

Comparison of Structural Change Tests in Linear Regression Models

Jaehee Kim¹

¹Department of Statistics, Duksung Women's University

(Received August 2011; accepted October 2011)

Abstract

The actual power performance of historical structural change tests are compared under various alternatives. The tests of interest are F , CUSUM, MOSUM, Moving Estimates and empirical distribution function tests with both recursive and ordinary least-squares residuals. Our comparison of the structural tests involves limiting distributions under the null hypothesis, the ability to detect the alternative hypotheses under one or double structural change, and smooth change in parameters. Even though no version is uniformly superior to the other, the knowledge about the properties of those tests and connections between these tests can be used in practical structural change tests and in further research on other change tests.

Keywords: Brownian bridge process, Brownian motion process, change-point model, CUSUM, empirical distribution functional test, MOSUM, moving estimates test, recursive residual, stochastic process, structural change.

1. Introduction

Methods to detect structural changes or parameter instabilities in the linear regression models have been developed in econometrics and statistics communities. Structural stability is of importance in regression models. A form of model misspecification as parameter nonconstancy, may have severe consequences on inference if left undetected. In particular, if the data generating process changes in some ways, then forecasts loss accuracy. In consequences, many econometricians routinely apply tests for parameter change. Significant effort has been devoted to obtaining convenient and powerful tests in a variety of modeling contexts.

Consider the following linear regression model

$$Y_i = \mathbf{x}_i' \boldsymbol{\beta}_i + \epsilon_i, \quad i = 1, \dots, T, \quad (1.1)$$

where $\boldsymbol{\beta}_i$ is a $k \times 1$ vector and the errors $\epsilon_1, \dots, \epsilon_T$ is a martingale difference sequence with respect to some sequence of σ -algebras $\{F_i\}$ with $E(\epsilon_i^2 | F_{i-1}) = \sigma^2$, and that \mathbf{x}_i is $\{F_{i-1}\}$ measurable and obeys the weak laws of large numbers.

This research was supported by the Basic Science Research Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology(2010-05-0014).

¹Professor, Department of Statistics, Duksung Women's University, Seoul 132-714, South Korea.

E-mail: jaehee@duksung.ac.kr

Tests on structural change are concerned with testing the null hypothesis of “no structural change”

$$H_0 : \beta_i = \beta_0, \quad \text{for all } i, \quad (1.2)$$

which means no change in the regression parameter, against the alternative that the coefficient vector varies over time.

Various tests have been developed for different alternatives and for different estimation techniques. These tests can be divided in three classes that are differently suitable for certain patterns of deviation from the null hypothesis. The first class is the tests from the F test framework (Hansen, 1992, Andrews, 1993) that assumes one breakpoint under the alternative. The second class is the tests from the generalized fluctuation test framework (Kuan and Hornik, 1995, Zeileis, 2005) that can detect various types of structural changes. The third class is a nonparametric tests based on the empirical distribution functions (Bai, 1996).

Chow (1960) was the first to suggest the F test if the potential change-point is known. Andrews (1993) and Andrews and Ploberger (1994) suggested three types of test statistics - sup F , ave F , exp F - that are based on different kinds of F -statistics for the one time alternative change. The asymptotic theory was established for models estimated by generalized methods of moments(GMM) by Andrews (1993). For unknown change-points F statistics can be calculated for an interval of potential change-points and their supremum can be used as the test statistic. Such a test rejects the null hypothesis if one of the computed F statistics increases over a certain critical value or if the path of F statistics crosses a constant boundary. The sequences of F statistics can be treated as an empirical fluctuation process.

Starting from the recursive CUSUM test of Brown *et al.* (1975), a large variety of fluctuation tests for structural change in linear regression models by ordinary least squares(OLS) have been suggested. These tests are typically derived without a particular pattern of deviation from the parameter stability and are emphasized as suitable as an explorative tool. Ploberger and Krämer (1992), motivated by the intuition that least-squares residuals can better approximate true disturbances under the null hypothesis, proposed a CUSUM test based on recursive and least-squares residuals. However, Bauer and Hackl (1978) found that cumulated sums of recursive residuals are not sensitive to certain types of parameter changes, and proposed a MOSUM test based on moving sums of recursive residuals. Chu *et al.* (1995a) proposed a least-squares-MOSUM test based on least-squares residuals and its limiting distribution is a standard Wiener process. In particular, fluctuation tests based on OLS residuals and MOSUM tests are popular because they are easy to compute and to interpret. The OLS-based CUSUM and MOSUM tests have similar properties as the corresponding estimate-based processes and under a single shift alternative the OLS-CUSUM path should have a peak and the OLS-MOSUM path a shift around the change-point. As an extension of invariance principle in the sequential testing, Horváth *et al.* (2004) proposed two monitoring schemes to detect a structural change. Aue *et al.* (2006) developed asymptotic theory for monitoring schemes to detect a change in the regression parameters.

Fluctuation tests are either based on estimates or on residuals. The idea of the estimates-based tests is, that if there is a structural change in the data, the estimate of the regression coefficients (based on all data) should be substantially different from the subsample estimates of the data that do not contain the structural changes; however, these estimates should be smaller if the true coefficients remain constant over time. In this case an empirical process can be computed by the differences of these subsample estimates with the overall estimate. The resulting empirical process under the null hypothesis is essentially governed by the functional central limit theorem; however,

under the alternative, this empirical process will ‘fluctuate’ in the sense that its behavior is not completely characterized by the functional central limit theorem. A test for parameter constancy can be obtained by assigning an appropriate functional to measure the fluctuation of the empirical process of parameter estimates; in addition, this test rejects the null hypothesis if this process fluctuates significantly.

