Bootstrap procedures for sequential change point analysis in autoregressive models

Zdeněk Hlávka

Charles University in Prague, Faculty of Mathematics and Physics,
Department of Statistics, Prague, Czech Republic

Marie Hušková

Charles University in Prague, Faculty of Mathematics and Physics,
Department of Statistics, Prague, Czech Republic

Claudia Kirch

Karlsruhe Institute of Technology,
Institut für Mathematische Stochastik, Karlsruhe, Germany,

Simos G. Meintanis

National and Kapodistrian University of Athens,
Department of Economics, Athens, Greece

Abstract: We compare numerically the behavior of several bootstrap procedures for monitoring changes in the error distribution of autoregressive time series. The proposed procedures include classical approaches based on the empirical distribution function as well as Fourier-type methods which utilize the empirical characteristic function, both functions being computed on the basis of properly estimated residuals. The Monte Carlo study incorporates different estimators and bootstrap plans and a variety of sampling situations with and without outliers.

2000 Mathematical Subject Class: 62G20, 62E20, 60F17

Keywords: empirical characteristic function, change point analysis

1. Introduction

The detection of structural changes in time series models has received much attention from researchers as it is well known that ignoring those changes could lead to biased out–of–sample predictions, and generally to incorrect statistical assessments. For autoregressive (AR) time series in particular, much work has been focused on the issue of parameter–change. This problem has been studied by following different approaches, and in various degrees of generality, by Lee et al. (2003), Carsoule and Franses (2003), Cai and Davies (2003), Hušková et al. (2007), Hušková et al. (2008), Horváth et al. (2008), Gombay and Serban (2009) and Na et al. (2010). As in most of these papers, our setting here would be sequential and not retrospective, i.e. we shall assume that the time horizon is not fixed, and that our data do not arrive at once, but one–by–one with each time unit. The task is then, in view of each new arrival to decide whether the current AR model is still valid for the data at hand or if some of its aspects might have changed in the meantime. In this paper we shall be concerned with structural breaks due to a change in the distribution of the observations, an issue which has not received so much attention neither in the i.i.d. case nor in the case of dependent observations.

Suppose that X_t are observations on an stationary AR(p) process

$$X_t - \mathbf{X}'_{t-1}\boldsymbol{\beta} = \varepsilon_t, \tag{1}$$

where $\mathbf{X}_t = (X_t, ..., X_{t-p+1})'$, and $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)'$. In (1) the errors ε_t are assumed to be independent with corresponding distribution function denoted by F_t, and that $\mathbb{E}(\varepsilon_t) = 0$ and $\mathbb{E}(\varepsilon_t^2) < \infty$. For some known $T < \infty$, we are interested in testing the hypothesis

$$\mathbb{H}_0 : F_t = F_0, \forall t \tag{2}$$

against the alternative

$$\mathbb{H}_1 : F_t = F_0,\ t \leq T + t_0;\ F_t = F^0,\ t > T + t_0, \tag{3}$$

where the distribution functions $F_0 \neq F^0$ as well as the time of change t_0 are assumed unknown.

In the hypothesis–testing problem exemplified by $\mathbb{H}_i, i = 0, 1$, we suppose that there exists a fixed set $X_1, ..., X_T$, of historic data which involve no change, i.e. that $F_1 = \cdots = F_T$. Based on this data set, we compute the estimate $\widehat{\boldsymbol{\beta}}_T := \widehat{\boldsymbol{\beta}}(X_1, ..., X_T)$ of $\boldsymbol{\beta}$ in model (1). Then, in view of the fact that the errors are unobserved, typically one calculates the residuals

$$\widehat{\varepsilon}_t = X_t - \mathbf{X}'_{t-1}\widehat{\boldsymbol{\beta}}_T, \tag{4}$$

and it is on the basis of these residuals that the null hypothesis $\mathbb{H}_0$ will be tested against $\mathbb{H}_1$. In this paper we consider bootstrap test statistics which are based on the empirical distribution function (EDF) of residuals as well as corresponding statistics which make use of the empirical characteristic function (ECF) of residuals. In particular we study the small sample behavior of Kolmogorov–Smirnov (KS) type tests initially put forward by Lee et al. (2009), and the family of the ECF–based tests suggested by Hlávka et al. (2012). The simulation study was designed so as to allow us to investigate the impact of: (i) Different estimators of parameters, (ii) different types of outliers, (iii) alternative bootstrap schemes, and (iv) different versions of violation of the null hypothesis such as change of scale and change in distribution. Moreover, and since the aforementioned tests are also appropriate for detection of changes in the value of the AR–parameter, we study this aspect of the methods as well.

Now recall that we are operating within an on–line monitoring scheme, so that the test statistic, say S_t, is computed sequentially at each time point, and that the null hypothesis should be rejected when the value of the statistic exceeds an appropriately chosen constant *for the first time*. Otherwise we should continue monitoring. The associated stopping rule is given by

$$\tau(T) = \begin{cases} \inf\{T < t < L_T : S_t > c\}, \\ +\infty \text{ if } S_t \leq c \text{ for all } T < t < L_T, \end{cases}$$

where L_T denotes the duration of the testing period. We distinguish open-end ($L_T = \infty$) and closed-end procedures (with $L_T = \lfloor NT \rfloor + 1$ for some $N > 1$). The closed-end procedures are sometimes also called curtailed or truncated.

The rest of the paper unfolds as follows. In Section 2, the test statistics are defined and discussed, while in Section 3 we address some aspects of the ECFbased methods. Numerical details on estimation and testing are provided in Section 4, while bootstrap procedures are analyzed in Section 5. Section 6 is devoted to a detailed study of the finite–sample behavior of the methods. Finally, our results are summarized in Section 7.

2. Test statistics

On the basis of the residuals $\widehat{\varepsilon}_t$ defined in (4), classical procedures for testing the null hypothesis $\mathbb{H}_0$ utilize the EDF,

$$\widehat{F}_{t_1,t_2}(z) = \frac{1}{t_2 - t_1} \sum_{t=t_1+1}^{t_2} \mathbb{I}(\widehat{\varepsilon}_t \leq z),$$

where $\mathbb{I}(\cdot)$ denotes the indicator function. In particular, the Kolmogorov–Smirnov (KS) statistic is defined by

$$KS_t(\gamma) := \sqrt{T} \left(\frac{t}{T+t} \right)^{(1+\gamma)/2} \sup_{-\infty<z<\infty} |\widehat{F}_{T,T+t}(z) - \widehat{F}_{p,T}(z)|, \tag{5}$$

$\gamma \in (0,1]$ being a normalizing parameter which controls the size of the test. The advantage of the KS statistic is that its limit distribution does not depend on the underlying error distribution (see for instance Lee et al., 2009), provided that this error distribution is sufficiently smooth.

