Working Paper 04-13
Statistics and Econometrics Series 05
Febrary 2004

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

VARIANCE CHANGES DETECTION IN MULTIVARIATE TIME SERIES

Pedro Galeano and Daniel Peña*

Abstract

This paper studies the detection of step changes in the variances and in the correlation structure of the components of a vector of time series. Two procedures are considered. The first is based on the likelihood ratio test and the second on cusum statistics. These two procedures are compared in a simulation study and we conclude that the cusum procedure is more powerful. The procedures are illustrated in two examples.

Keywords:

Heteroskedasticity; Step Changes; VARMA models; Likelihood ratio test statistic; Cusum statistic.

* Galeano, Departamento de Estadística, Universidad Carlos III de Madrid, c/Madrid, 126, 28903 Getafe (Madrid), e-mail:pgaleano@est-econ.uc3m.es. Peña, Departamento de Estadística, Universidad Carlos III de Madrid, c/Madrid, 126, 28903 Getafe (Madrid), e-mail:dpena@est-econ.uc3m.es. We acknowledge financial support from BEC2000-0167, MCYT, Spain.

VARIANCE CHANGES DETECTION IN MULTIVARIATE TIME SERIES

Pedro Galeano and Daniel Peña

Universidad Carlos III de Madrid

Abstract: This paper studies the detection of step changes in the variances and in the correlation structure of the components of a vector of time series. Two procedures are considered. The first is based on the likelihood ratio test and the second on cusum statistics. These two procedures are compared in a simulation study and we conclude that the cusum procedure is more powerful. The procedures are illustrated in two examples.

Key words and phrases: Heteroskedasticity; Step Changes; VARMA models; Likelihood ratio test statistic; Cusum statistic.

1. Introduction

The problem of detection of a sudden change in the marginal variance of a univariate time series has been extensively studied. For independent observations see Hinkley (1971), Hsu, Miller and Wichern (1974), Smith (1975), Hsu (1977), Menzefricke (1981) and Booth and Smith (1982). More recently Inclán (1993) studied variance changes in independent observations by means of a Bayesian procedure and Inclán and Tiao (1994) proposed a cumulative sums of squares statistic and an iterative procedure based on this statistic for the detection of several variance changes in Gaussian independent observations. Chen and Gupta (1997) considered an information theoretic approach based on the Bayesian Information Criteria (BIC) for this problem. For dependent observations, Wichern, Miller and Hsu (1976) considered a detection procedure for a variance change at an unknown position in a first order autoregressive model and Abraham and Wei (1984) analyzed the same problem under the Bayesian framework. Baufays and Rasson (1985) proposed an iterative algorithm for changes in autoregressive models. Tsay (1988) studied outliers, level shifts and variance changes in ARIMA models. Park, Lee and Jeong (2000) and Lee and Park (2001) extended the Inclán and Tiao approach to autoregressive and moving average models, respectively.

The case of multivariate sequences has, to the best of our knowledge, not been considered yet. In this article we study the detection of step changes in the variance and in the correlation structure of the components of a vector autoregressive moving average (VARMA) model. Two approaches are introduced and compared. The first is a likelihood ratio approach, which can be seen as a generalization of the univariate procedure due to Tsay (1988). The second is a cusum approach, which can be seen as a generalization of the univariate procedure due to Inclán and Tiao (1994).

The rest of this article is organized as follows. In section 2, we present the model for variance changes and two statistics that can be used for testing for variance changes when the parameters of the VARMA model are known. In section 3, we study two different procedures for detection and estimation of these changes. In section 4, we extend this approach for allowing changes in the correlation structure and present two statistics for testing for such a change. In section 5 the two procedures are compared in a Monte Carlo experiment for different models, sample sizes, number of changes and situation of the change points. Finally, in section 6, we illustrate the procedures by means of two real data examples. We conclude that the procedure based on the cusum statistic has an overall better performance than the one based on the likelihood ratio test.

2. Variance changes in multivariate time series

Let $x_t = (x_{1t}, ..., x_{kt})'$, $t = 1, ..., n$ be a $k-$dimensional vector of time series following a vector ARIMA model, given by

$$\Phi(B)\, x_t = c + \Theta(B)\, a_t, \tag{2.1}$$

where B is the backshift operator, $Bx_t = x_{t-1}$, $\Phi(B) = I - \Phi_1 B - ... - \Phi_p B^p$ and $\Theta(B) = I - \Theta_1 B - ... - \Theta_q B^q$, are $k \times k$ matrix polynomials of finite degrees p and q, c is a $k-$dimensional constant vector, and $a_t = (a_{1t}, ..., a_{kt})'$ is a sequence of independent and identically distributed (iid) Gaussian random vectors with zero mean and positive-definite covariance matrix, Σ. We assume that $\Phi(B)$ and $\Theta(B)$ are left coprime and that all the zeros of the determinants $|\Phi(B)|$ are on or outside the unit circle and those of $|\Theta(B)|$ outside the unit circle. The series

x_t is stationary if $|\Phi(z)| \neq 0$ for all $|z| = 1$ and is unit-root nonstationary if $|\Phi(1)| = 0$. The autoregressive representation of the model (2.1) is,

$$\Pi(B)x_t = c_\Pi + a_t, \tag{2.2}$$

where $\Pi(B) = \Theta(B)^{-1}\Phi(B) = I - \sum_{i=1}^{\infty}\Pi_i B^i$, and $c_\Pi = \Theta(1)^{-1} c$ is a vector of constants. In the stationary case we also have moving-average representation,

$$x_t = c_\Psi + \Psi(B)\, a_t, \tag{2.3}$$

where, $\Psi(B) = \Phi(B)^{-1}\Theta(B) = I + \sum_{i=1}^{\infty}\Psi_i B^i$, and $\Phi(1)\, c_\Psi = c$. We can also use this representation in the nonstationary case where now $\Psi(B)$ is defined by $\Phi(B)\Psi(B) = \Theta(B)$.

We generalize the variance change model in Tsay (1988) in a direct manner. Suppose that instead of observing x_t we observe a time series $y_t = (y_{1t}, ..., y_{kt})'$, defined as follows. Let $S_t^{(h)}$ be a step function such that $S_t^{(h)} = 0$, $t < h$ and $S_t^{(h)} = 1$, $t \geq h$. Let W a constant diagonal matrix of size $k \times k$ denoting the impact of the variance change. Then, we assume that the innovations affecting the series, e_t, is not a sequence of iid $N_k(0, \Sigma)$ variables because it has a change in the variance of the components at same point $t = h$, given by

$$e_t = a_t + W S_t^{(h)} a_t, \tag{2.4}$$

and, therefore, the observed vector time series $y_t = (y_{1t}, ..., y_{kt})'$ can be written as

$$\Phi(B)\, y_t = c + \Theta(B)\,(a_t + W S_t^{(h)} a_t),$$

and, by using (2.1), the relation between the observed series, y_t, and the unobserved vector ARIMA time series, x_t, is given by

$$y_t = x_t + \Psi(B)\, W S_t^{(h)} a_t. \tag{2.5}$$

The variance of e_t changes from Σ to $\Omega = (I + W)\,\Sigma\,(I + W)$ at the time point $t = h$. Without loss of generality we assume that $(I + W)$ is a positive defined matrix, so that the matrix W is well identified. For that, the spectral decompositions of the matrices Σ and Ω are given by $\Sigma = D_\Sigma R_\Sigma D_\Sigma$ and $\Omega = D_\Omega R_\Omega D_\Omega$ respectively, where R_Σ and R_Ω are the correlation matrices of Σ and

Ω which are assumed to be equal, and D_Σ and D_Ω are diagonal matrices whose elements are the standard deviations of each component. Then, by taking

$$W = D_\Omega D_\Sigma^{-1} - I, \tag{2.6}$$

we obtain that $\Omega = (I+W)\,\Sigma\,(I+W)$, and the matrix W is unique. We note that the variance change may affect one or several components and the elements different from 0 of W indicate the components with changing variance.

To test the significance of a variance change at $t = h$, suppose that the parameters of the ARIMA model are known and using them we compute the residuals:

$$e_t = y_t - \sum_{i=1}^{p} \Phi_i y_{t-i} - c + \sum_{j=1}^{q} \Theta_j e_{t-j}. \tag{2.7}$$

We want to test the hypothesis that these residuals are iid homoskedastic, versus the alternative hypothesis that they are heteroskedastic. Thus, we consider the null hypothesis $H_0 : W = 0$ versus the alternative hypothesis $H_1 : W \neq 0$. The most usual method for testing the homogeneity of the covariance matrices of two Gaussian populations is the likelihood ratio (LR) test, which is asymptotically the most powerful test. Let us define the three values $s\,(i) = \sum_{t=1}^{n} \left(e_{it}^2\right)/n$, $s_1^{h-1}\,(i) = \sum_{t=1}^{h-1} \left(e_{it}^2\right)/\,(h-1)$ and $s_h^n\,(i) = \sum_{t=h}^{n} \left(e_{it}^2\right)/\,(n-h+1)$. The likelihood ratio statistic of the residuals in (2.7) for a variance change after the time point $t = h$ is given by

$$LR_h = \log \frac{\left(s\,(1)\cdots s\,(k)\right)^n}{\left(s_1^{h-1}\,(1)\cdots s_1^{h-1}\,(k)\right)^{h-1}\left(s_h^n\,(1)\cdots s_h^n\,(k)\right)^{n-h+1}}, \tag{2.8}$$

and, under the null hypothesis of no variance change and assuming that the model is known, the LR_h statistic has an asymptotic chi-squared distribution with k degrees of freedom.

An alternative test statistic can be built as follows. Under the null hypothesis of homoskedasticity, the covariance matrix of e_t can be written as $\Sigma = D_\Sigma R_\Sigma D_\Sigma$. We define $b_t = D_\Sigma^{-1} e_t$ with $Cov\,(b_t) = R_\Sigma$. The principal components of the series b_t are given by $c_t = U_\Sigma b_t$, where U_Σ is the matrix whose columns are the eigenvectors of the matrix R_Σ, and $Cov\,(c_t) = \Lambda$, which is a diagonal matrix. The components of c_t are uncorrelated with variances equal to the diagonal elements of the matrix Λ. Let $A_m = \sum_{t=1}^{m} c_t' c_t$ be the multivariate cumulative sum of squares of the sequence $\{c_1, \ldots, c_m\}$ where m is any given value $1 \leq m \leq n$. Let,

$$B_m = \frac{A_m}{A_n} - \frac{m}{n}, \qquad m = 1, \ldots, n \tag{2.9}$$

where $B_1 = B_n = 0$, be the centered and normalized cumulative sum of squares of the sequence c_t. We study the asymptotic behavior of the statistic (2.9) under the hypothesis of homoskedasticity.

