Communications of the Korean Statistical Society
2012, Vol. 19, No. 3, 423–432

DOI: http://dx.doi.org/10.5351/CKSS.2012.19.3.423

Multiple Structural Change-Point Estimation in Linear Regression Models

Jaehee Kim[1,a]

[a]Department of Statistics, Duksung Women's University

Abstract

This paper is concerned with the detection of multiple change-points in linear regression models. The proposed procedure relies on the local estimation for global change-point estimation. We propose a multiple change-point estimator based on the local least squares estimators for the regression coefficients and the split measure when the number of change-points is unknown. Its statistical properties are shown and its performance is assessed by simulations and real data applications.

Keywords: Change-point model, local least squares estimator.

1. Introduction

Both statistics and econometrics literature contain a vast amount of change-point issues related to structural change. Most of the work in the literature is designed for a single change. The occurrence of a single change-point in real data is rare, as data in economics, finance, and biology display multiple changes. Thus, a statistical procedure able to reliably detect multiple changes is of practical interest.

To test the null hypothesis that regression coefficients are constant over time, Brown *et al.* (1975) introduced a CUSUM test based on cumulated sums of recursive residuals. Ploberger and Kr̈amer (1992) proposed a CUSUM test based on recursive and least-squares residuals. Bauer and Hackl (1978) proposed a MOSUM test based on moving sums of recursive residuals. Chu *et al.* (1995) proposed a least-squares-MOSUM test based on least-squares residuals. Andrews (1993) and Andrews and Ploberger (1994) proposed tests for parameter constancy in linear models.

After rejecting the null hypothesis with no change model, each change-point should be estimated. In the case of multiple changes, the problem is more intricate and few approaches are dedicated to this problem.

There are change-point researches with the Bayesian approach since the Bayesian framework enables the formal incorporation of uncertainty regarding the parameterization of theoretical models. Broemeling and Tsurumi (1987) studied Bayesian structural change models with one change-point and its application of MCMC and BIC criterion has been studied by Lavielle and Lebarbier (2001). Chib (1998) provided a Bayesian approach for a multiple change-points model using a latent discrete state variable. Koop and Porter (2004) developed a change-point model with regime duration using a MCMC sampler for conditional mean and variance. Recently Kim and Cheon (2010a) derived the Bayesian posterior distribution for multiple change-point estimation in several parametric distributions; in addition, Kim and Cheon (2010b) also developed a Bayesian regime-switching model

This research was supported by Basic Science Research Program through the National Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0011094).

[1] Professor, Department of Statistics, Duksung Women's University, Seoul 132-714, South Korea.
Email: jaehee@duksung.ac.kr

with multiple change-points. However, there are difficulties that arise in working with fully specified Bayesian models because the exact form of the associated likelihood function is generally unknown and the posterior analysis cannot be carried out analytically.

Dufour and Ghysels (1996) focused the importance of a structural change model that reflect dynamic state in econometric models. Loader (1996) proposed a change-point estimator using nonparametric regression. McDonald and Owen (1986) considered change-point estimation using three split linear fits at each point. Bai and Perron (1998, 2003) proposed a multiple change-point estimation based on the minimization of the residual sum of squares on possible partitions; however, their method cannot be applicable to other types of underlying functions. Perron and Zhu (2005) analyzed the limiting distributions of parameter estimates when the trend function exhibits a one slope change. Gregoire and Hamrouni (2002) considered estimating the jump in a smooth curve using local linear smoothing. Kim and Hart (2010) proposed one change-point estimation with the left and the right Local Fourier estimators. One change-point detection methods are provided in various techniques; however, multiple change-point estimation methods are relatively rare in regression models. The problem is more complicated when the number of changes is unknown, and a few papers are dedicated to this problem.

This paper provides a multiple change-point estimation method where there are several change-points in the linear regression models. We use the divergence measure to divide segments based on the difference of the local least squares estimators(LSE) for regression parameters and a split measure at each possible change-point.

Our method is a new simultaneous approach computationally simple and differs from most of the techniques suggested elsewhere (Bai and Perron, 1998, 2003). The novelty of our method is the use of local parameter coefficient estimators to detect changes in regression parameters. Our method is a universal approach because the change-points are simultaneously estimated and is also local in the sense that detection is based on local least squares estimators within the bandwidth.

