Pouliot, W. & Olmo, J. (2008). U-statistic Type Tests for Structural Breaks in Linear Regression Models (Report No. 08/15). London, UK: Department of Economics, City University London.

Department of Economics
School of Social Sciences

U-statistic Type Tests for Structural Breaks in Linear Regression Models

Jose Olmo[1]
City University

William Pouliot[2]
City University

Department of Economics
Discussion Paper Series
No. 08/15

[1] Department of Economics, City University, Northampton Square, London, EC1V 0HB, UK. Email: *j.olmo@city.ac.uk*

[2] Department of Economics, City University, Northampton Square, London, EC1V 0HB, UK. Email: *William.Pouliot.1@city.ac.uk*

U-statistic Type Tests for Structural Breaks in Linear Regression Models

WILLIAM POULIOT*, *Department of Economics, City University London.*
JOSE OLMO, *Department of Economics, City University London.*

November 2008

Abstract

This article introduces a U-statistic type process that is based on a kernel function which can depend on nuisance parameters. It is shown here that this process can accommodate very easily anti-symmetric kernels very useful for detecting changing patterns in the dynamics of time series. This theory is applied to structural break hypothesis tests in linear regression models. In particular, the flexibility of these processes will be exploited to introduce a simultaneous and joint test that exhibit statistical power against changes in either intercept or slope. In contrast to the literature, these tests are able to distinguish between rejections due to changes in intercept from rejections due to changes in slope; allow control of global errors rate; and are explicitly designed to have power when the distribution error is asymmetric. These tests can also incorporate different weight functions devised to detect changes early as well as later on in the sample, and show very good performance in small samples. These tests, therefore, outperform CUSUM type tests widely employed in this literature.

Keywords and Phrases: Change-Point tests; CUSUM test; Linear regression models; Stochastic processes; U-statistics

JEL codes: C12, C22, C52.

*Corresponding Address: Dept. Economics, City University. Northampton Square, EC1 V0HB, London. William Pouliot, E-mail: William.Pouliot.1@city.ac.uk

1 Introduction

Economics and finance frequently consider linear regression models (hereafter LRMs) with coefficients that are constant for all time periods. It is well-known that these parameters can, and do, change over time due, for example, to abrupt policy changes, to wars, to oil price or to technology shocks. This has led to considerable econometric research into methods that can detect if such exogenous events have caused parameters of linear regression models to change. One of the first papers published on this matter was by Chow (1960). He constructed two test statistics capable of detecting a one-time change in regression parameters at a known time. Work by Brown, Durbin and Evans (1975) (hereafter BDE) and Dufour (1988) extended Chow's test to accommodate multiple changes in regression parameters that may occur at unknown times. Other tests, called fluctuation tests, such as that of Ploberger, Kramer and Kontrus (1989) (hereafter PKK) have also been developed. An interesting contribution to this literature is that of Altissimo and Corradi (2003) who develop a statistic that tests for any number of break-points. This test as well as the other three, however, when applied to regression models are not devised to distinguish between changes in intercept or slope and, in turn, although informative about the number of break points are not very informative about the statistical cause of rejection.

A more recent contribution to this literature is Olmo and Pouliot (2008); they use U-statistic type processes to fashion a statistic that is capable of detecting a one-time change in parameters of a linear regression model that occurs early and later on in the sample. Their process, however, depends on a kernel defined by a function of a vector of nuisance parameters, that in most practical situations must be estimated. They show that although said substitution does not introduce estimation effects into the asymptotic distribution their kernel lacks desirable properties such as continuity. The research of Olmo and Pouliot (2008) does, nevertheless, confirm the importance of U-statistic based processes in detecting a one-time change in parameters of regression models.

This article extends this work in different directions. First, we introduce a U-statistic type process that is based on a general kernel that can depend on nuisance parameters and can accommodate multivariate random sampling. One of the main features of this process is that it can entertain anti-symmetric kernels which are very useful for detecting changing patterns in the dynamics of time series. The asymptotic theory of this process is derived, under the assumptions of known nuisance parameters and also when these parameters are estimated. The second contribution is to propose this family of processes for detecting structural breaks in linear regression models. In particular, we exploit the flexibility of these processes to introduce simultaneous and joint tests that exhibit statistical power to detect changes in either intercept or slope of linear regression models. This is in contrast with existing literature on the topic, see CUSUM type tests as introduced by BDE and fluctuation tests of PKK. More importantly, in contrast to these influential papers, our tests are able to distinguish between rejections due to changes in intercept from rejections due to changes in the slope parameter. These tests have the additional attraction of enabling control of global error rates in a similar fashion to ANOVA tests in the setting of testing for equality of k, $(k > 2)$ population means.

Another interesting feature of these tests is their explicit dependence on the third moment of the residuals of linear regression models. This characteristic idiosyncratic to our test implies significant improvements in terms of power when the distribution of the residuals shows some asymmetries about zero. The last contribution to the literature on change point detection and structural break tests is to show that simple modifications of the family of U-statistic type processes introduced in this paper given by suitable weight functions have more power against changes in the parameters in the linear regression model that occur early as well as later on in the sample. This is an important feature of this class of statistics not satisfied by CUSUM type tests which are unable to detect a change in parameters produced early/later on in the sample. It is also worth highlighting the good performance of both simultaneous and joint tests in small samples.

The paper is structured as follows. Section 2 introduces a family of processes based on U-statistics, and derives the corresponding asymptotic theory. Section 3 applies the findings of the previous section to derive a simultaneous and joint tests with power against deviations in either intercept or slope of linear regression models. The section also discusses suitable choices of weights in the test statistic that enhance the power against deviations of the process early or later on in the sample. Section 4 details the results from an extensive Monte Carlo exercise, studying nominal size and power of the test against alternatives that include a one-time change in intercept and slope. Section

5 concludes. Lastly, the limiting distribution of the weighted processes entertained here is tabulated and collected in the Appendix - the method used to simulate it follows that of Orasch and Pouliot (2004).

2 A New General Kernel function for U-statistic type processes

There are situations where the kernel upon which the U-statistic type process is fashioned from is differentiable with respect to nuisance parameters. One situation is that considered by Gombay, Horváth and Hušková, (hereafter GHH) (1996). They develop a statistic that can be used to test a sequence of i.i.d. random variables for constant variance. They consider the following setting; given a set of observations $\{X_1, \ldots, X_T\}$ for $T \geq 2, 3, \ldots$, one might be interested in testing for the presence of at most one change in variance at a distinct, yet unknown time. With positive constants σ and $\sigma^\star$, let

$$X_t = \begin{cases} \mu + \sigma\varepsilon_t, & 1 \leq t \leq t^*, \\ \mu + \sigma^\star\varepsilon_t, & t^* < t \leq T. \end{cases} \tag{1}$$

where

$$\varepsilon_t \text{ are independent and identically distributed with } \mathbb{E}\varepsilon_1 = 0,\, \mathbb{E}\varepsilon_t^2 = 1 \text{ and } \mathbb{E}|\varepsilon_t|^4 < \infty,\ t = 1, \ldots, T. \tag{2}$$

The values of the parameters $\mu,\ \sigma, \sigma^\star$ and $t^\star$ are unknown. Assuming that $\sigma \neq \sigma^\star$, the no change in variance null hypothesis can be formulated as

$$H_O : t^* \geq T$$

versus the at-most-one change (AMOC) in variance alternative

$$H_A : 1 \leq t^* < T.$$

To test the null hypothesis GHH use the change in mean framework to develop a statistic suited to testing for AMOC in the variance. Their statistic is reproduced below;

$$M_T^{(1)}(\tau) := T^{1/2}\tau(1 - \tau)\left\{\frac{1}{T\tau}\sum_{t=1}^{[(T+1)\tau]}(X_t - \mu)^2 - \frac{1}{T - T\tau}\sum_{t=[T\tau]+1}^{T}(X_t - \mu)^2\right\}, \quad 0 \leq \tau < 1 \tag{3}$$

which compares two estimators of the variance. One estimator is fashioned from the first $[(T+1)\tau]$ observations and then compared to the estimator constructed from the last $T - [(T+1)\tau]$ observations. After some simple algebra, the above process can be re-expressed as,

$$M_T^{(1)}(\tau) := T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]}(X_t - \mu)^2 - \tau\sum_{t=1}^{T}(X_t - \mu)^2\right\}, \quad 0 \leq \tau < 1. \tag{4}$$

This representation of $M_T^{(1)}(\tau)$ will be used in what follows as it is simpler to manipulate.

The kernel, $h(x, y)$, used to construct their process set $h(x, y) = (x - \mu)^2 - (y - \mu)^2$, which depends on the unknown parameter μ. GHH substitute $\bar{X}_T = \frac{\sum_{t=1}^{T} X_t}{T}$ for μ and arrive at,

$$\widetilde{M}_T^{(1)}(\tau) := T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]}(X_t - \bar{X}_T)^2 - \tau\sum_{t=1}^{T}(X_t - \bar{X}_T)^2\right\}, \quad 0 \leq \tau < 1. \tag{5}$$

They are able to derive the corresponding asymptotic distribution of (5) but their results apply only to the above process. It would seem natural, then, to see whether it is possible to make a more general statement that would hold in situations where the kernel, $h(\cdot,\cdot)$ is a more general function of nuisance parameters than the one used in their process.

Here, the partial sum process developed in (3) is extended to accommodate a kernel that is now a differentiable function of the nuisance parameters and can accommodate multivariate random samples.

Definition 2.1. *Let $\{\mathbf{X}_t\}_{t=1}^{T}$ be a sequence of multivariate random variables (hereafter rvs); let the kernel have the following representation $h(\mathbf{x},\mathbf{y};\boldsymbol{\theta}) = f(\mathbf{x};\boldsymbol{\theta}) - f(\mathbf{y};\boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \mathbb{R}^p$, $\mathbf{x},\mathbf{y} \epsilon \mathbb{R}^n$, $f(\cdot;\cdot)$ is a continuous function of $\boldsymbol{\theta}$ that is at least once differentiable - the derivative need not be continuous. This leads to the following modification of GHH's process,*

$$M_T^{(2)}(\tau) := T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]} f(\mathbf{X}_t;\boldsymbol{\theta}) - \tau\sum_{t=1}^{T} f(\mathbf{X}_t;\boldsymbol{\theta})\right\}, \quad 0 \le \tau < 1. \tag{6}$$

Some additional definitions and notation are required before any statement can be made regarding this sequence of partial sum processes. First, we introduce a class of functions Q.

