geman (1984) to approximate this distribution. The sampler Gibbs is an iterative algorithm that constructs a dependent sequence of parameter values whose distribution converges to the target joint posterior distribution (Hoff 2009).

The procedure used to implement the Gibbs sampler to the problem was:

1. Generate appropriate initial values for each of the 13 parameters to create the initial parameter vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{13})^T$.
2. Update the component $j = 1, \dots, 13$ of $\boldsymbol{\theta}$ generating a random observation for the parameter θ_j using the corresponding posterior distribution of Appendix and the subset of parameters of $\boldsymbol{\theta}$ present in the posterior distribution of θ_j.
3. Repeat step 2 a number of times until obtaining convergence in all the parameters.

3. Simulation Study

In this section we present the results of implementation of the Gibbs sampler for the model given in equations (3) and (4) under three different assumptions of the parameters. In the first case we analyze the model with a simulated dataset considering $\lambda = \sigma^2_{e_i}/\sigma^2_{u_i}$ known; in the second case we consider the variances $\sigma^2_{u_1}$ and $\sigma^2_{u_2}$ known and equal, and in the third case we consider the variances $\sigma^2_{e_1}$ and $\sigma^2_{e_2}$ known and equals. In adittion to the above cases we also analized the changepoint estimate of the model for different n values with the aim of observing the behavior of the estimate of k with respect to its true value.

3.1. λ Known

In Table 1 we present a dataset of $n = 60$ observations generated in R Development Core Team (2011) from the model given in equations (3) and (4) with the assumption of $\lambda = 1$ considering the following set of parameters: $k = 20$, $\alpha_1 = 2$, $\beta_1 = 2$, $\mu_1 = 1$, $\sigma^2_{x_1} = 1$, $\sigma^2_{e_1} = 1.5$, $\sigma^2_{u_1} = \sigma^2_{e_1}/\lambda$, $\alpha_2 = -1$, $\beta_2 = 4$, $\mu_2 = 5$, $\sigma^2_{x_2} = 2$, $\sigma^2_{e_2} = 2.5$ and $\sigma^2_{u_2} = \sigma^2_{e_2}/\lambda$. Figure 1 shows the scatter plot for the data generated and there is not clear indication of the changepoint in the model structure.

We used the Gibbs sampler to obtain estimates of the parameters. The prior distributions used to run the Gibbs sampler were as follows: $\alpha_1 \sim N(2, 15)$, $\beta_1 \sim N(2, 15)$, $\mu_1 \sim N(1, 15)$, $\sigma^2_{x_1} \sim GI(2, 5)$, $\sigma^2_{e_1} \sim GI(2, 5)$, $\sigma^2_{u_1} \sim GI(2, 5)$, $\alpha_2 \sim N(-1, 15)$, $\beta_2 \sim N(4, 15)$, $\mu_2 \sim N(5, 15)$, $\sigma^2_{x_2} \sim GI(2, 5)$, $\sigma^2_{e_2} \sim GI(2, 5)$ and $\sigma^2_{u_2} \sim GI(2, 5)$.

Table 1: Random sample of simulated data with $\lambda = 1$.

X	0.05	4.34	1.18	0.65	1.47	−0.57	2.42	0.78	1.63	−1.02	−0.78	0.48
Y	0.03	4.92	4.42	1.62	5.91	3.81	6.60	4.97	2.79	2.37	1.45	5.24
X	1.05	2.50	3.76	0.61	1.46	0.98	1.59	−0.95	4.89	2.28	7.09	7.18
Y	1.74	6.07	6.96	6.65	2.08	2.63	5.55	1.39	17.84	9.15	16.82	25.40
X	6.58	4.85	6.23	5.30	7.29	6.73	6.78	7.46	2.86	3.33	3.80	9.66
Y	18.51	18.86	13.12	19.57	17.16	22.71	22.16	26.90	21.30	12.82	27.43	27.96
X	3.63	3.66	5.70	5.64	2.15	3.13	9.10	9.88	4.73	7.48	2.55	11.11
Y	9.53	12.68	18.54	19.15	20.43	21.36	32.97	27.38	13.67	18.73	14.54	23.11
X	7.10	4.95	9.17	2.03	9.54	5.08	7.36	6.37	4.35	1.45	7.67	2.97
Y	25.54	21.00	27.77	9.75	27.62	22.29	19.45	17.34	23.68	13.99	25.15	11.38

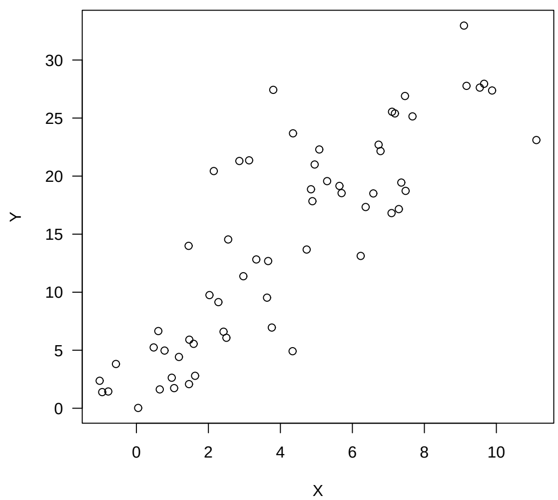

Figure 1: Scatter plot for simulated data with $\lambda = 1$.

We ran five chains of the Gibbs sampler. Each sequence was run for 11000 iterations with a burn-in of 1000 samples. The vectors of initial values for each of the chains were:

$$\boldsymbol{\theta}_1^{(0)} = (5, 1.886, 1.827, 2.4, 0.942, 1.134, 1.015, -1.5, 2.100, 1.3, 0.6, 0.893, 1.8)$$
$$\boldsymbol{\theta}_2^{(0)} = (10, 2.537, 1.225, 2.2, 1.404, 2.171, 0.552, 0.2, 3.500, 4.3, 1.1, 0.903, 3.4)$$
$$\boldsymbol{\theta}_3^{(0)} = (30, 1.856, 1.855, 2.6, 0.928, 1.087, 1.029, -0.3, 3.829, 4.5, 2.0, 0.900, 2.1)$$
$$\boldsymbol{\theta}_4^{(0)} = (40, 2.518, 1.242, 2.8, 1.386, 2.142, 0.571, -2.0, 3.829, 3.5, 2.8, 0.901, 1.4)$$
$$\boldsymbol{\theta}_5^{(0)} = (50, 2.516, 1.244, 1.8, 1.383, 2.138, 0.573, -1.3, 3.829, 2.5, 3.5, 0.899, 2.4)$$

The above vectors were obtained by the following procedure. For fixed values of $k = 5, 10, 30, 40, 50$ numerical methods were used to determine the values of $\boldsymbol{\theta}$ that maximize the likelihood function given in (5). These estimates were obtained using the function `optim` of R Development Core Team (2011), which uses optimization methods quasi-Newton such as the bounded limited-memory algorithm L-BFGS-B (Limited memory, Broyden- Fletcher-Goldfarb-Shanno, Bounded) proposed by Byrd, Lu, Nocedal & Zhu (1995).

