Testing for Structural Change of AR Model to Threshold AR Models

István Berkes[*], Lajos Horváth[†], Shiqing Ling[‡] and Johannes Schauer[*]

[*]Institute of Statistics, Graz University of Technology, Münzgrabenstraße 11, A-8010 Graz, Austria

[†]Department of Mathematics, University of Utah, 155 South 1440 East, Salt Lake City, UT 84112-0090 USA

[‡]Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong

Abstract

The purpose of this paper is to develop the likelihood ratio test for the structural change of an AR model to a threshold AR model. It is shown that the log-likelihood ratio test converges to the maxima of a two-parameter Gaussian process in distribution. This limiting distribution is novel and we tabulate the critical values. Some simulations are carried out to examine the finite-sample performance of this test statistic. This paper also includes a weak convergence of a two-parameter marked empirical process, which is of independent interest.

Keywords: AR model, threshold AR model, likelihood ratio test, structural change, marked empirical process

1 Introduction

Many events, like the great depression, oil price shocks, and abrupt policy changes, may cause a change in the structure of time series models used in economics and finance. When the model changes, the original model may perform poorly, either for forecasting purposes or for the purpose of analyzing the effect of policy changes. How to quickly and accurately detect such a change has been a challenging issue to econometricians and statisticians for a long time. The earliest references go back to Chow (1960) and Quandt (1960). For a survey on the history and early results, we refer to Csörgő and Horváth (1997). Davis, Huang and Yao (1995) considered the problem of testing whether or not a change has occurred in the parameter values of AR models. Bai, Lumsdaine, and Stock (1998) studied the change point tests for $I(0)$ and $I(1)$ multivariate time series. Bai and Perron (1998) considered issues related to multiple structural changes, occurring at unknown dates, in the linear regression model. They proposed a procedure which tests the null hypothesis of no change against the alternative of at least one change. Bai (1999) proposed a likelihood-ratio-type test for multiple structural changes in regression models. His model allows for lagged-dependent variables and trending regressors. He showed that asymptotic critical values can be obtained analytically. See also Hansen (2003), Yao and Davis (1986), Horváth (1993, 1995). Ling (2007) developed a general asymptotic theory on the Wald test for change-points in a general class of time series models under no change-point hypothesis and applied the results for the long-memory fractional ARIMA model.

All the test statistics in the previous papers are for the change of parameters in the models, while the form of the model is not changed under both null and alternative hypotheses. However, the form of the model may be changed in practice. In fact, it is not unusual to see that the forms of fitted models to different periods of a time series are different. For example, Li and Lam (1995) used the threshold AR-ARCH models to fit 11 non-overlapping two-year period Hong Kong Hang Seng index from 1970 to 1991, but Wong and Li (1997) found that two out of 11 periods should follow the AR-ARCH models. When the form of models is changed, the likelihood ratio test (LRT) for the standard change-point problem is not locally most powerful any more. How to efficiently detect the change of the structural forms of the time series models is a new issue. This paper is the first step for this issue. We study the LRT for testing the structural change of an AR model to a threshold AR model. The test statistic is shown to converge the maxima of a two-parameter Gaussian process in distribution. This limiting distribution is novel and never appears in the literature before. The critical values of the test statistic are tabulated. Some simulations are carried out to examine the finite-sample performance of this test statistic. This paper also includes a weak convergence of a two-parameter marked empirical process, which is of independent interest.

This paper is organized as follows. Section 2 presents the test statistic and its limiting distribution. Section 3 reports simulation results and gives one real data analysis. Section 4 gives the proof of our main results. Section 5 studies the weak convergence of a two-parameter marked empirical process.

2 Likelihood Ratio Test

We consider the null hypothesis

$$H_0: \quad y_i = \rho y_{i-1} + \varepsilon_i, \quad \text{if} \quad i = 1, \ldots, n$$

against the alternative

$$H_1: \quad \text{there is an integer} \;\; 1 \le k^* < n \quad \text{such that}$$

$$y_i = \begin{cases} \phi_1 y_{i-1} + \varepsilon_i, & \text{if} \;\; i = 1, \ldots, k^*, \\ \\ \phi_1 y_{i-1} + \phi_2 y_{i-1} I\{y_{i-1} \le r\} + \varepsilon_i, & \text{if} \;\; i = k^* + 1, \ldots, n, \end{cases}$$

where $|\rho| < 1$, r is called the threshold parameter and the errors, ε_i, are independent identically distributed random variables with mean 0 and variance $\sigma^2 > 0$. Under H_0, the quasi-likelihood function can be written as

$$L_n(\rho, \sigma^2) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \rho y_{i-1})^2 \right\}.$$

Under H_1, assuming that $k = k^*$, the time of change, and r, the threshold, are both known, the quasi-likelihood function can be written as

$$L_{1n}(\phi_1, \phi_2, \sigma^2)$$
$$= \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left\{ -\frac{1}{2\sigma^2} \left[\sum_{i=1}^{k} (y_i - \phi_1 y_{i-1})^2 + \sum_{i=k+1}^{n} (y_i - \phi_1 y_{i-1} - \phi_2 y_{i-1} I\{y_{i-1} \le r\})^2 \right] \right\}.$$

We should reject H_0 in favor of H_1 if the likelihood ratio

$$\sup_{\rho,\sigma} L_n(\rho, \sigma^2) \Big/ \sup_{\phi_1,\phi_2,\sigma} L_{1n}(\phi_1, \phi_2, \sigma^2)$$

is small. It is well known that

$$\sup_{\rho,\sigma} L_n(\rho, \sigma^2) = L_n(\hat{\rho}_n, \hat{\sigma}_n^2) = \left(\frac{1}{2\pi\hat{\sigma}_n^2} \right)^{n/2} \exp(-n/2),$$

where

$$\hat{\rho}_n = \left(\sum_{i=1}^{n} y_{i-1}^2\right)^{-1} \left(\sum_{i=1}^{n} y_{i-1} y_i\right)$$

and

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\rho}_n y_{i-1})^2$$

(cf. Brockwell and Davis (1991)). Since

$$\begin{aligned} &\sum_{i=1}^{k} (y_i - \phi_1 y_{i-1})^2 + \sum_{i=k+1}^{n} (y_i - \phi_1 y_{i-1} - \phi_2 y_{i-1} I\{y_{i-1} \le r\})^2 \\ &\quad = \sum_{i=1}^{n} \left(y_i - \phi_1 y_{i-1} - \phi_2 y_{i-1} I\{y_{i-1} \le r, i > k\}\right)^2, \end{aligned}$$

we get that for any fixed k and r the quasi-maximum likelihood estimators $\hat{\phi}_{1n}(k,r)$, $\hat{\phi}_{2n}(k,r)$, $\hat{\sigma}_n^2(k,r)$ are given by

$$\left[\, \hat{\phi}_{1n}(k,r),\ \hat{\phi}_{2n}(k,r) \,\right]' = \left(\sum_{i=1}^{n} A_{i-1} A_{i-1}'\right)^{-1} \left(\sum_{i=1}^{n} A_{i-1} y_i\right),$$

$$\hat{\sigma}_n^2(k,r) = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{\phi}_{1n}(k,r) y_{i-1} - \hat{\phi}_{2n}(k,r) y_{i-1} I\{y_{i-1} \le r, i > k\}\right)^2,$$

where $A_{i-1} = (1,\ I\{y_{i-1} \le r, i > k\})' y_{i-1}$ (x' denotes the transpose of vectors and matrices). Thus we get

$$\sup_{\phi_1, \phi_2, \sigma} L_{1n}(\phi_1, \phi_2, \sigma^2) = L_{1n}(\hat{\phi}_{1n}(k,r), \hat{\phi}_{2n}(k,r), \hat{\sigma}_n^2(k,r)) = \left(\frac{1}{2\pi \hat{\sigma}_n^2(k,r)}\right)^{n/2} \exp(-n/2).$$

Hence, for any fixed k and r, -2 times the log-likelihood ratio is

$$T_n(k,r) = n \left(\log \hat{\sigma}_n^2 - \log \hat{\sigma}_n^2(k,r)\right).$$

Next we consider the asymptotic behavior of $T_n([nt], r)$, $0 \le t \le 1, -\infty < r < \infty$, where $[nt]$ denotes the integer part of nt. Before we state our result, we need some further notations. Let $W(t,u), 0 \le t, u < \infty$ denote a two-parameter Wiener process, i.e. W is a continuous Gaussian process with $EW(t,u) = 0,\ E(W(t,u)W(s,v)) = \min(t,s)\min(u,v)$. For the existence and basic properties of the two-parameter Wiener process we refer to Csörgő and Révész (1981). Let $H(r) = E(y_0^2 I\{y_0 \le r\})$ and $H = \lim_{r\to\infty} H(r) = Ey_0^2$. Now we define the following Gaussian process

$$U(t,x) = x(1-t)W(1,H) - [W(1,x) - W(t,x)]H.$$

Theorem 2.1. *We assume that*

$$\{\varepsilon_i\} \text{ are independent, identically distributed random variables,} \tag{2.1}$$

$$E\varepsilon_0 = 0, E\varepsilon_0^2 = \sigma^2 > 0, E|\varepsilon_0|^{6+\delta} < \infty \text{ with some } \delta > 0$$

and

$$|P(a \le y_0 \le b)| \le C|a-b|^\alpha \quad \text{with some } C \text{ and } 0 < \alpha \le 1. \tag{2.2}$$

If H_0 holds, then

$$T_n([nt], r) \Longrightarrow \frac{(U(t,H(r)))^2}{(1-t)H(r)(H-(1-t)H(r))H} \quad \text{in } \mathcal{D}([t_1,t_2]\times[r_1,r_2]), \tag{2.3}$$

for any $0 < t_1 < t_2 < 1$ and $r_1 < r_2$ satisfying $0 < H(r_1) < H(r_2) < H$.

It is easy to see that the distribution of the limit on the right hand side of (2.3) is $\chi^2(1)$ for any fixed t and r. This is natural, since $T_n([nt], r)$ is -2 times the log-likelihood ratio test for any fixed t and r. The time of change and the threshold are unknown so it is natural to consider the supremum functional of $T_n([nt], r)$.

