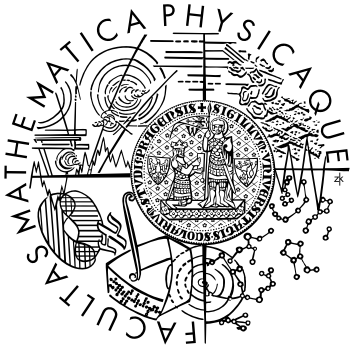
Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Miriam Marušiaková

Vícefázová regrese

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Prof. RNDr. Marie Hušková, DrSc.

Studijní program: Matematika, Pravděpodobnost,
matematická statistika a ekonometrie, Ekonometrie

Děkuji své školitelce Prof. RNDr. Marii Huškové, DrSc. za starostlivé vedení diplomové práce, za konzultace a rovněž za poskytnutí dat o teplotách v Klementinu.

Prohlašuji, že jsem svou diplomovou práci napsala samostatně a výhradně s použitím citovaných pramenů.

Souhlasím se zapůjčováním práce.

V Praze dne 20.4. 2005 Miriam Marušiaková

Contents

Název práce: *Vícefázová regrese*
Autor: *Miriam Marušiaková*
Katedra (ústav): *Katedra pravděpodobnosti a matematické statistiky*
Vedoucí diplomové práce: *Prof. RNDr. Marie Hušková, DrSc.*
e-mail vedoucího: *Marie.Huskova@mff.cuni.cz*

Abstrakt: Studujeme lineární regresní modely se strukturálními změnami, ke kterým dochází v neznámých časech. Věnujeme se testům založeným na F statistikách, které umožňují detekovat přítomnost změn. Pro výpočet aproximací příslušných kritických hodnot navrhujeme metodu založenou na aplikaci permutačního principu. Dokážeme, že užití této metody je opodstatněné. Výsledky simulací potvrzují, že daná metoda uspokojivě aproximuje kritické hodnoty v případech, kdy změny nejsou příliš velké. Na odhad počtu změn užijeme informační kritéria a sekvenční metody. Všechny zkoumané metody aplikujeme na průměrné roční teploty z Klementina.

Klíčová slova: body změny, strukturální změna, segmentovaná regrese, permutační princip

Title: *Multiphase regression*
Author: *Miriam Marušiaková*
Department: *Department of Probability and Mathematical Statistics*
Supervisor: *Prof. RNDr. Marie Hušková, DrSc.*
Supervisor's e-mail address: *Marie.Huskova@mff.cuni.cz*

Abstract: We consider linear regression models with structural changes occurring at unknown time points. We describe F type tests for detection of changes. For calculation of approximations to the corresponding critical values we suggest a method based on the application of the permutation principle. We prove that the method is applicable to the F type test statistics. The simulation study shows that the permutation arguments provide satisfactory approximations to the critical values when the change in parameters is not too large. For estimation of the number of changes present we use information criteria and a method based on sequential testing. All discussed methods are applied to Klementinum data.

Keywords: change points, structural change, segmented regression, permutation principle

Chapter 1

Introduction

In practical situations we often face a problem where a data sample cannot be well described by one relatively simple statistical model during the entire observational period. Various economic factors or human activities (deforestation, urbanisation, ...) may cause that the relationships among the variables change over time. In this case some of the parameters of the statistical model are subject to shifts. Time moments where a change occurs are usually called change points.

The change point problem has attracted attention of many researchers in recent years. This topic offers interesting theoretical problems and has many applications in economics, meteorology, hydrometeorology, environmental studies, biology and many other disciplines. Examples are US ex-post real interest rates or UK inflation rates (Bai and Perron 2003a), monthly water discharges in Načetínský Creek, rainfall departures in Sahel or total ozone amount measured in Hradec Králové (Jarušková 1997), temperature series from Klementinum in Prague or Nile river discharges (Antoch and Hušková 1998), segmentation of the DNA sequence of Bacteriophage λ (Braun, Braun, and Müller 2000) or analysis of cancer mortality and incidence data (Kim, Fay, Feuer, and Midthune 2000), among many others.

The main task is to test whether it is necessary to divide the time ordered data into segments in such a way that the same model can be applied to data in each segment or whether just to use one model for all data. If the data indicate some change, the next task is to estimate the unknown change

points and the total number of changes present. The estimators of the model parameters and their properties are also of interest.

There is a vast amount of literature considering the change point problem. However, most of it deals with just one single change, partially because estimating multiple change points typically requires intensive computation which could have been a problem some years ago. The literature addressing the issue of multiple change points is also rich. There exist different approaches to this subject - variety of methods and model settings were considered. It is impossible to include all of them in this short text and we will refer only to a small part of the existing literature.

A great amount of works covers changes in means of a sequence of independent observations. The observations are divided into segments in such a way that their mean is constant in each segment but varies across the segments. This type of the model is often called "location model". The problem of estimating abrupt change points was discussed e.g. in Antoch and Hušková (1998), who consider procedures based on maximum of the weighted partial sums of residuals and on moving sums of partial residuals to estimate the number and locations of changes. Venter and Steel (1995) propose normal and non-parametric tests based on ratios of optimal sums of squared residuals associated with k and $k+1$ changes, respectively. The tests produce a value for the number of changes present if the null hypothesis of no change is rejected. Chen and Gupta (2000) (in Chapter 2) apply a binary segmentation procedure combined with the Schwarz' information criterion (Schwarz 1978) for detection of changes in normal models. The analysed situations involve change points in means, in variances and changes in means and variances. The advantage of this procedure is that the change points are detected and estimated simultaneously.

Changes in regression parameters in a linear regression model are studied e.g. in the following papers. Bai and Perron (1998) deal with F type tests for multiple changes, namely tests of no change versus k changes where k is fixed or arbitrary with some upper bound, and tests of k versus $k+1$ changes. They consider a partial structural change model where not all parameters are subject to shifts, with quite general assumptions on errors and regressors. They also present a sequential test for estimation of the

number of changes. Bai and Perron (2003a) consider practical issues related to the empirical applications of the F type tests for multiple changes and present an efficient dynamic algorithm to obtain global minimisers of the sum of squared residuals. Bai (1998) deals with least absolute deviations estimation of a regression model with multiple change points. Kim, Fay, Feuer, and Midthune (2000) consider a segmented linear regression model with a continuity constraint at the change points.

In this work we consider multiple linear regression models with changes occurring at unknown time points. In Chapter 2 we introduce the model and notation and formulate the assumptions on regressors and errors. Chapter 3 is devoted to F type tests for detection of changes in linear regression. The approximations to the corresponding critical values are usually derived from a limit distribution of the test statistic under the null hypothesis (Bai and Perron 1998). In Chapter 4 we propose another possibility how to obtain them - we use the approximations based on the application of the permutation principle. After a short description of permutation test procedures based on F type test statistics we prove the asymptotic equivalence of both approaches for obtaining approximations to the critical values. Details of the proof are given in Appendix A. We present a number of simulation results and show that the permutation arguments provide satisfactory approximations to the critical values when the change in parameters is not too large. In Chapter 5 we discuss Schwarz' and modified Schwarz' information criteria and sequential methods to estimate the total number of changes present. In the last Chapter 6 we apply all the discussed methods on the temperature series from Klementinum, Prague. For all calculations we use the software R, see `http://www.r-project.org/`.

Chapter 2

Model and assumptions

We consider the following multiple linear regression model with m changes, i.e. $m+1$ segments

$$\begin{array}{ll} y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \boldsymbol{z}_t'\boldsymbol{\delta}_1 + e_t & t = 1, \ldots, t_1 \\ y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \boldsymbol{z}_t'\boldsymbol{\delta}_2 + e_t & t = t_1 + 1, \ldots, t_2 \\ \vdots & \vdots \\ y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \boldsymbol{z}_t'\boldsymbol{\delta}_{m+1} + e_t & t = t_m + 1, \ldots, n \end{array} \tag{2.1}$$

where t_j, $j = 1, \ldots, m$ are the change points, y_t is the observed dependent variable, $\boldsymbol{x}_t\,(p \times 1)$ and $\boldsymbol{z}_t\,(q \times 1)$ are the vectors of regressors, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_j$, $j = 1, \ldots, m+1$ are the corresponding regression coefficients and e_t is the error at time t.

The change points t_j, $j = 1, \ldots, m$ are in practice mostly unknown. The purpose is to estimate them together with the regression coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_j$, $j = 1, \ldots, m+1$, given the observations $(y_t, \boldsymbol{x}_t, \boldsymbol{z}_t)$, $t = 1, \ldots, n$. We do not impose any continuity constraint on the segmented regression model and so the change points are supposed to coincide with the observational times. The number of changes m is also treated as unknown and has to be estimated.

The regressors may be fixed or random in repeated samples. Since these variables are often not perfectly controlled in economics, we will assume the random design in this work. Non-random regressors (the fixed design) are

covered by the theory as well.

The model (2.1) is called a partial structural change model because only the vector of regression parameters $\boldsymbol{\delta}_j$ is subject to a change, $\boldsymbol{\beta}$ remains the same in all segments. The reason why we assume such a general model is that the vector $\boldsymbol{\beta}$ can be estimated from the entire sample. This is better than to reestimate it whenever a change occurs because we increase the efficiency of the estimator and the power of the tests as well. So if we know that some regression coefficients do not vary, we should include this knowledge in our model, especially when there are multiple changes. When $p = 0$, we obtain a pure structural change model where all parameters are subject to shifts:

$$y_t = \boldsymbol{z}_t'\boldsymbol{\delta}_j + e_t \qquad t = t_{j-1} + 1, \ldots, t_j \tag{2.2}$$

for the j-th segment, $j = 1, \ldots, m+1$, with the convention $t_0 = 0$ and $t_{m+1} = n$.

The model (2.1) can be rewritten in the matrix form as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \bar{\boldsymbol{Z}}\boldsymbol{\delta} + \boldsymbol{e}$$

where

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \ \boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1p} \\ x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{np} \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix},$$

$$\bar{\boldsymbol{Z}} = \begin{pmatrix} \boldsymbol{Z}_1 & \boldsymbol{0} & \ldots & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z}_2 & & \ldots & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \ldots & \ldots & \boldsymbol{0} & \boldsymbol{Z}_{m+1} \end{pmatrix}, \ \boldsymbol{\delta} = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \vdots \\ \boldsymbol{\delta}_{m+1} \end{pmatrix}, \ \boldsymbol{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

with

$$\boldsymbol{Z}_j = \begin{pmatrix} z_{t_{j-1}+1,1} & z_{t_{j-1}+1,2} & \ldots & z_{t_{j-1}+1,q} \\ z_{t_{j-1}+2,1} & z_{t_{j-1}+2,2} & \ldots & z_{t_{j-1}+2,q} \\ \vdots & & \ddots & \vdots \\ z_{t_j,1} & z_{t_j,2} & \ldots & z_{t_j,q} \end{pmatrix}.$$

Having the observations $(y_1, \boldsymbol{x}_1, \boldsymbol{z}_1), \ldots, (y_n, \boldsymbol{x}_n, \boldsymbol{z}_n)$ given, the goal is to estimate the change points t_j, $j = 1, \ldots, m$ and the regression coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_j$, assuming $\boldsymbol{\delta}_j \neq \boldsymbol{\delta}_{j+1}, j = 1, \ldots, m$. We assume for this moment that the number of changes m is known. We discuss possible methods of estimating it in Chapter 5. We also postpone the problem of testing for structural changes to Chapter 3.

The estimation of the regression coefficients is based on the least squares (LS) principle. For each m-partition $(t_1, \ldots, t_m)$ the associated LS estimates of the regression parameters $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_j$ are obtained by minimising the sum of squared residuals (SSR)

$$(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta} - \bar{\boldsymbol{Z}}\boldsymbol{\delta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta} - \bar{\boldsymbol{Z}}\boldsymbol{\delta}) = \sum_{j=1}^{m+1} \sum_{t=t_{j-1}+1}^{t_j} (y_t - \boldsymbol{x}_t'\boldsymbol{\beta} - \boldsymbol{z}_t'\boldsymbol{\delta}_j)^2 .$$

We denote the minimum of this sum by $S_n(t_1, \ldots, t_m)$ and the resulting LS estimates as $\hat{\boldsymbol{\beta}}(t_1, \ldots, t_m)$ and $\hat{\boldsymbol{\delta}}(t_1, \ldots, t_m)$. The change points are estimated as

$$(\hat{t}_1, \ldots, \hat{t}_m) = \arg\min_{t_1, \ldots, t_m} S_n(t_1, \ldots, t_m) \tag{2.3}$$

where the minimisation is taken over all m-partitions such that $t_{j+1} - t_j \geq h \geq q$, $j = 1, \ldots, m$, h is the minimal possible length of a segment. We find the estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_j$ as the LS estimates $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\hat{t}_1, \ldots, \hat{t}_m)$ and $\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{\delta}}(\hat{t}_1, \ldots, \hat{t}_m)$ associated with the best partition (2.3).

An efficient dynamic algorithm for obtaining the estimated change points from (2.3) is discussed in detail in Bai and Perron (2003a). We will briefly outline the idea of this algorithm. We consider a data sample of size n and the total number m of changes. We denote by $c_{i,j}^k$ the minimal SSR obtained by the best partition of a sample starting at time i and ending at time j into k segments. In the first step we calculate SSR of all possible segments $c_{i,j}^1$ with the minimal length h. For a sample size n, the upper bound to the number of segments is $n(n-1)/2$ (all combinations of two indices (i, j), $i < j$, $i, j = 1, \ldots, n$). SSR of any $(m+1)$–segment partition is calculated as a sum of SSR in individual segments. Therefore the algorithm is of order $O(n^2)$[1] for every number of changes $m > 0$.

[1] In comparison, the grid search algorithm is of order $O(n^m)$ for m changes.

The procedure is based on the recursive formula

$$c_{1,j}^{k+1} = \min_{kh \le l \le j-h} \left(c_{1,l}^{k} + c_{l+1,j}^{1} \right)$$

calculated for each possible ending time $j = (k+1)h, \ldots, n-(m-k)h$ of a sample partitioned into $k+1$ segments. We find the optimal $(m+1)$–segment partition of the whole sample as

$$c_{1,n}^{m+1} = \min_{mh \le j \le n-h} \left(c_{1,j}^{m} + c_{j+1,n}^{1} \right),$$

where the last segment is combined with all samples which have ending time j and are optimally partitioned into m segments. The partition which yields an overall minimal SSR is chosen.

We impose the following assumptions on the change points, regressors and errors which we will need in the following chapters. We adopt the convention $t_0 = 0$ and $t_{m+1} = n$.

Assumption 1

$$t_j = [n\lambda_j]^1, \quad 0 = \lambda_0 < \lambda_1 < \ldots < \lambda_{m+1} = 1, \quad \textit{for each} \quad j = 1, \ldots, m+1$$

Assumption 2

$$\frac{(\boldsymbol{X}_j, \boldsymbol{Z}_j)'(\boldsymbol{X}_j, \boldsymbol{Z}_j)}{t_j - t_{j-1}} \xrightarrow{p} \boldsymbol{C} > 0 \quad \textit{as} \quad t_j - t_{j-1} \to \infty, \quad \textit{for each} \quad j = 1, \ldots, m+1$$

where $\boldsymbol{X}_j$ are the rows of the matrix $\boldsymbol{X}$ corresponding to the j-th segment, the letter p means convergence in probability and $\boldsymbol{C}$ is a finite positive definite matrix.