The paper is organized as follows. We propose a new multiple structural change-point estimation method using local LSE in Section 2. The properties of the proposed estimator are studied via simulation in Section 3. As an illustration, we also include two examples in Section 4. Finally Section 5 concludes the paper.

2. Proposed Multiple Change-Point Estimation Method

2.1. Parameter change detection with local least squares estimators

We consider the one change-point linear regression model with the change-points at τ as follows:

$$Y_i = \begin{cases} \mathbf{x}_i'\boldsymbol{\beta}_1 + \epsilon_i, & i = 1, \ldots, \tau, \\ \mathbf{x}_i'\boldsymbol{\beta}_2 + \epsilon_i, & i = \tau + 1, \ldots, T. \end{cases} \tag{2.1}$$

where the explanatory vector $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \ldots, x_{ip})'$, and regression parameter vectors $\boldsymbol{\beta}_1 = (\beta_{10}, \beta_{11}, \ldots, \beta_{1p})'$, $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21}, \ldots, \beta_{2p})'$ and $\epsilon_i \sim iid\, N(0, \sigma^2)$. Let h be a positive bandwidth that is less than 1/2. At point t, $\hat{\boldsymbol{\beta}}_{1t}$ is defined as the local LSE of the data left at t within $[t - Th, t]$ with the design matrix $\mathbf{X}_1$ and $\hat{\boldsymbol{\beta}}_{2t}$ is the local LSE of the data right at t within $[t + 1, t + Th + 1]$ with the design matrix $\mathbf{X}_2$. The difference of the left and right local LSEs is

$$\mathbf{Z}_t = \hat{\boldsymbol{\beta}}_{1t} - \hat{\boldsymbol{\beta}}_{2t} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{Y}_1 - (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{Y}_2$$

and its variance is obtained as

$$\mathbf{V}_t = V\left(\hat{\boldsymbol{\beta}}_{1t} - \hat{\boldsymbol{\beta}}_{2t}\right) = V\left(\hat{\boldsymbol{\beta}}_{1t}\right) + V\left(\hat{\boldsymbol{\beta}}_{2t}\right)$$
$$= \left\{(\mathbf{X}_1'\mathbf{X}_1)^{-1} + (\mathbf{X}_2'\mathbf{X}_2)^{-1}\right\}\sigma^2.$$

Therefore, we consider the unbiased estimator of $\mathbf{V}_t$ as

$$\hat{\mathbf{V}}_t = \left\{(\mathbf{X}_1'\mathbf{X}_1)^{-1} + (\mathbf{X}_2'\mathbf{X}_2)^{-1}\right\}\hat{\sigma}^2.$$

One possible approach is to use the idea of differencing to remove trends for the variance estimator. Consider the first-order difference-based estimator proposed by Rice (1984) such as

$$\hat{\sigma}^2 = \frac{1}{2(T-1)}\sum_{i=2}^{T}(Y_i - Y_{i-1})^2. \tag{2.2}$$

Consider the following parameter divergence measure

$$Q_t = \mathbf{Z}_t'\hat{\mathbf{V}}_t^{-1}\mathbf{Z}_t \tag{2.3}$$

that follows the chi-squared distribution under $H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}$. At the true change-point $t = \tau$, Q_τ follows the noncentral chi-squared distribution with the noncentral parameter δ^2.

Without loss of generality, we assume that there are enough observations near the true change-points so that they can be identified by local estimation.

For detecting changes in regression parameters, we propose a change-point detector based on

$$D_t = Q_t + \mathrm{spl}_t. \tag{2.4}$$

Here Q_t measures the difference between the left and the right estimators of regression coefficients and spl_t is a derivative measure at the point t.