Definition 2.2. *Let Q be the class of positive functions on $(0,1)$ which are non-decreasing in a neighborhood of zero and non-increasing in a neighbourhood of one, where a function*

$q(\cdot)$ defined on (0,1) is called positive if

$$\inf_{\delta \leq \tau \leq 1-\delta} q(\tau) > 0 \qquad \textit{for all } \delta \in (0, 1/2).$$

Definition 2.3. *Let $q(\cdot)\ \epsilon\ Q$. Then $I(q,c) = \int_0^1 \frac{1}{\tau(1-\tau)} \exp^{-\frac{c}{(\tau(1-\tau))q^2(\tau)}} d\tau$ for some constant $c > 0$.*

As advertised in Section 1, the weight functions $q(\cdot)$ that are members of the set Q will play an important role in this and following sections. It is useful then to discuss the purpose that these functions serve. Weight functions are of interest here because they add some flexibility in the search for tests that are able to detect at most one-change in intercept or slope. To justify this statement one important source in the statistical literature, Mason and Scheunemeyer (1983) (hereafter MS), can be cited. MS study finite and large sample properties of the power of the Kolmogorov-Smirnov (hereafter KS) statistic. They conclude that KS statistic displays poor sensitivity to detect deviations from the hypothesized distribution that may occur on the tails: MS show the KS statistic is inconsistent for such deviations. A similar fate holds true for the statistics to be fashioned from the process dealt with here; they also display low power to detect such deviations. Including weight functions early on in this development offers returns in terms of higher power of our statistics fashioned from the weighted process. Section 3 will offer more details on the nature of these weight functions. More information regarding these weight functions and associated theory can be found in Csörgő and Horváth (1997) and relevant chapters therein.

In order to establish the asymptotic behaviour of the process detailed in (6), it will be necessary to extend a result of Szyszkowicz (Thereom 2.1, (1991)), to the multivariate setting. To do so, some additional notation and definitions must now be introduced. To begin, let $\{\mathbf{X}_t\}_{t=1}^{T}$ be a sequence of independent and identically distributed random vectors with

$$\mathbb{E}\mathbf{X}_t \ = \ \boldsymbol{\mu} \text{ and } \mathbb{E}[(\mathbf{X}_1 - \boldsymbol{\mu})(\mathbf{X}_1 - \boldsymbol{\mu})^{'}] \ = \ \boldsymbol{\Sigma}, \tag{7}$$

where $\boldsymbol{\Sigma}$ is nonsingular and all diagonal terms nonzero and less than infinity. Furthermore, let the function $h(\cdot,\cdot)$ be anti-symmetric: $h(\mathbf{x},\mathbf{y};\boldsymbol{\theta}) \ = \ -h(\boldsymbol{y},\mathbf{x};\boldsymbol{\theta})$, where $\mathbf{x},\mathbf{y} \ \epsilon \ \mathbb{R}^n$, $\tilde{h}(\mathbf{u},\boldsymbol{\theta}) = \mathbb{E}h(\mathbf{X}_1, u;\boldsymbol{\theta})$ and such that

$$\mathbb{E}h^2(\mathbf{X}_1,\mathbf{X}_2;\boldsymbol{\theta}) \ < \ \infty \text{ and } 0 < \sigma^2 \ = \ \mathbb{E}\tilde{h}^2(X_2;\boldsymbol{\theta}) < \infty. \tag{8}$$

Define the stochastic process $Z_{[(T+1)\tau]}$, given below, as

$$Z_{[(T+1)\tau]} \ = \ \sum_{i=1}^{[(T+1)\tau]} \ \sum_{j=[(T+1)\tau]+1}^{T} h(\mathbf{X}_i,\mathbf{X}_j;\boldsymbol{\theta}). \tag{9}$$

The next proposition details the statements which can be made regarding the stochastic process given in (9).

Proposition 2.1. *Let the function $h(\mathbf{X}_1,\mathbf{X}_2;\theta)$ satisfy (8); and let $q \in Q$. Then we can define a sequence of Brownian bridges $\{B_T(\tau);\ 0 \leq \tau \leq 1\}$ such that, as $T \to \infty$,* *(i)* $$\sup_{0<\tau<1} \frac{\left|\frac{1}{\sigma}T^{-3/2}Z_{[(T+1)\tau]} - B_T(\tau)\right|}{q(\tau)} = \begin{cases} o_P(1), & \textit{if and only if } I(q,c) < \infty \quad \textit{for all } \; c > 0 \\ O_P(1), & \textit{if and only if } I(q,c) < \infty \quad \textit{for some } c > 0, \end{cases}$$

(ii) $$\sup_{0<\tau<1} \frac{\left|\frac{1}{\sigma}T^{-3/2}Z_{[(T+1)\tau]}\right|}{q(\tau)} \xrightarrow{\mathcal{D}} \sup_{0<\tau<1} \frac{|B(\tau)|}{q(\tau)},$$

if and only if $I(q,c) < \infty$ for some c. $B(\tau)$ is a Brownian bridge.

Proof. The results follow from the fact that

$$T^{-3/2}Z_{[(T+1)\tau]} = U_{1,T} - U_{1,[(T+1)\tau]} - U_{[(T+1)\tau],T} \tag{10}$$

where

$$\begin{aligned} U_{1,T} &= \frac{\sum_{i=1}^{T-1}\sum_{j=i+1}^{T} h(\mathbf{X}_i, \mathbf{X}_j; \boldsymbol{\theta})}{\binom{T}{2}}, \\ U_{1,[(T+1)\tau]} &= \frac{\sum_{i=1}^{[(T+1)\tau]-1}\sum_{j=i+1}^{T} h(\mathbf{X}_i, \mathbf{X}_j; \boldsymbol{\theta})}{\binom{[(T+1)\tau]}{2}}, \\ U_{[(T+1)\tau],T} &= \frac{\sum_{i=[(T+1)\tau]+1}^{T-1}\sum_{j=i+1}^{T} h(\mathbf{X}_i, \mathbf{X}_j; \boldsymbol{\theta})}{\binom{T-[(T+1)\tau]}{2}} \end{aligned} \tag{11}$$

are U-statistics. From Theorem 5.4.1 of Koroljuk and Borovshich (1994), the following holds:

$$\begin{aligned} \max_{1<k\leq T} k^2 |U_{1,T} - \frac{2}{k}\sum_{i=1}^{T} \tilde{h}(\mathbf{X}_i;\boldsymbol{\theta})| &= O_P(T) \\ \max_{1<k\leq T} k^2 |U_{1,k} - \frac{2}{k}\sum_{i=1}^{k} \tilde{h}(\mathbf{X}_i;\boldsymbol{\theta})| &= O_P(T), \\ \max_{1<k\leq T} (T-k)^2 |U_{k,T} - \frac{2}{T-k}\sum_{i=k+1}^{T} \tilde{h}(\mathbf{X}_i;\boldsymbol{\theta})| &= O_P(T). \end{aligned} \tag{12}$$

From this, it can be concluded that

$$\begin{aligned} \sup_{0<\tau<1} |Z_{[(T+1)\tau]} - \frac{2}{[(T+1)\tau]}\sum_{i=1}^{T} \tilde{h}(\mathbf{X}_i;\boldsymbol{\theta}) - \frac{2}{[(T+1)\tau]}\sum_{i=1}^{[(T+1)\tau]} \tilde{h}(\mathbf{X}_i;\boldsymbol{\theta}) \\ - \frac{2}{T-k}\sum_{i=k+1}^{T} \tilde{h}(\mathbf{X}_i;\boldsymbol{\theta})| = O_P(T). \end{aligned} \tag{13}$$

The remaining steps in the proof follow along the lines of the proofs of Theorems 6.2.1 and 5.2.1 of Szyszkowicz (1992). With this, statement i) of the proposition follows. For statement ii), we appeal to Theorem 1.1 of Szyszkowicz (1997) and employ the symmetrization method used in the proof of Theorem 5.2.1 of Szyszkowicz (1992) as well as the result established in (13).

2.1 Parameters Known

Here, the properties of the statistic defined in (6) will be explored. Before this can be done, the following additional assumptions will be made.

$$\begin{aligned} \mathbb{E}f(\mathbf{X}_i;\boldsymbol{\theta}) &= \gamma \\ \mathbb{E}f^2(\mathbf{X}_i;\boldsymbol{\theta}) - \gamma^2 &= \Delta^2, \end{aligned} \tag{14}$$

with γ a parameter. Now, as a special case of Proposition 2.1, the following statements can be made regarding the process defined in (6), each is detailed in Proposition 2.2.

Proposition 2.2. *Let* $\{\mathbf{X}_t\}_{t=1}^{\infty}$ *be a sequence of i.i.d multivariate rvs that satisfy* (7)*; let* $f(\mathbf{X}_1;\theta)$ *satisfy* (14)*; and let* $q \in Q$*. Then we can define a sequence of Brownian bridges* $\{B_T(\tau);\ 0 \le \tau \le 1\}$ *such that, as* $T \to \infty$*,*

(i) $$\sup_{0<\tau<1} \frac{\left|\frac{1}{\Delta}M_T^{(2)}(\tau) - B_T(\tau)\right|}{q(\tau)} = \begin{cases} o_P(1), & \text{if and only if } I(q,c) < \infty \quad \text{for all } \ c > 0 \\ O_P(1), & \text{if and only if } I(q,c) < \infty \quad \text{for some } c > 0, \end{cases}$$

(ii) $$\sup_{0<\tau<1} \frac{|\frac{1}{\Delta}M_T^{(2)}(\tau)|}{q(\tau)} \xrightarrow{\mathcal{D}} \sup_{0<\tau<1} \frac{|B(\tau)|}{q(\tau)},$$

if and only if $I(q,c) < \infty$ *for some* c*.*

Proof: This follows from Proposition 2.1, and observing that $M_T^{(2)}(\tau)$ can be expressed in terms of $Z_{[(T+1)\tau]}$.

2.2 Parameters Unknown

In most situations the vector of parameters $\boldsymbol{\theta}$ is unknown and must be estimated by a consistent sequence of estimators $\{\widehat{\boldsymbol{\theta}}_T\}_{T=1}^{\infty}$. This leads to the following slightly altered process,

$$\widehat{M}_T^{(2)}(\tau) := T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]} f(\mathbf{X}_t;\widehat{\boldsymbol{\theta}}_T) - \tau\sum_{t=1}^{T} f(\mathbf{X}_t;\widehat{\boldsymbol{\theta}}_T)\right\}, \quad 0 \le \tau \le 1. \tag{15}$$

With such substitution, it would seem natural that the limiting distribution of the slightly altered process detailed in (15) would be different from that detailed in Proposition 2.2, but Proposition 2.3 reveals otherwise.

To establish the main proposition of this section, we first use a result from Buck [(1965), Lemma page 244] to establish the following equality,

$$\begin{aligned}\widehat{M}_T^{(2)}(\tau) &= M_T^{(2)}(\tau) + (\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'\frac{1}{T^{1/2}}\left\{\sum_{t=1}^{[(T+1)\tau]} D_{\boldsymbol{\theta}} f(\mathbf{X}_t;\theta)|_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}}\right.\\ &\qquad \left. - \tau\sum_{t=1}^{T} D_{\boldsymbol{\theta}} f(\mathbf{X}_t;\boldsymbol{\theta})|_{\theta=\widetilde{\boldsymbol{\theta}}}\right\}, \quad 0 \le \tau < 1, \end{aligned} \tag{16}$$

for some $\widetilde{\boldsymbol{\theta}}$. $D_{\boldsymbol{\theta}}$ represents the vector of partial derivatives with respect to $\boldsymbol{\theta}$ and superscript $'$ represents the transpose operation.