In order to verify the convergence of the chains we used the diagnostic indicator proposed by Brooks & Gelman (1998). The diagnostic value of R found in this case was 1.04; values close to 1 indicate convergence of the chains. Additionally, for each parameter, the posterior distribution was examined visually by monitoring the density estimates, the sample traces, and the autocorrelation function. We found not evidence of trends or high correlations. Figures 2 and 3 show the Highest Density Region (HDR) graphics for the parameters of the chain 1. These graphics show that the true values of the model parameters are in the Highest Density Regions.

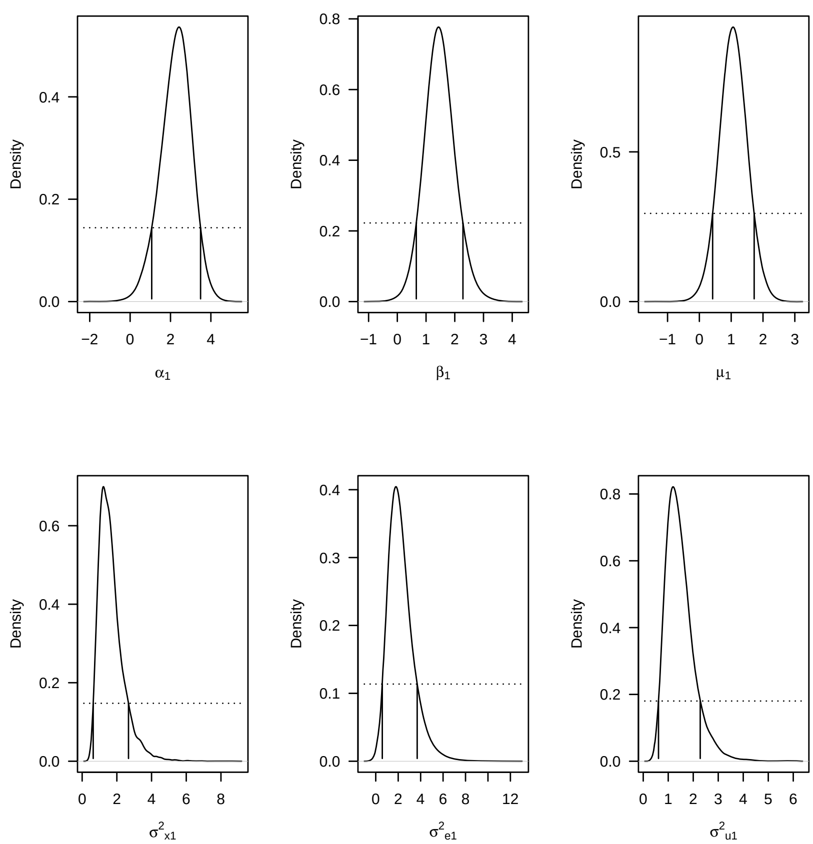

FIGURE 2: HDR plot for α_1, β_1, μ_1, $\sigma^2_{x_1}$,$\sigma^2_{e_1}$ and $\sigma^2_{u_1}$.

Table 2 presents the posterior mean and standard deviation (SD) for the model parameters and the 90% HDR interval. Note that the true values parameters are close to mean and are within the HDR interval.

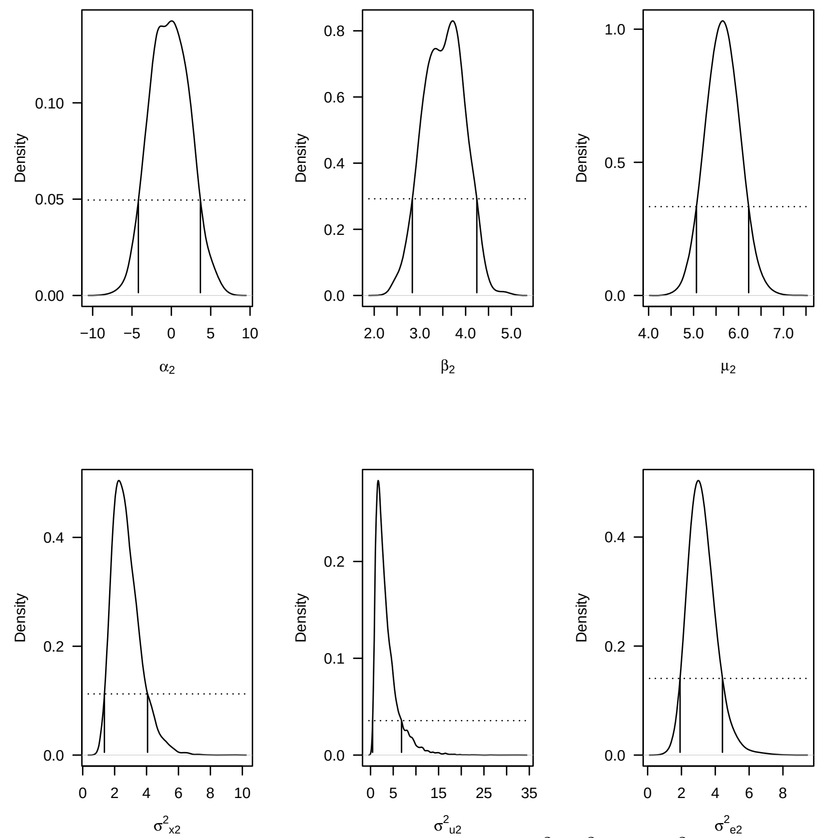

FIGURE 3: HDR plot for α_2, β_2, μ_2, $\sigma^2_{x_2}$,$\sigma^2_{e_2}$ and $\sigma^2_{u_2}$.

TABLE 2: Posterior mean, standard deviation (SD), HDRlower and HDRupper of parameters with $\lambda = 1$.