At t, let $\hat{Y}_{t+1}$ be the least squares fitted estimate at $t+1$ with the data in the right window and $\hat{Y}_t$ is the least squares fitted estimate at t with the data in the left window. With $\mathbf{Y}_{2t} = (Y_{t+1}, Y_{t+2}, \ldots, Y_{t+Th})'$ and its corresponding $\mathbf{X}_{2t}$ and $\mathbf{Y}_{1t} = (Y_t, Y_{t-1}, \ldots, Y_{t-Th-1})'$ and its corresponding $\mathbf{X}_{1t}$ we have

$$\hat{Y}_{t+1} = \mathbf{x}_{t+1}'\hat{\boldsymbol{\beta}}_{2t} = \mathbf{x}_{t+1}'\left(\mathbf{X}_{2t}'\mathbf{X}_{2t}\right)^{-1}\mathbf{X}_{2t}'\mathbf{Y}_{2t},$$

and

$$\hat{Y}_t = \mathbf{x}_t'\hat{\boldsymbol{\beta}}_{1t} = \mathbf{x}_t'\left(\mathbf{X}_{1t}'\mathbf{X}_{1t}\right)^{-1}\mathbf{X}_{1t}'\mathbf{Y}_{1t}.$$

We define spl_t as

$$\mathrm{spl}_t = \frac{d_t^2}{A_t\hat{\sigma}^2} \tag{2.5}$$

with

$$d_t = \hat{Y}_{t+1} - \hat{Y}_t$$

and

$$A_t = \mathbf{x}'_{t+1}\left(\mathbf{X}'_{2t}\mathbf{X}_{2t}\right)^{-1}\mathbf{x}_{t+1} + \mathbf{x}'_t\left(\mathbf{X}'_{1t}\mathbf{X}_{1t}\right)^{-1}\mathbf{x}_t.$$

Since the split measure should be maximized at the true change-point, we consider the split term at τ

$$\begin{aligned} d_\tau &= \hat{Y}_{\tau+1} - \hat{Y}_\tau \\ &= \mathbf{x}'_{\tau+1}\hat{\boldsymbol{\beta}}_{2\tau} - \mathbf{x}'_\tau\hat{\boldsymbol{\beta}}_{1\tau} \\ &= \mathbf{x}'_{\tau+1}\left(\hat{\boldsymbol{\beta}}_{2\tau} - \boldsymbol{\beta}_2\right) - \mathbf{x}'_\tau\left(\hat{\boldsymbol{\beta}}_{1\tau} - \boldsymbol{\beta}_1\right) + \mathbf{x}'_{\tau+1}\left(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_1\right) + \left(\mathbf{x}_{\tau+1} - \mathbf{x}_\tau\right)'\boldsymbol{\beta}_1. \end{aligned}$$

Let $\mathbf{x}_t = (1, x_{1t}, \ldots, x_{pt})'$ with $x_{kt} = t/T$, $k = 1, 2, \ldots, p$, and set $\boldsymbol{\beta}_2 = \boldsymbol{\beta}_1 + \boldsymbol{\Delta}$.

Consider expectation of d_τ and d_τ^2 as

$$\begin{aligned} E[d_\tau] &= E\left[\hat{Y}_{\tau+1} - \hat{Y}_\tau\right] \\ &= \mathbf{x}'_{\tau+1}\left(\mathbf{X}'_{2\tau}\mathbf{X}_{2\tau}\right)^{-1}\mathbf{X}'_{2\tau}\left(\mathbf{X}_{2\tau}\boldsymbol{\beta}_2 + E\boldsymbol{\epsilon}_{\tau+1}\right) - \mathbf{x}'_\tau\left(\mathbf{X}'_{1\tau}\mathbf{X}_{1\tau}\right)^{-1}\mathbf{X}'_{1\tau}(\mathbf{X}_{1\tau}\boldsymbol{\beta}_{1\tau} + E\boldsymbol{\epsilon}_\tau) \\ &= \mathbf{x}'_{\tau+1}\boldsymbol{\beta}_2 - \mathbf{x}'_\tau\boldsymbol{\beta}_1 \\ &= \mathbf{x}'_{\tau+1}(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_1) + (\mathbf{x}_{\tau+1} - \mathbf{x}_\tau)'\boldsymbol{\beta}_1 \end{aligned}$$

and

$$E\left[d_\tau^2\right] = \mathbf{x}'_{\tau+1}\boldsymbol{\Delta}'\boldsymbol{\Delta}\mathbf{x}_{\tau+1} + A_\tau\sigma^2 + O\left(\frac{1}{T}\right)$$

where

$$A_\tau = \mathbf{x}'_{\tau+1}\left(\mathbf{X}'_{2\tau}\mathbf{X}_{2\tau}\right)^{-1}\mathbf{x}_{\tau+1} + \mathbf{x}'_\tau\left(\mathbf{X}'_{1\tau}\mathbf{X}_{1\tau}\right)^{-1}\mathbf{x}_\tau.$$