Loynes (1980) studied the empirical distribution of generalized residuals that are residuals from generalised regression and gave a hint for the use of a goodness-of-fit test for a model based on the empirical distribution function of the residuals. Resembling the prototypical Kolmogorov-Smirnov two-sample test, there are tests based on the non-weighted sequential empirical processes of residuals that include Csörgö and Horváth (1987, 1997). Carlstein (1988) and Dümbgen (1991) proposed to estimate a change-point under the alternative hypothesis based on test statistics. Bai (1996) extended this type test to regression models with estimated parameters as an empirical distribution function approach. Kuan and Hornik (1995) gave a unifying view as fluctuation tests for parameter constancy without assuming a specific alternative. Zeileis (2005) also provided a unified approach to structural change tests by embedding the tests into the same framework and Zeileis *et al.* (2007) developed the R package for structural change tests. Silva and Teixeira (2008) surveyed structural change analysis which made an important tradition in economic theory and testified the recent rise of interest in structural change where technological issues gained increased relevance. There has been a strong impetus towards empirically led work and a demand for the appropriate treatment to detect changes.

In this paper we provide a comparison of some historical tests for structural change. We compare the F tests, several fluctuation tests and empirical distribution functional tests with their statistical properties as limiting distributions and empirical behaviors via simulation.

This paper provides the asymptotic properties and simulation results for a wide variety of structural change models that arise in econometric applications and provides tests that can accommodate different situations with structural change. The paper is organized as follows. We introduce the regression model and the structural change problems in Section 1. F tests are described in Section 2. Generalized fluctuation tests, CUSUM and MOSUM tests, and Moving Estimate tests are explained in Section 3. Empirical distribution function tests based on the nonparametric empirical distribution, are considered in Section 4. A numerical power comparison of these tests via simulation is provided and an illustration with a macroeconomic example is included in Section 5. Section 6 concludes the paper.

2. F Tests

Andrews (1993) and Andrews and Ploberger (1994) suggested three types of test statistics: $\text{Sup}F$, $\text{Avg}F$ and $\text{Exp}F$ statistics that are based on different kind of F statistics such as Wald, LM or LR statistics in a very general class of models fitted by GMM (generalized methods of moments). These statistics are easy to interpret as well as possess certain optimal properties against single shift alternatives. The other two require partial sample estimates before and after a hypothetical break point moved over a subset of the sample.

For every potential change-points a F statistic is computed. For this OLS model is fitted for the observations before and after the potential change-point. At t^{th} point, with k regressors, the error sum of squares (ESS_t) with OLS residuals from a segmented regression, and the residual sum of

squares(RSS) from the unsegmented model are computed as

$$F_T(t) = \frac{\text{RSS} - \text{ESS}_t}{\text{ESS}_t / (T - 2k)}.$$

Let $F_T(t)$ denote a Wald, Lagrange multiplier(LM), or likelihood ratio statistic of the hypothesis of no structural change ($\beta = \beta_0$) for given t , where $1 < t_1 < t \leq t_2 < T$.

$$\text{Sup}F = \sup_{t_1 \leq t \leq t_2} F_T(t). \quad (2.1)$$

The Andrews and Ploberger (1994) ‘‘Exp’’ and ‘‘Avg’’ tests are

$$\text{Exp}F = \ln \left(\frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} \exp \left(\frac{1}{2} F_T(t) \right) \right), \quad (2.2)$$

and

$$\text{Avg}F = \frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} F_T(t). \quad (2.3)$$

As shown by Andrews (1993) and Andrews and Ploberger (1994), under a wide class of regularity conditions, these statistics have the asymptotic null distributions as follows

$$\begin{aligned} \text{Sup}F_T &\rightarrow_d \sup_{\pi_1 \leq \tau \leq \pi_2} F(\tau), \\ \text{Exp}F_T &\rightarrow_d \ln \left(\frac{1}{\pi_2 - \pi_1} \int_{\pi_1}^{\pi_2} \exp \left(\frac{1}{2} F(\tau) \right) d\tau \right) \end{aligned}$$

and

$$\text{Avg}F_T \rightarrow_d \frac{1}{\pi_2 - \pi_1} \int_{\pi_1}^{\pi_2} F(\tau) d\tau,$$

where $\pi_1 = t_1/T$, $\pi_2 = t_2/T$ and

$$F(\tau) = \frac{(W(\tau) - \tau W(1))'(W(\tau) - \tau W(1))}{\tau(1 - \tau)}$$

here $W(\tau)$ is a vector Brownian motion.

Hansen (1997) suggested a more accurate calculation of asymptotic p -values.

3. Generalized Fluctuation Tests

Zeileis (2005) provided a unifying view on the structural change tests by generalized M-fluctuation test framework. The unified tests are based on the same FCLT and just use different functionals to compute test statistics. To assess the null hypothesis, the parameter β is estimated by M-estimation that includes ML, OLS, and other robust estimation techniques. The parameter estimate $\hat{\beta}$ is computed once for the full sample (assuming H_0 is true) along with a corresponding fluctuation process that captures departures from stability. The estimate and the corresponding fluctuation process depend on the choice of suitable estimating function $\psi(\cdot)$, which should have zero expectation

at the true parameters $E[\psi(Y_i, \mathbf{x}_i, \beta_i)] = 0$. Hence under the null hypothesis,

$$\sum_{i=1}^T \psi(Y_i, \mathbf{x}_i, \hat{\beta}_i) = 0,$$

and the decorrelated partial sums of the expression on the left can be used as the fluctuation process capturing structural changes over time. The resulting cumulative score process is referred to as the empirical fluctuation process $\text{efp}(\cdot)$ and is formally defined as

$$W_n(t, \hat{\theta}) = n^{-\frac{1}{2}} \sum_{i=1}^{[Th]} \psi(Y_i, \mathbf{x}_i, \hat{\beta}_i), \tag{3.1}$$

$$\text{efp}(t) = \hat{\mathbf{J}}^{-\frac{1}{2}} W_n(t, \hat{\theta}), \tag{3.2}$$

where $\hat{\mathbf{J}}$ is some suitable consistent estimate of the covariance matrix of the scores $\psi(Y_i, \beta)$.

Under the null hypothesis, an FCLT(functional central limit theorem) holds: the empirical fluctuation process $\text{efp}(\cdot)$ converges to a k -dimensional Brownian bridge $W^0(\cdot)$, which can be written as $W^0(t) = W(t) - tW(1)$, where $W(\cdot)$ is a standard k -dimensional Brownian motion. Under the alternative, the fluctuation process should typically exhibit peaks at the times changes in β occur.