An alternative approach which is suggested here utilizes the ECF of the residuals

$$\widehat{\varphi}_{t_1,t_2}(u) = \int_{-\infty}^{\infty} e^{iuz} d\widehat{F}_{t_1,t_2}(z) = \frac{1}{t_2 - t_1} \sum_{t=t_1+1}^{t_2} e^{iu\widehat{\varepsilon}_t},$$

and rejects the null hypothesis $\mathbb{H}_0$ for large values of

$$CF_t(\gamma) := T \left(\frac{t}{T+t} \right)^{1+\gamma} \int_{-\infty}^{\infty} |\widehat{\varphi}_{T,T+t}(u) - \widehat{\varphi}_{p,T}(u)|^2 w(u) du, \tag{6}$$

where $\gamma \in (0,1)$ and $w(u)$ is a weight function which is introduced in order to smooth out the periodic components of the ECF. This approach was suggested by Hlávka et al. (2012) that also studied the asymptotic properties of the resulting test statistic. We now outline some of these properties. In particular, the limit distribution of the ECF statistic does not require that the underlying true–error distribution is smooth, and it can be shown that it is also free of the estimator of $\boldsymbol{\beta}_T$, as long as

$$\sqrt{T}(\widehat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}) = O_P(1), \text{ as } T \to \infty. \tag{7}$$

A standard estimator is the least squares (LS) estimator, but more general estimators, such as M–estimators, also satisfy (7). Moreover given (7), the aforementioned limit distribution remains the same if we replace the residuals $\widehat{\varepsilon}_t$ by the true errors ε_t.

To get the limit distribution of the ECF test statistic with ε_t in the place of $\widehat{\varepsilon}_t$, it can be shown that for $k = \lfloor Ts \rfloor$, $s > 0$ fixed , $T \to \infty$, the limit behavior of

$$T\Big(\frac{k}{k+T}\Big)^{\gamma+1} \int_{-\infty}^{\infty} |\widehat{\varphi}_{T,T+k}(u) - \widehat{\varphi}_{p,T}(u)|^2 w(u) du$$

is the same as that of

$$\Big(\frac{k}{T+k}\Big)^{\gamma} \Big(\int w(t)dt - \mathbb{E}\left[h_w(\varepsilon_1, \varepsilon_2)\right] \Big) + \sum_{j=1}^{\infty} \lambda_j \frac{W_j^2(k/(T+k)) - k/(T+k)}{(k/(T+k))^{1-\gamma}},$$

where $h_w(x,y) := h_w(x-y) = \int \cos(u(x-y))w(u)du$, $W_{j,1}(\cdot)$, $j \geq 1$, are independent Wiener processes, and λ_j, $j \geq 1$, are eigenvalues which depend on the underlying distribution function F_t. For further details on the asymptotic behavior of the ECF statistic (asymptotic null distribution, behavior under alternatives, etc) the reader is referred to Hlávka et al. (2012), but the main message is clear: Since F_t is unknown, one should use a bootstrap procedure in order to actually carry out the test. Moreover, this bootstrap procedure should be appropriate for the abovementioned sequential setup under which we operate.

3. FLS estimation and ECF statistics

In this section we consider a more general type of estimation than LS termed functional least squares (FLS) estimator. In addition specific moment–based interpretations are provided for the FLS estimator and the ECF–based statistics.

3.1. **FLS estimation.** The FLS estimator was proposed by Heathcote and Welsh (1983). In particular, and for fixed $u \in \mathbb{R}$, this estimator is obtained by minimizing the loss function

$$L_T(\boldsymbol{\beta}, u) = -\frac{1}{u^2} \log |\widehat{\varphi}_{p,T}(u)|^2 . \tag{8}$$

Under standard assumptions the FLS method produces consistent and asymptotically normal estimators (that satisfy (7)), and which are robust to 'innovation' outliers, i.e., to outliers with respect to the error distribution; see Dhar (1993) and Meintanis and Donatos (1999). We also note that a data–driven FLS estimator which is efficient can be obtained by choosing the argument u so that a non–parametric estimate of the corresponding asymptotic variance is minimized; refer to Csörgő (1983), and to §4 below.

A drawback in FLS estimation is that since

$$|\widehat{\varphi}_{p,T}(u)|^2 = \frac{1}{(T-p)^2} \sum_{t,s=p+1}^{T} \cos\left[u(\widehat{\varepsilon}_t - \widehat{\varepsilon}_s)\right], \tag{9}$$

the FLS–criterion is location invariant and therefore the intercept can not be estimated. Despite the fact that Welsh (1985) proposed a variation of the FLS procedure in which the intercept can also be estimated, in what follows we shall only consider the case of no intercept.

It is illuminating to further investigate which factors play part in the FLS loss function $L_T(\boldsymbol{\beta}, u)$ in (8). To this end, from (9) and by simple Taylor expansions of $\cos(u)$ and $\log(1+u)$ one has after some algebra that

$$\log |\widehat{\varphi}_{p,T}(u)|^2 = -\frac{u^2}{2} M_2 + \frac{u^4}{24}\left(M_4 - 3M_2^2\right) + \frac{u^6}{720}\left(15M_2M_4 - M_6 - 30M_2^3\right) + o(u^6), \tag{10}$$

as $u \to 0$, where $M_k = (T-p)^{-2} \sum_{t,s=p+1}^{T} (\widehat{\varepsilon}_t - \widehat{\varepsilon}_s)^k$, $k = 2, 3, \ldots$. Equation (10) contains powers of u and associated contrasts incorporating empirical moments (in forms reminiscent of V–statistics) computed from the FLS residuals. It is particularly interesting that under the assumption of normally distributed errors, and as $T \to \infty$, the coefficients of u^k in (10) vanish identically for $k > 2$. This is because the aforementioned contrasts hold true for the normal distribution; for instance the coefficient of u^4 corresponds to the moment equation $\mathbb{E}(X^4) = 3(\mathbb{E}(X^2))^2$, while the coefficient of u^6 corresponds to the moment equation $\mathbb{E}(X^6) = 15\mathbb{E}(X^4)\mathbb{E}(X^2) - 30(\mathbb{E}(X^2))^2$, which are both valid under normality. Hence it becomes transparent that when estimating $\boldsymbol{\beta}$ on the basis of the loss function $L_T(\boldsymbol{\beta}, u)$, it is unnecessary, at least asymptotically, to go beyond u^2 in that Taylor expansion when the errors are normal. In fact by replacing (10) in (8) one has $\lim_{u\to 0} L_T(\boldsymbol{\beta}, u) = M_2$. Now M_2 coincides with the loss function of the LS–slope coefficients in deviation form, and this in turn shows that LS is a special case of the FLS estimator corresponding to $u = 0$. In addition, this fact renders the FLS with a sense of optimality as it is well known that under normal errors LS estimation is optimal.