Lemma 2.1. *Let $\{\mathbf{X}_t\}_{t=1}^{T}$ be i.i.d multivariate rvs that satisfy* (7)*; let $f(\mathbf{X}_1;\boldsymbol{\theta})$ satisfy* (14)*; and let $q(\cdot) \in Q$ with $I(q,c) < \infty$ for some $c > 0$. Then, as $T \to \infty$,*

$$\sup_{0<\tau<1} \frac{|\widehat{M}_T^{(2)}(\tau) - M_T^{(2)}(\tau)|}{q(\tau)} = o_P(1). \tag{17}$$

Proof. The following majorization can be obtained via equation (16).

$$\begin{aligned}\sup_{0<\tau<1}\frac{|\widehat{M}_T^{(2)}(\tau)-M_T^{(2)}(\tau)|}{q(\tau)} &\leq ||(\widehat{\boldsymbol{\theta}}_T-\boldsymbol{\theta})'||_{\mathbb{R}_p}\sup_{0<\tau<1}\left|\frac{1}{T^{1/2}}\sum_{t=1}^{[(T+1)\tau]}D_{\boldsymbol{\theta}}f(Y_t;\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}}\right. \\ &\quad \left. -\tau\frac{1}{T^{1/2}}\sum_{t=1}^{T}D_{\boldsymbol{\theta}}f(X_t;\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}}\right| \\ &= o_P(1)O_P(1), \qquad (18)\\ &= o_P(1),\ \text{as } T\to\infty,\end{aligned}$$

where the term $o_P(1)$ follows from consistency of the vector of estimators and $O_P(1)$ follows from Donsker's (1951) theorem restated on $D[0,\ 1]$. $||\cdot||_{\mathbb{R}_p}$ refers to Euclidean norm in $\mathbb{R}^p$.

Now, the main statement regarding the process developed in (15) can be made.

Proposition 2.3. *Let $\{\mathbf{X}_t\}_{t=1}^T$ be i.i.d rvs that satisfy (7); let $f(\mathbf{X}_1,\boldsymbol{\theta})$ satisfy (14); and let $q\in Q$. Then we can define a sequence of Brownian bridges $\{B_T(\tau);\ 0\leq\tau\leq 1\}$ such that, as $T\to\infty$,*

(i) $$\sup_{0<\tau<1}\frac{\left|\frac{1}{\Delta}\widehat{M}_T^{(2)}(\tau)-B_T(\tau)\right|}{q(\tau)}=\begin{cases} o_P(1), & \textit{if and only if } I(q,c)<\infty \quad \textit{for all } c>0\\ O_P(1), & \textit{if and only if } I(q,c)<\infty \quad \textit{for some } c>0,\end{cases}$$

(ii) $$\sup_{0<\tau<1}\frac{1}{\Delta}\frac{|\widehat{M}_T^{(2)}(\tau)|}{q(\tau)}\xrightarrow{\mathcal{D}}\sup_{0<\tau<1}\frac{|B(\tau)|}{q(\tau)},$$

only if $I(q,c)<\infty$ for some c.

Proof.

(i)

$$\begin{aligned}
\sup_{0<\tau<1} \frac{\left|\frac{1}{\Delta}\widehat{M}_T^{(2)}(\tau) - B_T(\tau)\right|}{q(\tau)} &\leq \sup_{0<\tau<1} \frac{\left|\frac{1}{\Delta}\widehat{M}_T^{(2)}(\tau) - M_T^{(2)}(\tau)\right|}{q(\tau)} \\
&\quad + \sup_{0<\tau<1} \frac{\left|\frac{1}{\Delta}M_T^{(2)}(\tau) - B_T(\tau)\right|}{q(\tau)} \\
&= o_P(1) + o_P(1) \text{ as } T \to \infty, \qquad (19)
\end{aligned}$$

where the last line in (19) follows from Lemma 2.1 and Proposition 2.2.

(ii)

This follows from statement ii) of Proposition 2.2 and Lemma 2.1 (Lemma 2.1 requires the integral condition to hold only for some $c > 0$). This implies that, as $T \to \infty$;

$$\begin{aligned}
&\left|\mathbb{P}\left\{\sup_{0<\tau<1} \frac{|\widehat{M}_T^{(2)}(\tau)|}{q(\tau)} \leq x\right\} - \mathbb{P}\left\{\sup_{0<\tau<1} \frac{|B(\tau)|}{q(\tau)} \leq x\right\}\right| \leq \left|\mathbb{P}\left\{\sup_{0<\tau<1} \frac{|\widehat{M}_T^{(2)}(\tau)|}{q(\tau)} \leq x\right\}\right. \\
&\quad \left. - \mathbb{P}\left\{\sup_{0<\tau<1} \frac{|M_T^{(2)}(\tau)|}{q(\tau)} \leq x\right\}\right| + \left|\mathbb{P}\left\{\sup_{0<\tau<1} \frac{|M_P^{(2)}(\tau)|}{q(\tau)} \leq x\right\}\right. \\
&\quad \left. -\mathbb{P}\left\{\sup_{0<\tau<1} \frac{|B(\tau)|}{q(\tau)} \leq x\right\}\right| \\
&= 0,
\end{aligned}$$

for all $x \in \mathbb{R}$.

3 Tests for Structural Change

The purpose of this section is to design a simultaneous and joint test that detect a change in intercept/slope and permit control of global error rates. It will be shown that the test uses OLS residuals directly rather than require calculation of recursive residuals as in the CUSUM test of BDE. An interesting by-product of the simultaneous test is that it exhibits power to detect a one-time change in intercept regardless of where it may occur in the sample - early, in the middle or later on.

Before these statistics can be fashioned, the appropriate metric must be constructed as well as some additional notation; let $D^2[0,\ 1] = D[0,\ 1]\ \mathrm{X}\ D[0,\ 1]$ and let the metric associated with this space be given by

$$\sup_{0<\tau<1} |x_1(\tau) - y_1(\tau)| + \sup_{0<\tau<1} |x_2(\tau) - y_2(\tau)|, \tag{20}$$

where $[x_1(\tau), x_2(\tau)]'$ and $[y_1(\tau), y_2(\tau)]'$ are elements of $D^2[0,\ 1]$.

3.1 Parameters Known

Consider the following process;

$$Y_t = \begin{cases} \beta_0^{(1)} + \boldsymbol{\beta}'^{(1)}\mathbf{X}_t + \sigma\varepsilon_t, & 1 \le t \le t^*, \\ \beta_0^{(2)} + \boldsymbol{\beta}'^{(2)}\mathbf{X}_t + \sigma\varepsilon_t, & t^* < t \le T. \end{cases} \tag{21}$$

where the ε_t's satisfy conditions detailed in (2). In addition, assume that at least one of the following holds: $\beta_0^{(1)} \neq \beta_0^{(2)}$ or $\boldsymbol{\beta}'^{(1)} \neq \boldsymbol{\beta}'^{(2)}$. The values of parameters $\beta_0^{(1)}$,

$\beta_0^{(2)}$, $\boldsymbol{\beta}'^{(1)}$, $\boldsymbol{\beta}'^{(2)}$, σ and $t^\star$ are all unknown. $\boldsymbol{\beta}'^{(1)}$ and $\boldsymbol{\beta}'^{(2)}$ are 1 x K vector of slope parameters. The data, $\{(Y_t, \mathbf{X}_t)'\}_{t=1}^{T}$, is a random sample.

The null and alternative hypothesis are as follows;

$$H_O : t^* \geq T$$

versus the alternative hypothesis of at-most-one change (AMOC) in intercept or slope;

$$H_A : 1 \leq t^* < T.$$

As advertised, the task here is to construct a test to detect such deviations. To construct such test two processes are required. These two processes will be constructed from two kernels each being unbiased for the intercept and variance parameters of LRMs, respectively. The kernel that is unbiased for the intercept sets $h(x, y; \beta_0, \boldsymbol{\beta}') = (y - \boldsymbol{\beta}'\mathbf{x}) - (y - \boldsymbol{\beta}'\mathbf{x})$; while that for the variance sets $h(\mathbf{x}, y; \beta_0, \boldsymbol{\beta}') = (y - \beta_0 - \boldsymbol{\beta}'\mathbf{x})^2 - (y - \beta_0 - \boldsymbol{\beta}'\mathbf{x})^2$. Even though the latter kernel is unbiased for the variance and not for slope parameters, it is employed here for two reasons. One, it closely corresponds to CUSUM of squares test of BDE and secondly, it will be shown in the section that studies the asymptotic behaviour of the statistics under one-time change in parameters (cf. Section 3.4), that a change in the slope will translate into a change in the variance of the residuals as long as a particular condition holds. Remark 3.3 will detail said condition.

Section 4 will confirm that the test statistic fashioned from the first kernel is particularly sensitive to a one-time change in the intercept, while the statistic fashioned from the second kernel is designed to detect deviations in the slope, and desirably, does not detect a one-time change in intercept when it occurs. With this in mind, these two kernels are now substituted into equation (6) which results in the following processes;

$$M_T^{(3)}(\tau) \;:=\; T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]}(Y_t-\beta_0^{'}-\boldsymbol{\beta}^{'}\mathbf{X}_t)^2-\tau\sum_{t=1}^{T}(Y_t-\beta_0^{'}-\boldsymbol{\beta}^{'}\mathbf{X}_t)^2\right\} \tag{22}$$

$$M_T^{(4)}(\tau) \;:=\; T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]}(Y_t-\boldsymbol{\beta}^{'}\mathbf{X}_t)-\tau\sum_{t=1}^{T}(Y_t-\boldsymbol{\beta}^{'}\mathbf{X}_t)\right\}. \tag{23}$$

As these processes remain a function of τ, they cannot be used in their present form to test the null hypothesis of no change in intercept/slope: that is, they are not yet statistics because of their dependency on τ. Here, interest centers on how large these processes can be for $0 < \tau < 1$. If there is in fact a change in one of the parameters: intercept or slope, the value of the supremum of the process that corresponds to the parameter that changed should be large. These considerations lead to the following test statistics:

$$\sup_{0<\tau<1}\frac{|M_T^{(i)}(\tau)|}{q(\tau)} \tag{24}$$

for $i = 3,\ 4$, where $q(\cdot)\ \in\ Q$. When $i = 3$ in (24), the test statistic can be used to test for AMOC in slope parameters; when $i = 4$ the statistic can be used to test for AMOC in intercept.

With the test statistics now defined, it is possible to make the following statement regarding their asymptotic behavior.