3.1. CUSUM tests based on residuals

Starting with Brown *et al.* (1975), residuals became one of the most important tools in change-point analysis for testing the constancy of regression relationships. In Brown *et al.* (1975) the recursive residual is defined as

$$w_n = \frac{Y_n - \mathbf{x}'_n \hat{\beta}_{n-1}}{\sqrt{1 + \mathbf{x}'_n (\mathbf{X}'_{n-1} \mathbf{X}_{n-1})^{-1} \mathbf{x}_n}}, \quad n = k + 1, \dots, T, \tag{3.3}$$

where $\mathbf{X}'_{n-1} = [\mathbf{x}'_1, \dots, \mathbf{x}'_{n-1}]$ and $\mathbf{Y}'_n = [Y_1, \dots, Y_n]$. Brown *et al.* (1975) showed that under H_0 , w_{k+1}, \dots, w_T are independent $N(0, \sigma^2)$. For $s = 1, \dots, T$, the s^{th} cumulated sum of recursive residuals is

$$R(s) = \frac{1}{\hat{\sigma} T^{\frac{1}{2}}} \sum_{n=k+1}^s w_n, \quad \text{for } s = k + 1, \dots, T, \tag{3.4}$$

where $\hat{\sigma}$ denotes the estimated standard deviation. For example the first-order difference-based estimator

$$\hat{\sigma}_R^2 = \frac{1}{2(n-1)} \sum_{i=2}^T (Y_i - Y_{i-1})^2 \tag{3.5}$$

proposed by Rice (1984) or $\hat{\sigma}^2 = \sum_{n=1}^T (Y_n - \mathbf{x}'_n \hat{\beta}_T)^2 / (T - k)$ can be used. It rejects for large values of

$$T_{recus} = \max_{1 \leq s < T} |R(s)|. \tag{3.6}$$

From the properties of the w_n 's under H_0 , the sequence $R(k + 1), \dots, R(T)$ is a sequence of approximately normal variables such that

$$E[R(s)] = 0, \quad V[R(s)] = s - k \quad \text{and} \quad C[R(s), R(r)] = \min(s, r) - k. \tag{3.7}$$

$R(s)$ is approximated by the continuous Gaussian process with the mean and covariance functions in (3.7), which is the Brownian motion process starting from zero at time $t = k$.

Ploberger and Krämer (1992) introduced the least-squares CUSUM test based on the maximal fluctuation of these cumulated sums, and showed that its asymptotic distribution. The ordinary least-squares residuals

$$e_n = Y_n - \mathbf{x}'_n \hat{\boldsymbol{\beta}}_T, \quad n = 1, \dots, T \quad (3.8)$$

are defined with the least squares parameter estimate $\hat{\boldsymbol{\beta}}_T$.

Note that there are T least-squares residuals, but only $T - k$ recursive residuals. Unlike recursive residuals, OLS residuals usually change whenever another observation is added to the sample. In addition, they are correlated and heteroskedastic even under H_0 . However, they better approximate the true disturbances under H_0 . The OLS residuals sum to zero so that the cumulated sums do not drift off after the structural change, as often happens with recursive residuals.

For $s = 1, \dots, T$, the s^{th} cumulated sum of least-squares residuals is

$$L(s) = \frac{1}{\hat{\sigma} T^{\frac{1}{2}}} \sum_{n=1}^s e_n \quad (3.9)$$

and the test statistic is

$$T_{ols\ cus} = \max_{1 \leq s < T} |L(s)|. \quad (3.10)$$

Its asymptotic distribution is given in Ploberger and Krämer (1992) as

$$L(s) \rightarrow B(z), \quad \text{in distribution as } T \rightarrow \infty,$$

where $B(z)$, $0 \leq z \leq 1$ is the standard Brownian bridge.

3.2. MOSUM tests

Bauer and Hackl (1978, 1980) found that the cumulated sums of recursive residuals are not very sensitive to certain type of parameter changes, and proposed the use of moving sums (MOSUMs) of recursively calculated residuals for testing the constancy of the parameters. While all past observations contribute to CUSUM test statistics, MOSUMs use the sum of a fixed number of past residuals being taken as a test criterion at each point of time. Chu *et al.* (1995a) considered the MOSUM test for the regression models and characterized the limiting process of moving sums of recursive residuals and those of least squares residuals in terms of the increments of a standard Wiener process.

MOSUM of recursive residual w_n in (3.3) is

$$M_{rec}(j, h) = \frac{1}{\hat{\sigma} [Th]^{\frac{1}{2}}} \sum_{n=j+1}^{j+[Th]} w_n. \quad (3.11)$$

The corresponding recursive-residual-MOSUM statistic is defined as

$$T_{rec\ mos} = \max_{1 \leq j < T - [Th]} |M_{rec}(j, h)|. \quad (3.12)$$

Chu *et al.* (1995a) showed that under H_0 ,

$$T_{rec\ mos} \rightarrow \max_{0 \leq z \leq 1-h} |S_h(z)| \equiv \max_{0 \leq t \leq \frac{1}{h}-1} |W(t+1) - W(t)|,$$

where $S_h(z) = h^{-1/2}\{W(z+h) - W(z)\}$ ($0 \leq z \leq 1-h$) and W is a standard Wiener process.

The significance limits for the simultaneous test of MOSUMs may be taken constant over time. For $j = 0, 1, \dots, T - [Th]$, the j^{th} moving sum of least-squares residual is

$$M_{ls}(j, h) = \frac{1}{\hat{\sigma}[Th]^{\frac{1}{2}}} \sum_{n=j+1}^{j+[Th]} e_n, \tag{3.13}$$

and the corresponding least-squares-MOSUM statistic is defined as

$$T_{rec\ mos} = \max_{1 \leq j < T-[Th]} |M_{ls}(j, h)|. \tag{3.14}$$

Chu *et al.* (1995b) showed that under H_0 ,

$$T_{rec\ mos} \rightarrow \max_{0 \leq z \leq 1-h} |S_h^0(z)| \equiv \max_{0 \leq t \leq \frac{1}{h}-1} |B(t+1) - B(t)|,$$

where $S_h^0(z) = h^{-1/2}\{B(z+h) - B(z)\}$ ($0 \leq z \leq 1-h$) and B is a Brownian bridge on $[0, 1]$.