Assumption 3 *The errors are independent and identically distributed (hereafter i.i.d.) with zero mean, nonzero finite variance σ^2 and some finite moment $\mathrm{E}|e_t|^{2+\Delta} > 0$ with some $\Delta > 0$.*

Assumption 4 *The regressors $\boldsymbol{x}_t = (x_{t1}, \ldots, x_{tp})$ and $\boldsymbol{z}_t = (z_{t1}, \ldots, z_{tq})$ are independent with the errors $e_{t'}$ for all t and all t'.*

[1] $[x]$ is the integer part of x

Assumption 1 is needed for the asymptotic theory. It allows the change points to be asymptotically distinct.

Assumption 2 is satisfied e.g. by i.i.d. regressors having a positive definite variance matrix. It rules out trending explanatory variables ($z_t = t$) that have an infinite matrix $\boldsymbol{C}$ in the limit, or vanishing explanatory variables ($z_t = \lambda^t$, $\lambda < 1$) with a singular matrix in the limit. Note, that the limit matrix $\boldsymbol{C}$ in Assumption 2 is the same for all indeces j.

For simplicity we do not allow any heteroscedasticity in the model (Assumption 3) or any correlations between regressors and errors (Assumption 4).

Chapter 3

Test statistics for multiple changes

We introduce tests that help us to decide if a structural change in a linear regression occurred or not. All tests are based on F type test statistics. In Section 3.1 we discuss a test of no change versus k changes, where k can be arbitrary but fixed. We also describe a test against an alternative hypothesis of unknown number of changes k with some upper bound for k. In Section 3.2 we consider a test of k versus $k+1$ changes. This test is particularly useful for determining the number of changes present. We work with the partial structural change model (2.1) where not all regression coefficients are subject to shifts.

3.1 A test of no change versus k changes

In the first part of this section we describe a test of no change against k changes where k is considered to be some fixed number. *First of all we assume that the change points* $t_1, \ldots, t_k$ such that $t_j = [n\lambda_j]$, $0 < \lambda_1 < \ldots < \lambda_k$ under the alternative hypothesis *are known.* The F type test statistic is then defined as

$$F_n(\lambda_1, \ldots \lambda_k; q) = \frac{n - (k+1)\, q - p}{kq} \, \frac{SSR_0 - SSR_k}{SSR_k} \tag{3.1}$$

with

$$SSR_0 = \sum_{t=1}^{n} \left(y_t - \boldsymbol{x}_t' \hat{\boldsymbol{\beta}}_0 - \boldsymbol{z}_t' \hat{\boldsymbol{\delta}}_0 \right)^2;$$

$$SSR_k = \sum_{j=1}^{k+1} \sum_{t=t_{j-1}+1}^{t_j} \left(y_t - \boldsymbol{x}_t' \hat{\boldsymbol{\beta}}(t_1, \ldots, t_k) - \boldsymbol{z}_t' \hat{\boldsymbol{\delta}}_j(t_1, \ldots, t_k) \right)^2.$$

SSR_0 is the minimal SSR under the null hypothesis $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_2 = \cdots = \boldsymbol{\delta}_{k+1} = \boldsymbol{\delta}_0$, SSR_k is the minimal SSR under the alternative hypothesis $H_A : \boldsymbol{\delta}_j \neq \boldsymbol{\delta}_{j+1}, \forall j = 1, \ldots, k$ with the known partition $(t_1, \ldots, t_k)$.

$$\hat{\sigma}_k^2 = \frac{SSR_k}{n - (k+1)\, q - p}$$

is a consistent estimator of the error variance σ^2 under H_A and H_0 (see Appendix in Bai and Perron (1998) or proof of Lemma 3 in Yao (1988) for a location model). There are $(k+1)q + p$ unknown regression parameters in the model under H_A and $p + q$ parameters under H_0. A large value of the test statistic (3.1) indicates that the null hypothesis of no change is violated.

We derive the limit distribution of the test statistic 3.1 under H_0. For ease of notation let us assume a special case with $p = 0$. Then the above test statistic can be rewritten using

$$\begin{aligned} SSR_0 - SSR_k = & - \left(\sum_{t=1}^{n} y_t \boldsymbol{z}_t \right)' \boldsymbol{C}_n^{-1} \left(\sum_{t=1}^{n} y_t \boldsymbol{z}_t \right) \\ & + \sum_{j=1}^{k+1} \left(\sum_{t=t_{j-1}+1}^{t_j} y_t \boldsymbol{z}_t \right)' \boldsymbol{C}_{t_{j-1}, t_j}^{-1} \left(\sum_{t=t_{j-1}+1}^{t_j} y_t \boldsymbol{z}_t \right) \end{aligned} \tag{3.2}$$

$$\begin{aligned} \hat{\sigma}_k^2 = & \frac{1}{n - (k+1)q} \left[\sum_{t=1}^{n} y_t^2 - \sum_{j=1}^{k+1} \left(\sum_{t=t_{j-1}+1}^{t_j} y_t \boldsymbol{z}_t \right)' \right. \\ & \left. \times \boldsymbol{C}_{t_{j-1}, t_j}^{-1} \left(\sum_{t=t_{j-1}+1}^{t_j} y_t \boldsymbol{z}_t \right) \right] \end{aligned} \tag{3.3}$$

where

$$\boldsymbol{C}_n = \sum_{t=1}^{n} \boldsymbol{z}_t \boldsymbol{z}_t'; \qquad \boldsymbol{C}_{t_{j-1},t_j} = \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \boldsymbol{z}_t', \quad j = 1, \ldots, k+1. \tag{3.4}$$

Under H_0 the formulas (3.2) and (3.3) also hold when y_t is replaced by $e_t = y_t - \boldsymbol{z}_t' \boldsymbol{\delta}_0$. The estimator $\hat{\sigma}_k^2$ converges in probability to σ^2. Hence, we can concentrate on the limit of (3.2). Applying the central limit theorem to a vector $\left(\sum_{t=t_0+1}^{t_1} e_t \boldsymbol{z}_t, \ldots, \sum_{t=t_k+1}^{t_{k+1}} e_t \boldsymbol{z}_t\right)$ and using Assumption 2 we arrive to the following result.

Theorem 3.1.1 *Under Assumptions 1-4 the limit distribution of the test statistic (3.1) under the null hypothesis is*

$$F(k;q) = \frac{\sum_{j=1}^{k+1} (1 - \lambda_j + \lambda_{j-1}) \chi_j^2(q)}{kq}$$

where $\chi_j^2(q)$ stands for independent chi-square distributions with q degrees of freedom.

More details about the proof are given in Appendix A.

Now we assume the change points $t_1, \ldots, t_k$ are unknown. For an asymptotic analysis we need to impose some restrictions on the possible values of the change points. We define a set

$$\Lambda_\varepsilon = \{(\lambda_1, \ldots \lambda_k)\,;\, \lambda_{j+1} - \lambda_j \geq \varepsilon,\ \forall\, j = 0, \ldots k\} \tag{3.5}$$

for some arbitrary small $\varepsilon > 0$, so called the trimming parameter. ε imposes the minimal possible length $h = n\varepsilon$ of a segment. The $\sup F$ type test statistic is defined as

$$\sup F_n(k;q) = \sup_{(\lambda_1, \ldots \lambda_k) \in \Lambda_\varepsilon} F_n(\lambda_1, \ldots \lambda_k; q) \tag{3.6}$$

for some arbitrary positive ε.

Since computing of the $\sup F_n(k;q)$ through all $(\lambda_1,\ldots,\lambda_k)$ in the set Λ_ε is rather inconvenient, Bai and Perron (2004) define an asymptotically equivalent version which is simpler to obtain:

$$\sup F_n^*(k;q) = F_n(\hat{\lambda}_1,\ldots,\hat{\lambda}_k;q) \tag{3.7}$$

where $\hat{\lambda}_j = \hat{t}_j/n$, $j = 1,\ldots,k$ and $\hat{t}_1,\ldots,\hat{t}_k$ are the estimated change points obtained as global minimisers of the SSR, see equation (2.3).

The limit distribution of the test $\sup F_n(k;q)$ (3.6) under H_0 is specified in Proposition 6 of Bai and Perron (1998) under quite general assumptions on errors and regressors. It depends on the value of the trimming parameter ε: as $\varepsilon \to 0$, the critical values of the test statistic diverge to infinity. The authors adopted $\varepsilon = 0.05$. Asymptotic critical values up to 9 changes ($1 \le k \le 9$) and for maximum of 10 changing regressors ($q \le 10$) are displayed in Table I of Bai and Perron (1998). Additional critical values for $\varepsilon = 0.10, 0.15, 0.20$ can be found in Bai and Perron (2003b).

So far we have tested the null hypothesis of no structural change versus the alternative assuming a particular number of changes. In practice, however, the number of changes is often unknown. Therefore it is more of interest to test the hypothesis of no change versus an unknown number of changes, given some upper bound M for the number of changes. A new test, so called a double maximum test (Bai and Perron 1998), is defined as

$$D\max F_n(M,q,a_1,\ldots a_M) = \max_{1\le k\le M}\left(a_k \sup_{(\lambda_1,\ldots\lambda_k)\in\Lambda_\varepsilon} F_n(\lambda_1,\ldots\lambda_k;q)\right)$$

for some weights $a_1,\ldots a_M$. If we have some prior knowledge about the likelihood of various numbers of changes, then the weights may be given in such a way that the more probable the number of changes is, the higher weight is selected.

The simplest case is to set all weights to unity:

$$UD\max F_n(M,q) = \max_{1\le k\le M} \sup_{(\lambda_1,\ldots\lambda_k)\in\Lambda_\varepsilon} F_n(\lambda_1,\ldots\lambda_k;q). \tag{3.8}$$

The asymptotically equivalent version is

$$UD\max F_n^*(M,q) = \max_{1\le k\le M} F_n(\hat{\lambda}_1,\ldots\hat{\lambda}_k;q) \tag{3.9}$$

where $\hat{\lambda}_j = \hat{t}_j/n$, $j = 1, \ldots k$ and $\hat{t}_j$ are again the estimated change points from (2.3).

For a fixed sample the critical values of the test (3.6) decrease as k increases and so the p-values also decrease with k (the null hypothesis is more often rejected even if it is true) and hence the test has less informative power if the number of changes is large. Therefore Bai and Perron (1998) specify some special weights such that the p–values equal for each k. Let $c(q, \alpha, k)$ be the asymptotic critical value of the test (3.6). Then the weights $a_1, \ldots a_M$ are defined as $a_1 = 1$ and $a_k = c(q, \alpha, 1)/c(q, \alpha, k)$ for $k > 1$. They depend on the value of q and on the significance level of the test α. This version of the test is denoted as

$$WD\max F_n(M, q) = \max_{1 \le k \le M} \frac{c(q, \alpha, 1)}{c(q, \alpha, k)} \sup_{(\lambda_1, \ldots \lambda_k) \in \Lambda_\varepsilon} F_n(\lambda_1, \ldots \lambda_k; q) \qquad (3.10)$$

and the asymptotically equivalent version is

$$WD\max F_n^*(M, q) = \max_{1 \le k \le M} \frac{c(q, \alpha, 1)}{c(q, \alpha, k)} F_n(\hat{\lambda}_1, \ldots \hat{\lambda}_k; q). \qquad (3.11)$$

Bai and Perron (1998) obtained the asymptotic critical values of the tests (3.8) and (3.10) for $M = 5$ and $\varepsilon = 0.05$. The critical values vary little for $M > 5$. Additional critical values for $\varepsilon = 0.10$ ($M = 5$), 0.15 ($M = 5$), 0.20 ($M = 3$) and 0.25 ($M = 2$) are tabulated in Bai and Perron (2003b).

3.2 A test of k versus $k + 1$ changes

Bai and Perron (1998) also consider a test of the null hypothesis of k changes against the alternative that an additional change is present. The test is based on testing of each from $k + 1$ segments for a presence of a change. The k change points $\hat{t}_1, \ldots \hat{t}_k$ under the null hypothesis are obtained by a global minimisation of SSR using the dynamic algorithm, see Chapter 2 for more information.

For each segment containing the observations $\hat{t}_{j-1}+1, \ldots \hat{t}_j$, $j = 1, \ldots k+1$ (with the convention $\hat{t}_0 = 0$, $\hat{t}_{k+1} = n$), the test of no change versus one

change is applied. If the overall minimum of the SSR of the sample with $k+1$ changes is sufficiently smaller than the SSR associated with k changes, then the null hypothesis is rejected and a new change point is added to that segment where SSR achieves the greatest reduction. The test is defined as

$$F_n(k+1|k) = \left\{ S_n\left(\hat{t}_1, \ldots, \hat{t}_k\right) - \min_{1 \le j \le k+1} \left(\inf_{\hat{t}_{j-1}+1+h \le \tau \le \hat{t}_j - h} S_n\left(\hat{t}_1, \ldots, \hat{t}_{j-1}, \tau, \hat{t}_j, \ldots, \hat{t}_k\right) \right\} / \hat{\sigma}_k^2 \right) \quad (3.12)$$

where $S_n\left(\hat{t}_1, \ldots, \hat{t}_k\right)$, $S_n\left(\hat{t}_1, \ldots, \hat{t}_{j-1}, \tau, \hat{t}_j, \ldots, \hat{t}_k\right)$ is the minimal SSR for a given partition $(\hat{t}_1, \ldots, \hat{t}_k)$, $(\hat{t}_1, \ldots, \hat{t}_{j-1}, \tau, \hat{t}_j, \ldots, \hat{t}_k)$, $h = n\varepsilon$ is the minimal possible length of a segment and $\hat{\sigma}_k^2 = S_n(\hat{t}_1, \ldots, \hat{t}_k)/n$ is a consistent estimator of the error variance σ^2 under the null hypothesis.

The limiting distribution of the test statistic (3.12) under the null hypothesis is specified in Proposition 7 of Bai and Perron (1998). The critical values for $\varepsilon = 0.05, 0.10, 0.15, 0.20, 0.25$ and $1 \le q \le 10$ can be found in Bai and Perron (1998, 2003b).

Bai (1999) introduced an alternative procedure to test k changes versus $k+1$ changes. Unlike the previous test (3.12), here the change points under null and also alternative hypothesis are obtained simultaneously via global minimisation of SSR. The test is based on the difference between the optimal SSR corresponding to k changes and that corresponding to $k+1$ changes. Let $\hat{t}_1, \ldots \hat{t}_k$ be the estimated change points under the null hypothesis and $\hat{t}_1^*, \ldots \hat{t}_{k+1}^*$ the estimated change points under the alternative. Then the test statistic, so called likelihood ratio test statistic, is defined as

$$LR_n(k+1|k) = \frac{S_n(\hat{t}_1, \ldots, \hat{t}_k) - S_n(\hat{t}_1^*, \ldots, \hat{t}_{k+1}^*)}{\hat{\sigma}_{k+1}^2} \quad (3.13)$$

where $\hat{\sigma}_{k+1}^2 = S_n(\hat{t}_1^*, \ldots, \hat{t}_{k+1}^*)/n$ is a consistent estimator of the error variance σ^2 under both hypothesises.

The limiting distribution of the test (3.13) is derived in Theorem 1 of Bai (1999). It has a known analytical density function and hence the critical values of the test can be easily computed from the formula in Corollary 1 of Bai (1999).