3.2. **ECF statistics.** By similar expansions as in (9) and (10) one has

$$
\begin{aligned}
|\widehat{\varphi}_{T,T+t}(u) - \widehat{\varphi}_{p,T}(u)|^2 &= u^2(m^{(1)}_{T,T+t} - m^{(1)}_{p,T})^2 \\
&+ \frac{u^4}{12}\left[3(m^{(2)}_{T,T+t} - m^{(2)}_{p,T})^2 - 4(m^{(1)}_{T,T+t} - m^{(1)}_{p,T})(m^{(3)}_{T,T+t} - m^{(3)}_{p,T})\right] \\
&+ \frac{u^6}{360}[6(m^{(1)}_{T,T+t} - m^{(1)}_{p,T})(m^{(5)}_{T,T+t} - m^{(5)}_{p,T}) + 10(m^{(3)}_{T,T+t} - m^{(3)}_{p,T})^2 \\
&- 15(m^{(2)}_{T,T+t} - m^{(2)}_{p,T})(m^{(4)}_{T,T+t} - m^{(4)}_{p,T})] + o(u^6),\ u \to 0,
\end{aligned}
$$

where $m^{(k)}_{t_1,t_2} = (t_2 - t_1)^{-1}\sum_{\tau=t_1+1}^{t_2} \widehat{\varepsilon}^k_\tau,\ k = 1, 2, \ldots$. It is transparent from the equation above that moment matching takes place in the ECF– statistic, and that this matching is between the sample moments computed from $\{\widehat{\varepsilon}_\tau\}_{\tau=T+1}^{T+t}$ (residuals following the training period) and the corresponding sample moments computed from $\{\widehat{\varepsilon}_\tau\}_{\tau=p+1}^{T}$ (residuals during the training period). In addition, by looking at eq. (6) it becomes clear that the role of the weight function $w(u)$ is to assign specific weights with which each of these moment–matching equations enters the test statistic. In this connection, a weight function with rapid rate of decay assigns little weight to higher sample moments, and the value of the test statistic is then dominated by matching the low–order moments of $\{\widehat{\varepsilon}_\tau\}_{\tau=T+1}^{T+t}$ and $\{\widehat{\varepsilon}_\tau\}_{\tau=p+1}^{T}$. In contrast a slower rate of decay of $w(u)$ allows moments of higher order to also have an significant impact on the test statistic. Standard choices such as $w(u) = \exp(-a|u|^m), a, m > 0$, in (6) yield (following some algebra), for $m = 1$ the limit statistic

$$
CF^{(1)}_{t,\infty} := \lim_{a\to\infty} a^3 CF^{(1)}_{t,a} = 4T\left(\frac{t}{T+t}\right)^{1+\gamma}(m^{(1)}_{T,T+t} - m^{(1)}_{p,T})^2,
$$

and for $m = 2$, the limit statistic

$$
CF^{(2)}_{t,\infty} := \lim_{a\to\infty} a^{3/2} CF^{(2)}_{t,a} = \frac{\sqrt{\pi}}{2}T\left(\frac{t}{T+t}\right)^{1+\gamma}(m^{(1)}_{T,T+t} - m^{(1)}_{p,T})^2,
$$

where $CF^{(m)}_{t,a},\ m = 1, 2$, denotes the test statistic in (6) with weight function $e^{-a|u|^m}$. These limit statistics underline the fact that extreme smoothing of the periodic component of the ECF leads to tests which are based solely on sample means computed from the residuals. Specifically we should reject the null hypothesis when large values are observed in the sequential matching between the sample means of the residuals before and after the end of the training period.

4. Computational issues

The FLS estimating equations are given by $\boldsymbol{\ell}_T(\boldsymbol{\beta}, u) = \mathbf{0}$ where $\boldsymbol{\ell}_T := \partial L_T/\partial\boldsymbol{\beta}$ is a vector of dimension p with m^{th} element

$$\ell_{T,m}(\boldsymbol{\beta}, u) = \frac{1}{(T-p)^2}\frac{1}{u}\sum_{t,s=p+1}^{T} X_{t-m} \sin\left[u(X_t - X_s) - u(\mathbf{X}'_{t-1} - \mathbf{X}'_{s-1})\boldsymbol{\beta}\right],$$

$m = 1, ..., p$. The FLS–equation is solved iteratively. In order to outline a numerical procedure for an efficient FLS estimate, write $\widehat{\boldsymbol{\beta}}_u$ for the FLS estimate and $V(\varphi(u), u)$ for the corresponding asymptotic FLS variance. The formula for this asymptotic variance in terms of the characteristic function of the underlying error distribution $\varphi(u)$ can be found in Meintanis and Donatos (1999). Then each step of the iteration can be outlined as follows:

- For fixed u, obtain the current value of the estimate $\widehat{\boldsymbol{\beta}}_u$
- Obtain the current residuals $\widehat{\varepsilon}_t(u) := \widehat{\varepsilon}_t(\widehat{\boldsymbol{\beta}}_u)$
- Compute the corresponding ECF $\widehat{\varphi}_{p,T}(u)$
- Estimate the asymptotic FLS–variance by $\widehat{V}_u = V(\widehat{\varphi}_{p,T}(u), u)$

Then the current FLS estimator is obtain as $\widehat{\boldsymbol{\beta}}_{u_{min}}$ with $u_{min} = \mathrm{argmin}_u \widehat{V}_u$, i.e. the efficient FLS estimator is computed at the argument which minimizes the estimate of the FLS limit variance over a specified interval.

The initial value for the above iterative algorithm is obtained by using function `lmrob()` in the R library `robustbase` (R Development Core Team; 2011; Rousseeuw et al.; 2011). The function `lmrob()` computes an MM-type regression estimator as described in Yohai (1987) and Koller and Stahel (2011).

As proposed in Meintanis and Donatos (1999), the asymptotic FLS-variance $\widehat{V}_u$ is minimized over $u \in [0.001, 1.0]$ and, for the purpose of FLS estimation, the time series X_t is standardized by the median of its absolute values, i.e., the FLS estimator is calculated from observations $X_t/\mathrm{med}|X_t|$, a fact that guarantees scale invariance.

4.1. **Recursive formula for ECF test statistic.** Computationally convenient expressions for the ECF test statistic may be obtained from (6) by straightforward algebra as, $CF_k(\gamma) = T(k/(T+k))^{1+\gamma}\Sigma_k$ where

$$\Sigma_k = \frac{1}{k^2}S_{1,k} + \frac{1}{(T-p)^2}S_{2,T} - 2\frac{1}{k(T-p)}S_{3,k}, \tag{11}$$

with

$$S_{1,k} = \sum_{t,s=T+1}^{T+k} h_w(\hat{\varepsilon}_t - \hat{\varepsilon}_s),\ S_{2,k} = \sum_{t,s=p+1}^{k} h_w(\widehat{\varepsilon}_t - \widehat{\varepsilon}_s),\ S_{3,k} = \sum_{t=T+1}^{T+k}\sum_{s=p+1}^{T} h_w(\widehat{\varepsilon}_t - \widehat{\varepsilon}_s).$$

In fact, some further algebra shows that the computation of the test statistic is facilitated by the recursive relations

$$S_{1,k+1} = S_{1,k} + 2\sum_{t=T+1}^{T+k} h_w(\widehat{\varepsilon}_t - \widehat{\varepsilon}_{T+k+1}) + h_w(0), \quad S_{3,k+1} = S_{3,k} + \sum_{t=p+1}^{T} h_w(\widehat{\varepsilon}_t - \widehat{\varepsilon}_{T+k+1}),$$

(for $S_{2,k}$ the recursive relation is obtained from that of $S_{1,k}$ by simply setting $T = 0$).

5. Bootstrap

The validity of the bootstrap approximation in Section 5.1 has been established in Hlávka et al. (2012, Theorem 3.1). These authors provide strong motivation for using the bootstrap approximation by establishing the fact (as it was already mentioned in Section 2) that the limit distribution of the ECF test statistics is not distribution free. In this section, we describe two simple bootstrap schemes that will be closer investigated in the simulation study in Section 6.