Parameter	Mean	SD	HDRlower	HDRupper
k	19.99	0.10	-	-
α_1	2.21	0.83	0.94	3.57
β_1	1.50	0.54	0.64	2.38
μ_1	1.09	0.40	0.44	1.74
$\sigma^2_{x_1}$	1.64	0.72	0.60	2.61
$\sigma^2_{e_1}$	2.28	1.14	0.62	3.77
$\sigma^2_{u_1}$	1.47	0.59	0.56	2.24
α_2	0.17	2.54	-3.97	4.46
β_2	3.44	0.45	2.71	4.20
μ_2	5.72	0.38	4.10	5.34
$\sigma^2_{x_2}$	2.86	0.94	1.39	4.22
$\sigma^2_{e_2}$	3.77	2.68	0.53	7.32
$\sigma^2_{u_2}$	3.18	0.81	1.86	4.42

3.2. $\sigma^2_{u_1}$ and $\sigma^2_{u_2}$ Known

In this case we consider the structural ME model with $\sigma^2_{u_1} = \sigma^2_{u_2} = 2$. Table 3 shows a dataset of size $n = 60$ generated from the model given in equations (3) and (4) with the following set of parameters: $\alpha_1 = 2$, $\beta_1 = 2$, $\mu_1 = 1$, $\sigma^2_{x_1} = 1$, $\sigma^2_{e_1} = 1.5$, $\alpha_2 = -1$, $\beta_2 = 4$, $\mu_2 = 5$, $\sigma^2_{x_2} = 2$ and $\sigma^2_{e_2} = 2.5$. Figure 4 shows the scatter plot for the simulated data.

Table 3: Random sample of data simulated with $\sigma^2_{u_1} = \sigma^2_{u_2} = 2$.

X	2.03	2.16	1.68	-0.07	1.00	-0.82	1.42	-0.42	3.36	0.88	0.12	0.80
Y	5.11	4.31	5.33	2.73	0.33	1.69	7.48	3.06	2.65	0.48	1.82	5.37
X	2.84	4.15	-1.54	0.84	1.55	0.99	-0.27	4.16	6.13	5.01	5.09	1.12
Y	0.42	5.20	3.75	4.88	3.87	0.73	6.01	8.41	20.03	18.82	15.29	8.10
X	5.40	3.28	8.06	5.78	5.68	3.26	2.48	3.72	2.85	6.13	2.85	8.47
Y	8.55	16.47	26.13	15.54	16.11	14.48	11.52	21.86	9.55	24.49	14.44	24.76
X	3.18	3.90	2.58	7.58	5.59	6.79	7.20	4.01	6.10	5.73	1.82	7.95
Y	20.74	15.84	4.54	20.84	20.96	24.59	23.95	11.74	18.99	15.13	9.98	29.01
X	4.42	4.01	7.72	9.25	4.60	4.73	0.52	0.46	2.76	5.44	7.22	3.33
Y	13.82	14.92	23.08	32.71	10.53	22.03	11.28	14.74	8.30	15.60	30.96	17.54

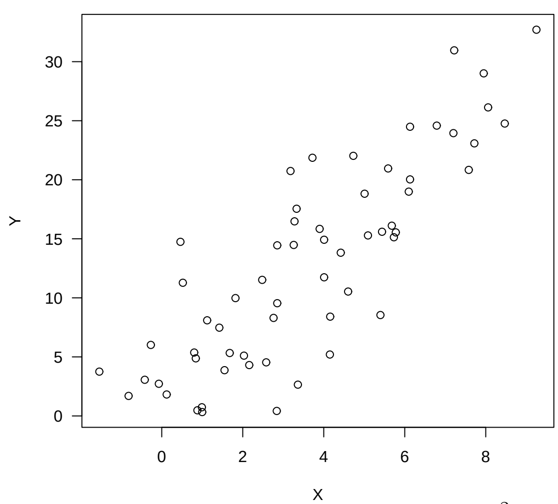

Figure 4: Scatter plot for the simulated data with $\sigma^2_{u_1} = \sigma^2_{u_2} = 2$.

The prior distributions for $\alpha_1, \beta_1, \mu_1, \sigma^2_{x_1}, \sigma^2_{e_1}, \alpha_2, \beta_2, \mu_2, \sigma^2_{x_2}$, and $\sigma^2_{e_2}$ were the same considered in the case of λ known. The vectors of initial values for the model parameters in each of the five Markov chain were as follows:

$$\boldsymbol{\theta}_1^{(0)} = (5, 1.0, 2.305, 1.063, 0.189, 3.0, 2, -2.0, 4.071, 4.727, 2.0, 1.984, 2)$$
$$\boldsymbol{\theta}_2^{(0)} = (10, 1.5, 2.800, 1.063, 0.189, 1.7, 2, -2.0, 4.071, 3.300, 3.0, 1.980, 2)$$
$$\boldsymbol{\theta}_3^{(0)} = (30, 3.1, 2.305, 0.300, 2.000, 2.0, 2, -0.5, 1.500, 3.200, 2.0, 2.700, 2)$$
$$\boldsymbol{\theta}_4^{(0)} = (40, 3.1, 1.900, 1.063, 1.400, 2.1, 2, -2.0, 4.071, 4.727, 1.4, 1.981, 2)$$
$$\boldsymbol{\theta}_5^{(0)} = (50, 2.9, 0.500, 1.063, 0.189, 2.6, 2, 0.7, 2.400, 1.500, 3.1, 1.981, 2).$$

The diagnostic value of convergence R was 1.01, indicating the convergence of the chains. A visual monitoring of the density estimates, the sample traces, and the correlation function for each parameter in each of the chains did not show any problem. In Figures 5 and 6 we present the HDR graphics for the parameters in the chain 1. The graphics show that true values of the parameters model are within the Highest Density Regions.

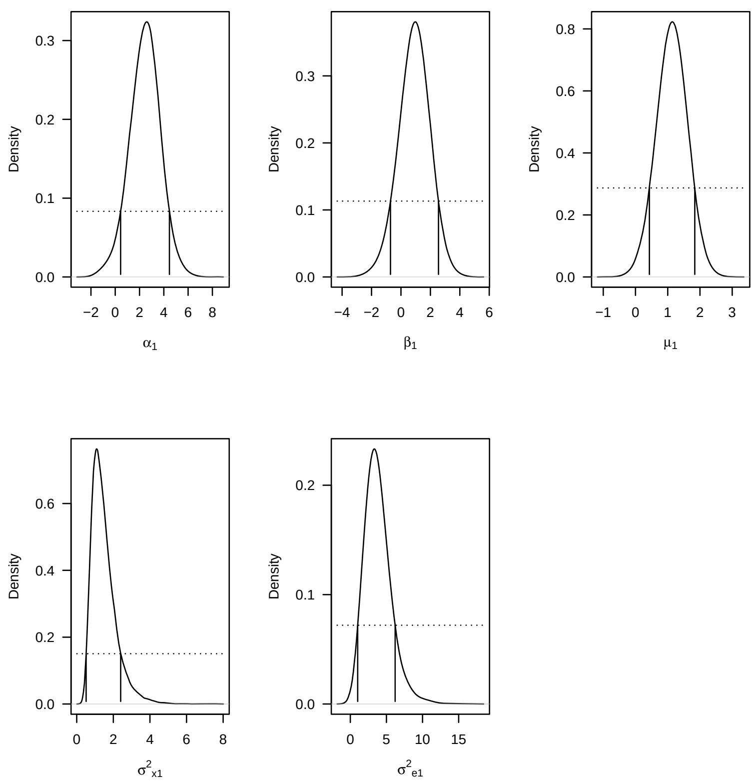

FIGURE 5: HDR plot for α_1, β_1, μ_1, $\sigma^2_{x_1}$ and $\sigma^2_{e_1}$.