Therefore, the split measure has the more weight at τ with

$$\mathrm{spl}_\tau = \frac{d_\tau^2}{A_\tau\hat{\sigma}^2}.$$

The proposed method requires a careful inspection of the sequence of $\{D_t\}$; however, it is difficult to automate in order to estimate unknown number of structural change-points. The idea of this method is to model how the sequence $\{D_t\}$ decreases when there is a change-point and to look for the possible change-point.

Theorem 1. *When there is no change-point, Q_t follows the chi-square distribution with the degree of freedom $p + 1$ as $T \to \infty$.*

Proof: When there is no change-point, Q_t follows the chi-square distribution with the degree of freedom $p + 1$ as $T \to \infty$.

$$\begin{aligned} Q_t &= \mathbf{Z}'_t\hat{\mathbf{V}}_t^{-1}\mathbf{Z}_t \\ &= \mathbf{Z}'_t\mathbf{V}_t^{-1}\mathbf{Z}_t + \mathbf{Z}'_t\left(\hat{\mathbf{V}}_t^{-1} - \mathbf{V}_t^{-1}\right)\mathbf{Z}_t \\ &= \left[\left(\hat{\boldsymbol{\beta}}_{1t} - \boldsymbol{\beta}\right) - \left(\hat{\boldsymbol{\beta}}_{2t} - \boldsymbol{\beta}\right)\right]'\mathbf{V}_t^{-1}\left[\left(\hat{\boldsymbol{\beta}}_{1t} - \boldsymbol{\beta}\right) - \left(\hat{\boldsymbol{\beta}}_{2t} - \boldsymbol{\beta}\right)\right] + O_p(1) \\ &\to (\mathbf{U}_1 - \mathbf{U}_2)'\mathbf{V}_t^{-1}(\mathbf{U}_1 - \mathbf{U}_2) + O_p(1), \end{aligned}$$

where $\mathbf{U}_1$ and $\mathbf{U}_2$ are iid multivariate normal distributed random variable with mean vector $\mathbf{0}$. Therefore $Q_t \to \chi^2_{p+1}$ as $T - p \to \infty$. □

If there is no change-point in the linear model

$$E[D_t] = E[Q_t] + E[\text{spl}_t] = p + 1 + O\left(\frac{1}{T}\right)$$

since there is no difference in the regression parameters. At a change-point τ

$$E[D_\tau] = E[Q_\tau] + E[\text{spl}_\tau] = p + 1 + \delta^2 + g(\mathbf{\Delta}'\mathbf{\Delta}) + O\left(\frac{1}{T}\right),$$

where δ^2 is the noncentral parameter of noncentral χ^2_p and $g(\mathbf{\Delta}'\mathbf{\Delta})$ is a function of $\mathbf{\Delta}'\mathbf{\Delta}$ and the value of g is positive.

We consider the k structural change-points in the linear regression model at $c_1 \le c_2 \le \cdots \le c_k$ with unknown k.

$$Y_i = \begin{cases} \mathbf{x}_i'\boldsymbol{\beta}_1 + \epsilon_i, & i = 1, \ldots, c_1, \\ \mathbf{x}_i'\boldsymbol{\beta}_2 + \epsilon_i, & i = c_1 + 1, \ldots, c_2, \\ \quad\vdots & \quad\vdots \\ \mathbf{x}_i'\boldsymbol{\beta}_k + \epsilon_i, & i = c_{k-1} + 1, \ldots, c_k, \\ \mathbf{x}_i'\boldsymbol{\beta}_{k+1} + \epsilon_i, & i = c_k + 1, \ldots, T. \end{cases} \tag{2.6}$$

Assuming that no other change-point occurs within the bandwidth of the estimated change-point, our multiple change-point estimation procedure is as follows:

Step 1. Calculate D_t for possible range of change-points.

Step 2. Arrange D_t in the decreasing order as $\{D_{(1)}, D_{(2)}, \ldots, D_{(T^*)}\}$, where T^* is the number of possible change-points. Searching is done in this order.