5.1. **Classical bootstrap.** The bootstrap scheme which employs estimated residuals based only on the training data has been used in a small simulation study in Hlávka et al. (2012): let $U_T(p+1), \ldots, U_T(\tilde{L}_T)$ be i.i.d. uniformly distributed on the set of indices $\{p+1, \ldots, T\}$ and independent of $\{X_t\}$, where we choose $\tilde{L}_T = L_T - 1 = \lfloor NT \rfloor$ in case of the closed-end procedure and $\tilde{L}_T/T \to \infty$ in case of the open-end procedure. Let us define bootstrap residuals

$$\varepsilon_t^* = \widehat{\varepsilon}_{U_T(t)}$$

for $t \geq 1$ with $\widehat{\varepsilon}_t$ defined in (4). The bootstrap critical value $c_\alpha(X_1, \ldots, X_T)$ is then chosen as the minimal value such that

$$P_T^* \left[\max_{\{T < t \leq \tilde{L}_T\}} \{CF_t^*(\gamma)\} \leq c_\alpha(X_1, \ldots, X_T) \right] \geq 1 - \alpha.$$

where $CF_t^*(\gamma)$ denotes the test statistic (6) calculated from the resampled residuals ε_t^* and the conditional distribution $P_T^*(\cdot) = P(\cdot \,|\, X_1, \ldots, X_T)$. This conditional distribution can be easily simulated by drawing B random realizations of $\{U_T(\cdot)\}$.

For this statistic we can use uncentered bootstrap residuals ε_t^*, because by (11) the value of the test statistic depends only on differences between residuals, so that centering does not change the value of the test statistic.

5.2. **Sequential bootstrap.** The classical bootstrap uses only the residuals $\widehat{\varepsilon}_{p+1}, \ldots, \widehat{\varepsilon}_T$ obtained from the starting sample of size T; see (4). The idea of sequential bootstrapping (Kirch; 2008; Hušková and Kirch; 2011) is to repeat the bootstrap procedure several times in order to use the increased knowledge obtained during the monitoring. Unfortunately, it would be computationally too expensive to completely (re-)generate all B bootstrap samples in order to update the bootstrap critical value after each new incoming observation. Therefore, Hušková and Kirch (2011) propose a compromise approach that calculates new critical values only after each L-th observation. Moreover, in each of these steps, only a small percentage, $(1/M) \times 100\%$, of "oldest" bootstrap samples is discarded and replaced by the more up-to-date bootstrap resamples.

The algorithm of the sequential bootstrap is just a minor modification of the classical bootstrap. The main differences are:

(1) during the sequential monitoring of the incoming time series, B/M oldest bootstrap samples are replaced after the arrival of each L-th observation,
(2) in j-th step of the sequential bootstrap algorithm, the estimator $\widehat{\beta}_{T+jL}$ is used instead of $\widehat{\beta}_T$ in order to obtain residuals $\widehat{\varepsilon}_{p+1}, \ldots, \widehat{\varepsilon}_{T+jL}$ that are used for the new bootstrap resamples.

This procedure leads to updated critical values after each new bootstrap step, so that the critical values now depend on the time point t as well.

In our simulation study in Section 6 we use the values $L = M = 5$ recommended by (Hušková and Kirch; 2011) with $B = 1000$, which means that we will replace one fifth (i.e., $B/M = 1000/5 = 200$) oldest bootstrap replicates after each 5-th new incoming observation.

6. Simulation study

The training sample is always an AR(1) process with the regression parameter $\beta = \beta_0 = 0.4$ and the error terms have standard deviation $\sigma_0 = 1$ (if the random error distribution has finite variance). The symbols $\beta^0 = \beta_0 + \delta$ and σ^0 denote the values of the same parameters after the change occurring at the change-point t_0. The results in Tables 1–5 are obtained using 2000 simulations, the bootstrap approximation uses $B = 2000$ bootstrap replicates and the parameters of the sequential bootstrap are set to $L = M = 5$.

We consider the following types of change:

- change in the regression coefficient,
- change in scale,
- change in the error distribution.

We use the following distributions of the error terms:

- standard Normal,
- standard Normal distribution with innovation outliers (with probability 0.1, the distribution of the error term is contaminated by N(0,100) distribution),
- standardized double exponential (Laplace) distribution,
- standardized χ^2-distribution,
- Student's t-distribution (standardized for >2 d.f. and centered for 2 d.f.)

Apart of the innovation outliers, we investigate also robustness with respect to additive outliers: in this case, we generate the time series with standard Normal random errors and contaminate the simulated X_t's by adding a random deviate following a $N(0, 100)$ distribution, with probability 0.1.

Tuning parameters. We compare the behavior of the EDF-based KS test statistic and ECF test statistics with weight functions $w_1(u) = \exp(-a_1|u|)$ and $w_2(u) = \exp(-a_2u^2)$. A preliminary simulation study, not included here, suggested to set $a_1 = 1$ and $a_2 = 1/2$. In order to adapt our procedure to a possibly different scale of observations, we calculate the sample standard deviation $\widehat{\sigma}_T$ from the residuals (4) obtained from the starting sample and set $a_1 = \widehat{\sigma}_T$ and $a_2 = \widehat{\sigma}_T^2/2$ in the test statistics CF_1 and CF_2, respectively. For the value of the remaining parameter γ in (5) and (6) we set $\gamma = 1$.

Apart from the test statistics (KS, CF_1, and CF_2), we compare also the classical and sequential bootstrap (CB and SB) described in Section 5, and two methods of the regression parameter estimation (LS and FLS). Note that, in our simulation study, the value of the tuning parameter u minimizing the estimated asymptotic variance of the FLS estimator, described in Section 4, is chosen only from the starting sample of size T and the parameter u is fixed during the sequential bootstrap algorithm.

Empirical level and achieved power. The results obtained for various starting sample sizes T, monitoring periods of length $(N-1)T$, and changepoints t_0 are summarized in Tables 1–5 containing empirical level and achieved power corresponding

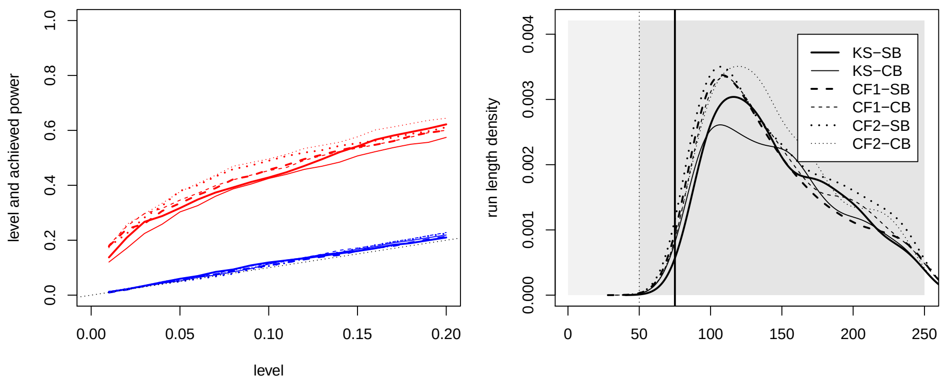

FIGURE 1. Change in the regression parameter, $T = 50$, $N = 5$, $t_0 = 25$, LS estimator.