Lemma 1 *Under the null hypothesis of no change in the covariance matrix of the sequence* $\{e_1, \ldots, e_n\}$ *in* (2.7)*, for a given value* $t = m$*,*

$$E[B_m] = o\left(n^{-1}\right).$$

Proof. The second order Taylor expansion of the ratio A_m/A_n about the value $(E[A_m], E[A_n])$ is:

$$E\left[\frac{A_m}{A_n}\right] = \frac{E[A_m]}{E[A_n]} - \frac{E[A_m A_n]}{E[A_n]^2} + \frac{E[A_m]\, E\left[A_n^2\right]}{E[A_n]^3} + o\left(n^{-1}\right).$$

Taking into account that $tr(\Lambda) = k$, where tr stands for trace, as $E[A_m] = mk$, and

$$E[A_m A_n] = E\left[\left(\sum_{t=1}^{m} c_t' c_t\right)\left(\sum_{l=1}^{n} c_l' c_l\right)\right] =$$

$$= \sum_{t=1}^{m} \sum_{l=1}^{n} E\left[\left(c_t' c_t\right)\left(c_l' c_l\right)\right] = m\left[2tr\left(\Lambda^2\right) + k^2\right],$$

the ratio $E[A_m/A_n]$ can be written as:

$$E\left[\frac{A_m}{A_n}\right] = \frac{m}{n} - \frac{m\left[2tr\left(\Lambda^2\right) + k^2\right]}{n^2 k^2} +$$

$$\frac{n\left[2tr\left(\Lambda^2\right) + k^2\right] mk}{n^3 k^3} + o\left(n^{-1}\right) = \frac{m}{n} + o\left(n^{-1}\right),$$

and therefore, $E[B_m] = o\left(n^{-1}\right)$. ■

Consequently, the mean of the statistic B_m is asymptotically 0 for every m. Let us study the asymptotic distribution of the statistic B_m under the hypothesis of no change in the covariance matrix for $t = 1, \ldots, n$. Let M a Brownian motion process verifying $E[M_r] = 0$, and $E[M_r M_s] = s$, where $0 \leq s < r \leq 1$. Let M^0

denote a Brownian bridge given by $M_r^0 = M_r - rM_1$, verifying $E\left[M_r^0\right] = 0$, $E\left[M_r^0 M_s^0\right] = s(1-r)$, $0 \leq s < r \leq 1$, and $M_0^0 = M_1^0 = 0$, with probability 1. The asymptotic distribution of the statistic B_m is obtained in the following theorem.

Theorem 2 *Let $\{e_1, \ldots, e_n\}$ be a sequence of independent identically distributed Gaussian random variables with zero mean and common covariance matrix Σ. Let $b_t = D_\Sigma^{-1} e_t$ with $Cov(b_t) = \Lambda$, where D_Σ is a diagonal matrix with elements the square root of the variances of the components of e_t and let c_t be the principal components of the series b_t. Let $B_m = A_m/A_n - m/n$, where $A_m = \sum_{t=1}^m c_t' c_t$. Therefore, $B_m^* = \sqrt{\frac{n}{2}} \frac{k}{\sqrt{tr(\Lambda^2)}} B_m \xrightarrow{D} M^0$.*

Proof. Let $\xi_m = c_m' c_m - k$, such that $E[\xi_m] = 0$, and,

$$\sigma^2 = E\left[\xi_m^2\right] = E\left[\left(c_m' c_m\right)^2\right] - k^2 = 2tr\left(\Lambda^2\right).$$

Let $X_n(r) = \frac{1}{\sigma\sqrt{n}} S_{[nr]} + (nr - [nr]) \frac{1}{\sigma\sqrt{n}} \xi_{[nr]+1}$, where $S_n = \sum_{i=1}^n \xi_i$. By Donsker's Theorem, $X_n \xrightarrow{D} M$, so $\{X_n(r) - rX_n(1)\} \xrightarrow{D} M_0$, see Billingsley (1968, Th. 10.1 and Th. 5.1). Let $nr = m$, $m = 1, \ldots, n$. Then,

$$X_n(r) - rX_n(1) = \frac{1}{\sigma\sqrt{n}} S_{[nr]} + (nr - [nr]) \frac{1}{\sigma\sqrt{n}} \xi_{[nr]+1} - r\frac{1}{\sigma\sqrt{n}} S_{[n]} =$$
$$= \frac{1}{\sigma\sqrt{n}} \left(S_m - \frac{m}{n} S_n\right).$$

As $S_n = \sum_{t=1}^n c_t' c_t - nk$ and $S_m = \sum_{t=1}^m c_t' c_t - mk$, we get,

$$S_m - \frac{m}{n} S_n = \left(\sum_{t=1}^n c_t' c_t\right) B_m.$$

Then, when $n \to \infty$,

$$\frac{1}{\sigma\sqrt{n}} \left(\sum_{t=1}^n c_t' c_t\right) B_m \underset{D}{\longrightarrow} M^0.$$

Therefore, as $\frac{1}{n} \sum_{t=1}^n c_t' c_t \to k$,

$$\frac{1}{\sigma\sqrt{n}} \left(\sum_{t=1}^n c_t' c_t\right) B_m = \sqrt{\frac{n}{2}} \frac{k}{\sqrt{tr\left(\Lambda^2\right)}} \frac{\frac{1}{n}\sum\limits_{t=1}^n c_t' c_t}{k} B_m \underset{D}{\longrightarrow} M^0$$

that proves the stated result. ■

We have proved that the asymptotic distribution of the statistic B_m^* under the hypothesis of no change in the covariance matrix is a Brownian Bridge. Thus, we may use the statistic B_{h-1}^* to test the presence of a change in the covariance matrix at $t = h$, and the asymptotic critical value of the distribution of a Brownian Bridge. The statistic B_m^* depends on $tr\left(\Lambda^2\right)$, which in practice is unknown. Let $\Lambda(i,i)$ be the i diagonal element of the matrix Λ. Under the assumption of no change, we can estimate $\Lambda(i,i)$ by means of $\widehat{\Lambda}(i,i) = \sum_{t=1}^{n}\left(c_{it}^2\right)/n$, which is a consistent estimator of $\Lambda(i,i)$. If $\widehat{\Lambda}_1^m(i,i) = \sum_{t=1}^{m}\left(c_{it}^2\right)/m$ and taking into account that $tr\left(\widehat{\Lambda}\right) = k$, then, we define the statistic, C_m, as follows,

$$C_m = \sqrt{\frac{1}{2n}}\frac{mk}{\sqrt{tr\left(\widehat{\Lambda}^2\right)}}\left(\frac{\widehat{\Lambda}_1^m(1,1) + \cdots + \widehat{\Lambda}_1^m(k,k)}{k} - 1\right), \tag{2.10}$$

Under the hypothesis of no change in the variances, as $\widehat{\Lambda}(i,i)$ is a consistent estimator of $\Lambda(i,i)$, the statistics B_m^* and C_m have the same asymptotic distribution.

The impact of a variance change is estimated as follows. Let $\Omega(i,i)$, $\Sigma(i,i)$ and $W(i,i)$ be the i diagonal elements of the matrices Ω, Σ and W, respectively. Then,

$$(1 + W(i,i))^2 = \frac{\Omega(i,i)}{\Sigma(i,i)}, \qquad i = 1,\ldots,k$$

and as the maximum likelihood estimates of Σ and Ω are given by $\widehat{\Sigma} = S_1^{h-1}$ and $\widehat{\Omega} = S_h^n$, we estimate $\widehat{W}(i,i)$ by:

$$\left(1 + \widehat{W}(i,i)\right)^2 = \frac{S_h^n(i,i)}{S_1^{h-1}(i,i)}, \qquad i = 1,\ldots,k$$

where $S_1^{h-1}(i,i)$ and $S_h^n(i,i)$ are the i elements of the diagonals of the matrices S_1^{h-1} and S_h^n, respectively. Under the null hypothesis of no variance change, $\left(1 + \widehat{W}(i,i)\right)^2$ is distributed as a F distribution with $(n-h, h-2)$ degrees of freedom. Therefore, we can test the null hypothesis of $W(i,i) = 0$ against the alternative of being different by means of the F distribution. As $1 + \widehat{W}(i,i)$ should be larger than 1, we can obtain the final estimate of $\widehat{W}(i,i)$ as:

$$\widehat{W}(i,i) = \sqrt{\frac{S_h^n(i,i)}{S_1^{h-1}(i,i)}} - 1, \qquad i = 1,\ldots,k \tag{2.11}$$

A confidence interval for $W(i,i)$ for a significant level α is given by:

$$1-\alpha = P\left(\frac{1+\widehat{W}(i,i)}{\sqrt{F^{1-\alpha/2}_{(n-h,h-2)}}} - 1 \le W(i,i) \le \frac{1+\widehat{W}(i,i)}{\sqrt{F^{\alpha/2}_{(n-h,h-2)}}} - 1\right)$$

where $F^{\alpha/2}_{(n-h,h-2)}$ and $F^{1-\alpha/2}_{(n-h,h-2)}$ are the critical values of the F distribution with $(n-h, h-2)$ degrees of freedom for the significance levels $\alpha/2$ and $1-\alpha/2$, respectively.

3. Procedures for variance changes detection

A series can be affected by several variance changes. In this case, we observe a time series $y_t = (y_{1t}, ..., y_{kt})'$, defined as follows:

$$y_t = x_t + \Psi(B)(I + W_r S_t^{(h_r)}) \cdots (I + W_1 S_t^{(h_1)}) a_t,$$

where $\{h_1, \ldots, h_r\}$ are the time of r change points and $W_1, \ldots, W_r$ are $k \times k$ diagonal matrices denoting the impact of the r changes. Assuming that the parameters are known, the filtered series of residuals is given by:

$$e_t = (I + W_r S_t^{(h_r)}) \cdots (I + W_1 S_t^{(h_1)}) a_t,$$

and the residual covariance matrix of e_t changes from Σ to $(I+W_1)\Sigma(I+W_1)$ at $t = h_1$, and to $(I+W_2)(I+W_1)\Sigma(I+W_1)(I+W_2)$ at $t = h_2, \ldots$

In practice, the parameters of the VARMA model, the number, location and the sizes of the variance changes are unknown. Let $\widehat{LR}_t$ and $\widehat{C}_t$ be the statistics (2.8) and (2.10) respectively computed using the estimated residuals which are obtained by (2.7). We define the maximum of these statistics in the sample as,

$$\Lambda_{\max}\left(h^{LR}_{\max}\right) = \max\left\{\left|\widehat{LR}_t\right|, 1 \le t \le n\right\}, \ \Gamma_{\max}\left(h^{C}_{\max}\right) = \max\left\{\left|\widehat{C}_t\right|, 1 \le t \le n\right\} \tag{3.1}$$

where $h^{LR}_{\max}$ and $h^{C}_{\max} + 1$ are the estimates of the time of the change using the LR test or the cusum statistic, respectively. The distribution of $\Lambda_{\max}$ in (3.1) is intractable and critical values should be obtained by simulation. The distribution of $\Gamma_{\max}$ in (3.1) is asymptotically the distribution of $\sup\left\{\left|M^0_r\right| : 0 \le r \le 1\right\}$ which is given by (see, Billingsley, pg. 85, 1968),

$$P\left\{\sup\left|M^0_r\right| \le a : 0 \le r \le 1\right\} = 1 + 2\sum_{i=1}^{\infty}(-1)^i \exp\left(-2i^2a^2\right)$$

and critical values can be obtained from this distribution. If several changes have occurred in the series, we propose two iterative procedures to detect them and estimate its impacts based on the statistics $\Lambda_{\max}$ and $\Gamma_{\max}$.