Step 3. Select $D_{c_1} = D_{(1)}$ and the first change-point is c_1. Let $m < [Th]$. Search the second change-point c_2 with $D_{c_2} \notin (D_{c_1-m}, D_{c_1+m})$. That is to avoid the change-point within $(c_1 - m, c_1 + m)$. If $D_{(2)} \notin (D_{c_1-m}, D_{c_1+m})$, then $D_{c_2} = D_{(2)}$. If $D_{(2)} \in (D_{c_1-m}, D_{c_1+m})$, then disregard $D_{(2)}$ and find D_{c_2} until $D_{c_2} \notin (D_{c_1-m}, D_{c_1+m})$. Likewise search the third change-point $D_{c_3} \notin (D_{c_1-m}, D_{c_1+m})$ and $\notin (D_{c_2-m}, D_{c_2+m})$. Repeat this process until the possible change-points are obtained. The search procedure is summarized as follows

$$\begin{aligned} \hat{c}_1 &= \arg\max_{t\in[m,T-m]}\{D_t\} \\ \hat{c}_2 &= \arg\max_{t\notin[\hat{c}_1-m,\hat{c}_1+m]}\{D_t\} \\ \hat{c}_3 &= \arg\max_{t\notin[\hat{c}_1-m,\hat{c}_1+m],\, t\notin[\hat{c}_2-m,\hat{c}_2+m]}\{D_t\} \\ &\vdots \\ \hat{c}_k &= \arg\max_{t\notin[\hat{c}_1-m,\hat{c}_1+m],\ldots,t\notin[\hat{c}_{k-1}-m,\hat{c}_{k-1}+m]}\{D_t\} \end{aligned}$$

where $m = [Th]$ or $m = [Th/2]$ in practice.

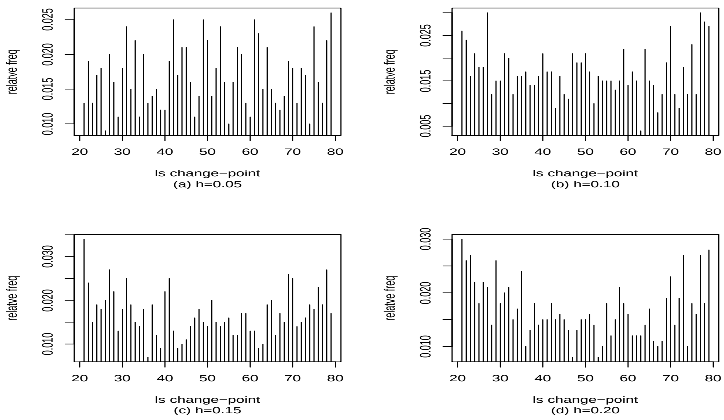

Figure 1: *Relative frequency of one change-point estimates in no change-point model with the bandwidth* $h = 0.05, 0.10, 0.15, 0.2$ *and* $\sigma = 0.5$

3. Simulation

In this section, some simulation experiments are performed to assess the performance of the proposed estimation method. The models considered are as follows:

(i) Simple model with no change with $\beta = 1$

$$Y_i = \beta x + \epsilon_i, \quad i = 1, \ldots, T.$$

(ii) Single structural change model with $\Delta_1 = 1.0$, $\Delta_2 = 0.3$, $\tau = 50$

$$Y_i = \begin{cases} \beta x_i + \epsilon_i, & i = 1, \ldots, \tau, \\ \Delta_1 + (\beta + \Delta_2) x_i + \epsilon_i, & i = \tau + 1, \ldots, T. \end{cases}$$

(iii) Two structural change-point model with $\Delta = 0.5$, $\tau_1 = 30$, $\tau_2 = 60$

$$Y_i = \begin{cases} \beta x_i + \epsilon_i, & i = 1, \ldots, \tau_1, \\ (\beta + \Delta) x_i + \epsilon_i, & i = \tau_1 + 1, \ldots, \tau_2, \\ \beta x_i + \epsilon_i, & i = \tau_2 + 1, \ldots, T, \end{cases}$$

where $x_i = i/T$ and $\epsilon \sim iid\, N(0, \sigma^2)$.