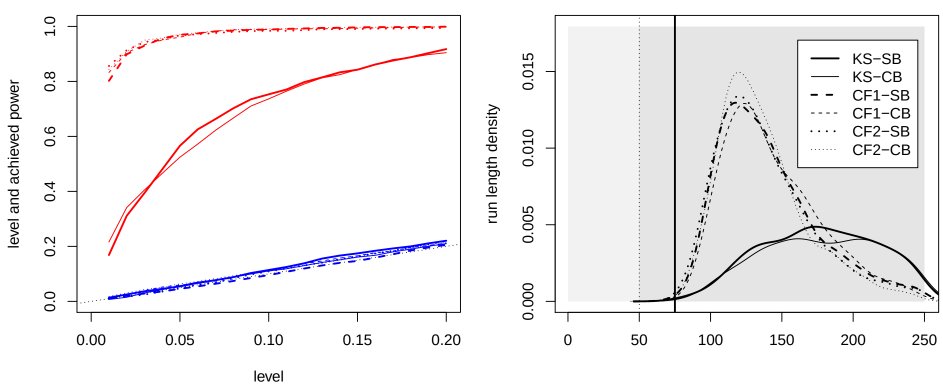

FIGURE 2. Change in scale, $T = 50$, $N = 5$, $t_0 = 25$, LS estimator.

to the nominal level $\alpha = 0.05$. Note that the so-called achieved power describes the size-corrected power, i.e., the empirical power of the test corresponding to the true level set to α. More precisely, we first establish the nominal level α^* such that the corresponding empirical level obtained in our simulation study is equal to $\alpha = 0.05$ and the achieved power is then the (usual) empirical power of the test with the adjusted nominal level α^*. We use the achieved power because it adjusts for different empirical significance levels and thus it allows a more straightforward comparison of two tests.

Change in scale or regression parameter. Tables 1 and 2 show that the achieved power of all tests increases for earlier changepoints and with increasing length of monitoring. All test statistics behave similarly against changes in the regression parameter but ECF test statistics have better power against changes in the scale (standard deviation) of the random errors. Using FLS leads to overly high

		t_0		emp. level LS CB	emp. level LS SB	emp. level FLS CB	emp. level FLS SB	ach. power LS CB	ach. power LS SB	ach. power FLS CB	ach. power FLS SB
$T=20$	$N=5$	10	KS	0.059	0.062	0.115	0.141	0.265	0.275	0.185	0.175
			CF_1	0.066	0.048	0.114	0.108	0.253	0.263	0.228	0.227
			CF_2	0.062	0.060	0.133	0.121	0.290	0.263	0.240	0.223
		40	KS	0.066	0.059	0.114	0.146	0.116	0.132	0.097	0.096
			CF_1	0.069	0.052	0.121	0.110	0.102	0.118	0.102	0.111
			CF_2	0.076	0.057	0.135	0.115	0.114	0.130	0.101	0.102
	$N=10$	10	KS	0.049	0.069	0.113	0.130	0.270	0.281	0.184	0.178
			CF_1	0.067	0.048	0.135	0.109	0.263	0.289	0.216	0.215
			CF_2	0.066	0.051	0.140	0.143	0.298	0.315	0.232	0.226
		40	KS	0.054	0.070	0.098	0.144	0.155	0.159	0.120	0.120
			CF_1	0.071	0.054	0.114	0.112	0.157	0.166	0.147	0.134
			CF_2	0.064	0.058	0.148	0.132	0.193	0.189	0.147	0.146
$T=50$	$N=5$	25	KS	0.051	0.059	0.080	0.079	0.303	0.317	0.277	0.281
			CF_1	0.056	0.057	0.082	0.079	0.346	0.335	0.311	0.308
			CF_2	0.060	0.050	0.084	0.071	0.377	0.380	0.354	0.352
		100	KS	0.051	0.048	0.070	0.091	0.141	0.141	0.122	0.120
			CF_1	0.052	0.044	0.083	0.077	0.163	0.150	0.136	0.123
			CF_2	0.061	0.050	0.089	0.077	0.155	0.150	0.137	0.131
	$N=10$	25	KS	0.049	0.057	0.078	0.091	0.322	0.322	0.268	0.254
			CF_1	0.057	0.048	0.086	0.083	0.381	0.381	0.338	0.338
			CF_2	0.062	0.058	0.085	0.068	0.385	0.406	0.345	0.414
		100	KS	0.040	0.052	0.076	0.084	0.224	0.206	0.168	0.174
			CF_1	0.054	0.051	0.083	0.067	0.240	0.233	0.224	0.219
			CF_2	0.060	0.055	0.092	0.081	0.235	0.237	0.207	0.208

TABLE 1. The regression parameter changes from $\beta_0 = 0.4$ to $\beta^0 = 0.8$, random errors with standard Normal distribution.

empirical levels (approx. 0.1–0.14) for $T = 20$ with a small improvement (0.056–0.106) for $T = 50$. For these sample sizes, the CB and the SB lead to comparable results.

For illustration, we display the results using LS residuals for $T = 50$, $N = 5$, and $t_0 = 25$ graphically in Figures 1 and 2. The left display shows the observed empirical level and the achieved power and the right display shows the estimated run-length density corresponding to the achieved level $\alpha = 0.05$.

Change in distribution. Table 3 shows that the empirical level for FLS improves for higher starting sample sizes and also that the SB leads to empirical levels that seem to be somewhat more stable and closer to the nominal value 0.05. For detecting a change in the error distribution, the achieved power of the CF test statistics is better than the achieved power of the KS test statistic.

Outliers. Table 4 and Figure 5 show results concerning change in the regression parameter in a situation with innovation outliers: The results suggest that the ECF based test statistics outperform the KS test and that the FLS seems to perform somewhat better than LS for $T = 50$.

		t_0		emp. level LS CB	emp. level LS SB	emp. level FLS CB	emp. level FLS SB	ach. power LS CB	ach. power LS SB	ach. power FLS CB	ach. power FLS SB
$T = 20$	$N = 5$	10	KS	0.055	0.070	0.110	0.134	0.217	0.199	0.105	0.115
			CF_1	0.066	0.050	0.126	0.117	0.532	0.541	0.398	0.306
			CF_2	0.065	0.050	0.141	0.123	0.601	0.549	0.347	0.351
		40	KS	0.062	0.071	0.111	0.134	0.073	0.084	0.065	0.062
			CF_1	0.070	0.051	0.112	0.110	0.172	0.136	0.103	0.107
			CF_2	0.067	0.050	0.138	0.118	0.172	0.156	0.095	0.091
	$N = 10$	10	KS	0.050	0.065	0.108	0.129	0.266	0.279	0.121	0.132
			CF_1	0.059	0.043	0.124	0.133	0.718	0.680	0.499	0.434
			CF_2	0.069	0.045	0.145	0.124	0.735	0.706	0.459	0.431
		40	KS	0.048	0.056	0.117	0.138	0.148	0.178	0.070	0.086
			CF_1	0.058	0.059	0.126	0.123	0.453	0.369	0.307	0.226
			CF_2	0.067	0.054	0.136	0.123	0.507	0.401	0.306	0.236
$T = 50$	$N = 5$	25	KS	0.057	0.054	0.078	0.087	0.524	0.566	0.383	0.385
			CF_1	0.059	0.045	0.068	0.085	0.961	0.970	0.956	0.910
			CF_2	0.064	0.047	0.077	0.088	0.972	0.968	0.967	0.934
		100	KS	0.051	0.061	0.084	0.079	0.129	0.116	0.102	0.089
			CF_1	0.052	0.059	0.082	0.070	0.438	0.344	0.344	0.302
			CF_2	0.073	0.057	0.080	0.076	0.426	0.416	0.360	0.304
	$N = 10$	25	KS	0.049	0.052	0.071	0.083	0.735	0.760	0.552	0.608
			CF_1	0.059	0.048	0.075	0.079	0.995	0.995	0.993	0.979
			CF_2	0.062	0.052	0.090	0.086	0.997	0.995	0.992	0.987
		100	KS	0.046	0.056	0.086	0.080	0.444	0.397	0.264	0.335
			CF_1	0.051	0.047	0.079	0.074	0.940	0.924	0.886	0.880
			CF_2	0.062	0.053	0.100	0.082	0.953	0.926	0.886	0.889

TABLE 2. The standard deviation of random errors changes from $\sigma_0 = 1$ to $\sigma^0 = 2$, random errors with standard Normal distribution.

