Granger Centre Discussion Paper Series

Robust methods for detecting multiple level breaks in autocorrelated time series

by

David I. Harvey, Stephen J. Leybourne and A. M. Robert Taylor

Granger Centre Discussion Paper No. 10/01

Robust Methods for Detecting Multiple Level Breaks in Autocorrelated Time Series*

David I. Harvey, Stephen J. Leybourne and A. M. Robert Taylor
Granger Centre for Time Series Econometrics
School of Economics, University of Nottingham

February 2010

Abstract

In this paper we propose tests for the null hypothesis that a time series process displays a constant level against the alternative that it displays (possibly) multiple changes in level. Our proposed tests are based on functions of appropriately standardized sequences of the differences between sub-sample mean estimates from the series under investigation. The tests we propose differ notably from extant tests for level breaks in the literature in that they are designed to be robust as to whether the process admits an autoregressive unit root (the data are $I(1)$) or stable autoregressive roots (the data are $I(0)$). We derive the asymptotic null distributions of our proposed tests, along with representations for their asymptotic local power functions against Pitman drift alternatives under both $I(0)$ and $I(1)$ environments. Associated estimators of the level break fractions are also discussed. We initially outline our procedure through the case of non-trending series, but our analysis is subsequently extended to allow for series which display an underlying linear trend, in addition to possible level breaks. Monte Carlo simulation results are presented which suggest that the proposed tests perform well in small samples, showing good size control under the null, regardless of the order of integration of the data, and displaying very decent power when level breaks occur.

Keywords: Level breaks; unit root; moving means; long run variance estimation; robust tests; breakpoint estimation.

JEL Classification: C22.

*The authors are grateful to the Co-Editor, Peter Robinson, and two anonymous referees for helpful comments on a previous version of this paper. Address correspondence to: Robert Taylor, School of Economics, University of Nottingham, Nottingham, NG7 2RD, U.K. E-mail addresses: dave.harvey@nottingham.ac.uk (D.I. Harvey), steve.leybourne@nottingham.ac.uk (S.J. Leybourne), robert.taylor@nottingham.ac.uk (A.M.R. Taylor)

1 Introduction

Testing for structural change is a long-standing area of important research in econometrics and statistics. A vast number of tests for the presence of structural change in the parameters of stationary ($I(0)$) regression models have been proposed in the literature; see, *inter alia*, the literature reviews in Stock (1994), Kuan and Hornik (1995) and Perron (2006).

Only recently in the econometrics literature have methods surfaced for structural change testing that are valid regardless of whether the regression errors are stationary or follow a unit root ($I(1)$) process. This is an important practical development because macroeconomic and financial series typically appear to be characterized by temporary ($I(0)$) or permanent ($I(1)$) shocks fluctuating around a broken trend function: see, *inter alia*, Stock and Watson (1996,1999) and Perron and Zhu (2005). Indeed, Perron (2006, p.279) emphasizes the intricate interaction between unit roots, stationarity and structural changes, and the need to distinguish between them in practical economic applications.

Developing structural change tests that are valid under both $I(0)$ and $I(1)$ errors is, like developing unit root tests which are robust to changes in the deterministic trend function, not an easy task, however, and only a small number of contributions exist with this property. Harvey *et al.* (2009a) develop tests for a one-time break in the slope of the deterministic trend function (of the form embodied in Models B and C of Perron, 1989, p.1364), while Perron and Yabu (2009b) [PY] and Saygınsoy and Vogelsang (2010) (building on the approaches of Perron and Yabu (2009a), and Vogelsang (1998) and Bunzel and Vogelsang (2005), respectively) also allow for the possibility of a one-time shift in just the intercept (Model A of Perron, 1989, p.1364). The importance of the need to control for the possibility of either $I(0)$ or $I(1)$ errors has also been recognized in the recent literature on testing for and detecting outliers, with Perron and Rodríguez (2003) and Burridge and Taylor (2006) developing procedures which allow for (multiple) outliers and are robust to the order of integration of the data.

All of the above tests for level and/or trend breaks, which have the property of being robust as to whether the errors are $I(0)$ or $I(1)$, have in common the feature that they only allow for the possibility of a one-time change in the parameters of the trend function. In a recent paper, Kejriwal and Perron (2009) [KP] extend the work of PY, developing a sequential approach which allows the practitioner to test the trend function for multiple breaks in slope or multiple simultaneous breaks in slope and level, but not multiple breaks solely in the level. In this paper our aim is to fill this important gap in the literature by providing tests which allow for multiple level breaks, which are robust as to whether the data are $I(0)$ or $I(1)$ and, subject to standard moment conditions holding, to the distribution of the innovations. This particular testing problem is especially relevant for the analysis of financial time series data where there is a continuing debate in the literature over whether such data is subject to multiple level shifts; see for example the discussion in Perron (2006, section 8.7) and the references therein. Testing for and dating such level breaks is clearly a very important exercise in its own right, but also as a first step in developing other testing procedures in the data, such as unit root tests, which are robust to the possible presence of multiple level breaks.

The tests which we propose in this paper are based around statistics formed from sequences of the differences between sub-sample mean estimates (modifications to allow for a linear trend in the data are also discussed) from the data and test the null hypothesis that no level breaks are present against the alternative of at least one break in level. These

statistics belong to the *generalized fluctuations* class of statistics for structural change; see Kuan and Hornik (1995), *inter alia*. We demonstrate that these statistics cannot be employed directly to test for level breaks in the presence of uncertainty over whether the data are $I(1)$ or $I(0)$ because in each case the standardization (both in the sample size and the long run variance) required to obtain pivotal non-degenerate limiting null distributions for the statistics differs according to whether the data are $I(0)$ or $I(1)$. This leads us to consider tests based on two statistics, one appropriate for the case of $I(0)$ data and the other appropriate for $I(1)$ data. We derive the large sample properties of these statistics in both $I(0)$ and (local to) $I(1)$ environments. This analysis reveals that the test statistic appropriate for the case of $I(1)$ errors converges in probability to zero under $I(0)$ errors, while the test appropriate for $I(0)$ errors is always under-sized when the errors are (local to) $I(1)$. These properties facilitate the construction of a size-controlled *union of rejections* testing approach whereby we reject the null hypothesis of no level breaks if either of the two tests rejects. We investigate both the asymptotic local power and the finite sample power properties of the three tests. Here it is shown that the union test essentially capitalizes on the superior power of the test designed for $I(0)$ errors when the errors are $I(0)$ and of the test designed for $I(1)$ errors when the errors are $I(1)$, losing very little in power relative to the better of the two tests in each environment.

