Power Properties of the Modified CUSUM Tests

Peiyun Jiang   Eiji Kurozumi

August 2017

Power Properties of the Modified CUSUM Tests [1]

Peiyun Jiang   Eiji Kurozumi

Department of Economics, Hitotsubashi University

August, 2017

Abstract

The CUSUM test has played an important role in theory and applications related to structural change, but its drawback is that it loses power when the break is orthogonal to the mean of the regressors. In this study, we consider two modified CUSUM tests that have been proposed, implicitly or explicitly, in the literature to detect such structural changes, and investigate the limiting power properties of these tests under a fixed alternative. We demonstrate that the modified tests are superior to the classic tests in terms of both asymptotic theory and in finite samples, when detecting the orthogonal structural shift.

1

1. Introduction

In this study we investigate the power properties of the modified CUSUM test. The original CUSUM test was introduced by Brown, Durbin, and Evans (1975), and has been widely used to test for parameter stability in practical analyses. It has also been investigated extensively and extended in various ways in the literature. With regard to the power properties, Garbade (1977) studied the finite sample performance of the CUSUM test under three patterns of coefficient variations. Then, based on Monte Carlo experiments, the results showed that the CUSUM test is quite weak in detecting parameter instability under the simulation settings. While this property has been repeatedly observed in the literature, Ploberger and Krämer (1990) conducted a theoretical investigation of the power of the test, where changes in the parameters are local to zero under the alternative. They showed that the limiting distribution of the test under the local alternative is expressed as a Brownian motion plus an additional term related to the interaction between the mean of the regressors and the direction of the structural change. This result implies that in the case of a simple shift in parameters, the power of the CUSUM test depends on the angle between the mean of the regressors and the direction of the shift, and the test loses power when this angle is perpendicular. Their result theoretically explains the poor performance of the CUSUM test in the study of Garbade (1977), in which the mean of the regressors is set to be orthogonal to the shift (the mean is equal to zero). While the original CUSUM test was proposed using recursive residuals, Ploberger and Krämer (1992) developed a CUSUM test based on ordinary least squares (OLS) residuals, and compared the local power of this test to that of the original test. On the other hand, Deng and Perron (2008) suggested investigating the power properties of both versions of the CUSUM test, not under a local alternative, but from a non-local perspective. They derived the limiting distribution of the test statistic under the fixed alternative, and confirmed that even in this case, the power of the test depends on the angle between the mean of the regressors and the direction of the change.

While these undesirable power properties of the CUSUM test have been noted in the literature, several modified versions of the test have been proposed, explicitly or implicitly, in order to overcome the problem. For example, Luger (2001) introduced a test statistic based

on the symmetrization of the absolute value of the recursive residuals. This modified test performs better than the original test does in terms of power when the angle between the mean of the regressor and the shift is perpendicular, although the original CUSUM test performs better when this angle decreases. Hušková and Koubková (2005) considered a quadratic form of the product of the regressors and the residuals for monitoring tests, while Xia, Guo and Zhao (2011) studied the model with endogenous regressors, and proposed a CUSUM test based on the weighted residuals from the GMM estimation. These studies concentrate mainly on the reactions of the tests to the location change and to the magnitude of the break, but do not analyze the impact of the angle on the power of the test. Further studies are still necessary to investigate the performance of this method in terms of detecting an orthogonal structural change.

Therefore, in this study, we investigate two versions of the CUSUM test modified to avoid losing power when the mean of the regressor is orthogonal to the shift. Our modified tests are based on the CUSUM of $x_{jt}\breve{u}_t$ instead of $\breve{u}_t$, where x_{jt} denotes the jth element of regressor x_t, and $\breve{u}_t$ denotes the OLS residuals or recursive residuals. The asymptotic distributions of the test statistics are investigated under the null hypothesis of parameter stability, as well as under the fixed alternative, and the powers of the modified tests are investigated. The finite sample properties of the modified tests are supported by the asymptotic theory, and we confirm that the modified tests are superior to the classic test, in terms of both the asymptotic theory and in finite samples, when detecting an orthogonal structural shift.

The remainder of the paper is organized as follows. Section 2 introduces the model and assumptions. Section 3 presents the asymptotic behaviors of the two modified tests considered in this study. We further extend the modified tests to models with serially correlated errors. The finite sample properties are investigated using Monte Carlo simulations in Section 4. Section 5 concludes the paper. The mathematical proofs are relegated to the Appendix.

2. Models and Assumptions

We consider the standard linear regression model given by

$$y_t = x_t'\beta_t + u_t \qquad (t = 1, 2, \cdots, T), \tag{1}$$

3

where $x_t = [x_{1t}, x_{2t}, \cdots, x_{kt}]'$ is a k-dimensional regressor, u_t is an unobservable stochastic disturbance, and β_t is a k-dimensional vector of coefficients. Since a constant term is typically included in a model, the first element of the regressor, x_{1t}, is unity for all t. We consider the testing problem given by

$$H_0: \ \beta_t = \beta \ \forall t \quad \text{vs.} \quad H_1: \ \beta_t = \beta + \delta 1_{(t > [T\lambda])},$$

where $\lambda \in (0, 1)$ represents the break fraction and $1_{(t > [T\lambda])}$ is an indicator function that takes the value one if $t > [T\lambda]$, and zero otherwise. Then, the parameters in (1) are stable under the null hypothesis, whereas we allow for a one-time change in the parameters under the alternative.

To investigate the asymptotic properties of the CUSUM test, we make the following assumptions:

Assumption 1 *The regressor x_t and the error term u_t are defined on a common probability space, and the following condition holds:*

$$\lim_{T \to \infty} \sup \frac{1}{T} \sum_{t=1}^{T} \| x_t \|^{2+\delta} < \infty, \quad \text{a.s. for some} \quad \delta > 0.$$

Assumption 2 *The following probability limits exist:*

$$p \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x_t = E[x_t] = c_1,$$

$$p \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x_t x_t' = E[x_t x_t'] = C,$$

$$p \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x_t x_t' \otimes x_t x_t' = E[x_t x_t' \otimes x_t x_t'] = \Lambda,$$

where c_1 is a $k \times 1$ vector and C and Λ are $k \times k$ and $k^2 \times k^2$ nonsingular and nonstochastic matrices, respectively.

We need Assumption 2 to investigate the power of the modified CUSUM tests. We denote

rows of C by c'_j for $i = 1, \cdots, k$. That is,

$$C = E[x_t x'_t] = E \begin{bmatrix} x_{1t} x'_t \\ x_{2t} x'_t \\ \vdots \\ x_{kt} x'_t \end{bmatrix} = \begin{bmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_k \end{bmatrix}.$$

The vector c_1 is called the mean regressor.

Assumption 3 *The disturbances $\{u_t\}$ are stationary and ergodic with*

$$E[u_t | U_{t-1}] = 0, \quad E[u_t^2 | U_{t-1}] = \sigma^2, \quad E[u_t^4] < \infty,$$

where U_{t-1} is the σ-field generated by $\{x_t, u_{t-1}, x_{t-1}, u_{t-2}, \cdots\}$.

Assumption 3 implies that the error term is a martingale difference sequence. We first proceed with this assumption, but it will be relaxed in a later section in order to investigate the effect of the serial correlation on the power of the tests.

3. The Modified CUSUM Tests

3.1 Preliminaries to the modified CUSUM tests

We first consider the standard CUSUM test in order to motivate the modification. The test statistic based on the OLS residuals is given by

$$CUSUM^{ols} = \sup_{0 \leq r \leq 1} \left| \frac{1}{\hat{\sigma} \sqrt{T}} \sum_{t=1}^{[Tr]} \hat{u}_t \right|,$$

where $\hat{u}_t$ are the OLS residuals and $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2$, while the recursive residuals based test statistic is

$$CUSUM^{rec} = \sup_{0 \leq r \leq 1} \left| \frac{\sum_{t=k+1}^{[Tr]} \tilde{u}_t}{\tilde{\sigma} \sqrt{T-k}} \right| \bigg/ \left(1 + 2 \frac{[Tr] - k}{T - k} \right),$$

where $\tilde{u}_t = (y_t - x'_t \hat{\beta}_{t-1})/f_t$ for $t = k+1, \cdots, T$ are the recursive residuals, $\hat{\beta}_t = (X'_t X_t)^{-1} X'_t Y_t$ with $X_t = [x'_1, x'_2, \cdots, x'_t]'$ and $Y_t = [y_1, y_2, \cdots, y_t]'$, $f_t = (1 + x'_t (X'_{t-1} X_{t-1})^{-1} x_t)^{1/2}$, and $\tilde{\sigma}^2 = (T-k)^{-1} \sum_{t=k+1}^{T} (\tilde{u}_t - \bar{\tilde{u}})^2$ with $\bar{\tilde{u}} = (T-k)^{-1} \sum_{t=k+1}^{T} \tilde{u}_t$.

