

TESTING FOR STRUCTURAL STABILITY IN THE WHOLE SAMPLE

JAVIER HIDALGO
(LSE)

AND

MYUNG HWAN SEO
(LSE)

Discussion paper
No. EM/2011/558
October 2011

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel: 020 7955 6674

Abstract

The paper examines a Lagrange Multiplier type test for the constancy of the parameter in general models with dependent data without imposing any artificial choice of the possible location of the break. In order to prove the asymptotic behaviour of the test, we extend a strong approximation result for partial sums of a sequence of random variables. We also present a Monte-Carlo experiment to examine the finite sample performance of the test and how it compares with tests which assume some knowledge of the possible location of the break..

Keywords:

Structural Stability. GMM estimation. Strong approximation. Extreme value distribution.

1. INTRODUCTION

Since the work of Chow (1960) and Quandt (1960), testing for structural stability has been a very active topic of theoretical and applied research. Most of the research assumes that a structural break occurred in a particular segment of the sample. For instance, extensively studied in econometrics is the scenario where the possible break lies in the “middle” of the sample, that is in the interval $\Pi = ([n\tau], n - [n\tau])$ for some $\tau > 0$, and where n denotes herewith the sample size. See for example Andrews (1993) or the latest review article by Perron (2006). More recently, there has been some interest to learn what it would be the behaviour of the tests for stability if the break occurs at the end of the sample, that is among the last \hat{t} observations for some finite value \hat{t} . The work in the latter scenario is less prolific than in the former one, although we can cite the works by Andrews (2003) or Andrews and Kim (2006) and references therein. Between the two aforementioned setups, however, there is an important gap in the theory. More specifically, when we allow the possible break to lie in the intervals $(\hat{t}, [n\tau])$ or $(n - [n\tau], n - \hat{t})$, or more generally when we do not wish to impose any prior knowledge on the location of the hypothesized break.

The paper considers the problem of testing for structural stability over the whole sample span $t = 1, \dots, n$. That is, when no previous information about the location of the break is available, or equivalently with $\tau = 0$. We will examine the testing problem in models useful in econometrics, such as nonlinear simultaneous equations and transformation models under general conditions on the dependence structure of the variables of the model. In doing so, we develop a general strong approximation for a near epoch dependent random vectors. This substantially extends Horvath’s (1993) results who only examines this problem for a break in the mean of a sequence of otherwise independent normally distributed random variables or Davis et al’s (1995) autoregression setup. While Ling (2007) considers the near epoch dependence, it also requires a martingale property for its strong approximation.

It is worth mentioning that Andrews (1993) looked briefly at our setup. He signalled that his sup-tests for breaks will diverge to infinity if the supremum is taken with $\tau = 0$. Thus, the restriction to take the supremum (or other continuous functionals) over a trimmed set Π was made not only for technical convenience but is crucial to obtain a proper asymptotic distribution. It also applies to the “optimal tests” proposed by Andrews and Ploberger (1994). In this way, our paper shows that (a) the reason for Andrews’s (1993) findings is because the normalization that we need for the tests to have a proper asymptotic distribution is different to the conventional “root- n ”, and (b) that the asymptotic distribution of the tests is completely different than that obtained when we focus our test for stability in the interval Π . More specifically, we show that, after appropriate normalization, the sup-tests with $\tau = 0$ converge to the Type I *Extreme Value Distribution* or *Gumbel* distribution.

The Monte-Carlo experiment in Section 5 suggests that when we compare the power of the conventional tests presented in Andrews (1993), Andrews and Ploberger (1994), or Brown, Durbin and Evans (1975) against that of our test in Sections 2, the assumptions made in the previous works are not innocuous. We discuss some

theoretical and practical issues regarding our test compared to them. From a theoretical point of view, our test is always consistent irrespective of the location of the break, whereas the conventional tests are not. More specifically, as Section 4 shows, the conventional tests are not consistent when the break occurs at time $t \leq \lfloor n^{1/2} \rfloor$ or $n - \lfloor n^{1/2} \rfloor < t$, whereas our tests are. In addition, we show that when the break falls in the region $t \in \left(\lfloor n^{1/2} \rfloor, n / (\log \log n)^{1/2} \right)$, the conventional tests have zero asymptotic efficiency compared to ours, in the sense that our tests are able to detect local alternatives that, for instance, the “optimal tests” would not detect. From a practical point of view, our test has at least two advantages. First, there is no arbitrariness associated with the choice of the trimming parameter τ . Second, our test can detect a break earlier than the conventional ones and it can detect a break that has occurred more recently. This would be particularly relevant in forecasting.

The remainder of the paper is organized as follows. Sections 3 and 4 develop the asymptotic distribution of \mathcal{LM} and \mathcal{LM}^* under the null and local alternatives, examining the asymptotic power properties of the tests. Section 5 presents a Monte-Carlo experiment to examine the finite sample performance of \mathcal{LM}^* and \mathcal{LM} and how they compare with \mathcal{LM}_τ for some values of τ . Finally, Section 6 gives the proofs of our results in Sections 3 and 4.

2. TESTING FOR A BREAK IN NON-LINEAR MODELS

Let y_t be a $N \times 1$ vector of endogenous variables, x_t a $K \times 1$ vector of explanatory variables, ν_t a $N \times 1$ vector of disturbance terms such that

$$(2.1) \quad \nu(y_t, x_t; \theta_{t0}) \triangleq \nu_t, \quad t = 1, \dots, n$$

for a vector-valued function $\nu(y, x; \theta)$, where θ_t is a p -dimensional vector of unknown parameters. As usual a subscript 0 to a parameter indicates its true value. Model (2.1) is general enough to accommodate nonlinear transformation models and nonlinear simultaneous equation models. By the former we mean that $\nu(y, x; \theta_t) = \nu_y(y; \theta_t) - \nu_x(x; \theta_t)$. Well-known transformations include the Box-Cox and the Burbidge, Magee, and Robb’s (1988) transformations, given respectively by $\nu_y(y; \theta_t) = (y^{\lambda_t} - 1) / \lambda_t$ and $\nu_y(y; \theta_t) = \operatorname{arcsinh}(\lambda_t y) / \lambda_t$, and where λ_t is an element of θ_t .

This paper is concerned with testing the constancy of θ_t within the framework of one-time structural break. That is, setting

$$\theta_{t0} = \theta_0 + \delta_0 \mathbf{1}(s < t), \quad 1 < s < n,$$

the null hypothesis of interest becomes

$$(2.2) \quad H_0 : \delta_0 = 0 \quad \text{for all } 1 < s < n,$$

whereas the alternative hypothesis H_1 is the negation of the null.

Under the assumption that $\mathbb{E}(\nu_t | x_t) = 0$, we have that for any $M \times 1$ vector-valued function $P(x)$, $\mathbb{E}[P(x_t) \otimes \nu_t] = 0$. Here “ \otimes ” denotes the Kronecker product. Thus under the null hypothesis and denoting

$$g_t(\theta) = P(x_t) \otimes \nu(y_t, x_t; \theta),$$

a common way to estimate the parameter θ_0 in (2.1) is by

$$(2.3) \quad \hat{\theta} = \operatorname{argmin}_{\theta \in \Theta} \left\{ \frac{1}{n} \sum_{t=1}^n g'_t(\theta) \right\} \Delta(\hat{\theta})^{-1} \left\{ \frac{1}{n} \sum_{t=1}^n g_t(\theta) \right\}.$$

Here $\tilde{\theta}$ denotes some preliminary $n^{1/2}$ -consistent estimator of θ_0 , say

$$\tilde{\theta} = \operatorname{argmin}_{\theta \in \Theta} \left\{ \frac{1}{n} \sum_{t=1}^n g'_t(\theta) \right\} \left(\frac{1}{n} \sum_{t=1}^n P(x_t) P'(x_t) \otimes I_N \right)^{-1} \left\{ \frac{1}{n} \sum_{t=1}^n g_t(\theta) \right\},$$

and, letting $I(\lambda; \theta)$ the periodogram matrix of the sequence $\{g_t(\theta)\}_{t=1}^n$,

$$(2.4) \quad \Delta(\tilde{\theta}) = \frac{1}{m} \sum_{j=1}^m I\left(\frac{2\pi j}{n}; \tilde{\theta}\right)$$

is the standard weighted periodogram estimator of

$$(2.5) \quad \Delta = \sum_{j=-\infty}^{\infty} \mathbb{E}(g_t(\theta_0) g'_{t+j}(\theta_0)).$$

Under the assumptions given in Section 3, it is known that $\Delta(\tilde{\theta}) \rightarrow_P \Delta$. Under Robinson's (1998) conditions, one can take $m = \lfloor n/2 \rfloor$ in (2.4). The estimator $\hat{\theta}$ in (2.3) is known as the generalized method of moments (GMM) estimator.

We now introduce our LM-type test for H_0 . To that end, define

$$(2.6) \quad \mathcal{LM}(s) = \frac{n}{n-s} \frac{1}{s} \operatorname{lm}(s)' \left(G_n(\hat{\theta}) \Delta(\hat{\theta})^{-1} G'_n(\hat{\theta}) \right)^{-1} \operatorname{lm}(s),$$

where $\operatorname{lm}(s) = G_n(\hat{\theta}) \Delta(\hat{\theta})^{-1} \sum_{t=1}^s g_t(\hat{\theta})$ with $G_n(\theta) = \frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \theta'} g_t(\theta)$. Also let $a_n = (2 \log_2 n)^{1/2}$, $b_n = 2 \log_2 n + \frac{p_1}{2} \log_3 n - \log \Gamma(p_1/2)$, where $\log_2 x = \log \log x$, $\log_3 x = \log \log \log x$ and $\Gamma(\cdot)$ is the gamma function. Then, our test for H_0 in (2.2) is given by either

$$(2.7) \quad (a) \quad \mathcal{LM}^* = a_n \left(\max_{p < s \leq n-p} \mathcal{LM}(s) \right)^{1/2} - b_n,$$

$$(2.8) \quad (b) \quad \mathcal{LM} = \left(\max_{p < s \leq n-p} \mathcal{LM}(s) - B_n \right) / A_n,$$

where $A_n = b_n/a_n^2$ and $B_n = b_n^2/a_n^2$. The difference between (2.7) and (2.8) is a negligible constant, see Davis et al. (1995) for more discussion.

We make some comparison to the existing literature. First, similar tests employed by Davis et al. (1995), Hušková et al. (2007), or Ling (2007) mainly concern an (linear or non-linear) autoregressive model, extending earlier results by Horváth (1993) in the context of testing the constancy of the mean of a sequence of independent normal variates. And thus, they do not cover a nonlinear regression model with general heteroskedasticity and serial dependence, commonly employed in econometric literature. For instance, many asset pricing models entails serially correlated $g_t(\theta_0)$, see e.g. Cochrane (2005). Second, as a consequence of the first, we consider the GMM estimation rather than the M -estimation. Third, existing asymptotic theory does not cover our general case, thus entailing a new approximation theory as in Proposition 1.

The tests (2.7) and (2.8) are more convenient than Wald type tests due to the possible computational burden arising from the estimation of the parameters, in this case, θ and δ . In addition, the estimators of the parameters θ and δ are very poor when s is either small or large, and thus the estimation of Δ . The tests \mathcal{LM} and \mathcal{LM}^* are in clear contrast with those which assumed some knowledge on the possible location of the break. That is, tests based on

$$\mathcal{LM}_\tau = \max_{s \in \mathfrak{Z}(\tau)} \mathcal{LM}(s),$$

where $\mathfrak{Z}(\tau) = [[\tau n], n - [\tau n]]$ for some fixed $\tau \in (0, 1/2)$, and which are widely used in setups like ours, see Andrews (1993). However, there are a few drawbacks, namely (i) the arbitrariness on the choice of the trimming value τ and (ii) the loss of (lack of) power when the break happens to be outside the set $\mathfrak{Z}(\tau)$. Sections 4 and 5 explore its consequences through the local power analysis and a Monte-Carlo study. In particular, we show that \mathcal{LM}_τ is not consistent and the outcome of the test in finite samples depends on the choice of τ . Moreover, to our knowledge there is no theoretical (or practical) guidance on how to choose τ .