A rejection by any of our proposed test procedures indicates that at least one level break is present and so we also outline an associated sequential procedure for detecting and dating possibly multiple breaks in level. Monte Carlo evidence suggests that this procedure does a very good job in practice. The statistics used in constructing the proposed tests and level break detection procedure require a choice of window width for constructing the sequences of sub-sample mean estimates, and this choice in turn impacts on the maximum number of breaks assumed to be present in the data. We present a detailed examination of the impact of this choice on the performance of our approach.

The paper is organized as follows. In section 2 we introduce our reference data generating process [DGP], a first-order autoregression, which embeds both the (near) unit root and stationary cases, driven by linear process innovations, and which allows for (multiple) level breaks occurring at unknown points through the sample. In sections 3 and 4 we develop our approach to testing for level breaks, the latter section including an analysis of the asymptotic local power properties of our proposed tests. Finite sample critical values, obtained under the assumption of normal errors, are reported in section 5, and this section also examines the robustness of the tests' finite sample size to non-normal errors, and also autocorrelated errors, as well as investigating the finite sample power properties of the tests. Our sequential procedure for determining the number and timing of the level breaks is detailed in section 6. Extensions of our approach to allow for the presence of a linear trend in the data are discussed in section 7. Section 8 offers some conclusions. Proofs are placed in an Appendix. In the following '$\lfloor \cdot \rfloor$' denotes the integer part of its argument; '$\xrightarrow{w}$' denotes weak convergence and '$\xrightarrow{p}$' convergence in probability, in each case as the sample size diverges to positive infinity; '$\mathbb{I}(\cdot)$' denotes the indicator function, and finally '$x := y$' ('$x =: y$') indicates that x is defined by y (y is defined by x).

2 The Multiple Level Breaks Model

Consider a time series process $\{y_t\}$ of T observations generated according to the DGP

$$y_t = \alpha + \sum_{i=1}^{n} \gamma_{i,T}^* DU_t(\lfloor \tau_i T \rfloor) + u_t, \quad t = 1, ..., T, \tag{1}$$

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad t = 2, ..., T, \tag{2}$$

where $DU_t(\lfloor \tau_i T \rfloor) := \mathbb{I}(t > \lfloor \tau_i T \rfloor)$ with $\lfloor \tau_i T \rfloor$ a potential level break point with associated break fraction τ_i and break magnitude $\gamma_{i,T}^*$. We assume that the break fractions, τ_i, $i = 1, ..., n$, are unknown but satisfy $\tau_i \in \Lambda$, where $\Lambda = [\tau_L, \tau_U]$ with $0 < \tau_L < \tau_U < 1$; the fractions τ_L and τ_U being trimming parameters below and above which no break is deemed allowable. We further assume that $\tau_1 < \tau_2 < ... < \tau_{n-1} < \tau_n$, without loss of generality.

The initial condition, u_1, is assumed to be such that $T^{-1/2} u_1 \xrightarrow{p} 0$, while the error process $\{\varepsilon_t\}$ is taken to satisfy the following conventional linear process assumption.

Assumption LP. *The stochastic process* $\{\varepsilon_t\}$ *is such that* $\varepsilon_t = C(L)\eta_t$, $C(L) := \sum_{j=0}^{\infty} C_j L^j$ *with* $C(1)^2 > 0$ *and* $\sum_{i=0}^{\infty} i|C_i| < \infty$, *and where* $\{\eta_t\}$ *is an IID sequence with mean zero, variance* σ_η^2 *and finite fourth moment. The long run variance of* ε_t *is defined as* $\omega_\varepsilon^2 := \lim_{T\to\infty} T^{-1} E(\sum_{t=1}^{T} \varepsilon_t)^2 = \sigma_\eta^2 C(1)^2$.

Remark 1. Under Assumption LP, $\{\varepsilon_t\}$ satisfies a functional central limit theorem [FCLT],

$$T^{-1/2} \sum_{t=1}^{\lfloor T\cdot \rfloor} \varepsilon_t \xrightarrow{w} \omega_\varepsilon W(\cdot) \tag{3}$$

where $W(\cdot)$ is a standard Brownian motion process on $[0, 1]$.

We consider two cases for the order of integration of the autoregressive process, u_t. The $I(1)$ case for u_t is represented by setting $\rho = \rho_T := 1 - c/T$ for $0 \leq c < \infty$ in (1), which permits (local to) unit root behaviour when $(c > 0)$ $c = 0$. Here we will also assume that $\gamma_{i,T}^* := \omega_\varepsilon T^{1/2} \gamma_i$, $i = 1, ..., n$. The $T^{1/2}$ scaling in $\gamma_{i,T}^*$ provides the appropriate Pitman drift, while scaling by ω_ε is a convenience device allowing it to be factored out of the limit distributions that arise later. The $I(0)$ case for u_t is represented by setting $|\rho| < 1$ in (1). In this situation the long run variance of u_t is given by $\omega_u^2 := \lim_{T\to\infty} T^{-1} E(\sum_{t=1}^{T} u_t)^2 = \sigma_\eta^2 C(1)^2/(1-\rho)^2$. Here we assume $\gamma_{i,T}^* := \omega_u T^{-1/2} \gamma_i$, $i = 1, ..., n$, with $T^{-1/2}$ now providing the appropriate Pitman drift, and scaling by ω_u again being used for convenience. For future brevity, the two cases are embodied in the following assumptions:

Assumption I(1). *Let Assumption LP hold. Also, let* $\rho = \rho_T := 1 - c/T$, $0 \leq c < \infty$, *and let* $\gamma_{i,T}^* = \omega_\varepsilon T^{1/2} \gamma_i$.

