Journal of Statistical Research *ISSN 0256 - 422 X*
2005, Vol. 39, No. 2, pp. 51–70
Bangladesh

MONITORING JUMP CHANGES IN LINEAR MODELS

Marie Hušková
Department of Statistics, Charles University of Prague
Sokolovská 83, CZ – 186 75 Praha 8, Czech Republic
Email: marie.huskova@karlin.mff.cuni.cz

Alena Koubková
Department of Statistics, Charles University of Prague
Sokolovská 83, CZ – 186 75 Praha 8, Czech Republic
Email: koubkova@karlin.mff.cuni.cz

SUMMARY

Two classes of sequential test procedures for detection of a change in linear regression model are developed and their properties are studied when a training set of data without any change is available. Asymptotic results are checked on simulation study.

Keywords and phrases: Sequential change point analysis

*AMS Classification:*62G20, 62E20, 60F17

1 Introduction

The paper concerns procedures for detection of a change in linear models when data arrive sequentially and training (historical) data with no change are available. Such problems occur in a number of applications, e.g., in economics and finance, geophysical sciences, statistical quality control, medical care.
We assume that the data follow the linear regression model

$$Y_i = \boldsymbol{X}_i^T \boldsymbol{\beta}_i + \epsilon_i, \quad 1 \le i < \infty, \tag{1.1}$$

and concern only the changes in p-dimensional regression parameters $\boldsymbol{\beta}_i$, $1 \le i < \infty$. The stability of the historical data, represented by observations $Y_1, \ldots, Y_m$, are described by the so called *noncontamination condition*

$$\boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m.$$

The sequence $\{\boldsymbol{X}_i, 1 \le i < \infty\}$ are p-dimensional regression random vectors and the sequence $\{\epsilon_i, 1 \le i < \infty\}$ represents the random errors. It is assumed that the data are arriving sequentially.

Our problem of detection of a change in the linear model can be formulated as a sequential hypothesis testing problem, where the null hypothesis corresponds to the model without any change:

$$H_0 : \boldsymbol{\beta}_i = \boldsymbol{\beta}_0, \quad 1 \le i < \infty$$

and the alternative hypothesis says that the model is changing at some unknown point:

$$\begin{aligned} &H_A : \text{there exists } k^* \ge 1 \text{ such that} \\ &\qquad \boldsymbol{\beta}_i = \boldsymbol{\beta}_0,\ 1 \le i < m + k^*, \quad \boldsymbol{\beta}_i = \boldsymbol{\beta}_0 + \boldsymbol{\delta}_m,\ m + k^* \le i < \infty, \quad \boldsymbol{\delta}_m \ne \mathbf{0}. \end{aligned}$$

where $\boldsymbol{\beta}_0, \boldsymbol{\delta}_m$ and k^* are unknown parameters, and data are arriving sequentially.

The considered problem can be described also as follows. At the beginning we have a training data $Y_1, \ldots, Y_m$ that follows the model (1.1) with $\boldsymbol{\beta}_i = \boldsymbol{\beta}_0$. Then the data are arriving sequentially one by one and we want to reveal a change in the regression parameter $\boldsymbol{\beta}$ as soon as it occurs. Such problems can be met not only in statistical quality control but also in econometrics, financial time series or applications in medical research.

In the following we assume either $\boldsymbol{\delta}_m = \boldsymbol{\delta}$ is a fixed nonzero vector or $\boldsymbol{\delta}_m$ can change with m (typically $\boldsymbol{\delta}_m \to \mathbf{0}$ as $m \to \infty$ at some rate). The aim of the paper is to develop and to study procedures based on weighted residuals that have a small probability of false alarm when there is no change and that detect the change as occurs.

Such problems have been considered in a few papers, e.g., Chu et al. [1996] consider test procedures based on CUSUM type test statistics calculated from recursive residuals (see below (2.5)) and a fluctuation test based on differences between estimates of the regression coefficients. The later test statistic was generalized in Leisch et al. [2000] to the so called generalized fluctuation test. Zeileis et al. [2005] suggest MOSUM type statistics based on last h ordinary residuals (see below (2.4)). They also compare three mentioned test statistics through a simulation study.

CUSUM type test statistics based on ordinary residuals (2.4) and on recursive residuals (2.5) are studied in Horváth et al. [2004] (see below for more details). The results are generalized in Aue [2003] to more complex data, but only for location model. Aue [2003] also derives a limit distribution of the delay between the true change-point and its detection for some particular cases. All the mentioned statistics use the OLS (ordinary least squares) estimators of the regression parameter $\boldsymbol{\beta}$. In the quoted papers, the limit distribution of the suggested statistics under the null hypothesis of no change, as well as its limit behavior if a change occurs are studied.

Koubková [2004] provides simulations to specify the behavior of the CUSUM type test statistics based on ordinary L_2-residuals and suggests statistics based on ordinary L_1-residuals.

Our procedures are related to those introduced in Horváth et al. [2004]. Here we introduce other CUSUM type test statistics based on partial sums of weighted ordinary (2.4) and recursive (2.5) residuals. In Section 2 test procedures are described and their limit properties are stated both under the null as well as under alternatives. Results of a simulation study are presented in Section 3. The proofs of assertions from Section 2 are contained in Appendix.

2 Main results

The test procedures for the testing problem H_0 against H_A in our sequential setup are described through the stopping time $\tau(m)$ defined as follows

$$\tau(m) = \begin{cases} \inf\{k \geq 1 : Q(m,k) \geq c\sqrt{m}q(k/m), \} \\ \infty \text{ if } Q(m,k) < c\sqrt{m}q(k/m) \text{ for all } k = 1,2,\ldots, \end{cases} \tag{2.1}$$

where $Q(m,k)$ are test statistics (detectors) based on $Y_1,\ldots,Y_{m+k},\ k=1,2,\ldots;\ q(t),\ t \in (0,\infty)$ is a (stopping) boundary function and the constant $c = c(\alpha)$ is chosen such that

$$\lim_{m\to\infty} P_{H_0}[\tau(m) < \infty] = \alpha \tag{2.2}$$

and

$$\lim_{m\to\infty} P_{H_A}[\tau(m) < \infty] = 1, \tag{2.3}$$

where $\alpha \in (0,1)$. In other words, our sequential test procedure rejects the null hypothesis and stop observations as soon as

$$Q(m,k) > c\sqrt{m}q(k/m)$$

and we stop only when the data indicate that H_0 is violated. In terms of hypothesis testing the condition (2.2) requires the level of the test equals to α while the condition (2.3) corresponds to the requirement that the probability of the type II error tends to zero as $m \to \infty$ (or equivalently, the power of the test tends to 1, as $m \to \infty$).

We assume that the random sequences $\{\epsilon_i,\ 1 \leq i < \infty\}$ and $\{\boldsymbol{X}_i^T,\ 1 \leq i < \infty\}$ satisfy the following conditions:

(A.1) $\{\epsilon_i\}_{i=1}^{\infty}$ is a sequence of independent identically distributed (i.i.d.) random variables with $\mathsf{E}\,\epsilon_1 = 0$, $0 < \mathsf{Var}\,\epsilon_1 = \sigma 2 < \infty$ and $\mathsf{E}\,|\epsilon_1|^\nu < \infty$ for some $\nu > 2$,

(A.2) $\{\boldsymbol{X}_i^T\}_{i=1}^{\infty}$ is a strictly stationary sequence of p-dimensional vectors $\boldsymbol{X}_i^T = (1, X_{2i}, \ldots, X_{pi})$, which is independent of $\{\epsilon_i, 1 \leq i < \infty\}$,

(A.3) there exist a positive definite matrix $\mathbf{C}$ and a constant $\tau > 0$ such that

$$\left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T - \mathbf{C} \right| = O(n^{-\tau}), \quad a.s.$$

where $|.|$ denotes a maximum norm of vectors and matrices.