Proposition 3.1. *Assume H_O; let $\{(Y_t, \mathbf{X}_t)\}_{t=1}^T$ be a sequence of i.i.d rvs; let the conditions detailed in* (2) *hold; and let $q(\cdot) \in Q$. Then, as $T \to \infty$,*

$$M_T := \begin{bmatrix} \frac{1}{\sigma^2\sqrt{Var(\varepsilon_1^2)}} \frac{M_T^{(3)}(\tau)}{q(\tau)} \\ \frac{1}{\sigma} \frac{M_T^{(4)}(\tau)}{q(\tau)} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{B^{(1)}(\tau)}{q(\tau)} \\ \frac{\rho B^{(1)}(\tau) \,+\, (1-\rho^2)^{-1/2} B^{(2)}(\tau)}{q(\tau)} \end{bmatrix},$$

only if $I(q, c) < \infty$ for all $c > 0$. $B^{(1)}(\tau)$ and $B^{(2)}(\tau)$ are independent Brownian bridges, $\rho = \frac{\mathbb{E}[\varepsilon_1^3]}{\sqrt{Var(\varepsilon_1^2)}}$, and $\Rightarrow$ refers to weak convergence.

Proof. Let $||\cdot||$ be the metric on $D^2[0,\ 1]$ as defined in (20). Define two sequences of Brownian bridges $\{B_T^{(i)}(\tau);\ 0 \leq \tau \leq 1\}$ for $i = 1,\ 2$. Then, via statement (i) of Proposition 2.2, $||M_T - \mathbf{B}_T(\tau)|| = o_P(1)$, as $T \to \infty$, where $\mathbf{B}_T(\tau) = [B_T^{(1)}(\tau), B_T^{(2)}(\tau)]'$, is a sequence of bivariate Brownian Bridges and $'$ refers to the transpose.

Proposition 3.1 characterizes the limiting behaviour of test statistics (22) and (23) in terms of a vector of Brownian bridges that depend on unknown parameters. Such dependence poses a practical problem as the distribution of the limiting stochastic process depends on the correlation between error and the squared error. The following corollary introduces an alternative reformulation of the above proposition that solves this problem.

Corollary 3.1. *Under the same assumptions of Proposition* 3.1*, the following holds, as*

$T \to \infty,$

$$\begin{bmatrix} \frac{1}{\sigma^2\sqrt{Var(\varepsilon_1^2)}} \frac{M_T^{(3)}(\tau)}{q(\tau)} \\ \frac{-\rho((1-\rho^2)\sigma^4 Var(\varepsilon_1^2))^{-\frac{1}{2}} M_T^{(3)}(\tau) \;+\; ((1-\rho^2)\sigma^2)^{-\frac{1}{2}} M_T^{(4)}(\tau)}{q(\tau)} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{B^{(1)}(\tau)}{q(\tau)} \\ \frac{B^{(2)}(\tau)}{q(\tau)} \end{bmatrix}. \tag{25}$$

Using Corollary 3.1 in conjunction with the continuous mapping theorem, it is possible to conclude with the following statement regarding the bivariate distribution of the supremum of the processes developed in (22) and (23).

Corollary 3.2. *Under the same assumptions of Proposition* 3.1*, the following holds, as* $T \to \infty,$

$$\begin{bmatrix} \sup_{0<\tau<1} \frac{1}{\sigma^2\sqrt{Var(\varepsilon_1^2)}} \frac{|M_T^{(3)}(\tau)|}{q(\tau)} \\ \sup_{0<\tau<1} \frac{|-\rho((1-\rho^2)\sigma^4 Var(\varepsilon_1^2))^{-\frac{1}{2}} M_T^{(3)}(\tau) \;+\; ((1-\rho^2)\sigma^2)^{-\frac{1}{2}} M_T^{(4)}(\tau)|}{q(\tau)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \sup_{0<\tau<1} \frac{|B^{(1)}(\tau)|}{q\tau)} \\ \sup_{0<\tau<1} \frac{|B^{(2)}(\tau)|}{q(\tau)} \end{bmatrix}. \tag{26}$$

The U-statistic type processes and the corresponding asymptotic theory introduced above enable us now to introduce two different test statistics for the above null hypothesis of no change in either intercept or slope in linear regression models.

In particular, our simultaneous test is defined by the hypothesis

$$H_{O,sim} : \beta_0^{(1)} \;=\; \beta_0^{(2)} \, and \, \boldsymbol{\beta}'^{(1)} = \boldsymbol{\beta}'^{(2)}, \tag{27}$$

versus

$$H_{A,sim} : \beta_0^{(1)} \;\neq\; \beta_0^{(2)} \, or \, \boldsymbol{\beta}'^{(1)} \neq \boldsymbol{\beta}'^{(2)}. \tag{28}$$

The appropriate test statistic is RM_T, the simultaneous test, and $R = [0 \quad 0; 0 \quad 1]$ is the selector matrix. By Corollary 3.2 and the continuous mapping theorem the asymptotic distribution of this test is $\sup_{0<\tau<1} \frac{|B^{(2)}(\tau)|}{q\tau)}$, that is parameter free, implying critical values for the test that can be universally tabulated via simulation.

The nature of the test, however, does not distinguish the source of the rejection, that is, whether intercept or slope have changed after some t^*. This can be corrected by exploiting the U-statistic type process $\frac{M_T^{(3)}(\tau)}{q(\tau)}$ and the bivariate distribution derived in Corollary 3.2. More specifically, it has been discussed that this process is devised to detect deviations of the slope and not from the intercept. We exploit this property by devising an auxiliary test $H_{O,slope} : \boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}^{(2)}$ versus $H_{A,slope} : \boldsymbol{\beta}^{(1)} \neq \boldsymbol{\beta}^{(2)}$, used to define a joint test that can be carried out in one step and that controls for the global error rate.

The joint test is $H_{O,joint} = H_{O,sim}$ versus

$$H_{A,joint} := \begin{cases} H_{A,intercept} : H_{O,slope} \bigcap H_{A,sim} \\ H_{A,slope} \end{cases}, \tag{29}$$

where $H_{A,slope}$ and $H_{A,intercept}$ define a test for change only in slope, and a change only in intercept, respectively. To be more specific, if the simultaneous rejects but the test statistic based on $\frac{M_T^{(3)}(\tau)}{q(\tau)}$ accepts, then $H_{A,intercept}$ holds; there is a change only in intercept. Otherwise, one can conclude only a slope parameter has changed, while no statement can be made regarding a change in intercept.

This joint test has several interesting features. First, the standardization provided in Corollary 3.1 guarantees that the marginal asymptotic distributions are independent and identically distributed. This implies that the critical values of each test at the same significance level are identical. Furthermore, one can control the global error rate by simply taking the product of one minus the error rate idiosyncratic to each marginal test and noting the global error rate is one minus this product; i.e, if the idiosyncratic error is 5% then then the global error rate is $1 - (1 - 0.05)^2 = 0.091$. Finally note that for symmetric error distributions, $\rho = 0$, and the joint test boils down to two independent hypothesis tests based on the marginal U-statistic type processes $\frac{M_T^{(3)}(\tau)}{q(\tau)}$ and $\frac{M_T^{(4)}(\tau)}{q(\tau)}$.

As mentioned above, in the joint test, if $\frac{M_T^{(3)}(\tau)}{q(\tau)}$ rejects there remains the question of whether the intercept has changed. In this situation a second layer of hypothesis testing must be considered. One would then run an individual test, based on $\sup_{0<\tau<1} \frac{|M_T^{(4)}(\tau)|}{q(\tau)}$, suited for changes in intercept. Its asymptotic distribution is detailed in Proposition 2.2 statement ii). $H'_{O,intercept} : \beta_0^{(1)} = \beta_0^{(2)}$ versus $H'_{A,intercept} : \beta_0^{1)} \neq \beta_0^{(2)}$. If this test rejects the null hypothesis of no change in intercept then one concludes there was a change in both slope and intercept. Otherwise one concludes only the slope has changed. Unfortunately in this case, we lose independence between the asymptotic distribution of this test and the marginal distributions of the joint test and therefore lose control of the global error rate.

3.2 Parameters Unknown

The processes defined in (22) and (23) depend on unknown parameters. OLS will produce consistent estimators of β_0 and $\boldsymbol{\beta}$ under $H_{O,sim}$; let these sequences of estimators be denoted $\{\widehat{\beta}_{T,0}\}_{T=1}^{\infty}$ and $\{\widehat{\boldsymbol{\beta}}_T\}_{T=1}^{\infty}$. When these sample estimates are substituted for the population parameters, this produces the following slightly altered sequence of partial sum processes;

$$\widehat{M}_T^{(3)}(\tau) := T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]}(Y_t-\widehat{\beta}_{T,0}'-\widehat{\boldsymbol{\beta}}_T'\mathbf{X}_t)^2-\tau\sum_{t=1}^{T}(Y_t-\widehat{\beta}_{T,0}'-\widehat{\boldsymbol{\beta}}_T'\mathbf{X}_t)^2\right\} \quad (30)$$

$$\widehat{M}_T^{(4)}(\tau) := T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]}(Y_t-\widehat{\boldsymbol{\beta}}_T'\mathbf{X}_t)-\tau\sum_{t=1}^{T}(Y_t-\widehat{\boldsymbol{\beta}}_T'\mathbf{X}_t)\right\}. \quad (31)$$

Proposition 3.2. *Assume H_O; let $\{(Y_t,\mathbf{X}_t)'\}_{t=1}^{T}$ be a sequence of i.i.d rvs; let the conditions detailed in (2) hold; and let $q(\cdot)\in Q$. Then, as $T\to\infty$,*

$$\begin{bmatrix} \frac{1}{\sigma^2\sqrt{Var(\varepsilon_1^2)}}\frac{\widehat{M}_T^{(3)}(\tau)}{q(\tau)} \\ \frac{1}{\sigma}\frac{\widehat{M}_T^{(4)}(\tau)}{q(\tau)} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{B^{(1)}(\tau)}{q(\tau)} \\ \frac{-\rho B^{(1)}(\tau)\;+\;(1-\rho)^{-\frac{1}{2}}B^{(2)}(\tau)}{q(\tau)} \end{bmatrix},$$

only if $I(q,c)<\infty$ for all $c>0$.

Proof. This follows from Propositions 2.3, and Lemma 2.1.

As in the case when the parameters were known, there is a similar statement to Corollary 3.1 that can be made regarding the sequence of partial sum processes detailed in (30) and (31).

Corollary 3.3. *Under the same assumptions of Proposition* 3.2*, the following holds, as* $T \to \infty$*,*

$$\begin{bmatrix} \frac{1}{\sigma^2\sqrt{Var(\varepsilon_1^2)}}\frac{\widehat{M}_T^{(3)}(\tau)}{q(\tau)} \\ \frac{-\rho((1-\rho^2)\sigma^4 Var(\varepsilon_1^2))^{-\frac{1}{2}}\widehat{M}_T^{(3)}(\tau)\ +\ ((1-\rho^2)\sigma^2)^{-\frac{1}{2}}\widehat{M}_T^{(4)}(\tau)}{q(\tau)} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{B^{(1)}(\tau)}{q(\tau)} \\ \frac{B^{(2)}(\tau)}{q(\tau)} \end{bmatrix}.$$

Using Corollary 3.3 in conjunction with the continuous mapping theorem, it is possible to conclude with the following statement regarding the bivariate distribution of the supremum of the processes developed in (30) and (31).