Both mentioned tests of k versus $k+1$ changes can be used for identification of the number of change points. We will describe a sequential procedure based on these tests in Section 5.2.

The F type test statistics mentioned above are applicable also under fairly general assumptions on regressors and errors, see any article of Bai and Perron in references. For example they can be applied to models allowing serial correlated errors and heteroscedasticity. In that case it is recommended to use larger trimming parameter ε to achieve tests with correct size in finite samples. The tests can also be constructed for different distribution of the errors and regressors across the segments. Bai and Perron (2003a) analyse various versions of the tests depending on the assumptions. Different specifications are considered in the case of pure and partial structural change models.

Chapter 4

Permutation test procedures

In this chapter we deal with approximations to the critical values of the test of no change versus k fixed changes. Bai and Perron (1998) use approximations based on the limit distribution of the test statistic (3.6) under H_0. Here we describe another possible approach based on the application of the permutation principle. In Section 4.1 we explain the theory concerning the permutation test procedures related to F type test statistics. We prove that the permutational method is applicable to our situation. In Section 4.2 we conduct various simulation experiments in order to demonstrate how the method works when applied to regression models with changes of different size.

4.1 Principle of permutation tests

We were inspired by Hušková (2004), Hušková and Antoch (2003) and Antoch and Hušková (2001) where the permutation test procedures were used for the approximations of the critical values of maximum type statistics based on partial weighted sums of residuals. The procedures were applied to location models or regression models with at most one change. In these cases approximations based on the limit behaviour of the considered test statistics under H_0 were not satisfactory because their convergence rate was rather small. Therefore the asymptotic critical values were far from reality when the sample size was not too large. The approximations based on the per-

mutation tests gave much better results, the obtained critical values were smaller than the asymptotic ones and hence also changes of smaller sizes could be detected.

In this section we apply the permutation principle to the F type test statistic of no change against k fixed changes. For simplicity we explain the permutation approach only for the statistic $F_n(k;q)$ (3.1) where the locations of changes under H_A, i.e. $\lambda_1 < \cdots < \lambda_k$ are assumed to be known. We consider the pure structural change model (2.2) with all regression coefficients subject to a change.[1] We add the following assumption on the regressors.

Assumption 5 *The regressors $\boldsymbol{z}_t$ are known constants and the first component is equal to 1, i.e. $z_{t1} = 1,\ t = 1, \ldots, n$.*

Under the null hypothesis $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_2 = \ldots = \boldsymbol{\delta}_{k+1} = \boldsymbol{\delta}_0$ the errors $e_t = y_t - \boldsymbol{z}_t'\boldsymbol{\delta}_0$, $t = 1, \ldots, n$ are i.i.d. random variables. Thus they are exchangeable and $(e_1, \ldots, e_n)$ has the same distribution as $(e_{R_1}, \ldots, e_{R_n})$ where $\boldsymbol{R} = (R_1, \ldots, R_n)$ is a random permutation of $(1, \ldots, n)$. Since the errors e_t are unknown, we replace them by their estimators under H_0, i.e. the residuals

$$\hat{e}_t = y_t - \boldsymbol{z}_t'\hat{\boldsymbol{\delta}}_0. \tag{4.1}$$

The main idea is to randomly permute the residuals and for every such permutation calculate the related test statistic. More exactly, recall that the statistic $F_n(k;q)$ (3.1) under H_0 can be written using formulas (3.2) and (3.3) with e_t instead of y_t. The permutation version of the statistic $F_n(k;q)$ has the form

$$F_n(k;q;\boldsymbol{R}) = \frac{SSR_0(\boldsymbol{R}) - SSR_k(\boldsymbol{R})}{kq\ \hat{\sigma}_k^2(\boldsymbol{R})} \tag{4.2}$$

where $SSR_0(\boldsymbol{R}) - SSR_k(\boldsymbol{R})$ and $\hat{\sigma}_k^2(\boldsymbol{R})$ are given by equations (3.2) and (3.3), respectively, with y_t replaced by $\hat{e}_{R_t}$.

The residuals $\hat{e}_t$ depend on the original observations y_t. We will study the conditional distribution of the statistic $F_n(k;q;\boldsymbol{R})$ (4.2), given $y_1, \ldots, y_n$, also

[1] The proofs concerning the application of the permutation principle related to the statistic $\sup F_n^*(k;q)$ (3.7) with unknown change points, while considering the partial structural change models (2.1), are beyond this work.

called permutation distribution. Its exact form is known because the distribution of random permutations $\boldsymbol{R}$ is known. However, it is computationally demanding to compute the statistics for all $n!$ permutations. Therefore we independently and randomly select N permutations where $N << n!$ is a reasonably large number to get satisfactory approximations. For these permutations we compute the statistic $F_n(k; q; \boldsymbol{R})$ (4.2).

For the purpose of examining the limit conditional distribution of the test $F_n(k; q; \boldsymbol{R})$, given $y_1, \ldots, y_n$, we can write

$$\begin{aligned} &SSR_0(\boldsymbol{R}) - SSR_k(\boldsymbol{R}) \\ &\quad = -\left(\sum_{j=1}^{k+1} \boldsymbol{S}_{j,n}\right)' \boldsymbol{C}_n^{-1} \left(\sum_{j=1}^{k+1} \boldsymbol{S}_{j,n}\right) + \sum_{j=1}^{k+1} \boldsymbol{S}'_{j,n} \boldsymbol{C}^{-1}_{t_{j-1},t_j} \boldsymbol{S}_{j,n} \end{aligned}$$

$$\hat{\sigma}_k^2(\boldsymbol{R}) = \frac{1}{n-(k+1)q} \left[\sum_{t=1}^{n} \hat{e}_t^2 - \sum_{j=1}^{k+1} \boldsymbol{S}'_{j,n} \boldsymbol{C}^{-1}_{t_{j-1},t_j} \boldsymbol{S}_{j,n}\right]$$

where

$$\boldsymbol{S}_{j,n} = \sum_{t=t_{j-1}+1}^{t_j} \hat{e}_{R_t} \boldsymbol{z}_t, \quad j = 1, \ldots, k+1; \quad \boldsymbol{S}_{0,n} = \sum_{t=1}^{n} \hat{e}_{R_t} \boldsymbol{z}_t = \sum_{j=1}^{k+1} \boldsymbol{S}_{j,n}$$

are vectors of linear rank statistics, given $y_1, \ldots, y_n$. $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ are known regression vectors and $a_n(t) = \hat{e}_t$ are the scores. Thus the study of the conditional limit distribution of the statistic $F_n(k; q; \boldsymbol{R})$ is reduced to the study of the limit distribution of vectors of linear rank statistics. It is sufficient to deal only with $SSR_0(\boldsymbol{R}) - SSR_k(\boldsymbol{R})$, because, similarly as in Section 3.1, $\hat{\sigma}_k^2(\boldsymbol{R})$ converges in probability to σ^2. Under Assumptions 2, 4 and 5 we can approximate the vectors of linear rank statistics $\boldsymbol{S}_{j,n}$ by vectors of weighted sums of independent random variables (Theorem 5.1 in Hušková and Antoch (2003))

$$\boldsymbol{T}_{j,n} = \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \left(a_n\left(\lfloor nU_t \rfloor + 1\right) - \bar{a}_n(U)\right) \; j = 1, \ldots, k+1$$

where $U = (U_1, \ldots, U_n)'$ is a sample from a uniform distribution on $(0, 1)$, $\boldsymbol{R} = (R_1, \ldots, R_n)'$ are the corresponding ranks and

$$\bar{a}_n(U) = \frac{1}{n} \sum_{t=1}^{n} a_n(\lfloor nU_t \rfloor + 1).$$

Using the multivariate central limit theorem we get that the vectors of linear rank statistics $(\boldsymbol{S}_{1,n}, \ldots, \boldsymbol{S}_{k+1,n})$ have asymptotically a normal distribution with zero mean and the variance matrix calculated in Appendix A. Further, using Assumptions 1, 2 and realizing that $\sum_{t=t_{j-1}}^{t_j} \boldsymbol{z}_t/(t_j - t_{j-1})$ converges to the first column of the limit matrix $\boldsymbol{C}$, we obtain the following result:

Theorem 4.1.1 *Let observations $(y_1, \boldsymbol{z}_1'), \ldots, (y_n, \boldsymbol{z}_n')$ follow the model (2.2) with no restrictions on the number of change points m. Under Assumptions 1-5 the conditional distribution of the test statistic $F_n(k; q; \boldsymbol{R})$ (3.1), given $y_1, \ldots, y_n$, converges in distribution to*

$$F_n(k; q; \boldsymbol{R}) \xrightarrow{\mathcal{D}} \frac{\sum_{j=1}^{k+1} (1 - \lambda_j + \lambda_{j-1}) \chi_j^2(q)}{kq}$$

where $\chi_j^2(q)$ stands for independent chi-square distributions with q degrees of freedom.

The derivation of the limit distribution in Theorem 4.1.1 is given in Appendix A. Notice, that the conditional limit distribution of the test $F_n(k; q; \boldsymbol{R})$ (4.2), given $\boldsymbol{y}$, does not depend on the original observations $\boldsymbol{y}$ and coincides with the limit distribution of the test statistic $F_n(k; q)$ (3.1) under the null hypothesis. Therefore the quantiles corresponding to the empirical conditional distribution of the statistic $F_n(k; q; \boldsymbol{R})$ can be good approximations to critical values corresponding to the test based on the statistic $F_n(k; q)$.

The situation is more complicated when the locations of changes $t_1, \ldots, t_k$ under the alternative hypothesis are unknown. Then the limit distribution of the test can be described via Wiener processes. We do not prove it but conduct a number of simulations which are discussed in the next section.

4.2 Simulated critical values

We would like to test the null hypothesis of no change versus some fixed number of changes k where the change points $t_1, \ldots, t_k$, (i.e. $0 < \lambda_1 < \ldots < \lambda_k$) under H_A are unknown. In the previous section we showed that we can get reasonable approximations to critical values of the test $F_n(k; q)$ with known change points under H_A by applying the permutation test procedures. Here we conduct various simulation experiments and apply the permutation arguments to the test statistic $\sup F_n^*(k; q)$ (3.7) assuming unknown change points under H_A. We want to show that the approximations to the critical values obtained through the permutation principle are quite stable whether the data follow the null hypothesis or the alternatives. We compare the empirical critical values with the asymptotic ones calculated by Bai and Perron (1998, 2003b).

We denote the permutational version of the test statistic $\sup F_n^*(k; q)$ by $\sup F_n^*(k; q; \boldsymbol{R})$. It is defined as

$$\begin{aligned} &\sup F_n^*(k; q; \boldsymbol{R}) \\ &\quad = \frac{n-(k+1)q}{kq} \, \frac{\sum_{t=1}^{n} \left(\hat{e}_{R_t} - \boldsymbol{z}_t' \hat{\boldsymbol{\delta}}_0\right)^2 - \sum_{j=1}^{k+1} \sum_{t=\hat{t}_{j-1}+1}^{\hat{t}_j} \left(\hat{e}_{R_t} - \boldsymbol{z}_t' \hat{\boldsymbol{\delta}}_j\right)^2}{\sum_{j=1}^{k+1} \sum_{t=\hat{t}_{j-1}+1}^{\hat{t}_j} \left(\hat{e}_{R_t} - \boldsymbol{z}_t' \hat{\boldsymbol{\delta}}_j\right)^2} \end{aligned} \tag{4.3}$$

with

$$\hat{\boldsymbol{\delta}}_0 = \boldsymbol{C}_n^{-1} \sum_{t=1}^{n} \boldsymbol{z}_t \hat{e}_{R_t}; \qquad \hat{\boldsymbol{\delta}}_j = \hat{\boldsymbol{\delta}}_j(\hat{t}_1, \ldots, \hat{t}_k) = \boldsymbol{C}_{\hat{t}_{j-1}, \hat{t}_j}^{-1} \sum_{t=\hat{t}_{j-1}+1}^{\hat{t}_j} \boldsymbol{z}_t \hat{e}_{R_t}$$

where $\hat{t}_1, \ldots, \hat{t}_k$ are obtained as global minimisers of SSR assuming the minimal length of a segment to be $h = n\varepsilon$. In our simulation experiments we used the trimming parameter $\varepsilon = 0.15$ which is also the default value in the function `breakpoints` in the program R.

We generate data from the model (2.2) considering

- sample sizes $n = 100, 160$;

- i.i.d. errors with normal or Laplace distribution with variance equal to one;
- $q = 2$; $\boldsymbol{z}_t = (z_{t1}, z_{t2})$ where $z_{t1} = 1$, $t = 1, \ldots, n$ and regressors z_{t2} are generated from a logarithmic normal distribution where logarithm of the distribution function has mean equal to 0 and standard deviation equal to 1;
- up to 2 change points ($m = 0, 1, 2$) with timing $t_1 = n/4$, $t_2 = 3n/4$;
- regression coefficients $\boldsymbol{\delta}_1' = (0, 1)$; all considered values of $\boldsymbol{\delta}_2$, $\boldsymbol{\delta}_3$ can be seen in any Table B.1 - B.6 in Appendix B.

The value of the regression coefficients in the first segment is always $\boldsymbol{\delta}_1' = (0, 1)$. In the second and third segment either the value of intercept or the slope or both may change. We consider models with no change ($\boldsymbol{\delta}_1 = \boldsymbol{\delta}_2 = \boldsymbol{\delta}_3 = \boldsymbol{\delta}_0$) or changes of sizes 0.5 and 1. Changes of greater size than 1 are easy to detect and there is no need to test whether they have occurred or not.

We proceed as follows. First we generate n independent errors e_t and regressors z_{t2}. For particular values of the coefficients $\boldsymbol{\delta}_j$, $j = 1, \ldots m + 1$, we calculate y_t. Having y_t and $\boldsymbol{z}_t$ we calculate the residuals $\hat{e}_t$ (4.1) from the model under the null hypothesis. We apply the permutation principle to these residuals: we generate a random permutation $\boldsymbol{r} = (r_1, \ldots, r_n)$ of $(1, \ldots, n)$ and calculate the permutation version of the statistic $\sup F_n^*(k; q; \boldsymbol{R})$ (4.3) for $\boldsymbol{R} = \boldsymbol{r}$. We repeat the last two steps for 10 000 random permutations $\boldsymbol{R}$. Finally we obtain the empirical distribution of $\sup F_n^*(k; q; \boldsymbol{R})$ and compute its corresponding empirical quantiles which we use as the approximations to critical values of the test $\sup F_n^*(k; q)$. The empirical distribution of $\sup F_n^*(k; q; \boldsymbol{R})$ for $k = 2, q = 2, \varepsilon = 0.15$ is plotted in Figure 4.1.