Assumption I(0). *Let Assumption LP hold. Also, let* $|\rho| < 1$ *and let* $\gamma_{i,T}^* = \omega_u T^{-1/2} \gamma_i$.

3 Detecting Multiple Level Breaks

The focus of this paper is on testing the null hypothesis of no level breaks, that is $H_0 : \gamma_{i,T}^* = 0$ for $i = 1, 2, ..., n$, against the alternative of at least one level break; that is $H_1 : \gamma_{i,T}^* \neq 0$

for at least one $i \in \{1, ..., n\}$. In implementing a test of such a hypothesis, consider a sequence of statistics of the form $M_{t,\lfloor mT \rfloor}$, for $t \in \Lambda_T := [\lfloor \tau_L T \rfloor, \lfloor \tau_U T \rfloor]$, where

$$M_{t,\lfloor mT \rfloor} := \frac{\sum_{i=1}^{\lfloor \frac{m}{2} T \rfloor} y_{t+i} - \sum_{i=1}^{\lfloor \frac{m}{2} T \rfloor} y_{t-i+1}}{\lfloor \frac{m}{2} T \rfloor} \tag{4}$$

which is the difference between the mean of the $\lfloor \frac{m}{2} T \rfloor$ observations $y_{t+1}, y_{t+2}, ..., y_{t+\lfloor \frac{m}{2} T \rfloor}$ and the mean of the $\lfloor \frac{m}{2} T \rfloor$ observations $y_t, y_{t-1}, ..., y_{t-\lfloor \frac{m}{2} T \rfloor + 1}$.

The statistic $M_{t,\lfloor mT \rfloor}$ of (4) belongs to the *generalized fluctuation* class of test statistics for structural change introduced in Kuan and Hornik (1995) and Leisch, Hornik and Kuan (2000), *inter alia*. We need to ensure that at most only one level break can occur in the data spanned by any $M_{t,\lfloor mT \rfloor}$; we therefore impose the restriction that $\tau_i - \tau_{i-1} \geq m$ for all $i = 2, ..., n$. Under this constraint, the DGP admits n level breaks occurring at unknown points across the interval Λ_T, with a minimum of $\lfloor mT \rfloor$ observations between breaks. Notice that n and m must satisfy the relation

$$n \leq 1 + \left\lfloor \frac{\tau_U - \tau_L}{m} \right\rfloor =: n_{\max}$$

which provides an upper bound for the maximum number of breaks assumed to be present for given choices of the *window width*, m, and the trimming parameters, τ_L and τ_U.

By way of motivation for our approach, consider a stylized example where the errors ε_t in (2) are Gaussian white noise and only a single mid-sample break is present (i.e. $\tau_1 = 0.5$). Suppose first that Assumption I(1) holds with $\rho = 1$, so that the process follows a random walk with Gaussian innovations. Then the optimal test of $\gamma^*_{1,T} = 0$ is based on the maximum likelihood (ML) (equivalently GLS) estimator $y_{\lfloor 0.5T \rfloor + 1} - y_{\lfloor 0.5T \rfloor} = \Delta y_{\lfloor 0.5T \rfloor + 1}$, i.e. $M_{\lfloor 0.5T \rfloor, \lfloor mT \rfloor}$ with $m = 2/T$. If, on the other hand, Assumption I(0) holds with $\rho = 0$, so that the process is simply Gaussian white noise, then ML considerations would lead to a test statistic based on $\lfloor 0.5T \rfloor^{-1} \sum_{i=\lfloor 0.5T \rfloor + 1}^{T} y_i - \lfloor 0.5T \rfloor^{-1} \sum_{i=1}^{\lfloor 0.5T \rfloor} y_i$, which is identical to $M_{\lfloor 0.5T \rfloor, \lfloor mT \rfloor}$ of (4) with $m = 1$.

As regards $M_{\lfloor 0.5T \rfloor, 2}$, only two observations are used, and it is therefore not possible to establish an invariance principle for $M_{\lfloor 0.5T \rfloor, 2}$ under Assumption I(1). Clearly this is also the case under Assumption I(0). This mitigates the use of $M_{\lfloor 0.5T \rfloor, 2}$, notwithstanding the fact that under Assumption I(0) it is far from being the optimal approach in any case. On the other hand, $M_{\lfloor 0.5T \rfloor, T}$ uses all observations and therefore no such limitations regarding an invariance principle apply. However, it is also obvious that under Assumption I(1), use of $M_{\lfloor 0.5T \rfloor, T}$ is completely at odds with what the optimal approach suggests.

These considerations prompt the use of statistics based on the *form* of $M_{t,\lfloor mT \rfloor}$ of (4), but also highlight the difficulties involved in devising a powerful and robust testing approach when the order of integration is unknown. A possible approach in this regard, which lies in between the $I(1)$- and $I(0)$-based extremes of using $m = 2/T$ (which, of course, implies that $m \to 0$ as $T \to \infty$) and $m = 1$, respectively, is to consider a statistic based on $M_{t,\lfloor mT \rfloor}$ but where we restrict m, the window width, to lie in the range $0 < m < 1$.

Whilst using a fraction of the data to detect breaks in level might at first sight appear inefficient (in the $I(0)$ case), a very important consideration arises in the current context of detecting multiple level breaks. Specifically, if we are attempting to identify the break point τ_i (with $\gamma^*_{i,T} \neq 0$), then if two neighbouring breaks exist (i.e. $\gamma^*_{i-1,T} \neq 0$ and $\gamma^*_{i+1,T} \neq 0$), the only data relevant for the purpose of identifying the break at τ_i is the subset of observations

contained in the interval $\lfloor \tau_{i-1} T \rfloor + 1 \leq t \leq \lfloor \tau_{i+1} T \rfloor$; the remaining observations are totally uninformative with regard to detecting that level break.