To motivate the proposed procedures let us consider a bivariate series from a first order vector autoregressive model. We consider three different situations which are illustrated in Figure 1. The three columns in this matrix of plots represents three different generating processes. The first column corresponds to the case of no variance changes. The second column corresponds to the case of a single change in the covariance matrix at $t = 250$, where the innovation covariance matrix goes from I to the matrix $3 \times I$. The third column corresponds to the case of two changes at $t = 166$, where the innovation covariance matrix goes from I to $3 \times I$, and $t = 333$, where the innovation covariance matrix goes back to I. The rows represent the two components of the bivariate time series and the two statistics introduced in the previous section. The first (second) rows in Figure 1 shows a sample of 500 observations of the first (second) component of this bivariate series, and the third and fourth rows show the LR statistic (2.8) and the cusum statistic (2.10) respectively computed with the bivariate time series in the same column. In the first column in Figure 1, no variance change case, the two statistics plotted in the third and fourth row are inside the two straight lines computed as explain next as for the 95% confidence interval of the distributions of $\Lambda_{\max}$ and $\Gamma_{\max}$. In the second column, a single change at $t = 250$, the maximum of both statistics in absolute value is around $t = 250$, and the maximum is larger than the critical value, so the hypothesis of no change is rejected. In the third column, two variance changes at $t = 166$ and $t = 333$, these changes appear as two significant extremes around the times of the changes $t = 166$ and $t = 333$.

3.1 LR procedure

1. Assuming no variance changes, a vector ARIMA model is specified for the observed series y_t. The maximum likelihood estimates of the model are obtained as well as the filtered series of residuals.

2. Compute the statistics LR_h, $h = d + 1, ..., n - d$, for a given value of an integer d, using the residuals obtained in Step 1. The number $d = k + m$

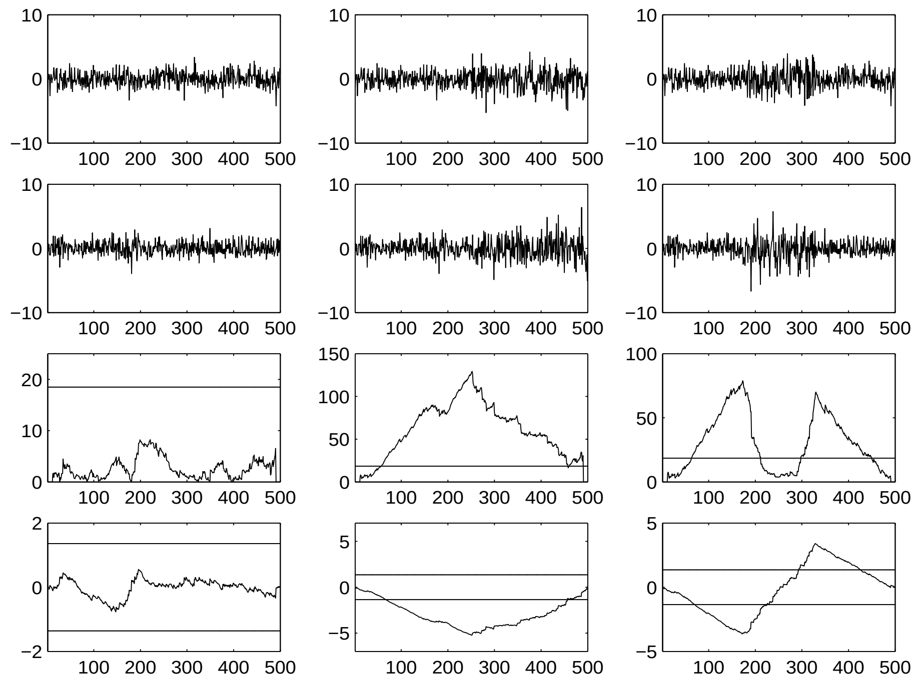

Figure 3.1: Bivariate series and statistics for variance change detection.

is a positive integer denoting the minimum number of residuals needed to estimate the covariance matrix. The value m can be fixed by the user, and in the examples and simulations we have taken $m = 10$. With them, the statistic $\Lambda_{\max}\left(h_{\max}^{LR}\right)$ in (3.1) is obtained.

3. Compare $\Lambda_{\max}\left(h_{\max}^{LR}\right)$ with a specified critical value C for a given critical level. If $\Lambda_{\max}\left(h_{\max}^{LR}\right) < C$, it is concluded that there is not a significant variance change and the procedure ends. If $\Lambda_{\max}\left(h_{\max}^{LR}\right) \geq C$, it is assumed that a variance change is detected at time $t = h_{\max}^{LR}$.

4. The matrix $\widehat{W}$ is estimated by (2.11) and a modified residual series is computed as follows:

$$e_t^* = \begin{cases} \widehat{e}_t & t < h_{\max}^{LR} \\ \left(I + \widehat{W}\right)^{-1} \widehat{e}_t & t \geq h_{\max}^{LR} \end{cases}$$

and, with this residual series, a corrected time series is defined by

$$y_t^* = \begin{cases} y_t & t < h_{\max}^{LR} \\ \widehat{c} + \widehat{\Phi}_1 y_{t-1}^* + \ldots + \widehat{\Phi}_p y_{t-p}^* + e_t^* - \widehat{\Theta}_1 e_{t-1}^* - \ldots - \widehat{\Theta}_q e_{t-q}^* & t \geq h_{\max}^{LR} \end{cases}$$

where the polynomials $\widehat{\Phi}(B)$ and $\widehat{\Theta}(B)$ are the maximum likelihood estimates of the parameters. Then go back to Step 1 considering y_t^* as the observed process.

5. When no more variance changes are detected, the parameters of the series and all the variance changes detected in the previous steps are estimated jointly, using the model

$$\Phi(B)\, y_t = c + \Theta(B)\,(I + W_r S_t^{(h_r)}) \cdots (I + W_1 S_t^{(h_1)}) a_t, \qquad (3.1)$$

This joint estimation is carried out in two steps. First, estimate the parameters assuming no variance changes and then estimate the matrices W_i. After that, correct the series, and repeat these two steps until convergence.

3.2 Cusum procedure

The following procedure is a generalization to the one proposed by Inclán and Tiao (1994). The algorithm is based on successive divisions of the series into two pieces when a change is detected and proceeds as follows:

1. Assuming no variance changes, a vector ARIMA model is specified for the observed series y_t. The maximum likelihood estimates of the model are obtained as well as the series of residuals. Then, obtain the principal components of the residual series, c_t, as in Section 2. Let $t_1 = 1$.

2. Obtain $\Gamma_{\max}\left(h_{\max}^C\right)$ for c_t in (3.1) for $t = 1, \ldots, n$. If $\Gamma_{\max}\left(h_{\max}^C\right) > C$, where C is the asymptotic critical value for a critical level, go to step 3. If $\Gamma_{\max}\left(h_{\max}^C\right) < C$, it is assumed that there is not a variance change in the series and the procedure ends.

3. Step 3 has three substeps:

   (a) Obtain $\Gamma_{\max}\left(h_{\max}^C\right)$ for $t = 1, \ldots, t_2$, where $t_2 = h_{\max}^C$. If $\Gamma_{\max}\left(h_{\max}^C\right) > C$, redefine $t_2 = h_{\max}^C$ and repeat Step 3(a) until $\Gamma_{\max}\left(h_{\max}^C\right) < C$.

When this happens, define $h_{first} = t_2$ where t_2 is the last value such that $\Gamma_{\max}\left(h^C_{\max}\right) > C$.

(b) Repeat a similar search in the interval $t_2 \leq t \leq n$, where t_2 is the point $h^C_{\max}$ obtained in Step 2. For that, define $t_1 = h^C_{\max} + 1$, where $h^C_{\max} = \arg\max\{C_t : t = t_1, \ldots, n\}$ and repeat it until $\Gamma_{\max}\left(h^C_{\max}\right) < C$. Define $h_{last} = t_1 - 1$, where t_1 is the last value such that $\Gamma_{\max}\left(h^C_{\max}\right) > C$.

(c) If $|h_{last} - h_{first}| < d$, there is just one change point and the algorithm ends here. Otherwise, keep both values as possible change points and repeat Steps 2 and 3 for $t_1 = h_{first}$ and $n = h_{last}$, until no more possible change points are detected. Then, go to step 4.

4. Define a vector $\ell = (\ell_1, \ldots, \ell_s)$ where $\ell_1 = 1$, $\ell_s = n$ and $\ell_2, \ldots, \ell_{s-1}$ are the points detected in Steps 2 and 3 in increasing order. Obtain the statistic C_t in each one of the intervals (ℓ_i, ℓ_{i+2}) and check if its maximum is still significant. If it is not, eliminate the corresponding point. Repeat Step 4 until the number of possible change points does not change, and the points found in previous iterations do not differ from those in the last one. The vector $(\ell_2 + 1, \ldots, \ell_{s-1} + 1)$ are the points of variance change.

5. Finally, estimate the parameters of the series and the variance changes detected in the previous steps jointly by using (3.1).

Some comments with regards to these algorithms are in order. First, the critical values in the LR algorithm have to be obtained by simulation as we will study in section 5, while the critical values used in the cusum procedure are the asymptotic critical values of the maximum of the absolute value of a Brownian Bridge. Second, in both algorithms we require a minimum distance between variance changes larger than d, so that the covariance matrix can be estimated. If several changes were found in an interval smaller than d, these changes will be considered as outliers and estimated by the procedure proposed by Tsay et al (2000). Third, the last step in the LR procedure is needed for avoiding bias in the size of the estimated variance changes. Note that in Step 4, the size of the variance change is estimated after detecting it. Thus, if there are two variance changes the impact of the first change detected is estimated without taking into account the second one. Therefore, a joint estimation is needed taking into account all the changes detected by the procedure.