Corollary 2.1. *If the conditions of Theorem 2.1 are satisfied, then*

$$\max_{nt_1 \le k \le nt_2} \sup_{r_1 \le r \le r_2} T_n(k,r) \xrightarrow{\mathcal{D}} \sup_{1-t_2 \le s \le 1-t_1} \sup_{u_1 \le u \le u_2} \frac{(suW(1,1) - W(s,u))^2}{su(1-su)}. \tag{2.4}$$

where $u_1 = \bar{H}(r_1)$ and $u_2 = \bar{H}(r_2)$ with $\bar{H}(x) = H(x)/H$.

Corollary 2.1 leads to distribution free tests. First we need an estimator for $\bar{H}(x)$. Let

$$\bar{H}_n(x) = \frac{\sum_{t=1}^{n} y_{t-1}^2 I\{y_{t-1} \le x\}}{\sum_{t=1}^{n} y_{t-1}^2}.$$

Clearly, $\bar{H}(x)$ and the empirical counterpart $\bar{H}_n(x)$ are distribution functions so by $\bar{H}_n^{-1}(x)$ we denote the empirical quantile function, the generalized inverse of $\bar{H}_n(x)$.

Corollary 2.2. *If the conditions of Theorem 2.1 are satisfied and the distribution function of y_0 is strictly increasing on its support, then for all $0 < t_1 < t_2 < 1$ and $0 < u_1 < u_2 < 1$ we have*

$$\max_{nt_1 \le k \le nt_2} \sup_{\hat{r}_1 \le r \le \hat{r}_2} T_n(k,r) \xrightarrow{\mathcal{D}} \sup_{1-t_2 \le s \le 1-t_1} \sup_{u_1 \le u \le u_2} \frac{(suW(1,1) - W(s,u))^2}{su(1-su)},$$

where $\hat{r}_1 = \bar{H}_n^{-1}(u_1)$ and $\hat{r}_2 = \bar{H}_n^{-1}(u_2)$ with $\bar{H}(x) = H(x)/H$.

It is natural to ask if Corollaries 2.1 and 2.2 remain true if the suprema are taken for all possible values of the arguments. The answer is no, since according to the law of iterated logarithm for W (cf. Csörgő and Révész (1981)) we have that

$$P\left\{\sup_{0<s<1} \sup_{0<u<1} \frac{(suW(1,1) - W(s,u))^2}{su(1-su)} = \infty\right\} = 1. \tag{2.5}$$

So truncation is needed or we have to work with weighted statistics. Since one term Taylor expansion yields that, under H_0, $T_n(k,r)$ is equivalent with $n(\hat{\sigma}_n^2(k,r) - \hat{\sigma}_n^2)/\hat{\sigma}_n^2$, we consider the weighted version of $n(\hat{\sigma}_n^2(k,r) - \hat{\sigma}_n^2)/\hat{\sigma}_n^2$:

$$R_n(k,r) = \frac{1}{n^2 \hat{\sigma}_n^2}(Z_n(Z_n S_n(k,r) - S_n^2(k,r)))(\hat{\sigma}_n^2 - \hat{\sigma}_n^2(k,r)),$$

where

$$Z_n = \sum_{i=1}^{n} y_{i-1}^2 \quad \text{and} \quad S_n(k,r) = \sum_{i=k+1}^{n} y_{i-1}^2 I\{y_{i-1} \le r\}.$$

Theorem 2.2. *If H_0, (2.1) and (2.2) hold, then*

$$R_n(nt,r) \Longrightarrow (U(t,H(r)))^2 \quad \text{in} \quad \mathcal{D}([0,1] \times [-\infty,\infty]).$$

Theorem 2.2 provides a statistic which is distribution free under the null hypothesis.

Corollary 2.3. *If the conditions of Theorem 2.2 are satisfied, then*

$$\bar{R}_n = \left(\frac{n}{Z_n}\right)^3 \max_{1\le k\le n} \sup_{-\infty<r<\infty} R_n(k,r) \xrightarrow{\mathcal{D}} \sup_{0\le s\le 1} \sup_{0\le u\le 1} (suW(1,1) - W(s,u))^2.$$

The two-parameter process $W(t,s) - stW(1,1)$ is similar to $W^*(u) - uW^*(1)$, the representation of the Brownian bridge in terms of W^*, a Wiener process (standard Brownian motion). If $t = 1$, then $W(1,s) - sW(1,1)$ is a Brownian bridge. Choosing $k = 1$ in $R_n(k,r)$, we are testing the AR(1) null hypothesis against a threshold AR(1) alternative. This problem was studied by Chan (1990) and Chan and Tong (1990) and $\max_r R_n(1,r)$ is the weighted version of the likelihood ratio test in their paper. More references on testing the threshold models can be found in Wong and Li (1997, 2000) and Ling and Tong (2005). If $s = 1$, then $W(t,1) - tW(1,1)$ is again a Brownian bridge. Rejecting for large values of $\max_k R_n(k,\infty)$ we are rejecting the AR(1) model null hypothesis in favor of the change in the parameter of the autoregressive process alternative. Our weighted test is designed to detect if the model changed into a threshold autoregression at an unknown time.

3 Simulations and an Application

In this section we study the finite sample behavior of the asymptotically distribution free test statistic $\bar{R}_n$ of Corollary 2.3. Corollary 2.3 provides the asymptotic distribution of the test statistic under H_0. Under H_0 the value of ρ is arbitrary with $|\rho| < 1$ and the errors ε_i are i.i.d. random variables with finite variance. We restrict ourselves to the case of normally distributed errors with mean 0 and variance 1 in our study. We will consider five different cases, $\rho \in \{-0.5, -0.25, 0, 0.25, 0.5\}$. We also want to check the test statistic's performance under different alternative hypotheses. This leads us to the values of $\phi_2 \in \{-0.75 - \phi_1, -0.5 - \phi_1, \ldots, 0.75 - \phi_1\} \cup \{-0.2, -0.15, \ldots, 0.2\}$ and $\phi_1 \in \{-0.5, -0.25, 0, 0.25, 0.5\}$. We note that each case with

$\phi_2 = 0$ corresponds to the null hypothesis case.

The threshold r can be arbitrary too. However for our simulations we fix its value at 0. In many practical situations this is a natural choice. Obviously the time of change k^* will influence the results, therefore we study the three cases $k^* \in \{n/4, n/2, 3n/4\}$ for a given sample size n. We will use $n \in \{200, 400\}$. Our main interest lies in empirical power curves for each setting of n, k^* and ϕ_1, showing us proportions of rejection of the null hypothesis under different alternative hypotheses (depending on the value of ϕ_2) for a fixed significance level $1 - \alpha$.

The asymptotic distribution of our test statistic is not known in a closed form. Thus we can only use Monte Carlo simulations to get the asymptotic critical values. We approximate the two-parameter Wiener process on a 500 by 500 grid with partial sums of i.i.d. normal variables. Repeating the procedure $N = 10000$ times we found the empirical quantiles as shown in Table 1.

Using the values in Table 1 we can construct the above described empirical power curves as shown in Figure 1. For each combination of k^* and ϕ_1 we show the curves for both values of $n \in \{200, 400\}$ corresponding to $\alpha = 0.05$.

We note that in all 15 situations the observed significance level of the proposed test under the null hypothesis is close to $1-\alpha = 0.95$ for sample sizes $n = 400$, while in the case of $n = 200$ it is slightly larger. The differences in the plots concerning the position of the change (for fixed ϕ_1) reveal an interesting property. Namely, the test generally seems to work better for early changes, in our test scenario $k^* = n/4$. The deterioration of the power is especially obvious in the late change $k^* = 3n/4$ when $\phi_1 > 0$. The increase in the power curves is slow when $\phi_1 > 0$, $k^* = 3n/4$ and ϕ_2 tends to -1. This behavior might be explained by the fact that the alternative hypothesis only really shows a change if the values fall below $r = 0$. Thus the test statistic seems to need considerably more time after the change to detect H_1.

The influence of the value of ϕ_1 on the power is quite small. While the behavior of the test statistic with negative ϕ_1 seems to be slightly better, the differences are not essential. We observe that for negative ϕ_1 the power curves corresponding to change-points $k^* = n/4$ and $k^* = n/2$ are nearly the same.

Now we study two data sets. Balcombe et al. (2007) models prices of agricultural products by threshold AR(1) models. Our first sample consists of monthly average corn prices and the second sample consists of monthly average soybeans prices achieved by farmers in Illinois from January 1960 until November 2008[1]. The prices are given in dollars per bushel. The sample size is 587. Figure 2 shows the two time series.

We are interested in the relative prices changes, thus for given prices y_j we study x_j defined by

$$x_j = \frac{y_j - y_{j-1}}{y_{j-1}} \approx \log\left(\frac{y_j}{y_{j-1}}\right).$$

To check if the data shows an indication of changing from an AR(1) to a threshold AR(1) model we use $\bar{R}_n$ of Corollary 2.3. The values of the test statistic $\bar{R}_n$ are 5.039 (corn) and 5.861 (soybeans), so using the critical values in Table 1 we reject the null hypothesis in both cases at the level $\alpha = 0.01$.

The location of the maximum of $\sup_r R_n(k, r)$ indicates that the changes occurred around July 1971 (corn) and October 1974 (soybeans). Our findings are summarized in Table 2.

Figure 3 shows the two time series of relative price changes together with the estimated change-points and thresholds. Note that for the soybeans data the model suggests that the observations x_j are nearly independent if $j \leq k^*$. To check the performance of our model we plot the residuals in Figure 4. The plots indicate that our model describes the data sets. Using the Ljung-Box test on the residuals we could not reject that the residuals are essentially independent.

[1] The data sets are available at `http://www.farmdoc.uiuc.edu/manage/pricehistory/price_history.html`.

4 Proofs

The first lemma provides a simple representation for the difference between the estimated variances under the null and the change in the structure at time k alternative.