Recall that in order to get the exact permutation distribution of the test statistic $\sup F_n^*(k; q; \boldsymbol{R})$ (4.3) we should calculate its values for all $n!$ permutations $\boldsymbol{R} = (R_1, \ldots, R_n)$. This is of course practically impossible unless n is very small ($n \leq 9$). 10 000 $<<$ $n!$ random permutations seem to be a reasonably large number for our simulations. In order to see how much the empirical quantiles change with the increasing number of permutations,

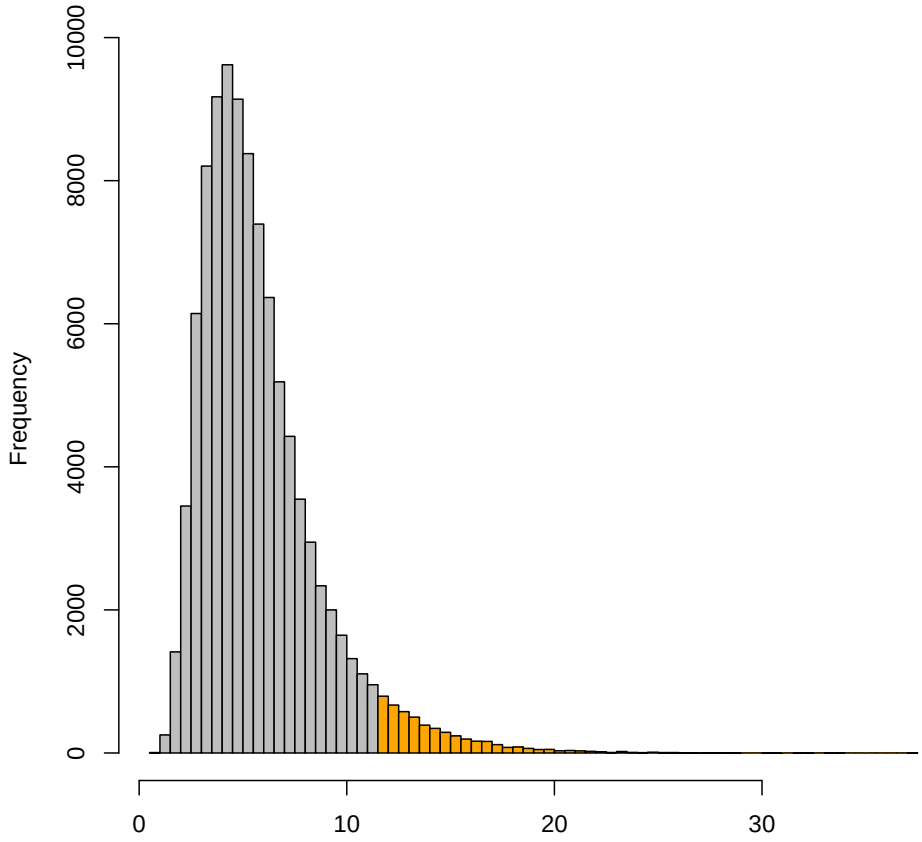


Figure 4.1: Histogram of the statistic $\sup F_n^*(k; q; \boldsymbol{R})$ for $k = 2$, $q = 2$ and $\varepsilon = 0.15$ calculated from 100 000 permutations. The orange bars in the graph represent values larger than the 95 % quantile. The original data sample followed the model with $m = 2$ changes and regression coefficients $\boldsymbol{\delta}_1' = (0, 1)$, $\boldsymbol{\delta}_2' = (0, 2)$, $\boldsymbol{\delta}_3' = (0, 3)$. The errors were generated from the standard normal distribution.

Table 4.1: Empirical quantiles x such that $P(\sup F_n^*(k;q) \leq x/q) = 1-\alpha$ calculated for the increasing number of random permutations N.

			$k=2$	$q=2$		$\varepsilon=0.15$			
N	0.10	0.05	0.025	0.01	N	0.10	0.05	0.025	0.01
10 000	9.46	11.60	13.48	15.88	60 000	9.50	11.51	13.41	15.82
20 000	9.51	11.54	13.55	16.09	70 000	9.52	11.52	13.45	15.95
30 000	9.47	11.51	13.45	15.95	80 000	9.52	11.52	13.45	15.93
40 000	9.49	11.53	13.46	15.92	90 000	9.51	11.51	13.45	15.97
50 000	9.51	11.53	13.46	15.95	100 000	9.50	11.49	13.44	15.96

Notes: $\alpha = 0.10,\ 0.05,\ 0.025,\ 0.01$.
Sample details: $n = 100$, $e_t \sim N(0,1)$, number of changes $m = 2$, regression coefficients $\boldsymbol{\delta}_1' = (0,1)$, $\boldsymbol{\delta}_2' = (0,2)$, $\boldsymbol{\delta}_3' = (0,3)$.

we generated up to 100 000 permutations and applied the related test statistics $\sup F_n^*(k;q;\boldsymbol{R})$ to a data sample following a model with two changes. The 90 %, 95 %, 97.5 % and 99 % empirical quantiles were calculated after $10\,000, 20\,000, 30\,000, \ldots, 100\,000$ permutations. The results are shown in Table 4.1. We see the values of the empirical quantiles stabilise already for $N = 10\,000$, the difference between the quantiles calculated from 10 000 permutations and those calculated from 100 000 permutations is at most 0.1.

The calculation of 10 000 values of the test statistic $\sup F_n^*(k;q;\boldsymbol{R})$ using $\varepsilon = 0.15$ took over 3 hours for the sample size $n = 100$ and about 9 hours for $n = 160$ (Pentium 4, 2.4 GHz).

In Tables B.1 - B.6 in Appendix B we present some of our simulation results. In each single table there are empirical quantiles from various data samples for the test of no change versus k changes. We assumed $k = 1, 2, 3$ changes under the alternative hypothesis. In the first three Tables B.1 - B.3 we considered sample size $n = 100$ and in Tables B.4 - B.6 size $n = 160$. We obtained quite satisfactory results in most of the simulations. The 90 %, 95 %, 97.5 % and 99 % empirical quantiles calculated from samples with $m = 1, 2$ changes are similar to those which were computed from samples following the null hypothesis ($m = 0$). The values are also close to the asymptotic critical values of the test $\sup F_n^*(k;q)$ (hereafter ACV) calculated by Bai and Perron

Table 4.2: Empirical quantiles x such that $P(\sup F_n^*(k;q) \leq x/q) = 1-\alpha$ calculated for different sets of 10 000 random permutations for a given data sample.

$k=2$				$q=2$			$\varepsilon = 0.15$
0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
9.455	11.599	13.482	15.884	9.465	11.304	13.070	15.274
9.560	11.471	13.642	16.181	9.665	11.669	13.694	16.470
9.380	11.439	13.246	15.615	9.482	11.502	13.467	15.915
9.575	11.583	13.535	15.813	9.450	11.423	13.414	16.162
9.558	11.535	13.468	16.065	9.442	11.372	13.373	15.892

See Table 4.1 for more sample details.

(1998, 2003b). However, there are situations for which the empirical quantiles are much higher than expected, especially for the significance levels $\alpha = 0.01$ and $\alpha = 0.025$. The empirical critical values corresponding to $\alpha = 0.10$ are all in a good agreement with empirical critical values for $m = 0$ and with ACV, the difference is in average $0.3 - 0.4$. The situation is a little better for a larger sample size $n = 160$. The empirical critical values for $n = 160$ seem to better agree with ACV, but there are still some exceptions with larger values than ACV. [2]

In order to save the computational time, the test statistics $\sup F_n^*(k;q;\boldsymbol{R})$ were calculated for all alternative hypotheses $k = 1, 2, 3$ using the same permutated residuals. Therefore it is more likely that the critical values not satisfactory for e.g. $k = 1$ will not be satisfactory for other k. The errors e_t and regressors z_{t2} were always newly generated for each considered value of regression coefficients $\boldsymbol{\delta}_j$.

When comparing the critical values obtained from the samples assuming normal errors and those assuming Laplace errors, we cannot say generally that the critical values are e.g. higher for Laplace errors or vice versa. The empirical quantiles do not indicate any dependence on the distribution of the errors.

[2] ACV and empirical critical values from samples with $m = 0$ are approximately the same

Table 4.3: Empirical quantiles x such that $P(\sup F_n^*(k;q) \le x/q) = 1-\alpha$ calculated 20 times, each from newly generated samples following the same model.

			$k=2$	$q=2$				$\varepsilon=0.15$	
	0.10	0.05	0.025	0.01		0.10	0.05	0.025	0.01
1	10.34	12.69	14.63	16.89	11	9.17	10.73	12.29	14.12
2	8.80	10.49	12.03	13.88	12	8.66	10.03	11.32	13.12
3	9.01	10.46	11.82	13.60	13	8.68	9.89	11.18	12.68
4	9.61	11.76	13.86	16.42	14	9.57	11.51	13.56	16.06
5	9.04	10.63	12.33	14.30	15	13.91	18.04	20.82	24.19
6	10.12	12.05	13.93	16.72	16	9.19	10.68	11.98	13.65
7	10.17	12.23	14.47	17.16	17	8.87	10.51	12.09	14.19
8	11.09	13.74	15.98	19.21	18	8.79	10.18	11.48	13.44
9	9.06	10.80	12.41	14.33	19	10.28	13.01	15.55	18.68
10	8.74	10.06	11.44	13.09	20	9.01	10.57	12.09	14.11

See Table 4.1 for more sample details.

Now let us examine the cases where the empirical critical values are too large when compared to ACV. The empirical critical values are more likely to exceed ACV when the data follow a model with at least one change in the slope parameter, especially when the change is large (1 is already considered to be a large change in the slope). We would not improve the results if we calculated the empirical distribution of the test statistic $\sup F_n^*(k;q;\boldsymbol{R})$ from more than 10 000 permutations, see the above discussed Table 4.1. The dependence on the randomness of selected random permutations is also negligible: for our chosen example

$$m = 2, \quad \boldsymbol{\delta}_1 = (0,1)', \quad \boldsymbol{\delta}_2 = (0,2)', \quad \boldsymbol{\delta}_3 = (0,3)' \tag{4.4}$$

we calculated the quantiles for different sets of 10 000 random permutations, see Table 4.2. We can see they are approximately the same.

To see how strongly the empirical conditional distribution of the statistic $\sup F_n^*(k;q;\boldsymbol{R})$ depends on the observations y_t, we repeated the simulations 20 times for the problematic example (4.4). The regressors and normal errors

Table 4.4: Empirical quantiles x such that $P(\sup F_n^*(k;q) \leq x/q) = 1-\alpha$ calculated 20 times, each from newly generated samples following the same model.

			$k=2$	$q=2$				$\varepsilon=0.15$	
	0.10	0.05	0.025	0.01		0.10	0.05	0.025	0.01
1	8.57	9.72	10.90	12.45	11	8.47	9.72	10.84	12.63
2	8.96	10.80	12.83	15.47	12	8.56	9.72	10.93	12.47
3	8.39	9.57	10.79	12.32	13	8.51	9.76	10.84	12.36
4	8.57	9.86	10.99	12.61	14	8.39	9.63	10.79	12.33
5	8.48	9.65	10.82	12.05	15	9.22	10.65	11.97	13.92
6	9.63	11.84	14.20	17.11	16	8.45	9.66	10.92	12.53
7	8.63	10.00	11.42	13.36	17	8.36	9.52	10.72	12.49
8	8.63	9.92	11.22	12.73	18	9.01	10.59	12.11	14.32
9	8.45	9.69	10.83	12.35	19	9.07	10.86	12.41	14.48
10	8.63	9.79	10.97	12.56	20	8.47	9.68	10.92	12.39

Sample details: $n = 100$, $e_t \sim N(0,1)$, number of changes $m = 2$, regression coefficients $\boldsymbol{\delta}_1' = (0,1)$, $\boldsymbol{\delta}_2' = (0,1.5)$, $\boldsymbol{\delta}_3' = (0,2)$.

were newly generated each time. The obtained results are in Table 4.3. We can conclude that in this case the empirical quantiles are not very stable and there might be some finite-sample dependence on the observations y_t.

To see if we get more reasonable approximations to critical values when the sample size is large, we simulated the critical values for the example (4.4) using $n = 300$. The calculation took about 28.5 hours. However, the critical values do not agree with ACV, they are again higher than we expected them to be. Their values are 9.399, 11.371, 13.849, 16.669 corresponding to 90 %, 95 %, 97.5 % and 99 % quantiles, respectively. The changes in the slope are perhaps too large.

For comparison we studied in detail another example where the generated data followed a model with 2 smaller changes of size 0.5. Similarly as above, we made our simulation experiments 20 times, see Table 4.4. Here the empirical quantiles are all satisfactory, only three cases out of twenty had a little larger values.

Applying the permutation principle we get satisfactory approximations to critical values of the test $\sup F_n^*(k;q)$ when the changes in regression parameters are small. In this case the conditional distribution of $\sup F_n^*(k;q;\boldsymbol{R})$ does not seem to be much influenced by the observations y_t even when sample size n is about 100.

Chapter 5

Estimating the number of changes

In this chapter we describe the possible methods to estimate the number of changes present in a segmented linear regression model. In Section 5.1 we consider information criteria such as Schwarz' criterion and modified Schwarz' criterion. In Section 5.2 we discuss an alternative method for determining the number of change points that is based on the sequential testing using the tests from Section 3.2.

5.1 Informational approach

We can always improve SSR of our model by allowing more and more change points, ending with every point as a change point. Such a solution is not satisfactory. Therefore we need to impose some penalty for the increased dimension of the model.

Suppose $M < n$ is a known upper bound for the number of change points. We are sure that the true number of changes m never exceeds M. For estimation of the number of parameters in the model (and thus also the number of change points) we will consider an information criterion. The information criteria are based on the log-likelihood function of the model penalised by a term that prevents favouring models with excessive parameters or too many change points.

Yao (1988) suggests to use Bayesian information criterion (Schwarz 1978), also called Schwarz' criterion, to estimate the number of changes for a special case of the model (2.1) with one regression parameter δ_j in the j-th segment:

$$y_t = \delta_j + e_t \qquad t = t_{j-1} + 1, \dots t_j \tag{5.1}$$

for $j = 1, \dots m+1$. Moreover, normality of the errors is assumed. There are together $2m + 2$ unknown parameters in this model: regression coefficients $\delta_1, \ \dots \delta_{m+1}$, the scale parameter σ^2 and the change points $t_1, \dots t_m$. The estimated number of changes $\hat{m}$ minimises the criterion

$$\underbrace{n + n \log 2\pi + n \log \hat{\sigma}_k^2}_{-2 \text{ log-likelihood function}} + \underbrace{(2k+2) \log n}_{\text{penalty term}}$$

subject to $k \leq M$. In the above equation n is the sample size and $\hat{\sigma}_k^2 = S_n(\hat{t}_1, \dots \hat{t}_k)/n$ is the maximum likelihood estimator of σ^2 if k changes are present, see Chapter 2 for the expression of $S_n(\hat{t}_1, \dots \hat{t}_k)$. The two leading terms $n + n \log 2\pi$ are often ignored because it makes no difference when we compare two models from the same sample. The Schwarz' criterion can be equivalently written as

$$BIC(k) = n \log \hat{\sigma}_k^2 + p^* \log n \tag{5.2}$$

for $k \leq M$, where p^* is the total number of fitted parameters. For normal sequence of means we have $p^* = 2k + 2$ parameters and for the regression model (2.1) $p^* = (k+1)q + p + k + 1$.

Yao (1988) further showed in his work that $\hat{m}$ is a consistent estimator of m in the normality case, i.e. the probability that we correctly estimate the number of changes tends to one with the increasing length of the sample.