Corollary 3.4. *Under the same assumptions of Proposition* 3.1*, the following holds, as* $T \to \infty$*,*

$$\begin{bmatrix} \sup\limits_{0<\tau<1}\frac{1}{\sigma^2\sqrt{Var(\varepsilon_1^2)}}\frac{|\widehat{M}_T^{(3)}(\tau)|}{q(\tau)} \\ \sup\limits_{0<\tau<1}\frac{|-\rho((1-\rho^2)\sigma^4 Var(\varepsilon_1^2))^{-\frac{1}{2}}\widehat{M}_T^{(3)}(\tau)\ +\ ((1-\rho^2)\sigma^2)^{-\frac{1}{2}}\widehat{M}_T^{(4)}(\tau)|}{q(\tau)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \sup\limits_{0<\tau<1}\frac{|B^{(1)}(\tau)|}{q(\tau)} \\ \sup_{0<\tau<1}\frac{|B^{(2)}(\tau)|}{q(\tau)} \end{bmatrix}. \tag{32}$$

Remark 3.1. *Let* $\{\widehat{\rho}_T\}_{T=1}^{\infty}$*,* $\{\widehat{\sigma}_T\}_{T=1}^{\infty}$ *and* $\{\widehat{Var(\varepsilon_1^2)}_T\}_{T=1}^{\infty}$ *be sequences of consistent estimators of* ρ*,* σ *and* $Var(\epsilon_1^2)$*. Note that Proposition* 3.2 *and Corollaries* 3.3 *and* 3.4 *continue to hold when the parameters are replaced by the above estimators.*

When the null hypothesis of no change in intercept or slope is rejected, it becomes important then to locate the value of the sample where this change occurred. One estimator that has been suggested is given below in equation (33). The properties of this estimator have been studied by GHH (1996) and others (cf. Antoch and Hušková (1995) and Ferger (2001)).

$$\hat{t}^{*i} \;=\; \min\left\{ t; \frac{|M_T^{(i)}(t)|}{q(\frac{t}{T})} \;=\; \min_{1\le t\le T} \frac{|M_T^{(i)}(t)|}{q(\frac{t}{T})} \right\} \tag{33}$$

for $i = 3,\ 4$. The information contained in this estimator allows the researcher to develop different LRMs; one for data up to and including $\hat{t}^{*i}$ and the other for data that occur after this estimate.

Up to now no discussion on the nature of the weight functions $q(\cdot)$ has been made. We correct this oversight now. One family of weight functions that has received some attention is due to GHH. This family of functions is given below;

$$q(\tau,\nu) := \{(\tau(1-\tau))^{\nu};\ 0 \le \nu < 1/2\}. \tag{34}$$

This class of functions has been shown to be sensitive to a change that occurs both early and later on in the sample (cf. Olmo and Pouliot (2008)). Moreover, this class is a member of Q and satisfies $I(q,c) < \infty$ for all $c > 0$. Since this condition holds for all $c > 0$, it is possible to construct the simultaneous and joint tests for AMOC in parameters, intercept or slope, of the linear model.

3.3 Dynamic LRM

The LRM model detailed in (21) is now altered to allow for lagged dependent variables. To accommodate this alteration, let

$$Y_t = \begin{cases} \beta_0^{(1)} + \gamma_1 Y_{t-1} + \cdots + \gamma_m Y_{t-m} + \boldsymbol{\beta}'^{(1)}\mathbf{X}_t + \sigma\varepsilon_t, & 1 \le t \le t^\star, \\ \beta_0^{(2)} + \gamma_1 Y_{t-1} + \cdots + \gamma_m Y_{t-m} + \boldsymbol{\beta}'^{(2)}\mathbf{X}_t + \sigma\varepsilon_t, & t^\star < t \le T. \end{cases} \tag{35}$$

where the ε_t's satisfy conditions detailed in (2). In addition, assume that at least one of the following holds: $\beta_0^{(1)} \neq \beta_0^{(2)}$ or $\boldsymbol{\beta}'^{(1)} \neq \boldsymbol{\beta}'^{(2)}$. The values of parameters $\beta_0^{(1)}$, $\beta_0'^{(2)}$, $\boldsymbol{\beta}'^{(1)}$, $\boldsymbol{\beta}'^{(2)}$, σ and $t^\star$ are all unknown. $\boldsymbol{\beta}'^{(1)}$ and $\boldsymbol{\beta}'^{(2)}$ are 1 x K vector of slope parameters. The data, $\{(Y_t, Y_{t-1}, \ldots, Y_{t-m}, \mathbf{X}_t)'\}_{t=m+1}^{T}$, is a random sample. With this notation, it is now possible to conclude with the following statement regarding the processes detailed in (22) and (23).

Corollary 3.5. *Assume H_O; let $\{(Y_t, Y_{t-1}, \ldots, Y_{t-m}, \mathbf{X}_t)'\}_{t=m+1}^{T}$ be a sequence of i.i.d rvs; let the conditions detailed in (2) hold; and let $q(\cdot) \in Q$ and the dynamic LRM specified in (35). Then we can define a sequence of Brownian bridges $\{B_T(\tau);\ 0 \le \tau \le 1\}$ for $i = 3, 4$ such that, as $T \to \infty$,*

(i) $$\sup_{0<\tau<1} \frac{\left|\frac{1}{\Delta^{(i)}} M_T^{(i)}(\tau) - B_T(\tau)\right|}{q(\tau)} = \begin{cases} o_P(1), & \textit{if and only if } I(q,c) < \infty \quad \textit{for all } \ c > 0 \\ O_P(1), & \textit{if and only if } I(q,c) < \infty \quad \textit{for some } c > 0, \end{cases}$$

(ii) $\sup_{0<\tau<1} \frac{1}{\Delta^{(i)}} \frac{|M_T^{(i)}(\tau)|}{q(\tau)} \xrightarrow{\mathcal{D}} \sup_{0<\tau<1} \frac{|B(\tau)|}{q(\tau)}$, *if and only if $I(q,c) < \infty$ for some $c > 0$, $i = 3,\ 4$.*

Proof. The corollary follows from the fact that processes (22) and (23) have, under the LRM specified in (35), the following representation:

$$
\begin{aligned}
M_T^{(3)}(\tau) &= T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]-m} \varepsilon_{t+m}^2 - \tau \sum_{t=1}^{T-m} \varepsilon_{t+m}^2\right\} \\
M_T^{(4)}(\tau) &= T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]-m} \varepsilon_{t+m} - \tau \sum_{t=1}^{T-m} \varepsilon_{t+m}\right\}.
\end{aligned}
$$

As a result of this representation one can directly apply Theorem 2.1 of Szyszkowicz (1991) to obtain the result detailed in i) and ii) of the proposition.

A similar result to Corollary 3.5 can be extended to processes (30) and (31) via the following lemma.

Lemma 3.1. *Under the same conditions as Corollary* 3.5 *along with the consistency of* $\{\widehat{\gamma}_{l,T}\}_{T=1}^{\infty}$ *for* $l = 1, \ldots, m$, $\{\widehat{\beta}_{T,0}\}_{T=1}^{\infty}$ *and* $\{\widehat{\boldsymbol{\beta}}_{T,0}\}_{T=1}^{\infty}$, *then as* $T \to \infty$,

$$
\sup_{0<\tau<1} \frac{|M_T^{(i)}(\tau) \; - \; \widehat{M}_T^{(i)}(\tau)|}{q(\tau)} \; = \; o_P(1),
$$

only if $I(q, c) < \infty$ *for some* $c > 0$ *and* $i = 3, 4$.

Proof: The lemma follows from the following decomposition of processes (30) and (31).

$$
\begin{aligned}
\widehat{M}_T^{(3)}(\tau) &= M_T^{(3)}(\tau) + o_P(1) \\
\widehat{M}_T^{(4)}(\tau) &= M_T^{(4)}(\tau) + (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_T)' T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]-m} \mathbf{X}_{t+m} \; - \; \tau \sum_{t=1}^{T-m} \mathbf{X}_{t+m}\right\} \\
&\quad + \sum_{l=1}^{m} (\gamma_l - \widehat{\gamma}_{T,l}) T^{-1/2}\left\{\sum_{t=1}^{[(T+1)\tau]-m} Y_{t-l+m} \; - \; \tau \sum_{t=1}^{T-m} Y_{t-l+m}\right\}.
\end{aligned}
$$

Remark 3.2. *As a result of Lemma* 3.1*, Proposition* 3.2 *and Corollaries* 3.3 *and* 3.4 *continue to hold for processes* (30) *and* (31) *in the dynamic LRM setting.*

3.4 Asymptotics Under the Alternative Hypothesis

Here, the asymptotics of statistics defined as supremum of (22) and (23) are studied. The first of two theorems to follow describes the distribution of statistics (22) under local alternatives of AMOC change in the slope.

Proposition 3.3. *Assume H_A, moment conditions* (2)*, equation* (21)*, $t^\star = [T\tau^\star]$, $\tau^\star \in (0, 1)$ hold and $\boldsymbol{\beta}^{(2)} = \boldsymbol{\beta}^{(1)} + \boldsymbol{\delta}$. Then $\sigma^{\star 2} = I\!E(Y_1 - \beta_0^{(1)} - \boldsymbol{\beta}^{(1)'}\mathbf{X}_1)^2 + \boldsymbol{\delta}' I\!E[\mathbf{X_1}\mathbf{X}_1']\boldsymbol{\delta}$, with $\boldsymbol{\delta} = \boldsymbol{\delta}(T) \to 0$, as $T \to \infty$. Let $q(\cdot) \in Q$ with $I(q, c) < \infty$ for some $c > 0$, then as $T \to \infty$,*

$$\frac{q(\tau^\star)}{\sqrt{\tau^\star(1-\tau^\star)}} \frac{1}{\sigma^2\sqrt{Var(\varepsilon_1^2)}} \left\{ \sup_{0<\tau<1} \frac{|M_T^{(3)}(\tau)|}{q(\tau)} - T^{1/2}\boldsymbol{\delta}' I\!E[\mathbf{X}_1\mathbf{X}_1']\boldsymbol{\delta}\frac{t^\star}{T}\frac{\left(1-\frac{t^\star}{T}\right)}{q(\frac{t^\star}{T})} \right\} \xrightarrow{\mathcal{D}} N(0,1).$$

Proof. This follows from Theorem 1.4 of GHH (1996).

Remark 3.3. *Proposition* 3.3 *reveals that a one-time change in slope parameters will cause a one-time change in variance if and only if the following condition holds:*

$$\boldsymbol{\delta}'\mathbf{X}_1 \neq 0.$$

A direct result of Proposition 3.3 is the consistency of this test for a one-time change in slope. This result is formally introduced in the next corollary.