Motivated by the above discussion, and given that we do not assume knowledge of the order of integration of the data, in this paper we examine statistics based on $M_{t,\lfloor mT \rfloor}$ of (4) with $0 < m < 1$ for detecting possibly multiple breaks in level. Since our focus is on detecting breaks at unknown points in time, consider the following prototypical statistic based on the maximum function of $|M_{t,\lfloor mT \rfloor}|$ (notice, therefore, that we assume the signs of the possible breaks are unknown) over all $t \in \Lambda_T$, i.e.

$$\mathcal{M} := \max_{t \in \Lambda_T} |M_{t,\lfloor mT \rfloor}|$$

cf. Andrews (1993). For a given value of m, this statistic therefore takes the largest (in absolute value) fluctuation measure $|M_{t,\lfloor mT \rfloor}|$ over all possible break points in Λ_T. Note that we require $\tau_L \geq m/2$ and $\tau_U \leq 1 - (m/2)$, to ensure $M_{t,\lfloor mT \rfloor}$ is only calculated from observed data. In the following theorem we establish the limiting distribution of $\mathcal{M}$ under both (near-) $I(1)$ and $I(0)$ environments.

Theorem 1 *Let y_t be generated according to (1) and (2) and let Assumption LP hold. Then,*

(a) Under Assumption I(1),

$$\omega_\varepsilon^{-1} T^{-1/2} \mathcal{M} \xrightarrow{w} \sup_{r \in \Lambda} |L_1(r, m, c) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})| ,$$

with

$$L_1(r, m, c) := 2m^{-1} \left\{ \int_r^{r+m/2} W_c(s) ds - \int_{r-m/2}^{r} W_c(s) ds \right\}$$

where $W_c(r) := \int_0^r e^{-(r-s)c} dW(s)$ denotes a standard Ornstein-Uhlenbeck (OU) process, $W(s)$ the standard Brownian motion arising from the FCLT in (3), and

$$K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma}) := \begin{cases} 0 & \tau_L \leq r \leq \tau_1 - m/2 \\ \gamma_1 (1 - \frac{|r - \tau_1|}{m/2}) & \tau_1 - m/2 \leq r \leq \tau_1 + m/2 \\ 0 & \tau_1 + m/2 \leq r \leq \tau_2 - m/2 \\ \gamma_2 (1 - \frac{|r - \tau_2|}{m/2}) & \tau_2 - m/2 \leq r \leq \tau_2 + m/2 \\ \vdots & \vdots \\ 0 & \tau_{n-1} + m/2 \leq r \leq \tau_n - m/2 \\ \gamma_n (1 - \frac{|r - \tau_n|}{m/2}) & \tau_n - m/2 \leq r \leq \tau_n + m/2 \\ 0 & \tau_n + m/2 \leq r \leq \tau_U \end{cases}$$

where $\boldsymbol{\tau} := [\tau_1, \tau_2, ..., \tau_n]$ and $\boldsymbol{\gamma} := [\gamma_1, \gamma_2, ..., \gamma_n]$.

(b) Under Assumption I(0),

$$\omega_u^{-1} T^{1/2} \mathcal{M} \xrightarrow{w} \sup_{r \in \Lambda} |L_0(r, m) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})|$$

with $L_0(r, m) := 2m^{-1} \{W(r + m/2) - 2W(r) + W(r - m/2)\}$, $W(s)$ again arising from the FCLT in (3), and where $K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})$ is as defined in part (a).

Remark 2. Under H_0, where $\gamma_i = 0$, $i = 1, 2, ..., n$, observe that $K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma}) = 0$, $\tau_L \leq r \leq \tau_U$, from which we obtain representations for the limiting null distribution of $\mathcal{M}$, when appropriately standardized for either $I(1)$ or $I(0)$ errors.

An immediate consequence of the large sample results in Theorem 1 is that in attempting to use $\mathcal{M}$ to test H_0, we would encounter two problems in practice. Firstly, under H_0, the appropriate scaling factor in the sample size, T, and choice of long run variance standardization (either ω_ε^2 or ω_u^2) to apply to $\mathcal{M}$ to obtain a non-degenerate and pivotal limiting null distribution both depend on whether the errors are $I(1)$ or $I(0)$, which will be unknown to the practitioner. Secondly, in practice we would also need to estimate either ω_ε^2 or ω_u^2 in order to yield a feasible testing procedure. In the next section we will explore solutions to both of these issues.

4 Feasible Robust Tests for Level Breaks

In this section we address the practical issues that exist in developing a feasible test for level breaks outlined at the end of the previous section. In section 4.1, we first consider the issue of long run variance estimation, and examine the behaviour of the estimators under both $I(1)$ and $I(0)$ errors. In section 4.2, we then use the results from section 4.1 to develop an operational test against level breaks in model (1)-(2) for the situation where the order of integration is unknown. Section 4.2 also presents an analysis of the asymptotic size properties of the proposed tests for a range of values of the window width, m, while section 4.3 gives asymptotic power results.

4.1 Long Run Variance Estimation

We now consider estimation of the long run variances ω_ε^2 (relevant under $I(1)$ errors) and ω_u^2 (relevant under $I(0)$ errors). Here, we focus on Berk (1974)-type estimators, initially assuming knowledge of the relevant order of integration.

4.1.1 Estimation of ω_ε^2

First we consider estimating ω_ε^2 when the errors are known to be $I(1)$. It is obviously desirable from a power standpoint that the long run variance estimator is influenced by the presence of the level breaks to the least degree possible, hence we need to remove the effects of the level breaks, bearing in mind that the number and timings of these breaks are unknown. Our first consideration is therefore estimation of the timings of the potential breaks. In the context of our reference level break model (1)-(2), we further assume that when there are n level breaks, that $|\gamma_1| > |\gamma_2| > ... > |\gamma_n|$. This ordering is adopted to expedite the arguments made below, and does not compromise the generality of the results.