While Yao (1988) used the Schwarz' criterion for selection of the number of changes, Braun, Braun, and Müller (2000) and Liu, Wu, and Zidek (1997) proposed modified Schwarz' criterion with a different penalty term. Braun, Braun, and Müller (2000) study a location model (5.1) as well but do not assume normality of the errors. They allow for heteroscedasticity of the errors: the variance may differ across the segments but remains the same within each segment. The number of change points m is estimated via minimising

$$B(k) = n \log \hat{\nu}_k^2 + kC_n \tag{5.3}$$

subject to $k \leq M$, where C_n is a penalty term of a form n^δ with $0 < \delta < 1$ and $\hat{\nu}_k^2$ is the minimised average quasi-deviance for a given number of the change points k. They suggest using $\delta = 0.23$ which showed reasonably good results in their simulation studies.

In an independent study Liu, Wu, and Zidek (1997) deal with segmented linear regression models

$$y_t = \boldsymbol{z}_t' \boldsymbol{\delta}_j + e_t \quad \text{if} \quad z_{td} \in (t_{j-1}, t_j], \quad j = 1, \ldots m+1, \;\; t = 1, \ldots, n$$

for the partitioning variable z_{td}, $d \leq q$. For estimation of the number of change points m they minimise the criterion

$$LWZ(k) = n \log \frac{S_n\left(\hat{t}_1, \ldots \hat{t}_k\right)}{n - p^*} + p^* \, c_0 \left(\log n\right)^{2+d_0} \tag{5.4}$$

subject to $k \leq M$, where c_0 and d_0 are some constants, p^* is the total number of fitted parameters and $S_n\left(\hat{t}_1, \ldots \hat{t}_k\right)$ is the optimal SSR with k changes present. They discuss a choice of the constants c_0 and d_0 under various conditions for small and moderate sample sizes. For example a heavier penalty is needed when the error distribution is not normal but heavy-tailed. As described in the article, there exists no best selection of the constants, the choice depends on the model itself and on the sample size. The values $d_0 = 0.1$ and $c_0 = 0.299$ seemed to work satisfactorily in most of their simulation experiments.

There are different information criteria considered in the above papers which forces us to ask a question which criterion is better in which situation. We can use Schwarz' criterion (5.2) for a normal sequence of random variables with shifts in mean, see model (5.1), or for a segmented linear regression (2.1) with normally distributed errors. In the case of nonnormal error distribution, especially when the distribution has heavy tails, it is better if we apply the modified Schwarz' criterion such as (5.4) which has a higher penalty term.

5.2 Sequential methods

Bai and Perron (1998) suggest to apply the tests of k changes versus $k+1$ changes (discussed in Section 3.2) for the selection of the number of changes

in a regression model (2.1). We will now summarise the procedure of testing. We begin with the test of no change versus a single change. If we reject the hypothesis then we proceed to test the null hypothesis of a single change versus two changes and so on, until we can no more reject the null hypothesis of k changes versus $k+1$ changes. The estimated number of changes equals the number of rejections. We can also start the process with testing the null hypothesis of some small number of changes k_0 versus k_0+1 changes, if we think that at least k_0 changes are necessary.

We can apply the tests $LR_n(k+1|k)$ from (3.13) or the tests $F_n(k+1|k)$ from (3.12). In the latter case we estimate the $(k+1)$th change point conditional on the first k changes estimated in previous steps. A sequential procedure based on these conditional tests is proposed in the following papers: theoretical aspects in Bai and Perron (1998), some applications in Bai and Perron (2003a) and simulation experiments in Bai and Perron (2004). The procedure is consistent under some fairly general assumptions on regressors and errors, see Proposition 8 in Bai and Perron (1998). The sequential method based on the likelihood ratio type tests $LR_n(k+1|k)$ is also consistent, for exact formulation of such statement see Theorem 2 in Bai (1999).

Bai (1999) compared the performance of these two sequential methods in finite samples via Monte Carlo simulations. He considered three types of models: simple linear regression, autoregression and linear trend, each with two changes and all coefficients allowing to change. He reported the percentage of rejections for testing the hypothesis of k changes versus $k+1$ changes for $k=0,1,2$ and the distribution of the estimated number of changes. In the reported simulations the test $LR_n(k+1|k)$ showed to have reasonable size and power properties. The results associated with this test were better than those related to the conditional test $F_n(k+1|k)$, but the conditional procedure also worked quite satisfactorily.

Bai and Perron (2004) made an extensive simulation study of the adequacy of the sequential procedure[1] and information criteria[2] used for selection of the number of changes. They also studied the size and power of the F type tests for multiple changes. They presented a variety of models, also allowing

[1] the procedure based on the conditional tests $F_n(k+1|k)$ (3.12) was applied

[2] the Schwarz' criterion (5.2) and the modified Schwarz' criterion (5.4) were used

for serial correlation or different distribution of the errors and regressors across the segments. They analysed cases with no change, one change and two changes and showed how well the procedure and the information criteria select the number of changes. The information criteria worked reasonably well in the absence of serial correlation but did not work very well in the presence of serial correlated errors and heterogeneity. The sequential methods seemed to work better in these cases, but a higher trimming parameter ε was needed.

In some situations the sequential procedure fails. The problem may occur e.g. when we have two changes present and the values of the coefficients after the second change return to their original values. In this case it is easy to identify two changes, but it can be more difficult to identify only one of them. Thus the hypothesis of no change versus 1 change is difficult to reject. In such cases Bai and Perron (2004) suggest to look at the tests $UD\max$ (3.8) or $WD\max$ (3.10) first which have higher power. If these tests indicate the presence of at least one change, then we can estimate the number of changes using the sequential procedure $F_n(k_0+1|k_0)$ starting from a particular $k_0 > 0$ and ignoring the test $F_n(1|0)$ which has a low power.

Chapter 6

Applications

In this chapter we apply the F type test statistics and the discussed methods for determining the number of changes present to a real data set. We consider the average annual temperature series from year 1775 to 1992 measured in Klementinum, Prague (for the data see Figure 6.1). The Klementinum data were often analysed in the literature and different methods were used, see e.g. Antoch and Hušková (1998), Jarušková (1997).

We will test whether the average annual temperatures changed over the passed 218 years, select an appropriate number of changes and finally estimate the parameters of the model. For the calculation we use the software R (version 2.0.1) and its package `strucchange`. For more information on the package see Zeileis, Leisch, Hornik, and Kleiber (2002) and for some applications using this package see Zeileis and Kleiber (2004) and Zeileis, Kleiber, Krämer, and Hornik (2003).

We know the neighboring temperature data may be correlated, but for this application we will assume they are independent. We consider the simple model (5.1)

$$y_t = \delta_j + e_t \qquad t = t_{j-1} + 1, \dots t_j$$

where j denotes the segment index, $j = 1, \dots m + 1$. Here y_t represents the temperature observations for $t = 1775, \dots, 1992$, δ_j is a temperature mean in the j-th period and m is the unknown number of changes in the temperature mean. It is questionable if this is a good model for the temperature series. From Figure 6.1 we can deduce that a linear segmented model would be

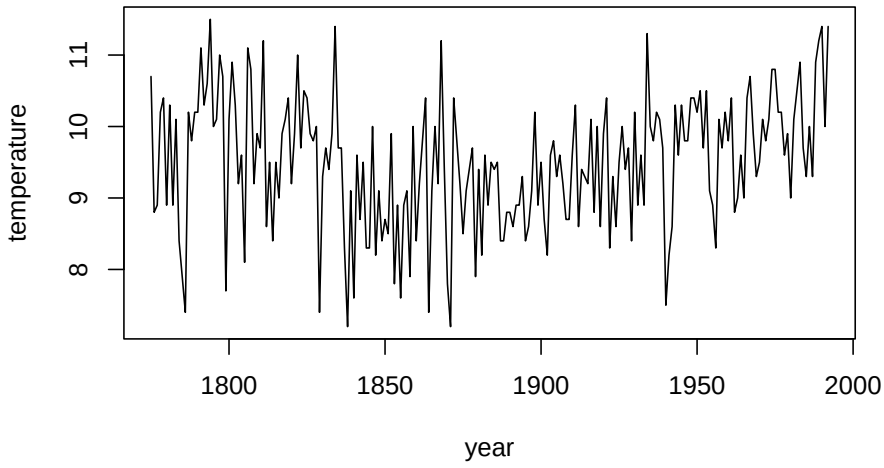


Figure 6.1: Annual temperature series measured in Klementinum, Prague, 1775-1992.

probably a better model, especially the data after 1900 could be well approximated by a linear function. But in this case z_t are trending regressors and the Assumption 2 in Chapter 2 is not satisfied. Some extensions to trending regressors are considered in Bai (1999). In order to have a simple application we remain by the previous model.

The first issue to be considered is to test for presence of structural changes. We consider a trimming parameter $\varepsilon = 0.15$[1], hence each segment has at least 33 observations. At most 5 changes are allowed with $n = 218$ and $\varepsilon = 0.15$.

First we apply the F type test $\sup F_n(1)$ (3.6) of no change versus one change. The values of statistics $F_n(1)$ are plotted in Figure 6.2. The supremum $\sup F_n(1)$ is reached for the year 1942 and has the value 20.4 which highly exceeds the critical value 8.6 at the significance level $\alpha = 0.05$. The tests of no change versus k changes for $k = 2, \ldots, 5$ reject the null hypothesis at the considered level $\alpha = 0.05$ as well. The results are presented in

[1]We obtain the same final results with smaller trimming parameters ε as well, such as $\varepsilon = 0.05$.

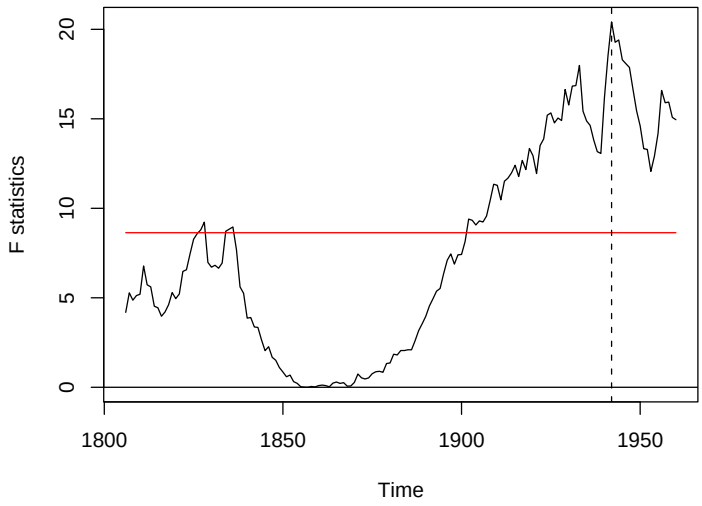


Figure 6.2: F type statistics $F_n(1)$ for Klementinum data. The red line indicates the boundary at the 5 % significance level.

Table 6.1.

We consider the double maximum tests UD max (3.9) and WD max (3.11) where the number of changes k under the alternative hypothesis is not specified, but an upper bound $M = 5$ for k is given. The tests are highly significant and indicate the presence of at least one change.

For selection of the appropriate number of changes we use the Bayesian information criterion BIC (5.2) and the criterion LWZ (5.4) which has a heavier penalty term. Both information criteria select two change points.

The same conclusion emerges from the sequential procedures using the tests $F_n(k+1|k)$ (3.12) and $LR_n(k+1|k)$ (3.13). In Table 6.2 we present the dates of the estimated change points for all optimal segmentations up to the maximal number of changes. The change points were computed using the function `breakpoints` in R. It is evident that the change points from Table 6.2 estimated via global minimisation of SSR (simultaneously) coincide with the change points estimated sequentially (the $(k+1)$-th change point is added conditional on the previous k change points) for up to three

Table 6.1: Empirical results for Klementinum data.

Specifications				
$z_t = \{1\}$	$q = 1$	$n = 218$	$\varepsilon = 0.15$	$M = 5$
Tests[1]				
$\sup F_n^*(1)$	$\sup F_n^*(2)$	$\sup F_n^*(3)$	$\sup F_n^*(4)$	$\sup F_n^*(5)$
20.41^*	25.02^*	19.87^*	15.11^*	10.56^*
$F_n(2\|1)$	$F_n(3\|2)$	$LR_n(1\|0)$	$LR_n(2\|1)$	$LR_n(3\|2)$
24.45^*	7.80	20.60^*	27.54^*	8.09
$UD\max$	$WD\max$			
25.02^*	21.73^*			
Number of changes selected[2]				
Sequential	2			
LWZ	2			
BIC	2			
Parameter estimates with two changes[3]				
$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{t}_1$	$\hat{t}_2$
9.79	9.10	9.99	1836	1942
(0.014)	(0.006)	(0.009)	(1821,1858)	(1936,1952)

[1] We used a 5 % significance level for all tests. The star * above the values of the test statistics indicates their significance at this level.

[2] Sequential methods based on the conditional tests $F_n(k+1|k)$ or tests $LR_n(k+1|k)$, both give the same results

[3] In parentheses are the standard errors for the estimates $\hat{\delta}_j$, $j = 1, 2, 3$ and the 95 % confidence intervals for the change points $\hat{t}_1$ and $\hat{t}_2$ calculated using the function `confint` in R. For theoretical details on confidence intervals see Bai and Perron (2003a).

changes. So the likelihood ratio type test $LR_n(k+1|k)$ and the conditional test $F_n(k+1|k)$ have the same values in their nominators for $k = 0, 1, 2$. The only difference is in their denominators: in the conditional test we insert SSR estimated under the null hypothesis with k changes whereas in the likelihood ratio type test we insert SSR under the alternative hypothesis with an additional change. This implies a little higher value of the test statistic $LR_n(k+1|k)$ than $F_n(k+1|k)$.

Table 6.2: Estimated change points

$m = 1$					1942
$m = 2$	1836				1942
$m = 3$	1836		1902		1942
$m = 4$	1836	1871		1909	1942
$m = 5$	1828	1860	1892	1924	1956

Notes: We considered the minimal length of a segment $h = \lceil 0.15\, n \rceil$. At most five changes are possible for $n = 218$ observations and $\varepsilon = 0.15$.

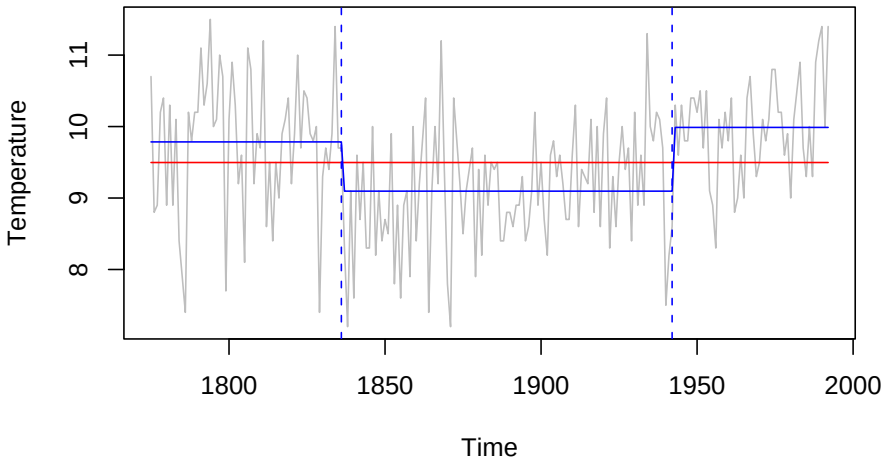


Figure 6.3: The fitted model with two changes ($t_1 = 1836$, $t_2 = 1942$) is represented by blue colour and the fitted model with no change by red colour.