Table 4 shows the posterior mean and standard deviation (SD) for each of the parameters model and the 90% HDR interval. It is noted again that the true values of the parameters are close to the mean and within the HDR intervals.

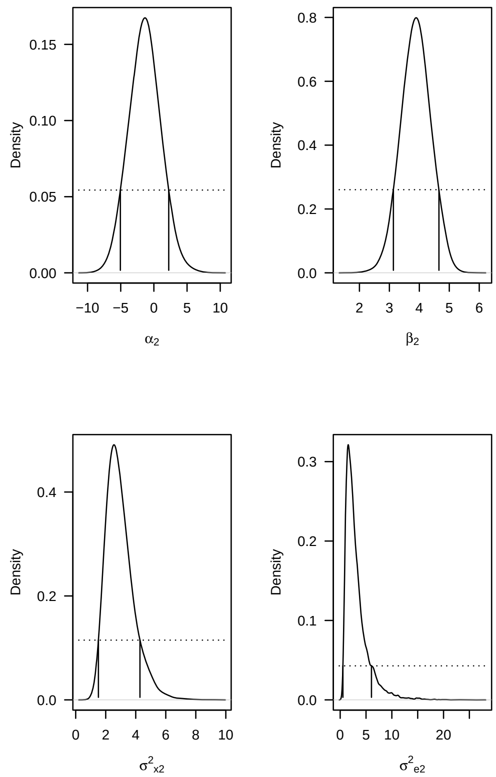

Figure 6: HDR plot for α_2, β_2, μ_2, $\sigma^2_{x_2}$ and $\sigma^2_{e_2}$.

Table 4: Posterior mean, standard deviation (SD), HDRlower and HDRupper of parameters with $\sigma^2_{u_1} = \sigma^2_{u_2}$.

Parameter	Mean	SD	HDRlower	HDRupper
k	19.25	0.61	-	-
α_1	2.48	1.24	0.48	4.47
β_1	0.89	0.99	-0.69	2.46
μ_1	1.15	0.43	0.44	1.86
$\sigma^2_{x_1}$	1.51	0.73	0.52	2.47
$\sigma^2_{e_1}$	3.85	1.72	1.05	6.18
α_2	-1.00	2.31	-4.62	2.91
β_2	3.82	0.46	3.04	4.55
μ_2	4.79	0.36	4.20	5.36
$\sigma^2_{x_2}$	3.02	0.94	1.53	4.39
$\sigma^2_{e_2}$	3.21	2.16	0.53	6.07

3.3. $\sigma^2_{e_1}$ and $\sigma^2_{e_2}$ Known

In this case we consider the structural ME model with $\sigma^2_{e_1} = \sigma^2_{e_2} = 2$. Table 5 shows a dataset of size $n = 60$ generated from the model given in equations (3) and (4) considering the same parameters values of the case λ known. Figure 7 presents the scatter plot of the simulated data.

Table 5: Random sample of the simulated data with $\sigma^2_{e_1} = \sigma^2_{e_2} = 2$.

X	-0.82	1.06	-2.25	0.76	1.71	-0.93	0.50	-0.08	1.12	-0.14	1.59	0.99
Y	2.72	2.75	1.88	8.66	5.12	2.51	7.00	7.64	4.55	7.50	2.89	2.55
X	-1.69	-0.14	2.32	-1.42	2.99	0.70	0.99	0.06	5.20	2.18	6.69	8.79
Y	2.47	0.43	5.32	3.10	6.63	1.99	5.57	3.39	11.10	13.80	17.71	32.28
X	4.98	8.05	6.29	5.99	5.32	5.80	7.73	5.45	4.27	2.84	7.69	10.61
Y	16.11	19.64	16.59	17.86	16.14	19.01	34.01	19.65	19.77	21.84	23.58	30.02
X	5.90	3.51	2.15	3.10	9.42	3.31	4.06	1.44	7.18	1.72	6.61	5.28
Y	21.45	11.51	17.44	17.75	29.13	18.25	20.95	8.24	24.11	8.12	29.31	19.54
X	4.68	1.88	5.02	1.76	7.67	5.31	6.42	7.79	4.32	1.20	3.22	3.37
Y	17.90	17.81	16.27	14.20	23.72	27.51	28.17	18.41	18.28	11.89	15.10	19.96

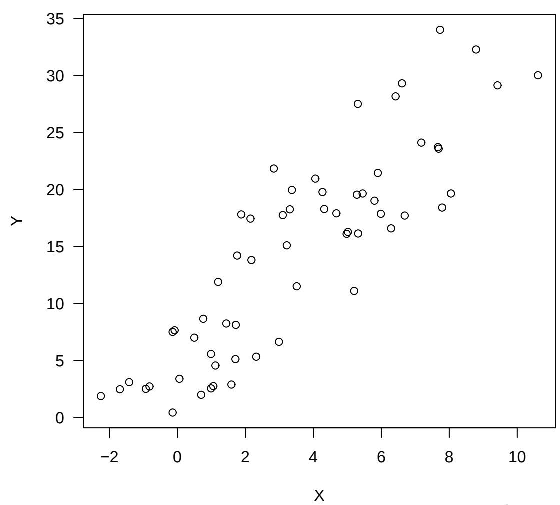

Figure 7: Scatter plot for the simulated data with $\sigma^2_{e_1} = \sigma^2_{e_2} = 2$.