Under Assumption I(1), if the errors ε_t in (2) are Gaussian white noise and only one break is present (at time $\lfloor \tau_1 T \rfloor$), the optimal test of $\gamma_{1,T}^* = 0$ is based on the ML estimator $\Delta y_{\lfloor \tau_1 T \rfloor + 1}$. It makes sense, therefore, under $I(1)$ errors, to consider $|\Delta y_t|$ to identify any break points. Consequently, let $\hat{t}_1 := (\arg\max_{t \in \Omega_T} |\Delta y_t|) - 1$ where $\Omega_T := [\lfloor \tau_L T \rfloor + 1, \lfloor \tau_U T \rfloor + 1]$ (bearing in mind that the outliers are observed one observation after a corresponding break point). Next, since we are assuming that the breaks are separated by at least $\lfloor mT \rfloor$ observations, we now wish to exclude the dates $[\hat{t}_1 - \lfloor mT \rfloor + 1, \hat{t}_1 + \lfloor mT \rfloor - 1]$, so now let $\hat{t}_2 := (\arg\max_{t \in \Omega_T - \Omega_{1,T}} |\Delta y_t|) - 1$ where $\Omega_{1,T} := [\hat{t}_1 - \lfloor mT \rfloor + 2, \hat{t}_1 + \lfloor mT \rfloor]$, then

$\hat{t}_3 := (\arg\max_{t\in\Omega_T-\Omega_{1,T}-\Omega_{2,T}} |\Delta y_t|) - 1$ where $\Omega_{2,T} := [\hat{t}_2 - \lfloor mT \rfloor + 2, \hat{t}_2 + \lfloor mT \rfloor]$, and so on, until $\Omega_{\bar{n}+1,T} = \varnothing$. This procedure identifies $\bar{n}$ break points, where it can be shown that

$$\left\lfloor \frac{\lfloor \tau_U T \rfloor - \lfloor \tau_L T \rfloor + \lfloor mT \rfloor}{2\lfloor mT \rfloor - 1} \right\rfloor \leq \bar{n} \leq n_{\max}. \tag{5}$$

Using the estimated break points, $\hat{t}_1, ..., \hat{t}_{\bar{n}}$, we then remove the effect of the level breaks on the Δy_t series by taking the residuals $\hat{\varepsilon}_t$ from the OLS regression

$$\Delta y_t = \sum_{i=1}^{\bar{n}} \hat{\gamma}_i^* D_t(\hat{t}_i) + \hat{\varepsilon}_t, \quad t = 2, ..., T \tag{6}$$

where the $D_t(\hat{t}_i) := \mathbb{I}(t = \hat{t}_i + 1)$, $i = 1, ..., \bar{n}$, are one-time dummy variables. The Berk (1974)-type autoregressive spectral density estimator of ω_ε^2 is then obtained as

$$\hat{\omega}_\varepsilon^2 := \frac{\hat{\sigma}^2}{\hat{\pi}^2}$$

which is based on estimating the OLS regression

$$\Delta\hat{\varepsilon}_t = \hat{\pi}\hat{\varepsilon}_{t-1} + \sum_{j=1}^{k-1} \hat{\psi}_j \Delta\hat{\varepsilon}_{t-j} + \hat{e}_t, \quad t = k+2, ..., T \tag{7}$$

with $\hat{\sigma}^2 := (T - 2k - 1)^{-1} \sum_{t=k+2}^{T} \hat{e}_t^2$. As is standard, we require that the lag truncation parameter, k, in (7) satisfies the condition that, as $T \to \infty$, $1/k + k^3/T \to 0$.

In Theorem 2 below, we now establish the large sample behaviour of $\hat{\omega}_\varepsilon^2$; since, in practice, the order of integration is unknown we detail the asymptotic properties of the estimator under both $I(1)$ and $I(0)$ errors.

Theorem 2 *Let the conditions of Theorem 1 hold. Then,*

(a) Under Assumption I(1),
(i) When $n > 0$: $\hat{t}_1 - \lfloor \tau_1 T \rfloor \xrightarrow{p} 0$, $\hat{t}_2 - \lfloor \tau_2 T \rfloor \xrightarrow{p} 0$, ..., $\hat{t}_n - \lfloor \tau_n T \rfloor \xrightarrow{p} 0$, and $\bar{n} \geq n$;
(ii) For $n \geq 0$: $\hat{\omega}_\varepsilon^2 \xrightarrow{p} \omega_\varepsilon^2$.

(b) Under Assumption I(0), $\hat{\omega}_\varepsilon^2 = O_p(k^{-2})$.

Remark 3. The break point consistency result in Theorem 2 (a)(i) relies on the break magnitudes under $I(1)$ errors being $O(T^{1/2})$. If the breaks are of fixed magnitude, consistency no longer pertains. However, in this case it is easily shown that the results in Theorem 2 (a)(ii) and (b) continue to hold.

4.1.2 Estimation of ω_u^2

Now consider estimating ω_u^2 in the case where the errors are known to be $I(0)$. Given the estimated break points, $\hat{t}_1, ..., \hat{t}_{\bar{n}}$, from section 4.1.1 above, we again account for the level breaks by taking the residuals $\hat{u}_t$ from the OLS regression

$$y_t = \hat{\alpha} + \sum_{i=1}^{\bar{n}} \hat{\gamma}_i^* DU_t(\hat{t}_i) + \hat{u}_t, \quad t = 1, ..., T \tag{8}$$

where $DU_t(\hat{t}_i) := \mathbb{I}(t > \hat{t}_i)$, $i = 1, ..., \bar{n}$. The estimator of ω_u^2 in this case is given by

$$\hat{\omega}_u^2 := \frac{\hat{\sigma}^2}{\hat{\pi}^2}$$

where $\hat{\pi}$ and $\hat{\sigma}$ are now obtained from the OLS regression

$$\Delta\hat{u}_t = \hat{\pi}\hat{u}_{t-1} + \sum_{j=1}^{k-1} \hat{\psi}_j \Delta\hat{u}_{t-j} + \sum_{j=0}^{k-1}\sum_{i=1}^{\bar{n}} \hat{\psi}_{j,i} D_{t-j}(\hat{t}_i) + \hat{e}_t, \quad t = k+1, ..., T, \tag{9}$$

with $\hat{\sigma}^2 := (T - (2 + \bar{n})k)^{-1} \sum_{t=k+1}^{T} \hat{e}_t^2$, and where k again satisfies the condition that, as $T \to \infty$, $1/k + k^2/T \to 0$. Notice that, for the reasons outlined in Perron and Vogelsang (1992), the regression in (9) augments the usual ADF-type regression with the $\bar{n}$ one-time dummy variables, $D_t(\hat{t}_i)$, $i = 1, ..., \bar{n}$, and the $(k-1)$ lagged values of each of these.