The simulation was done in 1,000 repetitions with the sample size $T = 100$ and $\sigma = 0.5$ and 1.0. For the local LSE, the bandwidth $h = 0.05$, 0.1, 0.15 and 0.2 are used to investigate the effects of the window size.

We do not consider the points within 20% of the right and the left edges of the points according to the prior information and the bandwidth constraint. The relative frequency in figures is calculated as the relative frequency of estimated change-points over 1,000 repetitions. Figure 1 shows the frequency with no change-point model (i) with $\sigma = 0.5$ and it does not give clear peaks or gives too many similar peaks. Figure 2 and Figure 4 show the relative frequency with the one change-point model (ii) with respectively $\sigma = 0.5$ and $\sigma = 1.0$ and they peak at the true change-point. Figure 3 with $\sigma = 0.5$ and Figure 5 with $\sigma = 1.0$ show the relative frequency with two change-point model (iii) and they

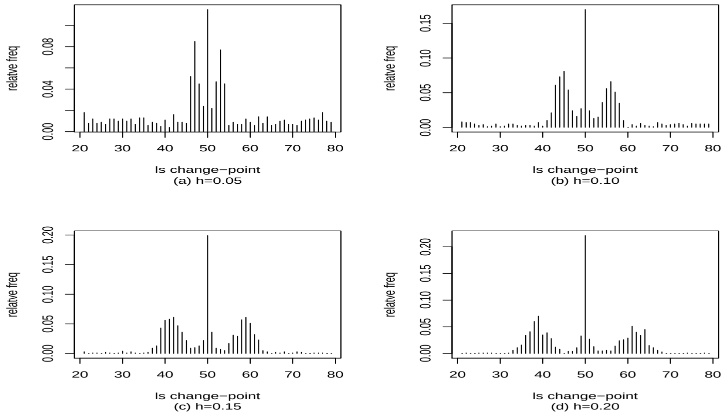

Figure 2: *Relative frequency of one change-point estimates in one change-point model (i) with the bandwidth* $h = 0.05, 0.10, 0.15, 0.2$, $\sigma = 0.5$ *and change-point at* $\tau = 50$

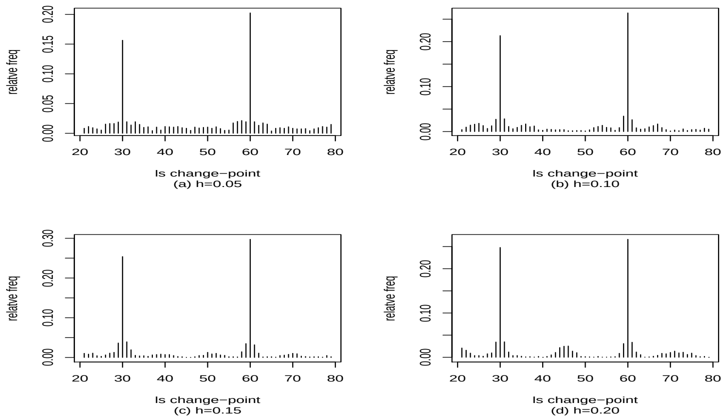

Figure 3: *Relative frequency of two change-point estimates in a two change-point model (ii) with the bandwidth* $h = 0.05, 0.10, 0.15, 0.2$, $\sigma = 0.5$ *and change-points at* $\tau = 30, 60$

give one high peak at the one true change-point and another peak at the other true change-point. The overall trend is similar with $\sigma = 0.5$ and $\sigma = 1.0$ except the more variability with the bigger variance. The size of the bandwidth does not significantly influence the change-point estimation result in these examples. However, the bandwidth selection should be considered carefully based on the data. We do not discuss further about the bandwidth problem.

4. Adaptation in the Example

We applied our method to find the change-point with annual GDP data in the U.S. from 1870 to 2009. For empirical analysis, the logarithm of the data was used. One change-point estimation with the bandwidth $h = 0.05, 0.10, 0.15, 0.2$ is tried and 1930 is estimated as a change-point with $h = 0.2$ which has the minimum MSE(mean squared error). From the U.S. Data. Figure 6 shows the data and its estimated change-point with the vertical line and two regression lines separated by the change-point. World War I would have significantly impacted the global economy as well as the Great Depression