The MOSUM test differs from the CUSUM test in that each moving sum contains a fixed number of residuals, whereas cumulated sums incorporate more residuals. An important issue in the MOSUM test is the choice of bandwidth h . If h is large, each moving sum includes too many residuals, and only a few moving sums are available to detect possible changes. Hence, moving sums with a large h are not very sensitive to parameter variation; however, if h is small, the sample variation in the moving sums is likely to be large, and the limit distribution may not be a good approximation.

3.3. Moving estimates tests

For the parameter constancy test, it is natural to consider the difference of parameter estimators at each possible change-point. Chu *et al.* (1995b) proposed the moving-estimates test which is asymptotically equivalent to the maximal likelihood ratio test under the alternative of double structural change in parameters. Moving estimate(ME) test is determined by the fluctuation of moving estimates computed from a sequence of subsamples of the same size. As moving estimates implement a locally weighted regression, the ME test can be interpreted as a nonparametric test for a non-constant mean function. The ME test can be robustified to allow for heteroskedasticity and serial correlation using the appropriate covariance computation.

$$T_{ME,h} = \max_{1 \leq n < T-[Th]} \frac{[Th]}{\hat{\sigma}\sqrt{T}} \left\| \mathbf{Q}_T^{\frac{1}{2}} \left(\tilde{\beta}_T(n, [Th]) - \hat{\beta}_T \right) \right\|, \tag{3.15}$$

where h is the bandwidth of moving windows characterizing the number of observations in the window, $\hat{\sigma}^2$ is an estimator of σ^2 , and $\|\cdot\|$ is some suitable norm. Also

$$\mathbf{Q}_T = T^{-1} \sum_{i=1}^T \mathbf{x}_i \mathbf{x}_i',$$

$$\tilde{\beta}(n, [Th]) = \left(\sum_{i=n+1}^{k+[Th]} \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \left(\sum_{i=n+1}^{k+[Th]} \mathbf{x}_i y_i \right), \quad n = 0, \dots, T - [Th],$$

here $\mathbf{Q}_T \rightarrow \mathbf{Q}$ almost surely and $\hat{\beta}_T$ is the full-sample OLS estimate of β .

Note that moving estimates are computed from windows (subsamples), each with $[Th]$ observations, moving across the whole sample. This is in contrast with the recursive estimates that are obtained from growing windows. That is, an OLS estimate from a potential shifting period with the duration $[Th]$ is just a moving estimate with the window size $[Th]$.

In the more general case where $\{\epsilon_i\}$ are heteroskedastic and/or serially correlated, the ME tests can be robustified by computing the covariance matrix properly as

$$T_{gME,h} = \max_{1 \leq n < T-[Th]} \frac{[Th]}{\sqrt{T}} \left\| \hat{\mathbf{D}}_T^{-\frac{1}{2}} \left(\tilde{\beta}_T(n, [Th]) - \hat{\beta}_T \right) \right\|, \quad (3.16)$$

where $\hat{\mathbf{D}}_T$ is a consistent estimator of $\mathbf{D} = \mathbf{Q}^{-1} \mathbf{V} \mathbf{Q}^{-1}$ such as $\hat{\mathbf{D}}_T^{-1/2} = \hat{\mathbf{V}}_T^{-1/2} \mathbf{Q}_T$, and

$$\mathbf{V} = \lim_{T \rightarrow \infty} \frac{1}{T} E \left[\left(\sum_{i=1}^T \mathbf{x}_i \epsilon_i \right) \left(\sum_{i=1}^T \mathbf{x}_i \epsilon_i \right)' \right].$$

For example, Newey and West (1987) proposed a consistent covariance estimator. By letting the bandwidth h of moving windows tend to zero at a suitable rate, one can obtain a consistency result for a general, unknown regression function.

Under the null hypothesis, if $\hat{\mathbf{V}}_T$ is consistent for \mathbf{V} ,

$$T_{ME,h} \Rightarrow \max \|M_h\| = \max \|W^0(t+h) - W^0(t)\|, \quad \text{for } 0 \leq t \leq 1-h,$$

where W^0 is the vector Brownian bridge process. Therefore the limiting distribution of the ME test is determined by the behavior of the limiting process M_h and the associated norm.

4. Empirical Distribution Functional Tests

The empirical distribution function (the cumulative distribution function associated with the empirical measure of the sample) estimates the true underlying cdf of the points in the sample. Loynes (1980) studied the empirical distribution of the generalized residuals in regression models and gave a hint for the use of goodness-of-fit test for a model based on the empirical distribution function of the residuals. Resembling the prototypical Kolmogorov-Smirnov two-sample test using empirical distribution functions, Csörgő and Horváth (1987, 1997) proposed change tests based on nonweighted sequential empirical processes of residuals. Carlstein (1988) and Dümbgen (1991) proposed to estimate a change-point under the alternative hypothesis based on these test statistics. Bai (1996) extended this type test to regression models with estimated parameters as an empirical distribution function approach.

For each fixed n , the empirical distribution function(edf) based on the first n residuals is defined as

$$\hat{F}_n(z) = \frac{1}{n} \sum_{t=1}^n I(\hat{\epsilon}_t \leq z)$$

and the edf based on the last $T - n$ residuals as

$$\hat{F}_{T-n}^*(z) = \frac{1}{T-n} \sum_{t=n+1}^T I(\hat{\epsilon}_t \leq z),$$

where $I(\cdot)$ is the indicator function. Let

$$T\left(\frac{n}{T}, z\right) = \frac{n}{T} \left(1 - \frac{n}{T}\right) \sqrt{n} \left(\hat{F}_n(z) - \hat{F}_{T-n}^*(z)\right). \tag{4.1}$$

The test statistic based on the edf's proposed by Bai (1996) is

$$T_{edf} = \max_{1 \leq n < T} \sup_z \left| T\left(\frac{n}{T}, z\right) \right|, \tag{4.2}$$

where $\|\mathbf{y}\|_\infty = \max\{|y_1|, \dots, |y_p|\}$ the supremum with respect to z is taken over the entire real line. This test T_{edf} looks for the maximum value of weighted Kolmogorov-Smirnov statistics for all possible sample splits. The test $T(n/T, z)$ is asymptotically distribution free and has nontrivial local power against changes in the scale parameter of the disturbances. If a trending regressor exists, T_{edf} will not be asymptotically distribution free. To circumvent these undesirable properties, Bai (1996) developed a test based on the regressor weighted empirical distribution function of residuals.