Ploberger and Krämer (1990, 1992) derived the limiting distributions of these test statistics under the local alternative, while Deng and Perron (2008) investigated the asymptotic properties of these statistics under the fixed alternative. Their results imply that the powers of these tests depend on the angle between the mean regressor and the direction of the break. To explain this dependence, we demonstrate that the power of the OLS-based test depends on $c_1'\delta$ by focusing on the fixed alternative. Given that the OLS estimator of β can be expressed as

$$\hat{\beta} = \beta + \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \sum_{t=1}^{T} \left(x_t x_t' \delta 1_{(t>[T\lambda])} + x_t u_t\right), \tag{2}$$

the OLS residuals are given by

$$\hat{u}_t = u_t + x_t'\delta 1_{(t>[T\lambda])} - x_t'(\hat{\beta} - \beta), \tag{3}$$

and thus

$$\frac{1}{T}\sum_{t=1}^{[Tr]} \hat{u}_t = \frac{1}{T}\sum_{t=1}^{[Tr]} u_t + \frac{1}{T}\sum_{t=1}^{[Tr]} x_t'\delta 1_{(t>[T\lambda])}$$
$$- \frac{1}{T}\sum_{t=1}^{[Tr]} x_t' \left(\frac{1}{T}\sum_{s=1}^{T} x_s x_s'\right)^{-1} \left[\frac{1}{T}\sum_{s=1}^{T} x_s x_s'\delta 1_{(s>[T\lambda])} + \frac{1}{T}\sum_{s=1}^{T} x_s u_s\right]. \tag{4}$$

It can be shown that for $r > \lambda$, the second and third terms on the right hand side of (4) converge in probability to $(r-\lambda)c_1'\delta$ and $-r(1-\lambda)c_1'\delta$, respectively. Therefore, we can see that the OLS-based CUSUM test loses power when $c_1'\delta = 0$.

To avoid the dependence of the power on $c_1'\delta$, we modify the CUSUM test such that it is not based on the residuals, but on the product of x_{jt} (for $j \neq 1$) and the residuals. Let $\hat{w}_{jt} = x_{jt}\hat{u}_t$ and $\tilde{w}_{jt} = x_{jt}\tilde{u}_t$. Then, the modified CUSUM test statistics are defined as

$$CUSUM_m^{ols} = \sup_{0 \leq r \leq 1} \left|\frac{1}{\hat{\sigma}_j\sqrt{T}}\sum_{t=1}^{[Tr]} \hat{w}_{jt}\right|,$$

$$CUSUM_m^{rec} = \sup_{0 \leq r \leq 1} \left|\frac{\sum_{t=k+1}^{[Tr]} \tilde{w}_{jt}}{\tilde{\sigma}_j\sqrt{T-k}}\right| \bigg/ \left(1 + 2\frac{[Tr]-k}{T-k}\right),$$

where

$$\hat{\sigma}_j^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{w}_{jt}^2 \quad \text{and} \quad \tilde{\sigma}_j^2 = \frac{1}{T-k}\sum_{t=k+1}^{T} (\tilde{w}_{jt} - \bar{\tilde{w}}_j)^2 \quad \text{with} \quad \bar{\tilde{w}}_j = \frac{1}{T-k}\sum_{t=k+1}^{T} \tilde{w}_{jt}.$$

6

Proposition 1 *Suppose that Assumptions 1–3 hold.*

(a) Under the null hypothesis,

$$CUSUM_m^{ols} \quad \Rightarrow \quad \sup_{0 \le r \le 1} |BB_j(r)|, \tag{5}$$

$$CUSUM_m^{rec} \quad \Rightarrow \quad \sup_{0 \le r \le 1} \left| \frac{W_j(r)}{1 + 2r} \right|, \tag{6}$$

where $BB_j(r)$ and $W_j(r)$ are the one-dimensional standard Brownian bridge and Brownian motion, respectively, and $\Rightarrow$ denotes weak convergence of the associated probability measures.

(b) Under the alternative hypothesis,

$$\frac{1}{\sqrt{T}} CUSUM_m^{ols} \quad \xrightarrow{p} \quad \frac{|c_j'\delta|\lambda(1-\lambda)}{\sqrt{\sigma^2 c_{jj} + \lambda(1-\lambda)\delta'\Lambda_{jj,0}\delta}}, \tag{7}$$

$$\frac{1}{\sqrt{T}} CUSUM_m^{rec} \quad \xrightarrow{p} \quad \frac{|c_j'\delta|q}{\sqrt{\sigma^2 c_{jj} + \lambda(1-\lambda)\delta'\Lambda_{jj,0}\delta - (c_j'\delta\lambda \log(\lambda))^2}}, \tag{8}$$

where c_{jj} and $\Lambda_{jj,0}$ are the (j,j) element of C and the (j,j) block of Λ, respectively; that is,

$$\frac{1}{T}\sum_{t=1}^{T} x_{jt}^2 \xrightarrow{p} c_{jj}, \quad and \quad \frac{1}{T}\sum_{t=1}^{T} x_{jt}^2 x_t x_t' \xrightarrow{p} \Lambda_{jj,0},$$

$$and \quad q = \sup_{0 \le r \le 1} \frac{\lambda \log \frac{r}{\lambda} 1_{(r>\lambda)}}{1 + 2r} = \begin{cases} \dfrac{\lambda \log \frac{\lambda^*}{\lambda}}{1 + 2\lambda^*} & 0 \le \lambda < e^{-\frac{3}{2}} \\[2mm] \dfrac{-\lambda \log \lambda}{3} & e^{-\frac{3}{2}} \le \lambda \le 1, \end{cases} \tag{9}$$

with $\lambda^* = \{\lambda^* : 0 \le \lambda^* \le 1 \text{ and } \log \lambda^* = 1 + \log \lambda + \frac{1}{2\lambda^*}\}$.

Proposition 1 shows that our modification could work well, even in the case of $c_1'\delta = 0$, for $c_j'\delta \ne 0$ and $j \ne 1$. Thus, we can avoid the loss of power caused by the orthogonal change. However, it is obvious that the modified test loses power if $c_j'\delta = 0$, and we do not know whether $c_j'\delta = 0$, for some $j = 1, \cdots, k$. We will discuss how to overcome this problem in the following subsection.

3.2 Modified CUSUM tests

Note that $C = E[x_t x_t']$ is positive definite, by Assumption 2. Therefore, we can easily see that $c_j'\delta \ne 0$ for at least one of $j = 1, \cdots, k$ if $\delta \ne 0$. Thus, it is natural to construct the test

statistics based on all $\hat{w}_{1t}, \cdots, \hat{w}_{kt}$ or $\tilde{w}_{1t}, \cdots, \tilde{w}_{kt}$ in order to avoid the potential loss of power caused by $c_j' \delta = 0$ for some j. One of the possible transformations used in the literature is to construct a quadratic form based on $x_t \hat{u}_t = [\hat{w}_{1t}, \cdots, \hat{w}_{kt}]'$ or $x_t \tilde{u}_t = [\tilde{w}_{1t}, \cdots, \tilde{w}_{kt}]'$ given by

$$Q^{ols} = \sup_{0 \le r \le 1} Q^{ols}(r) \quad \text{where} \quad Q^{ols}(r) = \frac{\left(\sum_{t=1}^{[Tr]} x_t \hat{u}_t\right)' \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{[Tr]} x_t \hat{u}_t\right)}{\hat{\sigma}^2},$$

$$Q^{rec} = \sup_{0 \le r \le 1} Q^{rec}(r) \quad \text{where} \quad Q^{rec}(r) = \frac{\left(\sum_{t=k+1}^{[Tr]} x_t \tilde{u}_t\right)' \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=k+1}^{[Tr]} x_t \tilde{u}_t\right)}{\tilde{\sigma}^2},$$

where $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ are defined as before. The following theorem gives the asymptotic properties of these test statistics.

Theorem 1 *Suppose that Assumptions 1-3 hold.*

(a) Under the null hypothesis,

$$Q^{ols} \quad \Rightarrow \quad \sup_{0 \le r \le 1} \|BB(r)\|^2, \tag{10}$$

$$Q^{rec} \quad \Rightarrow \quad \sup_{0 \le r \le 1} \|W(r)\|^2, \tag{11}$$

where $BB(r)$ and $W(r)$ are the k-dimensional standard Brownian bridge and Brownian motion, respectively.

(b) Under the alternative hypothesis,

$$\frac{1}{T} Q^{ols} \quad \xrightarrow{p} \quad \frac{\delta' C \delta \lambda^2 (1 - \lambda)^2}{\sigma^2 + \lambda(1 - \lambda)\delta' C \delta}, \tag{12}$$

$$\frac{1}{T} Q^{rec} \quad \xrightarrow{p} \quad \frac{\delta' C \delta (\lambda \log \lambda)^2}{\sigma^2 + \lambda(1 - \lambda)\delta' C \delta - (c_1' \delta \lambda \log(\lambda))^2}. \tag{13}$$

Theorem 1 clearly shows that the modified test statistics Q^{ols} and Q^{rec} can avoid the loss of power caused by $c_j' \delta = 0$ for some j, and are consistent because $\delta' C \delta \ne 0$.