Theorem 3 provides the large sample behaviour of $\hat{\omega}_u^2$ under both $I(1)$ and $I(0)$ errors.

Theorem 3 *Let the conditions of Theorem 1 hold. Then,*

(a) Under Assumption I(1),

$$T^{-2}\hat{\omega}_u^2 \xrightarrow{w} \omega_\varepsilon^2 Q(c, d, \tilde{\boldsymbol{\tau}})$$

where

$$Q(c, d, \tilde{\boldsymbol{\tau}}) := \frac{\{\int_0^1 H(r, c, d, \tilde{\boldsymbol{\tau}})^2 dr\}^2}{\{\int_0^1 H(r, c, d, \tilde{\boldsymbol{\tau}}) dW_c(r)\}^2}$$

and $H(r, c, d, \tilde{\boldsymbol{\tau}})$ is a continuous time residual from the projection of $W_c(r)$ onto the space spanned by $\{1, 1(r > \tilde{\tau}_1), 1(r > \tilde{\tau}_2), ..., 1(r > \tilde{\tau}_{\bar{n}})\}$ with $\tilde{\tau}_i := \lim_{T\to\infty} T^{-1}\hat{t}_i$, $i = 1, 2, ..., \bar{n}$ and $\tilde{\boldsymbol{\tau}} := [\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_{\bar{n}}]$. Here, as $T \to \infty$, $\bar{n} \geq n$ and:

(i) If $n = \bar{n}$ then $\tilde{\boldsymbol{\tau}} = [\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_n] = [\tau_1, \tau_2, ..., \tau_n] = \boldsymbol{\tau}$ (i.e. $\tilde{\boldsymbol{\tau}}$ is a non-stochastic argument of H);

(ii) If $n = 0$, which is the null case here, $\tilde{\boldsymbol{\tau}} = [\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_{\bar{n}}]$ is a vector of $\bar{n}$ dependent random variables (whose length $\bar{n}$ is stochastic but satisfies (5)), but whose distribution is the same for all ε_t and is independent of $W_c(r)$;

(iii) If $\bar{n} > n > 0$, $\tilde{\boldsymbol{\tau}} = [\tau_1, \tau_2, ..., \tau_n, \tilde{\tau}_{n+1}, .., \tilde{\tau}_{\bar{n}}]$ (whose length $\bar{n}$ is stochastic but satisfies (5)) where $\tilde{\tau}_{n+1}, .., \tilde{\tau}_{\bar{n}}$ are $\bar{n} - n$ dependent random variables whose distribution is the same for all ε_t and is also independent of $W_c(r)$.

(b) Under Assumption I(0), $\hat{\omega}_u^2 \xrightarrow{p} \omega_u^2$.

Remark 4. Observe from part (a) of Theorem 3 that the limiting distribution of $T^{-2}\hat{\omega}_u^2$ under $I(1)$ errors does not depend on the underlying break magnitudes, γ_i, $i = 1, ..., n$. This invariance arises from the presence of the one-time dummy variables, $D_t(\hat{t}_i)$, $i = 1, ..., \bar{n}$, and the $(k-1)$ lagged values of these, in (9); cf. Perron and Vogelsang (1992).

Remark 5. Under Assumption I(0), the level breaks are given by $\gamma_{i,T}^* = \omega_u T^{-1/2}\gamma_i$, $i = 1, ..., n$, and they have no asymptotic effect on long run variance estimation. It follows that the consistency result of $\hat{\omega}_u^2$ for ω_u^2 (i.e. Theorem 3 (b)) does not actually require the fitting of the dummy variables in (8). However, in finite samples it is to be expected that the presence of level breaks will have a non-negligible impact on the behaviour of the long run variance estimator, thus we continue to include the dummies in the estimation of $\hat{\omega}_u^2$ so as to minimize the impact of any level breaks on the estimates in finite samples.

4.2 Feasible Tests and Asymptotic Size

Having proposed suitable long run variance estimators and established their asymptotic properties, we are now in a position to define feasible statistics for detecting multiple level breaks. The results of Theorem 1, along with the properties of the long run variance estimators described in Theorems 2 and 3, suggest the following statistics, appropriate under $I(1)$ and $I(0)$ errors, respectively:

$$S_1 := \hat{\omega}_{\varepsilon}^{-1} T^{-1/2} \mathcal{M} \tag{10}$$

$$S_0 := \hat{\omega}_{u}^{-1} T^{1/2} \mathcal{M}. \tag{11}$$

Remark 6. It is useful for analysis in subsequent sections to note that S_1 and S_0 could equivalently be expressed as $S_1 := \max_{t \in \Lambda_T} S_{1,t,\lfloor mT \rfloor}$ and $S_0 := \max_{t \in \Lambda_T} S_{0,t,\lfloor mT \rfloor}$, where $S_{1,t,\lfloor mT \rfloor} := \hat{\omega}_{\varepsilon}^{-1} T^{-1/2} |M_{t,\lfloor mT \rfloor}|$ and $S_{0,t,\lfloor mT \rfloor} := \hat{\omega}_{u}^{-1} T^{1/2} |M_{t,\lfloor mT \rfloor}|$.

In the following lemma we now establish the large sample behaviour of the S_1 and S_0 statistics of (10) and (11), respectively, in both $I(1)$ and $I(0)$ environments.

Lemma 1 *Let the conditions of Theorem 1 hold. Then,*

(a) Under Assumption I(1),
(i) $S_1 \xrightarrow{w} \sup_{r \in \Lambda} |L_1(r, m, c) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})|$*;*
(ii) $S_0 \xrightarrow{w} \frac{\sup_{r \in \Lambda} |L_1(r,m,c) + K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma})|}{Q^{1/2}(c,d,\tilde{\boldsymbol{\tau}})}$.

(b) Under Assumption I(0),
(i) $S_1 = O_p(kT^{-1})$*;*
(ii) $S_0 \xrightarrow{w} \sup_{r \in \Lambda} |L_0(r, m) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})|$.