3. ASSUMPTIONS AND STATISTICAL PROPERTIES OF \mathcal{LM} AND \mathcal{LM}^* UNDER THE NULL

Let's introduce the following regularity assumptions.

A1: Set $\nu_t = \Sigma^{1/2}(x_t) u_t$ and $\tilde{P}(x_t) = P(x_t) \otimes \Sigma^{1/2}(x_t)$. The sequences $\{\tilde{P}(x_t)\}_{t \in \mathbb{Z}}$ and $\{u_t\}_{t \in \mathbb{Z}}$ are L^4 -NED of size $\iota > 2$. That is,

$$\begin{aligned} \left\| \tilde{P}(x_t) - \mathbb{E} \left(\tilde{P}(x_t) \mid \varepsilon_t, \dots, \varepsilon_{t-m} \right) \right\|_4 &= O(m^{-\iota}) \\ \left\| u_t - \mathbb{E} (u_t \mid \eta_t, \dots, \eta_{t-m}) \right\|_4 &= O(m^{-\iota}), \end{aligned}$$

where $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ and $\{\eta_t\}_{t \in \mathbb{Z}}$ are two zero mean sequences of independent random vectors such that $\mathbb{E}(\eta_t \eta_t') = \Sigma_\eta$ and $\mathbb{E}(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon > 0$. Furthermore, (u_t', η_t') is independent of (ε_s', x_s') for all t and s .

Herewith $\|\cdot\|$ denotes the Euclidean norm of a vector or a matrix and for a generic random vector w_t , $\|w_t\|_r^r = \mathbb{E} \|w_t\|_r^r$.

A2: (i) $\sup_t \|g_t(\theta_0)\|_r < \infty$ for some $r > 2$, and for Δ given in (2.5)

$$\Delta_{[\tau_1 n], [\tau_2 n]} =: \text{var} \left(([\tau_2 n] - [\tau_1 n])^{-1/2} \sum_{t=[\tau_1 n]+1}^{[\tau_2 n]} g_t(\theta_0) \right) \rightarrow \Delta.$$

(ii) $g_t(\theta)$ is continuously differentiable with probability one, $\mathbb{E} \left[\frac{\partial}{\partial \theta'} g_t(\theta_0) \right]$ is full rank and $\mathbb{E} \left[\sup_{\theta \in \Theta} \left\| \frac{\partial g_t(\theta)}{\partial \theta'} \right\|^{1+\epsilon} \right] < \infty$ for some $\epsilon > 0$.

(iii) $\frac{\partial}{\partial \theta'} g_t(\theta)$ is second order continuous, that is

$$\lim_{\delta \downarrow 0} \mathbb{E} \sup_{|\theta_1 - \theta_2| < \delta} \left\{ \left\| \frac{\partial}{\partial \theta'} g_t(\theta_1) - \frac{\partial}{\partial \theta'} g_t(\theta_2) \right\|^2 \right\} = 0.$$

Assumptions A1 and A2 allow for very general types of dependence and appear to be quite minimal. Assumption A2 (i) does not require that the sequence $g_t(\theta_0)$ be weakly stationary but only that it be asymptotically so and it yields a FCLT with A1, see Andrews (1993).

An important building block to derive the asymptotic distribution of \mathcal{LM} and \mathcal{LM}^* under H_0 is a strong approximation for partial sums of $g_t(\theta_0)$.

Proposition 1. *Assume A1 and A2 (i), and let $\Delta_{m,k}$ as defined in (??). Then, for each n and $m < n$ we can construct on a probability space a sequence of independent and identically distributed NM -dimensional standard normal random vectors \mathbf{n}_t such that for some $1 < \zeta < 2$ and $C > 0$,*

$$(3.1) \quad \Pr \left\{ \sup_{m \leq k \leq n} \left\| \sum_{t=m}^k g_t(\theta_0) - \Delta_{m,k}^{1/2} \sum_{t=m}^k \mathbf{n}_t \right\| > \epsilon \right\} \leq C \epsilon^{-4} (n-m)^{\frac{\zeta+2}{(\zeta+1)}}.$$

Proposition 1 extends previous strong approximation results. For instance, Kuelbs and Philipp (1980) imposes strong mixing condition and is subject to the Ling (2007)' critique regarding the backward sum. On the other hand, Ling relaxes the mixing condition but add the martingale difference assumption on g_t , developing a strong approximation results for the backward sum as well as the forward sum. We do not impose a mixing condition nor a martingale difference assumption, thus allowing for general serial dependence in g_t . This is made plausible by articulating a product of two sigma fields. Also, the backward sum issue discussed in Ling is resolved as a direct consequence of this proposition, see Corollary 1 in the Appendix.

Now we present the main result of this section.

Theorem 1. *Assuming A1 and A2, we have that under H_0 ,*

$$\Pr \{ \mathcal{LM}^* \leq x \} \rightarrow \exp(-2e^{-x}) \quad \text{and} \quad \Pr \{ \mathcal{LM} \leq x \} \rightarrow \exp(-2e^{-x/2}).$$

The limit distributions are (double) Gumble distributions, which are also known as extreme value distributions. The asymptotic critical values can be easily computable from the given distribution functions.

4. ASYMPTOTIC POWER OF \mathcal{LM}

We now examine the behavior of \mathcal{LM} under fixed and local alternatives, being the results for \mathcal{LM}^* similar to those of \mathcal{LM} . To that end, consider a sequence of alternatives

$$(4.1) \quad H_{1n} : \theta_{nt} = \theta_0 + \delta_n \mathbf{1}(t < s_0),$$

where δ_n is a p -dimensional column vector and $p < s_0 < n - p$ and $\mathbf{1}(\cdot)$ denotes the indicator function. For simplicity, we have chosen the one-time break model, although the (qualitative) conclusions hold true under more general types of breaks such as gradual or multiple ones.

First we consider the case of fixed alternatives. To that end, let's introduce,

A3: Under H_{1n} with $\delta_n = \delta \neq 0$, there exists θ_* such that $\hat{\theta} \xrightarrow{p} \theta_*$ and $\theta_* \neq \theta_0$ and A2 holds true with $g_t(\theta_0)$ replaced by $g_t(\theta_*) - \mathbb{E}g_t(\theta_*)$, where $\mathbb{E}g_t(\theta_*) \neq 0$.

Theorem 2. *Assume H_{1n} , A1 and A3. Then, for any $x \in \mathbb{R}$,*

$$\Pr \{ \mathcal{LM} \leq x \} \rightarrow 0,$$

provided that $h_n \leq s_0 \leq n - h_n$, where $h_n^{-1} \log_2 n = o(1)$.

Theorem 2 indicates that our results are in line with those found in Csörgo and Horváth (1997) who noted that the rate for h_n was necessary for the consistency of the test. The latter means that in practice \mathcal{LM} is consistent irrespective of the location of the break since $\log_2 n < p$ for the typical samples sizes n that we encounter in real examples. Notice that, for instance when $p = 3$, $\log_2 n > p$ if $n > 53 \times 10^7$, which is a sample size that we do not find even with financial data.

On the other hand, \mathcal{LM}_τ is not consistent. Indeed, as we can observe in the proof of Theorem 3 below, $\Lambda_n(s)$ has a centrality parameter (function), apart from multiplicative constants different than zero, equal to

$$(4.2) \quad \Lambda_n(s; s_0) = s_0 \left(\frac{n-s}{ns} \right)^{1/2} \mathbf{1}(s_0 \leq s) + (n-s_0) \left(\frac{s}{n(n-s)} \right)^{1/2} \mathbf{1}(s < s_0).$$

When $s_0 \in \mathfrak{S}(\tau)$, $\Lambda_n(s_0; s_0) = O(n^{1/2})$, which implies that \mathcal{LM}_τ is a consistent test when the break is in the middle of the sample. On the other hand, when $s_0 < [\tau n]$ or $n - [\tau n] < s_0$, this is not always the case. Indeed consider the case when $s_0 < [\tau n]$. Because $s_0 < s$ as $s \in \mathfrak{S}(\tau)$, the right side of (4.2) is

$$s_0 \left(\frac{n-s}{ns} \right)^{1/2} \mathbf{1}(s_0 \leq s) = O\left(\frac{s_0}{n^{1/2}} \right).$$

So if $s_0 = o(n^{1/2})$, the last displayed expression converges to zero uniformly in $s \in \mathfrak{S}(\tau)$. The latter implies that \mathcal{LM}_τ has the same asymptotic distribution as under H_0 . By symmetry, it is evident that we can draw the same conclusions when $n - s_0 = o(n^{1/2})$. Hence \mathcal{LM}_τ is inconsistent when $s_0 = o(n^{1/2})$ or $n - s_0 = o(n^{1/2})$, whereas \mathcal{LM} (and \mathcal{LM}^*) is still consistent in those regions.

We now examine the behaviour of our tests under local alternatives.

Theorem 3. *Assuming A1 and A2, under (2.1) and (4.1), we have that*

$$(a) \quad \Pr \{ \mathcal{LM} \leq x \} \rightarrow 0, \quad \text{if } (s_0 / \log_2 n)^{1/2} \delta_n \rightarrow \infty,$$

$$(b) \quad \Pr \{ \mathcal{LM} \leq x \} \rightarrow \exp(-2e^{-x/2}), \quad \text{if } (s_0 / \log_2 n)^{1/2} \delta_n \rightarrow 0.$$

Theorem 3 shows that as long as $\|\delta_n\|$ shrinks at a rate slower than $(s_0 / \log n)^{1/2}$, \mathcal{LM} rejects the null hypothesis with probability 1 as $n \nearrow \infty$, indicating that the test has non-trivial power against local alternatives H_{1n} such that

$$(4.3) \quad \delta_n = \delta (\log_2 n / s_0)^{1/2},$$

for some $\delta \neq 0$. This is in clear contrast with tests based on \mathcal{LM}_τ . Indeed, it can be shown that they have non-trivial power against local alternatives H_{1n} such that

$$(4.4) \quad \delta_n = \delta n^{1/2} / s_0.$$

So, when the break occurs in the “middle” of the sample, i.e. $s_0 \in \mathfrak{S}(\tau)$, we have the standard non-trivial power against local alternatives of order $O(n^{-1/2})$. On the other hand, when $s_0 < [n\tau]$ as $s \geq [n\tau]$, the noncentrality parameter in (4.2) satisfies

$$C^{-1} \left(s_0 / n^{1/2} \right) \|\delta_n\| \leq \|\Lambda_n(s; s_0)\| \leq C \left(s_0 / n^{1/2} \right) \|\delta_n\|,$$

which in turn yields the rate given in (4.4). The same argument applies to the case where $s_0 > n - [n\tau]$.

Comparison of (4.3) with (4.4) not only confirms our previous finding that \mathcal{LM}_τ is not consistent when s_0 , or $n-s_0$, grows slower than $n^{1/2}$, but also shows that \mathcal{LM}_τ has zero relative efficiency compared to ours when $s_0 = o\left(n/\log_2^{1/2} n\right)$. On the other hand when $s_0 \in \mathfrak{S}(\tau)$, \mathcal{LM}_τ is more efficient by a factor of $\log_2^{1/2} n$, however this loss appears not to be relevant in finite samples as next section illustrates.

5. MONTE CARLO EXPERIMENT

All throughout the experiment we have considered the following model:

$$(5.1) \quad y_t = \theta' z_t + \delta' z_t \mathbf{1}(t > s) + u_t, \quad t = 1, \dots, n$$

where z_t is either $1, x_t$, or $(1, x_t)'$, and

$$(5.2) \quad \begin{aligned} x_t &= \rho_1 x_{t-1} + e_{1t}, \\ u_t &= \rho_2 u_{t-1} + e_{2t}, \quad t = 1, \dots, n. \end{aligned}$$

The variables $\{e_{1t}\}_{t=1}^n$ and $\{e_{2t}\}_{t=1}^n$ were generated as two mutually independent sequences of independent standard normal random variables. We have set $\rho_1 = \rho_2 = \rho = 0, \pm 0.5, \pm 0.9$, and the sample sizes were $n = 100, 250, 500$ and 1000 . We considered several scenarios for the time of the break s_0 and the size of change δ as explained below.