4. A generalization for allowing variance and correlation changes

Suppose now that instead of observing x_t we observe a time series $y_t = (y_{1t}, ..., y_{kt})'$, defined as follows. Let W a constant lower triangular matrix of size $k \times k$. Then, we assume that the innovations affecting the series, e_t, has a change in the variance of the components at same point $t = h$ given by (2.4) with W lower triangular and therefore the observed vector time series $y_t = (y_{1t}, ..., y_{kt})'$ can be written as in (2.5). The variance of e_t at the time point $t = h$ changes from Σ to $\Omega = (I + W)\,\Sigma\,(I + W)'$. Without loss of generality, it is assumed that $(I + W)$ is a positive defined matrix so that the matrix W is well identified. For that, let $\Sigma = L_\Sigma L_\Sigma'$ and $\Omega = L_\Omega L_\Omega'$ be the Cholesky decompositions of Σ and Ω, respectively. Then, by taking,

$$W = L_\Omega L_\Sigma^{-1} - I \tag{4.1}$$

with W lower triangular, we obtain that $\Omega = (I + W)\,\Sigma\,(I + W)'$, and as the Cholesky decomposition of a matrix is unique, the matrix W is also unique.

As in the previous case, to test the significance of a change at $t = h$, suppose that the parameters of the ARIMA model are known and using then we compute the residuals as in (2.7). We consider the null hypothesis $H_0 : W = 0$ versus the alternative hypothesis $H_1 : W \neq 0$. Let us define the three matrices $S = \sum_{t=1}^{n} (e_t e_t') / n$, $S_1^{h-1} = \sum_{t=1}^{h-1} (e_t e_t') / (h-1)$ and $S_h^n = \sum_{t=h}^{n} (e_t e_t') / (n-h+1)$. The likelihood ratio statistic of the residuals in (2.7) for a variance change after the time point $t = h$ is given by

$$LR_h = \log \frac{|S|^n}{\left|S_1^{h-1}\right|^{h-1} \left|S_h^n\right|^{n-h+1}} \tag{4.2}$$

and under the null hypothesis of no variance change and assuming that the model is known, the LR_h statistic has an asymptotic chi-squared distribution with $\frac{1}{2}k\,(k+1)$ degrees of freedom.

An alternative cusum test statistic can be built as follows. Let $A_m = \sum_{t=1}^{m} e_t' \Sigma^{-1} e_t$ be the multivariate cumulative sum of squares of $\{e_1, \ldots, e_m\}$

where m is any given value $1 \leq m \leq n$. Let,

$$B_m = \frac{A_m}{A_n} - \frac{m}{n}, \qquad m = 1, \ldots, n \tag{4.3}$$

where $B_1 = B_n = 0$, be the centered and normalized cumulative sum of squares of the sequence e_t. The asymptotic distribution of the statistic (4.3) under the hypothesis of homoskedasticity can be obtained similarly to the case of changing variance. We state the following Lemma and Theorem which proofs are similar to the first case and are not shown here.

Lemma 3 *Under the null hypothesis of no change in the covariance matrix of the sequence $\{e_1, \ldots, e_n\}$ in (2.7), for a given value $t = m$,*

$$E\left[B_m\right] = o\left(n^{-1}\right).$$

The asymptotic distribution of the statistic B_m is obtained in the following theorem.

Theorem 4 *Let $\{e_1, \ldots, e_n\}$ be a sequence of independent identically distributed Gaussian random variables with zero mean and common covariance matrix Σ. Let $B_m = A_m/A_n - m/n$, where $A_m = \sum_{t=1}^{m} e_t'\Sigma^{-1}e_t$. Therefore, the statistic $B_m^* = \sqrt{nk/2}B_m \xrightarrow{D} M^0$.*

The asymptotic distribution of the statistic $B_m^* = (nk/2)^{\frac{1}{2}} B_m$ under the hypothesis of no change in the covariance matrix is a Brownian Bridge. Thus, we may use the statistic B_{h-1}^* to test the presence of a change in the covariance matrix at $t = h$, and the asymptotic critical value of the distribution of a Brownian Bridge. The statistic B_m^* depends on the covariance matrix Σ, which in practice can be estimated consistently by means of $S = \sum_{t=1}^{n}\left(e_t e_t'\right)/n$. Then, we define the statistic C_m as follows,

$$C_m = \sqrt{\frac{k}{2n}} m \left(\frac{trace\left(S^{-1}S_1^m\right)}{k} - 1 \right), \tag{4.4}$$

where $S_1^m = \sum_{t=1}^{m}\left(e_t e_t'\right)/m$. Under the hypothesis of no change, the statistics B_m^* and C_m have the same asymptotic distribution.

The impact of a covariance change is estimated using (4.1) by means of

$$\widehat{W} = L_{S_h^n} L_{S_1^{h-1}}^{-1} - I \tag{4.5}$$

Table 5.1: Models for the simulation study.

$k=2$		$k=3$	
Φ	Σ	Φ	Σ
$\begin{vmatrix} 0.6 & 0.2 \\ 0.2 & 0.4 \end{vmatrix}$	$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$	$\begin{Vmatrix} 0.6 & 0.2 & 0 \\ 0.2 & 0.4 & 0 \\ 0.6 & 0.2 & 0.5 \end{Vmatrix}$	$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$

When several changes are present, the LR and cusum procedures for variance changes are directly applied to the case of covariance changes where the matrix W is estimated with (4.5). The maximum statistics (3.1) are defined in the same way for the statistics (4.2) and (4.4).

5. Simulation study

The simulations in this section and the analysis of real datasets in the next section have been done using MATLAB (developed by The MathWorks, Inc.) by means of various routines written by the authors which can be downloaded from http://halweb.uc3m.es/esp/Personal/personas/dpena/esp/perso.html. We first obtain critical values for the statistic $\Lambda_{\max}$ in (3.1) for W diagonal and W lower triangular by simulating from the vector AR(1) models in Table 5.1, where $k = 2$, 3 and sample sizes $n = 100$, 200, 500 and 1000. For each model and sample size, we generate 10000 realizations and estimate a vector AR(1) model, obtain the residuals, $\widehat{e}_t$, and compute the statistics (3.1). Table 5.2 provides some quantiles of the distribution of $\Lambda_{\max}$ for both models and different sample sizes under the null hypothesis of no variance change in the sample. Note that the quantiles depend on the time series dimension. The asymptotic distribution of the statistic $\Gamma_{\max}$ is known but we also study the finite sample behavior of the quantiles of this statistic and Table 5.2 provide these quantiles. As we can see, the finite sample quantiles are always smaller than the asymptotic ones implying that the use of the asymptotic quantile is a conservative decision and therefore, the type I error will not increase. Note also that the quantiles do not depend on k.

First, we consider the case of variance changes and make a simulation study

Table 5.2: Empirical quantiles of the $\Lambda_{\max}$ and $\Gamma_{\max}$ statistics based on 10000 realizations.

W diagonal										
	Probability-LR					**Probability-CUSUM**				
k=2	**50%**	**90%**	**95%**	**97.5%**	**99%**	**50%**	**90%**	**95%**	**97.5%**	**99%**
n=100	9.02	13.94	16.08	17.96	20.02	0.73	1.12	1.27	1.41	1.55
n=200	9.74	15.10	16.76	18.80	21.27	0.76	1.13	1.28	1.42	1.56
n=500	10.87	15.91	17.82	19.92	21.70	0.80	1.18	1.31	1.46	1.61
n=1000	11.24	16.65	18.55	20.27	22.65	0.81	1.19	1.33	1.47	1.62
n=∞	-	-	-	-	-	0.82	1.22	1.35	1.48	1.62
k=3	**50%**	**90%**	**95%**	**97.5%**	**99%**	**50%**	**90%**	**95%**	**97.5%**	**99%**
n=100	11.97	17.76	19.82	21.55	24.68	0.72	1.08	1.22	1.32	1.45
n=200	13.03	18.89	20.79	22.44	25.04	0.77	1.17	1.31	1.41	1.56
n=500	14.34	20.10	22.16	24.13	26.06	0.80	1.19	1.32	1.43	1.57
n=1000	14.71	20.47	22.40	24.40	26.68	0.81	1.20	1.34	1.45	1.59
n=∞	-	-	-	-	-	0.82	1.22	1.35	1.48	1.62
W lower triangular										
	Probability-LR					**Probability-CUSUM**				
k=2	**50%**	**90%**	**95%**	**97.5%**	**99%**	**50%**	**90%**	**95%**	**97.5%**	**99%**
n=100	11.07	16.87	18.93	20.67	23.66	0.75	1.16	1.30	1.39	1.55
n=200	11.73	17.29	19.34	21.57	23.97	0.77	1.17	1.31	1.41	1.57
n=500	12.69	18.36	20.60	22.31	25.07	0.79	1.18	1.32	1.42	1.58
n=1000	13.16	19.01	21.44	23.56	26.52	0.80	1.19	1.33	1.44	1.60
n=∞	-	-	-	-	-	0.82	1.22	1.35	1.48	1.62
k=3	**50%**	**90%**	**95%**	**97.5%**	**99%**	**50%**	**90%**	**95%**	**97.5%**	**99%**
n=100	17.37	24.66	27.30	29.15	32.32	0.75	1.13	1.26	1.38	1.51
n=200	18.24	24.94	27.38	29.18	32.68	0.76	1.14	1.29	1.44	1.54
n=500	19.26	25.88	28.47	30.65	34.78	0.78	1.17	1.33	1.44	1.56
n=1000	19.96	26.86	28.86	30.80	34.95	0.80	1.20	1.34	1.46	1.60
n=∞	-	-	-	-	-	0.82	1.22	1.35	1.48	1.62

in order to study the size and power of the two procedures. For that, we consider the models in Table 5.1 for $n = 100$, 200 and 500. For the case of one variance change, for each n, we consider three locations of the change point, $h = [0.25n]$, $[0.50n]$ and $[0.75n]$. The changes are introduced by transforming the original covariance matrix, $\Sigma = I$, into $\Omega = (I + W)(I + W)$, where W is a diagonal matrix. We consider three possible matrices of the form $W = \alpha I$, where α takes three possible values: $\alpha = 0$, in the case of no variance change, $\alpha = \sqrt{2} - 1$, so that the covariance matrix is multiplied by 2, and $\alpha = \sqrt{3} - 1$, and the covariance matrix is multiplied by 3. For each case, we generate 5000 realizations. Then, we apply the two procedures with the 95% critical values from Table 5.2. The results are shown in Tables 5.3 and 5.4, where columns 4 to 6 and 9 to 11 report the number of variance changes detected by the algorithms and columns 7, 8, 12 and 13 show the median and the mean absolute deviation of the estimates of the change points for each case. The cases with $\alpha = 0$ indicate the type I error of the procedures, which is around 5 % in all the sample sizes considered. From these two tables we conclude that when $n = 100$ the cusum procedure appears to work better than the LR procedure. For $n = 200$ and for a small change, $\alpha = 2$, the cusum procedure is slightly better than the LR one, but for a larger change $\alpha = 3$, the LR seems to be slightly more powerful. The estimates of the time of the change, h, are similar for both procedures.