The other possible transformation is to take the maximum of the absolute values of the elements of $x_t \hat{u}_t$ or $x_t \tilde{u}_t$. Let

$$\left[M_1^{ols}(r), \cdots, M_k^{ols}(r)\right]' = \frac{1}{\hat{\sigma}} \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1/2} \sum_{t=1}^{[Tr]} x_t \hat{u}_t,$$

$$\left[M_1^{rec}(r), \cdots, M_k^{rec}(r)\right]' = \frac{1}{\tilde{\sigma}} \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1/2} \sum_{t=k+1}^{[Tr]} x_t \tilde{u}_t,$$

8

and consider the following test statistics:

$$M^{ols} = \max \left\{ \sup_{0 \le r \le 1} \left| M_1^{ols}(r) \right|, \cdots, \sup_{0 \le r \le 1} \left| M_k^{ols}(r) \right| \right\}$$

$$M^{rec} = \max \left\{ \sup_{0 \le r \le 1} \left| M_1^{rec}(r) \right|, \cdots, \sup_{0 \le r \le 1} \left| M_k^{rec}(r) \right| \right\}.$$

Theorem 2 *Suppose that Assumptions 1–3 hold.*

(a) Under the null hypothesis,

$$M^{ols} \Rightarrow \max \left\{ \sup_{0 \le r \le 1} |BB_1(r)|, \cdots, \sup_{0 \le r \le 1} |BB_k(r)| \right\}, \tag{14}$$

$$M^{rec} \Rightarrow \max \left\{ \sup_{0 \le r \le 1} |W_1(r)|, , \cdots, \sup_{0 \le r \le 1} |W_k(r)| \right\}, \tag{15}$$

where $\{BB_j(r)\}$ and $\{W_j(r)\}$ for $j = 1, \cdots, k$ are the independent one-dimensional standard Brownian bridge and Brownian motion, respectively.

(b) Under the alternative hypothesis,

$$\frac{1}{\sqrt{T}} M^{ols} \xrightarrow{p} \frac{\max\{|v_1|, \cdots, |v_k|\} \lambda (1-\lambda)}{\sqrt{\sigma^2 + \lambda(1-\lambda)\delta' C\delta}}, \tag{16}$$

$$\frac{1}{\sqrt{T}} M^{rec} \xrightarrow{p} \frac{\max\{|v_1|, \cdots, |v_k|\}(-\lambda \log \lambda)}{\sqrt{\sigma^2 + \lambda(1-\lambda)\delta' C\delta - (c_1' \delta \lambda \log(\lambda))^2}}, \tag{17}$$

where v_j is the k-th element of $C^{1/2}\delta$ for $j = 1, \cdots, k$; that is, $[v_1, \cdots, v_k]' = C^{1/2}\delta$.

Again, we can see from Theorem 2 that the maximum-type tests, M^{ols} and M^{rec}, are consistent irrespective of whether $c_j'\delta = 0$ for some j.

The critical values of the null-limiting distributions of the quadratic-type and maximum-type tests are obtained by approximating a standard Brownian motion using 2000 independent normal random variables with 1000000 replications (see Panel (a) of Table 1). In addition, it is sometimes the case that the test statistics are constructed by removing the first and last 100ε % observations. In this case, we have

$$\sup_{\varepsilon \le r \le 1-\varepsilon} Q^{ols}(r) \Rightarrow \sup_{\varepsilon \le r \le 1-\varepsilon} \|BB(r)\|^2, \quad \sup_{\varepsilon \le r \le 1-\varepsilon} Q^{rec}(r) \Rightarrow \sup_{\varepsilon \le r \le 1-\varepsilon} \|W(r)\|^2, \tag{18}$$

$$\max_{1 \le j \le k} \left\{ \sup_{\varepsilon \le r \le 1-\varepsilon} \left| M_j^{ols}(r) \right|, \ j = 1, \cdots, k \right\} \Rightarrow \max_{1 \le j \le k} \left\{ \sup_{\varepsilon \le r \le 1-\varepsilon} |BB_j(r)|, \ j = 1, \cdots, k \right\}, \tag{19}$$

9

$$\max_{1\leq j\leq k}\left\{\sup_{\varepsilon\leq r\leq 1-\varepsilon}\left|M_j^{rec}(r)\right|,\ j=1,\cdots,k\right\}\Rightarrow\max_{1\leq j\leq k}\left\{\sup_{\varepsilon\leq r\leq 1-\varepsilon}|W_j(r)|,\ j=1,\cdots,k\right\}. \quad (20)$$

The critical values for these distributions with $\varepsilon=0.15$ are given in Panel (b) of Table 1.

We next investigate the limiting properties of the tests under the alternative based on Theorems 1 and 2. Since the probability limits under the alternative depend on several parameters in the model, we focus on a simple case where $\sigma^2=1$ and $x_t=[1,z_t]'$ with $z_t\sim i.i.d.N(1,1)$. Furthermore, the change in the coefficients is specified by $\delta=[1,-1]'$ and $\delta=[1,1]'$, which correspond to $c_1'\delta=0$ and $c_1'\delta\neq 0$, respectively. Given that the power properties depend not only on the probability limits under the alternative, but also on the critical values used in the tests, we compare these limits divided by the asymptotic 5% critical values.

Figures 1(a) and (b) show the probability limits of the quadratic-type tests given by (12) and (13), respectively, divided by the corresponding critical values. We can see that the limit of the OLS-based version is maximized at the midpoint, whereas that of the recursive-based version is skewed to the right. As expected from the power analysis of Ploberger and Krämer (1990) and Deng and Perron (2008), Q^{rec} is more powerful than Q^{ols} when the break occurs early in the sample, whereas the reserved relation is observed when λ is closer to one. A similar tendency is observed for the maximum-type test, as shown in Figures 1(c) and (d). The two types of tests, the quadratic-type and maximum-type tests based on OLS, are compared in Figures 1(e) and (f). However, neither version is uniformly superior to the other, because the powers depend on many factors, such as the number of regressors k and $\delta'C\delta$, among others. For instance, in the case of $k=2$, the quadratic-type test outperforms the maximum-type test under our setting, as shown in Figure 1(e). However, Figure 1(f) implies that the latter performs better in the case of $k=3$.

3.3. Modified CUSUM tests with serially correlated errors

In practice, it is sometimes the case that the error term is not a martingale difference sequence, but instead is serially correlated. As shown by, for example, Tang and MacNeill (1993), the serial correlation in the error term can produce striking effects on the distribution. Therefore, when the error term is possibly serially correlated, we need to construct the test

statistics taking the serial correlation into account. In this case, we replace Assumption 3 with the following assumption.

Assumption 4 *The following functional central limit theorem holds:*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} x_t u_t \Rightarrow \Omega^{1/2} W(r)$$

uniformly for $0 \leq r \leq 1$, *where* $\Omega = \sum_{p=-\infty}^{\infty} \Gamma_p$ *with* $\Gamma_p = Cov(x_t u_t, x_{t-p} u_{t-p})$.

The test statistics, Q^{ols}, Q^{rec}, M^{ols}, and M^{rec} are defined as before with

$$Q^{ols}(r) = \frac{1}{T} \left(\sum_{t=1}^{[Tr]} x_t \hat{u}_t \right)' \hat{\Omega}^{-1} \left(\sum_{t=1}^{[Tr]} x_t \hat{u}_t \right), \tag{21}$$

$$Q^{rec}(r) = \frac{1}{T} \left(\sum_{t=k+1}^{[Tr]} x_t \tilde{u}_t \right)' \tilde{\Omega}^{-1} \left(\sum_{t=k+1}^{[Tr]} x_t \tilde{u}_t \right), \tag{22}$$

$$\left[M_1^{ols}(r), \cdots, M_k^{ols}(r) \right]' = \frac{1}{\sqrt{T}} \hat{\Omega}^{-1/2} \sum_{t=1}^{[Tr]} x_t \hat{u}_t, \tag{23}$$

$$[M_1^{rec}(r), \cdots, M_k^{rec}(r)]' = \frac{1}{\sqrt{T}} \tilde{\Omega}^{-1/2} \sum_{t=k+1}^{[Tr]} x_t \tilde{u}_t, \tag{24}$$

where $\hat{\Omega}$ and $\tilde{\Omega}$ are the consistent estimators of Ω based on $x_t \hat{u}_t$ and $x_t \tilde{u}_t$, respectively. In practice, it is often the case that Ω is estimated nonparametrically, such that

$$\hat{\Omega} = \hat{\Gamma}_0 + \sum_{p=1}^{m} k(p,m) \left(\hat{\Gamma}_p + \hat{\Gamma}_p' \right) \quad \text{where} \quad \hat{\Gamma}_p = \frac{1}{T} \sum_{t=p+1}^{T} x_t x_{t-p}' \hat{u}_t \hat{u}_{t-p},$$

$k(p,m) = 1 - p/(m+1)$ is the Bartlett kernel, and the bandwidth m is selected based on Andrews (1991) such that

$$m = [1.1447 \times (a(\delta)T)^{1/3}] \quad \text{where} \quad a(\delta) = \frac{\sum_{j=1}^{k} 4\hat{\rho}_j^2 \hat{\sigma}_j^4 / [(1-\hat{\rho}_j)^6 (1+\hat{\rho}_j)^2]}{\sum_{j=1}^{k} \hat{\sigma}_j^4 / (1-\hat{\rho}_j)^4}$$

with $\hat{\rho}_j$ obtained by regressing $\hat{w}_{jt}$ on $\hat{w}_{jt-1}$, and $\hat{\sigma}_j^2$ defined as before. Then, $\tilde{\Omega}$ is defined similarly using the recursive residuals.