To estimate θ , we set $g_t(\theta) = z_t(y_t - \theta' z_t)$, that is, $\hat{\theta}$ is the least square estimator. In the first experiment, we set $\rho = 0$. Then, $\mathcal{LM}(s)$ is given by (2.6) with $G_n(\theta) = \frac{1}{n} \sum_{t=1}^n z_t z_t'$ and $\Delta(\hat{\theta}) = \hat{\sigma}^2 G_n(\hat{\theta})$, where $\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{\theta}' z_t)^2$.

FIGURE 1 and TABLE 1 ABOUT HERE

First we compare the sizes of \mathcal{LM}^* with \mathcal{LM} in finite samples. Table 1 reports the rejection frequencies of both tests at the 5% significance level under various scenarios along with \mathcal{LM}_τ for $\tau = 0.05$ and 0.15 , Brown et al.'s (1975) CUSUM and the Chow test with $s_0 = n/2$. The test \mathcal{LM}^* exhibits severe size distortions at this significance level, while the test \mathcal{LM} performs reasonably well overall. The remaining tests appear to be more conservative than the \mathcal{LM} . Figure 1 depicts the qq plots of both \mathcal{LM} and \mathcal{LM}^* statistics for $n = 100$ and $n = 1000$, when $z_t = 1$. The other specifications of z_t yield similar results and thus not reported here. Similar to the results reported in Davis et al. (1995), the \mathcal{LM} follows more closely the 45 degree line than the \mathcal{LM}^* .

Because of the size distortions of the \mathcal{LM}^* test, especially when compared with \mathcal{LM} , we have decided to exclude it when comparing the power of the different tests. To examine the power, we have explored two situations depending on the time of the break s_0 . In the first one, the break lies in the ‘‘middle’’ of the sample, that is $s_0 = [\pi n]$ for $\pi = 1/2, 3/4, 9/10$ and $19/20$, whereas in the second scenario, we have considered the break towards the end of the sample, i.e. $s_0 = n - \pi n^{1/2}$, with $\pi = 1, 2/3, 1/2$ and $1/3$. The parameter δ is set as 0.1 or 0.5 (and $(0.1, 0.1)'$ or $(0.5, 0.5)'$ in case $z_t = (1, x_t)'$). Tables 2a-2c and 3a-3c present the outcomes from the two scenarios.

TABLES 2 AND 3 ABOUT HERE

We now comment on Table 2 and 3. These two tables suggest that the power function is independent of z_t and it increases with the size of the break, as one would expect that the power should increase as the alternative hypothesis becomes far apart from the null. More specifically, Table 2 suggests that our test does not perform worse than \mathcal{LM}_τ when the break is in the “middle” and the power of \mathcal{LM} is comparable to that of $\mathcal{LM}_{.05}$ for all “ n ” and values of π . However, we notice that when the break is towards the end of the sample, \mathcal{LM} performs better than $\mathcal{LM}_{.05}$. On the other hand, \mathcal{LM} seems to perform much better than $\mathcal{LM}_{.15}$ even with moderate sample sizes when $\pi \geq 9/10$, being the deterioration of the power of the latter even bigger when $\pi = 19/20$. Observe that when $\pi = 9/10$, $s_0 \notin \mathfrak{S}(.15)$. The latter might indicate that, in general, \mathcal{LM}_τ is not a very useful statistic to detect a break when $s_0 \notin \mathfrak{S}(\tau)$.

Next, Table 3 suggests that \mathcal{LM} outperforms $\mathcal{LM}_{.15}$ and also $\mathcal{LM}_{.05}$ when the break is toward the “end” of the sample. Notice also that the power function is smaller with $\mathcal{LM}_{.15}$ than with $\mathcal{LM}_{.05}$. But more importantly, as we can expect from the results of Section 4, the power function of \mathcal{LM}_τ remains constant with the sample size, so that it does not converges to 1. On the other hand, the power of \mathcal{LM} increases with the sample size which corroborates the consistency of \mathcal{LM} even when the break occurs towards the end (or beginning) of the sample. So, the main conclusion that we could draw from Table 2 and 3 is that the \mathcal{LM} appears to be more desirable than \mathcal{LM}_τ , not only because its power behaviour appears to be superior but also as we do not need to choose τ to compute the test, avoiding the unpleasant feature that depending on the choice of τ we might obtain different conclusions with the same data set.

In the second set of experiments, we consider $\rho = 0.5$ and 0.9 in (5.2), allowing for dependence in the data. The results with negative ρ 's are similar to the positive ones and thus they are not showing here. Also, we only report the case where $z_t = x_t$ in (5.1). The serial correlation of $x_t u_t$ entails the HAC estimation of Δ . As $\mathbb{E}(x_t) = 0$ and x_t and u_t are independent, we have employed Robinson's (1998) estimator of Δ , namely

$$\hat{\Delta} = \hat{\gamma}^x(0) \hat{\gamma}^u(0) + \sum_{j=1}^{n-1} (1 - j/n) (\hat{\gamma}^x(j) + \hat{\gamma}^{x'}(j)) \hat{\gamma}^u(j),$$

where $\hat{\gamma}^x(j) = \frac{1}{n} \sum_{t=1}^{n-j} x_t x_{t+j}$ and $\hat{\gamma}^u(j) = \frac{1}{n} \sum_{t=1}^{n-j} \hat{u}_t \hat{u}_{t+j}$ with $\hat{u}_t = y_t - \hat{\theta}' x_t$. The latter avoids the need to choose the bandwidth parameter m . The other variables are identically defined as in the first experiment.

TABLE 4 ABOUT HERE

Table 4 reports the size property of \mathcal{LM} and \mathcal{LM}^* compared with \mathcal{LM}_τ . As in Table 1, \mathcal{LM}^* exhibits rather large size distortion compared to the other tests, so that as before we would not report the power for the \mathcal{LM}^* test. Noting this, we compare power properties of \mathcal{LM} , $\mathcal{LM}_{.05}$ and $\mathcal{LM}_{.15}$ and the results are reported in Table 5 for the middle-of-sample break case and Table 6 for the end-of-sample break case. As the tables indicate the results are qualitatively similar to those in Table 2 and 3. While the test \mathcal{LM} is not outperformed much by \mathcal{LM}_τ when the break occurs in the middle of the sample, it demonstrates significant power gains

when the break occurs in the end of the sample, that is when the trimmed set does not contain the true break point. For instance, look at the last column in Table 5 where the power of $\mathcal{LM}_{.15}$ drops dramatically as the break point move from 750 to 900 and then 950 relative to the other tests. Also Table 6 shows that the power of $\mathcal{LM}_{.15}$ or $\mathcal{LM}_{.05}$ does not increase much as the sample size increases, as expected from the discussion of Section 3.

TABLES 5 AND 6 ABOUT HERE

6. PROOFS OF THEOREMS

6.1. Proof of Proposition 1.

Recall that $g_t(\theta_0) = P(x_t) \otimes \nu_t = \tilde{P}_t \otimes u_t$ with $\tilde{P}_t = P(x_t) \otimes \Sigma^{1/2}(x_t)$. Herewith $\tilde{t} = t - Ct^{1/q}$ with $q = 6l - 3$ and we denote $v_t = \tilde{P}_t \otimes u_t$ and $\dot{v}_t(s) = \dot{P}_t(s) \otimes \dot{u}_t(s)$, where $\dot{u}_t(s) = \mathbb{E}(u_t | \eta_t, \dots, \eta_s)$ and $\dot{P}_t(s) = \mathbb{E}(\tilde{P}_t | \varepsilon_t, \dots, \varepsilon_s)$ for some $s < t$. Note that $\dot{u}_t(\tilde{t})$ is independent of $\dot{u}_r(\tilde{r})$ if $\tilde{t} > r$. Denoting $\ddot{v}_t(s) = v_t - \dot{v}_t(s)$, $\ddot{u}_t(s) = u_t - \dot{u}_t(s)$ and $\ddot{P}_t(s) = \tilde{P}_t - \dot{P}_t(s)$, we have that

$$(6.1) \quad \ddot{v}_t(s) = \tilde{P}_t \otimes \ddot{u}_t(s) + \ddot{P}_t(s) \otimes u_t - \dot{P}_t(s) \otimes \dot{u}_t(s).$$

For notational simplicity, we will assume that $n - m =: \dot{n} = 2^d$ for some integer d . We shall first show that

$$(6.2) \quad \left(\mathbb{E} \sup_{m \leq k \leq n} \left\| \sum_{t=m}^k \ddot{v}_t(\tilde{t}) \right\|^4 \right)^{1/4} = o\left(\dot{n}^{\frac{\zeta+2}{4(\zeta+1)}}\right).$$

By the triangular inequality, it suffices to show (6.2) with $\ddot{v}_t(\tilde{t})$ replaced by each of the terms on the right of (6.1). We shall only handle explicitly the contribution due to $\tilde{P}_t \otimes \ddot{u}_t(s)$ into the left of (6.2) being for the other terms similarly handled.

Now because $\sup_t \mathbb{E} \|\tilde{P}_t\|^4 < \infty$ and that x_t and u_t are independent by A1, Wu's (2007) Proposition 1 implies that

$$(6.3) \quad \left(\mathbb{E} \sup_{m \leq k \leq n} \left\| \sum_{t=m}^k \tilde{P}_t \otimes \ddot{u}_t(\tilde{t}) \right\|^4 \right)^{1/4} \leq C \sum_{j=0}^d \left[\sum_{r=1}^{2^{d-j}} \mathbb{E} \left\| \sum_{t=2^j(r-1)+m}^{2^j r} \ddot{u}_t(\tilde{t}) \right\|^4 \right]^{1/4}.$$

Now abbreviating $\sum_{t=2^j(r-1)+m}^{2^j r}$ by \sum'_t , $\mathbb{E} \|\sum'_t \ddot{u}_t(\tilde{t})\|^4$ is bounded by

$$(6.4) \quad \begin{aligned} & 3 \left(\left\{ \sum'_{t_1 < \tilde{t}_2 \leq t_2} + \sum'_{\tilde{t}_2 \leq t_1 \leq t_2} \right\} \left| \mathbb{E} \left(\ddot{u}_{t_1}(\tilde{t}_1)' \ddot{u}_{t_2}(\tilde{t}_2) \right) \right| \right)^2 \\ & \left\{ \sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3 \leq t_4} + \sum'_{t_1 < \tilde{t}_3 \leq t_2 < t_3 \leq t_4} \right\} \left\| \text{cum}(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(\tilde{t}_3), \ddot{u}_{t_4}(\tilde{t}_4)) \right\| \\ & = C \left(\sum'_t t^{-1/3} \right)^2 := a_r \end{aligned}$$

as we now show, and where we have abbreviated \cdot . Because for $t_1 < \tilde{t}_2$, A1 implies that $\mathbb{E} \left(\ddot{u}_{t_1}(\tilde{t}_1)' (\ddot{u}_{t_2}(t_1) - \ddot{u}_{t_2}(\tilde{t}_2)) \right) = 0$ as $\ddot{u}_{t_2}(t_1) - \ddot{u}_{t_2}(\tilde{t}_2)$ only depends on $\{\varepsilon_s\}_{s=t_2}^{t_1+1}$, we have that $\mathbb{E} \left(\ddot{u}_{t_1}(\tilde{t}_1)' \ddot{u}_{t_2}(\tilde{t}_2) \right) = \mathbb{E} \left(\ddot{u}_{t_1}(\tilde{t}_1)' \ddot{u}_{t_2}(t_1) \right)$, and thus because $\left| \mathbb{E} \left(\ddot{u}_{t_1}(\tilde{t}_1)' \ddot{u}_{t_2}(t_1) \right) \right| \leq C t_1^{-\nu/q} (t_2 - t_1)^{-\nu}$ by A1, the first term on the left of (6.4) is bounded by

$$\begin{aligned} & C \left(\sum'_{t_1 < \tilde{t}_2 \leq t_2} t_1^{-\nu/q} (t_2 - t_1)^{-\nu} \right)^2 + C \left(\sum'_{\tilde{t}_2 \leq t_1 \leq t_2} t_1^{-\nu/q} t_2^{-\nu/q} \right)^2 \\ & \leq C \left(\sum'_{t_1 < \tilde{t}_2 \leq t_2} t_1^{-(2\nu-1)/q} (t_2 - t_1)^{-1} \right)^2 + C \left(\sum'_{t_2} t_2^{-(2\nu-1)/q} \right)^2 = a_r \end{aligned}$$

as $q = 6\nu - 3$ and $\tilde{t} = t - C t^{1/q}$.