Corollary 3.6. *Under the conditions of Proposition* 3.3*, and as* $T \to \infty$*,*

$$\frac{1}{T^{1/2}\boldsymbol{\delta}' \mathbb{E}[\mathbf{X}_1\mathbf{X}_1']\sigma\boldsymbol{\delta}} \sup_{0<\tau<1} \frac{|M_T^{(3)}(\tau)|}{q(\tau)} \xrightarrow{P} \frac{\tau^\star(1-\tau^\star)}{q(\tau^\star)}.$$

The next proposition details the asymptotic distribution of the AMOC in intercept statistic (cf. (23)).

Proposition 3.4. *Assume* H_A*, moment conditions* (2), (21), $t^\star = [T\tau^\star]$, $\tau^\star \in (0, 1)$ *and* $\beta_0^{(2)} = \beta_0^{(1)} + \Lambda$ *hold. Then for* $q(\cdot) \in Q$ *with* $I(q, c) < \infty$ *for some* $c > 0$*, and as* $T \to \infty$*,*

$$\frac{q(\tau^\star)}{\sigma\sqrt{\tau^\star(1-\tau^\star)}} \left\{ \sup_{0<\tau<1} \frac{|M_T^{(4)}(\tau)|}{q(\tau)} \; - \; T^{1/2}\Lambda\frac{t^\star}{T}\frac{\left(1-\frac{t^\star}{T}\right)}{q(\frac{t^\star}{T})} \right\} \xrightarrow{\mathcal{D}} N(0,1).$$

Proof. Without loss of generality, let $\frac{t^\star}{T} > \tau$, $t^\star = [(T+1)\tau^\star]$ and assume $\delta(T) \to 0$, as

$T \to 0$ and $\delta(T)T \to 0$, as $T \to \infty$. Then

$$\begin{aligned}
\sup_{\frac{t^\star}{T}-\delta(T)<\tau<\frac{t^\star}{T}+\delta(T)} \frac{1}{q(\tau)}\frac{|M^{(4)}(\tau)|}{q(\tau)} &= T^{-1/2} \sup_{\frac{t^\star}{T}-\delta(T)<\tau<\frac{t^\star}{T}+\delta(T)} \Bigg| \sum_{t=1}^{[(T+1)\tau]} (Y_t - \boldsymbol{\beta}' X_t) - \tau \sum_{t=1}^{t^\star} (Y_t - \boldsymbol{\beta}' X_t) \\
&\qquad - \tau \sum_{t=t^\star+1}^{T} (Y_t - \boldsymbol{\beta}' X_t) \Bigg| \\
&= T^{-1/2} \sup_{\frac{t^\star}{T}-\delta(T)<\tau<\frac{t^\star}{T}+\delta(T)} \frac{1}{q(\tau)} \Bigg| \sum_{t=1}^{[(T+1)\tau]} (Y_t - \boldsymbol{\beta}' X_t) - \tau \sum_{t=1}^{t^\star} (Y_t - \boldsymbol{\beta}' X_t) \\
&\qquad - \tau \sum_{t=t^\star+1}^{T} (Y_t - \boldsymbol{\beta}' X_t) \Bigg| \\
&= T^{-1/2} \sup_{\frac{t^\star}{T}-\delta(T)<\tau<\frac{t^\star}{T}+\delta(T)} \frac{1}{q(\tau)} \Bigg| \sum_{t=1}^{[(T+1)\tau]} ((Y_t - \boldsymbol{\beta}' X_t) - \beta_0^{(1)}) - \tau \sum_{t=1}^{t^\star} ((Y_t - \boldsymbol{\beta}' X_t) - \beta_0^{(1)}) \\
&\qquad - \tau \sum_{t=t^\star+1}^{T} ((Y_t - \boldsymbol{\beta}' X_t) - \beta_0^{(2)}) \;+\; ([(T+1)\tau] - \tau t^\star)\beta_0^{(1)} - \tau (T - t^\star)\beta_0^{(2)} \Bigg| \\
&= T^{-1/2} \sup_{\frac{t^\star}{T}-\delta(T)<\tau<\frac{t^\star}{T}+\delta(T)} \left| \sigma \left(\sum_{t=1}^{[(T+1)\tau]} \varepsilon_t - \tau \sum_{t=1}^{T} \varepsilon_i \right) \;-\; \tau (T - t^\star)\Lambda \right| \\
&= \left| \frac{\sigma}{T^{1/2}} \left(\sum_{t=1}^{t^\star} \varepsilon_t - \tau \sum_{t=1}^{T} \varepsilon_i \right) \;-\; T^{1/2} \frac{t^\star}{T} \left(1 - \frac{t^\star}{T} \right) \Lambda \right| \qquad (36)
\end{aligned}$$

Lemma 3.2, found below, will be needed to establish the proposition. The absolute value in equation (36) can be removed as it has no effect on the limiting distribution: that is, when inside the absolute value is negative, simply multiply by -1 and remove the absolute value. Hence, we have

$$\frac{\sigma}{T^{1/2}} \left(\sum_{t=1}^{t^\star} \varepsilon_t - \tau \sum_{t=1}^{T} \varepsilon_t \right) \;+\; T^{1/2} \frac{t^\star}{T} \left(1 - \frac{t^\star}{T} \right) \Lambda. \qquad (37)$$

Now, Lemma 3.2 and (36) establish the above proposition.

Lemma 3.2. *Under the same conditions as specified in Proposition* 3.3*, and as* $T \to \infty$*,*

$$\begin{bmatrix} \frac{\sum_{t=1}^{t^\star} \varepsilon_t}{T^{1/2}} \\ \\ \frac{\sum_{t=1}^{T} \varepsilon_t}{T^{1/2}} \end{bmatrix} \xrightarrow{\mathcal{D}} N \left[\begin{pmatrix} 0 \\ \\ 0 \end{pmatrix}, \ \Psi \right],$$

where

$$\Psi = \sigma_1^2 \begin{bmatrix} \tau^{\star 2} & \tau^\star \\ \\ \tau^\star & 1 \end{bmatrix}.$$

Proof. This follows from the bivariate version of the Lindberg-Levy Central Limit Theorem.

A corollary similar to Corollary 3.6 holds here as well and is a direct consequence of Proposition 3.4.

Corollary 3.7. *Under the conditions of Proposition* 3.4*, and as* $T \to \infty$*,*

$$\frac{1}{\Lambda \ T^{1/2}} \sup_{0<\tau<1} \frac{|M_T^{(4)}(\tau)|}{q(t)} \xrightarrow{P} \frac{\tau^\star(1-\tau^\star)}{q(\tau^\star)}.$$

4 Monte Carlo Simulation

This section is concerned with the comparison of the power of statistics developed from (15) with the CUSUM test of BDE and fluctuation test of PKK (1989). The comparison of power of each test will be done via simulation. Even though power is an important criteria for comparison, the accuracy of the nominal size of the tests should also be considered. Both criteria, power and accuracy of nominal coverage, were adopted by Kramer, Ploberger and Alt (hereafter KPA) (1988, page 1359) to evaluate performance of the BDE CUSUM with their fluctuation test within a dynamic LRM. As in their study, both criteria will be adopted here as well.

For the purpose of this simulation, the entertained model is given by

$$Y_t = \begin{cases} \beta_0^{(1)} + \beta^{(1)} X_t + \sigma\varepsilon_t, & 1 \le t \le t^\star, \\ \beta_0^{(2)} + \beta^{(2)} X_t + \sigma\varepsilon_t, & t^\star < t \le T. \end{cases} \tag{38}$$

where the ε_t's satisfy conditions detailed in (2), and the corresponding change point hypothesis test is

$$H_O : t^\star \ge T$$

versus the one-time change alternative,

$$H_A : 1 \le t^\star < T.$$

Interest here is with alternatives that involve a small change in intercept as well as slope; in small sample sizes; and in detecting a change in either intercept/slope when it occurs early and later on in the sample.

Within the LRM specified in (38), the specific alternatives considered for the intercept are $\beta^{(1)} = \beta^{(2)} = 1$, while $\beta_0^{(2)} = 1.25, 1.5, 1.75, 2$. The change in intercept considered in this simulation increased from a 125% - a small change, to 175% - a moderate change, to 200% - a large change. The sample size considered here ranged from $T = 75$ - a small size, to $T = 100$ - a moderate size and then $T = 125$ - a large size. Since interest is also with the skew of the distribution that generates the errors, it will first be assumed that $\varepsilon_t \overset{\mathcal{D}}{=} \chi^2$ with 1 degree of freedom, for $t = 1, \ldots, T$ and a second simulation will then assume $\varepsilon_t \overset{\mathcal{D}}{=} N(0, 1)$ for $t = 1, \ldots, T$. Note that the structure of our test statistics based on U-statistic type processes implies that the tests are invariant to location transformations of the error distribution. For simplicity, then, we only study the χ^2, defined in the positive domain.

A third and fourth simulation will explore a one-time change in slope parameter allowing the errors of the LRM to follow the two distributions specified in the first two simulations: that is, χ^2 and standard normal. The LRM for this part of the Monte Carlo exercise is specified as follows: the LRM detailed in (38) with $\beta_0^{(1)} = \beta_0^{(2)} = 1$ and $\beta^{(1)} = 3$ and under the alternative $\beta^{(2)} = 3 + \delta$ where $\delta = 0.75,\ 1.5,\ 2.25,\ 3.$

As the CUSUM test statistic of BDE and the fluctuation test of PKK are the competitors here, a brief introduction to each will be provided below. The CUSUM test of BDE is based on recursive residuals, standardized appropriately. In particular, the cumulative sum of recursive residuals is given by

$$W^{(r)} = \frac{1}{\hat{\sigma}} \sum_{t=K+2}^{r} w_t, \tag{39}$$

where w_t is the recursive residual. This leads to an equivalent test statistic detailed by the following formula

$$\text{CUSUM Test} := \max_{K+1<r\leq T} \frac{\frac{|W_t^{(r)}|}{\sqrt{T-K-1}}}{1 + 2\frac{r-K-1}{T-K-1}}. \tag{40}$$

In this formula, T refers to the sample size and K, the number of slope parameters. The null hypothesis of parameter constancy is rejected whenever BDE statistic exceeds some critical value.

The fluctuation test of PKK (1989) is based on estimates of the parameters from a LRM. Define $\mathbf{X}^{(t)} = [\mathbf{x}_1, \ldots, \mathbf{x}_t]'$, $\mathbf{Y}^{(t)} = [Y_1, \ldots, Y_t]'$, $t = 1, \ldots, T$ and $\boldsymbol{\beta}^{(t)} = (\mathbf{X}^{(t)\prime}\mathbf{X}^{(t)})^{(-1)}\mathbf{X}^{(t)}\mathbf{y}^{(t)}$

for $t = K, \ldots, T$. The test statistic is defined as

$$S^{(T)} = \max_{t=K,\ldots,T} \frac{t}{\widehat{\sigma} T} ||(\mathbf{X}^{(T)\prime}\mathbf{X}^{(T)})^{1/2}(\widehat{\boldsymbol{\beta}}^{(t)} - \widehat{\boldsymbol{\beta}}^{(T)}||_{\infty}, \tag{41}$$

where $||\widehat{\boldsymbol{\beta}}^{(t)} - \widehat{\boldsymbol{\beta}}^{(T)}||_{\infty} = \max_{t=1,\ldots,K} |\widehat{\boldsymbol{\beta}}^{(t)} - \boldsymbol{\beta}^{(T)}|$. The test statistic $S^{(T)}$ rejects H_O, given below, of a one-time change in $\boldsymbol{\beta}$ of the LRM whenever it is too large, i.e. the parameter estimates fluctuate too much.