The fitted models with two changes and no change are shown in Figure 6.3. The second change around the year 1942 is disputable. We can observe a linear trend in the temperature series since 1900. Moreover, there was one extremely high and one extremely low temperature measured in the years near the second estimated change. These two years can be considered as outliers. The LS approach used for the estimation of the change points

is not robust to outliers. Therefore, the more appropriate approach could be e.g. the least absolute deviations method (Bai 1998).

Last, we find the approximations to critical values of the tests $\sup F_n^*(k)$, $k = 1, 2, 3$ based on the permutation principle, see Table 6.3. We can see that the empirical critical values agree very well with the asymptotic ones calculated by Bai and Perron (2003b). The permutation test procedures provided a very good approximation to the critical values in this special case.

Table 6.3: Comparison of empirical and asymptotic critical values for the test of no change versus k changes for the Klementinum data

	Empirical CV				Asymptotic CV			
k	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
1	7.01	8.48	10.03	12.23	7.04	8.58	10.18	12.29
2	6.16	7.13	8.03	9.14	6.28	7.22	8.14	9.36
3	5.21	5.93	6.62	7.66	5.21	5.96	6.72	7.60

Notes: Critical values of the test statistics $\sup F_n^*(k)$ correspond to $\varepsilon = 0.15$, $q = 1$ and $k = 1, 2, 3$. The empirical critical values were calculated using 10 000 random permutations.

Conclusion

In this work we studied linear regression models with regression parameters that might change at unknown time points. Our task was to test if some changes in the parameters occurred, to estimate the appropriate number of changes present as well as their locations together with the regression coefficients of the model.

For detection of changes we used F type test statistics based on the difference of sums of squared residuals under the null and alternative hypothesises. We considered tests of no change against k changes where k was either fixed or arbitrary with some upper bound and tests of k changes against $k + 1$ changes.

We proposed an alternative approach to calculate the approximations to the critical values of the test of no change versus k changes with k fixed, namely the approach based on the application of the permutation principle. Usually the approximations based on the limit behaviour of the test statistic under the null hypothesis are used. We proved the asymptotic equivalence of both methods for obtaining the approximations to the critical values. We conducted extended simulation experiments for a number of data samples following models with changes of different size. We showed that the obtained empirical critical values were satisfactory when the data indicated small changes in the parameters.

For estimation of the number of changes we used Schwarz' and modified Schwarz' information criteria. We also described a sequential method based on the application of the tests of k changes against $k + 1$ changes.

All discussed methods were applied to the average annual temperature series from Klementinum, Prague, measured during the period 1775 and 1992. The software R and its package `strucchange` were used.

Appendix A

Proofs

Proof of Theorem 3.1.1. The nominator (3.2) of the test statistic (3.1) under H_0 can be written in the matrix form as

$$SSR_0 - SSR_k = \left(\boldsymbol{Y}'_{1,n}, \boldsymbol{Y}'_{2,n}, \ldots, \boldsymbol{Y}'_{k+1,n}\right) \boldsymbol{D}_n \begin{pmatrix} \boldsymbol{Y}_{1,n} \\ \vdots \\ \boldsymbol{Y}_{k+1,n} \end{pmatrix} \tag{A.1}$$

where

$$\boldsymbol{Y}_{j,n} = \sum_{t=t_{j-1}+1}^{t_j} e_t \boldsymbol{z}_t$$

and

$$d_{ii} = \boldsymbol{C}^{-1}_{t_{i-1},t_i} - \boldsymbol{C}^{-1}_n, \qquad d_{ij} = -\boldsymbol{C}^{-1}_n \quad i \neq j, \qquad i,j = 1,\ldots,k+1 \tag{A.2}$$

are the elements of matrix $\boldsymbol{D}_n$, matrices $\boldsymbol{C}_{t_{i-1},t_i}, \boldsymbol{C}_n$ are defined in (3.4). The errors e_t are i.i.d. and hence in the limit the vectors $\boldsymbol{Y}_{j,n}$ have normal distribution with zero mean and variance matrix

$$\operatorname{Var}\left(\boldsymbol{Y}'_{1,n}, \boldsymbol{Y}'_{2,n}, \ldots, \boldsymbol{Y}'_{k+1,n}\right) = \sigma^2 \begin{pmatrix} \boldsymbol{C}_{t_0,t_1} & \boldsymbol{0} & \cdots & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_{t_1,t_2} & \boldsymbol{0} & \cdots & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \cdots & \boldsymbol{0} & \boldsymbol{C}_{t_k,t_{k+1}} \end{pmatrix}.$$

Using Assumption 2 and then Assumption 1, we get

$$\begin{aligned}
&\mathrm{Var}\left(\boldsymbol{Y}'_{1,n}, \boldsymbol{Y}'_{2,n}, \ldots, \boldsymbol{Y}'_{k+1,n}\right)\boldsymbol{D}_n \\
&\quad = \sigma^2 \begin{pmatrix} (1-\lambda_1)\boldsymbol{I}_q & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & (1-\lambda_2+\lambda_1)\boldsymbol{I}_q & & \vdots \\ \vdots & & & \vdots \\ \vdots & & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & & \lambda_k \boldsymbol{I}_q \end{pmatrix}
\end{aligned} \tag{A.3}$$

where $\boldsymbol{I}_q$ is an identity matrix $q \times q$. Since $\mathrm{Var}\left(\boldsymbol{Y}'_{1,n}, \boldsymbol{Y}'_{2,n}, \ldots, \boldsymbol{Y}'_{k+1,n}\right)$ is a regular matrix and $\left(\boldsymbol{Y}'_{1,n}, \boldsymbol{Y}'_{2,n}, \ldots, \boldsymbol{Y}'_{k+1,n}\right)$ have asymptotically normal distribution, in the limit we can express the quadratic form (A.1) as (Imhof 1961)

$$\sum_{j=1}^{k+1} \sigma^2(1-\lambda_j+\lambda_{j-1})\chi_j^2(q).$$

Here $\sigma^2(1-\lambda_{j-1}+\lambda_j)$ are distinct nonzero characteristic roots of the matrix $\mathrm{Var}\left(\boldsymbol{Y}'_{1,n}, \boldsymbol{Y}'_{2,n}, \ldots, \boldsymbol{Y}'_{k+1,n}\right)\boldsymbol{D}_n$ with multiplicity q and $\chi_j^2(q)$ are independent chi-square variables with q degrees of freedom. Hence the statistic $F_n(k;q)$ (3.1) has asymptotically the distribution

$$\frac{\sum_{j=1}^{k+1}(1-\lambda_j+\lambda_{j-1})\chi_j^2(q)}{kq}.$$

qed

Proof of Theorem 4.1.1. Similarly as in the proof of Theorem 3.1.1 the nominator of the test statistic $F_n(k;q;\boldsymbol{R})$ (4.2) can be written in the matrix form as

$$SSR_0(\boldsymbol{R}) - SSR_k(\boldsymbol{R}) = \left(\boldsymbol{S}'_{1,n}, \boldsymbol{S}'_{2,n}, \ldots, \boldsymbol{S}'_{k+1,n}\right)\boldsymbol{D}_n \begin{pmatrix} \boldsymbol{S}_{1,n} \\ \vdots \\ \boldsymbol{S}_{k+1,n} \end{pmatrix} \tag{A.4}$$

where vectors $\boldsymbol{S}_{j,n}$ are defined in (3.4) and the matrix $\boldsymbol{D}_n$ in (A.2). We calculate mean[1] and variance of $\boldsymbol{S}_j$, $j = 1, \ldots, k+1$ (Anděl 2002)

$$\mathrm{E}\boldsymbol{S}_j = \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \mathrm{E}\hat{e}_{R_t} = \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t = \boldsymbol{0}$$

$$\begin{aligned}
\mathrm{Var}\, \boldsymbol{S}_j &= \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \boldsymbol{z}_t' \mathrm{Var}\, \hat{e}_{R_t} + \sum_{t=t_{j-1}+1}^{t_j} \sum_{\substack{s=t_{j-1}+1 \\ t \neq s}}^{t_j} \boldsymbol{z}_t \boldsymbol{z}_s' \mathrm{Cov}\left(\hat{e}_{R_t}, \hat{e}_{R_s}\right) \\
&= \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t^2 \left(\sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \boldsymbol{z}_t' - \frac{1}{n-1} \sum_{t=t_{j-1}+1}^{t_j} \sum_{\substack{s=t_{j-1}+1 \\ t \neq s}}^{t_j} \boldsymbol{z}_t \boldsymbol{z}_s' \right) \\
&= \frac{1}{n(n-1)} \sum_{t=1}^{n} \hat{e}_t^2 \left[n \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \boldsymbol{z}_t' - \left(\sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \right) \left(\sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \right)' \right] \\
&= \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t^2 \left(\boldsymbol{C}_{t_{j-1}, t_j} - \frac{1}{n} \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \sum_{s=t_{j-1}+1}^{t_j} \boldsymbol{z}_s' \right)
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}\, \left(\boldsymbol{S}_j, \boldsymbol{S}_v\right) &= \mathrm{Cov}\, \left(\sum_{t=t_{j-1}+1}^{t_j} \hat{e}_{R_t} \boldsymbol{z}_t, \sum_{s=t_{v-1}+1}^{t_v} \hat{e}_{R_t} \boldsymbol{z}_s' \right) \\
&= \sum_{t=t_{j-1}+1}^{t_j} \sum_{s=t_{v-1}+1}^{t_v} \boldsymbol{z}_t \boldsymbol{z}_s' \mathrm{Cov}\, \left(\hat{e}_{R_t}, \hat{e}_{R_s}\right) \\
&= -\frac{1}{n-1} \sum_{t=t_{j-1}+1}^{t_j} \sum_{s=t_{v-1}+1}^{t_v} \boldsymbol{z}_t \boldsymbol{z}_s' \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t^2 \\
&\approx -\frac{1}{n} \sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t \sum_{s=t_{v-1}+1}^{t_v} \boldsymbol{z}_s' \frac{1}{n} \sum_{t=1}^{n} \hat{e}_t^2
\end{aligned}$$

[1]Note that $\sum_{t=1}^{n} \hat{e}_t = 0$ because of Assumption 5

We have

$$\mathrm{Var}\left(\boldsymbol{S}'_{1,n}, \boldsymbol{S}'_{2,n}, \ldots, \boldsymbol{S}'_{k+1,n}\right) = \frac{1}{n}\sum_{t=1}^{n} \hat{e}_t^2 \left[\begin{pmatrix} \boldsymbol{C}_{t_0,t_1} & & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_{t_1,t_2} & & \vdots \\ \vdots & & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{C}_{t_k,t_{k+1}} \end{pmatrix} \right.$$
$$\left. - \frac{1}{n} \begin{pmatrix} \sum_{t=t_0+1}^{t_1} \boldsymbol{z}_t \\ \vdots \\ \vdots \\ \sum_{t=t_k+1}^{t_{k+1}} \boldsymbol{z}_t \end{pmatrix} \left(\sum_{t=t_0+1}^{t_1} \boldsymbol{z}'_t, \cdots, \sum_{t=t_k+1}^{t_{k+1}} \boldsymbol{z}'_t \right) \right]$$

Next we calculate $\mathrm{Var}\left(\boldsymbol{S}'_{1,n}, \boldsymbol{S}'_{2,n}, \ldots, \boldsymbol{S}'_{k+1,n}\right)\boldsymbol{D}_n$. Using Assumption 2 and realizing that the vectors

$$\sum_{t=t_{j-1}+1}^{t_j} \boldsymbol{z}_t/(t_j - t_{j-1}) \qquad \text{and} \qquad \sum_{t=1}^{n} \boldsymbol{z}_t/n$$

converge to the first column of the matrix $\boldsymbol{C}$ with elements c_{kl}, $k, l = 1, \ldots, q$ and that

$$\boldsymbol{C}^{-1} \begin{pmatrix} 1 \\ c_{12} \\ \vdots \\ c_{1q} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

we obtain the same matrix (A.3) as in the proof of Theorem 3.1.1. Since $\left(\boldsymbol{S}'_{1,n}, \boldsymbol{S}'_{2,n}, \ldots, \boldsymbol{S}'_{k+1,n}\right)$ have asymptotically normal distribution and their variance matrix is regular, the end of the proof is the same as the end of the previous one. Thus we get that the statistic (4.2) has same limit distribution as the statistic (3.1). *qed*

Appendix B

Tables of simulated critical values

In the following pages we present our simulation results. In Tables B.1-B.6 there are 90 %, 95 %, 97.5 % and 99 % empirical quantiles corresponding to the empirical distribution of the test $\sup F_n^*(k; q; \boldsymbol{R})$ computed using 10 000 random permutations. These quantiles may serve as the approximations to the critical values of the test $\sup F_n^*(k; q)$ (3.7) of no change versus k changes.

At the top of the header of each table we specify the values of the parameters k, q, ε on which the limit distribution of the test $\sup F_n^*(k; q)$ depends. Some specifications related to the sample generations are listed below. Only the sample size n varies, other parameters remain the same in all tables. In the columns there are (ordered from left): the number of actual changes m in the generated data sample, the values of the regression coefficients $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3$ and the empirical quantiles for the situations with normal and Laplace errors. To the end of each table we added the asymptotic critical values (Bai and Perron 1998, Bai and Perron 2003b) of the test $\sup F_n^*(k; q)$ corresponding to the parameters k, q, ε specified in the header of the table.

The obtained results are explained in Section 4.2 beginning at the page 27.

Table B.1: Empirical quantiles x such that $P(\sup F_n^*(k;q) \leq x/q) = 1-\alpha$

				$k=1$ $q=2$ $\varepsilon=0.15$							
			$n=100$	$t_1=n/4$ $t_2=3n/4$ $z_t \sim logN(0,1)$							
				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
0	(0,1)	(0,1)	(0,1)	9.94	11.77	13.37	15.83	10.01	12.10	13.96	16.60
1	(0,1)	($\frac{1}{2}$,1)	($\frac{1}{2}$,1)	9.42	11.20	12.89	15.38	10.02	11.91	13.89	16.39
1	(0,1)	(1,1)	(1,1)	9.84	11.66	13.90	16.62	9.49	11.57	13.73	16.65
1	(0,1)	(0,$\frac{3}{2}$)	(0,$\frac{3}{2}$)	9.96	11.71	13.50	15.95	10.27	12.39	14.65	17.47
1	(0,1)	(0,2)	(0,2)	9.69	11.63	13.95	16.89	9.63	11.56	13.31	15.87
1	(0,1)	($\frac{1}{2}$,$\frac{3}{2}$)	($\frac{1}{2}$,$\frac{3}{2}$)	9.50	11.37	13.09	15.44	9.35	11.79	14.73	19.95
1	(0,1)	(1,2)	(1,2)	10.23	12.02	13.92	16.44	9.88	12.05	14.03	16.88
1	(0,1)	($\frac{1}{2}$,2)	($\frac{1}{2}$,2)	9.75	11.66	13.63	16.11	9.76	12.37	15.79	21.80
1	(0,1)	(1,$\frac{3}{2}$)	(1,$\frac{3}{2}$)	9.49	11.25	12.95	15.49	9.86	11.76	13.61	15.86
2	(0,1)	($\frac{1}{2}$,1)	(0,1)	9.26	11.10	12.83	15.08	9.28	11.06	12.60	15.09
2	(0,1)	(1,1)	(0,1)	9.91	11.68	13.38	15.73	10.14	12.09	14.32	17.12
2	(0,1)	($\frac{1}{2}$,1)	(1,1)	9.62	11.43	13.14	15.59	9.54	11.22	13.00	15.18
2	(0,1)	(1,1)	(2,1)	9.36	11.11	12.81	15.45	9.16	10.91	12.92	15.12
2	(0,1)	(-1,1)	(-$\frac{1}{2}$,1)	9.53	11.39	13.19	15.70	9.69	11.35	13.10	15.42
2	(0,1)	($\frac{1}{2}$,1)	(-$\frac{1}{2}$,1)	9.26	10.97	12.57	14.75	10.09	12.25	14.56	17.62
2	(0,1)	(0,$\frac{3}{2}$)	(0,1)	9.79	11.52	13.39	15.65	10.42	12.88	15.72	19.35
2	(0,1)	(0,2)	(0,1)	10.68	13.49	16.25	20.79	9.63	11.33	13.24	15.61

Continued on Next Page...