The assumptions (A.2) and (A.3) are satisfied, e.g., $\{X_{ji}\}_i$, $j = 2, \dots, p$, are autoregressive sequences with finite moments of order higher than 2. In this case the elements c_{jv}, $j, v = 1, \dots, p$ of $\boldsymbol{C}$ are

$$c_{11} = 1, \quad c_{1j} = E\, X_{ji}, \quad c_{jv} = E\, X_{ji} X_{vi}, \quad 1 < j, v \le p.$$

We consider the following two classes of the boundary functions q:

(B.1) $q(t) = q_\gamma(t) = (1+t)\,(t/(t+1))^\gamma$, $t \in (0, \infty)$, where γ is a tuning constant taking values from the interval $[0, \min\{1/2, \tau\})$,

(B.2) q is a positive continuous function on $(0, \infty)$ such that

$$\lim_{t \to 0} \frac{t^\gamma}{q(t)} = 0, \qquad \textit{with} \qquad 0 \le \gamma < \min\{\tau, 1/2\}$$

and

$$\limsup_{t \to \infty} \frac{(t \log \log t)^{1/2}}{q(t)} < \infty.$$

The quantity γ plays a role of a tuning parameter that influences the stopping rule, particularly it modifies the ability of the test procedure to detect better early or late changes in the following way: $\gamma = 0$ is convenient when a late change is expected, while γ close to $1/2$ is appropriate when an early change is expected.

For the considered testing problem, Horváth et al. [2004] propose standardized CUSUM (cumulative sums) test procedure based on L_2-residuals of the form either

$$\widehat{e}_i = Y_i - \boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_m, \tag{2.4}$$

or

$$\widetilde{e}_i = Y_i - \boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_{i-1}, \tag{2.5}$$

where $\widehat{\boldsymbol{\beta}}_n$ is the least square estimator of $\boldsymbol{\beta}$ based on the first n observations, i.e.,

$$\widehat{\boldsymbol{\beta}}_n = \left(\sum_{i=1}^n \boldsymbol{X}_i \boldsymbol{X}_i^T \right)^{-1} \sum_{i=1}^n \boldsymbol{X}_i Y_i.$$

Notice that the ordinary residuals $\widehat{e}_i$ are based on the estimator of $\boldsymbol{\beta}$ calculated from the historical data while the recursive residuals $\widetilde{e}_i$ based on all $i-1$ previous observations, $i = m+1, \dots$. The related standardized CUSUM test statistics are defined as

$$\widehat{Q}(m,k) = \frac{1}{\widehat{\sigma}_m} \sum_{i=m+1}^{m+k} \widehat{e}_i \quad \text{and} \quad \widetilde{Q}(m,k) = \frac{1}{\widehat{\sigma}_m} \sum_{i=m+1}^{m+k} \widetilde{e}_i, \tag{2.6}$$

respectively, where the estimator, $\widehat{\sigma}_m^2$ of the variance of ϵ_i is calculated from the historical period as

$$\widehat{\sigma}_m^2 = \frac{1}{m-p} \sum_{i=1}^{m} (Y_i - \boldsymbol{X}_i^T \widehat{\boldsymbol{\beta}}_m)2. \tag{2.7}$$

Horváth et al. [2004] show that under the assumptions (A.1) – (A.3) and (B.1) and under the null hypothesis

$$\lim_{m\to\infty} P\left[\sup_{1\le k<\infty} \frac{\left|\widehat{Q}(m,k)\right|}{\widehat{\sigma}_m \sqrt{m} q_\gamma(k/m)} \le c\right] = P\left[\sup_{0\le t\le 1} \frac{|W(t)|}{t^\gamma} \le c\right], \tag{2.8}$$

holds for all $c > 0$, where $\{W(t),\, 0 \le t \le 1\}$ denotes a Wiener process. Moreover, they showed that under the alternative H_A and additional assumption on the amount of change $\boldsymbol{\delta}_m = \boldsymbol{\delta}$

$$\boldsymbol{\delta}^T \mathbf{c}_1 \ne 0, \tag{2.9}$$

where $\boldsymbol{c}_1$ denotes the first row of the matrix $\boldsymbol{C}$, the procedure with the detectors $\widehat{Q}(m,k)$ has the desired property (2.3). Horváth et al. [2004] derived similar results also for the procedure based on $\widetilde{Q}(m,k)$.

The assumption (2.9) is quite restrictive, e.g., if $\{\boldsymbol{X}_i\}_i$ is a stationary sequence with $EX_{ji} = 0$, $EX_{ji}^2 < \infty, j = 2,\ldots,p$, $i = 1,\ldots,$ then $\boldsymbol{c}_1^T = (1,0,\ldots,0)$ and the above procedures are not sensitive w.r.t. $\boldsymbol{\delta} \ne \mathbf{0}$ with the zero first component.

In the present paper we investigate procedures based on quadratic forms of partial sums of weighted residuals (2.4) or (2.5) that have the property (2.3) without additional assumption (2.9). Particularly, we consider here procedures based on either

$$\widehat{V}(m,k) = \Big(\sum_{i=m+1}^{m+k} \boldsymbol{x}_i \widehat{e}_i\Big)^T \boldsymbol{C}_m^{-1} \Big(\sum_{i=m+1}^{m+k} \boldsymbol{x}_i \widehat{e}_i\Big) \tag{2.10}$$

or

$$\widetilde{V}(m,k) = \Big(\sum_{i=m+1}^{m+k} \boldsymbol{x}_i \widetilde{e}_i\Big)^T \boldsymbol{C}_m^{-1} \Big(\sum_{i=m+1}^{m+k} \boldsymbol{x}_i \widetilde{e}_i\Big), \tag{2.11}$$

where $\widehat{e}_i$ and $\widetilde{e}_i$ are defined by (2.4) and (2.5), respectively, and

$$\mathbf{C}_k = \sum_{i=1}^{k} \mathbf{X}_i \mathbf{X}_i^T, \quad k = 1,2,\ldots.$$

Notice that $\widehat{V}(m,k)$ can be equivalently expressed as a quadratic form of differences of estimators of $\boldsymbol{\beta}$ based on $Y_{m+1},\ldots,Y_{m+k}$ and on $Y_1,\ldots,Y_m$, particularly,

$$\widehat{V}(m,k) = \Big(\widehat{\boldsymbol{\beta}}_{m+k,m} - \widehat{\boldsymbol{\beta}}_m\Big)^T \Big(\boldsymbol{C}_{m+k} - \boldsymbol{C}_m\Big)^T \boldsymbol{C}_m^{-1} \Big(\boldsymbol{C}_{m+k} - \boldsymbol{C}_m\Big)\Big(\widehat{\boldsymbol{\beta}}_{m+k,m} - \widehat{\boldsymbol{\beta}}_m\Big),$$

where

$$\widehat{\boldsymbol{\beta}}_{m+k,m} = \left(\sum_{i=m+1}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \right)^{-1} \sum_{i=m+1}^{m+k} \mathbf{X}_i Y_i.$$

Hence one can expect the procedures based on $\widehat{V}(m,k)$ will be sensitive w.r.t. to changes in regression parameters and that their large values indicate that the null hypothesis is violated.

The decison rule related to $\widehat{V}(m,k)$ is defined as follows:
having observed $Y_1, \ldots, Y_{m+k}$, the null hypothesis H_0 is rejected as soon as

$$\widehat{V}(m,k)\widehat{\sigma}_m^{-2} \geq c q_\gamma^2(k/m),$$

otherwise we continue with observations. Here $c = \widehat{c}_p(\alpha,\gamma)$ is determined in such a way that (2.2) holds true and the stopping boundary function q_γ satisfies (B.1).