We omit results for a change in scale with innovation outliers that lead to similar conclusions as Table 2.

We omit also all results for additive outliers: In this situation, the impact of additive outliers is drastic on the empirical level of all tests for both LS and FLS–based statistics (approximately 0.11–0.16 for the LS and 0.11–0.19 for the FLS), even for the larger starting sample size $T = 50$.

Change in scale for leptokurtic and skew distributions. Table 5 contains simulation results for several distributions with various values of skewness and kurtosis and a change in scale alternative. The proposed test seems to work even for Cauchy distribution (t_1) although the achieved power is quite low—best results are obtained by using the CF_1 test statistic. The next set of simulation results for the random errors having standardized t_3 distribution confirms our findings observed in Table 2, i.e., that the ECF based tests have better power than the KS test against the "change in scale" alternatives. It should also be pointed out that the KS test achieves the best power against the same type of alternatives if the random errors

	t_0		emp. level LS CB	SB	FLS CB	SB	ach. power LS CB	SB	FLS CB	SB
χ^2_4	100	KS	0.051	0.051	0.065	0.065	0.502	0.504	0.459	0.463
		CF_1	0.043	0.041	0.065	0.064	0.908	0.925	0.866	0.878
		CF_2	0.048	0.052	0.048	0.059	0.835	0.837	0.828	0.814
	400	KS	0.047	0.060	0.048	0.056	0.161	0.152	0.148	0.149
		CF_1	0.051	0.053	0.064	0.051	0.338	0.351	0.325	0.369
		CF_2	0.057	0.051	0.064	0.063	0.252	0.293	0.255	0.247
t_5	100	KS	0.046	0.059	0.054	0.066	0.122	0.106	0.114	0.108
		CF_1	0.049	0.056	0.059	0.057	0.210	0.237	0.197	0.230
		CF_2	0.055	0.051	0.062	0.064	0.141	0.205	0.140	0.157
	400	KS	0.041	0.059	0.054	0.057	0.064	0.059	0.071	0.063
		CF_1	0.050	0.045	0.051	0.059	0.079	0.088	0.086	0.092
		CF_2	0.052	0.051	0.071	0.060	0.066	0.077	0.059	0.079
Lap.	100	KS	0.055	0.052	0.059	0.060	0.286	0.301	0.281	0.277
		CF_1	0.046	0.051	0.061	0.057	0.559	0.607	0.463	0.561
		CF_2	0.059	0.051	0.064	0.050	0.284	0.382	0.269	0.407
	400	KS	0.054	0.048	0.064	0.063	0.103	0.111	0.095	0.097
		CF_1	0.047	0.051	0.064	0.057	0.137	0.148	0.116	0.155
		CF_2	0.053	0.063	0.054	0.066	0.080	0.093	0.083	0.097

TABLE 3. Change from Normal to another distribution of random errors (the regression parameter and the standard deviation of random errors do not change), $T = 200$, $N = 5$.

have a standardized χ^2_1 distribution with the CF_2 test giving worst results for this type of random errors; cf. Figure 6. The last part of Table 5 contains simulations with random errors following a standardized χ^2_4 distribution. These simulation results again confirm the findings of Table 2, i.e., that the ECF based tests behave better against "change in scale" alternatives; see also Figure 6.

7. Conclusion

We have studied the finite–sample behavior of several tests for structural breaks in an AR series. This study included the Kolmogorov–Smirnov statistic as well as test statistics based on the ECF, with changes in the distribution of the random errors as well as breaks due to parameter change, both with and without outliers. The results of our simulation study suggest that:

- ECF based tests have similar power as KS test against "change in regression parameter" alternatives but ECF based tests are better in the same situation with innovation outliers.
- ECF based tests are superior to KS test against "change in scale" and "change in distribution" alternatives for symmetric or "less skew" distributions.
- Additive outliers are worse than innovation outliers.

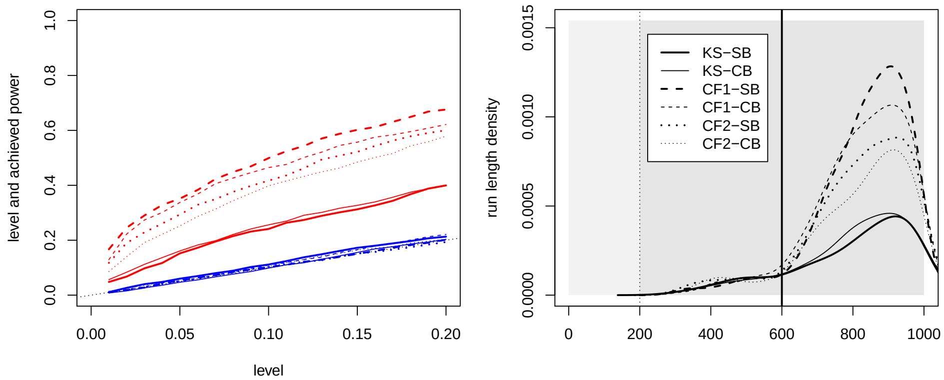

Figure 3. Normal distribution changes to standardized χ^2_4, $T = 200$, $N = 5$, $t_0 = 400$, LS estimator.

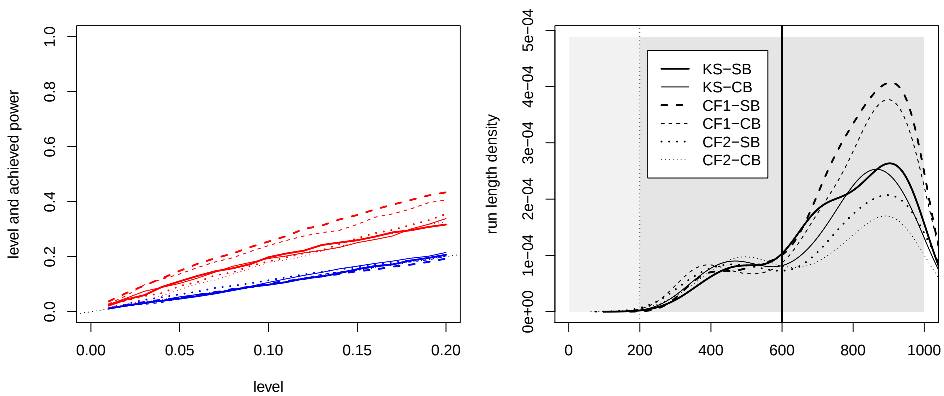

Figure 4. Normal distribution changes to double exponential (Laplace) distribution, $T = 200$, $N = 5$, $t_0 = 400$, LS estimator.