Lemma 4.1. *For any k and r we have*

$$n\left(\hat{\sigma}_n^2 - \hat{\sigma}_n^2(k,r)\right) \tag{4.1}$$

$$= \left(\sum_{i=1}^n A_{i-1}\varepsilon_i\right)' \left(\sum_{i=1}^n A_{i-1}A_{i-1}'\right)^{-1} \left(\sum_{i=1}^n A_{i-1}\varepsilon_i\right) - \left(\sum_{i=1}^n y_{i-1}^2\right)^{-1} \left(\sum_{i=1}^n y_{i-1}\varepsilon_i\right)^2$$

and

$$n\left(\hat{\sigma}_n^2 - \hat{\sigma}_n^2(k,r)\right) = \frac{\left(S_n(k,r)\sum_{i=1}^n y_{i-1}\varepsilon_i - Z_n \sum_{i=k+1}^n y_{i-1}\varepsilon_i I\{y_{i-1} \le r\}\right)^2}{Z_n(Z_n S_n(k,r) - S_n^2(k,r))}. \tag{4.2}$$

Proof: The proof of (4.1) can be found, for example, in Seber and Lee (2003). Using the definition of A_{t-1}, (4.2) follows from (4.1) via long but elementary calculations.

Lemma 4.2. *If the conditions of Theorem 2.1 are satisfied, then*

$$\frac{1}{n} Z_n \to E y_0^2 \quad a.s. \tag{4.3}$$

and

$$\max_{1\le k\le n} \sup_{-\infty<r<\infty} \left| \frac{1}{n} S_n(k,r) - \left(1 - \frac{k}{n}\right) E(y_0^2 I\{y_0 \le r\}) \right| \xrightarrow{P} 0. \tag{4.4}$$

Proof: The ergodic theorem implies (4.3). Next we show that

$$\max_{1\le k\le n-n\delta} \sup_{-\infty<r<\infty} \left| \frac{1}{n} S_n(k,r) - \left(1 - \frac{k}{n}\right) E(y_0^2 I\{y_0 \le r\}) \right| \to 0 \text{ a.s.} \tag{4.5}$$

for all $0 < \delta < 1$. Let $T > 0$. Clearly,

$$\sup_{T\le r<\infty} |E y_0^2 I\{y_0 \le r\} - E y_0^2| = E y_0^2 I\{y_0 > T\}$$

and

$$\sup_{-\infty<r\le -T} Ey_0^2 I\{y_0 \le r\} = Ey_0^2 I\{y_0 \le -T\}$$

hold. Similarly we find

$$\sup_{T\le r<\infty} \sum_{i=k+1}^{n} y_0^2 I\{y_0 > r\} = \sum_{i=k+1}^{n} y_0^2 I\{y_0 > T\}$$

and

$$\sup_{-\infty<r\le -T} \sum_{i=k+1}^{n} y_0^2 I\{y_0 \le r\} = \sum_{i=k+1}^{n} y_0^2 I\{y_0 \le -T\}.$$

Hence by the ergodic theorem

$$\limsup_{n\to\infty} \max_{1\le k\le n-n\delta} \sup_{-\infty<r\le -T} \left| \frac{1}{n} S_n(k,r) - \left(1-\frac{k}{n}\right) E(y_0^2 I\{y_0 \le r\}) \right|$$
$$\le 2E(y_0^2 I\{y_0 \le -T\})$$

and

$$\limsup_{n\to\infty} \max_{1\le k\le n-n\delta} \sup_{T\le r<\infty} \left| \frac{1}{n} S_n(k,r) - \left(1-\frac{k}{n}\right) E(y_0^2 I\{y_0 \le r\}) \right|$$
$$\le 2E(y_0^2 I\{y_0 \ge T\}).$$

By (2.2) we have that $Ey_0^2 I\{y_0 \le r\}$ is a continuous function, so the ergodic theorem with Kaczor and Nowak (2001, p. 85) yields for all $T>0$ that

$$\max_{1\le k\le n-n\delta} \sup_{-T\le r\le T} \left| \frac{1}{n} S_n(k,r) - \left(1-\frac{k}{n}\right) E(y_0^2 I\{y_0 \le r\}) \right| \to 0 \quad \text{a.s.}$$

Since we can choose T as large as we wish, (4.5) is proven.

Due to stationarity and the ergodic theorem we have

$$\frac{1}{n} \max_{n-n\delta\le k\le n} S_n(k,r) \le \frac{1}{n} \sum_{i=n-[n\delta]}^{n} y_{i-1}^2 \stackrel{\mathcal{D}}{=} \frac{1}{n} \sum_{i=1}^{[n\delta]+1} y_{i-1}^2 \to \delta E y_0^2 \quad \text{a.s.}$$

Since $\delta>0$ can be chosen as small as possible, (4.4) is proven.

Lemma 4.3. *If the conditions of Theorem 2.1 are statisfied, then*

$$\frac{1}{n^{3/2}}\left(S_n([nt],r)\sum_{i=1}^{n} y_{i-1}\varepsilon_i - Z_n \sum_{i=[nt]+1}^{n} y_{i-1}\varepsilon_i I\{y_{i-1}\le r\}\right) \Longrightarrow \sigma U(t,H(r)) \tag{4.6}$$

in $\mathcal{D}([0,1]\times[-\infty,\infty])$.

Proof: The result follows immediately from Theorem 5.1 (proved in Section 5) and Lemma 4.2.

Lemma 4.4. *If the conditions of Theorem 2.1 are satisfied, then*

$$\hat{\sigma}_n^2 \quad \xrightarrow{P} \quad \sigma^2 \tag{4.7}$$

and

$$\max_{n\delta\le k\le n-n\delta}\max_{r_1\le r\le r_2}|\hat{\sigma}_n^2(k,r)-\sigma^2| \quad \xrightarrow{P} \quad 0 \tag{4.8}$$

for all $0<\delta<1$ *and* $r_1<r_2$ *satisfying* $0<H(r_1)\le H(r_2)<1$.

Proof: The proof of (4.7) can be found, for example, in Brockwell and Davis (1991). By (4.2) and Lemmas 4.2 and 4.3 we have that

$$\max_{n\delta\le k\le n-n\delta}\max_{r_1\le r\le r_2}|\hat{\sigma}_n^2(k,r)-\hat{\sigma}_n^2| \quad \xrightarrow{P} \quad 0$$

for all $0<\delta<1$ and $r_1<r_2$ satisfying $0<H(r_1)\le H(r_2)<1$ and therefore (4.8) follows from (4.7).

Proof of Theorem 2.1: Using a two-term Taylor expansion we have that

$$n(\log\hat{\sigma}_n^2-\log\hat{\sigma}_n^2(k,r)) = \frac{1}{\hat{\sigma}_n^2}n(\hat{\sigma}_n^2-\hat{\sigma}_n^2(k,r)) + \frac{1}{2\xi_n^2(k,r)}n(\hat{\sigma}_n^2-\hat{\sigma}_n^2(k,r))^2,$$

where $\xi_n(k,r)$ is between $\hat{\sigma}_n^2(k,r)$ and $\hat{\sigma}_n^2$. Lemmas 4.2, 4.3 and (4.2) yield that

$$n(\hat{\sigma}_n^2-\hat{\sigma}_n^2([nt],r)) \Longrightarrow \frac{\sigma^2\left(U(t,H(r))\right)^2}{(1-t)H(r)(H-(1-t)H(r))H} \tag{4.9}$$

in $\mathcal{D}([t_1, t_2] \times [r_1, r_2])$. It follows from Lemma 4.4 that

$$\max_{n\delta \le k \le n - n\delta} \max_{r_1 \le r \le r_2} \frac{1}{\xi_n^2(k, r)} = O_P(1),$$

and therefore (4.9) implies

$$\max_{n\delta \le k \le n - n\delta} \max_{r_1 \le r \le r_2} \frac{1}{\xi_n^2(k, r)} n(\hat{\sigma}_n^2 - \hat{\sigma}_n^2(k, r))^2 = O_P\left(\frac{1}{n}\right).$$

Now (4.7) and (4.9) yield Theorem 2.1.

Proof of Corollary 2.1: Theorem 2.1 implies that

$$\max_{nt_1 \le k \le nt_2} \sup_{r_1 \le r \le r_2} T_n(k, r) \xrightarrow{\mathcal{D}} \sup_{t_1 \le t \le t_2} \sup_{r_1 \le r \le r_2} \frac{(U(t, H(r)))^2}{(1-t)H(r)(H - (1-t)H(r))H}.$$

According to the scale transformation of the two-parameter Wiener process for any $S > 0$

$$\{S^{-1/2} W(t, Ss), 0 \le t, s\} \stackrel{\mathcal{D}}{=} \{W(t, s), 0 \le t, s\}, \tag{4.10}$$

which shows that with $\bar{H}(r) = H(r)/H$

$$\begin{aligned}
&\sup_{t_1 \le t \le t_2} \sup_{r_1 \le r \le r_2} \frac{(U(t, H(r)))^2}{(1-t)H(r)(H - (1-t)H(r))H} \\
&\quad = \sup_{t_1 \le t \le t_2} \sup_{r_1 \le r \le r_2} \frac{(H(r)(1-t)W(1, H) - (W(1, H(r)) - W(t, H(r)))H)^2}{(1-t)H(r)(H - (1-t)H(r))H} \\
&\quad \stackrel{\mathcal{D}}{=} \sup_{t_1 \le t \le t_2} \sup_{r_1 \le r \le r_2} \frac{\left(\bar{H}(r)(1-t)W(1, 1) - (W(1, \bar{H}(r)) - W(t, \bar{H}(r))\right)^2 H^3}{(1-t)\bar{H}(r)(1 - (1-t)\bar{H}(r))H^3} \\
&\quad = \sup_{t_1 \le t \le t_2} \sup_{u_1 \le u \le u_2} \frac{((1-t)uW(1, 1) - (W(1, u) - W(t, u)))^2}{(1-t)u(1 - (1-t)u)},
\end{aligned}$$

where $u_1 = \bar{H}(r_1)$ and $u_2 = \bar{H}(r_2)$. Computing the covariance functions, one can easily verify that

$$\{W(1, s) - W(t, s), 0 \le t \le 1, 0 \le s < \infty\} \stackrel{\mathcal{D}}{=} \{W(1-t, s), 0 \le t \le 1, 0 \le s < \infty\}, \tag{4.11}$$

which immediately implies

$$\begin{aligned}
&\sup_{t_1 \le t \le t_2} \sup_{u_1 \le u \le u_2} \frac{((1-t)uW(1, 1) - (W(1, u) - W(t, u)))^2}{(1-t)u(1 - (1-t)u)} \\
&\quad \stackrel{\mathcal{D}}{=} \sup_{1-t_2 \le s \le 1-t_1} \sup_{u_1 \le u \le u_2} \frac{(suW(1, 1) - W(s, u))^2}{su(1 - su)}.
\end{aligned}$$

The proof of Corollary 2.2 is based on the following lemma.