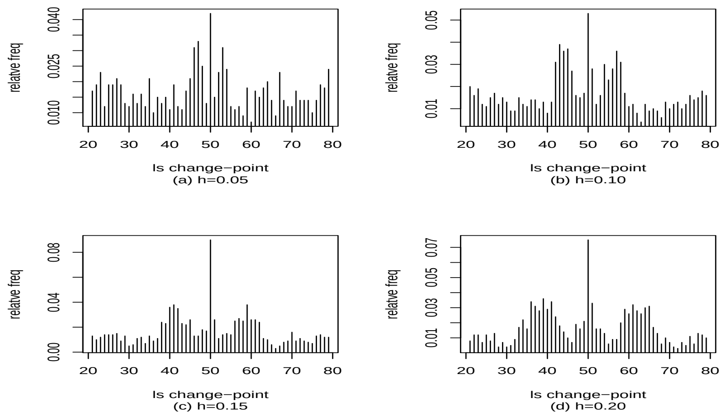

Figure 4: *Relative frequency of one change-point estimates in one change-point model (i) with the bandwidth* $h = 0.05, 0.10, 0.15, 0.2$, $\sigma = 1.0$ *and change-point at* $\tau = 50$

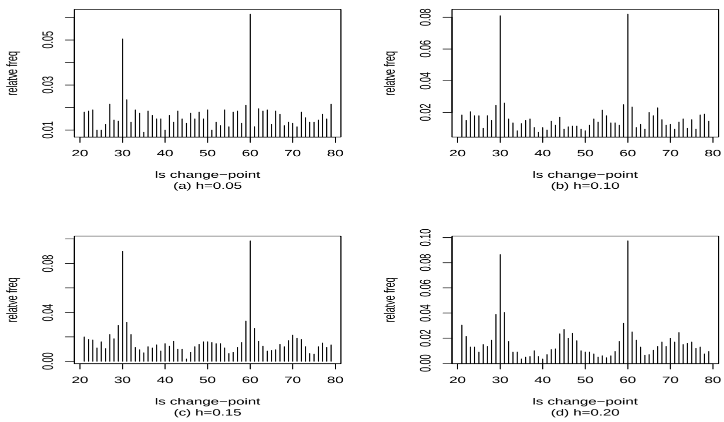

Figure 5: *Relative frequency of two change-point estimates in two change-point model (ii) with the bandwidth* $h = 0.05, 0.10, 0.15, 0.2$, $\sigma = 1.0$ *and change-points at* $\tau = 30, 60$

that lasted until the mid-1930s. The economy showed a distinct trend of growth after World War II, which seemed to cause a structural transformation in the U.S. economy.

We also applied our method to the data on the annual volume of the Nile River from 1871 to 1970. Cobb (1978) estimated one change-point as 1898 with the Nile River data. Among the bandwidth $h = 0.05, 0.10, 0.15, 0.2$, $h = 0.2$ gives two change-points at 1898 and 1938 with the least MSE in case of two possible change-points. Figure 7 shows the Nile River data and its estimated two change-points with the vertical lines.

5. Concluding Remarks

A multiple structural change-point estimation procedure is proposed in the linear regression models based on the difference of local least squares estimators and a split measure. Our method is a simultaneous change-point estimation approach. Even though the clear stopping rule is not provided, the proposed change-point search procedure is applicable to the data when the possible number of

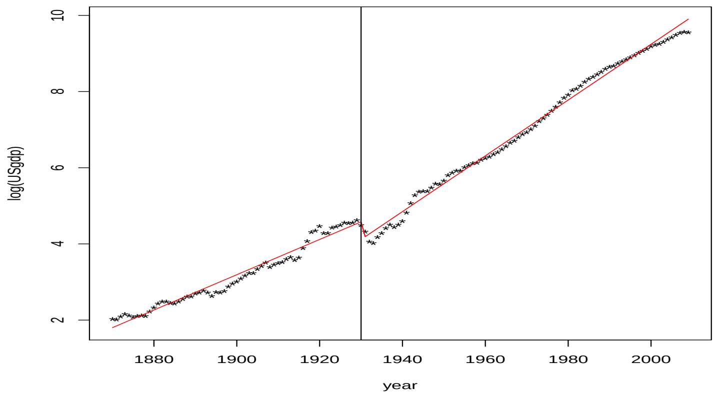

Figure 6: *One change-point estimate and two regression fitting lines with US GDP data from 1870 to 1970 with the bandwidth* $h = 0.2$