It remains to show that the second on the left of (6.4) is also a_r . By standard algebra, the contribution due to $\sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3 \leq t_4}$ is bounded by

$$\begin{aligned} & \sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3 \leq t_4} \left\{ \left\| \text{cum} \left(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(t_2), \ddot{u}_{t_4}(t_2) \right) \right\| \right. \\ (6.5) \quad & + \left\| \text{cum} \left(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(t_2) - \ddot{u}_{t_3}(\tilde{t}_3), \ddot{u}_{t_4}(\tilde{t}_4) \right) \right\| \\ & \left. + \left\| \text{cum} \left(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(t_2), \ddot{u}_{t_4}(t_2) - \ddot{u}_{t_4}(\tilde{t}_4) \right) \right\| \right\}. \end{aligned}$$

Recalling that

$$(6.6) \quad \left| \text{cum}(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \right| \leq \prod_{j=1}^4 \mathbb{E}^{1/4}(\epsilon_j^4),$$

A1 implies that the first term of (6.5) is bounded by

$$\begin{aligned} & C \sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3 \leq t_4} \frac{t_1^{-\nu/q} t_2^{-\nu/q}}{(t_3 - t_2)^\nu (t_4 - t_2)^\nu} \leq C \sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3} \frac{t_1^{-\nu/q} t_2^{-\nu/q}}{(t_3 - t_2)^{(2\nu-1)}} \\ & \leq C \sum'_{t_1 \leq t_2 < t_3} \frac{t_1^{-(2\nu-1)/q} t_2^{-(2\nu-1)/q}}{t_3 - t_2} = a_r \end{aligned}$$

because $t_2 < \tilde{t}_3$ implies that $t_3 - t_2 \geq t_3^{1/q} \geq t_2^{1/q} \geq t_1^{1/q}$. Using that if a_1 is independent of a_2, a_3, a_4 , $\text{cum}(a_1, a_2, a_3, a_4) = 0$, when $t_3 < \tilde{t}_4$

$$\begin{aligned} & \text{cum} \left(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(t_2), \ddot{u}_{t_4}(t_2) - \ddot{u}_{t_4}(\tilde{t}_4) \right) \\ & = \text{cum} \left(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(t_2), \ddot{u}_{t_4}(t_2) - \ddot{u}_{t_4}(t_3) \right), \end{aligned}$$

so that by standard algebra, the third term of (6.5) is

$$\begin{aligned} & C \sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3 < \tilde{t}_4 < t_4} \frac{t_1^{-\nu/q} t_2^{-\nu/q}}{(t_3 - t_2)^\nu (t_4 - t_3)^\nu} + C \sum'_{t_1 \leq t_2 < \tilde{t}_3 < \tilde{t}_4 < t_3 \leq t_4} \frac{t_1^{-\nu/q} t_2^{-\nu/q} t_4^{-\nu/q}}{(t_3 - t_2)^\nu} \\ & = a_r + C \sum'_{t_1 \leq t_2 < \tilde{t}_3 < \tilde{t}_4 < t_3 \leq t_4} \frac{t_1^{-\nu/q} t_2^{-\nu/q} t_4^{-(\nu-1)/q}}{(t_3 - t_2)^\nu (t_4 - t_3)} = a_r, \end{aligned}$$

after observing that $\tilde{t}_4 < t_3$ implies that $t_4^{-1/q} \leq (t_4 - t_3)^{-1}$.

Because $\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2)$ is independent of $\ddot{u}_{t_3}(t_2) - \ddot{u}_{t_3}(\tilde{t}_3), \ddot{u}_{t_4}(\tilde{t}_4) - \ddot{u}_{t_4}(t_2)$, the second term of (6.5) is

$$\begin{aligned} & \sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3 \leq t_4} \left\| cum(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(t_2) - \ddot{u}_{t_3}(\tilde{t}_3), \ddot{u}_{t_4}(t_2)) \right\| \\ & \leq C \sum'_{t_1 \leq t_2 < \tilde{t}_3 < t_3 \leq t_4} \frac{t_1^{-\nu/q} t_2^{-\nu/q} t_3^{-\nu/q}}{(t_4 - t_2)^\nu} = a_r \end{aligned}$$

after routine algebra.

Finally, we examine the contribution due to $\sum'_{t_1 < \tilde{t}_3 \leq t_2 < t_3 \leq t_4}$ of the second term on the left of (6.4), which is

$$(6.7) \quad \left\{ \sum'_{t_1 < \tilde{t}_3 < t_2 \leq t_3 < \tilde{t}_4 < t_4} + \sum'_{t_1 < \tilde{t}_3 < t_2 < \tilde{t}_4 < t_3 \leq t_4} \right\} \left\| cum(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(\tilde{t}_3), \ddot{u}_{t_4}(\tilde{t}_4)) \right\|.$$

The first term of (6.7) is bounded by

$$C \sum'_{t_1 < \tilde{t}_3 < t_2 \leq t_3 < \tilde{t}_4 < t_4} \frac{t_1^{-\nu/q} t_2^{-\nu/q} t_3^{-\nu/q}}{(t_4 - t_3)^\nu} \leq C \sum'_{t_1 < \tilde{t}_3 < t_2 \leq t_3 < \tilde{t}_4 < t_4} \frac{t_1^{-(2\nu-1)/q} t_2^{-(2\nu-1)/q}}{(t_4 - t_3)(t_3 - t_2)} = a_r$$

because if $\tilde{t}_3 < t_2$, we have that $t_3 - t_2 \leq t_3^{1/q}$, and $t_3 < \tilde{t}_4$ implies that

$$cum(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(\tilde{t}_3), \ddot{u}_{t_4}(\tilde{t}_4)) = cum(\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(\tilde{t}_3), \ddot{u}_{t_4}(t_3)),$$

because $\ddot{u}_{t_4}(\tilde{t}_4) - \ddot{u}_{t_4}(t_3)$ is independent of $\ddot{u}_{t_1}(\tilde{t}_1), \ddot{u}_{t_2}(\tilde{t}_2), \ddot{u}_{t_3}(\tilde{t}_3)$, whereas the second term of (6.7) is bounded by

$$C \sum'_{t_1 < \tilde{t}_3 < t_2 < \tilde{t}_4 < t_3 \leq t_4} t_1^{-\nu/q} t_2^{-\nu/q} t_3^{-\nu/q} t_4^{-\nu/q} \leq C \sum'_{t_1 \leq t_2 \leq t_3 \leq t_4} \frac{t_1^{-(2\nu-1)/q} t_2^{-(2\nu-1)/q}}{(t_3 - t_2)(t_4 - t_3)} = a_r$$

by similar arguments to those used for the first term. Thus, $\mathbb{E} \left\| \sum_t \ddot{u}_t(\tilde{t}) \right\|^4 = a_r$, which implies that the left side of (6.2) is $O(n^{1/3} \log n) = o\left(n^{\frac{\zeta+2}{4(\zeta+1)}}\right)$ because $\zeta < 2$.

To show that (3.1) holds for $\{\dot{v}_t\}_{t \in \mathbb{Z}}$, we employ standard blocking arguments. To that end, for $\ell \geq 2$ let $n_1 = m$, and $n_\ell = n_{\ell-1} + \ell^{1/\zeta} + \ell^{1/q}$ and consider blocks $A_\ell = \{t : n_{\ell-1} < t \leq n_{\ell-1} + \ell^{1/\zeta}\}$ and $B_\ell = \{t : n_{\ell-1} + \ell^{1/\zeta} < t \leq n_\ell\}$. Define

$$\xi_\ell = \sum_{t \in A_\ell} \dot{v}_t; \quad \epsilon_\ell = \sum_{t \in B_\ell} \dot{v}_t.$$

Observe that $\sum_{t=m}^{n_\ell} \dot{v}_t = \sum_{j=1}^{\ell} (\xi_j + \epsilon_j)$. Let $\bar{\ell}$ be the smallest integer such that $n_{\bar{\ell}-1} < n \leq n_{\bar{\ell}}$, so that $C^{-1} \leq \bar{\ell} n^{-\zeta/(\zeta+1)} \leq C$. We first show that we can find a sequence of *iid* standard normal random vectors $\{\mathbf{n}_j\}_{j \geq 1}$ such that

$$(6.8) \quad \Pr \left\{ \sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{t=m}^{n_\ell} \dot{v}_t - \Delta_{m, n_\ell}^{1/2} \sum_{j=1}^{\ell} \mathbf{n}_j \right\| > z \right\} \leq C z^{-4} \bar{n}^{\frac{\zeta+2}{\zeta+1}}.$$

By construction and A1, $\{\xi_\ell\}_{\ell \geq 1}$ is a sequence of independent random variables with finite 4th moments. So, by Götze and Zaitsev's (2007) Theorem 4 (Proposition 1),

we can find a sequence of *iid* normal random variables $\{\mathbf{n}_j\}_{j \geq 1}$ such that

$$\begin{aligned} \Pr \left\{ \sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{j=1}^{\ell} \xi_j - \sum_{j=1}^{\ell} \mathbb{E}^{1/2}(\xi_j \xi_j') \mathbf{n}_j \right\| > z \right\} &\leq Cz^{-4} \sum_{j=1}^{\bar{\ell}} \mathbb{E} \|\xi_j\|^4 \\ &= Cz^{-4} \bar{\ell}^{\frac{\zeta+2}{\zeta+1}} \end{aligned}$$

because $\mathbb{E} \|\xi_j\|^4 = O(j^{2/\zeta})$ and $\bar{\ell} \leq C\bar{n}^{\zeta/(\zeta+1)}$. Next, because $\{\epsilon_j\}_{j \geq 1}$ is a sequence of independent random variables with finite fourth moments, by the law of iterated logarithms, $\overline{\lim}_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{j=1}^{\ell} \epsilon_j / \ell^{(q+1)/2q} \log_2 \ell \right\| = 1$ a.s., see for instance Shao's (1995) Theorem 3.2. So, $\sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{j=1}^{\ell} \epsilon_j \right\| = o\left(\bar{n}^{\frac{\zeta+2}{4(\zeta+1)}}\right)$ a.s. and we conclude that

$$(6.9) \quad \Pr \left\{ \sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{t=m}^{n_{\ell}} \dot{v}_t - \sum_{j=1}^{\ell} \mathbb{E}^{1/2}(\xi_j \xi_j') \mathbf{n}_j \right\| > z \right\} \leq Cz^{-4} \bar{n}^{\frac{\zeta+2}{\zeta+1}}.$$