Table B.1 – Continued

				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
2	(0,1)	$(0,\frac{3}{2})$	(0,2)	9.84	11.79	13.68	16.15	9.77	11.57	13.52	15.74
2	(0,1)	(0,2)	(0,3)	9.97	13.04	17.29	24.61	10.08	12.47	15.11	18.00
2	(0,1)	$(0,\frac{1}{2})$	(0,1)	9.99	12.49	15.30	18.79	9.92	11.72	13.79	16.20
2	(0,1)	$(0,\frac{1}{2})$	$(0,-\frac{1}{2})$	9.51	11.48	13.42	16.23	9.66	11.69	14.29	17.36
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(0,1)	9.62	11.42	13.21	15.76	9.49	11.30	13.07	15.57
2	(0,1)	(1,2)	(0,1)	9.58	11.39	13.61	16.19	9.84	11.80	13.83	16.60
2	(0,1)	$(\frac{1}{2},\frac{3}{2})$	(1,2)	9.39	11.36	13.10	15.79	9.62	11.48	13.16	15.43
2	(0,1)	$(1,\frac{3}{2})$	(2,2)	9.94	11.63	13.31	15.58	9.89	11.78	13.76	15.87
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(1,1)	10.12	11.78	13.33	15.67	9.78	11.74	13.65	16.42
2	(0,1)	$(1,\frac{3}{2})$	(1,2)	10.31	12.82	15.30	18.58	9.68	12.08	14.80	18.35
2	(0,1)	$(1,\frac{1}{2})$	$(\frac{1}{2},\frac{3}{2})$	10.03	11.94	13.81	16.03	9.94	12.23	14.75	18.22
2	(0,1)	(-1,1)	$(-\frac{1}{2},\frac{3}{2})$	9.86	11.79	13.47	15.85	9.71	11.58	13.33	15.83
BP ACV				9.81	11.47	12.96	15.37	9.81	11.47	12.96	15.37

Table B.2: Empirical quantiles x such that $P(\sup F_n^*(k;q) \le x/q) = 1-\alpha$

				$k=2$ $q=2$ $\varepsilon=0.15$							
				$n=100$ $t_1=n/4$ $t_2=3n/4$ $z_t \sim logN(0,1)$							
				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
0	(0,1)	(0,1)	(0,1)	8.59	9.72	10.80	12.29	8.71	10.08	11.33	12.92
1	(0,1)	($\frac{1}{2}$,1)	($\frac{1}{2}$,1)	8.34	9.53	10.66	12.47	8.56	9.88	11.10	12.77
1	(0,1)	(1,1)	(1,1)	8.61	9.94	11.13	12.93	8.60	10.08	11.47	13.34
1	(0,1)	(0,$\frac{3}{2}$)	(0,$\frac{3}{2}$)	8.58	9.81	11.05	12.32	8.87	10.25	11.55	13.23
1	(0,1)	(0,2)	(0,2)	8.87	10.32	11.68	13.54	8.50	9.71	10.88	12.40
1	(0,1)	($\frac{1}{2}$,$\frac{3}{2}$)	($\frac{1}{2}$,$\frac{3}{2}$)	8.30	9.56	10.73	12.45	9.12	10.93	12.86	15.24
1	(0,1)	(1,2)	(1,2)	8.84	10.20	11.38	12.88	8.80	10.19	11.35	13.16
1	(0,1)	($\frac{1}{2}$,2)	($\frac{1}{2}$,2)	8.62	9.82	11.09	12.70	9.62	11.84	14.03	16.84
1	(0,1)	(1,$\frac{3}{2}$)	(1,$\frac{3}{2}$)	8.52	9.65	10.82	12.24	8.60	9.90	11.13	12.78
2	(0,1)	($\frac{1}{2}$,1)	(0,1)	8.27	9.50	10.69	12.27	8.32	9.49	10.62	12.07
2	(0,1)	(1,1)	(0,1)	8.60	9.78	10.99	12.46	8.82	10.30	11.61	13.26
2	(0,1)	($\frac{1}{2}$,1)	(1,1)	8.47	9.60	10.82	12.30	8.46	9.63	10.78	12.24
2	(0,1)	(1,1)	(2,1)	8.24	9.45	10.63	11.98	8.32	9.53	10.58	12.03
2	(0,1)	(-1,1)	(-$\frac{1}{2}$,1)	8.42	9.55	10.77	12.37	8.48	9.63	10.69	12.10
2	(0,1)	($\frac{1}{2}$,1)	(-$\frac{1}{2}$,1)	8.33	9.34	10.35	11.67	9.11	10.59	12.07	13.88
2	(0,1)	(0,$\frac{3}{2}$)	(0,1)	8.45	9.62	10.72	11.94	9.17	10.83	12.53	14.59
2	(0,1)	(0,2)	(0,1)	9.83	11.94	13.77	16.18	8.48	9.68	10.87	12.18

Continued on Next Page...

Table B.2 – Continued

				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
2	(0,1)	$(0,\frac{3}{2})$	(0,2)	8.74	9.95	11.08	12.36	8.62	9.83	11.07	12.43
2	(0,1)	(0,2)	(0,3)	9.63	12.15	14.68	17.29	9.12	10.72	12.36	14.38
2	(0,1)	$(0,\frac{1}{2})$	(0,1)	9.10	10.81	12.57	14.69	8.66	9.98	11.21	12.73
2	(0,1)	$(0,\frac{1}{2})$	$(0,\text{-}\frac{1}{2})$	9.02	10.35	11.72	13.64	9.33	10.87	12.72	15.06
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(0,1)	8.42	9.63	10.89	12.47	8.59	9.85	11.06	12.44
2	(0,1)	(1,2)	(0,1)	8.65	9.96	11.31	12.71	8.61	9.96	11.16	12.80
2	(0,1)	$(\frac{1}{2},\frac{3}{2})$	(1,2)	8.63	9.90	11.25	12.98	8.43	9.60	10.79	12.07
2	(0,1)	$(1,\frac{3}{2})$	(2,2)	8.44	9.60	10.81	12.07	8.60	9.84	11.07	12.61
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(1,1)	8.61	9.77	11.03	12.36	8.68	9.98	11.20	12.84
2	(0,1)	$(1,\frac{3}{2})$	(1,2)	9.24	10.82	12.33	14.47	9.45	11.28	13.15	15.70
2	(0,1)	$(1,\frac{1}{2})$	$(\frac{1}{2},\frac{3}{2})$	8.61	9.82	11.09	12.37	9.43	11.75	13.84	16.29
2	(0,1)	(-1,1)	$(\text{-}\frac{1}{2},\frac{3}{2})$	8.59	9.82	10.89	12.39	8.48	9.55	10.76	12.18
BP ACV				8.63	9.75	10.75	12.15	8.63	9.75	10.75	12.15

Table B.3: Empirical quantiles x such that $P(\sup F_n^*(k;q) \leq x/q) = 1-\alpha)$

				$k=3$ $q=2$ $\varepsilon=0.15$							
				$n=100$ $t_1=n/4$ $t_2=3n/4$ $z_t \sim logN(0,1)$							
				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
0	(0,1)	(0,1)	(0,1)	7.57	8.54	9.35	10.40	7.47	8.51	9.44	10.79
1	(0,1)	($\frac{1}{2}$,1)	($\frac{1}{2}$,1)	7.29	8.24	9.07	10.43	7.48	8.53	9.41	10.68
1	(0,1)	(1,1)	(1,1)	7.40	8.42	9.41	10.56	7.41	8.52	9.61	11.03
1	(0,1)	(0,$\frac{3}{2}$)	(0,$\frac{3}{2}$)	7.45	8.38	9.38	10.56	7.67	8.76	9.92	11.33
1	(0,1)	(0,2)	(0,2)	7.66	8.76	9.71	11.24	7.41	8.42	9.36	10.70
1	(0,1)	($\frac{1}{2}$,$\frac{3}{2}$)	($\frac{1}{2}$,$\frac{3}{2}$)	7.22	8.24	9.15	10.29	7.70	9.19	10.56	12.58
1	(0,1)	(1,2)	(1,2)	7.61	8.68	9.61	10.83	7.56	8.59	9.60	10.91
1	(0,1)	($\frac{1}{2}$,2)	($\frac{1}{2}$,2)	7.46	8.48	9.47	10.77	7.99	9.53	11.19	13.29
1	(0,1)	(1,$\frac{3}{2}$)	(1,$\frac{3}{2}$)	7.45	8.47	9.39	10.62	7.52	8.56	9.56	10.65
2	(0,1)	($\frac{1}{2}$,1)	(0,1)	7.28	8.23	9.18	10.45	7.33	8.22	9.17	10.37
2	(0,1)	(1,1)	(0,1)	7.46	8.44	9.40	10.50	7.69	8.76	9.76	11.08
2	(0,1)	($\frac{1}{2}$,1)	(1,1)	7.38	8.37	9.29	10.25	7.38	8.31	9.21	10.34
2	(0,1)	(1,1)	(2,1)	7.27	8.22	9.10	10.40	7.32	8.19	9.04	10.36
2	(0,1)	(-1,1)	(-$\frac{1}{2}$,1)	7.36	8.30	9.27	10.57	7.40	8.33	9.22	10.41
2	(0,1)	($\frac{1}{2}$,1)	(-$\frac{1}{2}$,1)	7.27	8.12	8.96	9.91	7.80	9.06	10.15	11.45
2	(0,1)	(0,$\frac{3}{2}$)	(0,1)	7.38	8.32	9.15	10.29	7.70	8.93	10.18	11.77
2	(0,1)	(0,2)	(0,1)	8.28	9.82	11.19	12.87	7.39	8.38	9.28	10.34

Continued on Next Page...

Table B.3 – Continued

				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
2	(0,1)	$(0,\frac{3}{2})$	(0,2)	7.58	8.56	9.49	10.56	7.51	8.55	9.37	10.71
2	(0,1)	(0,2)	(0,3)	7.94	9.78	11.75	13.96	7.78	8.93	10.16	11.66
2	(0,1)	$(0,\frac{1}{2})$	(0,1)	7.67	8.99	10.27	11.68	7.52	8.54	9.51	10.62
2	(0,1)	$(0,\frac{1}{2})$	$(0,\text{-}\frac{1}{2})$	7.70	8.81	10.00	11.51	7.84	9.08	10.45	12.13
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(0,1)	7.33	8.27	9.24	10.37	7.48	8.50	9.44	10.57
2	(0,1)	(1,2)	(0,1)	7.42	8.40	9.47	10.66	7.45	8.40	9.29	10.59
2	(0,1)	$(\frac{1}{2},\frac{3}{2})$	(1,2)	7.45	8.46	9.52	10.90	7.36	8.29	9.24	10.39
2	(0,1)	$(1,\frac{3}{2})$	(2,2)	7.36	8.32	9.19	10.28	7.48	8.38	9.31	10.56
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(1,1)	7.43	8.37	9.32	10.36	7.57	8.58	9.56	10.76
2	(0,1)	$(1,\frac{3}{2})$	(1,2)	7.83	9.08	10.20	11.84	7.97	9.48	10.79	12.60
2	(0,1)	$(1,\frac{1}{2})$	$(\frac{1}{2},\frac{3}{2})$	7.51	8.51	9.39	10.78	7.83	9.53	11.12	12.94
2	(0,1)	(-1,1)	$(\text{-}\frac{1}{2},\frac{3}{2})$	7.56	8.52	9.46	10.41	7.43	8.29	9.16	10.27
BP ACV				7.54	8.36	9.15	10.27	7.54	8.36	9.15	10.27

Table B.4: Empirical quantiles x such that $P(\sup F_n^*(k;q) \leq x/q) = 1-\alpha$

				$k=1$ $q=2$ $\varepsilon=0.15$							
			$n=160$	$t_1=n/4$ $t_2=3n/4$ $z_t \sim logN(0,1)$							
				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
0	(0,1)	(0,1)	(0,1)	9.53	11.23	12.84	14.77	9.45	11.37	13.16	15.59
1	(0,1)	($\frac{1}{2}$,1)	($\frac{1}{2}$,1)	9.30	11.07	12.77	15.00	10.29	12.38	14.32	17.04
1	(0,1)	(1,1)	(1,1)	9.07	10.83	12.58	14.96	9.60	11.41	13.38	15.84
1	(0,1)	(0,$\frac{3}{2}$)	(0,$\frac{3}{2}$)	9.39	11.16	13.18	15.77	9.54	11.47	13.23	15.29
1	(0,1)	(0,2)	(0,2)	9.71	11.51	13.28	15.35	9.88	12.04	14.22	17.32
1	(0,1)	($\frac{1}{2}$,$\frac{3}{2}$)	($\frac{1}{2}$,$\frac{3}{2}$)	9.59	11.36	13.28	15.54	10.26	12.38	14.32	17.04
1	(0,1)	(1,2)	(1,2)	10.17	12.60	15.57	19.92	9.41	11.29	13.08	15.18
1	(0,1)	($\frac{1}{2}$,2)	($\frac{1}{2}$,2)	10.27	12.64	15.44	20.84	9.88	11.88	13.92	16.51
1	(0,1)	(1,$\frac{3}{2}$)	(1,$\frac{3}{2}$)	9.45	11.26	12.96	15.54	9.96	11.93	13.90	16.89
2	(0,1)	($\frac{1}{2}$,1)	(0,1)	9.60	11.49	13.04	15.51	9.42	11.15	12.94	15.13
2	(0,1)	(1,1)	(0,1)	9.95	11.80	13.42	15.53	9.51	11.23	13.15	15.42
2	(0,1)	($\frac{1}{2}$,1)	(1,1)	9.47	11.36	13.28	15.73	9.60	11.36	13.25	15.75
2	(0,1)	(1,1)	(2,1)	9.78	11.54	13.16	15.65	9.83	12.09	14.41	17.38
2	(0,1)	(-1,1)	(-$\frac{1}{2}$,1)	9.57	11.36	13.00	15.18	9.31	11.18	13.03	15.26
2	(0,1)	($\frac{1}{2}$,1)	(-$\frac{1}{2}$,1)	9.51	11.14	12.85	14.80	9.24	10.98	12.68	15.10
2	(0,1)	(0,$\frac{3}{2}$)	(0,1)	9.77	11.43	13.00	15.11	9.84	11.97	13.91	16.21
2	(0,1)	(0,2)	(0,1)	10.01	11.98	14.39	17.36	8.96	11.39	13.79	17.63

Continued on Next Page...