The prior distributions for $\alpha_1, \beta_1, \mu_1, \sigma^2_{x_1}, \sigma^2_{u_1}, \alpha_2, \beta_2, \mu_2, \sigma^2_{x_2}$ and $\sigma^2_{u_2}$ were the same considered in the case of λ known. The vectors of the initial values for each of the five Markov chains were as follows:

$$\boldsymbol{\theta}_1^{(0)} = (5, 3.264, 2.650, 0.366, 0.437, 2.001, 1.276, 4.287, 3.0, 5.106, 3.793, 2.001, 2.394)$$
$$\boldsymbol{\theta}_2^{(0)} = (10, 1.000, 1.904, 1.500, 1.370, 2.001, 0.500, 4.512, 5.0, 2.000, 2.700, 2.001, 1.000)$$
$$\boldsymbol{\theta}_3^{(0)} = (30, 2.500, 2.650, 1.000, 0.437, 2.001, 1.500, 4.286, 2.0, 4.000, 3.792, 2.001, 0.700)$$
$$\boldsymbol{\theta}_4^{(0)} = (40, 0.500, 1.904, 0.900, 1.369, 2.000, 2.000, 4.512, 1.8, 5.500, 4.000, 2.001, 2.100)$$
$$\boldsymbol{\theta}_5^{(0)} = (50, 3.600, 2.652, 1.900, 0.437, 2.000, 0.700, 4.287, 4.1, 4.200, 3.792, 2.001, 2.000)$$

The diagnostic value of convergence was of 1.04, indicating the convergence of the chains. In Figures 8 and 9 we present the HDR graphics of the parameters for the chain 1. Note that the true values of the parameters model are within the Highest Density Regions.

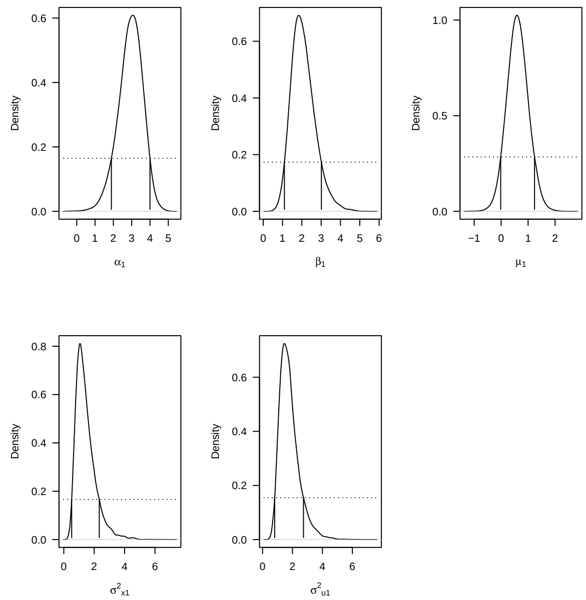

FIGURE 8: HDR plot for α_1, β_1, μ_1, $\sigma^2_{x_1}$ and $\sigma^2_{u_1}$.

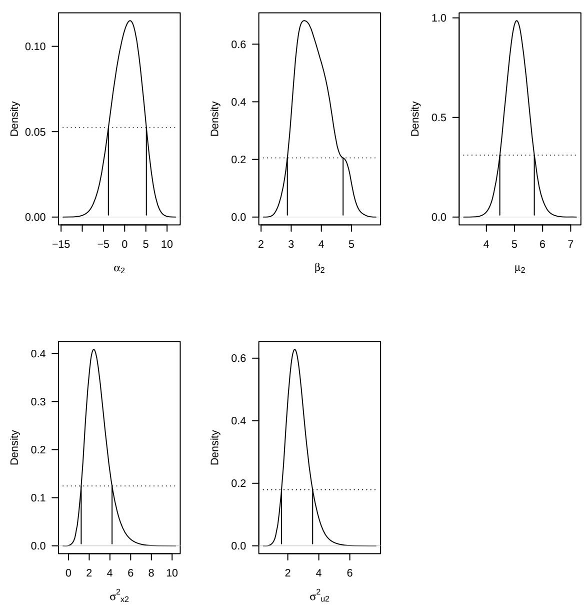

FIGURE 9: HDR plot for α_2, β_2, μ_2, $\sigma^2_{x_2}$ and $\sigma^2_{u_2}$.

Table 6 shows the posterior mean and the standard deviation (SD) for the model parameters and the 90% HDR interval. As in the previous cases the posterior means are close to the true values and are within the HDR interval.

Table 6: Posterior mean, standard deviation (SD), HDRlower and HDRupper of parameters with $\sigma^2_{e_1} = \sigma^2_{e_2}$.

Parameter	Mean	SD	HDRlower	HDRupper
k	20.22	0.47	-	-
α_1	2.95	0.62	0.96	2.98
β_1	2.09	0.61	1.09	3.02
μ_1	0.59	0.37	-0.01	1.20
$\sigma^2_{x_1}$	1.41	0.64	0.53	2.26
$\sigma^2_{u_1}$	1.76	0.64	0.80	2.64
α_2	0.60	2.62	-2.01	0.97
β_2	3.70	0.49	2.93	4.49
μ_2	5.16	0.35	4.57	5.73
$\sigma^2_{x_2}$	2.85	0.93	1.41	4.19
$\sigma^2_{u_2}$	2.62	0.63	1.59	3.53

3.4. Constant Sample Size and Variable Changepoint

In this case our objective was determine if the estimated changepoint of the model given in equations (3) and (4) differs from its true value when $n = 60$ is fixed. We generated nine random samples of size $n = 60$ based on the structure considered in Section 3.1; the values of the parameters were the same ones used in this section. The changepoint k for each nine random samples had different values, and the values were $k = 3, 5, 10, 20, 30, 40, 50, 55$ and 58. For each of the random samples were run five Markov chains of size 150000 with a burn in of 15000. Table 7 presents for the estimated changepoint k the posterior mean, standard deviation and percetiles of 10% and 90% when $n = 60$. Note that posterior mean of the changepoint is very close to the true value and the standard deviation tends to increase as the changepoint approach to the extreme values. Also the Table 7 shows that the distance between the percentiles 10% and 90% is at most 1%, which indicates that the posterior distribution for the parameter k is highly concentrated in one or two possible values and they match with the true value of k.

3.5. Sample Size and Changepoint Variable

In this case our objective was to determine if the estimated changepoint of the model given in equations (3) and (4) differs from its true value for differents values of n. As in the previous case we generated nine dataset with the structure of the Section 3.1. Each of the dataset had samples sizes of $n = 20, 30, 40, 50, 60, 70, 80, 90$ and 100. The true value for k in each of the nine set was $k = n/2$. Table 8 presents the posterior mean, standard deviation and 10% and 90% percentiles for the estimated changepoint and the true values of k. Again we see that the posterior mean of k is very close to the true values of k; it is also noted that the standard deviation tends to increase as the size sample n decreases; this means that if we have fewer observations the posterior distribution for k tends to have greater variability. As in the previous case the distance between the percentiles 10% and 90% is at most 1%, which means that the posterior distribution for the parameter k is highly concentrated in one or two possible values and they match the true value of k.