Denote $\mathbb{E} \left(\sum_{t \in A_j} v_t \sum_{t \in A_j} v_t' \right)$ by $\mathbb{E} \left(\sum_{t \in A_j} v_t \right)^2$. Because proceeding as with the proof of (6.4),

$$\left| \mathbb{E}(\xi_j \xi_j') - \mathbb{E} \left(\sum_{t \in A_j} v_t \right)^2 \right| \leq C \log j \sum_{t=n_{j-1}+1}^{n_{j-1}+j^{1/\zeta}} t^{-1/3} \leq C j^{(2-\zeta)/3\zeta} \log j$$

as $n_j = \sum_{h=1}^j (h^{1/\zeta} + h^{1/q}) \leq C j^{(\zeta+1)/\zeta}$, we have that

$$\Pr \left\{ \sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{j=1}^{\ell} \left\{ \mathbb{E}^{1/2}(\xi_j \xi_j') - \mathbb{E}^{1/2} \left(\sum_{t \in A_j} v_t \right)^2 \right\} \mathbf{n}_j \right\| > z \right\} \leq Cz^{-4} \bar{n}^{\frac{\zeta+2}{\zeta+1}}$$

using that $(a-b)^2 \leq a^2 - b^2$ for $a > b > 0$ and Levy's inequality, which also implies that $\Pr \left\{ \sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{j=1}^{\ell} \mathbb{E}^{1/2}(\epsilon_j \epsilon_j') \mathbf{n}_j \right\| > z \right\} \leq Cz^{-4} \bar{n}^{\frac{\zeta+2}{\zeta+1}}$. So, we conclude that

in (6.9) we can replace $\mathbb{E}^{1/2}(\xi_j \xi_j')$ by $\left(\mathbb{E} \left(\sum_{t \in A_j} v_t \right)^2 + \mathbb{E} \left(\sum_{t \in B_j} v_t \right)^2 \right)^{1/2}$ and standard arguments imply that

$$\Pr \left\{ \sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{t=m}^{n_{\ell}} \dot{v}_t - \sum_{j=1}^{\ell} \left(\mathbb{E} \left(\sum_{t \in A_j \cup B_j} v_t \right)^2 \right)^{1/2} \mathbf{n}_j \right\| > z \right\} \leq Cz^{-4} \bar{n}^{\frac{\zeta+2}{\zeta+1}}.$$

However because A1 implies that $\mathbb{E} \left(\sum_{t \in A_j \cup B_j} v_t \right)^2 = \Delta_{m, j^{1/q} + j^{1/\zeta}}$ and $\|\Delta_{m, n_j} - \Delta_{m, j^{1/q} + j^{1/\zeta}}\| = O(j^{-1/2\zeta})$, by Levy's inequality again,

$$\begin{aligned} &\Pr \left\{ \sup_{1 \leq \ell \leq \bar{\ell}} \left\| \sum_{j=1}^{\ell} \left\{ \left(\mathbb{E} \left(\sum_{t \in A_j \cup B_j} v_t \right)^2 \right)^{1/2} - \Delta_{m, n_{\ell}}^{1/2} \left(j^{1/q} + j^{1/\zeta} \right)^{1/2} \right\} \mathbf{n}_j \right\| > z \right\} \\ &\leq Cz^{-4} \bar{n}^{\frac{\zeta+2}{\zeta+1}} \end{aligned}$$

and because $n_j = \sum_{h=1}^j (h^{1/\zeta} + h^{1/q})$, we obtain that

$$\Delta_{m, n_\ell}^{1/2} \sum_{j=1}^{\ell} \left(j^{1/q} + j^{1/\zeta} \right)^{1/2} \mathbf{n}_j \stackrel{d}{=} \Delta_{m, n_\ell}^{1/2} \sum_{j=1}^{n_\ell} \mathbf{n}_j,$$

where “ $\stackrel{d}{=}$ ” denotes “distributed as”. This concludes the proof of (6.8), and that of (3.1), when the supremum is taken over values of k such that there exists a \bar{k} satisfying $n_{\bar{k}} = k$.

Hence, to finish the proof we need to examine the approximation when $n_\ell < k < n_{\ell+1}$. But by Csörgo and Révész's (1981) Theorem 1.2, we know that

$$\max_{n_{\ell-1} \leq j \leq n_\ell} \left| \sum_{t=n_{\ell-1}}^j \mathbf{n}_t \right| = O \left(\ell^{1/2\zeta} \log_2^{1/2} \ell \right) \quad a.s.$$

and because $\mathbb{E} \left\| \sum_{s \leq t \leq r} v_t \right\|^4 = O \left((r-s)^2 \right)$, the Borel Cantelli's theorem implies that $\max_{n_{\ell-1} \leq j \leq n_\ell} \left\| \sum_{j \leq t \leq n_\ell} v_t \right\| = O \left(\dot{n}^{\frac{\zeta+2}{4(\zeta+1)}} \right)$ a.s.. This concludes the proof. \square

We first introduce two corollaries.

Corollary 1. *Suppose $\Pr \left\{ \sup_{m \leq k \leq n} \left\| \sum_{t=m}^k x_t - \sum_{t=m}^k y_t \right\| > z \right\} \leq C_{n-m}(z)$. Then,*

$$\Pr \left\{ \sup_{m \leq k \leq n} \left\| \sum_{t=k+1}^n x_t - \sum_{t=k+1}^n y_t \right\| > z \right\} \leq C_{n-m}(z/2).$$

Proof. The proof follows after observing that

$$\begin{aligned} \sup_{m \leq k \leq n} \left\| \sum_{t=k+1}^n x_t - \sum_{t=k+1}^n y_t \right\| &= \sup_{m \leq k \leq n} \left\| \sum_{t=m}^n (x_t - y_t) - \sum_{t=m}^k (x_t - y_t) \right\| \\ &\leq 2 \sup_{m \leq k \leq n} \left\| \sum_{t=m}^k (x_t - y_t) \right\|. \end{aligned}$$

\square

Corollary 2. *Let $\{v_t\}_{t \in \mathbb{Z}}$ be a sequence as in Proposition 1. Then, for $\ell = \ell(n)$,*

$$\begin{aligned} (a) \quad \sup_{1 \leq k \leq \ell} \left\| \frac{1}{k^{1/2}} \sum_{t=1}^k v_t \right\| &= O_p \left(\log_2^{1/2} \ell \right) \\ (b) \quad \sup_{n-\ell \leq k < n} \left\| \frac{1}{(n-k)^{1/2}} \sum_{t=k}^n v_t \right\| &= O_p \left(\log_2^{1/2} \ell \right). \end{aligned}$$

Proof. We begin with (a). Because for an *iid* sequence $\{\mathbf{n}_t\}_{t \in \mathbb{Z}}$ of normal random variables, $\sup_{1 \leq k \leq \ell} \left\| k^{-1/2} \sum_{t=1}^k \mathbf{n}_t \right\| = O_p \left(\log_2^{1/2} \ell \right)$, it suffices to show that

$$(6.10) \quad \sup_{1 \leq k \leq \ell} \left\| \frac{1}{k^{1/2}} \sum_{t=1}^k (v_t - \mathbf{n}_t) \right\| = o_p \left(\log_2^{1/2} \ell \right).$$

Take $\ell = 2^d$ for some integer s . Now,

$$\begin{aligned} \Pr \left\{ \sup_{1 \leq k \leq \ell} \left\| \frac{1}{k^{1/2}} \sum_{t=1}^k (v_t - \mathbf{n}_t) \right\| > C \right\} &\leq \Pr \left\{ \sup_{1 \leq c \leq d} \sup_{2^{c-1} \leq k < 2^c} \left\| \frac{1}{k^{1/2}} \sum_{t=1}^k (v_t - \mathbf{n}_t) \right\| > C \right\} \\ &\leq \sum_{c=1}^d \Pr \left\{ \sup_{2^{c-1} \leq k < 2^c} \left\| \frac{1}{k^{1/2}} \sum_{t=1}^k (v_t - \mathbf{n}_t) \right\| > C \right\} \\ &\leq \sum_{c=1}^d \Pr \left\{ \sup_{2^{c-1} \leq k < 2^c} \left\| \sum_{t=1}^k (v_t - \mathbf{n}_t) \right\| > 2^{(c-1)/2} C \right\}. \end{aligned}$$

But now by Proposition 1, we have that the right side of the last displayed inequality is bounded by $C \sum_{c=1}^d 2^{-2c} 2^{3c/2} \leq C$. So, (6.10) holds true. This concludes the proof of part (a) of the corollary.

Part (b) follows similarly. Indeed, as

$$\sup_{n-\ell \leq k < n} \left\| \frac{1}{(n-k)^{1/2}} \sum_{t=k+1}^n \mathbf{n}_t \right\| = O_p \left(\log_2^{1/2} \ell \right),$$

it remains to show that

$$(6.11) \quad \sup_{n-\ell \leq k < n} \left\| \frac{1}{(n-k)^{1/2}} \sum_{t=k+1}^n (v_t - \mathbf{n}_t) \right\| = o_p \left(\log_2^{1/2} \ell \right).$$

However, (6.11) follows by the same arguments for (6.10) using Corollary 1. \square

Let's introduce the following notation.

$$\varkappa_n(s) = \left(\frac{n-s}{ns} \right)^{1/2}; \quad \lambda_n(s) = \left(\frac{n}{(n-s)s} \right)^{1/2}.$$

6.2. Proof of Theorem 1.

Let $G_0 = \mathbb{E} \frac{\partial}{\partial \theta'} g_t(\theta_0)$. Then A2(iii) implies that $G_n(\hat{\theta}) = G_0 + o_p(1)$ as $\hat{\theta} - \theta_0 = O_p(n^{-1/2})$. So using that $\Delta(\hat{\theta}) \rightarrow_p \Delta$ and standard manipulations,

$$(6.12) \quad \lambda_n(s) l_m(s) = \lambda_n(s) \left(G(\theta_0) \Delta^{-1} + o_p(1) \right) \left\{ \sum_{t=1}^s g_t(\theta_0) - \frac{s}{n} \sum_{t=1}^n g_t(\theta_0) \right\}.$$

We first notice that Corollary 2 implies that

$$(6.13) \quad a_n \left\{ \sup_{p < s < \log n} \left\| \lambda_n(s) \left\{ \sum_{t=1}^s g_t(\theta_0) - \frac{s}{n} \sum_{t=1}^n g_t(\theta_0) \right\} \right\| - b_n \xrightarrow{P} -\infty. \right.$$

Therefore,

$$\mathcal{LM}^* = a_n \sup_{\log n \leq s \leq n - \log n} \mathcal{LM}(s)^{1/2} - b_n + o_p(1).$$

However, due to Proposition 1 and Lemma 1, there exists a sequence $\{\mathbf{n}_t\}$ of independent standard normal vectors of dimension p such that

$$\begin{aligned} & \left(G(\theta_0) \Delta^{-1} G'(\theta_0) \right)^{-1/2} G(\theta_0) \Delta^{-1} \frac{1}{s^{1/2}} \left\{ \sum_{t=1}^s g_t(\theta_0) - \frac{s}{n} \sum_{t=1}^n g_t(\theta_0) \right\} \\ &= \frac{1}{s^{1/2}} \left\{ \sum_{t=1}^s \mathbf{n}_t - \frac{s}{n} \sum_{t=1}^n \mathbf{n}_t \right\} \left(1 + o_p \left(\frac{1}{\log_2^{1/2} n} \right) \right), \end{aligned}$$

uniformly in $\log n \leq s \leq n - \log n$. Thus, we conclude from this, (6.13) and (6.12) that the asymptotic distribution of \mathcal{LM}^* is that of

$$(6.14) \quad a_n \max_{\log n \leq s \leq n - \log n} \left\| \lambda_n(s) \left\{ \sum_{t=1}^s \mathbf{n}_t - \frac{s}{n} \sum_{t=1}^n \mathbf{n}_t \right\} \right\| - b_n.$$

On the other hand, as $\{\mathbf{n}_t\}$ is an *iid* sequence, the standard functional central limit theorem and the continuous mapping theorem imply that for any $\varrho > 0$

$$(6.15) \quad a_n \sup_{\log n \leq s < n - [n\varrho]} \left\| \varkappa_n(n-s) \sum_{t=s+1}^n \mathbf{n}_t \right\| - b_n \xrightarrow{P} -\infty$$

$$(6.16) \quad a_n \sup_{[n\varrho] < s \leq n - \log n} \left\| \varkappa_n(s) \sum_{t=1}^s \mathbf{n}_t \right\| - b_n \xrightarrow{P} -\infty.$$