Let $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$ and

$$\mathbf{A}_n = (\mathbf{X}'\mathbf{X})^{-\frac{1}{2}} (\mathbf{X}'_n \mathbf{X}_n) (\mathbf{X}'\mathbf{X})^{-\frac{1}{2}}.$$

Bai (1996) suggested analogously $k \times 1$ vector process

$$T^*\left(\frac{n}{T}, z\right) = (\mathbf{X}'\mathbf{X})^{-\frac{1}{2}} \sum_{t=1}^n \mathbf{x}_t I(\hat{\epsilon}_t \leq z) - \mathbf{A}_n (\mathbf{X}'\mathbf{X})^{-\frac{1}{2}} \sum_{t=1}^T \mathbf{x}_t I(\hat{\epsilon}_t \leq z) \tag{4.3}$$

and the test statistic

$$T_{edf}^* = \max_{1 \leq n < T} \sup_z \left\| T^*\left(\frac{n}{T}, z\right) \right\|_\infty. \tag{4.4}$$

Bai (1996) proved that under H_0 ,

$$T\left(\frac{n}{T}, z\right) \rightarrow B(\cdot, F(\cdot)) \quad \text{and} \quad T^*\left(\frac{n}{T}, z\right) \rightarrow B^*(\cdot, F(\cdot)),$$

where $B(u, v)$ is a Gaussian process on $[0, 1]^2$ with zero mean and covariance function

$$E[B(r, u)B(s, v)] = (r \wedge s - rs)(u \wedge v - uv),$$

and $B^* = (B_1, B_2, \dots, B_k)$ is a vector of p independent two-parameter Brownian bridges on $[0, 1]^2$.

5. Numerical Results

5.1. Simulation

Before applying these monitoring procedures to real-world data, a Monte Carlo study is conducted to study size and power properties. The empirical critical values are calculated with the model (1) and the empirical sizes under $\alpha = 0.10, 0.05$ are calculated and shown in the second column of Table 5.1. The powers of the tests are computed with the number of time points $T = 100$ in 1,000 repetitions. For the moving sums estimators, the bandwidth 0.1, 0.2, 0.3, 0.4, 0.5 are used to investigate the effects of the window size. The data-generating processes considered are as follows.

(1) Simple model with $\beta = 1$ under H_0 : no change

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

(2) Single structural change model with $\beta = 1$, $\Delta = 1.0$, $\lambda = 0.3, 0.4, 0.5, 0.6, 0.7$

$$Y_i = \begin{cases} \beta x_i + \epsilon_i, & i = 1, \dots, T\lambda, \\ (\beta + \Delta)x_i + \epsilon_i, & i = T\lambda + 1, \dots, T. \end{cases}$$

(3) Double structural change model with $\beta = 1$, $\Delta = 0.5$, $\lambda_1 = 0.3$, $\lambda_2 = 0.5, 0.6, 0.7$

$$Y_i = \begin{cases} \beta x_i + \epsilon_i, & i = 1, \dots, T\lambda_1, \\ (\beta + \Delta)x_i + \epsilon_i, & i = T\lambda_1 + 1, \dots, T\lambda_2, \\ \beta x_i + \epsilon_i, & i = T\lambda_2 + 1, \dots, T. \end{cases}$$

(4) Double structural change model with $\beta = 1$, $\Delta = 0.5$, $\lambda_1 = 0.3$, $\lambda_2 = 0.5, 0.6, 0.7$

$$Y_i = \begin{cases} \beta x_i + \epsilon_i, & i = 1, \dots, T\lambda_1, \\ -\beta x_i + \epsilon_i, & i = T\lambda_1 + 1, \dots, T\lambda_2, \\ \beta x_i + \epsilon_i, & i = T\lambda_2 + 1, \dots, T. \end{cases}$$

(5) Level structural change model with $\beta = 1$, $\Delta = 1.0$, $\lambda = 0.5, 0.6, 0.7$

$$Y_i = \begin{cases} \beta x_i + \epsilon_i, & i = 1, \dots, T\lambda, \\ \beta x_i + \Delta + \epsilon_i, & i = T\lambda + 1, \dots, T. \end{cases}$$

(6) Smooth change model with $\Delta = 1.0$, $\lambda = 0.4, 0.5, 0.6$

$$Y_i = \begin{cases} \cos(3\pi x_i) + \epsilon_i, & i = 1, \dots, T\lambda, \\ \cos(3\pi x_i) + \Delta + \epsilon_i, & i = T\lambda + 1, \dots, T. \end{cases}$$

(7) Smooth and level change model with $\lambda = 0.5$

$$Y_i = \begin{cases} \cos(\pi x_i) + \epsilon_i, & i = 1, \dots, T\lambda, \\ x_i + \epsilon_i, & i = T\lambda + 1, \dots, T. \end{cases}$$

(8) Level and smooth change model with $\Delta = 1.0$, $\lambda = 0.5$

$$Y_i = \begin{cases} x_i + \epsilon_i, & i = 1, \dots, T\lambda, \\ \cos(\pi x_i) + \Delta + \epsilon_i, & i = T\lambda + 1, \dots, T, \end{cases}$$

where ϵ_i are *i.i.d.* normal errors having mean 0 and variance 1.