Table B.4 – Continued

				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
2	(0,1)	$(0,\frac{3}{2})$	(0,2)	9.71	11.49	13.24	15.81	9.45	11.49	13.49	15.95
2	(0,1)	(0,2)	(0,3)	9.61	12.06	14.79	18.00	9.39	11.42	13.43	16.33
2	(0,1)	$(0,\frac{1}{2})$	(0,1)	9.65	11.46	13.21	15.28	9.71	11.57	13.42	16.08
2	(0,1)	$(0,\frac{1}{2})$	$(0,-\frac{1}{2})$	9.52	11.33	13.14	15.45	9.60	11.65	13.56	15.87
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(0,1)	9.26	11.12	12.80	15.40	9.82	11.71	13.55	15.96
2	(0,1)	(1,2)	(0,1)	10.08	11.94	13.86	16.68	9.72	11.54	13.39	15.95
2	(0,1)	$(\frac{1}{2},\frac{3}{2})$	(1,2)	9.75	11.76	13.76	16.23	9.87	11.84	13.81	16.28
2	(0,1)	$(1,\frac{3}{2})$	(2,2)	9.61	11.43	13.13	15.50	9.45	11.18	12.91	15.13
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(1,1)	9.54	11.20	13.10	15.40	9.95	11.97	14.12	17.35
2	(0,1)	$(1,\frac{3}{2})$	(1,2)	9.66	11.73	13.95	16.99	9.84	12.00	13.85	16.18
2	(0,1)	$(1,\frac{1}{2})$	$(\frac{1}{2},\frac{3}{2})$	10.12	12.21	14.13	16.58	9.55	11.73	14.04	18.31
2	(0,1)	(-1,1)	$(-\frac{1}{2},\frac{3}{2})$	9.20	11.12	13.17	15.55	9.82	11.84	13.80	16.53
BP ACV				9.81	11.47	12.96	15.37	9.81	11.47	12.96	15.37

Table B.5: Empirical quantiles x such that $P(\sup F_n^*(k;q) \leq x/q) = 1-\alpha$

				$k=2$ $q=2$ $\varepsilon=0.15$							
				$n=160$ $t_1=n/4$ $t_2=3n/4$ $z_t \sim logN(0,1)$							
				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
0	(0,1)	(0,1)	(0,1)	8.34	9.46	10.49	11.78	8.48	9.77	10.86	12.26
1	(0,1)	($\frac{1}{2}$,1)	($\frac{1}{2}$,1)	8.39	9.48	10.51	11.80	8.83	10.16	11.49	13.40
1	(0,1)	(1,1)	(1,1)	8.42	9.56	10.72	12.10	8.44	9.62	10.81	12.24
1	(0,1)	(0,$\frac{3}{2}$)	(0,$\frac{3}{2}$)	8.52	9.83	11.22	12.81	8.43	9.48	10.55	11.95
1	(0,1)	(0,2)	(0,2)	8.48	9.59	10.69	12.40	8.91	10.46	11.96	13.89
1	(0,1)	($\frac{1}{2}$,$\frac{3}{2}$)	($\frac{1}{2}$,$\frac{3}{2}$)	8.34	9.52	10.56	11.86	8.71	10.20	11.51	13.31
1	(0,1)	(1,2)	(1,2)	9.21	11.00	12.82	15.08	8.53	9.74	10.81	12.19
1	(0,1)	($\frac{1}{2}$,2)	($\frac{1}{2}$,2)	9.06	11.25	14.37	17.82	8.67	10.05	11.50	13.56
1	(0,1)	(1,$\frac{3}{2}$)	(1,$\frac{3}{2}$)	8.44	9.63	10.69	12.19	8.65	9.97	11.24	13.05
2	(0,1)	($\frac{1}{2}$,1)	(0,1)	8.28	9.39	10.39	11.70	8.49	9.60	10.83	12.21
2	(0,1)	(1,1)	(0,1)	8.53	9.65	10.78	12.08	8.38	9.55	10.71	12.07
2	(0,1)	($\frac{1}{2}$,1)	(1,1)	8.47	9.62	10.67	12.17	8.55	9.78	10.92	12.40
2	(0,1)	(1,1)	(2,1)	8.38	9.59	10.71	12.00	8.79	10.19	11.50	13.13
2	(0,1)	(-1,1)	(-$\frac{1}{2}$,1)	8.36	9.53	10.63	11.92	8.42	9.56	10.65	12.13
2	(0,1)	($\frac{1}{2}$,1)	(-$\frac{1}{2}$,1)	8.30	9.36	10.41	11.68	8.30	9.48	10.51	11.90
2	(0,1)	(0,$\frac{3}{2}$)	(0,1)	8.37	9.48	10.56	11.91	8.62	9.89	11.31	12.96
2	(0,1)	(0,2)	(0,1)	9.12	10.62	12.09	14.15	8.90	10.53	12.14	14.65

Continued on Next Page...

Table B.5 – Continued

				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
2	(0,1)	$(0,\frac{3}{2})$	(0,2)	8.49	9.65	10.77	12.30	8.70	10.07	11.47	13.05
2	(0,1)	(0,2)	(0,3)	9.67	11.53	13.14	15.29	9.02	10.58	12.16	13.93
2	(0,1)	$(0,\frac{1}{2})$	(0,1)	8.38	9.54	10.50	11.98	8.64	9.91	11.23	12.72
2	(0,1)	$(0,\frac{1}{2})$	$(0,\text{-}\frac{1}{2})$	8.57	9.80	10.95	12.62	8.77	10.20	11.64	13.44
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(0,1)	8.41	9.63	10.86	12.45	8.57	9.74	10.96	12.53
2	(0,1)	(1,2)	(0,1)	8.58	9.80	10.90	12.41	8.59	9.82	11.01	12.65
2	(0,1)	$(\frac{1}{2},\frac{3}{2})$	(1,2)	8.71	10.08	11.25	12.86	8.61	9.83	10.87	12.41
2	(0,1)	$(1,\frac{3}{2})$	(2,2)	8.38	9.57	10.73	12.13	8.30	9.49	10.50	11.84
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(1,1)	8.32	9.49	10.74	12.10	8.66	10.15	11.44	12.99
2	(0,1)	$(1,\frac{3}{2})$	(1,2)	8.73	10.13	11.41	13.20	8.75	10.00	11.24	12.96
2	(0,1)	$(1,\frac{1}{2})$	$(\frac{1}{2},\frac{3}{2})$	8.67	9.91	10.94	12.30	8.84	10.74	12.83	15.45
2	(0,1)	(-1,1)	$(\text{-}\frac{1}{2},\frac{3}{2})$	8.50	9.92	11.17	12.95	8.84	10.15	11.37	13.19
BP ACV				8.63	9.75	10.75	12.15	8.63	9.75	10.75	12.15

Table B.6: Empirical quantiles x such that $P(\sup F_n^*(k;q) \le x/q) = 1-\alpha$

				$k=3$ $q=2$ $\varepsilon=0.15$							
				$n=160$ $t_1=n/4$ $t_2=3n/4$ $z_t \sim logN(0,1)$							
				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
0	(0,1)	(0,1)	(0,1)	7.30	8.16	8.97	9.81	7.43	8.32	9.17	10.27
1	(0,1)	($\frac{1}{2}$,1)	($\frac{1}{2}$,1)	7.35	8.29	9.12	10.06	7.49	8.56	9.64	11.04
1	(0,1)	(1,1)	(1,1)	7.35	8.30	9.10	10.15	7.27	8.24	9.08	10.10
1	(0,1)	(0,$\frac{3}{2}$)	(0,$\frac{3}{2}$)	7.35	8.41	9.38	10.59	7.31	8.21	8.98	10.01
1	(0,1)	(0,2)	(0,2)	7.36	8.26	9.12	10.31	7.47	8.66	9.88	11.38
1	(0,1)	($\frac{1}{2}$,$\frac{3}{2}$)	($\frac{1}{2}$,$\frac{3}{2}$)	7.28	8.22	9.06	10.13	7.42	8.49	9.49	10.98
1	(0,1)	(1,2)	(1,2)	7.81	9.08	10.31	11.88	7.44	8.38	9.24	10.20
1	(0,1)	($\frac{1}{2}$,2)	($\frac{1}{2}$,2)	7.56	9.28	11.28	14.09	7.35	8.50	9.68	11.02
1	(0,1)	(1,$\frac{3}{2}$)	(1,$\frac{3}{2}$)	7.29	8.17	9.02	9.93	7.34	8.37	9.36	10.57
2	(0,1)	($\frac{1}{2}$,1)	(0,1)	7.19	8.02	8.86	9.95	7.38	8.33	9.14	10.19
2	(0,1)	(1,1)	(0,1)	7.35	8.24	9.11	10.14	7.27	8.18	9.00	10.06
2	(0,1)	($\frac{1}{2}$,1)	(1,1)	7.35	8.30	9.08	10.08	7.37	8.33	9.13	10.21
2	(0,1)	(1,1)	(2,1)	7.30	8.15	9.06	10.29	7.49	8.59	9.64	10.86
2	(0,1)	(-1,1)	(-$\frac{1}{2}$,1)	7.21	8.07	8.92	10.02	7.26	8.17	9.06	10.14
2	(0,1)	($\frac{1}{2}$,1)	(-$\frac{1}{2}$,1)	7.21	8.09	8.95	9.98	7.16	8.06	8.93	9.90
2	(0,1)	(0,$\frac{3}{2}$)	(0,1)	7.26	8.20	9.04	9.92	7.36	8.42	9.50	10.68
2	(0,1)	(0,2)	(0,1)	7.69	8.91	10.04	11.66	7.55	8.77	10.07	11.59

Continued on Next Page...

Table B.6 – Continued

				$e_t \sim N(0,1)$				$e_t \sim DE(0,1/\sqrt{2})$			
m	δ_1'	δ_2'	δ_3'	0.10	0.05	0.025	0.01	0.10	0.05	0.025	0.01
2	(0,1)	$(0,\frac{3}{2})$	(0,2)	7.27	8.32	9.23	10.42	7.48	8.49	9.50	10.70
2	(0,1)	(0,2)	(0,3)	7.93	9.44	10.59	12.21	7.54	8.74	9.79	11.35
2	(0,1)	$(0,\frac{1}{2})$	(0,1)	7.29	8.16	9.04	10.18	7.44	8.45	9.41	10.60
2	(0,1)	$(0,\frac{1}{2})$	$(0,\text{-}\frac{1}{2})$	7.34	8.29	9.24	10.37	7.47	8.61	9.64	10.89
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(0,1)	7.34	8.24	9.17	10.38	7.35	8.32	9.25	10.47
2	(0,1)	(1,2)	(0,1)	7.34	8.32	9.24	10.16	7.39	8.35	9.22	10.47
2	(0,1)	$(\frac{1}{2},\frac{3}{2})$	(1,2)	7.46	8.48	9.48	10.72	7.38	8.25	9.18	10.33
2	(0,1)	$(1,\frac{3}{2})$	(2,2)	7.23	8.18	9.07	10.25	7.18	8.14	8.95	9.86
2	(0,1)	$(\frac{1}{2},\frac{1}{2})$	(1,1)	7.20	8.16	9.04	10.21	7.41	8.50	9.53	10.71
2	(0,1)	$(1,\frac{3}{2})$	(1,2)	7.44	8.48	9.45	10.82	7.48	8.57	9.58	10.76
2	(0,1)	$(1,\frac{1}{2})$	$(\frac{1}{2},\frac{3}{2})$	7.41	8.33	9.34	10.35	7.59	9.00	10.46	12.77
2	(0,1)	(-1,1)	$(\text{-}\frac{1}{2},\frac{3}{2})$	7.33	8.37	9.45	10.69	7.52	8.49	9.50	11.02
BP ACV				7.54	8.36	9.15	10.27	7.54	8.36	9.15	10.27

Bibliography

Anděl, J. (2002): *Základy matematické statistiky*. Charles University, Prague, Preprint.

Antoch, J., and M. Hušková (1998): "Estimators of changes," in *Asymptotics, nonparametrics, and time series*, vol. 158 of *Statist. Textbooks Monogr.*, pp. 533–577. Dekker, New York.

Antoch, J., and M. Hušková (2001): "Permutation tests in change point analysis," *Statist. Probab. Letters*, 53, 37–46.

Bai, J. (1998): "Estimation of multiple-regime regressions with least absolutes deviation," *J. Statist. Plann. Inference*, 74(1), 103–134.

——— (1999): "Likelihood ratio tests for multiple structural changes," *J. Econometrics*, 91(2), 299–323.

Bai, J., and P. Perron (1998): "Estimating and testing linear models with multiple structural changes," *Econometrica*, 66(1), 47–78.

——— (2003a): "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, 18(1), 1–22.

——— (2003b): "Critical values for multiple structural change tests," *Econom. J.*, 6(1), 72–78.

——— (2004): "Multiple Structural Change Models: A Simulation Analysis," *Submitted to Econometric Essays*.

Braun, J. V., R. K. Braun, and H.-G. Müller (2000): "Multiple changepoint fitting via quasilikelihood, with application to DNA sequence segmentation," *Biometrika*, 87(2), 301–314.

Chen, J., and A. K. Gupta (2000): *Parametric Statistical Change Point Analysis*. Birkhäuser, Boston.

Hušková, M. (2004): "Permutation Principle and Bootstrap in Change Point Analysis," in *Asymptotic Methods in Stochastics*, vol. 44 of *Fields Institute Communications*, pp. 273–291.

Hušková, M., and J. Antoch (2003): "Detection of Structural Changes in Regression," *Tatra Mountains Mathematical Publications*, 26(2), 201–216.

Imhof, J. P. (1961): "Computing the distribution of quadratic forms in normal variables," *Biometrica*, 48, 419–426.

Jarušková, D. (1997): "Some problems with application of change-point detection methods to environmental data," *Environmetrics*, 8(5), 469–483.

Kim, H.-J., M. P. Fay, E. J. Feuer, and D. N. Midthune (2000): "Permutation tests for joinpoint regression with applications to cancer rates," *Statistics in medicine*, 19(3), 335–351.

Liu, J., S. Wu, and J. V. Zidek (1997): "On segmented multivariate regression," *Statist. Sinica*, 7(2), 497–525.

Schwarz, G. (1978): "Estimating the dimension of a model," *Ann. Statist.*, 6(2), 461–464.

Venter, J. H., and S. J. Steel (1995): "Finding multiple abrupt change points," *Comput. Statist. Data Anal.*, 22(5), 481–504.

Yao, Y.-C. (1988): "Estimating the number of change-points via Schwarz' criterion," *Statist. Probab. Lett.*, 6(3), 181–189.

Zeileis, A., and C. Kleiber (2004): "Validating Multiple Structural Change Models – A Case Study," Report 2, Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Research Report Series.

Zeileis, A., C. Kleiber, W. Krämer, and K. Hornik (2003): "Testing and Dating of Structural Changes in Practice," *Computational Statistics and Data Analysis*, 44(1–2), 109–123.

Zeileis, A., F. Leisch, K. Hornik, and C. Kleiber (2002): "strucchange: An R Package for Testing for Structural Change in Linear Regression Models," *Journal of Statistical Software*, 7(2), 1–38.