Lemma 4.5. *If the conditions of Corollary 2.2 are satisfied, then*

$$\sup_{-\infty<r<\infty} |\bar{H}_n(r) - \bar{H}(r)| \quad \to \quad 0 \;\; a.s. \tag{4.12}$$

and

$$\sup_{\bar{H}(r_1)\le t\le \bar{H}(r_2)} |\bar{H}_n^{-1}(t) - \bar{H}^{-1}(t)| \quad \to \quad 0 \;\; a.s. \tag{4.13}$$

for all $0 < \bar{H}(r_1) < \bar{H}(r_2) < 1$.

Proof: The ergodic theorem yields that for any r

$$\bar{H}_n(r) \quad \to \quad \bar{H}(r) \quad \text{a.s.} \tag{4.14}$$

The limit $\bar{H}(r)$ is bounded, monotone and continuous, so standard arguments show that (4.14) implies (4.12). By (4.12) we have

$$\sup_{0\le u\le 1} |\bar{H}_n(\bar{H}^{-1}(u)) - u| \quad \to \quad 0 \;\; \text{a.s.} \tag{4.15}$$

Using Horváth (1984) (cf. also Csörgő and Horváth (1993)) we conclude from (4.15) that

$$\sup_{0\le u\le 1} |\bar{H}(\bar{H}_n^{-1}(u)) - u| \quad \to \quad 0 \;\; \text{a.s.} \tag{4.16}$$

Since $\bar{H}$ is strictly increasing and continous, $\bar{H}^{-1}$ is a continous strictly increasing one-to-one mapping, and therefore (4.16) implies (4.13).

Proof of Corollary 2.2: It is an immediate consequence of Corollary 2.1 and Lemma 4.5.

Proof of Theorem 2.2: By Lemma 4.1 we have that

$$R_n(k,r) = \frac{1}{n^3\hat{\sigma}_n^2}\left(S_n(k,r)\sum_{i=1}^{n} y_{i-1}\varepsilon_i - Z_n \sum_{i=k+1}^{n} y_{i-1}\varepsilon_i I\{y_{i-1}\le r\}\right)^2,$$

so the theorem follows from (4.7) and Lemma 4.3.

Proof of Corollary 2.3: The result follows from Theorem 2.2 combined with (4.10) and (4.11).

5 Weak Convergence of a Two-Parameter Marked Empirical Process

It is immediate from Lemma 4.1 that the statistics discussed in Section 2 are closely connected to the two-parameter marked empirical process

$$G_n(t,r) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} y_{i-1}\varepsilon_i I\{y_{i-1} \le r\}.$$

Since we work with the stationary solution under the null hypothesis,

$$y_i = \sum_{k=0}^{\infty} \rho^k \varepsilon_{i-k}, \ -\infty < i < \infty \text{ with some } |\rho| < 1. \tag{5.1}$$

Throughout this section C stands for a generic constant whose value may change from line to line.

Theorem 5.1. *If* (2.1), (2.2) *and* (5.1) *hold, then*

$$G_n(t,r) \Longrightarrow \sigma W(t, H(r)) \ \text{in } \mathcal{D}([0,1] \times [-\infty, \infty]),$$

where W is a two-parameter Wiener process.

Remark: Although Theorem 5.1 suffices for the purposes of the present paper, it is worth pointing out that the result remains valid for a much larger class of linear processes, e.g., for processes

$$y_i = \sum_{k=0}^{\infty} \beta_k \varepsilon_{i-k}, \ -\infty < i < \infty,$$

where $\beta_k = O(\rho^k)$ for some $|\rho| < 1$. The proof requires only trivial changes.

The proof of Theorem 5.1 is based on the following tightness criterion.

Lemma 5.1. *Let $\{\zeta_i(s), 0 \le s \le 1, i \ge 1\}$ be non-decreasing processes in $\mathcal{D}[0,1]$, let $\zeta(s)$, $0 \le s \le 1$, be a non-decreasing function and define*

$$K_n(t,s) = \frac{1}{n^{1/2}} \sum_{i=1}^{[nt]} (\zeta_i(s) - \zeta(s)).$$

If there exist a $\tau > 2$, $C > 0$ and a sequence a_n such that $a_n\sqrt{n} \to 0$ and

$$\begin{aligned} &E(K_n(t_2, s) - K_n(t_1, s))^6 \leq C|t_2 - t_1|^\tau \quad \text{if} \quad |t_2 - t_1| \geq a_n, \\ &E(K_n(t, s_2) - K_n(t, s_1))^6 \leq C|s_2 - s_1|^\tau \quad \text{if} \quad |s_2 - s_1| \geq a_n \end{aligned} \tag{5.2}$$

and

$$n^{1/2} \sup_{|s_2 - s_1| \leq a_n} |\zeta(s_2) - \zeta(s_1)| \;\to\; 0, \tag{5.3}$$

then $K_n(t, s)$ is tight.

Proof. By a classical tightness criterion (see e.g. Billingsley (1968)) it suffices to show that for any $\varepsilon > 0$ and $\delta > 0$ there is an L such that

$$\limsup_{n\to\infty} P\left\{ \sup_{|t-t'|\leq 1/L} \sup_{|s-s'|\leq 1/L} |K_n(\Delta(t, t', s, s')| > \varepsilon \right\} < \delta, \tag{5.4}$$

where $\Delta(t, t', s, s')$ denotes the rectangle $[t, t'] \times [s, s']$ and for any rectangle Δ, $K_n(\Delta)$ denotes the increment of the function K_n over Δ. (Here, and in the rest of this proof, all rectangles have sides parallel with the coordinate axes.) Let $K_n(t, s) = K_n^{(1)}(t, s) - K_n^{(2)}(t, s)$, where

$$K_n^{(1)}(t, s) = n^{-1/2} \sum_{i=1}^{[nt]} \zeta_i(s) \quad \text{and} \quad K_n^{(2)}(t, s) = n^{-1/2} \sum_{i=1}^{[nt]} \zeta(s).$$

To simplify the formulas, we will assume that $N = 1/a_n$ is an integer; the proof in the general case requires only trivial changes. Let $\mathcal{P}_n$ denote the set of points in the unit square of which both coordinates belong to the set $\mathcal{Q}_n = \{ia_n, 0 \leq i \leq N\}$. Let Δ^* be the largest rectangle with vertices in $\mathcal{P}_n$ contained in $\Delta(t, t', s, s')$. Clearly, the difference $\Delta(t, t', s, s') \setminus \Delta^*$ is the union of

4 rectangles with one side not exceeding $1/L$ and the other not exceeding a_n and thus

$$\begin{aligned}
&\sup_{|t-t'|\le 1/L}\ \sup_{|s-s'|\le 1/L} |K_n(\Delta(t,t',s,s')| \\
&\quad\le \sup_{\substack{|t-t'|\le 1/L\\ t,t'\in\mathcal{Q}_n}}\ \sup_{\substack{|s-s'|\le 1/L\\ s,s'\in\mathcal{Q}_n}} |K_n(\Delta(t,t',s,s')| \\
&\qquad + 2 \sup_{|t-t'|\le 1/L}\ \sup_{\substack{|s-s'|\le a_n\\ s \text{ or } s'\in\mathcal{Q}_n}} |K_n(\Delta(t,t',s,s')| \\
&\qquad + 2 \sup_{\substack{|t-t'|\le a_n\\ t \text{ or } t'\in\mathcal{Q}_n}}\ \sup_{|s-s'|\le 1/L} |K_n(\Delta(t,t',s,s')| \\
&\quad =: I_1 + I_2 + I_3.
\end{aligned}$$

The increments in I_2 and I_3 are small. To see this, we let Δ be a rectangle in the unit square $[0,1]\times[0,1]$ as in the suprema in I_2 and I_3. Then we denote by Δ' the smallest rectangle with vertices belonging to $\mathcal{P}_n$ covering Δ. Clearly, the length of one side of the rectangle Δ' is equal to a_n, the other is $< 2/L$ for sufficiently large n. Thus we get, using the monotonicity of $\zeta(s)$ and $\zeta_i(s)$, that

$$|K_n(\Delta)| \le |K_n^{(1)}(\Delta)| + |K_n^{(2)}(\Delta)| \le |K_n^{(1)}(\Delta')| + |K_n^{(2)}(\Delta)| \le |K_n(\Delta')| + 2|K_n^{(2)}(\Delta')|.$$