Asymptotic null critical values for S_1 under $I(1)$ errors with $c = 0$, and S_0 under $I(0)$ errors, are reported in Table 1 for $m = \{0.10, 0.15, 0.20, 0.25, 0.30\}$, for the settings $\tau_L = 0.15$ and $\tau_U = 0.85$ (the symmetric interval commonly employed in the breaks testing literature), and for the significance levels $\xi = 0.10$, 0.05 and 0.01.[1] Our choices of m, τ_L and τ_U imply that the maximum number of possible breaks that are assumed to be admitted in the model are $n_{\max} = \{8, 5, 4, 3, 3\}$, respectively, for the values of m given above. The numerical results were obtained by simulation of the asymptotic distributions given in Lemma 1, setting $K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma}) = K(r, m, \boldsymbol{\tau}, \mathbf{0}) = 0$, approximating the functionals $L_1(r, m, 0)$ and $L_0(r, m)$ by normalized sums of 5,000 steps using normal $IID(0, 1)$ random variates. In the simulations here and in the remainder of the paper, unless stated otherwise, we use 50,000 Monte Carlo replications for computing critical values and sizes, and 20,000 replications for powers. All simulations were programmed in Gauss 9.0.

Tables 1 and 2 about here

It is also of interest, given lack of knowledge concerning the order of integration, to examine the asymptotic size properties of S_1 when $c > 0$, and also S_0 under both $c = 0$ and $c > 0$. These results are provided in Table 2, again obtained via direct simulation of

[1]The values that we consider for m are constrained to be no greater than 0.30, given that we require $\tau_L \geq m/2$ and $\tau_U \leq 1 - (m/2)$, with $\tau_L = 0.15$ and $\tau_U = 0.85$. Finite sample considerations lead us to use $m = 0.10$ as a lower bound, so that no less than 10% of the observations are used in the window over which $M_{t,\lfloor mT \rfloor}$ is computed.

$L_1(r, m, c)$ and $Q(c, d, \tilde{\boldsymbol{\tau}})$ in Lemma 1 (a). The S_1 test becomes increasingly under-sized as c increases; essentially, this is because the OU process $W_c(r)$ is distributed $N(0, V_{c,r})$, where $V_{c,r} := (2c)^{-1}(1 - e^{-2rc})$, is a monotonically decreasing function of c. Therefore, employing critical values which are appropriate for $c = 0$ will result in an under-sized test when $c > 0$. Other things being equal, the under-sizing becomes more apparent as m is increased.

Of particular interest is the behaviour of S_0 in the (local to) $I(1)$ case. Consider the nominal 0.05 significance level (i.e. Panel B). For $m = 0.10$, we see that the asymptotic size of S_0 is never above 0.003 across all c. As m increases, the maximum sizes are increasing in m but remain well below 0.05 (reaching a maximum across c of 0.036 when $m = 0.30$). Similar comments apply to tests conducted at the nominal 0.10 and 0.01 significance levels.

We now turn to consideration of a feasible test that can be applied in the absence of knowledge concerning the order of integration. Our approach deliberately exploits the under-sizing phenomenon seen in the S_0 test in the (local to) $I(1)$ world, and is based on the *union of rejections* approach advocated by Harvey *et al.* (2009b) in a unit root testing context. Specifically, we consider the union of rejections decision rule

$$U : \text{Reject } H_0 \text{ if } \{S_1 > \kappa_\xi cv_\xi^1 \text{ or } S_0 > \kappa_\xi cv_\xi^0\}$$

where cv_ξ^1 and cv_ξ^0 denote the ξ significance level asymptotic critical values of S_1 under $I(1)$ ($c = 0$) errors and S_0 under $I(0)$ errors, respectively, and κ_ξ is a positive scaling constant whose role is made precise below. Note that U can equivalently be expressed as

$$U : \text{Reject } H_0 \text{ if } \max\left\{S_1, \left(\frac{cv_\xi^1}{cv_\xi^0}\right) S_0\right\} > \kappa_\xi cv_\xi^1.$$

If the U decision rule was to be applied with $\kappa_\xi = 1$ (i.e. without any adjustment to the asymptotic critical values used for the constituent tests in U), then the testing strategy would be asymptotically correctly sized under $I(0)$ errors, as $S_1 \xrightarrow{p} 0$. In the $I(1)$ case, the Bonferroni inequality along with the size results for S_1 and S_0 reported in Table 2, show that such a strategy could only ever be (modestly) asymptotically over-sized when $c = 0$; indeed, for $m = 0.10$ the maximum possible asymptotic sizes at the 0.10, 0.05 and 0.01 nominal significance levels are, respectively, 0.104, 0.053 and 0.011, such that the size distortions will be almost non-existent. However, to ensure that U is an asymptotically conservative testing strategy (i.e. asymptotically exactly correctly sized in the case of $I(0)$ errors and $I(1)$ errors when $c = 0$, and always asymptotically under-sized elsewhere), we can avoid any size distortions by suitably choosing κ_ξ.

Noting that the maximum size of U is realized when $c = 0$, choosing κ_ξ such that U has an asymptotic size of ξ in this case ensures that the procedure will be conservative. We therefore obtain κ_ξ by simulating the limit distribution of $\max\{S_1, (cv_\xi^1/cv_\xi^0)S_0\}$, calculating the ξ-level critical value for this distribution, say $cv_\xi^{\max}$, and then computing $\kappa_\xi := cv_\xi^{\max}/cv_\xi^1$. Values of κ_ξ for different m and ξ are shown in Table 1. Hereafter, reference to the decision rule U assumes the κ_ξ adjustment values from Table 1 are used.

Table 2 also provides asymptotic size results for U. As expected, the testing strategy is correctly sized for $I(1)$ errors when $c = 0$. When the errors are $I(1)$ with $c > 0$, U is conservative, in line with the size properties of the constituent tests S_1 and S_0 discussed above. It is also conservative when the errors are $I(0)$.