Let us define $\gamma_{jj,p}$ and $\Lambda_{jj,p}$ as the probability limits of

$$\frac{1}{T}\sum_{t=p+1}^{T} x_{jt}x_{jt-p}u_t u_{t-p} \xrightarrow{p} \gamma_{jj,p} \quad \text{and} \quad \frac{1}{T}\sum_{t=p+1}^{T} x_{jt}x_{jt-p}x_t x'_{t-p} \xrightarrow{p} \Lambda_{jj,p}.$$

Theorem 3 *Suppose that Assumptions 1, 2, and 4 hold and that the quadratic-type and maximum-type test statistics are constructed using (21)–(24).*

(a) Under the null hypothesis, Q^{ols}, Q^{rec}, M^{ols}, and M^{rec} have the same limiting distributions as those in Theorems 1(a) and 2(a).

(b) Under the alternative hypothesis, if $\delta'(\Lambda_{jj,1} - \Lambda_{jj,0})\delta \to 0$ as $|\delta| \to \infty$ for some j, then

$$\frac{1}{T^{2/3}}Q^{ols} = O_p(\|\delta\|^{-4/3}), \qquad \frac{1}{T^{2/3}}Q^{rec} = O_p(\|\delta\|^{-4/3}),$$
$$\frac{1}{T^{1/3}}M^{ols} = O_p(\|\delta\|^{-2/3}), \qquad \frac{1}{T^{1/3}}M^{rec} = O_p(\|\delta\|^{-2/3}),$$

whereas if $\delta'(\Lambda_{jj,1} - \Lambda_{jj,0})\delta \nrightarrow 0$ as $|\delta| \to \infty$ for all of j, then

$$\frac{1}{T^{2/3}}Q^{ols} = O_p(1), \qquad \frac{1}{T^{2/3}}Q^{rec} = O_p(1),$$
$$\frac{1}{T^{1/3}}M^{ols} = O_p(1), \qquad \frac{1}{T^{1/3}}M^{rec} = O_p(1).$$

We can see that $\delta'(\Lambda_{jj,1} - \Lambda_{jj,0})\delta \to 0$ if x_t consists of only a constant ($x_t = 1$). In this case, the tests suffer from the so-called non-monotonic power problem as investigated by Vogelsang (1999) and several methods have been proposed to overcome this problem by Crainiceanu and Vogelsang (2007), Kejriwal (2009), Juhl and Xiao (2009), Shao and Zhang (2010), Yang and Vogelsang (2011), and Yamazaki and Kurozumi (2015) among others. We also note that even if the above condition does not hold, the divergence rates of the test statistics are reduced in the case of serially correlated errors compared to the case in Section 3.2. That is, the modifications robust to serial correlation result in the reduction of power, as is often observed in the literature.

4. Finite Sample Properties

In this section, we investigate the finite sample performance of the tests considered in this study. The data generating process (DGP) we consider is given by

$$y_t = x'_t(\beta_t + \delta 1_{(t>[T\lambda])}) + u_t, \qquad u_t = \rho u_{t-1} + \epsilon_t,$$

12

where $x_t = [1, z_t]'$, $\beta = [1, 1]'$, and $\{\epsilon_t\} \sim i.i.d.N(0, (1 - \rho)^2)$. The settings for δ and λ are explained later. The stochastic regressor z_t is an AR(1) process with mean 1 and variance 1, given by

$$z_t = 0.5 + 0.5z_{t-1} + e_t, \quad \{e_t\} \sim i.i.d.N(0, 0.75),$$

where $\{e_t\}$ is independent of $\{\epsilon_t\}$. We set $\rho = 0$ to investigate the performance of the truncated versions of the tests in Section 3.2 given in (18)–(20), while $\rho = 0.4$ and 0.8 are used for the tests that are robust to serial correlation developed in Section 3.3. The sample size T is 100 and 200, the number of replications is 5000, and all computations are conducted using the GAUSS matrix language.

We first investigate the finite sample performance of the tests in Section 3.2 with $\rho = 0$. From Panel (a) of Table 2, we can see that the sizes of all the tests are relatively well controlled, although they tend to be slightly conservative. Since the empirical sizes of the tests are different, we investigate the finite sample properties of the tests under the alternative using the size adjusted powers. We set a one-time shift in the coefficient to $\delta = b[1, 1]'$ (non-orthogonal change with $c_1'\delta = 1$) and $\delta = b[-1, 1]'$ (orthogonal change with $c_1'\delta = 0$), in which the magnitude of the change is controlled by $b = 0, 0.5, 1.0, 1.5,$ and 2.0, and the break fraction λ is set to 0.5. Figures 2(a) and (b) show that the difference in power is relatively small among the three tests based on the same (OLS or recursive) residuals when $c'\delta = 1$. However, as is seen in Figures 2(c) and (d), when $c'\delta = 0$, the modified tests are more powerful than the original tests. When we focus on either the quadratic-type or maximum-type tests, the OLS-based test is more powerful than the recursive-based test. We also investigate the effect of the location of the break on the performance of the quadratic-type and maximum-type tests by changing λ from 0.2 to 0.8. Figures 2(e)–(h) show that the effect of the location of the change in finite samples is consistent with the theoretical result given in Section 3. For example, the modified tests using the OLS residuals are maximized at $\lambda = 0.5$, whereas for an early break, the tests using recursive residuals outperform those using OLS residuals and vice versa for a late break.

In the case where the error term is possibly serial correlated, we should use the tests proposed in Section 3.3, the empirical sizes of which are summarized in Panels (b) and (c) of

13

Table 2. We can see that the tests based on the OLS residuals tend to suffer from under-size distortion, particularly when the serial correlation is strong, with $\rho = 0.8$. With regards to power in the case of $\rho = 0.4$, the relative performance of the tests seems to be preserved, but we also observe from Figures 3(a) and (b) that the tests suffer from the so-called non monotonic power.[2] Figures 3(g) and (h) show that the effect of the location of a change on the tests is similar to the case of serially uncorrelated errors. We obtain the similar tendency in the case of $\rho = 0.8$ and omit details.

5. Conclusion

When a structural change is orthogonal to the mean of the regressors, the standard CUSUM tests lose power. As a result, several modified tests have been proposed, explicitly or implicitly, in the literature. We investigated the asymptotic properties of such modified tests, and found that they can successfully reject the null hypothesis, even in the case of an orthogonal structural change. In this sense, the modified tests could complement the standard test in empirical analyses.

---

[2]Since the main purpose of this study is to investigate the modified CUSUM tests developed to overcome the loss of power caused by an orthogonal shift of parameters, we do not pursue this problem further here.

References

[1] Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817-858.

[2] Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society* B, 37, 149-163.

[3] Crainiceanu, C. M. and Vogelsang, T. J. (2007). Nonmonotonic power for tests of a mean shift in a time series. *Journal of Statistical Computation and Simulation* 77, 457-476.

[4] Deng, A. and Perron, P. (2008). A non-local perspective on the power properties of the CUSUM and CUSUM of squares tests for structural change. *Journal of Econometrics* 142, 212-240.

[5] Garbade, K. (1977). Two methods for examining the stability of regression coefficients. *Journal of the American Statistical Association* 72:357, 54-63.

[6] Horváth, L., Hušková, M., Kokoszka, P., and Steinebach, J. (2004). Monitoring changes in linear models. *Journal of Statistical Planning and Inference* 126, 225-251.

[7] Hušková, M. and Koubková, A. (2005). Monitoring jump changes in linear models. *Journal of Statistical Research* 39:2, 51-70.

[8] Hušková, M., Prášková, Z., and Steinebach, J. (2007). On the detection of changes in autoregressive time series I. Asymptotics. *Journal of Statistical Planning and Inference* 137, 1243-1259.

[9] Hušková, M., Kirch, C., Prášková, Z., and Steinebach, J. (2008). On the detection of changes in autoregressive time series, II. Resampling procedures. *Journal of Statistical Planning and Inference* 138, 1697-1721.

[10] Juhl, T. and Xiao, Z. (2009). Tests for changing mean with monotonic power.*Journal of Econometrics* 148, 14-24.

[11] Kejriwal, M. (2009). Tests for a mean shift with good size and monotonic power. *Economics Letters* 102, 78-82.

[12] Luger, R. (2001). A Modified CUSUM test for orthogonal structural changes. *Economics Letters* 73, 301-306.

[13] Ploberger, W. and Krämer, W. (1990). The local power of the CUSUM and CUSUM of squares tests. *Econometric Theory* 6, 335-347.

[14] Ploberger, W. and Krämer, W. (1992). The CUSUM test with OLS residuals. *Econometrica* 60, 271-285.

[15] Sen, P. K. (1982). Invariance principles for recursive residuals. *The Annals of Statistics* 10, 307-312.