Again by the monotonicity of ζ, relation (5.3) and $a_n\sqrt{n} \to 0$, the increment of $K_n^{(2)}(\Delta')$ is $o_n(1)$, uniformly over all rectangles on $\mathcal{P}_n$ with one side having length a_n. Thus we have

$$I_2 + I_3 \le 4 \sup_{\substack{|t-t'|<2/L\\ t,t'\in\mathcal{Q}_n}}\ \sup_{\substack{|s-s'|<2/L\\ s,s'\in\mathcal{Q}_n}} |K_n(\Delta(t,t',s,s')| + o_n(1)$$

and hence it suffices to prove (5.4) in the case when all vertices of the rectangle $\Delta(t,t',s,s')$

belong to $\mathcal{P}_n$. Let $s_i = ia_n$, $t_j = ja_n$, $0 \le i, j \le N$, and

$$\begin{aligned}
\tau_{i,j} &= K_n(\Delta(t_i, t_{i+1}, s_j, s_{j+1})),\\
\mathcal{T}_{i,j} &= \left\{(p,q) \in \mathbb{N}_0^2 : \frac{i}{L} \le pa_n \le \frac{(i+1)}{L}, \frac{j}{L} \le qa_n \le \frac{(j+1)}{L}\right\},\\
S(\mathcal{T}) &= \sum_{(i,j)\in\mathcal{T}} \tau_{i,j},\\
M(\mathcal{T}) &= \max\{|S(\mathcal{T}')| : \mathcal{T}' \subseteq \mathcal{T}\},
\end{aligned}$$

where $\mathcal{T}, \mathcal{T}'$ are rectangles in the unit square. Geometrically, the set $\{(t_i, s_j),\ 0 \le i, j \le N\}$ is a lattice with $(N+1)^2$ points and the sets $\mathcal{T}_{i,j}$ give a partition of this lattice into L^2 'square formed' sublattices. Clearly, any rectangle Δ in the original lattice with sides $< 2/L$ intersects at most 9 of the the L^2 sublattices $\mathcal{T}_{i,j}$ and thus

$$P\left\{\sup_{\substack{|t-t'|\le 1/L\\ t,t'\in\mathcal{Q}_n}} \sup_{\substack{|s-s'|\le 1/L\\ s,s'\in\mathcal{Q}_n}} |K_n(\Delta(t,t',s,s')| \ge \varepsilon\right\} \le \sum_{0\le i,j\le L-1} P\left\{M(\mathcal{T}_{i,j}) \ge \varepsilon/9\right\}.$$

Now (5.2) and Móricz (1977) yield

$$\sum_{0\le i,j\le L-1} P\{M(\mathcal{T}_{i,j}) \ge \varepsilon/9\} \le \sum_{0\le i,j\le L-1} \frac{1}{(\varepsilon/9)^6} EM^6(\mathcal{T}_{i,j}) \le \frac{C'}{\varepsilon^6} L^2 (1/L)^\tau$$

for some constant C', completing the proof. □

Lemma 5.1 can also be derived from Corollary 1 in Davydov and Zitikis (2008) when the convergence of the finite dimensional distributions is also assumed.

Proof of Theorem 5.1. We must show the convergence of the finite dimensional distributions and tightness. We use Lemma 5.1 to establish the tightness of $G_n(t,r)$. Their result can be used for partial sums of differences of non-decreasing random functions. Hence we write, as the first step in the proof, $G_n(t,r)$ in this form.

Let $x^+ = xI\{x \ge 0\}$ and $x^- = xI\{x \le 0\}$. We write

$$\varepsilon_i y_{i-1} = \varepsilon_i^+ y_{i-1}^+ + \varepsilon_i^+ y_{i-1}^- + \varepsilon_i^- y_{i-1}^+ + \varepsilon_i^- y_{i-1}^-.$$

We use the decomposition

$$
\begin{aligned}
G_n(t,r) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} y_{i-1}\varepsilon_i I\{y_{i-1} \le r\} \qquad (5.5)\\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left(y_{i-1}^{+}\varepsilon_i^{+} I\{y_{i-1} \le r\} - E(y_{i-1}^{+}\varepsilon_i^{+} I\{y_{i-1} \le r\})\right)\\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left(y_{i-1}^{+}\varepsilon_i^{-} I\{y_{i-1} \le r\} - E(y_{i-1}^{+}\varepsilon_i^{-} I\{y_{i-1} \le r\})\right)\\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left(y_{i-1}^{-}\varepsilon_i^{+} I\{y_{i-1} \le r\} - E(y_{i-1}^{-}\varepsilon_i^{+} I\{y_{i-1} \le r\})\right)\\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left(y_{i-1}^{-}\varepsilon_i^{-} I\{y_{i-1} \le r\} - E(y_{i-1}^{-}\varepsilon_i^{-} I\{y_{i-1} \le r\})\right).
\end{aligned}
$$

Note that y_{i-1} and ε_i are independent and thus $E(y_{i-1}\varepsilon_i I\{y_{i-1} \le r\}) = 0$. All terms in the right hand side of (5.5) are in the required form of Lemma 5.1. We show the tightness of the first term, the tightness of the other three terms can be established in the same way.

Since $F(x) = P\{y_0 \le x\}$ is continuous by (2.2), it is enough to estimate the increments of the partial sum process $\sum_{i=1}^{[nt]} \left(y_{i-1}^{+}\varepsilon_i^{+} I\{F(y_{i-1}) \le s\} - c(s)\right)\right), 0 \le t, s \le 1$, where the function $c(s)$ is defined by $c(s) = E(\varepsilon_i^{+} y_{i-1}^{+} I\{F(y_{i-1}) \le s\})$. Let $d(s) = E(y_{i-1}^{+} I\{F(y_{i-1}) \le s\})$. We split the increments of

$$
G_n^{+}(t,s) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left(y_{i-1}^{+}\varepsilon_i^{+} I\{F(y_{i-1}) \le s\} - E(y_{i-1}^{+}\varepsilon_i^{+} I\{F(y_{i-1}) \le s\})\right)
$$

in s into two terms:

$$
G_n^{+}(t,s_2) - G_n^{+}(t,s_1) = A_n(t;s_1,s_2) + B_n(t;s_1,s_2), \qquad (5.6)
$$

where

$$
A_n(t;s_1,s_2) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} (\varepsilon_i^{+} - E\varepsilon_i^{+}) y_{i-1}^{+} I\{s_1 < F(y_{i-1}) \le s_2\}
$$

and

$$
B_n(t;s_1,s_2) = \frac{E\varepsilon_0^{+}}{\sqrt{n}} \sum_{i=1}^{[nt]} \left[y_{i-1}^{+} I\{s_1 < F(y_{i-1}) \le s_2\} - E(y_{i-1}^{+} I\{s_1 < F(y_{i-1}) \le s_2\})\right].
$$

For any $s_1 \le s_2$ we have by the Hölder inequality that

$$d(s_2) - d(s_1) = E\left(y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\}\right) \le \left(E(y_{i-1}^+)^6\right)^{1/6} \left(P(s_1 < F(y_0) \le s_2)\right)^{5/6} \quad (5.7)$$
$$= \left(E(y_{i-1}^+)^6\right)^{1/6} (s_2 - s_1)^{5/6}.$$

For the increments we write

$$\varepsilon_i^+ y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\} - (c(s_2) - c(s_1))$$
$$= (\varepsilon_i^+ - E\varepsilon_i^+) y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\} + (E\varepsilon_i^+)\left[y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\} - (d(s_2) - d(s_1))\right].$$

It is easy to see that

$$E\left\{(\varepsilon_i^+ - E\varepsilon_i^+)^2 (y_{i-1}^+)^2 I\{s_1 < F(y_{i-1}) \le s_2\} | \mathcal{F}_{i-1}\right\}$$
$$= \left[E(\varepsilon_0^+ - E\varepsilon_0^+)^2\right] (y_{i-1}^+)^2 I\{s_1 < F(y_{i-1}) \le s_2\},$$

where $\mathcal{F}_i = \sigma(\varepsilon_j, -\infty < j \le i)$. So using the Rosenthal inequality for martingales (cf. Hall and Heyde (1980), pp. 23-24) we conclude

$$E\left(\sum_{i=1}^{[nt]} (\varepsilon_i^+ - E\varepsilon_i^+) y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\}\right)^6 \quad (5.8)$$
$$\le C\left\{\sum_{i=1}^{n} E|\varepsilon_i^+ - E\varepsilon_i^+|^6 E(y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\})^6\right.$$
$$\left. + E\left(\sum_{i=1}^{n} (E(\varepsilon_0^+)^2)(y_{i-1}^+)^2 I\{s_1 < F(y_{i-1}) \le s_2\}\right)^3\right\}.$$

Applying Hölder's inequality we obtain that

$$\sum_{i=1}^{n} E|\varepsilon_i^+ - E\varepsilon_i^+|^6 E(y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\})^6 = nE|\varepsilon_0^+ - E\varepsilon_0^+|^6 E(y_0^+ I\{s_1 < F(y_0) \le s_2\})^6$$
$$\le Cn(s_2 - s_1)^{\delta/(6+\delta)}.$$

Introducing $g(s) = E[(y_{i-1}^+)^2 I\{F(y_{i-1}) \le s\}]$ and

$$\eta_i(s_1, s_2) = (y_{i-1}^+)^2 I\{s_1 < F(y_{i-1}) \le s_2\} - (g(s_2) - g(s_1)),$$

we have that

$$E\left(\sum_{i=1}^{n}(y_{i-1}^{+})^2 I\{s_1 < F(y_{i-1}) \le s_2\}\right)^3 \tag{5.9}$$

$$= E\left(\sum_{i=1}^{n}\eta_i(s_1,s_2) + \sum_{i=1}^{n}(g(s_2)-g(s_1))\right)^3$$

$$\le 4E\left|\sum_{i=1}^{n}\eta_i(s_1,s_2)\right|^3 + 4[n(g(s_2)-g(s_1))]^3.$$

Since $E(y_i^+)^{6+\delta} < \infty$, we get by the Hölder inequality that

$$\begin{aligned} g(s_2)-g(s_1) =& E[(y_{i-1}^{+})^2 I\{s_1 < F(y_{i-1}) \le s_2\}] \\ &\le \left(E(y_{i-1}^{+})^{6+\delta}\right)^{1/(3+\delta/2)} (s_2-s_1)^{(4+\delta)/(6+\delta)}. \end{aligned}$$

Now we write

$$E\left|\sum_{i=1}^{n}\eta_i(s_1,s_2)\right|^3 \le \sum_{1\le i,\,j,\,k\le n} \left|E(\eta_i(s_1,s_2)\eta_j(s_1,s_2)\eta_k(s_1,s_2))\right|. \tag{5.10}$$