In this paper, we consider fixed- h tests for the sake of simplicity for the window size. For MOSUM and ME tests, the bandwidth $h = 0.1, 0.2, 0.3, 0.4, 0.5$ are considered. In tables, test r.mos.01 notes the MOSUM test with the recursive residuals and $h = 0.1$. Likewise test me.01 means the ME test with $h = 0.1$. The tests are considered on the possible range of change-points, from $\lambda = 0.15$ to 0.85.

For one structural change, the locations of change-point $\lambda = 0.3, 0.4, 0.5, 0.6, 0.7$ are considered corresponding to a shift early, midway and late in the sample period. At level $\alpha = 0.05$, the empirical powers under the null hypothesis are given in the second column of in Table 5.1. The powers for one structural change-point cases are shown in Table 5.1. Table 5.1 shows that F tests

Table 5.1. Empirical powers of tests based on $T = 100$ in 1,000 Repetitions with *i.i.d.* normal errors with $sd = 1.0$ in various one change-point models with $\Delta = 1.0$ at level $\alpha = 0.05$

Method	Model (1)	Model (2)				
	null	$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
sup F	0.036	0.049	0.085	0.144	0.218	0.347
ave F	0.046	0.058	0.079	0.133	0.303	0.462
exp F	0.041	0.053	0.081	0.140	0.262	0.406
rec.cus	0.054	0.047	0.069	0.098	0.112	0.103
ols.cus	0.035	0.049	0.074	0.126	0.250	0.441
r.mos.01	0.032	0.044	0.061	0.101	0.125	0.199
r.mos.02	0.041	0.056	0.092	0.148	0.216	0.330
r.mos.03	0.051	0.066	0.096	0.172	0.247	0.365
r.mos.04	0.044	0.069	0.089	0.180	0.269	0.272
r.mos.05	0.045	0.063	0.092	0.166	0.229	0.220
l.mos.01	0.050	0.062	0.077	0.119	0.149	0.195
l.mos.02	0.040	0.068	0.083	0.162	0.245	0.325
l.mos.03	0.040	0.054	0.081	0.163	0.283	0.390
l.mos.04	0.030	0.050	0.062	0.130	0.303	0.433
l.mos.05	0.041	0.054	0.069	0.111	0.295	0.455
me.01	0.056	0.064	0.067	0.057	0.073	0.110
me.02	0.046	0.063	0.067	0.092	0.131	0.142
me.03	0.037	0.052	0.080	0.109	0.159	0.205
me.04	0.040	0.051	0.074	0.145	0.209	0.251
me.05	0.043	0.049	0.063	0.164	0.266	0.321
rec.edf	0.042	0.058	0.055	0.088	0.106	0.059
ols.edf	0.053	0.054	0.063	0.074	0.140	0.330

and MOSUM tests have better power overall in the Model (2). The powers for two structural change-point cases are shown in Table 5.2 with the Model (3) and Model (4). For double structural change Model (3), the ME tests with the larger bandwidth perform better than F tests, CUSUM and MOSUM tests. When two break points are close, the ME tests perform better. The power of the tests increases with a longer duration of the regime since a shift rather late in the sample period means that each test has much time to pick up. With the abrupt structural change Model (5), the powers of ME tests are higher than others. Smooth structural models with one change-point are considered with the different change-point locations and the their powers are provided in Table 5.3. With the Model (6), each test gives the similar low powers due to the fluctuation of the functions. There needs a test to be developed for detecting this type of change. With the Model (7) and Model (8), MOSUM tests with the recursive residuals perform slightly better.

5.2. Applications to real data

We investigate the evidence of structural change is US real GDP from 1870 to 2009. The logarithm of the series is shown in Figure 5.1. There is a dip between 1930 and 1940 (the Great Depression) then accelerated growth during World War II. Wang and Zivot (2000) estimated two structural change-points at 1930 and 1945 with very high posterior probabilities. The test statistics are calculated as shown in Table 5.4. All the tests reject the null hypothesis with the significance level $\alpha = 0.05$. Some structural test processes are shown in Figure 5.2. They give a hint for the location of change-point. Therefore, the change-point estimators can be made based on the change test for parameter change models.

Table 5.2. Empirical powers of tests based on $T = 100$ in 1,000 Repetitions in two change-point models with $\Delta = 1.0$ at level $\alpha = 0.05$

Method	Model (3) $\lambda_1 = 0.3$			Model (4) $\lambda_1 = 0.3$		
	$\lambda_2 = 0.5$	$\lambda_2 = 0.6$	$\lambda_2 = 0.7$	$\lambda_2 = 0.5$	$\lambda_2 = 0.6$	$\lambda_2 = 0.7$
sup F	0.158	0.264	0.398	0.532	0.862	0.967
ave F	0.171	0.340	0.525	0.494	0.891	0.981
exp F	0.178	0.301	0.469	0.548	0.897	0.976
rec.cus	0.076	0.086	0.083	0.140	0.189	0.197
ols.cus	0.163	0.318	0.492	0.519	0.900	0.984
r.mos.01	0.118	0.158	0.236	0.345	0.636	0.824
r.mos.02	0.150	0.241	0.350	0.501	0.794	0.934
r.mos.03	0.171	0.280	0.417	0.517	0.844	0.971
r.mos.04	0.165	0.277	0.331	0.505	0.855	0.897
r.mos.05	0.170	0.218	0.248	0.517	0.687	0.739
l.mos.01	0.152	0.173	0.209	0.517	0.687	0.739
l.mos.02	0.212	0.330	0.383	0.688	0.891	0.944
l.mos.03	0.192	0.393	0.477	0.597	0.951	0.978
l.mos.04	0.170	0.341	0.557	0.478	0.908	0.993
l.mos.05	0.144	0.325	0.520	0.392	0.842	0.977
me.01	0.083	0.081	0.098	0.274	0.352	0.357
me.02	0.116	0.154	0.177	0.436	0.649	0.724
me.03	0.156	0.218	0.253	0.563	0.769	0.839
me.04	0.197	0.306	0.307	0.639	0.885	0.910
me.05	0.209	0.360	0.417	0.639	0.924	0.967
rec.edf	0.068	0.068	0.071	0.147	0.166	0.204
ols.edf	0.083	0.168	0.354	0.211	0.547	0.912