Now, we formulate assertions on the limit behavior of the test statistics based on weighted suprema of $\widehat{V}(m,k)$, $k = 1, 2, \ldots$, under H_0 as well as under alternatives. The former one provides an approximation for $\widehat{c}_p(\alpha,\gamma)$.

Theorem 1. Let $Y_1, Y_2, \ldots,$ follow the model (1.1). Let the assumptions (A.1) – (A.3) and (B.1) be satisfied. Then under H_0

$$\lim_{m\to\infty} P\left(\sup_{1\leq k<\infty} \frac{\widehat{V}(m,k)}{\widehat{\sigma}_m^2 q_\gamma^2(k/m)} \leq x \right) = P\left(\sup_{0\leq t\leq 1} \frac{\sum_{i=1}^p W2_i(t)}{t^{2\gamma}} \leq x \right)$$

for all x, where $\{W_i(t);\, 0 \leq t \leq 1\}$, $i = 1, \ldots, p$ are independent Wiener processes.

Theorem 2. Let $Y_1, Y_2, \ldots,$ follow the model (1.1). Let the assumptions (A.1) – (A.3) and (B.1) be satisfied. Then under the alternatives H_A with

$$\lim_{m\to\infty} m \boldsymbol{\delta}_m^T \boldsymbol{\delta}_m = \infty$$

we have, as $m \to \infty$

$$\sup_{1\leq k<\infty} \frac{\widehat{V}(m,k)}{\widehat{\sigma}_m^2 q_\gamma^2(k/m)} \xrightarrow{P} \infty.$$

The proofs are postponed to Section 4.

The critical values $\widehat{c}_p(\alpha,\gamma)$ can be found as a solution of the equation (w.r.t. c):

$$P\left(\sup_{0\leq t\leq 1} \frac{\sum_{i=1}^p W2_i(t)}{t^{2\gamma}} \leq c \right) = 1 - \alpha. \tag{2.12}$$

The explicit form for the distribution of $\sup_{0\leq t\leq 1} \sum_{i=1}^p W2_i(t) t^{-2\gamma}$ is known only for $\gamma = 0$, otherwise its approximation as well as an approximation for critical values $\widehat{c}_p(\alpha,\gamma)$ can be obtained through simulations of Wiener processes. See Table 1 for simulated critical values $\widehat{c}_p(\alpha,\gamma)$, $p = 2$ and selected values of α and γ. Theorems 1 and 2 imply that for $c = \widehat{c}_p(\alpha,\gamma)$

defined in (2.12) the test procedure based on $\widehat{V}(m,k)$ has the desired properties (2.2) and (2.3).

Next, we shall have a look at *the test procedures based on* $\widetilde{V}(m,k)$. The corresponding the decision rule is defined as follows:
having observed $Y_1, \ldots, Y_{m+k}$ the null hypothesis H_0 is rejected and the observation is stopped as soon as

$$\widetilde{V}(m,k)\widehat{\sigma}_m^{-2} \geq cq2(k/m),$$

and continue with observations otherwise. Here $c = \widetilde{c}(\alpha, q)$ is determined in such a way that (2.2) holds true and q is a boundary function satisfying (B.2). The limit behavior of the test procedures based on statistics $\widetilde{V}(m,k)$ under the null and alternative hypotheses is stated in Theorems 3 and 4.

Theorem 3. Let $Y_1, Y_2, \ldots,$ follow the model (1.1). Let the assumptions (A.1) – (A.3) and (B.2) be satisfied. Then under H_0

$$\lim_{m\to\infty} P\left(\sup_{1\leq k<\infty} \frac{\widetilde{V}(m,k)}{\widehat{\sigma}_m 2q2(k/m)} \leq x\right) = P\left(\sup_{0\leq t<\infty} \frac{\sum_{i=1}^{p} W2_i(t)}{q2(t)} \leq x\right)$$

for all x, where $\{W_i(t);\, 0 \leq t \leq 1\}$, $i = 1, \ldots, p$ are independent Wiener processes.

Theorem 4. Let $Y_1, Y_2, \ldots,$ follow the model (1.1). Let the assumptions (A.1) – (A.3) and (B.2) be satisfied. Then under the alternatives H_A with

$$\lim_{m\to\infty} m\boldsymbol{\delta}_m^T\boldsymbol{\delta}_m = \infty$$

we have, as $m \to \infty$,

$$\sup_{1\leq k<\infty} \frac{\widetilde{V}(m,k)}{\widehat{\sigma}_m^2 q2(k/m)} \xrightarrow{P} \infty.$$

The proofs are postponed to Section 4.

The critical values $\widetilde{c}_p(\alpha, q)$ can be found as a solution of the equation (w.r.t. c):

$$P\left(\sup_{0\leq t<\infty} \frac{\sum_{i=1}^{p} W2_i(t)}{q2(t)} \leq c\right) = 1 - \alpha. \tag{2.13}$$

The explicit form for the distribution of $\sup_{0\leq t<\infty} \sum_{i=1}^{p} W2_i(t)q^{-2}(t)$ is known only for $p = 1$ and some particular choices of q, for more information see, Chu et al. [1996]. Approximations can be obtained through the simulations of Wiener processes. Theorems 3 and 4 imply that for $c = \widetilde{c}_p(\alpha, q)$ defined in (2.13) the test procedure based on $\widetilde{V}(m,k)$ has the desired properties (2.2) and (2.3).

3 Simulations

A small simulation study was performed in order to obtain approximations for critical values of the introduced test procedures and to get a picture on their behavior under various alternatives.

Table 1 provides approximations to the critical values $\widehat{c}_p(\alpha, \gamma)$ with $p = 2$ defined as

$$P\left(\sup_{1 \leq t \leq 1} \frac{\sum_{i=1}^{p} W_i 2(t)}{t^{2\gamma}} > \widehat{c}_p(\alpha, \gamma)\right) = \alpha,$$

for $\alpha = 0.1, 0.05, 0.025, 0.01$ and various γ. They are based on 10 000 repetitions of simulation from the Wiener process. The supremum over the interval [0,1] is approximated through maximum of 10 000 grid points. The obtained critical values are rounded to 4 decimal places.

$\boldsymbol{\gamma \setminus \alpha}$	0.1	0.05	0.025	0.01
0.00	5.8554	7.1540	8.5462	10.0883
0.05	5.9222	7.3184	8.6439	10.4928
0.10	5.9637	7.2987	8.7480	10.5202
0.15	6.0038	7.3815	8.7675	10.3285
0.20	6.3110	7.7812	9.3889	11.3457
0.25	6.3977	7.7922	9.1359	10.8234
0.30	6.6555	8.0667	9.5175	11.5386
0.35	7.1613	8.6912	10.1804	11.8058
0.40	7.8245	9.3454	10.6790	12.5857
0.49	10.8873	12.4327	14.1495	16.1948

Table 1: Simulated critical values $\widehat{c}_2(\alpha, \gamma)$

Tables 2, 3 and 4 compare behavior of the test procedures based on the test statistics $\widehat{Q}(m, k)$ developed in Horváth et al. [2004] and those based on $\widehat{V}(m, k)$. We consider the simple linear regression model

$$Y_i = \beta_{0,0} + \beta_{1,0} X_i + \delta_0 I_{\{i > k^*\}} + \delta_1 X_i I_{\{i >> k^*\}} + \epsilon_i, \quad i = 1, 2, \ldots$$

where $\{X_i\}$, $i = 1, 2, \ldots$ is either

(a) a sequence of i.i.d. random variables uniformly distributed in the interval $[-\sqrt{3}, \sqrt{3}]$ or

(b) a simple autoregression sequence

$$X_i = 0.5 X_{i-1} + \varepsilon_i,$$

where $\varepsilon_i \sim N(0,1)$ are i.i.d. random variables.