For any $c > 0$ we define the truncated sums

$$y_{i,n} = \sum_{0\le k\le c\log n} \rho^k \varepsilon_{i-k}.$$

First we show that for any $\kappa > 0$, $c = c(\kappa)$ in the previous sum can be chosen such that for all $1 \le i, j, k \le n$

$$|E(\eta_i(s_1,s_2)\eta_j(s_1,s_2)\eta_k(s_1,s_2)) - E(\eta_{i,n}(s_1,s_2)\eta_{j,n}(s_1,s_2)\eta_{k,n}(s_1,s_2))| \le Cn^{-\kappa}, \tag{5.11}$$

with

$$\eta_{i,n}(s_1,s_2) = (y_{i-1,n}^{+})^2 I\{s_1 < F(y_{i-1,n}) \le s_2\} - E\left((y_{i-1,n}^{+})^2 I\{s_1 < F(y_{i-1,n}) \le s_2\}\right),$$

where C does not depend on s_1, s_2. By Hölder's inequality we have

$$\begin{aligned}
&E|\eta_i(s_1,s_2)-\eta_{i,n}(s_1,s_2)|^3\\
&\quad\le 4E|(y_{i-1}^+)^2-(y_{i-1,n}^+)^2|^3+4E[(y_{i-1,n}^+)^2|I\{s_1<F(y_{i-1})\le s_2\}-I\{s_1<F(y_{i-1,n})\le s_2\}|]^3\\
&\quad\le C\Bigg\{E\Bigg(\sum_{c\log n<k<\infty}\rho^k|\varepsilon_{-k}|\Bigg)^3\\
&\qquad+[E(y_{i-1,n}^+)^{6+\delta}]^{6/(6+\delta)}[E|I\{s_1<F(y_{i-1})\le s_2\}-I\{s_1<F(y_{i-1,n})\le s_2\}|^{3(6+\delta)/\delta}]^{\delta/(6+\delta)}\Bigg\}.
\end{aligned}$$

Using (2.1), for any κ there is c such that

$$E\Bigg(\sum_{c\log n<k<\infty}\rho^k|\varepsilon_{-k}|\Bigg)^3\le Cn^{-\kappa}.$$

Also, by (2.1) there is a C such that $E(y_{i,n}^+)^{6+\delta}\le C$ for all c, n and i. Next we combine the definition of $y_{i,n}$ with (2.2), to see that for any given κ there is a c such that

$$[E|I\{s_1<F(y_{i-1})\le s_2\}-I\{s_1<F(y_{i-1,n})\le s_2\}|^{3(6+\delta)/\delta}]^{\delta/(6+\delta)}\le Cn^{-\kappa}.$$

Thus the proof of (5.11) can be completed via Hölder's inequality.

Using the definition of $y_{i,n}$ we observe that

$$E(\eta_{i,n}(s_1,s_2)\eta_{j,n}(s_1,s_2)\eta_{k,n}(s_1,s_2))=0,$$

if there is at least one index i, j, k which differs from the other two with more than $c\log n+1$. So to estimate (5.10) we need to consider the terms when the difference between the indices is less than $c\log n+1$. The number of these terms cannot be more than $n(c\log n+1)^2$. We claim that for such indices

$$E|\eta_{i,n}(s_1,s_2)\eta_{j,n}(s_1,s_2)\eta_{k,n}(s_1,s_2)|\le C(|s_2-s_1|^{\delta/(6+\delta)}+n^{-\kappa}). \tag{5.12}$$

To prove (5.12) we note that the $\eta_{i,n}$'s are differences and we compute the product resulting in a sum of eight terms. All the elements of the sum are three-term-products. One of the terms is

$(y_{i,n}^+)^2 I\{s_1 < F(y_{i,n}) \le s_2\}(y_{j,n}^+)^2 I\{s_1 < F(y_{j,n}) \le s_2\}(y_{k,n}^+)^2 I\{s_1 < F(y_{k,n}) \le s_2\}$ for which we have, on account of Hölder's inequality and (2.2),

$$\begin{aligned}
&E\left((y_{i,n}^+)^2 I\{s_1 < F(y_{i,n}) \le s_2\}(y_{j,n}^+)^2 I\{s_1 < F(y_{j,n}) \le s_2\}(y_{k,n}^+)^2 I\{s_1 < F(y_{k,n}) \le s_2\}\right)\\
&\quad \le E\left((y_{i,n}^+)^2 I\{s_1 < F(y_{i,n}) \le s_2\}\right)^3\\
&\quad = E\left((y_i^+)^6 I\{s_1 < F(y_i) \le s_2\}\right) + O(n^{-\kappa})\\
&\quad \le \left(E(y_i^+)^{6+\delta}\right)^{6/(6+\delta)} (s_2 - s_1)^{\delta/(6+\delta)} + O(n^{-\kappa}).
\end{aligned}$$

Similar inequalities can be established for all the other 7 terms, proving (5.12).

Putting together (5.8)-(5.12) we conclude that for all $0 \le t, s_1, s_2 \le 1$ that

$$\begin{aligned}
&EA_n^6(t; s_1, s_2) \qquad (5.13)\\
&\quad \le C\left\{n^{-\kappa} + \frac{1}{n^2}(s_2 - s_1)^{\delta/(6+\delta)} + (s_2 - s_1)^{3(4+\delta)/(6+\delta)} + \frac{1}{n^2}(\log n)^2 (s_2 - s_1)^{\delta/(6+\delta)}\right\}\\
&\quad \le C\left\{n^{-\kappa} + (s_2 - s_1)^{3(4+\delta)/(6+\delta)} + \frac{1}{n^2}(\log n)^2 (s_2 - s_1)^{\delta/(6+\delta)}\right\},
\end{aligned}$$

since we can assume without loss of generality that $0 < \delta < 1$.

Next we prove an upper bound for $EB_n^6(t; s_1, s_2)$. Similarly to $\eta_i(s_1, s_2)$ and $\eta_{i,n}(s_1, s_2)$ we define $\xi_i(s_1, s_2) = y_{i-1}^+ I\{s_1 < F(y_{i-1}) \le s_2\} - (d(s_2) - d(s_1))$ and

$$\xi_{i,n}(s_1, s_2) = y_{i-1,n}^+ I\{s_1 < F(y_{i-1,n}) \le s_2\} - E(y_{i-1,n}^+ I\{s_1 < F(y_{i-1,n}) \le s_2\}).$$

We note that

$$\begin{aligned}
EB_n^6(t; s_1, s_2) &\le \frac{C}{n^3} \sum_{1 \le i_1, \ldots, i_6 \le nt} \left| E\prod_{j=1}^6 \xi_{i_j}(s_1, s_2) - E\prod_{j=1}^6 \xi_{i_j,n}(s_1, s_2)\right|\\
&\quad + \frac{C}{n^3} \sum_{1 \le i_1, \ldots, i_6 \le nt} \left| E\prod_{j=1}^6 \xi_{i_j,n}(s_1, s_2)\right|\\
&= D_{1,n}(t; s_1, s_2) + D_{2,n}(t; s_1, s_2).
\end{aligned}$$

Following the arguments leading to (5.11), one can show that for any $\kappa > 0$ there is $c = c(\kappa)$ such that

$$\left| E \prod_{j=1}^{6} \xi_{i_j}(s_1, s_2) - E \prod_{j=1}^{6} \xi_{i_j,n}(s_1, s_2) \right| \le Cn^{-\kappa} \tag{5.14}$$

and therefore

$$D_{1,n}(t; s_1, s_2) \le Cn^{-\kappa}.$$

Now we estimate $E \prod_{j=1}^{6} \xi_{i_j,n}(s_1, s_2)$. Let us divide the indices $i_1, \dots, i_6$ into blocks so that the differences between indices within a block are less than $c \log n$ and between blocks are larger than $c \log n$. Clearly, if there is at least one block containing a single element, the expected value of the product is 0. So it suffices to consider the cases when all blocks contain at least two elements. This allows the cases of one single block with 6 elements, two blocks with $3 + 3$ or $4 + 2$ elements and finally 3 blocks with 2 elements each.

If there is only one block then we use again Hölder's inequality to conclude that

$$\left| E \prod_{j=1}^{6} \xi_{i_j,n}(s_1, s_2) \right| \le C(|s_2 - s_1|^{\delta/(6+\delta)} + n^{-\kappa}).$$

The number of such terms in $D_{2,n}(t; s_1, s_2)$ is at most $Cn(\log n)^5$. If there are two blocks in the product, then the blocks are independent and the numbers of terms in the blocks are 2 and 4 or 3 and 3. Using the independence and then Hölder's inequality for the terms in the same block we get

$$\left| E \prod_{j=1}^{6} \xi_{i_j,n}(s_1, s_2) \right| \le C(|s_2 - s_1|^{(6+2\delta)/(6+\delta)} + n^{-\kappa}).$$

The number of two-block terms in $D_{2,n}(t; s_1, s_2)$ is not more than $Cn^2(\log n)^4$. Finally, consider the case of three blocks of two elements each. As before, using Hölder's inequality we have

$$\left| E \prod_{j=1}^{6} \xi_{i_j,n}(s_1, s_2) \right| \le C(|s_2 - s_1|^{3(4+\delta)/(6+\delta)} + n^{-\kappa}).$$

The number of terms in $D_{2,n}(t;s_1,s_2)$ with 3 blocks is $Cn^3(\log n)^3$. Hence we obtained the following upper bound for $B_n(t;s_1,s_2)$:

$$\begin{aligned} & EB_n^6(t;s_1,s_2) \\ & \quad \le C\Big\{n^{-\kappa} + \frac{1}{n^2}(\log n)^5(s_2-s_1)^{\delta/(6+\delta)} + \frac{1}{n}(\log n)^4(s_2-s_1)^{(6+2\delta)/(6+\delta)} \\ & \qquad + (\log n)^3(s_2-s_1)^{3(4+\delta)/(6+\delta)}\Big\}. \end{aligned} \tag{5.15}$$