Then, (6.15) – (6.16) imply that the (asymptotic) distribution of (6.14) is that of

$$a_n \max \left\{ \sup_{\log n \leq s < [n\varrho]} \left\| \varkappa_n(s) \sum_{t=1}^s \mathbf{n}_t \right\|; \sup_{\log n \leq s < [n\varrho]} \left\| \varkappa_n(s) \sum_{t=1}^s \mathbf{n}_t^* \right\| \right\} - b_n$$

using Proposition 1 and where the sequences $\{\mathbf{n}_t\}_{t \geq 1}$ and $\{\mathbf{n}_t^*\}_{t \geq 1}$ are independent sequences of mutually independent standard normal random vectors. But

$$(6.17) \quad \sup_{n/\log n \leq s < [n\varrho]} \left\| \varkappa_n(s) \sum_{t=1}^s \mathbf{n}_t \right\| = O_p(\log_3 n),$$

for $\{\mathbf{n}_t\}_{t \geq 1}$ a sequence of iid Gaussian random vectors, which implies that

$$\mathcal{LM}^* = a_n \max \left\{ \sup_{\log n \leq s < \frac{n}{\log n}} \left\| \varkappa_n(s) \sum_{t=1}^s \mathbf{n}_t \right\|; \sup_{\log n \leq s < \frac{n}{\log n}} \left\| \varkappa_n(s) \sum_{t=1}^s \mathbf{n}_t^* \right\| \right\} - b_n + o_p(1).$$

But, $\sup_{\log n \leq s < n/\log n} \left| \left(\left(\frac{n-s}{n} \right)^{1/2} \right) - 1 \right| = O(\log^{-1} n)$, so we can conclude that

$$\mathcal{LM}^* = a_n \max \left\{ \sup_{\log n \leq s < \frac{n}{\log n}} \left\| \frac{1}{s^{1/2}} \sum_{t=1}^s \mathbf{n}_t \right\|; \sup_{\log n \leq s < \frac{n}{\log n}} \left\| \frac{1}{s^{1/2}} \sum_{t=1}^s \mathbf{n}_t^* \right\| \right\} - b_n + o_p(1).$$

Denoting by $U(s^*)$ the Ornstein-Uhlenbeck process, by the change of time $s \rightarrow e^{s^*}$, we obtain that

$$\sup_{\log n \leq s < n/\log n} \frac{1}{s^{1/2}} \sum_{t=1}^s \mathbf{n}_t = \sup_{\log_2 n < s^* \leq \log n - \log_2 n} U(s^*).$$

From here the proof follows by Lemma 2.2 of Horváth (1993).

The proof for \mathcal{LM} is obvious after observing that \mathcal{LM} is a continuous mapping of \mathcal{LM}^* . \square

For expositional purpose, the proof of Theorem 2 comes after that of Theorem 3. Before proving Theorem 3, we shall give a lemma.

Lemma 1. *Under A1 and A2 and H_{1n} , if $\sup_{p < s < [n/2]} |h_n(s)| = o(b_n^{-1/2})$,*

$$\Pr \left\{ a_n \sup_{p < s < [\frac{n}{2}]} |U_n(s) + h_n(s)| - b_n < x \right\} = \Pr \left\{ a_n \sup_{p < s < [\frac{n}{2}]} |U_n(s)| - b_n < x \right\} + o(1).$$

Proof. The proof follows after we observe that

$$a_n \sup_{p < s < [n/2]} |U_n(s) + h_n(s)| - b_n = b_n \left(\sup_{p < s < [n/2]} \left| \frac{a_n}{b_n} U_n(s) + o(b_n^{-1}) \right| - 1 \right).$$

Now, use the inequality $|a| - |b| \leq |a + b| \leq |a| + |b|$ to conclude the proof. \square

6.3. Proof of Theorem 3.

By symmetry, we can focus when $s_0 \leq [n/2]$. We shall prove parts (a) and (b) simultaneously. From the proof of Theorem 1 and noticing that as $\|\delta_n\| \rightarrow 0$ we have that $\widehat{\theta} - \theta_0 = o_p(1)$, it suffices to examine the behaviour of

$$(6.18) \quad Q_n = a_n \sup_{p < s \leq n-p} \|Q_n(s)\| - b_n,$$

where $Q_n(s) = (U_n(s) + \Lambda_n(s; s_0))$ with

$$(6.19) \quad U_n(s) = \lambda_n(s) \left(\sum_{t=1}^s g_t(\theta_{nt}) - \frac{s}{n} \sum_{t=1}^n g_t(\theta_{nt}) \right),$$

where by an application of the mean value expansion of $g_t(\theta_{nt})$ at θ_0 ,

$$(6.20) \quad \Lambda_n(s; s_0) = \widetilde{\delta}_n s_0 \varkappa_n(s) \mathbf{1}(s_0 \leq s) + \widetilde{\delta}_n (n - s_0) \varkappa_n(n - s) \mathbf{1}(s < s_0),$$

where $\widetilde{\delta}_n = G_0 \delta_n$.

The proof will be done when (i) $[n\tau] \leq s_0 < [n/2]$, (ii) $[n/\log n] \leq s_0 < [n\tau]$ and when (iii) $\log_2 n \leq s_0 < [n/\log n]$. We begin with part (i). Because by standard arguments, $\sup_{[n\tau] \leq s < [n/2]} |U_n(s)| = O_p(1)$, and using (6.20),

$$\sup_{[n\tau] \leq s < [n/2]} |U_n(s) + \Lambda_n(s; s_0)| = O_p(1) + O\left(a_n^{1/2} c_n\right),$$

when $\|\delta_n\|^{-1} = o\left(s_0^{1/2}/\log_2^{1/2} n\right)$ and where $c_n^{-1} = o(1)$. So, $Q_n^{-1} \xrightarrow{p} 0$, which implies that the test rejects with probability 1 as n increases to infinity.

On the other hand, when $\|\delta_n\| = o\left(\log_2^{1/2} n/s_0^{1/2}\right)$ we obtain that

$$\sup_{[n\tau] \leq s < [n/2]} |Q_n(s)| = O_p(1) + o\left(a_n^{1/2}\right)$$

so that $Q_n \rightarrow_P -\infty$, which implies that the ‘‘max’’ in the definition of Q_n is when $s < [n\tau]$. Now because $s < [n\tau]$ and hence $s_0 > s$, we obtain that

$$(6.21) \quad \Lambda_n(s; s_0) = O\left(\|\delta_n\| \left(\frac{s}{n}\right)^{1/2} \frac{n - s_0}{(n - s)^{1/2}}\right) = o\left(\left(\frac{s \log_2 n}{n}\right)^{1/2}\right).$$

So, $a_n \sup_{[n/\log n] < s < [n\tau]} |\Lambda_n(s; s_0)| = o(b_n)$ and proceeding as in the proof of Theorem 1, cf. (6.17), we have that the ‘‘max’’ of Q_n is achieved when $s < [n/\log n]$. But in this region, uniformly in s , (6.21) = $o\left((\log_2 n/\log n)^{1/2}\right) = o\left(b_n^{-1/2}\right)$, so by Lemma 1, we conclude part (b) of the theorem. This completes the proof of (i).

Next we examine case (ii). By definition of $\Lambda_n(s; s_0)$, we have that

$$\sup_{p < s < [n/\log n]} |\Lambda_n(s; s_0)| \geq C \|\delta_n\| \frac{s_0}{n^{1/2}} \log^{1/2} n.$$

So, when $\|\delta_n\|^{-1} = o\left(s_0^{1/2}/\log_2^{1/2} n\right)$ the last displayed inequality implies that

$$\|\delta_n\|^{-1} \frac{n^{1/2}}{s_0 \log^{1/2} n} = o\left(\left(\frac{n}{s_0 \log n \log_2 n}\right)^{1/2}\right) = o\left(b_n^{-1/2}\right).$$

Hence $|a_n \Lambda_n(s; s_0)|^{-1} = o(b_n^{-1})$, which implies that the test rejects with probability 1 as n increases to infinity. Next, when $\|\delta_n\|^{-1} = o\left(\log_2^{1/2} n/s_0^{1/2}\right)$, we have that $\sup_{[n\tau] < s < [n/2]} |\Lambda_n(s; s_0)| = o\left(\|\delta_n\| \frac{s_0}{n^{1/2}}\right) = o\left(b_n^{1/2}\right)$ as is $\sup_{[n/\log n] < s \leq [n\tau]} |\Lambda_n(s; s_0)|$ as we now show. Indeed,

$$|\Lambda_n(s; s_0)| = o\left(\left(\frac{s_0}{s}\right) \log_2^{1/2} n \mathbf{1}(s_0 < s) + \left(\frac{s}{n}\right)^{1/2} \left(\frac{n-s_0}{s_0}\right)^{1/2} \log_2^{1/2} n \mathbf{1}(s \leq s_0)\right),$$

so that $\sup_{[n/\log n] < s \leq [n\tau]} |\Lambda_n(s; s_0)| = o\left(b_n^{1/2}\right)$. Thus, proceeding as in the proof of Theorem 1, cf. (6.17), we have that the “max” in Q_n is when $p < s < [n/\log n]$. But in that region, $\sup_{p^* < s < [n/\log n]} |\Lambda_n(s; s_0)| = o\left(b_n^{-1/2}\right)$ and by Lemma 1, we conclude part (b) of the theorem.

Finally, we examine case (iii). If $s_0 \leq s$, we know that $|\Lambda_n(s; s_0)| = C \|\delta_n\| s_0/s^{1/2}$. So, we have that when $\|\delta_n\|^{-1} = o\left(s_0^{1/2}/\log_2^{1/2} n\right)$, $|\Lambda_n(s; s_0)|^{-1} = o\left((s/s_0)^{1/2} \log_2^{-1/2} n\right)$ which implies that $\inf |\Lambda_n(s; s_0)|^{-1} = o\left(\log_2^{-1/2} n\right) = o\left(b_n^{-1/2}\right)$. So $Q_n^{-1} \rightarrow_P 0$ and hence the test rejects with probability 1 as $n \nearrow \infty$. Next, when $\|\delta_n\| = o\left(\log_2^{1/2} n/s_0^{1/2}\right)$, the proof proceeds as in cases (i) or (ii) and so it is omitted. \square

6.4. Proof of Theorem 2.

First of all, proceeding as in the proof of Theorem 1 and ignoring the constant terms, it suffices to examine the behaviour of Q_n in (6.18) but with $U_n(s)$ in (6.19) replaced by $U_n(s) = \lambda_n(s) \left(\sum_{t=1}^s (g_t^* - E(g_t^*)) - \frac{s}{n} \sum_{t=1}^n (g_t^* - E(g_t^*))\right)$. Since $\Lambda_n(s_0; s_0) = \gamma s_0 (n - s_0)/n$, $s_0^{-1} = o\left(\log_2^{-1} n\right)$ and $(n - s_0)^{-1} = o\left(\log_2^{-1} n\right)$ imply that $a_n \sup_{p < s \leq n-p} |\Lambda_n(s; s_0)| - b_n \rightarrow \infty$. From here, the conclusion is standard. \square

REFERENCES

- [1] ANDREWS, D.W.K. (1993): “Tests for parameter instability and structural change with unknown change point,” *Econometrica*, **61**, 821-856.
- [2] ANDREWS, D.W.K. (2003): “End-of-sample instability tests,” *Econometrica*, **71**, 1661-1694.
- [3] ANDREWS, D.W.K. AND KIM, J.-Y. (2006): “End-of-sample cointegration breakdown tests,” *Journal of Business and Economic Statistics*, **24**, 379-394.
- [4] 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. Ser. B.*, **37**, 149-192.
- [5] BURBIDGE, J.B., MAGEE, L. AND ROBB, A.L. (1988): “Alternative transformations to handle extreme values of the dependent variable,” *Journal of the American Statistical Association*, **83**, 123-127.
- [6] CHOW, G.C. (1960): “Tests of equality between sets of coefficients in two linear regressions,” *Econometrica*, **28**, 591-605.
- [7] COCHRANE, J.H. (2005): *Asset Pricing*. Princeton University Press.
- [8] CSÖRGO, M. AND HORVÁTH, L. (1997): *Limit Theorems in Change-point Analysis*. Wiley.
- [9] CSÖRGO, M. AND RÉVÉSZ, P. (1981): *Strong approximations in probability and statistics*. Academic Press.