The sequence ϵ_i consists of i.i.d. random variables with normal and Laplace distribution with zero mean and unit variance.

The other parameters were chosen as follows

1. $m = 100$ – size of training data set,
2. $\gamma = 0,\ 0.25,\ 0.49$,
3. change points: $k^* = m/2, m, 2m, 5m$,
4. alternatives

   (i) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (2,1)^T$ – change in the intercept only
   (ii) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (1,1.5)^T$ – small change in the slope only
   (iii) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (1,2)^T$ – moderate change in the slope only
   (iv) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (1,3)^T$ – large change in the slope only.

   Change in both, in the intercept and in the slope, would be detected if either a change in the intercept or a change in the slope would be detected.

For each of the combinations of the parameters we simulated a random sequence of size 10 000, evaluate the two test statistics and calculate the corresponding stopping times $\tau(m)$. We made 2500 repetitions for each of the combinations. In the following tables the selected results for $\alpha = 0.05$ are presented through the summaries of the simulations, i.e., the extremes, median and both quartiles.

In Tables 2 and 3, there are the summaries of the stopping times for selected alternatives, when the sequence $\{X_i\}$ satisfies condition (a). In Table 2 are the results for alternative (i) and in Table 3 the results for alternative (iii).

Both the test statistics, $\widehat{Q}(m,k)$ and $\widehat{V}(m,k)$ behave very well for changes in location, although the statistic $\widehat{Q}(m,k)$ detects the change a bit earlier. One can see that in the case of Laplace error distribution, the stopping times are not so concentrated around their median as are the stopping times when the errors follow the normal distribution. This invites the idea to use another test statistic for the nonnormal errors, for example the test statistic based on L_1-residuals suggested in Koubková [2004].

The results also confirm the pattern of the tuning constant γ mentioned already in the beginning of the paper. Since we do not choose very early change, $\gamma = 0.49$ gives never the best results.

In the case of the alternatives (ii), (iii) and (iv), the changes were detected only with the test statistics $\widehat{V}(m,k)$, not with the statistics $\widehat{Q}(m,k)$. For the small change in the slope (alternative (ii)) neither the test statistics $\widehat{V}(m,k)$ behaves satisfactorily. Such a change is too small to be easily detect.

On the other side, the large change in the slope only (alternative (iv)) is detected by the statistics $\widehat{V}(m,k)$ even better. The test statistics $\widehat{Q}(m,k)$ begin to react to a change in slope only, when the size of the change is somewhere around 6.

For completeness we add also results for the case when the sequence $\{X_i\}$ satisfies condition (b) and for alternative (iii) (see Table 4). The conclusion about the results remains the same.

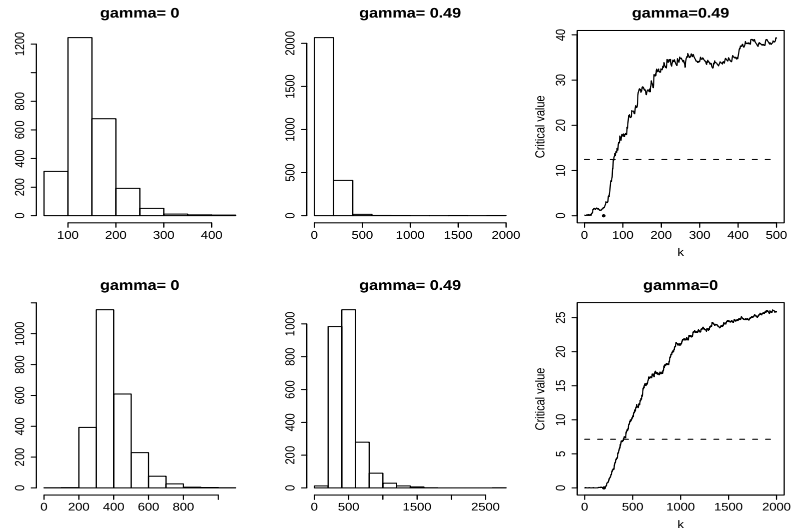

Figure 1: Histograms of stopping times $\tau(m)$ (the time 1 is the first nonhistorical observation) for the alternative (iii), early and late changes and the sequence $\{X_i\}$ satisfying the condition (a) and the behavior of the test statistic $\widehat{V}(m,k)$

To see better how a change is detected, we present also several plots in Figure 1. In the plots we compare influence of small and large γ for alternatives with early and late change. In all the plots we use the parameters $m = 100$, $\alpha = 0.05$, normal error distribution and the sequence $\{X_i\}$ satisfying the condition (a). The upper row corresponds to the early change ($k^* = m/2$) and in the lower row the late change ($k^* = 2m$) was examined. The first and the second columns present the histograms of the stopping times $\tau(m)$ for $\gamma = 0$ and for $\gamma = 0.49$, respectively. The last column shows, how the test statistics $\widehat{V}(m,k)q_\gamma^{-2}(k/m)$ react to the changes. The upper row relates to $\gamma = 0.49$ while the lower one to $\gamma = 0$. In both plots the dashed line denotes the critical value and the dot in the bottom of the plot stands at the time k^*.

Appendix: Proofs of the Results.

In the following, the historical data $Y_1, \ldots, Y_m$ and all the regressors available at the actual time point i, i.e., $\mathbf{X}_1, \ldots, \mathbf{X}_i$ are assumed to be known and so fixed.

We start with several lemmas. The first one contains useful properties of the matrices $\mathbf{C}_m$. In the second and the third ones weak invariance principles for functionals of $\widehat{V}(m,k)$ and $\widetilde{V}(m,k)$, respectively, are formulated. They provide a key tool for the proofs of the theorems from Section 2.

Lemma 3.1. *Let $\boldsymbol{X}_i^T = (1, x_{2i}, \ldots, x_{pi})$, $1 \le i < \infty$, be a strictly stationary sequence of p-dimensional random vectors satisfying (A.3). Then, as $k \to \infty$,*

(a) $|(k\mathbf{C}_k^{-1} - \mathbf{C}^{-1})| = O(k^{-\tau}) \qquad a.s.$

(b) $\mathbf{C}_{m+k} - \mathbf{C}_m = k\mathbf{C} + O(k^{1-\tau})$ *a.s. for each m*

(c) $\limsup_{k\to\infty} k^{-1}|\mathbf{C}_k| < \infty$, *a.s.*

Proof. The assertion (a) follows easily from the assumption (A.3). Towards the assertion (b) it suffices to realize that by the strict stationarity of the sequence of $\boldsymbol{X}_i$, $i = 1, 2, \dots$ we have for any m that

$$(\boldsymbol{C}_{k+m} - \boldsymbol{C}_m,\ k = 1, 2, \dots) \stackrel{D}{=} (\boldsymbol{C}_k,\ k = 1, 2, \dots).$$