The upper bounds in (5.13) and (5.15) yield

$$\begin{aligned} & E(G_n^+(t,s_2) - G_n^+(t,s_1))^6 \\ & \quad \le C\Big\{n^{-\kappa} + \frac{1}{n^2}(\log n)^5(s_2-s_1)^{\delta/(6+\delta)} + \frac{1}{n}(\log n)^4(s_2-s_1) \\ & \qquad + (\log n)^3(s_2-s_1)^{3(4+\delta)/(6+\delta)}\Big\}. \end{aligned} \tag{5.16}$$

Next we consider the increments of $G_n^+(t,s)$ in t. By stationarity,

$$E\left(\sum_{i=[nt_1]+1}^{[nt_2]} (\varepsilon_i^+ y_{i-1}^+ I\{F(y_{i-1}) \le s\} - c(s))\right)^6 = E\left(\sum_{i=1}^{[nt_2]-[nt_1]} (\varepsilon_i^+ y_{i-1}^+ I\{F(y_{i-1}) \le s\} - c(s))\right)^6.$$

Next Rosenthal's inequality for martingale differences and the independence of ε_i and y_{i-1} yield

$$\begin{aligned} & E\left(\sum_{i=1}^{[nt]}(\varepsilon_i^+ - E\varepsilon_i^+)y_{i-1}^+ I\{F(y_{i-1}) \le s\}\right)^6 \\ & \quad \le C\left\{\sum_{i=1}^{[nt]} E\left[(\varepsilon_i^+ - E\varepsilon_i^+)^6 (y_{i-1}^+)^6\right] + E\left(\sum_{i=1}^{[nt]}(y_{i-1}^+)^2 E(\varepsilon_i^+ - E\varepsilon_i^+)^2\right)^3\right\} \\ & \quad \le C\left\{[nt] + [nt]^3 + E\left(\sum_{i=1}^{[nt]}\left((y_{i-1}^+)^2 - E(y_{i-1}^+)^2\right)\right)^3\right\} \\ & \quad \le C\left\{[nt] + [nt]^3 + [nt](\log n)^3 + n^{-\kappa}\right\}. \end{aligned}$$

In the last step we estimated the third moment in the second line above by using the truncated variables $y_{i-1,n}$.

Following the arguments leading to (5.15) one could easily verify that

$$E\left(\sum_{i=1}^{[nt_2]-[nt_1]+1}(y_{i-1}^{+}I\{F(y_{i-1})\le s\}-d(s))\right)^6$$
$$\le C\left\{n^{-\kappa}+([nt_2]-[nt_1]+1)(\log n)^5+([nt_2]-[nt_1]+1)^2(\log n)^4\right.$$
$$\left.+([nt_2]-[nt_1]+1)^3(\log n)^3\right\}.$$

Thus we have for all $0\le t_1\le t_2\le 1$ and $0\le s\le 1$

$$E(G_n^{+}(t_2,s)-G_n^{+}(t_1,s))^6$$
$$\le C\frac{1}{n^3}\left\{n^{-\kappa}+([nt_2]-[nt_1]+1)(\log n)^5+([nt_2]-[nt_1]+1)^2(\log n)^4\right. \tag{5.17}$$
$$\left.+([nt_2]-[nt_1]+1)^3(\log n)^3\right\}.$$

Let $a_n=n^{-\gamma}$ with $0.6<\gamma<1$. Choose β such that $\gamma<\beta<1$. If $s_2-s_1\ge a_n$, then

$$\frac{(\log n)^5}{n^2}(s_2-s_1)^{\delta/(6+\delta)}\le C(s_2-s_1)^{2/\beta},$$

$$\frac{(\log n)^5}{n}(s_2-s_1)\le C(s_2-s_1)^{1+1/\beta},$$

$$\frac{(\log n)^3}{n}(s_2-s_1)^{3(4+\delta)/(6+\delta)}\le C(s_2-s_1)^{1+1/\beta}.$$

Since $3(4+\delta)/(6+\delta)>2$, we can choose $\kappa>2$, such that by (5.16) there is $\tau>2$ such that for all $0\le t\le 1$ and $|s_2-s_1|\ge a_n$

$$E(G_n^{+}(t,s_2)-G_n^{+}(t,s_1))^6\le C(s_2-s_1)^{\tau}.$$

Similarly, by (5.17) we have for all $0\le s\le 1$ and $|t_2-t_1|\ge a_n$

$$E(G_n^{+}(t_2,s)-G_n^{+}(t_1,s))^6\le C(t_2-t_1)^{\tau}.$$

Since $\gamma>0.6$, it follows from (5.7) that

$$n^{1/2}\sup_{|s_2-s_1|\le a_n}E(y_0^{+}I\{s_1\le F(y_0)\le s_2\})\to 0.$$

The tightness of $G_n^+(t,s)$ follows from Lemma 5.1. By (5.5) we have the tightness of $G_n(t, F^{-1}(s))$ as well.

Since the partial sums of $\sum_{i=1}^{N} y_{i-1}\varepsilon_i I\{F(y_{i-1}) \leq s\},\ N \geq 1$ is also a martingale for all s, the convergence of the finite dimensional distributions of $G_n(t, F^{-1}(s))$ is a consequence of the central limit theorem for martingale differences (cf. Hall and Heyde (1980)).

Acknowledgements: Our research was supported by FWF grant S9603-N23 and OTKA grants K 61052 and K 67961 (István Berkes), NFS grants DMS 0604670 and DMS 0652420 (Lajos Horváth) and the Hong Kong Research Grants commission HKUST 6428/06H, HKUST 6016/07P and HKUST602609 (Shiqing Ling).

References

Bai, J. (1998) Likelihood ratio tests for multiple structural changes. *J. Econometrics* **91**, 299–323.

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

Bai, J. and Ng, S. (2005) Tests for skewness, kurtosis, and normality for time series data. *J. Bus. Econom. Statist.* **23**, 49–60.

Bai, J., Lumsdaine, R. L. and Stock, J. H. (1998) Testing for and dating common breaks in multivariate time series. *Rev. Econom. Stud.* **65**, 395–432.

Balcombe, K., Bailey, A. and Brooks, J. (2007) Threshold effects in price transmissions: the case of Brazilian wheat, maize, and soya prices. *Amer. J. Agr. Econ.* **89**, 308–323.

Berkes, I. and Horváth, L. (2001) Strong approximation for the empirical process of a GARCH sequence. *Ann. Applied Probab.* **11**, 789–809.

Berkes, I., Horváth, L., Kokoszka, P. and Shao, Q. M. (2006) On discriminating between long-range dependence and changes in the mean. *Annals of Statistics* **34**, 1140–1165.

Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods.* Second edition. Springer Series in Statistics. Springer-Verlag, New York.

Chan, K. S. (1990) Testing for threshold autoregression. *Annals of Statistics* **18**, 1886–1894.

Chan, K. S. and Tong, H. (1990) On likelihood ratio tests for threshold autoregression. *J. Roy. Statist. Soc. Ser. B* **52**, 469–476.

Chow, G. C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **28** 591–604.

Csörgő, M. and Horváth, L. (1993) *Weighted Approximations in Probability and Statistics*, Wiley, Chichester, U.K.

Csörgő, M. and Horváth, L. (1997) *Limit Theorems in Change-point Analysis.* Wiley, Chichester, U.K.

Csörgő, M. and Révész, P. (1981) *Strong Approximations in Probability and Statistics.* Academic Press, New York.

Davis, R. A., Huang, D. W. and Yao, Y. C. (1995) Testing for a change in the parameter values and order of an autoregressive model. *Annals of Statistics* **23**, 282–303.

Davydov, Y. and Zitikis, R. (2008). On the weak convergence of random fields. *Annals of the Institute of Statistical Mathematics* **60**, 345–365.

Hall, P. G. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Application.* Academic Press, New York.

Hansen, B. E. (2003) Tests for parameter instability in regressions with I(1) processes. Twentieth anniversary commemorative issue. *J. Bus. Econom. Statist.* **20**, 45–59.

Horváth, L. (1984) Strong approximation of renewal processes. *Stochastic Process. Appl.* **18**, 127–138.

Horváth, L. (1993) The maximum likelihood method for testing changes in the parameters of normal observations. *Ann. Statist.* **21**, 671–680.

Horváth, L. (1995) Detecting changes in linear regressions. *Statistics* **26**, 189–208.

Horváth, L., Kokoszka, P. and Teyssiére, G. (2001) Empirical processes of the squared residuals of an ARCH sequence. *Annals of Statistics* **29**, 445–469.

Kaczor, W. J. and Nowak, M. T. (2001) *Problems in Mathematical Analysis II.* AMS, Providence, RI.

Li, W. K. and Lam, K. (1995) Modelling asymmetry in stock returns by a threshold autoregressive conditional heteroscedastic model. *J. Roy. Statist. Soc. Ser. D* **44**, 333–341.

Ling, S. (2007) Limiting theorems for NED sequences and testing for change-points in time series models. *Annals of Statistics* **35**, 994–1012.

Ling, S. and Tong, H. (2005) Testing for a linear MA model against threshold MA models. *Annals of Statistics* **33**, 2529–2552.

Móricz, F. (1977) Moment inequalities for the maximum of partial sums of random fields. *Acta Sci. Math. (Szeged)* **39**, 353–366.

Quandt, R. E. (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes. *J. Amer. Statist. Assoc.* **55**, 324–330.

Seber, G. A. F. and Lee, A. J. (2003) *Linear Regression Analysis*, Wiley, Chichester, U.K.

Wong, C. S. and Li, W. K. (1997) Testing for threshold autoregression with conditional heteroscedasticity. *Biometrika* **84**, 407–418.

Wong, C. S. and Li, W. K. (2000) Testing for double threshold autoregressive conditional heteroscedastic model. *Statist. Sinica* **10**, 173–189.

Yao, Y. C. and Davis, R. A. (1986) The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates. *Sankhya Ser. A* **48**, 339–353.