- Using FLS may be recommended only for larger starting sample sizes ($T \geq 50$) but the improvement in power is not large even when innovation outliers are present.
- Sequential bootstrap may help to control the level of the test in a situation with larger starting sample size ($T \geq 200$).

It seems that the CF_1 test using LS estimator and classical bootstrap may be recommended in most situations.

Acknowlegements

The work of Z. Hlávka was partially supported by GAČR P201/12/0083, the work of M. Hušková was partially supported by the grant GAČR 201/09/0755, and the work of S. Meintanis by grant number KA: 70/4/7658 of the program 'Kapodistrias' of the special account for research of the National and Kapodistrian

		t_0		emp. level LS CB	emp. level LS SB	emp. level FLS CB	emp. level FLS SB	ach. power LS CB	ach. power LS SB	ach. power FLS CB	ach. power FLS SB
$T=20$	$N=5$	10	KS	0.069	0.094	0.103	0.126	0.445	0.407	0.397	0.386
			CF_1	0.076	0.063	0.090	0.102	0.463	0.520	0.464	0.452
			CF_2	0.058	0.072	0.089	0.114	0.483	0.473	0.452	0.461
		40	KS	0.074	0.083	0.112	0.123	0.187	0.199	0.144	0.155
			CF_1	0.058	0.082	0.097	0.120	0.224	0.187	0.161	0.161
			CF_2	0.058	0.095	0.091	0.122	0.227	0.167	0.177	0.147
	$N=10$	10	KS	0.065	0.083	0.102	0.126	0.475	0.480	0.395	0.412
			CF_1	0.067	0.112	0.103	0.120	0.524	0.499	0.457	0.485
			CF_2	0.070	0.097	0.098	0.124	0.529	0.506	0.476	0.458
		40	KS	0.066	0.085	0.096	0.112	0.280	0.259	0.261	0.264
			CF_1	0.061	0.089	0.092	0.132	0.361	0.337	0.341	0.293
			CF_2	0.062	0.102	0.097	0.136	0.360	0.290	0.321	0.283
$T=50$	$N=5$	25	KS	0.055	0.082	0.053	0.077	0.642	0.596	0.684	0.652
			CF_1	0.059	0.086	0.067	0.073	0.781	0.764	0.818	0.810
			CF_2	0.053	0.084	0.065	0.073	0.754	0.727	0.743	0.762
		100	KS	0.066	0.062	0.058	0.082	0.259	0.238	0.259	0.224
			CF_1	0.062	0.076	0.065	0.072	0.403	0.326	0.395	0.345
			CF_2	0.064	0.078	0.060	0.075	0.377	0.336	0.433	0.335
	$N=10$	25	KS	0.057	0.061	0.079	0.071	0.699	0.695	0.675	0.723
			CF_1	0.056	0.086	0.063	0.084	0.823	0.824	0.874	0.895
			CF_2	0.051	0.079	0.059	0.084	0.806	0.793	0.844	0.854
		100	KS	0.067	0.064	0.069	0.078	0.423	0.451	0.448	0.449
			CF_1	0.061	0.083	0.070	0.083	0.665	0.625	0.720	0.669
			CF_2	0.055	0.092	0.057	0.082	0.635	0.568	0.674	0.641

Table 4. The regression parameter changes from $\beta_0 = 0.4$ to $\beta^0 = 0.8$, random errors have standard Normal distribution with 10% contamination with innovation outliers ($N(0, 100)$).

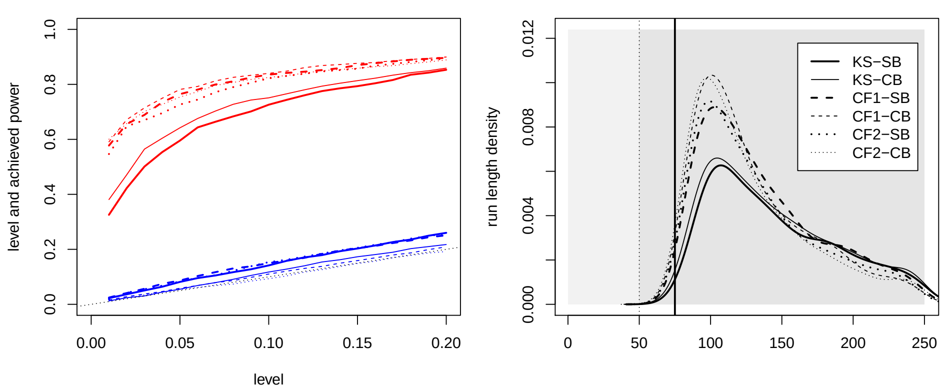

Figure 5. Change in the regression parameter with innovation outliers, $T = 50$, $N = 5$, $t_0 = 25$, LS estimator.

				emp. level				ach. power			
				LS		FLS		LS		FLS	
		t_0		CB	SB	CB	SB	CB	SB	CB	SB
t_1	$T=50$	50	KS	0.081	0.090	0.081	0.081	0.080	0.080	0.073	0.104
			CF_1	0.091	0.092	0.088	0.102	0.149	0.103	0.147	0.100
			CF_2	0.084	0.089	0.077	0.104	0.145	0.104	0.149	0.081
	$T=100$	50	KS	0.074	0.075	0.076	0.067	0.146	0.136	0.153	0.198
			CF_1	0.075	0.083	0.062	0.080	0.274	0.194	0.300	0.173
			CF_2	0.079	0.084	0.054	0.088	0.237	0.166	0.267	0.149
		200	KS	0.072	0.068	0.063	0.073	0.071	0.070	0.073	0.073
			CF_1	0.069	0.095	0.061	0.081	0.144	0.086	0.138	0.095
			CF_2	0.070	0.085	0.065	0.087	0.123	0.088	0.111	0.071
t_3	$T=50$	50	KS	0.073	0.064	0.077	0.074	0.094	0.099	0.076	0.101
			CF_1	0.068	0.056	0.081	0.075	0.276	0.224	0.238	0.196
			CF_2	0.070	0.051	0.080	0.097	0.290	0.247	0.284	0.176
	$T=100$	50	KS	0.054	0.064	0.069	0.063	0.267	0.263	0.228	0.249
			CF_1	0.056	0.060	0.061	0.068	0.689	0.625	0.700	0.605
			CF_2	0.052	0.068	0.070	0.059	0.715	0.571	0.697	0.609
		200	KS	0.050	0.057	0.056	0.064	0.097	0.092	0.093	0.093
			CF_1	0.054	0.051	0.069	0.059	0.248	0.210	0.202	0.183
			CF_2	0.051	0.048	0.072	0.074	0.261	0.210	0.233	0.177
χ_1^2	$T=50$	50	KS	0.060	0.049	0.074	0.071	0.927	0.947	0.944	0.955
			CF_1	0.059	0.062	0.077	0.076	0.746	0.671	0.669	0.645
			CF_2	0.070	0.066	0.072	0.059	0.511	0.472	0.483	0.500
	$T=100$	50	KS	0.040	0.066	0.056	0.074	1.000	1.000	1.000	1.000
			CF_1	0.058	0.051	0.060	0.062	0.997	0.998	0.998	0.997
			CF_2	0.059	0.062	0.061	0.059	0.969	0.946	0.960	0.958
		200	KS	0.065	0.064	0.055	0.070	0.950	0.947	0.980	0.970
			CF_1	0.059	0.059	0.061	0.051	0.651	0.635	0.651	0.682
			CF_2	0.068	0.056	0.060	0.064	0.435	0.443	0.424	0.378
χ_4^2	$T=50$	50	KS	0.049	0.057	0.086	0.092	0.201	0.168	0.139	0.114
			CF_1	0.059	0.056	0.082	0.082	0.479	0.457	0.448	0.418
			CF_2	0.056	0.062	0.076	0.085	0.501	0.388	0.413	0.346
	$T=100$	50	KS	0.047	0.064	0.065	0.067	0.832	0.802	0.759	0.787
			CF_1	0.059	0.058	0.064	0.058	0.968	0.966	0.966	0.959
			CF_2	0.051	0.057	0.079	0.053	0.957	0.928	0.931	0.931
		200	KS	0.046	0.053	0.059	0.068	0.182	0.159	0.126	0.120
			CF_1	0.048	0.053	0.069	0.066	0.463	0.387	0.366	0.352
			CF_2	0.057	0.060	0.062	0.064	0.385	0.344	0.416	0.344