Relation (c) is a consequence of (a). □

Lemma 3.2. *Let the assumptions of Theorem 1 be satisfied. Then for each m there exist two independent p-dimensional Wiener processes* $\{\mathbf{W}_{1,m}(t),\ t \in [0,\infty)\}$ *and* $\{\mathbf{W}_{2,m}(t),\ t \in [0,\infty)\}$, *with independent components, such that, as* $m \to \infty$,

$$\sup_{1\le k<\infty} \frac{1}{q_\gamma^2(k/m)} \left| \hat{\sigma}_m^{-2} \widehat{V}(m,k) - m^{-1} \left|\left| \mathbf{W}_{1,m}(k) - \frac{k}{m}\mathbf{W}_{2,m}(m) \right|\right|^2 \right| = o_p(1),$$

where $||.||$ *denotes the Euclidean norm.*

Proof. Since, as $m \to \infty$,

$$\widehat{\sigma}2_m - \sigma 2 = O_P(m^{-1/2})$$

it suffices to show that the assertion of the lemma holds true if $\widehat{\sigma}2_m$ is replaced by $\sigma 2$. Denoting

$$\mathbf{Z}_{0,m} = \sum_{i=1}^{m} \mathbf{X}_i \epsilon_i, \quad \mathbf{Z}_{m,k} = \sum_{i=m+1}^{m+k} \mathbf{X}_i \epsilon_i, \quad k = 1, 2, \dots, \tag{3.1}$$

we realize that under H_0 the test statistics $\widehat{V}(m,k)$ can be written in the form

$$\widehat{V}(m,k) = \left|\left| \mathbf{C}_m^{-1/2} \left(\mathbf{Z}_{m,k} - (\mathbf{C}_{m+k} - \mathbf{C}_m)\mathbf{C}_m^{-1}\mathbf{Z}_{0,m} \right) \right|\right|^2, \ k = 1, 2, \dots. \tag{3.2}$$

By Lemma 3.1.7 in Csörgő and Horváth [1997] we know that if, as $n \to \infty$,

$$\left| n^{-1} \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T - \mathbf{A} \right| = o(\ln^{-\nu} n) \tag{3.3}$$

holds for some symmetric positive definite matrix $\mathbf{A}$ and for some constant $\nu > 0$ and $\{\epsilon_i\}$ is a sequence satisfying (A.1), then there exists a sequence of i.i.d. random normal vectors $\mathbf{N}_i$, $i = 1, 2, \dots$ with $\mathsf{E}\mathbf{N}_i = \mathbf{0}$ and $\mathsf{Cov}\,\mathbf{N}_i = \sigma 2\mathbf{A}$, such that

$$\left|\left| \sum_{i=1}^{n} \mathbf{X}_i \epsilon_i - \sum_{i=1}^{n} \mathbf{N}_i \right|\right| = o(n^{1/2} \ln^{-\lambda} n), \qquad \text{a.s.},$$

with some $\lambda > 1$.

In our case the assumption (A.3) is stronger than (3.3) and therefore for each m we can find a sequence of i.i.d. p-dimensional random vectors $\mathbf{N}_{i,m}$, $i = 1, 2, \dots$ distributed as $N(\mathbf{0}, \mathbf{I}_p)$ such that, as $n \to \infty$,

$$\left|\left| \sum_{i=1}^{n} \mathbf{X}_i \epsilon_i - \sigma \mathbf{C}^{1/2} \sum_{i=1}^{n} \mathbf{N}_{i,m} \right|\right| = o(n^{1/2} \ln^{-\lambda} n), \qquad \text{a.s.} \tag{3.4}$$

Then regarding our assumptions, particularly (A.2), for each m there exist two independent p-dimensional Wiener processes $\{\mathbf{W}_{1,m}(t),\, t \in [0,\infty)\}$ and $\{\mathbf{W}_{2,m}(t),\, t \in [0,\infty)\}$ with independent components such that, as $m \to \infty$,

$$\left|\left|\sum_{i=1}^{m} \mathbf{X}_i\epsilon_i - \sigma\mathbf{C}^{1/2}\mathbf{W}_{2,m}(m)\right|\right| = o(m^{1/2}\ln^{-\lambda} m), \qquad \text{a.s.,} \tag{3.5}$$

and

$$\sup_{1\le k<\infty} k^{-1/2}\ln^{\lambda} k \left|\left|\sum_{i=m+1}^{m+k} \mathbf{X}_i\epsilon_i - \sigma\mathbf{C}^{1/2}\mathbf{W}_{1,m}(k)\right|\right| = o_p(1). \tag{3.6}$$

Hence by Lemma 1 we get, as $m \to \infty$,

$$\sup_{1\le k<\infty} \left|\left|\mathbf{C}_m^{-1/2}\left(\mathbf{Z}_{m,k} - \sigma\mathbf{C}^{1/2}\mathbf{W}_{1,m}(k) - (\mathbf{C}_{m+k} - \mathbf{C}_m)\mathbf{C}_m^{-1}(\mathbf{Z}_{0,m} - \sigma\mathbf{C}^{1/2}\mathbf{W}_{2,m}(m))\right)\right|\right|^2$$
$$\times\left(\sqrt{m/k}\ln^{\lambda} k + (m/k)\ln^{\lambda} m\right)2 = o_P(1).$$

Therefore, to finish the proof, it suffices to show that, as $m \to \infty$,

$$\sup_{1\le k<\infty} B2(m,k)/q2_{\gamma}(k/m) = o_P(1) \tag{3.7}$$

and

$$\sup_{1\le k<\infty} \frac{1}{q_{\gamma}^2(k/m)}\frac{1}{m}\left|\left|\mathbf{W}_{1,m}(k) - \frac{k}{m}\mathbf{W}_{2,m}(m)\right|\right|^2 = O_p(1).$$

where

$$B(m,k) = \sqrt{k/m}\ln^{-\lambda} k + (k/m)\ln^{-\lambda} m, \quad k = 1,2,\ldots.$$

The later relation is a consequence of standard properties of Wiener processes. Elementary calculations give that, as $m \to \infty$,

$$\sup_{1\le k\le m} B2(m,k)/q_{\gamma}^2(k/m) = O_P\Big(\sup_{1\le k\le m}\big\{(k/m)^{1/2-\gamma}\ln^{-\lambda} k + (k/m)^{1-\gamma}\ln^{-\lambda} m\big\}^2\Big) = o_P(1)$$

and

$$\sup_{m\le k<\infty} B2(m,k)/q_{\gamma}^2(k/m) = O_P\Big(\sup_{m\le k<\infty}\big\{(m/k)\ln^{-\lambda} k + \ln^{-\lambda} m\big\}^2\Big) = o_P(1).$$

Hence (3.7) holds true and the proof is finished. □

Lemma 3.3. *Let the assumptions of Theorem 3 be satisfied. Then for each m there exist two independent p-dimensional Wiener processes $\{\mathbf{W}_{1,m}(k),\, k \in [0,\infty)\}$ and $\{\mathbf{W}_{2,m}(k),\, k \in [0,\infty)\}$ with independent components such that, as $m \to \infty$,*

$$\sup_{1\le k<\infty} \frac{1}{q2(k/m)}\Bigg|\widetilde{V}(m,k)$$

$$- m^{-1}\sigma\left|\left|\mathbf{W}_{1,m}(k) - \sum_{i=m+1}^{m+k}\frac{1}{i-1}\mathbf{W}_{1,m}(i-m) - \ln\left(1+\frac{k}{m}\right)\mathbf{W}_{2,m}(m)\right|\right|^2\Bigg| = o_P(1)$$

Proof. Notice that

$$\sum_{i=m+1}^{m+k} \mathbf{X}_i \widetilde{e}_i = \sum_{i=m+1}^{m+k} \mathbf{X}_i \epsilon_i - \sum_{i=m+1}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}_{i-1}^{-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j.$$

For the second term on the r.h.s. one gets, as $m \to \infty$,

$$\begin{aligned} &\sum_{i=m+1}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}_{i-1}^{-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j \\ &\quad = \sum_{i=m+1}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}^{-1} \frac{1}{i-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j + O\left(\sum_{i=m+1}^{m+k} |\mathbf{X}_i \mathbf{X}_i^T| i^{-1-\tau} \left| \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j \right| \right), \end{aligned} \tag{3.8}$$

uniformly in k. The later term on the r.h.s. of (3.8) does not influence the limit behavior of the whole sum (see Horváth et al. [2004], the proof of Lemma 6.1).