A Tables

$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
2.343	2.758	3.604

Table 1: Simulated critical values

	Corn	Soybeans
$\hat{\phi}_1$	0.131	0.026
$\hat{\phi}_2$	0.344	0.430
r	0.076	0.090

Table 2: Parameter estimators

B Figures

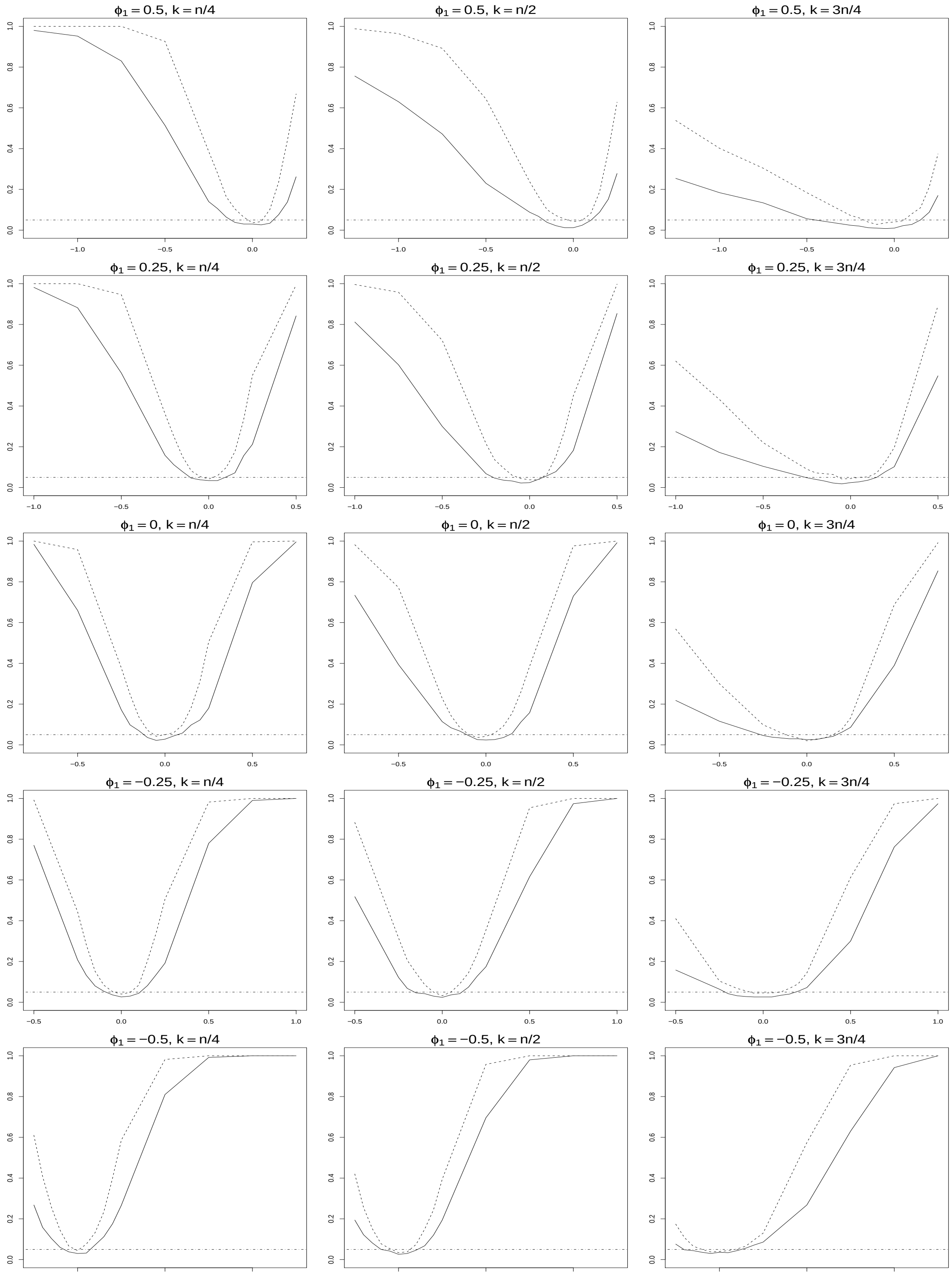

Figure 1: Empirical power curves with $\alpha = 0.05$ (reference line), $n = 200$ (solid) and $n = 400$ (dashed), where the x-axis corresponds to the values of ϕ_2.

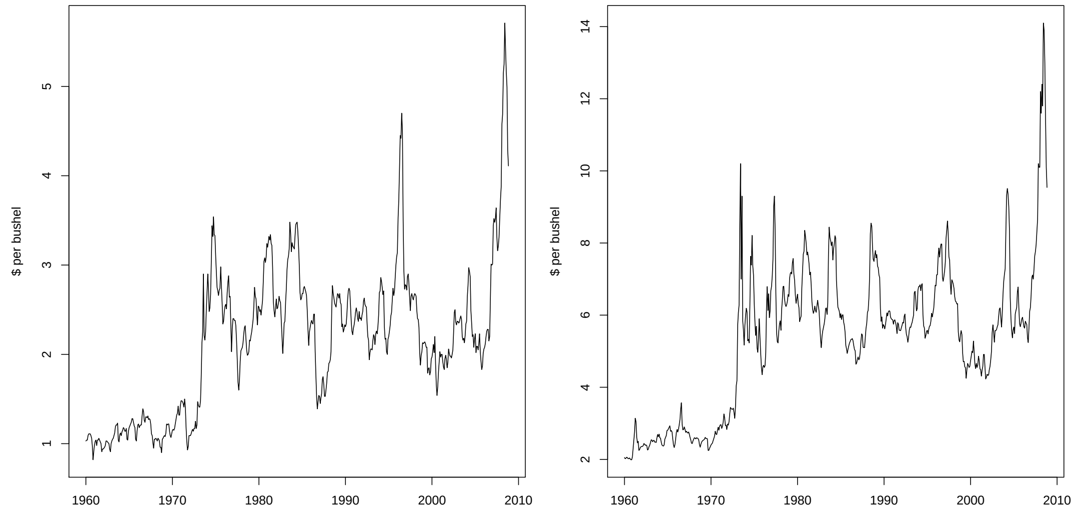

Figure 2: Monthly average corn (left) and soybeans (right) prices from 1960-2008

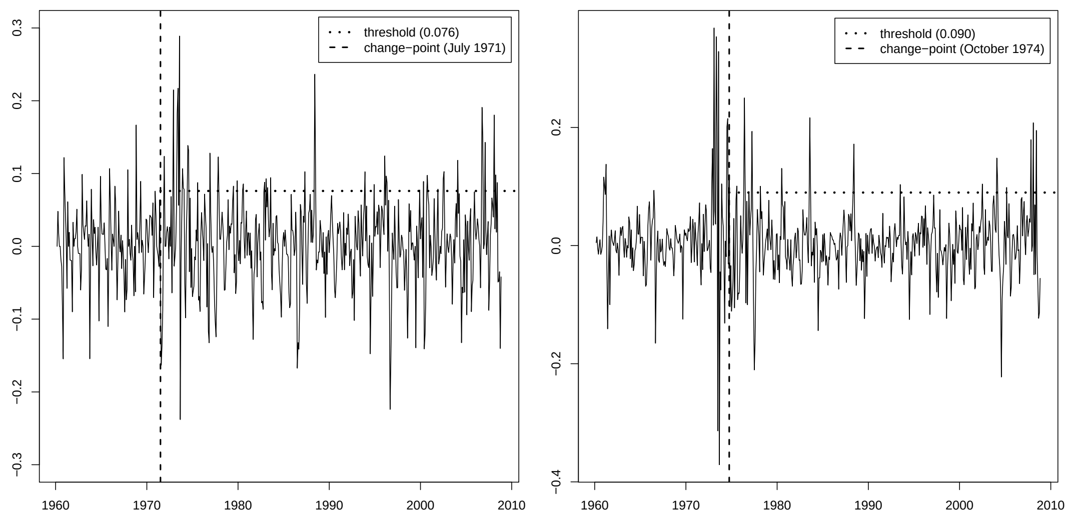

Figure 3: Relative monthly average corn (left) and soybeans (right) price changes from 1960-2008

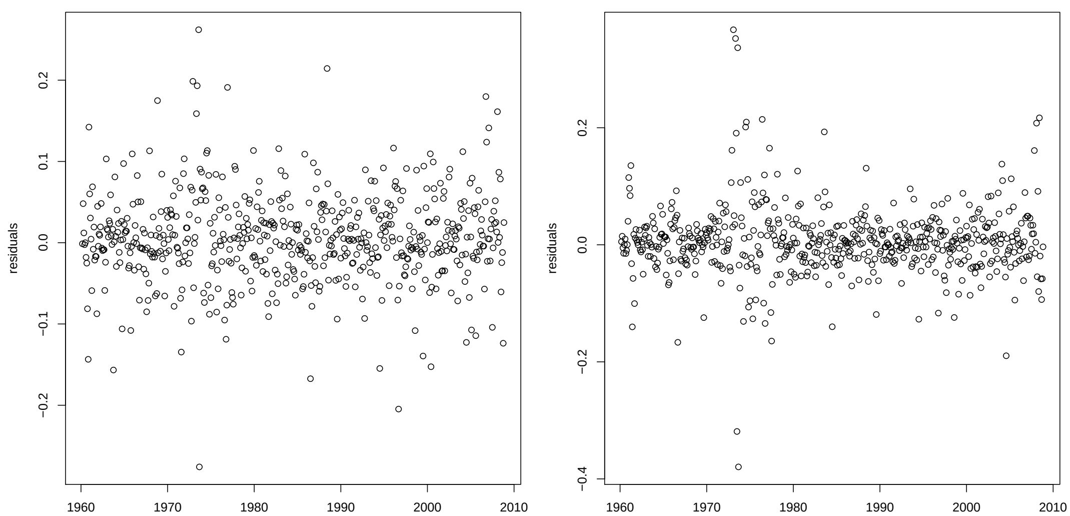

Figure 4: Residual plots of corn data set (left) and soybeans data set (right)