4.1 Estimation Effects

In this section the LRM (cf. (38)) is estimated and the statistic calculated first under the assumption that $t^{\star} \geq T$ which provides an estimate of nominal coverage of these tests and then with a one-time change in the intercept β_0. This will allow a more realistic assessment of the ability of the newly fashioned statistics to detect a change in intercept and follows closely the criteria used by PKA. The first simulation considered here sets the distribution of the error term in the LRM to be a χ^2 with one degree of freedom and then a second simulation sets the errors as standard normal rvs. The results from the simulation with the first choice of error distribution are tabulated under the null hypothesis of no change and are recorded in Table I, while Table II records results for a one-time change in intercept. Table I reveals that the nominal coverage of all the test statistics under study, except the fluctuation test of PKK, achieve a nominal coverage of 8% or less - the significance level throughout the simulations will be 5%. The nominal coverage of the fluctuation tests of PKK was over 20% (cf Table I bold numbers) for all sample sizes considered here. As a result of this consideration and employing the first criteria of KPA, the fluctuation test is not appropriate for the sample sizes entertained here. On the contrary, the three statistics constructed from the U-statistic process (cf. Section

3) and weighted by function $q(\tau, \nu = \frac{15}{128})$ perform very well in terms of nominal coverage as the coverage is less than or equal 8%. As a by-product of this research, it was found that $\nu = \frac{15}{128}$ performed better - in terms of nominal coverage - than other choices of ν.

Table I			
	T = 75	T = 100	T=125
$\sup_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.02	0.02	0.02
$\sup_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.046	0.046	0.038
CUSUM Test	0.044	0.026	0.044
FLUCTUATION	**0.192**	**0.18**	**0.152**
SIMULT	0.08	0.08	0.07

Table II, found below, details the results from the simulation under $H_{A,intercept}$ with a one-time change in intercept. Since the errors of the LRM were generated from a χ^2 distribution with 1 degree of freedom, the third moment is not 0. This has a positive effect on the finite sample properties of the simultaneous test (cf. SIMULT in Table II) as the empirical power for a one-time change in intercept that occurs on the middle of the sample ranges from a low of 0.26 when the sample size is only 75 and intercept is increased by 25% to a high of 1. For a change in intercept that occurs early or later on in the sample, the simultaneous test exhibits low empirical power when the sample size is small (T=75). Otherwise, the simultaneous test's power reaches a high of 0.80 and 0.67 when the change occurs early and then later on in the sample, respectively. When the simultaneous test is compared to the CUSUM test, its performance in terms of empirical power is strikingly better for all values of the sample size

Table II

MIDDLE OF SAMPLE ($\tau^\star = 0.5$)												
	$\beta_0^{(2)} = 1.25$			$\beta_0^{(2)} = 1.5$			$\beta_0^{(2)} = 1.75$			$\beta_0^{(2)} = 2$		
Statistic	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup\limits_{0<t<1} \frac{\|M_T^{(3)}(\tau)\|}{q(\tau,\nu=\frac{15}{218})}$	0.02	0.02	0.02	0.03	0.01	0.04	0.03	0.01	0.03	0.04	0.03	0.02
$\sup\limits_{0<t<1} \frac{\|M_T^{(4)}(\tau)\|}{q(\tau,\nu=\frac{15}{218})}$	0.04	0.05	0.08	0.13	0.18	0.26	0.34	0.53	0.63	0.61	0.76	0.88
CUSUM Test	0.06	0.07	0.08	0.10	0.13	0.16	0.21	0.26	0.34	0.4	0.45	0.58
FLUCTUATION	0.23	0.20	0.22	0.29	0.39	0.43	0.53	0.62	0.71	0.76	0822	0.90
SIMULT	0.2	0.17	0.22	0.42	0.58	0.70	0.82	0.93	0.98	0.95	0.99	1.0
LATE DETECTION ($\tau^\star = 0.85$)												
	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup\limits_{0<t<1} \frac{\|M_T^{(3)}(\tau)\|}{q(\tau,\nu=\frac{15}{218})}$	0.02	0.03	0.02	0.03	0.02	0.04	0.04	0.05	0.03	0.03	0.04	0.05
$\sup\limits_{0<t<1} \frac{\|M_T^{(4)}(\tau)\|}{q(\tau,\nu=\frac{15}{218})}$	0.05	0.03	0.03	0.07	0.06	0.08	0.10	0.10	0.18	0.16	0.19	0.30
CUSUM Test	0.05	0.04	0.04	0.05	0.04	0.05	0.04	0.04	0.04	0.04	0.06	0.06
FLUCTUATION	0.25	0.17	0.16	0.23	0.23	0.19	0.28	0.33	0.27	0.31	0.36	0.5
SIMULT	0.06	0.05	0.05	0.08	0.12	0.16	0.16	0.27	0.36	0.31	0.45	0.67
EARLY DETECTION ($\tau^\star = 0.15$)												
	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup\limits_{0<t<1} \frac{\|M_T^{(3)}(\tau)\|}{q(\tau,\nu=\frac{15}{218})}$	0.02	0.02	0.02	0.04	0.02	0.04	0.04	0.03	0.02	0.05	0.05	0.05
$\sup\limits_{0<t<1} \frac{\|M_T^{(4)}(\tau)\|}{q(\tau,\nu=\frac{15}{218})}$	0.03	0.03	0.04	0.04	0.05	0.04	0.04	0.06	0.11	0.08	0.19	0.26
CUMSUM Test	0.06	0.06	0.05	0.12	0.16	0.20	0.25	0.31	0.41	0.47	0.56	0.61
FLUCTUATION	0.22	0.21	0.17	0.23	0.22	0.20	0.22	0.26	0.23	0.35	0.36	0.37
SIMULT	0.07	0.06	0.08	0.09	0.16	0.21	0.25	0.42	0.60	0.54	0.72	0.80

and change in intercept.

Table III records the nominal coverage of the entertained tests when the errors of LRM are standard normal random variables. As the third moment of a standard normal distribution is zero, there should be little gain in efficiency from using the simultaneous test but this test still permits control of global error rates. All tests, except the fluctuation test, have nominal coverage probabilities that are less than 6% - the fluctuation test exceeds 20%.

Table III			
	T = 75	T = 100	T=125
$\sup_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.030	0.028	0.026
$\sup_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.044	0.050	0.046
CUSUM Test	0.050	0.056	0.040
FLUCT	**0.224**	**0.224**	**0.208**
SIMULT	0.052	0.046	0.048

Table IV details the empirical power under $H_{A,intercept}$ for a one-time change in intercept. The symmetry of the standard normal distribution decreases the empirical power of the simultaneous test but increases the empirical power of the CUSUM test.

The simultaneous test as well as statistic $\sup_{0<t<1} \frac{|M_T^{(4)}(\tau)|}{q(\tau,\nu=\frac{15}{128})}$ outperform the CUSUM test for changes that occur in the middle and later on in the sample. When the change in intercept occurs early on in the sample, the CUSUM test has an empirical power of 80% while the empirical power of the simultaneous and $\sup_{0<t<1} \frac{|M_T^{(4)}(\tau)|}{q(\tau,\nu=\frac{15}{128})}$ empirical power is 53% and 56% respectively. It

Table IV

MIDDLE OF SAMPLE $(\tau^\star = 0.5)$												
	$\beta_0^{(2)} = 1.25$			$\beta_0^{(2)} = 1.5$			$\beta_0^{(2)} = 1.75$			$\beta_0^{(2)} = 2$		
Statistic	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{218})}$	0.03	0.03	0.03	0.02	0.01	0.04	0.03	0.01	0.03	0.02	0.04	0.02
$\sup_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{218})}$	0.06	0.11	0.15	0.27	0.46	0.54	0.65	0.81	0.92	0.91	0.99	1.0
CUSUM Test	0.06	0.07	0.03	0.08	0.17	0.20	0.24	0.37	0.41	0.55	0.58	0.73
FLUCT	0.33	0.28	0.25	0.44	0.53	0.64	0.76	0.86	0.95	0.95	0.99	1.0
SIMULT	0.25	0.10	0.11	0.26	0.29	0.57	0.62	0.81	0.92	0.89	0.98	1.0
EARLY DETECTION $(\tau^\star = 0.15)$												
	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{218})}$	0.03	0.03	0.05	0.05	0.04	0.06	0.08	0.08	0.01	0.15	0.18	0.23
$\sup_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{(\tau,\nu=\frac{15}{218})}$	0.04	0.05	0.04	0.05	0.07	0.12	0.10	0.20	0.28	0.24	0.44	0.56
CUSUM Test	0.09	0.009	0.09	0.16	0.24	0.31	0.39	0.47	0.61	0.64	0.74	0.84
FLUCT	0.25	0.25	0.19	0.29	0.28	0.25	0.34	0.39	0.44	0.46	0.59	0.68
SIMULT	0.03	0.04	0.03	0.05	0.06	0.12	0.09	0.16	0.26	0.21	0.37	0.53
LATE DETECTION $(\tau^\star = 0.85)$												
	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{218})}$	0.04	0.03	0.05	0.02	0.03	0.04	0.04	0.04	0.06	0.05	0.10	0.13
$\sup_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{218})}$	0.03	0.04	0.07	0.06	0.08	0.13	0.17	0.23	0.33	0.34	0.55	0.69
CUSUM Test	0.05	0.06	0.04	0.04	0.06	0.05	0.08	0.07	0.07	0.07	0.08	0.11
FLUCT	0.25	0.25	0.19	0.29	0.28	0.25	0.34	0.39	0.44	0.46	0.59	0.69
SIMULT	0.03	0.04	0.03	0.06	0.06	0.11	0.15	0.22	0.31	0.29	0.41	0.62

is evident that the CUSUM test is significantly affected by the asymmetry of the distribution of the errors.

The last simulation undertaken was to determine the ability of the simultaneous test to detect a one-time change in slope of the LRM detailed in (38). Table V summaries the nominal coverage under $H_{O,joint}$ of the five test statistics considered here. Regarding the nominal coverage, again the tests designed via U-statistic type processes performed well. The simultaneous test for a one-time change in slope did have a slightly higher nominal coverage of 9% when the sample was 75. The nominal coverage fell to 6.6% when the sample size increased to 100 and then to 125. Again, the fluctuation test of PKK had a much larger nominal coverage at 20% (cf. bold numbers in Table V).