Table 5.3. Empirical powers of tests based on $T = 100$ in 1,000 Repetitions in various structural change models with $\Delta = 1.0$ at level $\alpha = 0.05$

Method	Model (5)			Model (6)			Model (7)	Model (8)
	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.5$	$\lambda = 0.5$
sup F	0.338	0.355	0.437	0.099	0.090	0.078	0.405	0.498
ave F	0.229	0.314	0.534	0.096	0.068	0.064	0.554	0.626
exp F	0.338	0.368	0.501	0.096	0.085	0.079	0.489	0.589
rec.cus	0.135	0.114	0.145	0.060	0.061	0.068	0.162	0.257
ols.cus	0.318	0.388	0.558	0.084	0.065	0.068	0.450	0.535
r.mos.01	0.240	0.271	0.320	0.062	0.059	0.065	0.203	0.269
r.mos.02	0.323	0.357	0.441	0.088	0.075	0.063	0.326	0.455
r.mos.03	0.294	0.346	0.470	0.081	0.073	0.066	0.399	0.557
r.mos.04	0.260	0.319	0.340	0.084	0.080	0.081	0.459	0.588
r.mos.05	0.236	0.276	0.293	0.084	0.081	0.072	0.468	0.626
l.mos.01	0.226	0.243	0.260	0.052	0.071	0.058	0.259	0.304
l.mos.02	0.354	0.362	0.440	0.084	0.081	0.081	0.400	0.492
l.mos.03	0.381	0.399	0.483	0.104	0.068	0.081	0.468	0.580
l.mos.04	0.290	0.370	0.538	0.097	0.058	0.073	0.531	0.599
l.mos.05	0.111	0.324	0.555	0.082	0.072	0.058	0.510	0.560
me.01	0.129	0.121	0.143	0.038	0.063	0.051	0.126	0.165
me.02	0.190	0.198	0.228	0.056	0.057	0.047	0.216	0.242
me.03	0.267	0.265	0.297	0.061	0.058	0.064	0.289	0.339
me.04	0.328	0.298	0.348	0.079	0.062	0.056	0.370	0.461
me.05	0.347	0.368	0.404	0.082	0.072	0.064	0.495	0.568
rec.edf	0.239	0.186	0.080	0.101	0.075	0.079	0.068	0.145
ols.edf	0.121	0.117	0.322	0.085	0.058	0.057	0.316	0.302

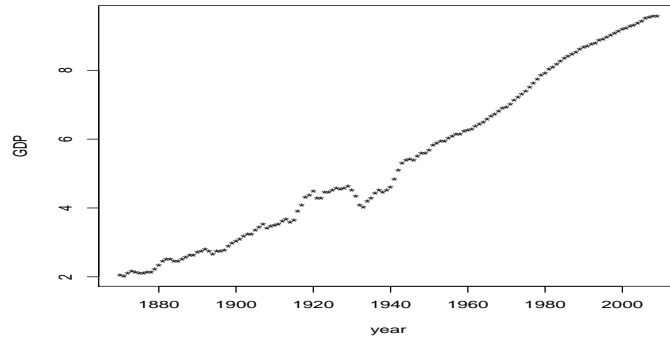


Figure 5.1. Scatter plot of US GDP data from 1870 to 2009

Table 5.4. Test statistics for structural change with the US GDP data

CUSUM	F		rec		ols		ME	stat	EDF	stat
	stat		MOSUM	stat	MOSUM	stat				
supF	594.920		r.mos.01	2.697	l.mos.01	2.313	me.01	2.324	rec.edf	0.982
aveF	251.182		r.mos.02	4.937	l.mos.02	3.190	me.02	3.521	ols.edf	2.224
expF	292.866		r.mos.03	6.552	l.mos.03	4.018	me.03	4.442		
rec.cus	3.239		r.mos.04	7.959	l.mos.04	4.165	me.04	4.984		
ols.cus	2.511		r.mos.05	9.201	l.mos.05	4.018	me.05	5.565		

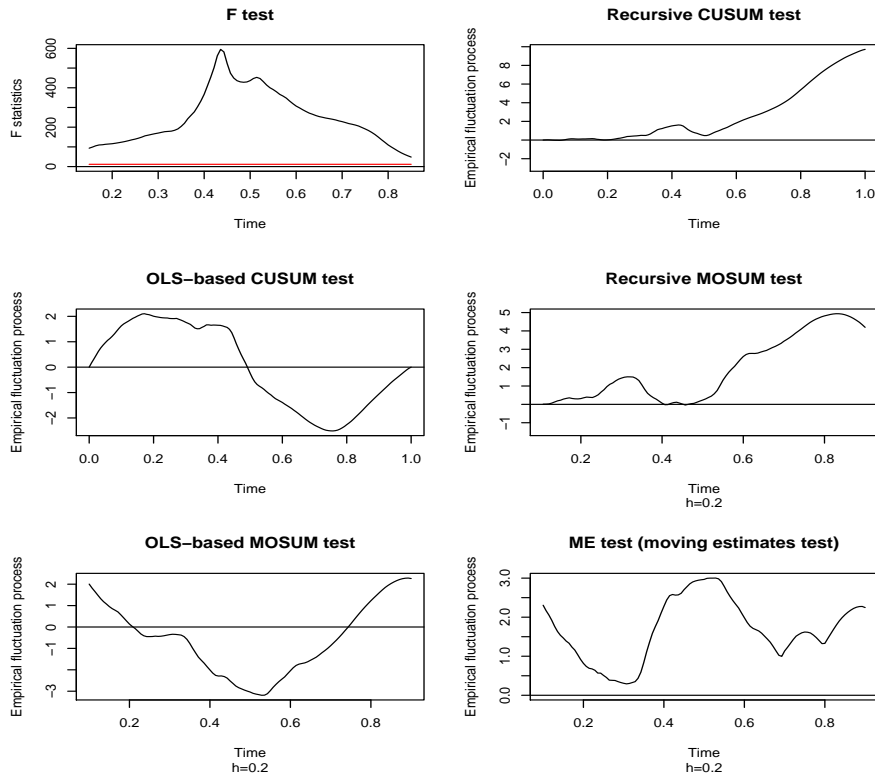


Figure 5.2. Structural test processes for US GDP data

6. Concluding Remarks

We have considered tests for structural change in the regression parameters. Our proposed test based on local Fourier estimators is generally applicable to detect changes in regression models. No version of the structural change tests considered is uniformly superior to the other. The CUSUM and MOSUM tests are considered and compared with the proposed test. In power comparison, tests appear to have nontrivial power against general alternatives of parameter nonconstancy because rejecting the null hypothesis does not necessarily imply one-time structural change. We applied our method to determine the empirical evidence for structural changes in US GDP data. There are significant changes in the data and the next step must be to estimate the location where structural change occurs.