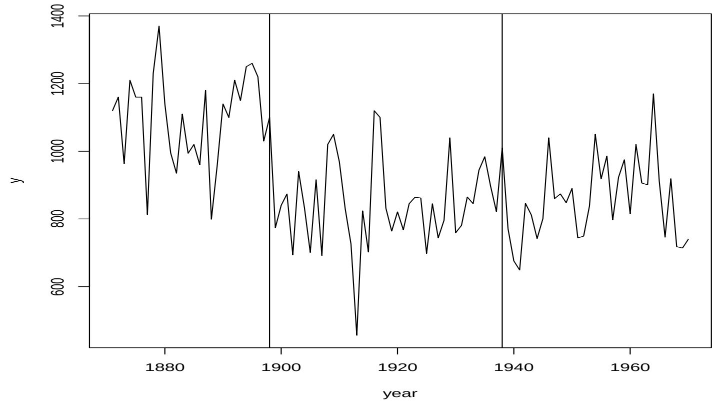

Figure 7: *Two change-point estimates with Nile River data from 1871 to 1970 with the bandwidth* $h = 0.2$

change-points is guessed. Further research with the proposed procedure is expected with the stopping rule, the bandwidth selection and weights for the split measure.

References

Andrews, D. W. K. (1993). Tests for parameter instability and structural changes with unknown change point, *Econometrica*, **61**, 821–856.

Andrews D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, **62**, 1383–1414.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes, *Econometrica*, **66**, 47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, **18**, 1–22.

Bauer, P. and Hackl, P. (1978). The use of MOSUMs for quality control, *Technometrics*, **20**, 431–436.

Broemeling, L. D. and Tsurumi, H. (1987). *Econometrics and Structural Change*, Dekker, New York.

Brown, L. R., Durbin, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time, *Journal of Royal Statistical Society B*, **37**, 149–192.

Chu, C. J., Hornik, K. and Kuan, C. (1995). MOSUM tests for parameter constancy, *Biometrika*, **82**, 603–617.

Chib, S. (1998). Estimation and comparison of multiple change-point models, *Journal of Econometrics*, **86**, 221–241.

Cobb, G. W. (1978). The problem of Nile: Conditional solution to a change-point problem, *Biometrika*, **65**, 243–251.

Dufour, J.-M. and Ghysels, E. (1996). Recent developments in the econometrics of structural change, *Journal of Econometrics*, **70**, 1–8.

Gregoire, G. and Hamrouni, Z. (2002). Change-point estimation by local linear smoothing, *Journal of Multivariate Analysis*, **83**, 56–83.

Kim, J. and Cheon, S. (2010a). Bayesian multiple change-point estimation with annealing stochastic approximation Monte Carlo, *Computational Statistics*, **25**, 215–239.

Kim, J. and Cheon, S. (2010b). A Bayesian regime-switching time series model, *Journal of Time Series Analysis*, **31**, 365–378.

Kim, J. and Hart, J. D. (2010). A change-point estimator using local Fourier series, *Journal of Nonparametric Statistics*, **23**, 83–98.

Koop, G. M. and Porter, S. M. (2004). Forecasting and estimating multiple change-point models with an unknown number of change-points, U of Leicester, Working paper No. 04/31.

Lavielle, M. and Lebarbier, E. (2001). An application of MCMC methods for the multiple change-points problem, *Signal Processing*, **81**, 39–53.

Loader, C. R. (1996). Change point estimation using nonparametric regression, *Annals of Statistics*, **24**, 1667–1678.

McDonald, J. A. and Owen, A. B. (1986). Smoothing with split linear fits, *Technometrics*, **28**, 195–208.

Perron, P. and Zhu, X. (2005). Structural breaks with deterministic and stochastic trends, *Journal of Econometrics*, **129**, 65–119.

Ploberger, W. and Kr̈amer, W. (1992). The cusum test with OLS residuals, *Econometrica*, **60**, 271–285.

Rice, J. (1984). Bandwidth choice for nonparametric regression, *Annals of statistics*, **12**, 1215–1230.

Received February 16, 2012; Revised March 10, 2012; Accepted April 2, 2012