[16] Shao, X. and Zhang,X (2010). Testing for change points in time series. *Journal of American Statistical Association* 105, 1228-1240.

[17] Tang, S. M. and MacNeill, I. B. (1993). The effect of serial correlation on tests for parameter change at unknown time. *Annals of Statistics* 21, 552-575.

[18] Vogelsang, T. J. (1999). Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. *Journal of Econometrics* 88, 283-299.

[19] Xia, Z. M., Guo, P. J., and Zhao, W. Z. (2011). CUSUM methods for monitoring structural changes in structural equations. *Communications in Statistics-Theory and Methods* 40:6, 1109-1123.

[20] Yang, J. and Vogelsang, T. J. (2011). Fixed-b analysis of LM-type tests for a shift in mean. *Econometrics Journal* 14, 438-456.

[21] Yamazaki, D. and Kurozumi, E. (2015). Improving the finite sample performance of tests for a shift in mean. *Journal of Statistical Planning and Inference* 167, 144-173.

Appendix

Proof of Proposition 1: (a) In the case of the OLS residuals, from the functional central limit theorem (FCLT), the weak law of large numbers (WLLN), and the continuous mapping theorem (CMT), we have

$$\sqrt{T}(\hat{\beta} - \beta) \Rightarrow C^{-1}B(1), \tag{25}$$

where $B(r) = [B_1(r), \cdots, B_k(r)]'$ for $0 \leq r \leq 1$ is a k-dimensional Brownian motion with variance $\sigma^2 C$. Since $\hat{u}_t = u_t - x_t'(\hat{\beta} - \beta)$, we have

$$
\begin{aligned}
\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \hat{w}_{jt} &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} x_{jt} u_t - \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} x_{jt} x_t'(\hat{\beta} - \beta) \\
&\Rightarrow B_j(r) - r c_j' C^{-1} B(1) \\
&= B_j(r) - r B_j(1),
\end{aligned}
$$

where the last equality holds because $c_j' C^{-1} = [0, \cdots, 0, 1, 0, \cdots, 0]$. Similarly, we have

$$\hat{\sigma}_j^2 = \frac{1}{T} \sum_{t=1}^{T} x_{jt}^2 u_t^2 + o_p(1) \xrightarrow{p} \sigma^2 c_{jj}.$$

Since $B_j(r) - r B_j(1) =_d \sigma c_{jj}^{1/2} BB_j(r)$ where $BB_j(r)$ is a standard Brownian bridge, we obtain (5).

For the CUSUM test based on the recursive residuals, following Ploberger and Krämer (1988), we have

$$\frac{1}{\sqrt{T-k}} \sum_{t=k+1}^{[Tr]} \tilde{w}_{jt} = \frac{1}{\sqrt{T-k}} \sum_{t=k+1}^{[Tr]} x_{jt} \tilde{u}_t \Rightarrow \sigma c_{jj}^{1/2} W_j(r),$$

and $T^{-1} \sum_{t=1}^{T} \tilde{w}_{jt}^2 \xrightarrow{p} \sigma^2 c_{jj}$, where $W_j(r)$ for $0 \leq r \leq 1$ is a standard Brownian motion. We then obtain (6).

(b) Since $\hat{\beta}$ is expressed as (2) under the alternative, we have

$$\hat{\beta} - \beta \xrightarrow{p} (1 - \lambda)\delta.$$

Then, using expression (3), the numerator of the test statistic becomes

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{[Tr]}\hat{w}_{jt} &= \frac{1}{T}\sum_{t=1}^{[Tr]}x_{jt}u_t + \frac{1}{T}\sum_{t=1}^{[Tr]}x_{jt}x_t'\delta 1_{(t>[T\lambda])} - \frac{1}{T}\sum_{t=1}^{[Tr]}x_{jt}x_t'(\hat{\beta}-\beta) \\
&\xrightarrow{p} 0 + c_j'\delta(r-\lambda)1_{(r>\lambda)} - rc_j'(1-\lambda)\delta \\
&= c_j'\delta\left[(r-\lambda)1_{(r>\lambda)} - r(1-\lambda)\right] \tag{26}
\end{aligned}
$$

uniformly over $0 \leq r \leq 1$, the absolute value of which is maximized at $r = \lambda$, which is equal to $|c_j'\delta|\lambda(1-\lambda)$. Similarly, we have

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\hat{w}_{jt}^2 &= \frac{1}{T}\sum_{t=1}^{T}x_{jt}^2 u_t^2 + \frac{1}{T}\sum_{t=1}^{T}\left(x_{jt}x_t'\delta 1_{(t>[T\lambda])}\right)^2 + \frac{1}{T}\sum_{t=1}^{T}\left(x_{jt}x_t'(\hat{\beta}-\beta)\right)^2 \\
&\quad + \frac{2}{T}\sum_{t=1}^{T}x_{jt}^2 u_t x_t'\delta 1_{(t>[T\lambda])} - \frac{2}{T}\sum_{t=1}^{T}x_{jt}^2 u_t x_t'(\hat{\beta}-\beta) \\
&\quad - \frac{2}{T}\sum_{t=1}^{T}x_{jt}^2 1_{(t>[T\lambda])}\delta' x_t x_t'(\hat{\beta}-\beta) \\
&\xrightarrow{p} \sigma^2 c_{jj} + (1-\lambda)\delta'\Lambda_{jj,0}\delta + (1-\lambda)^2\delta'\Lambda_{jj,0}\delta + 0 - 0 - 2(1-\lambda)^2\delta'\Lambda_{jj,0}\delta \\
&= \sigma^2 c_{jj} + \lambda(1-\lambda)\delta'\Lambda_{jj,0}\delta. \tag{27}
\end{aligned}
$$

From (26) and (27), we obtain (7).

The proof in the case with the recursive residuals is analogous to the OLS case. We first note that since $f_t \xrightarrow{p} 1$, it can be shown that $T^{-1}\sum_{t=1}^{T}(1/f_t - 1)x_t x_t' \xrightarrow{p} 0$, which implies that $T^{-1}\sum_{t=1}^{T}x_t x_t'/f_t \xrightarrow{p} C$. This result will be used repeatedly in our proofs below.

Since the recursive residuals are written under the alternative as $\tilde{u}_t = [u_t + x_t'\delta 1_{t>[(T\lambda)]} - x_t'(\hat{\beta}_{t-1}-\beta)]/f_t$ and $\hat{\beta}_{t-1}-\beta = \left(\sum_{s=1}^{t-1}x_s x_s'\right)^{-1}\left(\sum_{s=1}^{t-1}x_s x_s'\delta 1_{t>[T\lambda]} + \sum_{s=1}^{t-1}x_s u_s\right)$, $\tilde{w}_{jt}$ can be expressed as

$$
\tilde{w}_{jt} = \frac{1}{f_t}x_{jt}u_t + \frac{1}{f_t}x_{jt}x_t'\delta 1_{(t>[T\lambda])} - \frac{1}{f_t}x_{jt}x_t'\left(\sum_{s=1}^{t-1}x_s x_s'\right)^{-1}\left(\sum_{s=1}^{t-1}x_s x_s'\delta 1_{(s>[T\lambda])} + \sum_{s=1}^{t-1}x_s u_s\right).
$$

Then, we have

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{[Tr]}\tilde{w}_{jt} &\xrightarrow{p} 0 + c_j'\int_0^r \delta 1_{(v>\lambda)}dv - c_j'\int_0^r\left[(vC)^{-1}\int_0^v C\delta 1_{(w>\lambda)}dw\right]dv \\
&= c_j'\delta\lambda[\log(r) - \log(\lambda)]1_{(r>\lambda)} \tag{28}
\end{aligned}
$$

18

uniformly over $0 \le r \le 1$. Using standard calculus, we have that $|c_j'\delta|\lambda[\log(r)-\log(\lambda)]1_{(r>\lambda)}/(1+2r)|$ takes the maximum value $|c_j'\delta|q$, where q is defined as (9). Furthermore, similarly to (27), we have

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\tilde{w}_{jt}^2 \quad\xrightarrow{p}\quad & \sigma^2 c_{jj} + (1-\lambda)\delta'\Lambda_{jj,0}\delta \\
& + \int_0^1 tr\left[\frac{1}{v}C^{-1}\left(\int_0^v C\delta 1_{(w>\lambda)}dw\right)\left(\int_0^v C\delta 1_{(w>\lambda)}dw\right)'\frac{1}{v}C^{-1}\Lambda_{jj,0}\right]dv \\
& + 0 - 0 - 2\int_0^1 tr\left[\delta 1_{(v>\lambda)}\int_0^v \delta'1_{(w>\lambda)}dw\frac{1}{v}\Lambda_{jj,0}\right]dv \\
= \quad & \sigma^2 c_{jj} + \lambda(1-\lambda)\delta'\Lambda_{jj,0}\delta. \tag{29}
\end{aligned}
$$

From (28) and (29), we can see that

$$
\tilde{\sigma}_j^2 \xrightarrow{p} \sigma^2 c_{jj} + \lambda(1-\lambda)\delta'\Lambda_{jj,0}\delta - \left(c_j'\delta\lambda\log(\lambda)\right)^2.
$$

We then obtain (8).∎

Proof of Theorem 1: (a) From (25), we have

$$
\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]}x_t\hat{u}_t \Rightarrow B(r) - rB(r),
$$

and thus

$$
\left(\sum_{t=1}^{[Tr]}x_t\hat{u}_t\right)'\left(\sum_{t=1}^{T}x_t x_t'\right)^{-1}\left(\sum_{t=1}^{[Tr]}x_t\hat{u}_t\right) \Rightarrow \sigma^2\|BB(r)\|^2.
$$

Since $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$ under the null hypothesis, (10) is obtained.