Structural stability is of importance in regression models. A form of model misspecification as parameter nonconstancy, may have severe consequences on inference if left undetected. In particular, if the data generating process changes in some ways, then forecasts lose accuracy. In consequence, many econometricians routinely apply tests for parameter change. Significant effort has been devoted to obtain convenient and powerful tests in a variety of modeling contexts. We provided a comparison of some historical tests for structural change.

Our simulations show that different tests have quite different behavior under various alternatives and that no test is superior.

The three F tests - $\text{Sup}F$, $\text{Avg}F$, and $\text{Exp}F$ - are all tests of the same null hypothesis but differ in their choice of alternative hypothesis. Even though they have power in similar directions, the choice of the test may be made on a particular application for the purpose of the test. If the desire is to discover whether there was a swift shift in the regime, then the $\text{Sup}F$ test is appropriate. The CUSUM test can do well based on OLS residuals. It reacts to structural shifts that occur late in the sample and are likely to go unnoticed by the other F tests.

Since power depends on the bandwidth, the sample size, the magnitude of change, the location of the change-points, and the standard deviation of disturbances, the knowledge about the properties of these historical tests and connections between these tests can be used in practical structural change tests and in extension to other change tests.

The assumption that the disturbances are independent is restrictive and this assumption can be weakened to linear processes such as $\text{AR}(p)$. Further research is expected with a change test for more change types and dependent errors. Kim and Hart (1998) considered the change test for dependent data and Kim and Hart (2011) proposed a change-point estimator that can be used for independent data and extensively used for dependent data. Their tests and estimators can be further studied for linear structural change and time series structural change.

References

- Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point, *Econometrica*, **61**, 821–856.
- Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, **62**, 1383–1414.
- Aue, A., Horváth, L., Hušková, M. and Kokoszka, P. (2006). Change-point monitoring in linear models, *Econometric Journal*, **9**, 373–403.
- Bai, J. (1996). Testing for parameter constancy in n linear regressions: An empirical distribution function approach, *Econometrica*, **64**, 597–622.

- Bauer, P. and Hackl, P. (1978). The use of MOSUMs for quality control, *Technometrics*, **20**, 431–436.
- Bauer, P. and Hackl, P. (1980). An extension of the MOSUM technique for quality control, *Technometrics*, **22**, 1–7.
- Brown, L. R., Durbin, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time, *Journal of Royal Statistical Society B*, **37**, 149–192.
- Carlstein, E. (1988). Nonparametric change point estimation, *Annals of Statistics*, **16**, 188–197.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions, *Econometrica*, **28**, 591–604.
- Chu, C. J., Hornik, K. and Kuan, C. (1995a). MOSUM tests for parameter constancy, *Biometrika*, **82**, 603–617.
- Chu, C. J., Hornik, K. and Kuan, C. (1995b). The moving-estimates test for parameter stability, *Econometric Theory*, **11**, 699–720. *Biometrika*, **65**, 243–251.
- Csörgö, M. and Horváth, L. (1987). Nonparametric Tests for the Change-point Problem, *Journal of Statistical Planning and Inference*, **17**, 1–9.
- Csörgö, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*, Wiley, New York.
- Dümbgen, L. (1991). The asymptotic behavior of some nonparametric change point estimators, *Annals of Statistics*, **19**, 1471–1495.
- Hansen, B. E. (1992). Tests for parameter instability in regressions with $I(1)$ processes, *Journal of Business & Econometric Statistics*, **10**, 321–335.
- Hansen, B. E. (1997). Approximate asymptotic P values for structural-change tests, *Journal of Business & Econometric Statistics*, **15**, 60–67.
- Horváth, L., Hušková, M., Kokoszka, P. and Steinebach, J. (2004). Monitoring changes in linear models, *Journal of Statistical Planning and Inference*, **126**, 225–251.
- Kim, J. and Hart, J. D. (1998). Tests for change in a mean function when the data are dependent, *Journal of Time Series Analysis*, **19**, 399–424.
- Kim, J. and Hart, J. D. (2011). A change-point estimator using local fourier series, *Journal of Nonparametric Statistics*, **23**, 83–98.
- Kuan, C. M. and Hornik, K. (1995). The generalized fluctuation test: A unifying view, *Econometric Reviews*, **14**, 135–161.
- Loynes, R. M. (1980). The empirical distribution function of residuals from generalized regression, *Annals of Statistics*, **8**, 285–298.
- Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, **55**, 703–708.
- Ploberger, W. and Krämer, W. (1992). The CUSUM test with OLS residuals, *Econometrica*, **60**, 271–285.
- Rice, J. (1984). Bandwidth choice for nonparametric regression, *Annals of statistics*, **12**, 1215–1230.
- Silva, E. G. and Teixeira, A. A. C. (2008). Surveying structural change: Seminal contributions and a bibliometric account, *Structural Change and Economic Dynamics*, **19**, 273–300.
- Wang, J. and Zivot, E. (2000). A bayesian time series model of multiple structural changes in level, trend and variance, *Journal of Business & Economic Statistics*, **18**, 374–386.
- Zeileis, A. (2005). A unified approach to structural change tests based on ML scores, F statistics, and OLS residuals, *Econometric Reviews*, **23**, 445–466.
- Zeileis, A., Leisch, L., Hansen, B., Hornik, K. and Kleiber, C. (2007). The strucchange Package, R manual.