TABLE 5. The standard deviation of random errors changes from $\sigma_0 = 1$ to $\sigma^0 = 1.5$ for various distributions of random errors, $N = 5$.

University of Athens. The position of C. Kirch was financed by the Stifterverband für die Deutsche Wissenschaft by funds of the Claussen-Simon-trust and the research supported by the DFG grant KI 1443/2-2.

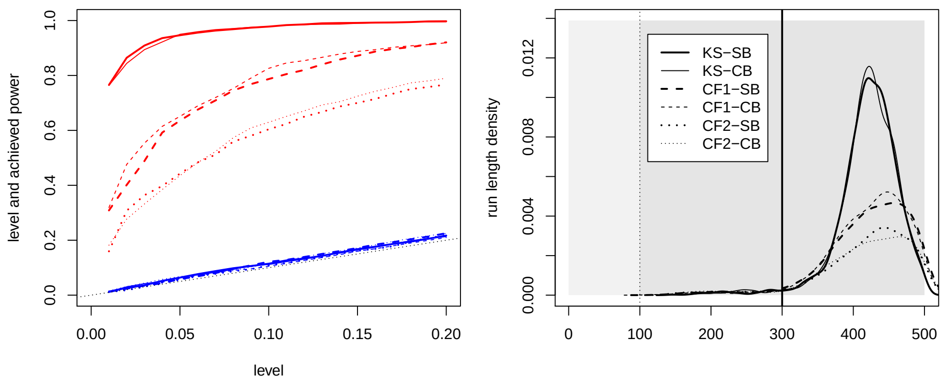

FIGURE 6. Change in scale with standardized χ_1^2 distribution, $T = 100$, $N = 5$, $t_0 = 200$, LS estimator.

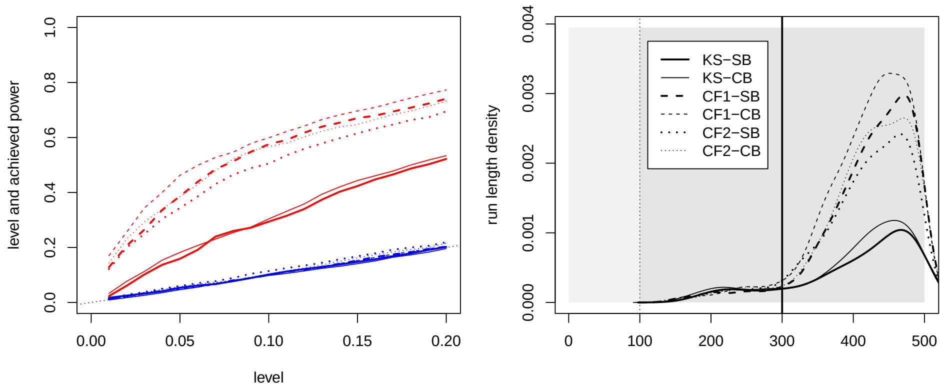

FIGURE 7. Change in scale with standardized χ_4^2 distribution, $T = 100$, $N = 5$, $t_0 = 200$, LS estimator.

REFERENCES

Cai, Y. and Davies, N. (2003). Monitoring the parameter changes in general ARIMA time series models, *J. Appl. Stat.* **30**(9): 983–1001.

Carsoule, F. and Franses, P. H. (2003). A note on monitoring time-varying parameters in an autoregression, *Metrika* **57**(1): 51–62.

Csörgő, S. (1983). The theory of functional least squares, *J. Austral. Math. Soc. Ser. A* **34**(3): 336–355.

Dhar, S. K. (1993). Computation of certain minimum distance estimators in AR[k] model, *J. Amer. Statist. Assoc.* **88**(421): 278–283.

Gombay, E. and Serban, D. (2009). Monitoring parameter change in AR(p) time series models, *J. Multivariate Anal.* **100**(4): 715–725.

Heathcote, C. R. and Welsh, A. H. (1983). The robust estimation of autoregressive processes by functional least squares, *J. Appl. Probab.* **20**(4): 737–753.

Hlávka, Z., Hušková, M., Kirch, C. and Meintanis, S. G. (2012). Monitoring changes in the error distribution of autoregressive models based on Fourier methods, *TEST* . To appear.

Horváth, L., Kühn, M. and Steinebach, J. (2008). On the performance of the fluctuation test for structural change, *Sequential Anal.* **27**(2): 126–140.

Hušková, M., Kirch, C., Prášková, Z. and Steinebach, J. (2008). On the detection of changes in autoregressive time series. II. Resampling procedures, *J. Statist. Plann. Inference* **138**(6): 1697–1721.

Hušková, M., Prášková, Z. and Steinebach, J. (2007). On the detection of changes in autoregressive time series. I. Asymptotics, *J. Statist. Plann. Inference* **137**(4): 1243–1259.

Hušková, M. and Kirch, C. (2011). Bootstrapping sequential change-point tests for linear regression, *Metrika* . To appear.

Kirch, C. (2008). Bootstrapping sequential change-point tests, *Sequential Anal.* **27**(3): 330–349.

Koller, M. and Stahel, W. A. (2011). Sharpening Wald-type inference in robust regression for small samples, *Comput. Statist. Data Anal.* **55**(8): 2504–2515.

Lee, S., Ha, J., Na, O. and Na, S. (2003). The cusum test for parameter change in time series models, *Scand. J. Statist.* **30**(4): 781–796.

Lee, S., Lee, Y. and Na, O. (2009). Monitoring distributional changes in autoregressive models, *Comm. Statist. Theory Methods* **38**(16-17): 2969–2982.

Meintanis, S. G. and Donatos, G. S. (1999). Finite-sample performance of alternative estimators for autoregressive models in the presence of outliers., *Comput. Stat. Data Anal.* **31**(3): 323–339.

Na, O., Lee, J. and Lee, S. (2010). Monitoring parameter changes for random coefficient autoregressive models, *J. Korean Statist. Soc.* **39**(3): 281–288.

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

Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M. and Maechler, M. (2011). *robustbase: Basic Robust Statistics*. R package version 0.8-0, http://CRAN.R-project.org/package=robustbase.

Welsh, A. H. (1985). An angular approach for linear data, *Biometrika* **72**(2): 441–450.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression, *Ann. Statist.* **15**(2): 642–656.