The null limiting distribution of Q^{rec} can be derived similarly.

(b) In the same way as (26), we have

$$
\frac{1}{T}\sum_{t=1}^{[Tr]}x_t\hat{u}_t \xrightarrow{p} C\delta\left[(r-\lambda)1_{(r>\lambda)} - r(1-\lambda)\right] \tag{30}
$$

uniformly over $0 \le r \le 1$, and, thus,

$$
\left(\frac{1}{T}\sum_{t=1}^{[Tr]}x_t\hat{u}_t\right)'\left(\frac{1}{T}\sum_{t=1}^{T}x_t x_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{[Tr]}x_t\hat{u}_t\right) \xrightarrow{p} \delta'C\delta\left[(r-\lambda)1_{(r>\lambda)} - r(1-\lambda)\right]^2
$$

uniformly over $0 \leq r \leq 1$, which achieves a maximum at $r = \lambda$, while $\hat{\sigma}^2 \xrightarrow{p} \sigma^2 + \lambda(1-\lambda)\delta' C\delta$, as proved by Deng and Perron (2008). We then obtain (12).

In the case of recursive residuals, in the same way as in (28), we have

$$\frac{1}{T} \sum_{t=k+1}^{[Tr]} x_t \tilde{u}_t \xrightarrow{p} C\delta\lambda \left(\log(r) - \log(\lambda)\right) 1_{(r>\lambda)}$$

uniformly over $0 \leq r \leq 1$, and, thus,

$$\left(\frac{1}{T} \sum_{t=k+1}^{[Tr]} x_t \hat{u}_t\right)' \left(\frac{1}{T} \sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\frac{1}{T} \sum_{t=k+1}^{[Tr]} x_t \hat{u}_t\right) \xrightarrow{p} \delta' C\delta \left[\lambda(\log(r) - \log(\lambda))1_{(r>\lambda)}\right]^2,$$

while $\tilde{\sigma}^2 \xrightarrow{p} \sigma^2 + \lambda(1-\lambda)\delta' C\delta - (c_1' \delta\lambda \log(\lambda))^2$ is shown by Deng and Perron (2008), which implies (13).∎

Proof of Theorem 2: (a) Since

$$\left[M_1^{ols}, \cdots, M_k^{ols}\right]' = \frac{1}{\hat{\sigma}} \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1/2} \sum_{t=1}^{[Tr]} x_t \hat{u}_t \xrightarrow{d} BB(r),$$

(14) is obtained. (15) can be proved similarly to (11).

(b) Under the alternative, we have

$$\left[M_1^{ols}, \cdots, M_k^{ols}\right]' \xrightarrow{p} \frac{C^{1/2}\delta \left[(r-\lambda)1_{(r>\lambda)} - r(1-\lambda)\right]}{\sqrt{\sigma^2 + \lambda(1-\lambda)\delta' C\delta}}$$

and, therefore,

$$\left[\sup_{0\leq r\leq 1} |M_1^{ols}|, \cdots, \sup_{0\leq r\leq 1} |M_k^{ols}|\right]' \xrightarrow{p} \frac{1}{\sqrt{\sigma^2 + \lambda(1-\lambda)\delta' C\delta}} \left[|v_1|\lambda(1-\lambda), \cdots, |v_k|\lambda(1-\lambda)\right]',$$

where v_j is defined as in Theorem 2. We then have

$$\frac{1}{\sqrt{T}} \max_{1\leq j\leq k} \sup_{0\leq r\leq 1} \left|M_j^{ols}(r)\right| \xrightarrow{p} \frac{\max_{1\leq j\leq k}\{|v_1|, \cdots, |v_k|\}\lambda(1-\lambda)}{\sqrt{\sigma^2 + \lambda(1-\lambda)\delta' C\delta}}.$$

For the recursive residuals, we can prove (17) similarly to the proof of Theorem 1(b).∎

Proof of Theorem 3

(a) The null limiting distributions can be obtained in the same way as in Theorems 1(a) and 2(a).

(b) We first derive the divergence rate of the bandwidth m. Since $\hat{u}_t$ is expressed as (3), in the same way as (27), we have

$$\frac{1}{T}\sum_{t=p+1}^{T}\hat{w}_{jt}\hat{w}_{jt-p} \xrightarrow{p} \gamma_{jj,p} + \lambda(1-\lambda)(\delta'\Lambda_{jj,p}\delta) \tag{31}$$

for a given p. Then, we have

$$\hat{\rho}_j \xrightarrow{p} \frac{\gamma_{jj,1} + \lambda(1-\lambda)\delta'\Lambda_{jj,1}\delta}{\gamma_{jj,0} + \lambda(1-\lambda)\delta'\Lambda_{jj,0}\delta},$$

and hence

$$1-\hat{\rho}_j \xrightarrow{p} \frac{\gamma_{jj,0} - \gamma_{jj,1} + \lambda(1-\lambda)\delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta}{\gamma_{jj,0} + \lambda(1-\lambda)\delta'\Lambda_{jj,0}\delta} = \begin{cases} O_p\left(\|\delta\|^{-2}\right) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \to 0 \\ O_p(1) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \nrightarrow 0 \end{cases}.$$

Since $\hat{\sigma}_j^2 = O_p(\|\delta\|^2)$ by (27), we can see that $a(\delta) = O_p\left(\|\delta\|^4\right)$ if $\delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \to 0$, for some j, and is $O_p(1)$ otherwise. Hence,

$$m = \begin{cases} O_p\left(\|\delta\|^{4/3}T^{1/3}\right) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \to 0 \quad \exists j \\ O_p(T^{1/3}) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \nrightarrow 0 \quad \forall j \end{cases}.$$

Using this result, we next derive the divergence order of $\hat{\Omega}$. In the same way as (31),

$$\hat{\Gamma}_p \xrightarrow{p} \Gamma_p + \lambda(1-\lambda)\text{plim}\frac{1}{T}\sum_{t=p+1}^{T}x_t x_t' \delta\delta' x_{t-p} x_{t-p}' = O_p(\|\delta\|^2).$$

Then, since $\sum_{p=1}^{m} k(p,m) = O(m)$,

$$\|\hat{\Omega}\| = \left\| \hat{\Gamma}_0 + \sum_{p=1}^{m} k(p,m)\left(\hat{\Gamma}_p + \hat{\Gamma}_p'\right) \right\|$$

$$\leq O(m)O_p(\|\delta\|^2) = \begin{cases} O_p\left(\|\delta\|^{10/3}T^{1/3}\right) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \to 0 \quad \exists j \\ O_p(\|\delta\|^2 T^{1/3}) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \nrightarrow 0 \quad \forall j \end{cases}. \tag{32}$$

Since it can be shown that (30) holds under Assumption 4, we can see that

$$Q^{ols} = \begin{cases} O_p\left(\|\delta\|^{-4/3}T^{2/3}\right) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \to 0 \quad \exists j \\ O_p(T^{2/3}) & : & \delta'(\Lambda_{jj,0} - \Lambda_{jj,1})\delta \nrightarrow 0 \quad \forall j \end{cases}.$$

Thus, we obtain the result.

The order of M^{ols} is obtained similarly using (30) and (32).

Since we can obtain the result of Q^{rec} and M^{rec} in the same way as in the case of the OLS residuals, we omit the proof. ∎

Table 1: Asymptotic critical values

| | 1% | | | | 5% | | | | 10% | | | |
| | Quadratic | | Max | | Quadratic | | Max | | Quadratic | | Max | |
	ols	rec	ols	rec	ols	rec	ols	rec	ols	rec	ols	rec
(a) $r \in [0,1]$												
k=1	1.465	3.792	1.210	1.947	1.806	4.958	1.344	2.227	2.608	7.795	1.615	2.792
k=2	2.077	5.776	1.341	2.215	2.469	7.179	1.465	2.478	3.347	10.406	1.718	3.002
k=3	2.584	7.473	1.412	2.366	3.009	9.033	1.532	2.616	3.945	12.553	1.773	3.122
k=4	3.038	9.035	1.461	2.468	3.491	10.738	1.578	2.712	4.488	14.486	1.813	3.207
k=5	3.464	10.522	1.498	2.546	3.942	12.340	1.613	2.784	4.999	16.312	1.842	3.270
(b) $r \in [0.15, 0.85]$												
k=1	1.463	3.218	1.210	1.794	1.805	4.217	1.344	2.054	2.608	6.636	1.615	2.576
k=2	2.077	4.902	1.340	2.043	2.469	6.100	1.465	2.283	3.347	8.842	1.718	2.770
k=3	2.584	6.350	1.412	2.179	3.009	7.680	1.532	2.412	3.945	10.680	1.773	2.883
k=4	3.038	7.675	1.461	2.274	3.491	9.127	1.578	2.500	4.488	12.298	1.813	2.959
k=5	3.464	8.932	1.498	2.346	3.942	10.489	1.613	2.567	4.999	13.886	1.842	3.018