For two change points, we consider the same sample sizes and the change points at $(h_1, h_2) = ([0.33n], [0.66n])$. Each change point is associated with two matrices, Ω_1 and Ω_2, which give the residual covariance matrices after each change. Six combinations are considered. For each case, we generate 5000 realizations with the corresponding changes. Then, we apply the two procedures with the 95% critical values from Table 5.2. The results are shown in Tables 5.5 and 5.6. Columns 6 to 9 in these tables are the number of variance changes detected by the algorithms, and columns 10 to 13 show the median and the mean absolute deviation of the estimates of the change points. For two change points, the advantage of the cusum procedure over the LR one is clearer. Note that, first, the detection frequency of two change points are larger for the cusum procedure in almost all the cases, and, second, the LR procedure suffers of an overestimation of the number of changes in some situations. In general, except when $k = 2$,

Table 5.3: Results for model 1 and one variance change.

			LR procedure					**Cusum Procedure**				
Ω	**n**	**h**	**frequency**			$\widehat{\mathbf{h}}$		**frequency**			$\widehat{\mathbf{h}}$	
			0	1	≥ 2	Med.	Mad	0	1	≥ 2	Med.	Mad
I	100	—	**95.8**	4.2	0	—	—	**96.6**	3.4	0	—	—
I	200	—	**95.4**	4.4	0.2	—	—	**95.6**	4.2	0.2	—	—
I	500	—	**95.6**	4.2	0.2	—	—	**96.2**	3.6	0.2	—	—
		25	55.7	**43.5**	0.8	26	4	52.3	**47.1**	0.6	34	7
$2 \times I$	100	50	37.7	**61.5**	0.8	51	5	14.2	**84.8**	1.0	52	3
		75	49.5	**50.3**	0.2	75	4	25.3	**74.7**	0	74	3
		50	22.8	**75.0**	2.2	51	6	13.4	**83.2**	3.4	57	7
$2 \times I$	200	100	7.6	**90.6**	1.8	100	5	1.2	**96.2**	2.6	101	3
		150	14.8	**84.2**	1.0	150	4	4.8	**93.8**	1.4	150	4
		125	0	**98.6**	1.4	126	4	0.2	**95.4**	4.4	130	6
$2 \times I$	500	250	0	**96.8**	3.2	251	3	0	**94.0**	6.0	252	3
		375	0.2	**97.4**	2.4	376	4	0	**95.2**	4.8	375	4
		25	9.6	**88.8**	1.6	25	2	9.4	**88.4**	2.2	28	3
$3 \times I$	100	50	2.2	**97.0**	0.8	50	2	0.2	**98.0**	1.8	51	1
		75	6.4	**92.6**	1.0	75	2	1.6	**97.0**	1.4	75	2
		50	0	**97.4**	2.6	50	2	0	**95.8**	4.2	52	2
$3 \times I$	200	100	0	**96.8**	3.2	100	2	0	**96.6**	3.4	101	1
		150	0.2	**97.4**	2.4	150	1	0.2	**95.6**	4.2	150	1
		125	0	**95.4**	4.6	125	1	0	**92.0**	8.0	126	2
$3 \times I$	500	250	0	**97.2**	2.8	250	1	0.2	**91.8**	8.0	251	1
		375	0	**96.6**	3.4	375	2	0	**93.4**	6.6	375	1

Table 5.4: Results for model 2 and one variance change.

			LR procedure					**Cusum Procedure**				
Ω	**n**	**h**	**frequency**			$\widehat{\mathbf{h}}$		**frequency**			$\widehat{\mathbf{h}}$	
			0	1	≥ 2	Med.	Mad	0	1	≥ 2	Med.	Mad
I	100	—	**95.8**	4.0	0.2	—	—	**96.0**	4.0	0	—	—
I	200	—	**93.4**	6.2	0.4	—	—	**95.4**	4.4	0.2	—	—
I	500	—	**97.0**	3.0	0	—	—	**95.8**	3.8	0.4	—	—
		25	39.7	**59.7**	0.6	25	3	30.5	**68.1**	1.4	30	5
$2 \times I$	100	50	23.4	**76.0**	0.6	50	4	5.4	**92.6**	2.0	51	2
		75	41.1	**58.3**	0.6	75	4	14.4	**84.8**	0.8	75	3
		50	8.4	**90.4**	1.2	50	4	1.2	**96.2**	2.6	53	4
$2 \times I$	200	100	1.2	**96.6**	2.2	100	3	0.2	**97.2**	2.6	101	2
		150	5.4	**92.8**	1.8	150	3	1.2	**96.8**	2.0	150	2
		125	0	**97.8**	2.2	125	2	0	**92.0**	8.0	127	4
$2 \times I$	500	250	0	**97.8**	2.2	250	2	0.2	**94.4**	5.4	251	2
		375	0	**98.2**	1.8	375	3	0.2	**92.6**	7.2	375	2
		25	1.2	**97.6**	1.2	25	1	0.8	**97.8**	1.4	26	2
$3 \times I$	100	50	0	**98.2**	1.8	50	1	0	**96.4**	3.6	50	1
		75	2.4	**96.4**	1.2	75	1	0.4	**96.8**	2.8	75	1
		50	0	**98.4**	1.6	50	1	0	**96.0**	4.0	51	1
$3 \times I$	200	100	0	**95.6**	4.4	100	1	0.2	**92.2**	7.6	101	1
		150	0	**97.6**	2.4	150	1	0	**97.0**	3.0	150	1
		125	0	**97.8**	2.2	125	1	0.2	**93.0**	6.8	126	2
$3 \times I$	500	250	0	**96.4**	3.6	250	1	0	**92.8**	7.2	251	1
		375	0	**97.8**	2.2	375	1	0	**95.2**	4.8	375	1

the sample size is small ($n = 100$) and small changes ($\Omega_1 = 2 \times I$, $\Omega_2 = I$), the detection frequency is low: 17.2 % and 34.9 % for $k = 2$ and $k = 3$, respectively. In the rest of the cases, the cusum procedure works quite well, with several cases over the 90 % of detection frequency. As in the previous case, as the sample size increase, the change is larger and the number of components increase, the procedure works better. It also appears that the estimate of the second change point has smallest mad, suggesting that the procedure detect more precisely the change at the end of the series. The median of the estimates are quite approximated to the real change points except with the smallest sample size and the smallest changes.

Now, we study the case of both changes in variances and correlations. We make a simulation study in order to study the power of the proposed procedures for the case of a single change. For that, we consider the same models and sample sizes for $k = 2$ that in the previous case. The changes are introduced by transforming the original covariance matrix, $\Sigma = I$, into $\Omega = (I + W)(I + W)'$, where W is a lower triangular matrix. We consider two possible matrices W associated with two matrices Ω_1 and Ω_2 that represent the situation in which the variances of each component is multiplied by 2 and the covariances pass from 0 to 0.5 and -0.5 respectively. For each case, we generate 5000 realizations. Then, we apply the two procedures with the 5% critical values from Table 5.2. The results are shown in Table 5.8, with the same design as before. The case with $\Omega = I$ shows the type I error of the procedures, which is around the 5 % in all the sample sizes considered. When $n = 100$, the cusum procedure appears to work better than the LR procedure. For $n = 200$ the cusum procedure is slightly better than the LR one, but when $n = 500$, the detection frequency of one change point is larger than 90 % and there is a little increase of the detection of two changes in the cusum procedure. The estimates of the time of the change, h, are quite similar for both procedures.

Finally, we study the power of the statistics when there is also a change in the parameter matrices, which will be called a structural change. Let y_t a series generated by the following model:

$$\left\{ \begin{array}{cc} \Phi_1(B)\, y_t = c_1 + \Theta_1(B)\, a_t & t < h \\ \Phi_2(B)\, y_t = c_2 + \Theta_2(B)\left(a_t + W S_t^{(h)} a_t\right) & t \geq h \end{array} \right. ,$$

Table 5.5: Results for model 1 and two variance changes.

LR procedure												
$\boldsymbol{\Omega}_1$	$\boldsymbol{\Omega}_2$	**n**	$\mathbf{h}_1$	$\mathbf{h}_2$	**frequency**				$\widehat{\mathbf{h}}_1$		$\widehat{\mathbf{h}}_2$	
					0	1	2	≥ 3	Med.	Mad	Med.	Mad
		100	33	66	75.2	18.2	**6.6**	0	33	4	68	4
$2 \times I$	I	200	66	133	51.1	17.8	**30.7**	0.4	66	3	133	4
		500	166	333	6.6	3.6	**85.0**	4.8	166	4	333	4
		100	33	66	33.9	14.8	**50.9**	0.4	33	2	66	2
$3 \times I$	I	200	66	133	2.6	2.0	**92.4**	3.0	66	2	133	2
		500	166	333	0	0	**82.6**	17.4	166	1	333	1
		100	33	66	9.4	67.5	**22.8**	0.2	32	3	66	1
$2 \times I$	$1/2 \times I$	200	66	133	3.8	37.7	**57.9**	0.6	65	4	133	1
		500	166	333	17.4	2.4	**76.4**	3.8	166	4	333	1
		100	33	66	12.4	86.0	**1.6**	0	21	4	66	1
$1/2 \times I$	$2 \times I$	200	66	133	9.2	71.1	**19.0**	0.6	59	7	133	1
		500	166	333	21.4	9.0	**59.9**	9.6	163	5	334	1
		100	33	66	3.6	23.4	**72.1**	0.8	33	2	66	0
$3 \times I$	$1/3 \times I$	200	66	133	4.6	3.8	**88.6**	3.0	66	2	133	0
		500	166	333	0	1.0	**81.2**	17.8	166	1	333	0
		100	33	66	3.8	83.0	**12.4**	0.8	26	4	66	1
$1/3 \times I$	$3 \times I$	200	66	133	9.0	19.8	**63.3**	7.8	63	3	133	0
		500	166	333	0.8	1.6	**78.6**	19.0	164	2	333	0
Cusum Procedure												
$\boldsymbol{\Omega}_1$	$\boldsymbol{\Omega}_2$	**n**	$\mathbf{h}_1$	$\mathbf{h}_2$	**frequency**				$\widehat{\mathbf{h}}_1$		$\widehat{\mathbf{h}}_2$	
					0	1	2	≥ 3	Med.	Mad	Med.	Mad
		100	33	66	75.4	7.4	**17.2**	0	35	2	64	2
$2 \times I$	I	200	66	133	31.3	2.8	**64.9**	1.0	68	2	132	2
		500	166	333	0.6	0.2	**91.8**	7.4	167	3	331	3
		100	33	66	27.5	3.2	**68.9**	0.4	34	1	65	1
$3 \times I$	I	200	66	133	0.2	0	**95.0**	4.8	67	1	132	2
		500	166	333	0	0.2	**91.4**	8.4	167	1	331	2
		100	33	66	20.8	31.5	**47.5**	0.2	35	3	65	1
$2 \times I$	$1/2 \times I$	200	66	133	26.5	5.2	**65.5**	2.8	67	3	132	1
		500	166	333	23.4	0.6	**70.5**	5.4	168	4	332	1
		100	33	66	12.6	72.7	**14.0**	0.6	31	2	67	1
$1/2 \times I$	$2 \times I$	200	66	133	30.9	22.6	**41.7**	4.8	64	3	134	1
		500	166	333	23.4	3.6	**65.3**	7.6	165	3	334	1
		100	33	66	16.2	4.4	**78.4**	1.0	34	1	65	1
$3 \times I$	$1/3 \times I$	200	66	133	8.2	2.2	**86.0**	3.6	67	1	132	1
		500	166	333	0.2	0.8	**92.0**	7.0	167	2	332	1
		100	33	66	18.4	43.9	**34.5**	3.2	32	1	67	1
$1/3 \times I$	$3 \times I$	200	66	133	9.2	7.4	**78.6**	4.8	65	2	134	1
		500	166	333	0.8	4.0	**87.8**	7.4	165	2	334	1

Table 5.6: Results for model 2 and two variance changes.