- [10] DAVIS, R.A., HUANG, D. AND YAO, Y. (1995): "Testing for a Change in the Parameter Values and Order of an Autoregressive Model," *Annals of Statistics*, **23**, 282-304.
- [11] GÖTZE, F. AND ZAITSEV, A.Y. (2007): "Bounds for the rate of strong approximation in the multidimensional invariance principle," Preprint.
- [12] HORVÁTH, L. (1993): "The maximum likelihood method for testing change in the parameters of normal observations," *Annals of Statistics*, **21**, 671-680.
- [13] 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.
- [14] KUELBS, J. AND PHILIPP, W. (1980): "Almost sure invariance principles for partial sums of mixing B-valued random variables," *Annals of Probability*, **8**, 1003-1036.
- [15] LING, S. (2007): "Testing for change-points in time series models and limiting theorems for NED sequences," *Annals of Statistics*, **35**, 1213-1237.
- [16] PERRON, P. (2006): "Dealing with structural breaks," Forthcoming in Pelgrave Handbook of Econometrics, Vol. 1: Econometric Theory.
- [17] QUANDT, R.E. (1960): "Test of the hypothesis that a linear regression system obeys two separate regimes," *Journal of the American Statistical Association*, **55**, 324-330.
- [18] ROBINSON, P. M. (1998): "Inference-without-smoothing in the presence of nonparametric autocorrelation," *Econometrica*, **66**, 1163-1182.
- [19] SHAO, Q.-M. (1995): "Strong Approximation theorems for independent random variables and their applications," *Journal of Multivariate Analysis*, **52**, 107-130.
- [20] WU, W. B. (2007): "Strong invariance principles for dependent random variables," *Annals of Probability*, **35**, 2294-2320.

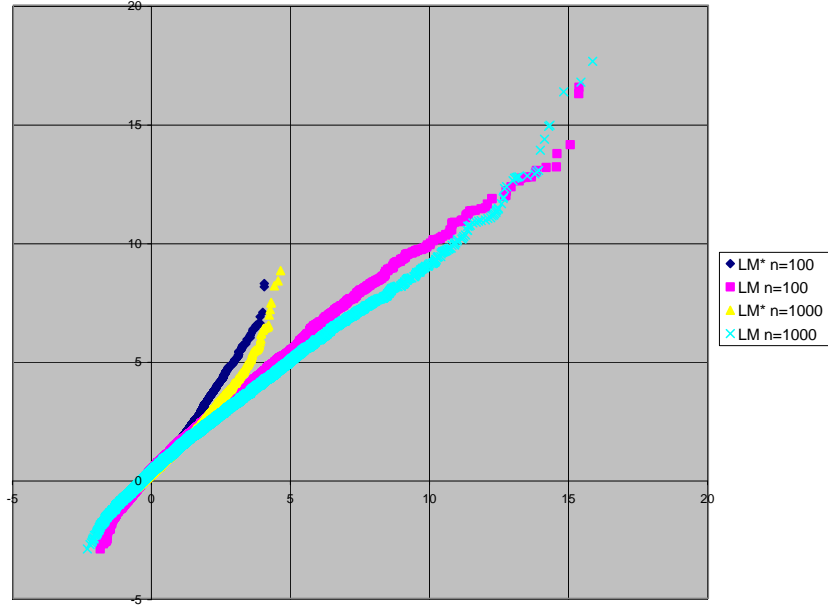


Figure 1
qq plot

n	\mathcal{LM}^*	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	CUSUM	CHOW $_{n/2}$
$z = 1$						
100	0.000	0.037	0.027	0.034	0.031	0.044
250	0.000	0.038	0.034	0.040	0.044	0.048
500	0.000	0.044	0.041	0.047	0.041	0.055
1000	0.000	0.049	0.047	0.051	0.048	0.047
$z = x$						
100	0.000	0.060	0.035	0.030	0.031	0.043
250	0.000	0.049	0.035	0.034	0.044	0.054
500	0.001	0.074	0.052	0.042	0.041	0.046
1000	0.003	0.067	0.030	0.033	0.048	0.045
$z = (1, x)'$						
100	0.000	0.055	0.029	0.025	0.030	0.041
250	0.000	0.076	0.046	0.036	0.046	0.056
500	0.000	0.077	0.045	0.045	0.046	0.054
1000	0.001	0.074	0.048	0.040	0.047	0.043

Table 1

Rejection frequencies under H_0 at 5% level.

s	δ	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	CUSUM	CHOW
50	0.1	0.051	0.041	0.049	0.030	0.087
	0.5	0.447	0.408	0.481	0.164	0.672
75	0.1	0.046	0.033	0.047	0.033	0.060
	0.5	0.325	0.286	0.350	0.046	0.576
90	0.1	0.043	0.028	0.039	0.031	0.051
	0.5	0.134	0.107	0.091	0.032	0.312
95	0.1	0.036	0.026	0.037	0.031	0.050
	0.5	0.080	0.060	0.048	0.031	0.194
125	0.1	0.078	0.080	0.088	0.059	0.130
	0.5	0.889	0.896	0.918	0.526	0.973
188	0.1	0.073	0.069	0.070	0.044	0.109
	0.5	0.771	0.778	0.815	0.146	0.928
225	0.1	0.053	0.048	0.048	0.043	0.092
	0.5	0.413	0.403	0.277	0.047	0.673
238	0.1	0.045	0.039	0.042	0.045	0.064
	0.5	0.196	0.166	0.091	0.042	0.416
250	0.1	0.106	0.115	0.129	0.070	0.190
	0.5	0.998	0.997	0.999	0.881	1.000
375	0.1	0.091	0.097	0.096	0.049	0.159
	0.5	0.977	0.977	0.987	0.341	0.999
450	0.1	0.065	0.064	0.059	0.043	0.080
	0.5	0.707	0.726	0.511	0.065	0.919
475	0.1	0.056	0.056	0.043	0.042	0.084
	0.5	0.370	0.340	0.137	0.045	0.669
500	0.1	0.173	0.193	0.237	0.099	0.343
	0.5	1.000	1.000	1.000	0.997	1.000
750	0.1	0.130	0.145	0.180	0.061	0.271
	0.5	1.000	1.000	1.000	0.752	1.000
900	0.1	0.090	0.091	0.080	0.050	0.167
	0.5	0.975	0.978	0.846	0.104	0.998
950	0.1	0.066	0.063	0.065	0.048	0.100
	0.5	0.734	0.735	0.253	0.052	0.933

Table 2a

Rejection frequencies under H_1 ; $z = 1$

When the break occurs in the middle of the sample.

s	δ	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	CUSUM	CHOW
50	0.1	0.078	0.051	0.057	0.031	0.076
	0.5	0.441	0.393	0.457	0.022	0.663
75	0.1	0.074	0.047	0.048	0.031	0.075
	0.5	0.335	0.300	0.328	0.025	0.543
90	0.1	0.068	0.041	0.037	0.031	0.075
	0.5	0.192	0.159	0.112	0.030	0.294
95	0.1	0.067	0.037	0.031	0.031	0.075
	0.5	0.136	0.099	0.050	0.030	0.214
125	0.1	0.079	0.063	0.073	0.045	0.129
	0.5	0.880	0.883	0.914	0.037	0.973
188	0.1	0.080	0.059	0.066	0.043	0.103
	0.5	0.730	0.729	0.767	0.037	0.907
225	0.1	0.065	0.047	0.040	0.044	0.073
	0.5	0.405	0.384	0.242	0.042	0.606
238	0.1	0.058	0.038	0.031	0.044	0.067
	0.5	0.232	0.176	0.078	0.043	0.372
250	0.1	0.125	0.104	0.122	0.041	0.187
	0.5	0.995	0.996	0.998	0.037	1.000
375	0.1	0.098	0.088	0.098	0.041	0.158
	0.5	0.968	0.975	0.982	0.038	0.996
450	0.1	0.086	0.061	0.050	0.041	0.106
	0.5	0.706	0.711	0.514	0.039	0.900
475	0.1	0.081	0.052	0.045	0.041	0.080
	0.5	0.429	0.376	0.145	0.040	0.625
500	0.1	0.188	0.189	0.221	0.047	0.331
	0.5	1.000	1.000	1.000	0.040	1.000
750	0.1	0.149	0.135	0.150	0.048	0.242
	0.5	1.000	1.000	1.000	0.040	1.000
900	0.1	0.095	0.069	0.065	0.048	0.150
	0.5	0.949	0.955	0.833	0.043	0.998
950	0.1	0.084	0.050	0.044	0.048	0.100
	0.5	0.686	0.668	0.270	0.044	0.892

Table 2b

Rejection frequencies under H_1 ; $z = x$
When the break occurs in the middle of the sample.

s	δ	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	CUSUM	CHOW
50	0.1	0.078	0.043	0.048	0.030	0.094
	0.5	0.676	0.614	0.665	0.142	0.864
75	0.1	0.073	0.043	0.041	0.030	0.090
	0.5	0.540	0.469	0.505	0.040	0.733
90	0.1	0.067	0.037	0.034	0.030	0.068
	0.5	0.302	0.243	0.141	0.031	0.418
95	0.1	0.059	0.035	0.028	0.031	0.075
	0.5	0.194	0.138	0.054	0.029	0.295
125	0.1	0.123	0.095	0.092	0.060	0.152
	0.5	0.994	0.993	0.996	0.493	0.999
188	0.1	0.112	0.086	0.077	0.048	0.131
	0.5	0.949	0.943	0.955	0.130	0.992
225	0.1	0.094	0.061	0.045	0.045	0.082
	0.5	0.644	0.605	0.406	0.047	0.832
238	0.1	0.084	0.048	0.037	0.046	0.067
	0.5	0.372	0.289	0.108	0.043	0.556
250	0.1	0.170	0.140	0.159	0.064	0.262
	0.5	1.000	1.000	1.000	0.857	1.000
375	0.1	0.140	0.106	0.122	0.052	0.204
	0.5	1.000	1.000	1.000	0.320	1.000
450	0.1	0.103	0.068	0.042	0.046	0.104
	0.5	0.928	0.914	0.728	0.062	0.987
475	0.1	0.090	0.051	0.038	0.046	0.092
	0.5	0.637	0.554	0.193	0.048	0.832
500	0.1	0.301	0.291	0.344	0.098	0.484
	0.5	1.000	1.000	1.000	0.996	1.000
750	0.1	0.231	0.209	0.224	0.059	0.354
	0.5	1.000	1.000	1.000	0.703	1.000
900	0.1	0.129	0.109	0.079	0.049	0.209
	0.5	0.998	0.998	0.979	0.101	1.000
950	0.1	0.098	0.066	0.046	0.048	0.120
	0.5	0.926	0.913	0.379	0.053	0.990

Table 2c

Rejection frequencies under H_1 ; $z = (1, x)'$

When the break occurs in the middle of the sample.