Table 7: Posterior mean, standard deviation (SD) and 10% and 90% percentiles of k estimated when $n = 60$.

k	Mean	SD	10%	90%
3	3.16	0.67	3.00	4.00
5	5.03	0.43	5.00	5.00
10	9.95	0.30	10.00	10.00
20	19.89	0.31	19.00	20.00
30	29.97	0.19	30.00	30.00
40	39.99	0.12	40.00	40.00
50	49.98	0.13	50.00	50.00
55	54.98	0.14	55.00	55.00
58	57.97	0.18	58.00	58.00

Table 8: Posterior mean, standard deviation (SD) and 10% and 90% percentiles of k.

n	k	Mean	SD	10%	90%
20	10	9.94	0.29	10.00	11.00
30	15	14.96	0.27	15.00	16.00
40	20	19.98	0.21	19.00	20.00
50	25	24.98	0.16	24.00	24.00
60	30	30.17	0.12	30.00	30.00
70	35	34.99	0.10	35.00	35.00
80	40	39.99	0.10	39.00	39.00
90	45	44.99	0.11	44.00	45.00
100	50	50.00	0.06	49.00	49.00

4. Application

This section illustrates the proposed procedure for the structural ME model with changepoint using a dataset of imports in the French economy.

Malinvaud (1968) provided the data of imports, gross domestic product (GDP), and other variables in France from 1949-1966. The main interest is to forecast the imports given the gross domestic product of the country. Chatterjee & Brockwell (1991) analyzed these data by the principal component method and found two patterns in the data; they argued that the models before and after 1960 must be different due to the fact the European Common Market began operations in 1960. Maddala (1992) considered a functional ME model; however, he ignored the possibility that some changes in the data may arise. Chang & Huang (1997) considered a structural ME model with changepoint using the likelihood ratio test based on the maximum Hotelling T^2 for the test of no change against the alternative of exactly one change and concluded that the changepoint ocurred in 1962. Table 9 presents the import data (Y) and gross domestic product (X).

Table 9: Imports and gross domestic product data from January 1949 to November 1966.

Year	1949	1950	1951	1952	1953	1954
GDP	149.30	161.20	171.50	175.50	180.80	190.70
Imports	15.90	16.40	19.00	19.10	18.80	20.40
Year	1955	1956	1957	1958	1959	1960
GDP	202.10	212.40	226.10	231.90	239.00	258.00
Imports	22.70	26.50	28.10	27.60	26.30	31.10
Year	1961	1962	1963	1964	1965	1966
GDP	269.80	288.40	304.50	323.40	336.80	353.90
Imports	33.30	37.00	43.30	49.00	50.30	56.60

The data were reanalized under a Bayesian perspective by adopting the structural ME model with changepoint. We considered non informative prior distributions for all parameters. Again as in the previous cases, we built five chains with different initial values of size 11000 with a burn in of 1000 samples to avoid correlations problems. We found the value $R = 1.03$, indicating the convergence of the chains.

Figure 10 shows the high concentration in the value 14 for the posterior distribution for the parameter k. The mean for this distribution is 13.92, which is the same obtained by Chang & Huang (1997), indicating that the data present a changepoint for the year 1962. Table 10 presents estimates for the remaining parameters of the model. The values are also close to the results obtained by Chang & Huang (1997). It is also noted that the means for β_1 and β_2 were 0.14 and 0.16, which indicates no significant changes in the slope for the trend lines before and after $k = 14$. The means obtained for the parameters α_1 and α_2 were -5.53 and -2.23, not being to close these values, which indicate that the trend lines before and after the change have different changepoints; this can be seen clearly in the Figure 11.

5. Conclusions

This paper proposes the Bayesian approach to study the structural ME model with changepoint. Through the simulation study was shown that the proposed procedure identifies correctly the point where the change comes to structure; note also that the variability in the posterior distribution of k decreases as the number of observations in the dataset increases. Another important aspect is that the variability of the posterior distribution for k increases as the true value of k is

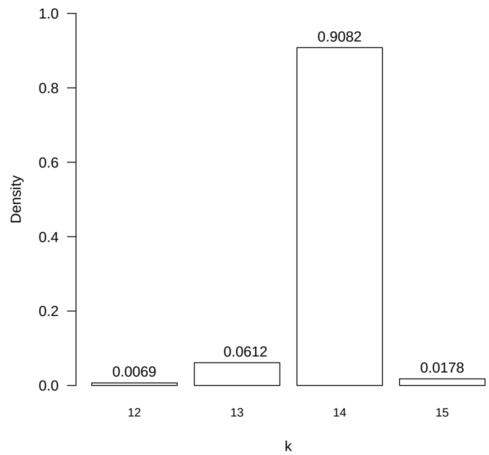

FIGURE 10: Posterior density for k.

TABLE 10: Posterior summary results.

Parameter	Mean	SD	HDRlower	HDRupper
α_1	-5.53	1.95	-8.77	-2.50
β_1	0.14	0.01	0.13	0.16
μ_1	16.33	3.88	9.99	22.67
$\sigma^2_{x_1}$	34601.81	13191.57	13641.42	51588.36
$\sigma^2_{e_1}$	1.93	0.89	0.72	3.07
$\sigma^2_{u_1}$	4.37	5.18	0.37	8.66
α_2	-2.23	3.78	-8.35	3.96
β_2	0.16	0.01	0.14	0.18
μ_2	5.39	3.89	-1.02	11.82
$\sigma^2_{x_2}$	70266.22	48727.80	1141.85	120582.77
$\sigma^2_{e_2}$	5.43	4.61	0.72	10.16
$\sigma^2_{u_2}$	5.14	13.70	-	-

close to 1. For the other parameters the proposed procedure generated posterior distributions with means very close to the real parameters in all cases considered. The proposed procedure generates chains that converge to the true parameters, regardless of whether or not identifiability assumptions.

Possible future works could consider other prior distributions such as non informative and skew normal and also introduce multiple changepoints in Y and X.