LR procedure												
$\mathbf{\Omega}_1$	$\mathbf{\Omega}_2$	**n**	$\mathbf{h}_1$	$\mathbf{h}_2$	**frequency**				$\widehat{\mathbf{h}}_1$		$\widehat{\mathbf{h}}_2$	
					0	1	2	≥ 3	Med.	Mad	Med.	Mad
		100	33	66	67.9	21.8	**10.2**	0	32	4	67	3
$2 \times I$	I	200	66	133	3.6	6.4	**80.4**	9.6	64	2	133	0
		500	166	333	1.2	0.2	**93.8**	4.8	166	4	333	3
		100	33	66	19.4	13.2	**66.5**	0.8	33	2	66	1
$3 \times I$	I	200	66	133	0.2	0.2	**95.0**	4.6	66	1	133	1
		500	166	333	0	0	**73.3**	26.7	166	1	333	1
		100	33	66	2.8	65.1	**32.1**	0	32	3	66	1
$2 \times I$	$1/2 \times I$	200	66	133	3.8	25.1	**68.3**	2.8	65	4	133	1
		500	166	333	7.6	1.0	**85.8**	5.6	166	3	333	1
		100	33	66	3.0	93.0	**4.0**	0	24	6	66	0
$1/2 \times I$	$2 \times I$	200	66	133	7.0	61.3	**29.9**	1.8	61	5	134	1
		500	166	333	10.2	3.0	**71.1**	15.6	163	4	334	1
		100	33	66	3.6	13.2	**82.4**	0.8	33	1	66	0
$3 \times I$	$1/3 \times I$	200	66	133	1.6	1.2	**90.9**	6.2	66	1	133	0
		500	166	333	0	0	**78.8**	21.2	166	1	333	0
		100	33	66	4.8	76.4	**18.4**	0.4	29	4	66	0
$1/3 \times I$	$3 \times I$	200	66	133	4.8	5.6	**78.6**	11.0	64	2	133	0
		500	166	333	0	0.4	**81.0**	18.6	165	1	333	0
Cusum Procedure												
$\mathbf{\Omega}_1$	$\mathbf{\Omega}_2$	**n**	$\mathbf{h}_1$	$\mathbf{h}_2$	**frequency**				$\widehat{\mathbf{h}}_1$		$\widehat{\mathbf{h}}_2$	
					0	1	2	≥ 3	Med.	Mad	Med.	Mad
		100	33	66	54.9	9.8	**34.9**	0.4	34	2	65	1
$2 \times I$	I	200	66	133	3.6	4.6	**88.0**	3.8	65	1	133	0
		500	166	333	0	0.2	**92.6**	7.2	167	2	332	2
		100	33	66	6.4	0.8	**90.6**	2.2	34	1	65	1
$3 \times I$	I	200	66	133	0.2	0	**95.2**	4.6	66	1	132	1
		500	166	333	0	0	**92.0**	8.0	167	1	333	1
		100	33	66	15.0	18.8	**65.7**	0.4	34	1	65	1
$2 \times I$	$1/2 \times I$	200	66	133	20.8	1.4	**74.3**	3.4	67	2	132	1
		500	166	333	10.0	0.4	**81.4**	8.2	167	2	332	1
		100	33	66	14.4	59.5	**23.8**	2.2	32	2	66	0
$1/2 \times I$	$2 \times I$	200	66	133	19.4	9.0	**64.9**	6.6	65	2	133	0
		500	166	333	12.8	3.0	**78.8**	5.4	165	2	334	1
		100	33	66	13.4	0.8	**84.4**	1.4	33	1	66	0
$3 \times I$	$1/3 \times I$	200	66	133	2.0	0.8	**93.4**	3.8	67	1	133	0
		500	166	333	0	0	**90.6**	9.4	167	1	333	0
		100	33	66	12.4	24.6	**58.7**	4.2	33	1	66	0
$1/3 \times I$	$3 \times I$	200	66	133	4.6	3.0	**85.0**	7.4	65	1	133	0
		500	166	333	0	0.8	**90.8**	8.4	165	1	333	0

Table 5.7: Models for the simulation study with structural changes.

Π_1	Π_2	Ω_1	Ω_2
$\begin{vmatrix} 0.6 & 0.2 \\ 0.2 & 0.4 \end{vmatrix}$	$\begin{vmatrix} 0.3 & 0.4 \\ 0.4 & 0.7 \end{vmatrix}$	$\begin{vmatrix} 2 & 0.5 \\ 0.5 & 2 \end{vmatrix}$	$\begin{vmatrix} 2 & -0.5 \\ -0.5 & 2 \end{vmatrix}$

Table 5.8: Results for model 1 with variances and correlation changes.

			LR procedure					**Cusum Procedure**				
$\mathbf{\Omega}$	**n**	**h**	**frequency**			$\widehat{\mathbf{h}}$		**frequency**			$\widehat{\mathbf{h}}$	
			0	1	≥ 2	Med.	Mad	0	1	≥ 2	Med.	Mad
I	100	—	**95.4**	4.6	0	—	—	**94.6**	5.4	0	—	—
I	200	—	**95.0**	5.0	0	—	—	**95.4**	4.4	0.2	—	—
I	500	—	**94.6**	5.2	0.2	—	—	**94.6**	5.0	0.4	—	—
		25	63.3	**36.3**	0.4	26	6	60.7	**38.1**	1.2	34	7
Ω_1	100	50	45.9	**53.3**	0.8	50	5	18.0	**81.2**	0.8	52	3
		75	58.7	**40.7**	0.6	75	3	31.3	**67.9**	0.8	74	3
		50	26.1	**71.9**	2.0	51	4	21.0	**77.6**	1.4	56	6
Ω_1	200	100	9.2	**89.0**	1.8	101	5	1.6	**95.8**	2.6	103	4
		150	16.6	**81.8**	1.6	151	4	5.4	**91.8**	2.8	149	4
		125	0.2	**96.4**	3.4	126	4	0.2	**93.0**	6.8	129	6
Ω_1	500	250	0	**97.4**	2.6	251	3	0	**94.0**	6.0	252	4
		375	0	**96.8**	3.2	376	3	0	**93.0**	7.0	375	4
		25	65.3	**34.1**	0.6	26	5	60.7	**38.1**	1.2	36	7
Ω_2	100	50	47.1	**52.3**	0.6	51	6	20.8	**78.6**	0.6	52	3
		75	53.9	**45.3**	0.8	75	4	31.3	**68.1**	0.6	74	3
		50	20.2	**77.8**	2.0	51	5	15.4	**81.8**	2.8	56	6
Ω_2	200	100	9.4	**87.8**	2.8	101	4	0.4	**97.6**	2.0	102	4
		150	15.8	**83.0**	1.2	151	5	5.4	**92.2**	2.4	150	4
		125	0.4	**98.0**	1.6	126	4	0.2	**95.4**	4.4	130	6
Ω_2	500	250	0	**97.2**	2.8	251	3	0	**94.8**	5.2	252	4
		375	0	**96.4**	3.6	376	3	0.2	**93.6**	6.2	375	4

Table 5.9: Results for structural changes.

			LR procedure					Cusum Procedure				
Ω	**n**	**h**	**frequency**			$\widehat{\mathbf{h}}$		**frequency**			$\widehat{\mathbf{h}}$	
			0	1	≥ 2	Med.	Mad	0	1	≥ 2	Med.	Mad
		25	74.1	**25.3**	0.6	24	6	73.3	**26.5**	0.2	41	10
Ω_1	100	50	54.3	**45.3**	0.4	50	5	28.1	**71.5**	0.4	52	3
		75	54.1	**45.7**	0.2	75	4	30.5	**69.1**	0.4	74	3
		50	43.7	**54.9**	1.4	50	8	43.9	**54.5**	1.6	63	12
Ω_1	200	100	15.8	**81.8**	2.4	101	8	5.0	**92.6**	2.4	102	5
		150	14.2	**80.4**	5.4	152	5	5.4	**91.6**	3.0	149	4
		125	3.2	**91.2**	5.6	126	6	2.0	**94.0**	4.0	132	11
Ω_1	500	250	0	**91.6**	8.4	251	4	0	**93.2**	6.8	252	5
		375	0.2	**72.5**	27.3	377	4	0	**94.0**	6.0	375	5
		25	73.7	**26.1**	0.2	25	5	75.8	**23.8**	0.4	38	10
Ω_2	100	50	53.9	**44.9**	1.2	49	8	31.5	**67.7**	0.8	52	4
		75	48.7	**50.5**	0.8	75	4	31.3	**67.9**	0.8	74	4
		50	44.5	**53.1**	2.4	50	8	40.9	**58.3**	0.8	62	12
Ω_2	200	100	12.4	**85.6**	2.0	101	7	3.0	**93.0**	4.0	102	5
		150	11.2	**83.4**	5.4	151	4	5.0	**92.8**	2.2	149	5
		125	5.8	**91.4**	2.8	126	8	3.2	**93.4**	3.4	132	10
Ω_2	500	250	0	**90.2**	9.8	251	5	0	**96.0**	4.0	253	5
		375	0.2	**71.9**	27.9	377	4	0	**92.0**	8.0	375	4

such that the covariance matrix as well as the polynomials of the model change at time $t = h$. The polynomials $\Phi_1(B)$, $\Phi_2(B)$, $\Theta_1(B)$ and $\Theta_2(B)$ are assumed to verify the conditions for stationarity and invertibility in section 2. If the procedures had good power properties for detecting a covariance change under a structural change, they could be an useful tool for detecting both covariance changes and structural changes. Consider the models in Table 5.7 and the sample sizes $n = 100$, 200 and 500, for $k = 2$. The changes are introduced by transforming the original covariance matrix, $\Sigma = I$, into Ω_1 and Ω_2, and the autoregressive polynomial Π_1 into Π_2. For each case, we generate 5000 realizations. Then, we apply the two procedures with the 5% critical values from Table 5.2. The results are shown in Table 5.9. We conclude that both procedures have a small decrease in power for small samples sizes, specially in the case in which $h = [0.25n]$, but they do not lose power for big sample sizes, here $n = 500$.