Next we proceed as in the proof of Lemma 6.2, Horváth et al. [2004]. By Abel's summation formula we decompose the first term on the r.h.s of (3.8) into four parts

$$\begin{aligned} &\sum_{i=m+1}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \frac{1}{i-1} \mathbf{C}^{-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j \\ &\quad = \sum_{i=m+1}^{m+k} \mathbf{C}_i \mathbf{C}^{-1} \frac{1}{i(i-1)} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j - \sum_{i=m+1}^{m+k} \mathbf{C}_i \mathbf{C}^{-1} \frac{1}{i} \mathbf{X}_i \epsilon_i \\ &\quad + \mathbf{C}_{m+k} \mathbf{C}^{-1} \frac{1}{m+k} \sum_{j=1}^{m+k} \mathbf{X}_j \epsilon_j - \mathbf{C}_m \mathbf{C}^{-1} \frac{1}{m} \sum_{j=1}^{m} \mathbf{X}_j \epsilon_j. \end{aligned}$$

Denoting four terms on the r.h.s. by $Z_1(m,k), Z_2(m,k), Z_3(m,k)$ and $Z_4(m,k)$ and following Horváth et al. [2004] we obtain, as $m \to \infty$,

$$\begin{aligned} Z_1(m,k) &= \sum_{i=m+1}^{m+k} \frac{1}{i-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j + O_P\left(\sum_{i=m+1}^{m+k} \frac{i^{-\tau}}{i-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j \right) \\ Z_2(m,k) &= - \sum_{i=m+1}^{m+k} \mathbf{X}_i \epsilon_i + O_P\left(\sum_{i=m+1}^{m+k} i^{-\tau} \mathbf{X}_i \epsilon_i \right) \\ Z_3(m,k) &= \sum_{i=1}^{m+k} \mathbf{X}_i \epsilon_i + O_P\left((m+k)^{-\tau} \sum_{j=1}^{m+k} \mathbf{X}_j \epsilon_j \right) \\ Z_4(m,k) &= - \sum_{i=1}^{m} \mathbf{X}_i \epsilon_i + O_P\left((m)^{-\tau} \sum_{j=1}^{m} \mathbf{X}_j \epsilon_j \right) \end{aligned}$$

uniformly in k. Here $\tau > 0$ is from the assumption (A.3). Then standard arguments give

$$\sup_{1 \le k \le \infty} |Z_2(m,k) + Z_3(m,k) + Z_4(m,k)| ((m+k)^{-1/2} m^{\tau}) = O_P(1)$$

and

$$\sup_{1 \le k \le \infty} \left| Z_1(m,k) - \sum_{i=m+1}^{m+k} \frac{1}{i-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j \right| \left((m+k)^{-1/2} m^{\tau} \right) = O_P(1).$$

Hence, as $m \to \infty$,

$$\sup_{1\le k\le\infty}\left|\sum_{i=m+1}^{m+k}\mathbf{X}_i\widetilde{e}_i-\sum_{i=m+1}^{m+k}\mathbf{X}_i\epsilon_i-\sum_{i=m+1}^{m+k}\frac{1}{i-1}\sum_{j=1}^{i-1}\mathbf{X}_j\epsilon_j\right|((m+k)^{-1/2}m^{\tau})=O_P(1).$$

Therefore it suffices to prove the assertion of the Lemma with $\widehat{\sigma}_m^{-2}\widetilde{V}(m,k)$ replaced by

$$\sigma^{-2}\left|\left|\mathbf{C}_m^{-1/2}\sum_{i=m+1}^{m+k}\left(\mathbf{X}_i\epsilon_i-\frac{1}{i-1}\sum_{j=1}^{i-1}\mathbf{X}_j\epsilon_j\right)\right|\right|^2. \tag{3.9}$$

Toward this we write

$$\sum_{i=m+1}^{m+k}\left(\mathbf{X}_i\epsilon_i-\frac{1}{i-1}\sum_{j=1}^{i-1}\mathbf{X}_j\epsilon_j\right)=\sum_{i=m+1}^{m+k}\left(\mathbf{X}_i\epsilon_i-\frac{1}{i-1}\sum_{j=m+1}^{i-1}\mathbf{X}_j\epsilon_j\right)-\sum_{i=m+1}^{m+k}\frac{1}{i-1}\sum_{j=1}^{m}\mathbf{X}_j\epsilon_j.$$

Now, similarly as in the proof of Lemma 2 we apply Lemma 3.1.7 in Csörgő and Horváth [1997] and Lemma 1 and we get that for each m there exist two independent p-dimensional Wiener processes $\{\mathbf{W}_{1,m}(k),\, k\in[0,\infty)\}$ and $\{\mathbf{W}_{2,m}(k),\, k\in[0,\infty)\}$ with independent components satisfying the equations (3.5) and (3.6) as $m\to\infty$.

Then proceeding similarly as in the proof of Lemma 6.4 in Horváth et al. [2004] and after a few standard steps we finish the proof of our lemma. □

Now, we prove Theorems 1-4.

Proof of Theorem 1. Proceeding as in the proof of Theorem 2.1 in Horváth et al. [2004], however we have the multidimensional version, we find that, as $m\to\infty$,

$$m^{-1}\left|\left|\mathbf{W}_{1,m}(k)-\frac{k}{m}\mathbf{W}_{2,m}(m)\right|\right|^2\overset{D}{=}\left|\left|\mathbf{W}_{1,m}(k/m)-\frac{k}{m}\mathbf{W}_{2,m}(1)\right|\right|^2,\quad k=1,2,\ldots,$$

$$\sup_{0\le k\le T}\left|\left|\mathbf{W}_{1,m}(k/m)-\frac{k}{m}\mathbf{W}_{2,m}(1)\right|\right|^2 q_\gamma^{-2}(k/m)\;\to\;\sup_{0\le t\le T}||\mathbf{W}_1(t)-t\mathbf{W}_2(1)||^2\,q_\gamma^{-2}(t)$$

for each $T>0$ and

$$\sup_{0\le t\le\infty}||\mathbf{W}_1(t)-t\mathbf{W}_2(1)||^2\,q_\gamma^{-2}(t)\overset{D}{=}\sup_{0\le t\le 1}||\mathbf{W}(t)||^2t^{-2\gamma},$$

where $\{\mathbf{W}_i(t),\, 0\le t<\infty\}$, $i=1,2$ are independent p-dimensional Wiener processes with independent components and $\{\mathbf{W}(t),\, 0\le t<\infty\}$ is also a p-dimensional Wiener process with independent components. These relations together with Lemma 1 imply that for all x

$$\lim_{m\to\infty}P\left(\sup_{1\le k<\infty}\frac{\widehat{V}(m,k)}{\widehat{\sigma}_m^2 q_\gamma^2(k/m)}\le x\right)=P\left(\sup_{0\le t\le 1}\frac{\sum_{i=1}^{p}W2_i(t)}{t^{2\gamma}}\le x\right)$$

and so the Theorem 1 is proved.