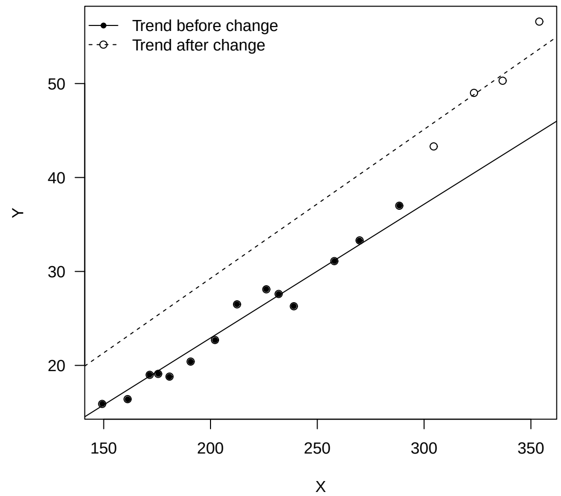

Figure 11: Scatter plot for the application.

6. Acknowledgements

The authors are grateful to the editor and referees for their valuable comments and helpful suggestions which led to improvements in the paper, and the professor Carlos Alberto de Bragança Pereira (IME-USP) for the motivation to write this work.

During the course of this work the authors received financial support from Coordenação de aperfeiçoamento de pessoal de nível superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Universidad de Antioquia.

[Recibido: octubre de 2010 — Aceptado: enero de 2011]

References

Aigner, D. J., Hsiao, C., Kapteyn, A. & Wansbeek, T. J. (1984), Latent variable models in econometrics, *in* Z. Griliches & M. D. Intriligator, eds, 'Handbook of Econometrics', Vol. 2, Elsevier Science, Amsterdam, pp. 1321–1393.

Bowden, R. J. (1973), 'The theory of parametric identification', *Econometrica* **41**, 1069–1174.

Brooks, S. P. & Gelman, A. (1998), 'General methods for monitoring convergence of iterative simulations', *Journal of Computational and Graphical Statistics* **7**(4), 434–455.

Byrd, R. H., Lu, P., Nocedal, J. & Zhu, C. (1995), 'A limited memory algorithm for bound constrained optimization', *SIAM Journal on Scientific Computing* **16**, 1190–1208.

Carlin, B. P., Gelfand, A. E. & Smith, A. F. M. (1992), 'Hierarchical Bayesian analysis of changepoint problems', *Applied of Statistics* **41**, 389–405.

Chang, Y. P. & Huang, W. T. (1997), 'Inferences for the linear errors-in-variables with changepoint models', *Journal of the American Statistical Association* **92**(437), 171–178.

Chatterjee, E. & Brockwell, P. J. (1991), *Regression Analysis by Example*, Wiley, New York.

Cheng, C. L. & Van Ness, J. W. (1999), *Statistical Regression with Measurement Error*, Arnold, London.

Deistler, M. & Seifert, H. G. (1978), 'Identifiability and consistent estimability in econometrics models', *Econometrica* **46**, 969–980.

Fuller, W. A. (1987), *Measurement Error Models*, Wiley, New York.

Geman, S. & Geman, D. (1984), 'Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741.

Hernandez, F. & Usuga, O. C. (2011), 'Analisis bayesiano para la distribucion lognormal generalizada aplicada a modelos de falla con censura', *Revista Colombiana de Estadística* **34**(1), 95–100.

Hinkley, D. V. (1970), 'Inference about the change-point in a sequence of random variables', *Biometrika* **57**, 1–17.

Hoff, P. (2009), *A First Course in Bayesian Statistical Methods*, Springer, New York.

Kiuchi, A. S., Hartigan, J. A., Holford, T. R., Rubinstein, P. & Stevens, C. E. (1995), 'Change points in the series of T4 counts prior to AIDS', *Biometrics* **51**, 236–248.

Lange, N., Carlin, B. & Gelfand, A. E. (1994), 'Hierarchical Bayes models for the progression of HIV infection using longitudinal CD4 T-cell numbers (with discussion)', *Journal of the American Statistical Association* **87**, 615–632.

Maddala, G. S. (1992), *Introduction to Econometrics*, Macmillan, New York.

Malinvaud, E. (1968), *Statistical Methods of Econometrics*, Rand-McNally, Chicago.

R Development Core Team (2011), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org

Reiersol, O. (1950), 'Identifiability of a linear relation between variables which are subject to errors', *Econometrica* **18**, 375–389.

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

Appendix. Conditional Posterior Distributions

1. Conditional posterior distribution of k

$$P(K=k \mid \boldsymbol{\theta}_{-k}, \boldsymbol{x}, \boldsymbol{X}, \boldsymbol{Y}) = \frac{L^*(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{Y})}{\sum\limits_{k=1}^{n} L^*(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{Y})}$$

where $L^*(\boldsymbol{\theta} \mid \boldsymbol{X}, \boldsymbol{Y})$ is given in equation (6).

2. Conditional posterior distribution of α_1

$$\pi\left(\alpha_1 \mid \boldsymbol{\theta}_{\{-\alpha_1\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim N\left(\frac{\sigma^2_{\alpha_1}\sum\limits_{i=1}^{k}(Y_i-\beta_1 x_i)+\alpha_{01}\sigma^2_{e_1}}{k\sigma^2_{\alpha_1}+\sigma^2_{e_1}}, \frac{\sigma^2_{e_1}\sigma^2_{\alpha_1}}{k\sigma^2_{\alpha_1}+\sigma^2_{e_1}}\right)$$

where $\boldsymbol{\theta}_{\{-\theta_i\}}$ is the vector $\boldsymbol{\theta}$ without considering the parameter θ_i.

3. Conditional posterior distribution of α_2

$$\pi\left(\alpha_2 \mid \boldsymbol{\theta}_{\{-\alpha_2\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim N\left(\frac{\sigma^2_{\alpha_2}\sum\limits_{i=k+1}^{n}(Y_i-\beta_2 x_i)+\alpha_{02}\sigma^2_{e_2}}{(n-k)\sigma^2_{\alpha_2}+\sigma^2_{e_2}}, \frac{\sigma^2_{e_2}\sigma^2_{\alpha_2}}{(n-k)\sigma^2_{\alpha_2}+\sigma^2_{e_2}}\right)$$

4. Conditional posterior distribution of β_1

$$\pi\left(\beta_1 \mid \boldsymbol{\theta}_{\{-\beta_1\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim N\left(\frac{\sigma^2_{\beta_1}\sum\limits_{i=1}^{k}(Y_i-\alpha_1)x_i+\beta_{01}\sigma^2_{e_1}}{\sigma^2_{\beta_1}\sum\limits_{i=1}^{k}x_i^2+\sigma^2_{e_1}}, \frac{\sigma^2_{e_1}\sigma^2_{\beta_1}}{\sigma^2_{\beta_1}\sum\limits_{i=1}^{k}x_i^2+\sigma^2_{e_1}}\right)$$