6. Illustrative examples

6.1 Example: Flour data

We consider the trivariate series of the logarithms of monthly flour price indices from three U.S. cities over the period August 1972 through November 1980. This vector series was analyzed by Tiao and Tsay (1989), Grubb (1992) and Lütkepohl and Poskitt (1996) and is shown in Figure 2. Tiao and Tsay (1989) fitted a restricted vector ARMA(1,1) to the series whereas Grubb (1992), by using the Akaike Information Criteria (AIC), chose a restricted VAR(2) model. Lütkepohl and Poskitt (1996) investigate cointegration in these series using the Johansen's test in a VAR(2) model, rejecting the null hypothesis of cointegration. Then, they fitted a VAR(1) for the differenced series, which is showed in the second row in Table 10.

Using this model, we will apply the LR and cusum procedures for variance changes to the data. Table 11 summarizes the results. No variance changes are detected by both procedures. Then, we apply the LR and cusum procedures for variances and correlation changes for the data, and both procedures detect one change point at $t = 33$ (April, 1975). The estimation of W is done as in (4.5). Taking into account that there is no evidence of variance changes, apparently the change happen in the correlation between the components of the series.

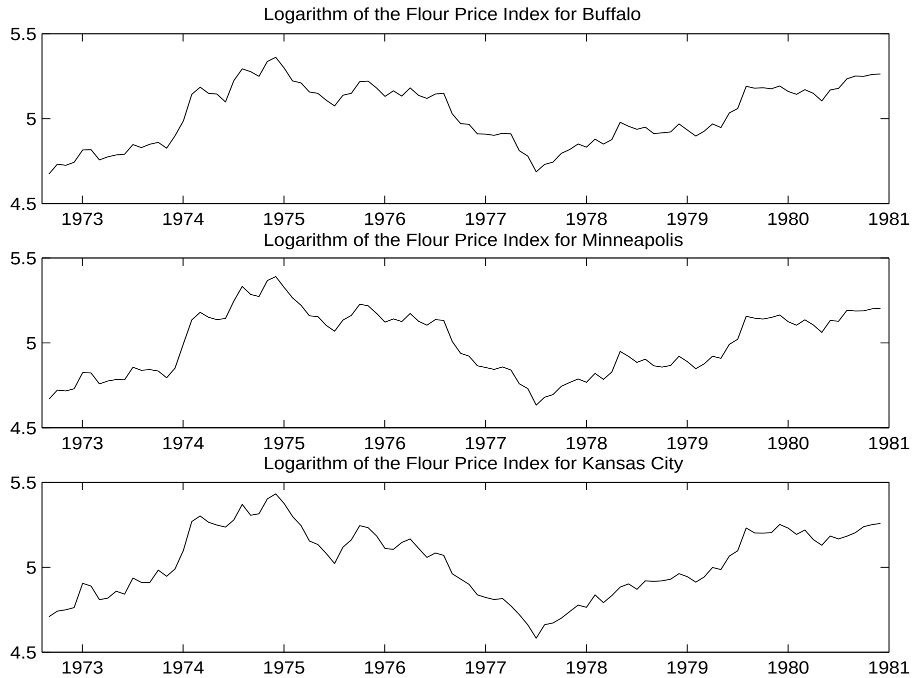

Figure 6.2: Monthly Flour Price Indices for Three U.S. Cities.

We include in the table the values of the Akaike and Bayesian information criteria (AIC and BIC) for each model, given by, $-(2/n)\log$ (maximized likelihood)$+$ (c/n)(number of parameters), where $c = 2$ for AIC and $c = \log(n)$ for BIC. Note that the value of both criteria is reduced when the covariance change is introduced. Both criteria indicate that the model with one covariance change at $t = 33$ (April, 1975) appears to be most appropriate for the data. The final estimated model is shown in the third row in Table 10.

Finally, we estimate a VAR(1) model to the subsamples 1-32 and 33-100. The two estimated models with their standard errors are given in the fourth and fifth rows in Table 10. As we can see, the parameters of the model and the sample residual covariance matrices are different in both models so that we conclude that the series has an structural change at $t = 33$. The model for the second part of the series is apparently a random walk and the sample residual covariance matrix has smaller values than the obtained in the first part of the series.

Table 6.10: AR parameter matrix ($\widehat{\Pi}$) and estimated covariance matrix ($\widehat{\Sigma}$) for three models fitted to the flour series. Standard errors of the coefficients are under parenthesis.

Model	$\widehat{\Pi}$	$10^2\widehat{\Sigma}$
VAR(1)	$\begin{Vmatrix} \underset{(0.17)}{-0.86} & \underset{(0.18)}{1.01} & 0 \\ \underset{(0.17)}{-0.43} & \underset{(0.19)}{0.62} & 0 \\ 0 & \underset{(0.10)}{0.25} & 0 \end{Vmatrix}$	$\begin{vmatrix} 0.20 & 0.21 & 0.20 \\ 0.21 & 0.24 & 0.22 \\ 0.20 & 0.22 & 0.27 \end{vmatrix}$
1 Variance Change	$\begin{Vmatrix} \underset{(0.15)}{-0.83} & \underset{(0.16)}{0.96} & 0 \\ \underset{(0.13)}{-0.48} & \underset{(0.14)}{0.64} & 0 \\ 0 & \underset{(0.11)}{0.21} & 0 \end{Vmatrix}$	$\begin{vmatrix} 0.20 & 0.21 & 0.20 \\ 0.21 & 0.23 & 0.22 \\ 0.20 & 0.22 & 0.26 \end{vmatrix}$
First part (1-32)	$\begin{Vmatrix} \underset{(0.14)}{-0.61} & \underset{(0.20)}{0.95} & 0 \\ 0 & \underset{(0.16)}{0.40} & 0 \\ 0 & \underset{(0.18)}{0.43} & 0 \end{Vmatrix}$	$\begin{vmatrix} 0.24 & 0.26 & 0.27 \\ 0.26 & 0.30 & 0.31 \\ 0.27 & 0.31 & 0.37 \end{vmatrix}$
Second part (33-100)	$\begin{Vmatrix} \underset{(0.11)}{-0.28} & \underset{(0.11)}{0.26} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$	$\begin{vmatrix} 0.19 & 0.26 & 0.17 \\ 0.20 & 0.21 & 0.19 \\ 0.17 & 0.19 & 0.21 \end{vmatrix}$

Table 6.11: Summary of the LR and cusum procedures for the flour data.

Method	VAR(1)	LR W diag.	Cusum W diag.	LR W triang.	CUSUM W triang.
h	—	28	60	33	33
$\Lambda_{\max}/\Gamma_{\max}$	—	13.26	0.63	28.95	1.78
$\widehat{W}$	—	—	—	$\begin{Vmatrix} -0.14 & 0 & 0 \\ 0.40 & -0.53 & 0 \\ -0.02 & -0.17 & -0.09 \end{Vmatrix}$	$\begin{Vmatrix} -0.14 & 0 & 0 \\ 0.40 & -0.53 & 0 \\ -0.02 & -0.17 & -0.09 \end{Vmatrix}$
AIC	-14.01	-14.01	-14.01	-27.39	-27.39
BIC	-13.77	-13.77	-13.77	-26.99	-26.99

6.2 Example: Wheat data

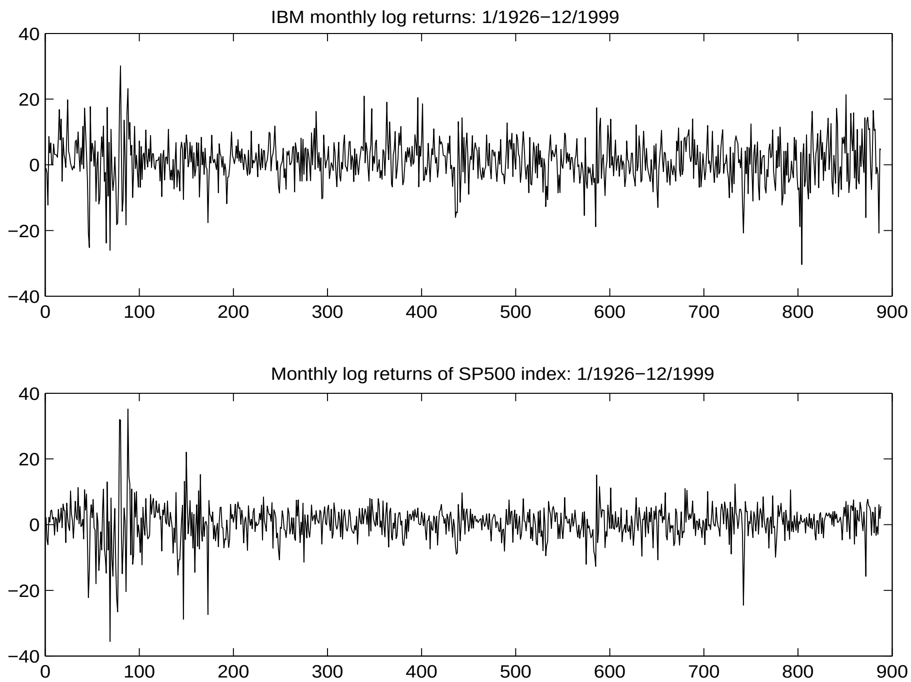

Figure 6.3: Monthly Wheat Price Indices for Five Provinces in Castillia, Spain.

We consider the series of the logarithms of the monthly wheat price indices from five provinces in Castillia, Spain, over the period July 1880 through December 1890. This vector series was analyzed in Peña and Box (1987) and is shown in Figure 3. We investigate cointegration in these series using the Johansen's test and by using the BIC we chose a VAR(1) model with three cointegration relationships. This is in agreement with the two factors found by Peña and Box (1987). Then, we apply the LR and cusum procedures for variance changes to the data assuming first, an unrestricted VAR(1) model, and second, a vector correction model with three cointegration relationships. As the results obtained with these two models were the same, we only report here the one with the estimated unrestricted VAR(1) model which is shown in the second row in Table 12. Table 13 summarizes the results. The LR procedure detects one change at $h = 22$ (April, 1882) and another at $h = 90$ (December, 1888). The cusum procedure detects three changes at $h = 22$ (April, 1882), $h = 41$ (November, 1883) and $h = 90$

Table 6.12: AR parameter matrix ($\widehat{\Pi}$) and estimated covariance matrix ($\widehat{\Sigma}$) for two models fitted to the wheat series. Standard errors of the coefficients are under parenthesis.