Table V

	T = 75	T = 100	T=125
$\sup_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.024	0.03	0.034
$\sup_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.056	0.058	0.052
CUSUM Test	0.038	0.03	0.034
FLUCT	**0.202**	**0.208**	**0.158**
SIMULT	0.094	0.076	0.066

Under the alternative hypothesis, the simultaneous test performed very well for a one-time change in slope when it occurs in the middle of the sample (cf. Table VI). This test is able to detect a 25% change in slope 13% of time when the sample is only 75. This rises to 40% of the time when the sample is 125 and the change is 75%. Moreover, the simultaneous test does better than each individual test as it exploits the asymmetry in the distribution which is χ^2 with 1 degree of freedom. For example, when the change in slope is 50% and the sample is 125, the individual tests detect the change 3% and 27%, respectively, of the time, while the simultaneous test detects it 34% of the time. Table VI reveals that the simultaneous test does not perform well when the change occurs early or later on in the sample - at best it detects the change 9% of the time. But the individual test statistic $\sup\limits_{0<t<1} \frac{|M_T^{(3)}(\tau)|}{q(\tau,\nu=\frac{15}{128})}$ performs very well and this is one reason to employ the joint test; that is, both the simultaneous test as well as the individual test for change in slope should be calculated. As these tests are asymptotically independent, the global error rate is easy to control. For example, if the global error rate is set at 10%, then the critical level used for the joint test would be approximately 5%. If this procedure is followed instead of using only the simultaneous test, this joint test performs much better than the CUSUM test of BDE.

Table VI

MIDDLE OF SAMPLE $(\tau^\star = 0.5)$												
	$\beta^{(2)} = 3.75$			$\beta^{(2)} = 4.5$			$\beta^{(2)} = 5.25$			$\beta^{(2)} = 6$		
Statistic	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup\limits_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.016	0.022	0.024	0.028	0.024	0.03	0.036	0.03	0.03	0.032	0.036	0.03
$\sup\limits_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.096	0.186	0.26	0.18	0.274	0.328	0.186	0.274	0.432	0.204	0.282	0.398
CUSUM Test	0.028	0.05	0.044	0.046	0.066	0.07	0.038	0.066	0.102	0.052	0.058	0.094
FLUCT	1	1	1	1	1	1	1	1	1	1	1	1
SIMULT	0.124	0.22	0.272	0.192	0.27	0.342	0.192	0.278	0.434	0.204	0.276	0.394
EARLY DETECTION $(\tau^\star = 0.15)$												
	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup\limits_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.67	0.792	0.822	0.964	0.99	0.994	0.986	0.996	1	0.99	1	1
$\sup\limits_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.09	0.12	0.15	0.234	0.25	0.296	0.242	0.276	0.308	0.266	0.33	0.374
CUSUM Test	0.144	0.168	0.224	0.252	0.28	0.314	0.238	0.282	0.354	0.264	0.306	0.38
FLUCT	0.998	1	1	1	1	1	1	1	1	1	1	1
SIMULT	0.11	0.212	0.248	0.056	0.086	0.106	0.048	0.076	0.086	0.088	0.098	0.108
LATE DETECTION $(\tau^\star = 0.85)$												
	N = 75	N = 100	N=125	N=75	N = 100	N=125	N=75	N = 100	N=125	N = 75	N = 100	N=125
$\sup\limits_{0<t<1} \frac{\lvert M_T^{(3)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.536	0.702	0.782	0.93	0.968	1	0.968	0.996	1	0.978	0.998	0.998
$\sup\limits_{0<t<1} \frac{\lvert M_T^{(4)}(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})}$	0.1	0.124	0.154	0.172	0.24	0.278	0.224	0.252	0.302	0.232	0.266	0.34
CUSUM Test	0.016	0.01	0.014	0.004	0.014	0.006	0.002	0.004	0.006	0.004	0.006	0.01
FLUCT	0.904	0.958	0.99	0.98	0.992	1	0.996	1	1	0.996	0.998	1
SIMULT	0.056	0.062	0.082	0.038	0.056	0.058	0.04	0.058	0.05	0.046	0.06	0.064

5 Conclusion

One simultaneous test statistic and two individual test statistics have been developed which can be used to test the hypothesis of a one-time change in intercept or slope in LRMs. The simultaneous is preferred because it controls the global error rate and exploits additional information which permits improvements in the power. This is particularly important when the distribution of the errors follow a distribution that is highly skewed. Moreover, the simultaneous test also performs very well when compared to CUSUM test of BDE. There is one drawback from the simultaneous test; the simulation revealed it is unable to detect a change in slope that occurs early and later on in the sample but this can be adjusted for by employing the joint test that calculates the simultaneous test as well as the statistic that detects a one-time change in slope. As these statistics are asymptotically independent, the global error rate can be controlled.

6 Appendix

$G(x) = \mathbb{P}\left\{\sup_{0<\tau<1} \frac{\lvert B(\tau)\rvert}{q(\tau,\nu=\frac{15}{128})} \le x\right\}$									
x	$G(x)$	x	$G(x)$	x	$G(x)$	x	$G(x)$	x	$G(x)$
0.534	0.01	0.792	0.21	0.934	0.41	1.087	0.61	1.290	0.81
0.568	0.02	0.799	0.22	0.941	0.42	1.096	0.62	1.303	0.82
0.597	0.03	0.806	0.23	0.950	0.43	1.104	0.63	1.317	0.83
0.615	0.04	0.813	0.24	0.958	0.44	1.112	0.64	1.333	0.84
0.632	0.05	0.820	0.25	0.966	0.45	1.120	0.65	1.349	0.85
0.648	0.06	0.827	0.26	0.974	0.46	1.128	0.66	1.368	0.86
0.661	0.07	0.834	0.27	0.981	0.47	1.136	0.67	1.385	0.87
0.673	0.08	0.840	0.28	0.988	0.48	1.144	0.68	1.400	0.88
0.684	0.09	0.847	0.29	0.994	0.49	1.153	0.69	1.422	0.89
0.695	0.10	0.854	0.30	1.001	0.50	1.163	0.70	1.447	0.90
0.707	0.11	0.861	0.31	1.008	0.51	1.175	0.71	1.471	0.91
0.720	0.12	0.868	0.32	1.014	0.52	1.186	0.72	1.496	0.92
0.729	0.13	0.875	0.33	1.021	0.53	1.196	0.73	1.531	0.93
0.738	0.14	0.882	0.34	1.029	0.54	1.206	0.74	1.563	0.94
0.746	0.15	0.890	0.35	1.037	0.55	1.216	0.75	1.609	0.95
0.753	0.16	0.897	0.36	1.045	0.56	1.228	0.76	1.668	0.96
0.761	0.17	0.904	0.37	1.053	0.57	1.240	0.77	1.723	0.97
0.769	0.18	0.912	0.38	1.061	0.58	1.252	0.78	1.797	0.98
0.778	0.19	0.919	0.39	1.069	0.59	1.264	0.79	1.916	0.99
0.785	0.20	0.926	0.40	1.078	0.60	1.277	0.80	2.040	1.00

References

[1] Altissimo, F. and Corradi, V. (2003). *Strong Rules for Detecting the Number of Breaks in a Time Series*. Journal of Econometrics, vol 117, pp. 207-244.

[2] Antoch, J. and Hušková, M. (1995). *Change-Point Problem and Bootstrap*. Journal of Non-parametric Statistics, Vol. 5 pp. 123-144.

[3] Brown, R., Durbin, J. and Evans, J. (1975). *Techniques for Testing the Constancy of Regression Coefficients over time*, Journal of The Royal Statistical Society Series B (Methodology), vol. 37, No. 2, pp. 149-192.

[4] Buck, E. (1965). Advanced Calculus, Publisher: McGraw-Hill, New York, Second edition.

[5] Chow, G.C. (1960). *Tests of Equality between Sets of Coefficients in two Linear Regression Models*, Econometrica, 28, pp. 591-605.

[6] Csörgő, M. and Horváth, L. (1997). Limit Theorems in Change-Point Analysis. Publisher: John Wiley, Chichester.

[7] Csörgő, M. and Horváth, L. (1988a). *Nonparametric Methods for Changepoint Problems.* Handbook of Statistics, Vol. **7**, 403-425. Elsevier Science Publishers B.V. (North-Holland).

[8] Donsker, M. (1951). *An Invariance Principle for Certain Probability Limit Theorems. Four papers on probability.* Mem. Amer. Math. Statist. **23**, 277-283.

[9] Dufour, J.M. (1988). *Recursive Stability Analysis of Linear Regression Relationships*, Journal of Econometrics, 19, pp. 31-76.

[10] Ferger, D. (2001). *Analysis of Change-point Estimators Under the Null Hypothesis.* Bernoulli Vol. 7, pp. 487-506.

[11] Gombay E. and Horváth L. and Hušková, M. (1996). *Estimators and Tests for Change in Variances.* Statistics & Decisions, vol. 14, 145 - 159.

[12] Koroljuk, V. and Borovshich, V. (1994). Theory of U-Statistics, Publisher: Kluwer Academic Publishers, Dordrecht, the Netherlends.

[13] Kramer, W., Ploberger, W., and Alt, R. (1988). *A Test for Structural Change Dynamic Models*, Econometrica, vol. 56, pp 1355-1369.

[14] Mason, D. M. and Schuenemeyer, J. H. (1983). *A Modified Kolmogorov-Smirnov Test Sensitive to Tail Alternatives.* The Annals of Statistics, Vol. 11, No. 3.

[15] Olmo, J and Pouliot, W. (2008) *Early Detection for Market Risk Failure*, Department of Economics, City University, London working paper 08/09.

[16] Orasch, M. and Pouliot, W. (2004). *Tabulating Weighted Sup-norm Functionals Used in Change-point Analysis.* Journal of Statistical Computation and Simulation, Vol. 74, April, pp. 249 - 276.

[17] Ploberger, W., Kramer, W. and Alt, R. (1988). *Testing for Structural Change in Dynamic Models.* Econometrica, vol. 56, pp. 1355 - 1369.

[18] Ploberger, W., Kramer, W. and Kontrus, K. (1989). *A new Test for Structural Stability in the Linear Regression Model*, Journal of Econometrics, Vol. 40, pp. 307 - 318.

[19] Szyszkowicz, B. (1991). *Weighted Stochastic Processes Under Contiguous Alternatives.* C.R. Math. Rep. Acad. Sci. Canada, Vol. XIII, No. 5, 211-216.

[20] Szyszkowicz, B. (1992). Weak Convergence of Stochastic Processes in Weighted Metrics and their Applications to Contiguous Changepoint Analysis. Ph.D. thesis Carleton University, Ottawa, Canada.

[21] Szyszkowicz, B. (1997). *Weighted Approximations of Partial Sum Processes in* $D[0,\infty)$*. II.* Studia Sci. Math. Hungar., **33**, 305-320.