Table 2: Empirical sizes under H_0

| | CUSUM | | Quadratic | | Max | |
	ols	rec	ols	rec	ols	rec
(a) $\rho = 0$						
$T = 100$	0.035	0.039	0.033	0.039	0.032	0.040
$T = 200$	0.039	0.040	0.037	0.039	0.040	0.040
(b) $\rho = 0.4$						
$T = 100$	0.042	0.050	0.019	0.059	0.020	0.049
$T = 200$	0.048	0.050	0.033	0.057	0.034	0.052
(c) $\rho = 0.8$						
$T = 100$	0.006	0.044	0.002	0.072	0.002	0.048
$T = 200$	0.029	0.038	0.006	0.057	0.009	0.039
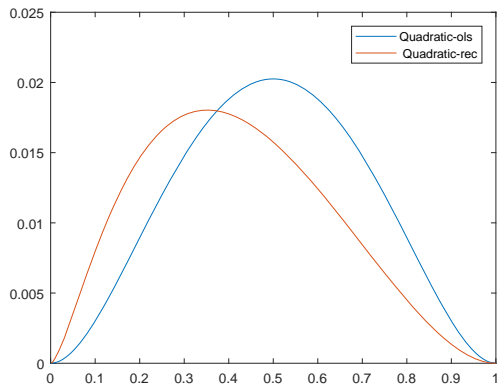
(a) Quadratic of ols and rec ($c'\delta = 0$)
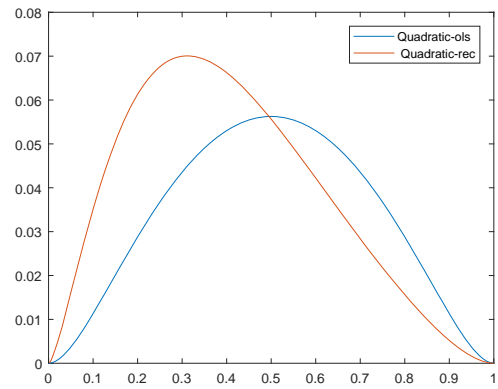
(b) Quadratic of ols and rec ($c'\delta \neq 0$)
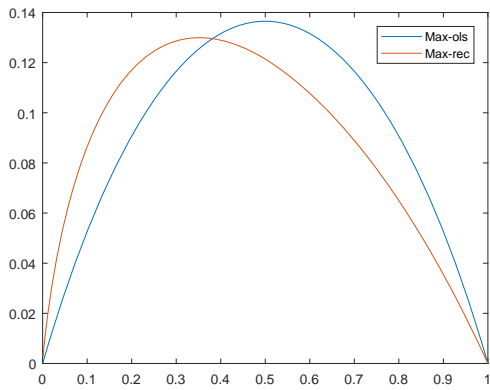
(c) Max of ols and rec ($c'\delta = 0$)
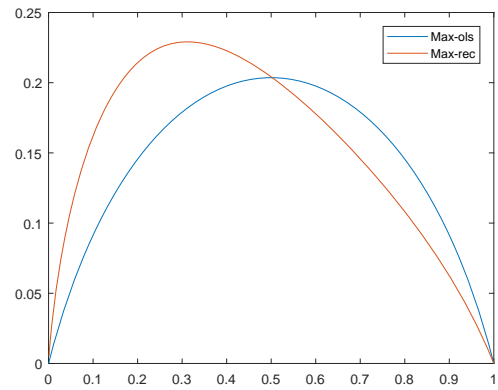
(d) Max of ols and rec ($c'\delta \neq 0$)
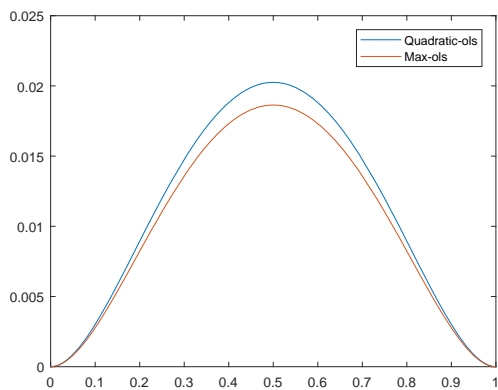
(e) Quadratic and Max of ols ($k = 2$)
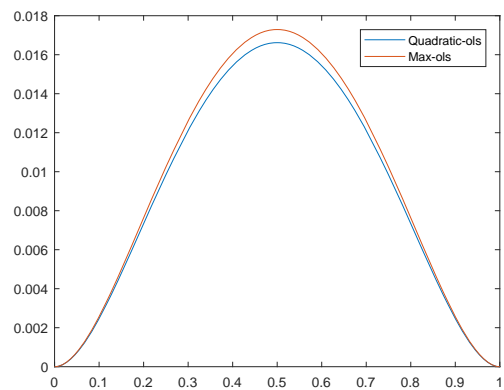
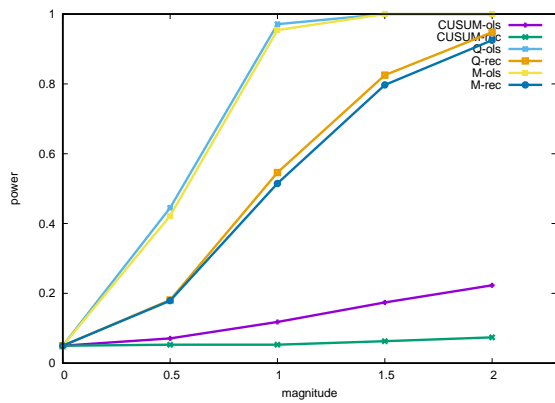(f) Quadratic and Max of ols ($k = 3$)
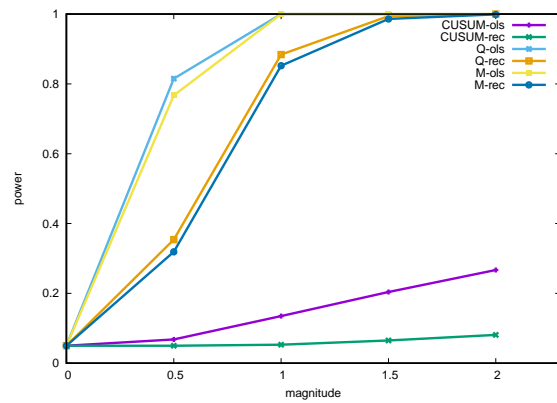
Figure 1: Asymptotic power properties
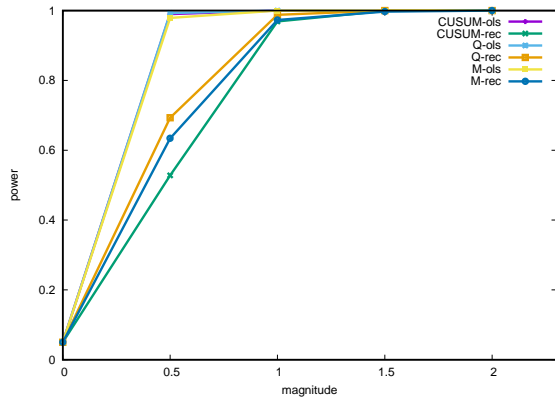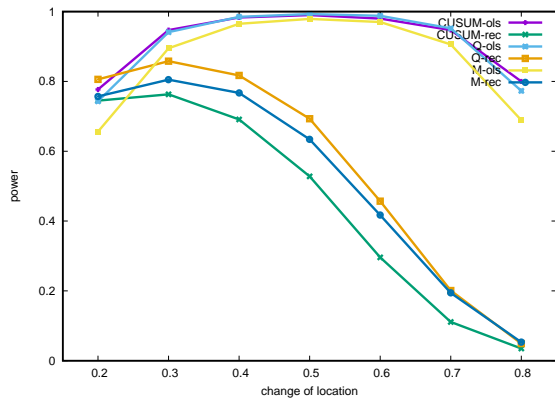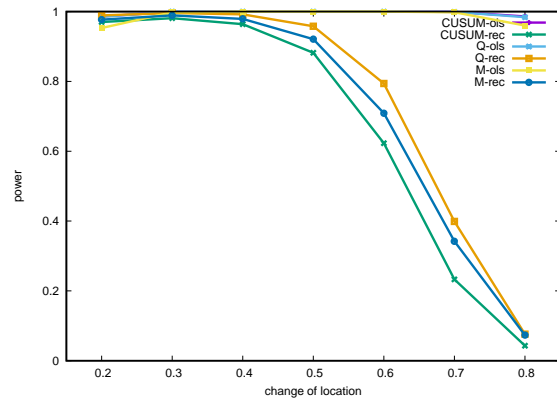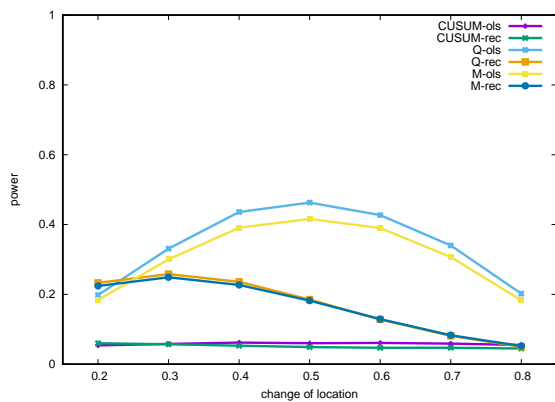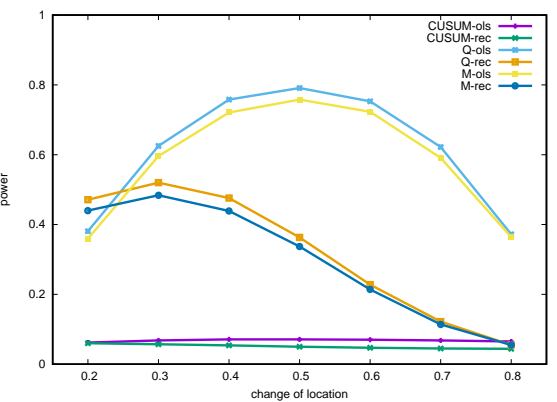
(a) $T = 100$, $c'\delta = 1$
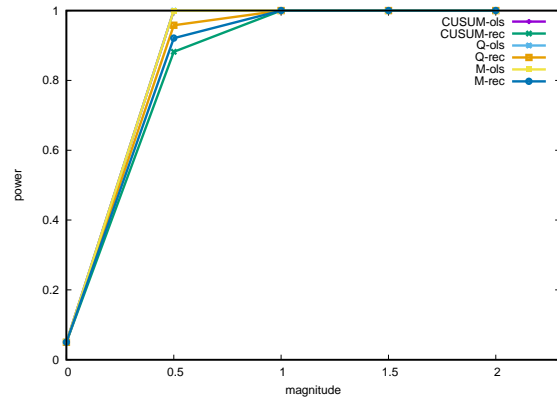
(b) $T = 200$, $c'\delta = 1$