5. Conditional posterior distribution of β_2

$$\pi\left(\beta_2 \mid \boldsymbol{\theta}_{\{-\beta_2\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim N\left(\frac{\sigma^2_{\beta_2}\sum\limits_{i=k+1}^{n}(Y_i-\alpha_2)x_i+\beta_{02}\sigma^2_{e_2}}{\sigma^2_{\beta_2}\sum\limits_{i=k+1}^{n}x_i^2+\sigma^2_{e_2}}, \frac{\sigma^2_{e_2}\sigma^2_{\beta_2}}{\sigma^2_{\beta_2}\sum\limits_{i=k+1}^{n}x_i^2+\sigma^2_{e_2}}\right)$$

6. Conditional posterior distribution of μ_1

$$\pi\left(\mu_1 \mid \boldsymbol{\theta}_{\{-\mu_1\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim N\left(\frac{\sigma^2_{\mu_1}\sum\limits_{i=1}^{k} x_i + \mu_{01}\sigma^2_{x_1}}{k\sigma^2_{\mu_1} + \sigma^2_{x_1}}, \frac{\sigma^2_{x_1}\sigma^2_{\mu_1}}{k\sigma^2_{\mu_1} + \sigma^2_{x_1}}\right)$$

7. Conditional posterior distribution of μ_2

$$\pi\left(\mu_2 \mid \boldsymbol{\theta}_{\{-\mu_2\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim N\left(\frac{\sigma^2_{\mu_2}\sum\limits_{i=k+1}^{n} x_i + \mu_{02}\sigma^2_{x_2}}{(n-k)\sigma^2_{\mu_2} + \sigma^2_{x_2}}, \frac{\sigma^2_{x_2}\sigma^2_{\mu_2}}{(n-k)\sigma^2_{\mu_2} + \sigma^2_{x_2}}\right)$$

8. Conditional posterior distribution of $\sigma^2_{u_1}$

$$\pi\left(\sigma^2_{u_1} \mid \boldsymbol{\theta}_{\{-\sigma^2_{u_1}\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim GI\left(\frac{k}{2} + a_{u_1}, \frac{1}{2}\sum_{i=1}^{k}(X_i - x_i)^2 + b_{u_1}\right)$$

9. Conditional posterior distribution of $\sigma^2_{u_2}$

$$\pi\left(\sigma^2_{u_2} \mid \boldsymbol{\theta}_{\{-\sigma^2_{u_2}\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim GI\left(\frac{(n-k)}{2} + a_{u_2}, \frac{1}{2}\sum_{i=k+1}^{n}(X_i - x_i)^2 + b_{u_2}\right)$$

10. Conditional posterior distribution of $\sigma^2_{e_1}$

$$\pi\left(\sigma^2_{e_1} \mid \boldsymbol{\theta}_{\{-\sigma^2_{e_1}\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim GI\left(\frac{k}{2} + a_{e_1}, \frac{1}{2}\sum_{i=1}^{k}(Y_i - \alpha_1 - \beta_1 x_i)^2 + b_{e_1}\right)$$

11. Conditional posterior distribution of $\sigma^2_{e_2}$

$$\pi\left(\sigma^2_{e_2} \mid \boldsymbol{\theta}_{\{-\sigma^2_{e_2}\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim GI\left(\frac{(n-k)}{2} + a_{e_2}, \frac{1}{2}\sum_{i=k+1}^{n}(Y_i - \alpha_2 - \beta_2 x_i)^2 + b_{e_2}\right)$$

12. Conditional posterior distribution of $\sigma^2_{x_1}$

$$\pi\left(\sigma^2_{x_1} \mid \boldsymbol{\theta}_{\{-\sigma^2_{x_1}\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim GI\left(\frac{k}{2} + a_{x_1}, \frac{1}{2}\sum_{i=1}^{k}(x_i - \mu_1)^2 + b_{x_1}\right)$$

13. Conditional posterior distribution of $\sigma^2_{x_2}$

$$\pi\left(\sigma^2_{x_2} \mid \boldsymbol{\theta}_{\{-\sigma^2_{x_2}\}}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}\right) \sim GI\left(\frac{(n-k)}{2} + a_{x_2}, \frac{1}{2}\sum_{i=k+1}^{n}(x_i - \mu_2)^2 + b_{x_2}\right)$$

14. Conditional posterior distribution of x_i, with $\boldsymbol{x}_{-i} = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$

$$\pi\left(x_i \mid \boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}_{-i}\right) \sim N\left(\mu_{x_i}, Var(\mu_{x_i})\right)$$

where,

$$\mu_{x_i} = \frac{(Y_i - \alpha_1)\beta_1\sigma^2_{u_1}\sigma^2_{x_1} + X_i\sigma^2_{e_1}\sigma^2_{x_1} + \mu_1\sigma^2_{e_1}\sigma^2_{u_1}}{\beta_1^2\sigma^2_{u_1}\sigma^2_{x_1} + \sigma^2_{e_1}\sigma^2_{x_1} + \sigma^2_{e_1}\sigma^2_{u_1}}$$

and

$$Var(\mu_{x_i}) = \frac{\sigma^2_{e_1}\sigma^2_{u_1}\sigma^2_{x_1}}{\beta_1^2\sigma^2_{u_1}\sigma^2_{x_1} + \sigma^2_{e_1}\sigma^2_{x_1} + \sigma^2_{e_1}\sigma^2_{u_1}}$$

15. Conditional posterior distribution of x_i, with $\boldsymbol{x}_{-i} = (x_{k+1}, x_{k+2}, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$

$$\pi\left(x_i \mid \boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{x}_{-i}\right) \sim N\left(\mu_{x_i}, Var(\mu_{x_i})\right)$$

where,

$$\mu_{x_i} = \frac{(Y_i - \alpha_2)\beta_2\sigma^2_{u_2}\sigma^2_{x_2} + X_i\sigma^2_{e_2}\sigma^2_{x_2} + \mu_2\sigma^2_{e_2}\sigma^2_{u_2}}{\beta_2^2\sigma^2_{u_2}\sigma^2_{x_2} + \sigma^2_{e_2}\sigma^2_{x_2} + \sigma^2_{e_2}\sigma^2_{u_2}}$$

and

$$Var(\mu_{x_i}) = \frac{\sigma^2_{e_2}\sigma^2_{u_2}\sigma^2_{x_2}}{\beta_2^2\sigma^2_{u_2}\sigma^2_{x_2} + \sigma^2_{e_2}\sigma^2_{x_2} + \sigma^2_{e_2}\sigma^2_{u_2}}$$