Proof of Theorem 2. We can write

$$\sum_{i=m+1}^{m+k}\mathbf{X}_i\widehat{e}_i=\mathbf{Z}_{m,k}-(\mathbf{C}_{m+k}-\mathbf{C}_m)\mathbf{C}_m^{-1}\mathbf{Z}_{0,m}+\sum_{i=m+k^*}^{m+k}\mathbf{X}_i\mathbf{X}_i^T\boldsymbol{\delta}_m I\{k>k^*\},$$

where $\mathbf{Z}_{m,k}$ and $\mathbf{Z}_{0,m}$ are defined in (3.1). The first two terms on the r.h.s. coincide with those under the H_0 and therefore by Theorem 1

$$\max_{1\le k\le T}\left|\left|\mathbf{C}_m^{-1/2}(\mathbf{Z}_{m,k}-(\mathbf{C}_{m+k}-\mathbf{C}_m)\mathbf{C}_m^{-1}\mathbf{Z}_{0,m})\right|\right|^2 q_\gamma^{-2}(k/m)=O_P(1),$$

for each $T>0$. Hence it suffices to show that, as $m\to\infty$,

$$F(m,k_m)=\left(\boldsymbol{\delta}_m^T(\mathbf{C}_{m+k}-\mathbf{C}_{m+k^*})^T\mathbf{C}_m^{-1}(\mathbf{C}_{m+k}-\mathbf{C}_{m+k^*})\boldsymbol{\delta}_m\right)q_\gamma^{-2}(k_m/m)\xrightarrow{P}\infty \tag{3.10}$$

for some $k_m>k^*$. We show that it holds for

$$k_m=2mI\{k^*<m\}+2k^*I\{k^*>m\}.$$

In such a situation Lemma 1 can be applied and we get

$$\begin{aligned}F(m,k_m)&=q_\gamma^{-2}(k_m/m)\\&\times\boldsymbol{\delta}_m^T((k_m-k^*)\mathbf{C}+O(m^{1-\tau}))^T(m^{-1}\mathbf{C}^{-1}+O(m^{-1-\tau}))((k_m-k^*)\mathbf{C}+O(m^{1-\tau}))\boldsymbol{\delta}_m.\end{aligned}$$

For $k^*<m$, $m\to\infty$,

$$F(m,k_m)=\frac{m^{-1}(k_m-k^*)2\boldsymbol{\delta}_m^T\mathbf{C}^T\boldsymbol{\delta}_m}{32\left(\frac{2}{3}\right)^{2\gamma}}+O(m^{1-\tau}||\boldsymbol{\delta}_m||^2)\ge z_1m\boldsymbol{\delta}_m^T\mathbf{C}\boldsymbol{\delta}_m(1+o(1))$$

with some positive constant z_1, and for $k^*>m$

$$F(m,k_m)=\frac{m^{-1}(k_m-k^*)2\boldsymbol{\delta}_m^T\mathbf{C}^T\boldsymbol{\delta}_m}{\left(1+\frac{k_m}{m}\right)2\left(\frac{k_m}{k_m+m}\right)^{2\gamma}}+O(m^{1-\tau}||\boldsymbol{\delta}_m||^2)\ge z_2m\boldsymbol{\delta}_m^T\mathbf{C}\boldsymbol{\delta}_m(1+o(1)),$$

with some positive constant z_2. Hence (3.10) holds true and the proof is finished.

In the proofs of Theorems 3 and 4 we follow the proofs of Theorems 3.1 and 3.2 in Horváth et al. [2004] and concentrate only on differences.

Proof of Theorem 3. Due to Lemma 3 it suffices to derive the limit distribution of

$$\sup_{1\le k<\infty}\frac{1}{q2(k/m)}\left|\left|\mathbf{C}_m^{-1/2}\mathbf{C}^{1/2}\left(\mathbf{W}_{1,m}(k)-\sum_{i=m+1}^{m+k}\frac{1}{i-1}\mathbf{W}_{1,m}(i-m)-\sum_{i=m+1}^{m+k}\frac{1}{i-1}\mathbf{W}_{2,m}(m)\right)\right|\right|^2, \tag{3.11}$$

where the Wiener processes $\mathbf{W}_{i,m}$, $i=1,2$ are defined in Lemma 3. By standard properties of Wiener processes, as $m\to\infty$,

$$\sup_{1\le k<\infty}\left|\left|\mathbf{W}_{1,m}(k)-\sum_{i=m+1}^{m+k}i^{-1}\mathbf{W}_{1,m}(i-m)-\ln\left(1+\frac{k}{m}\right)\mathbf{W}_{2,m}(m)\right|\right|^2(m^{-1}q^{-2}(k/m))$$

converges in distribution to

$$\sup_{0\le t<\infty}\frac{1}{q2(t)}\left|\left|\mathbf{W}(t+1)-\mathbf{W}(1)-\int_0^t\frac{1}{u+1}(\mathbf{W}(u+1)-\mathbf{W}(1))du-\mathbf{W}(1)\ln(1+t)\right|\right|^2,$$

where $\{\mathbf{W}(t),\, 0 \le t < \infty\}$ is a p-dimensional Wiener process with independent components. By direct calculations we find that

$$\mathbf{W}(t+1) - \mathbf{W}(1) - \int_0^t \frac{1}{u+1}(\mathbf{W}(u+1) - \mathbf{W}(1))du - \mathbf{W}(1)\ln(1+t) \stackrel{D}{=} \mathbf{W}^*(t)$$

whenever $0 \le t < \infty$ and where $\{\mathbf{W}^*(t),\, 0 \le t < \infty\}$ is a p-dimensional Wiener process with independent components. The proof is finished.

Proof of Theorem 4. We have

$$\sum_{i=m+1}^{m+k} \mathbf{X}_i \widetilde{e}_i = \sum_{i=m+1}^{m+k} \mathbf{X}_i \epsilon_i - \sum_{i=m+1}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}_{i-1}^{-1} \sum_{j=1}^{i-1} \mathbf{X}_j \epsilon_j$$
$$+ \sum_{m+k^*}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}_{i-1}^{-1} \mathbf{C}_{m+k^*} \boldsymbol{\delta}_m I\{k > k^*\}, \quad k = 1, 2, \ldots.$$

By Theorem 3 we obtain, as $m \to \infty$,

$$\sup_{1 \le k < \infty} \left|\left| \sum_{i=m+1}^{m+k} \mathbf{X}_i \epsilon_i - \sum_{i=m+1}^{m+k} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}_{i-1}^{-1} \sum_{j=1}^{i-1} \; boldX_j \epsilon_j \right|\right|^2 q^{-2}(k/m) = O_P(1).$$

Thus it suffices to show that, as $m \to \infty$,

$$\left|\left| \mathbf{C}_m^{-1/2} \sum_{m+k^*}^{m+k_m} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}_{i-1}^{-1} \mathbf{C}_{m+k^*} \boldsymbol{\delta}_m \right|\right|^2 q^{-2}(k_m/m) \xrightarrow{P} \infty \tag{3.12}$$

for for suitably chosen $k_m > k^*$. Noticing that for $k_m = 2mI\{k^* < m\} + 2k^* I\{k^* > m\}$, we have

$$\left|\left| \mathbf{C}_m^{-1/2} \sum_{m+k^*}^{m+k_m} \mathbf{X}_i \mathbf{X}_i^T \mathbf{C}_{i-1}^{-1} \mathbf{C}_{m+k^*} \boldsymbol{\delta}_m \right|\right|^2 q^{-2}(k_m/m) \le \left|\left| \mathbf{C}_m^{-1/2} (\mathbf{C}_{m+k_m} - \mathbf{C}_{m+k^*}) \boldsymbol{\delta}_m \right|\right|^2 q^{-2}(k_m/m)$$

and the proof can be finished similarly as that of Theorem 2.

Acknowldgements

The work was supported by grants GAČR 201/03/0945, MSM 113200008, NATO PST.EAP.CLG980599 and GAČR 205/05/H007.

References

[1] Aue, A. (2003). *Sequential Change-Point Analysis Based on Invariance Principles*, PhD Thesis written on der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln.

[2] Chu, C.-S., Stinchcombe, M. and White H. (1996). *Monitoring Structural Change*, Econometrica, Vol. **64**, No. **5**, 1045-1065.

[3] Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics.* John Willey & Sons, Chichester.