(c) $T = 100$, $c'\delta = 0$

(d) $T = 200$, $c'\delta = 0$
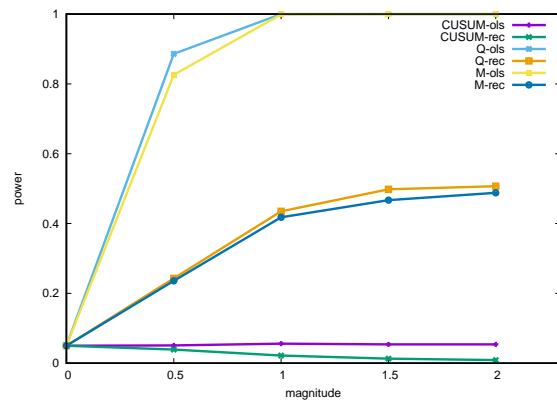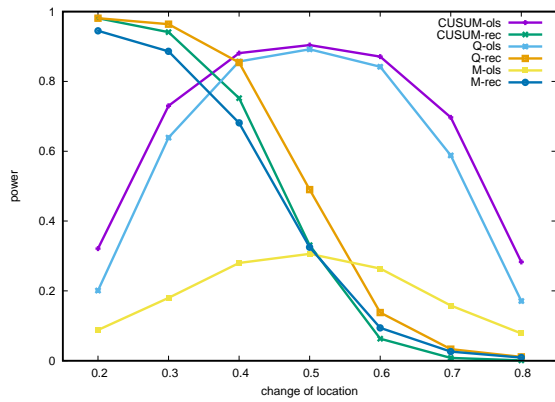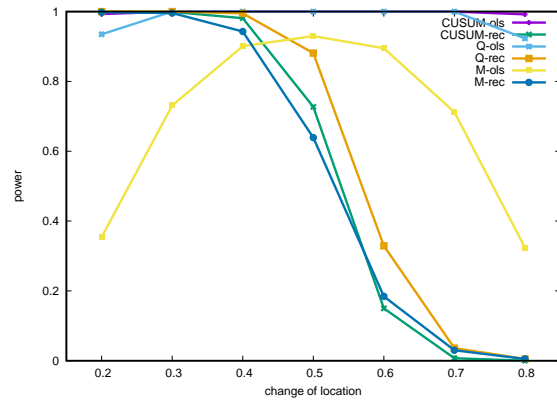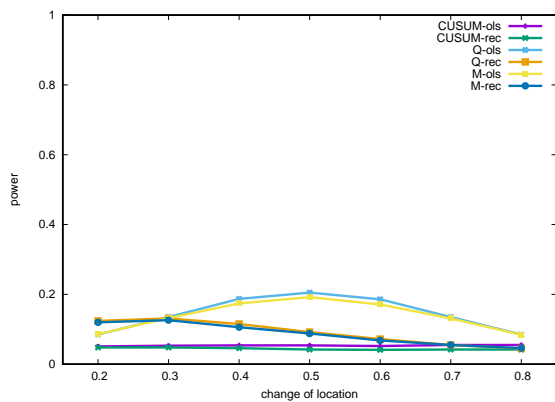
(e) $T = 100$, $c'\delta = 1$

(f) $T = 200$, $c'\delta = 1$

(g) $T = 100$, $c'\delta = 0$

(h) $T = 200$, $c'\delta = 0$

Figure 2: Size-adjusted powers ($\rho = 0$)
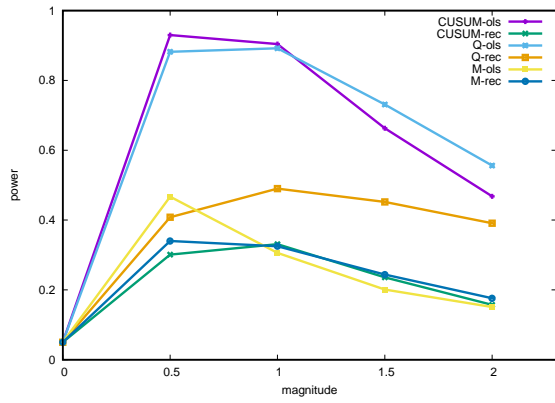
(a) $T = 100$, $c'\delta = 1$
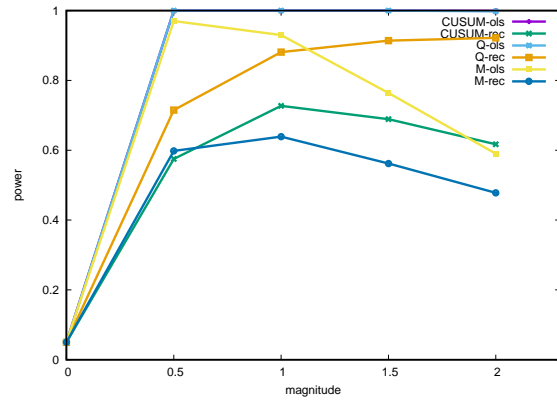
(b) $T = 200$, $c'\delta = 1$
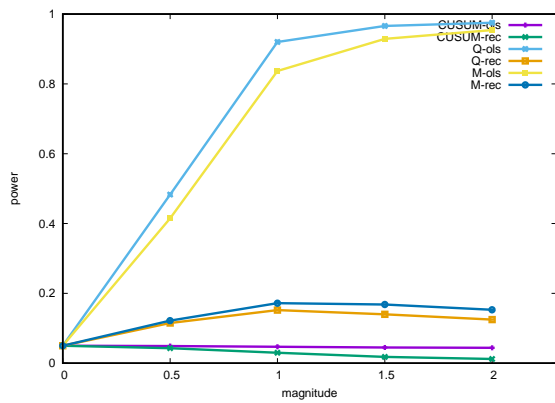
(c) $T = 100$, $c'\delta = 0$
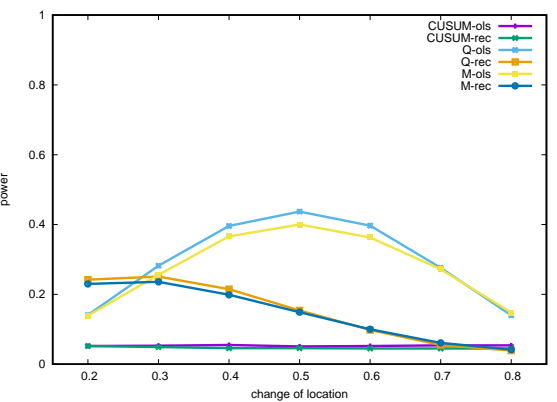
(d) $T = 200$, $c'\delta = 0$

(e) $T = 100$, $c'\delta = 1$

(f) $T = 200$, $c'\delta = 1$

(g) $T = 100$, $c'\delta = 0$

(h) $T = 200$, $c'\delta = 0$

Figure 3: Size-adjusted powers ($\rho = 0.4$)