Model	$\widehat{\Pi}_1$	$10^3\widehat{\Sigma}$
VAR(1)	$\begin{vmatrix} \underset{(0.07)}{0.87} & 0 & 0 & \underset{(0.12)}{0.44} & 0 \\ \underset{(0.05)}{0.13} & \underset{(0.07)}{0.32} & 0 & \underset{(0.09)}{0.43} & 0 \\ 0 & \underset{(0.13)}{-0.26} & \underset{(0.09)}{0.40} & \underset{(0.16)}{0.82} & 0 \\ \underset{(0.05)}{0.17} & \underset{(0.08)}{-0.21} & \underset{(0.06)}{0.15} & \underset{(0.10)}{0.85} & 0 \\ 0 & 0 & 0 & \underset{(0.14)}{0.46} & \underset{(0.08)}{0.47} \end{vmatrix}$	$\begin{vmatrix} 1.44 & 0.46 & 0.59 & 0.68 & 0.59 \\ 0.46 & 0.84 & 0.48 & 0.32 & 0.40 \\ 0.59 & 0.48 & 2.39 & 0.59 & 0.62 \\ 0.68 & 0.32 & 0.59 & 0.91 & 0.50 \\ 0.59 & 0.40 & 0.62 & 0.50 & 2.00 \end{vmatrix}$
3 V. C.	$\begin{vmatrix} \underset{(0.05)}{0.92} & \underset{(0.10)}{-0.34} & 0 & \underset{(0.09)}{0.48} & 0 \\ \underset{(0.03)}{0.10} & \underset{(0.05)}{0.39} & \underset{(0.03)}{0.06} & \underset{(0.05)}{0.37} & 0 \\ 0 & \underset{(0.18)}{-0.73} & \underset{(0.09)}{0.41} & \underset{(0.16)}{1.10} & 0 \\ \underset{(0.05)}{0.17} & \underset{(0.10)}{-0.50} & \underset{(0.05)}{0.14} & \underset{(0.09)}{1.05} & 0 \\ 0 & 0 & \underset{(0.06)}{0.14} & \underset{(0.10)}{0.55} & \underset{(0.11)}{0.54} \end{vmatrix}$	$\begin{vmatrix} 0.64 & 0.21 & 0.34 & 0.30 & 0.35 \\ 0.21 & 0.18 & 0.31 & 0.24 & 0.26 \\ 0.34 & 0.31 & 1.94 & 0.32 & 0.71 \\ 0.30 & 0.24 & 0.32 & 0.61 & 0.37 \\ 0.35 & 0.26 & 0.71 & 0.37 & 0.77 \end{vmatrix}$

Table 6.13: Summary of the LR and cusum procedures for variance changes for the wheat data.

Method	VAR(1)	LR		Cusum		
h	—	23	90	22	41	90
$\Lambda_{\max}/\Gamma_{\max}$	—	30.99	90.93	1.61	1.71	1.70
$\underset{I.C.}{\widehat{W}(1,1)}$	—	0	$\underset{(-0.57,-0.37)}{-0.48}$	$\underset{(0.02,1.22)}{0.54}$	$\underset{(-0.60,-0.13)}{-0.38}$	$\underset{(-0.52,-0.21)}{-0.38}$
$\underset{I.C.}{\widehat{W}(2,2)}$	—	$\underset{(-0.46,-0.21)}{0.71}$	$\underset{(-0.46,-0.21)}{-0.34}$	$\underset{(0.43,2.10)}{1.16}$	0	$\underset{(-0.49,-0.15)}{-0.33}$
$\underset{I.C.}{\widehat{W}(3,3)}$	—	$\underset{(-0.52,-0.04)}{-0.30}$	$\underset{(-0.66,-0.50)}{-0.58}$	0	0	$\underset{(-0.64,-0.40)}{-0.53}$
$\underset{I.C.}{\widehat{W}(4,4)}$	—	0	$\underset{(-0.42,-0.15)}{-0.29}$	0	$\underset{(-0.59,-0.11)}{-0.37}$	0
$\underset{I.C.}{\widehat{W}(5,5)}$	—	$\underset{(0.03,1.05)}{0.51}$	$\underset{(-0.42,-0.14)}{-0.29}$	$\underset{(0.40,2.04)}{1.12}$	$\underset{(-0.63,-0.20)}{-0.43}$	0
AIC	-19.84	-32.85		-37.52		
BIC	-19.52	-32.28		-36.98		

Table 6.14: Summary of the LR and cusum procedures for variance and correlation changes for the wheat data.

Method	h	$\Lambda_{\max}/\Gamma_{\max}$	$\widehat{W}$	AIC	BIC
VAR(2)	—	—	—	-19.84	-19.52
LR	88	82.79	$\begin{array}{ccccc} -0.50 & 0 & 0 & 0 & 0 \\ -0.17 & 0.09 & 0 & 0 & 0 \\ 0.07 & -0.42 & -0.53 & 0 & 0 \\ -0.07 & 0.00 & -0.03 & -0.20 & 0 \\ -0.20 & -0.33 & -0.16 & 0.72 & -0.32 \end{array}$	-29.33	-27.36
Cusum	22	2.35	$\begin{array}{ccccc} 0.74 & 0 & 0 & 0 & 0 \\ -0.51 & 2.00 & 0 & 0 & 0 \\ 0.53 & -1.67 & 0.31 & 0 & 0 \\ 0.81 & -2.55 & 0.35 & 0.50 & 0 \\ 0.22 & -2.35 & -0.33 & 0.06 & 1.48 \end{array}$	-38.85	-37.43
	35	2.26	$\begin{array}{ccccc} -0.41 & 0 & 0 & 0 & 0 \\ 0.07 & -0.42 & 0 & 0 & 0 \\ 0.11 & 0.32 & -0.33 & 0 & 0 \\ -0.12 & 0.30 & -0.05 & -0.41 & 0 \\ 0.00 & 0.25 & 0.19 & -0.06 & -0.28 \end{array}$		
	90	2.03	$\begin{array}{ccccc} -0.39 & 0 & 0 & 0 & 0 \\ 0.03 & -0.32 & 0 & 0 & 0 \\ -0.08 & -0.21 & -0.40 & 0 & 0 \\ -0.04 & 0.05 & -0.06 & 0.00 & 0 \\ -0.08 & -0.02 & -0.33 & 0.61 & -0.32 \end{array}$		

(December, 1888). The estimation of the changes and their confidence intervals appears in Table 13. If one of the changes in one component is not significant, we represented it by 0. The minimum of the values of both the AIC and BIC is obtained in the model proposed by the cusum with three changes.

Then, we apply the LR and cusum procedures for variances and correlation changes for the data which are summarized in Table 14. The LR procedure detects now one change at $h = 88$ (October, 1888), while the cusum procedure detects three changes at $h = 22$ (April, 1882), $h = 35$ (May, 1883) and $h = 90$ (December, 1888). The estimation of the changes appears in Table 14. The minimum of the values of both the AIC and BIC corresponds to the model proposed by the cusum with three changes.

The final model is selected by the BIC (although AIC gives the same results) and is the one obtained by the cusum procedure allowing variances and correlation changes. The estimated parameters for this VAR(1) model are shown in the third row in Table 12.

Acknowledgment

This paper is dedicated to Professor George C. Tiao. We acknowledge support from Grant BEC 2000-0167, Spain.

References

Abraham, B. and Wei, W. W. S. (1984) "Inferences About the Parameters of a Time Series Model with Changing Variance", *Metrika, 31, 183-194.*

Baufays, P. and Rasson, J. P. (1985) "Variance Changes in Autoregressive Models", *Time Series. Theory and Methods,* 2nd Ed. Springer, New York.

Booth, N. B. and Smith, A. F. M. (1982) "A Bayesian Approach to retrospective identification of the change-points", *Journal of Econometrics, 19, 7-22.*

Chen, J. and Gupta, A. K. (1997) "Testing and Locating Variance Changepoints with Application to Stock Prices", *Journal of the American Statistical Association, 92, 438, 739-747.*

Grubb, H. (1992) "A Multivariate Time Series Analysis of Some Flour Price Data", *Applied Statistics, 26, 279-284.*

Hinkley, D. V. (1971) “Inference about the change-point from cumulative sum tests”, *Biometrika, 58, 509-523.*

Hsu, D. A. (1977) “Tests for Variance Shift at an Unknown Time Point”, *Applied Statistics, 26, 279-284.*

Hsu, D. A., Miller, R. B. and Wichern, D. W. (1974) “On the stable Paretian behavior of stock-market prices”, *Journal of the American Statistical Association, 69, 108-113.*

Inclán, C. (1993) “Detection of Multiple Tests of Variances Using Posterior Odds”, *Journal of Business and Economic Statistics, 11, 189-200.*

Inclán, C. and Tiao, G. C. (1994) “Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance”, *Journal of the American Statistical Association, 89, 427, 913-923.*

Lee, S. and Park, S. (2001) “The Cusum of Squares Test for Scale Changes in Infinite Order Moving Average Models”, *Scandinavian Journal of Statistics, 28, 4, 625-644.*

Lütkepohl, H. and Poskitt, D. S. (1996) “Specification of Echelon-Form VARMA models”, *Journal of the Business and Economic Statistics, 14, 1, 69-79.*

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

Park, S., Lee, S. and Jeon J. (2000) “The Cusum of Squares Test for Variance Changes in Infinite Order Autoregressive Models”, *Journal of the Korean Statistical Society, 29, 3, 351-361.*

Peña, D. and Box, G. E. P. (1987) “Identifying a Simplifying Structure in Time Series”, *Journal of the American Statistical Association, 82, 399, 836-843.*

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

Tiao, G. C. and Tsay, R. S. (1989) "Model Specification in Multivariate Time Series (with discussion)", *Journal of the Royal Statistical Society B, 51, 2, 157-213.*

Tsay, R. S. (1988) "Outliers, Level Shifts and Variance Changes in Time Series", *Journal of Forecasting, 7, 1-20.*

Tsay, R. S., Peña, D. and Pankratz, A. E. (2000) "Outliers in Multivariate Time Series", *Biometrika, 87, 789-804.*

Wichern, D. W., Miller, R. B. and Hsu, D. A. (1976) "Changes of Variance in First Order Autoregressive Time Series Models - With An Application", *Applied Statistics, 25, 248-256.*

Pedro Galeano.

Departamento de Estadística, Universidad Carlos III de Madrid, Spain.

E-mail: pgaleano@est-econ.uc3m.es

Daniel Peña.

Departamento de Estadística, Universidad Carlos III de Madrid, Spain.

E-mail: dpena@est-econ.uc3m.es