Koubková, A. (2004). *Critical values for monitoring changes in linear regression models*, COMPSTAT 2004 Proceedings, Springer Verlag, 1345-1352.

[4] Horváth, L., Hušková, M., Kokoszka P. and Steinebach J. (2004). *Monitoring Changes in Linear Models*, Journal of statistical Planning and Inference, **126**, 225-251.

[5] Leisch, F., Hornik, K. and Kuan, Ch.-M. (2000). *Monitoring Structural Changes with the Generalized Fluctuation Test*, Econometric Theory, **16**, 835-854.

[6] Zeileis, A., Leisch, F., Kleiber, Ch. and Hornik, K. (2005). *Monitoring Structural Change in Dynamic Econometric Models* J. Appl. Econ. **20**, 99 – 121.

The alternative (i) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (2,1)^T$, normal error distribution											
k^*	γ	$\widehat{Q}(m,k)$					$\widehat{V}(m,k)$				
		min	1^{st}Q	med	3^{st}Q	max	min	1^{st}Q	med	3^{st}Q	max
$m/2$	0	60	104	119	137.3	256	68	118	136	160	342
	0.25	55	94	110	128.0	253	45	104.0	123.0	145	317
	0.49	55	102.0	121	145.0	356	1	113.0	136	168.0	459
m	0	117	169	192	218.0	437	116	188	216	249	582
	0.25	108	163	186	212.0	442	68	178.0	205.0	238	582
	0.49	117	182.0	212	248.0	718	1	200.0	238	285.3	924
$2m$	0	218	302	338	380.3	716	134	327	372	423	894
	0.25	212	301	338	382.0	731	126	321.8	367.0	419	908
	0.49	8	340.0	389	452.3	1076	1	372.0	433	511.0	2023
$5m$	0	176	701	777	870.0	1633	55	749	844	951	2040
	0.25	103	711	791	888.3	1784	27	754.0	853.5	965	2141
	0.49	411	813.8	924	1058.0	2485	1	881.8	1021	1191.0	5016

The alternative (i) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (2,1)^T$, Laplace error distribution											
		$\widehat{Q}(m,k)$					$\widehat{V}(m,k)$				
		min	1^{st}Q	med	3^{st}Q	max	min	1^{st}Q	med	3^{st}Q	max
$m/2$	0	62	116	140	176	1289	68	134	166.0	212	4770
	0.25	10	105	129	163.0	2353	10	118.0	148	193.0	5920
	0.49	1	117.0	150	206.0	10000	1	132.0	177	254.0	10000
m	0	55	185	222	273	1925	85	207	254.0	318	4008
	0.25	20	178	216	268.0	2570	13	195.0	241	306.3	5327
	0.49	1	205.0	259	339.3	10000	1	224.0	296	407.0	10000
$2m$	0	119	327	384	466	5368	78	361	432.0	540	3906
	0.25	28	326	385	472.3	9525	2	355.0	427	540.0	9238
	0.49	1	383.0	470	618.3	10000	1	422.8	542	744.3	10000
$5m$	0	100	745	876	1059	6794	53	808	968.5	1204	10000
	0.25	48	760	900	1102.0	10000	1	815.8	985	1240.0	10000
	0.49	1	904.8	1110	1455.0	10000	1	988.0	1272	1754.0	10000

Table 2: Simulated quantiles of the stopping times $\tau(m)$ (the time 1 is the first nonhistorical observation), for the alternative (i), when the sequence $\{X_i\}$ satisfies condition (a)
.

The alternative (iii) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (1,2)^T$, normal error distribution											
k^*	γ	$\widehat{Q}(m,k)$					$\widehat{V}(m,k)$				
		min	1^{st}Q	med	3^{st}Q	max	min	1^{st}Q	med	3^{st}Q	max
$m/2$	0	77	10000	10000	10000	10000	63	113	137	167.3	448
	0.25	56	10000	10000	10000	10000	55	101	123	151.3	448
	0.49	56	10000	10000	10000	10000	1	108.0	138.0	179.0	1924
m	0	154	10000	10000	10000	10000	113	182	216	258.0	745
	0.25	87	10000	10000	10000	10000	103	173	205	248.0	748
	0.49	11	10000	10000	10000	10000	1	192.0	237.0	302.3	1852
$2m$	0	111	10000	10000	10000	10000	98	320	371	439.0	1098
	0.25	55	10000	10000	10000	10000	23	314	367	435.0	1160
	0.49	55	10000	10000	10000	10000	1	360.0	431.5	536.0	2662
$5m$	0	115	10000	10000	10000	10000	127	739	846	984.0	2859
	0.25	104	10000	10000	10000	10000	66	742	854	1000.0	3005
	0.49	5	10000	10000	10000	10000	1	865.8	1025.0	1244.0	10000

The alternative (iii) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (1,2)^T$, Laplace error distribution											
		$\widehat{Q}(m,k)$					$\widehat{V}(m,k)$				
		min	1^{st}Q	med	3^{st}Q	max	min	1^{st}Q	med	3^{st}Q	max
$m/2$	0	54	10000	10000	10000	10000	58	130.0	168	226	10000
	0.25	26	10000	10000	10000	10000	1	114.0	151	206.0	10000
	0.49	1	10000	10000	10000	10000	1	124	177	276.0	10000
m	0	76	10000	10000	10000	10000	100	201.0	254	330	10000
	0.25	14	10000	10000	10000	10000	1	191.0	240	319.0	10000
	0.49	1	10000	10000	10000	10000	1	215	296	442.0	10000
$2m$	0	114	10000	10000	10000	10000	113	349.0	432	558	10000
	0.25	13	10000	10000	10000	10000	5	341.0	427	557.3	10000
	0.49	1	10000	10000	10000	10000	1	402	541	793.5	10000
$5m$	0	110	10000	10000	10000	10000	76	786.8	958	1234	10000
	0.25	19	10000	10000	10000	10000	12	793.8	978	1273.0	10000
	0.49	1	10000	10000	10000	10000	1	947	1249	1850.0	10000

Table 3: Simulated quantiles of the stopping times $\tau(m)$ (the time 1 is the first nonhistorical observation), for the alternative (iii), when the sequence $\{X_i\}$ satisfies the condition (a).

The alternative (iii) $\boldsymbol{\beta}_0 = (1,1)^T, \boldsymbol{\beta}_1 = (1,2)^T$, normal error distribution											
k^*	γ	$\widehat{Q}(m,k)$					$\widehat{V}(m,k)$				
		min	1^{st}Q	med	3^{st}Q	max	min	1^{st}Q	med	3^{st}Q	max
$m/2$	0	83	10000	10000	10000	10000	52	102.0	125	156	480
	0.25	55	10000	10000	10000	10000	27	90	112.0	142	480
	0.49	55	10000	10000	10000	10000	1	94	121	162.0	2445
m	0	131	10000	10000	10000	10000	78	169.0	201	243	688
	0.25	78	10000	10000	10000	10000	6	160	192.0	234	693
	0.49	134	10000	10000	10000	10000	1	175	217	276.0	1149
$2m$	0	135	10000	10000	10000	10000	80	301.0	346	408	905
	0.25	77	10000	10000	10000	10000	21	295	341.0	403	1116
	0.49	81	10000	10000	10000	10000	1	330	392	486.3	5360
$5m$	0	365	10000	10000	10000	10000	67	704.8	809	938	2449
	0.25	122	10000	10000	10000	10000	19	708	816.5	954	2746
	0.49	1532	10000	10000	10000	10000	1	799	954	1165.0	10000

Table 4: Simulated quantiles of the stopping times $\tau(m)$ (the time 1 is the first nonhistorical observation), for the alternative (iii), when the sequence $\{X_i\}$ satisfies the condition (b).