s	δ	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	CUSUM	CHOW
90	0.1	0.043	0.028	0.039	0.031	0.051
	0.5	0.134	0.107	0.091	0.032	0.312
93	0.1	0.039	0.027	0.037	0.031	0.061
	0.5	0.108	0.081	0.057	0.032	0.232
95	0.1	0.036	0.026	0.037	0.031	0.050
	0.5	0.080	0.060	0.048	0.031	0.194
97	0.1	0.037	0.026	0.036	0.031	0.055
	0.5	0.056	0.039	0.036	0.031	0.146
234	0.1	0.048	0.041	0.045	0.043	0.072
	0.5	0.282	0.263	0.129	0.044	0.525
239	0.1	0.044	0.038	0.042	0.045	0.066
	0.5	0.177	0.138	0.077	0.043	0.396
242	0.1	0.043	0.037	0.042	0.045	0.056
	0.5	0.143	0.085	0.057	0.044	0.307
245	0.1	0.042	0.035	0.041	0.044	0.043
	0.5	0.092	0.053	0.048	0.045	0.211
478	0.1	0.055	0.053	0.043	0.042	0.083
	0.5	0.322	0.263	0.109	0.043	0.598
485	0.1	0.054	0.048	0.045	0.042	0.076
	0.5	0.239	0.139	0.077	0.042	0.465
489	0.1	0.051	0.046	0.046	0.041	0.067
	0.5	0.165	0.083	0.060	0.043	0.360
493	0.1	0.048	0.042	0.047	0.041	0.060
	0.5	0.119	0.058	0.047	0.042	0.259
968	0.1	0.061	0.052	0.064	0.048	0.086
	0.5	0.511	0.318	0.119	0.048	0.793
979	0.1	0.055	0.051	0.057	0.048	0.075
	0.5	0.324	0.142	0.081	0.049	0.607
984	0.1	0.051	0.049	0.053	0.048	0.064
	0.5	0.243	0.098	0.071	0.047	0.506
989	0.1	0.048	0.047	0.050	0.048	0.066
	0.5	0.167	0.067	0.066	0.047	0.370

Table 3a

Rejection frequencies under H_1 ; $z = 1$
When the break occurs in the end of the sample.

s	δ	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	CUSUM	CHOW
90	0.1	0.068	0.041	0.037	0.031	0.075
	0.5	0.192	0.159	0.112	0.030	0.294
93	0.1	0.067	0.037	0.033	0.031	0.065
	0.5	0.159	0.121	0.068	0.030	0.241
95	0.1	0.067	0.037	0.031	0.031	0.075
	0.5	0.136	0.099	0.050	0.030	0.214
97	0.1	0.061	0.035	0.029	0.031	0.076
	0.5	0.111	0.062	0.041	0.031	0.156
234	0.1	0.061	0.039	0.040	0.043	0.059
	0.5	0.276	0.235	0.109	0.042	0.465
239	0.1	0.058	0.038	0.031	0.043	0.062
	0.5	0.215	0.150	0.073	0.042	0.354
242	0.1	0.059	0.038	0.032	0.044	0.061
	0.5	0.175	0.103	0.059	0.042	0.280
245	0.1	0.055	0.037	0.033	0.043	0.068
	0.5	0.130	0.059	0.043	0.043	0.195
478	0.1	0.079	0.052	0.043	0.041	0.078
	0.5	0.377	0.289	0.116	0.040	0.571
485	0.1	0.079	0.050	0.041	0.041	0.068
	0.5	0.279	0.138	0.074	0.042	0.450
489	0.1	0.078	0.049	0.041	0.041	0.070
	0.5	0.225	0.090	0.053	0.041	0.339
493	0.1	0.077	0.049	0.041	0.041	0.054
	0.5	0.162	0.065	0.046	0.041	0.247
968	0.1	0.073	0.039	0.041	0.048	0.088
	0.5	0.490	0.298	0.114	0.047	0.715
979	0.1	0.067	0.033	0.036	0.048	0.073
	0.5	0.338	0.129	0.070	0.046	0.557
984	0.1	0.067	0.033	0.035	0.048	0.057
	0.5	0.273	0.081	0.058	0.046	0.455
989	0.1	0.063	0.033	0.033	0.048	0.065
	0.5	0.214	0.060	0.047	0.046	0.329

Table 3b

Rejection frequencies under H_1 ; $z = x$
When the break occurs in the end of the sample.

s	δ	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	CUSUM	CHOW
90	0.1	0.067	0.037	0.034	0.030	0.068
	0.5	0.302	0.243	0.141	0.031	0.418
93	0.1	0.063	0.037	0.031	0.030	0.070
	0.5	0.239	0.182	0.076	0.029	0.346
95	0.1	0.059	0.035	0.028	0.031	0.075
	0.5	0.194	0.138	0.054	0.029	0.295
97	0.1	0.057	0.033	0.027	0.030	0.000
	0.5	0.124	0.071	0.033	0.030	0.000
234	0.1	0.089	0.053	0.038	0.045	0.074
	0.5	0.464	0.405	0.174	0.042	0.669
239	0.1	0.084	0.048	0.037	0.046	0.075
	0.5	0.350	0.250	0.095	0.044	0.511
242	0.1	0.080	0.049	0.034	0.046	0.069
	0.5	0.273	0.139	0.068	0.044	0.399
245	0.1	0.079	0.048	0.034	0.046	0.071
	0.5	0.197	0.076	0.046	0.045	0.291
478	0.1	0.086	0.048	0.039	0.046	0.090
	0.5	0.588	0.442	0.156	0.045	0.771
485	0.1	0.084	0.048	0.040	0.045	0.080
	0.5	0.454	0.217	0.076	0.046	0.613
489	0.1	0.082	0.047	0.043	0.045	0.075
	0.5	0.354	0.127	0.053	0.045	0.499
493	0.1	0.080	0.048	0.044	0.046	0.068
	0.5	0.240	0.075	0.044	0.044	0.363
968	0.1	0.087	0.054	0.046	0.047	0.113
	0.5	0.755	0.492	0.163	0.048	0.915
979	0.1	0.076	0.050	0.042	0.047	0.091
	0.5	0.541	0.196	0.087	0.047	0.788
984	0.1	0.073	0.047	0.043	0.047	0.083
	0.5	0.432	0.127	0.064	0.047	0.645
989	0.1	0.073	0.047	0.041	0.047	0.076
	0.5	0.332	0.078	0.051	0.048	0.494

Table 3c

Rejection frequencies under H_1 ; $z = (1, x)'$
When the break occurs in the end of the sample.

n	\mathcal{LM}^*	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$
$\rho = 0.5$				
100	0.018	0.061	0.043	0.027
250	0.032	0.067	0.042	0.032
500	0.024	0.076	0.042	0.033
1000	0.032	0.075	0.040	0.040
$\rho = 0.9$				
100	0.017	0.060	0.044	0.039
250	0.021	0.050	0.043	0.034
500	0.024	0.053	0.044	0.034
1000	0.026	0.057	0.050	0.036

Table 4

Rejection frequencies of the LM tests under H_0 with 5% level

s	δ	ρ			ρ		
		\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$
50	0.5	0.272	0.236	0.275	0.080	0.068	0.062
	1.0	0.817	0.785	0.833	0.168	0.147	0.164
75	0.5	0.216	0.184	0.209	0.088	0.066	0.062
	1.0	0.683	0.645	0.692	0.184	0.145	0.161
90	0.5	0.148	0.121	0.070	0.081	0.066	0.048
	1.0	0.402	0.348	0.224	0.127	0.109	0.087
95	0.5	0.120	0.088	0.040	0.077	0.059	0.044
	1.0	0.258	0.208	0.081	0.114	0.095	0.063
125	0.5	0.632	0.636	0.681	0.112	0.105	0.106
	1.0	0.999	0.999	0.999	0.358	0.352	0.414
188	0.5	0.483	0.474	0.525	0.116	0.114	0.101
	1.0	0.974	0.976	0.985	0.287	0.284	0.304
225	0.5	0.258	0.236	0.161	0.095	0.088	0.057
	1.0	0.718	0.706	0.524	0.205	0.193	0.122
238	0.5	0.180	0.130	0.062	0.074	0.063	0.042
	1.0	0.442	0.376	0.155	0.142	0.121	0.064
250	0.5	0.927	0.935	0.960	0.162	0.173	0.213
	1.0	1.000	1.000	1.000	0.693	0.714	0.775
375	0.5	0.796	0.811	0.848	0.156	0.161	0.167
	1.0	1.000	1.000	1.000	0.538	0.565	0.607
450	0.5	0.437	0.421	0.283	0.113	0.099	0.059
	1.0	0.939	0.948	0.816	0.302	0.298	0.202
475	0.5	0.272	0.213	0.091	0.091	0.066	0.042
	1.0	0.712	0.679	0.292	0.191	0.163	0.070
500	0.5	0.999	0.999	0.999	0.410	0.466	0.523
	1.0	1.000	1.000	1.000	0.972	0.983	0.991
750	0.5	0.991	0.994	0.995	0.302	0.329	0.368
	1.0	1.000	1.000	1.000	0.877	0.896	0.918
900	0.5	0.777	0.782	0.582	0.169	0.166	0.118
	1.0	0.999	0.999	0.994	0.495	0.512	0.378
950	0.5	0.456	0.417	0.159	0.114	0.099	0.061
	1.0	0.945	0.942	0.558	0.305	0.268	0.118

Table 5

Rejection frequencies under H_1 with 5% level.
When the break occurs in the middle of the sample.

s	δ	ρ			ρ		
		\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$	\mathcal{LM}	$\mathcal{LM}_{.05}$	$\mathcal{LM}_{.15}$
90	0.5	0.148	0.121	0.070	0.081	0.066	0.048
	1.0	0.402	0.348	0.224	0.127	0.109	0.087
93	0.5	0.132	0.100	0.049	0.077	0.063	0.047
	1.0	0.317	0.276	0.119	0.124	0.100	0.073
95	0.5	0.120	0.088	0.040	0.077	0.059	0.044
	1.0	0.258	0.208	0.081	0.114	0.095	0.063
97	0.5	0.100	0.066	0.034	0.072	0.052	0.044
	1.0	0.186	0.124	0.048	0.099	0.068	0.048
234	0.5	0.197	0.163	0.085	0.080	0.074	0.045
	1.0	0.530	0.498	0.242	0.172	0.153	0.081
239	0.5	0.171	0.122	0.056	0.073	0.062	0.042
	1.0	0.416	0.332	0.141	0.136	0.113	0.062
242	0.5	0.145	0.079	0.045	0.068	0.057	0.040
	1.0	0.336	0.212	0.096	0.123	0.090	0.050
245	0.5	0.119	0.062	0.035	0.066	0.052	0.038
	1.0	0.252	0.110	0.065	0.104	0.064	0.045
478	0.5	0.243	0.169	0.080	0.086	0.063	0.039
	1.0	0.655	0.569	0.230	0.172	0.130	0.065
485	0.5	0.191	0.100	0.048	0.074	0.055	0.038
	1.0	0.497	0.289	0.124	0.149	0.094	0.053
489	0.5	0.167	0.074	0.039	0.071	0.048	0.038
	1.0	0.406	0.181	0.079	0.125	0.075	0.046
493	0.5	0.137	0.059	0.037	0.068	0.046	0.037
	1.0	0.286	0.092	0.058	0.106	0.062	0.042
968	0.5	0.316	0.173	0.094	0.098	0.069	0.052
	1.0	0.814	0.621	0.258	0.232	0.138	0.080
979	0.5	0.241	0.091	0.068	0.090	0.058	0.046
	1.0	0.634	0.311	0.127	0.184	0.094	0.064
984	0.5	0.213	0.070	0.059	0.084	0.055	0.042
	1.0	0.529	0.183	0.089	0.163	0.077	0.059
989	0.5	0.158	0.054	0.051	0.079	0.051	0.044
	1.0	0.395	0.099	0.066	0.134	0.067	0.047

Table 6

Rejection frequencies under H_1 with 5% level.
When the break occurs in the end of the sample.

LONDON SCHOOL OF ECONOMICS, HOUGHTON STREET, LONDON WC2A 2AE, UK
E-mail address: f.j.hidalgo@lse.ac.uk

LONDON SCHOOL OF ECONOMICS, HOUGHTON STREET, LONDON WC2A 2AE, UK
E-mail address: m.seo@lse.ac.uk