Remark 7. It is important to note that the union of rejections procedure is only rendered viable due to the specific behaviour of the Berk-type estimator $\hat{\omega}_u^2$ under $I(1)$ errors, in

that it diverges at a rate T^2; see Theorem 3 (a). This ensures that S_0 is $O_p(1)$. If a typical kernel-based (e.g. Bartlett) long run variance estimator with bandwidth ℓ, say, growing at rate smaller than T was used, then under Assumption I(1), it is easy to show that $\hat{\omega}_u^2$ diverges at a rate less than T^2, so that S_0 diverges to ∞. In such a case, a union of rejections approach is clearly precluded, because, regardless of the choice of κ_ξ, its size would approach one in the limit under $I(1)$ errors.[2]

4.3 Asymptotic Power

Table 3 shows asymptotic local powers of S_1, S_0 and U, conducted at the nominal 0.05-level. We consider the same settings of m (and hence $n_{\max}$) as in Table 2, and the same error specifications (i.e. $I(1)$ errors with $c \geq 0$ and $I(0)$ errors). As regards the break(s) in mean, we consider four different specifications: Panel A provides results for a single mid-point break, while Panels B, C and D present results for two, three and four equally spaced breaks in the DGP, respectively. The break magnitudes are common across both the break dates and the number of breaks in the DGP, i.e. $\gamma_1 = \gamma_2 = ... = \gamma_n = \gamma$, and benchmarked so that the powers of S_1 for $c = 0$ in the $I(1)$ case, and S_0 in the $I(0)$ case, are equal to 0.50 when $m = 0.10$ and there is a single break in the generated data.[3]

Table 3 about here

Consider first the behaviour of S_1. For $I(1)$ errors and a given c, power decreases monotonically as m increases, as might be expected in view of the discussion of the stylized example in section 3. In addition, for a given m and number of breaks, power is monotonically decreasing as c increases, while for a given m and c, power is increasing in the number of breaks, as would be expected since the test essentially now has an increasing number of opportunities to detect a level break. It can also be seen that the power losses that accrue across m do so at a faster rate for larger values of c. For example, in the two break case of Panel B, when $c = 0$, power falls from 0.736 when $m = 0.10$ to 0.324 when $m = 0.30$, but when $c = 40$, power falls from 0.694 to zero across this same range in m. Note also that for $I(0)$ errors, the power of S_1 is always zero, in line with the results of Lemma 1 (b) (i).

The power of S_0 under $I(0)$ errors is seen to increase monotonically as m increases, for example in the one break case of Panel A, rising from the benchmarked power of 0.50 for $m = 0.10$ to a power of almost one for $m = 0.30$, which is again consistent with the discussion in section 3. When the errors are $I(1)$, the power of S_0 rises monotonically in c for a given value of m. When $c = 0$ (with the exception of $m = 0.30$ when only a single break is present), S_0 has lower power than S_1; indeed, the power displayed is often trivial. As c increases, this ranking becomes reversed, with S_0 being substantially more powerful than S_1 for all m and all numbers of breaks once $c = 40$. This feature is somewhat surprising, given that in these cases S_0 was seen to be markedly under-sized. For a given value of c, the power of S_0 is broadly increasing in m, although this pattern is not monotone.

Inspection of the power performance of U shows that it essentially capitalizes on the relatively high power of S_1 for $I(1)$ errors when c is zero (or small) while simultaneously

[2]As pointed out by a referee, an alternative to the autoregressive spectral density long run variance estimator we adopt here would be to use the fixed-bandwidth kernel-based estimator suggested by Kiefer and Vogelsang (2005), since this is also T^2 divergent.

[3]Note that no results are reported for $m = 0.30$ in the case of three breaks, or $m = 0.25$ and $m = 0.30$ in the case of four breaks, since, given our chosen dates for the breaks in the DGP, these settings of m would violate our assumption that no more than one break can occur within any $\lfloor mT \rfloor$ observations.

capturing the superior power properties of S_0 for larger c and also $I(0)$ errors. For a given DGP and choice of m, U generally displays power very close to the better power of the two individual tests S_1 and S_0. Some minor power losses inevitably arise due to the adjustment parameter κ_ξ included to ensure that U is an asymptotically conservative testing strategy. These losses are not surprisingly at their largest when the adjustment factor is furthest from one, i.e. when $m = 0.30$ where a power loss of 0.058 is seen for the two break case when $c = 10$. Conversely, they are at their smallest when the adjustment factor is closest to one, i.e. when $m = 0.10$. On the other hand, there are many instances where the power of U exceeds that of either of the constituent tests S_1 and S_0, resulting from the fact that the rejections from S_1 and S_0 need not be perfectly correlated. These power gains can be quite substantial, for example in the two break case with $c = 10$, a gain of 0.128 is evident over the best of S_1 and S_0 when $m = 0.15$. The robust power performance of U relative to S_1 and S_0, whose powers are sensitive to particular properties of the DGP, therefore makes a strong case for using the modified union of rejections approach in practice.

5 Finite Sample Analysis

In order to ensure that the robust strategy U retains decent size control in finite samples, in Table 4 we report finite sample null critical values for the S_1 and S_0 tests, again for $m = \{0.10, 0.15, 0.20, 0.25, 0.30\}$, together with finite sample variants of the adjustment factors κ_ξ. These are calculated by simulation using the DGP (1)-(2) with $y_t = u_t$ (without loss of generality). To obtain the critical values for S_1 we generate $I(1)$ data by setting $\rho = 1$, with $u_1 = \varepsilon_1$, while for S_0 we generate $I(0)$ data with $\rho = 0$. In both cases we generate $\varepsilon_t = \eta_t$ as $IIDN(0, 1)$. In the computation of the test statistics here and in all subsequent finite sample results, the long run variance estimators $\hat{\omega}^2_\varepsilon$ and $\hat{\omega}^2_u$ use values of k (in (7) and (9), respectively) determined according to the BIC criterion with $k_{\max} = \lfloor 4(T/100)^{1/4} \rfloor$. We consider the sample sizes $T = \{150, 300, 600, 1200\}$. Because the tests are based on only modest fractions of the data (m), it is perhaps no surprise to see that the speed of convergence of the finite sample critical values to their limiting counterparts in Table 1 is fairly slow. This makes a *prima facie* case for employing the finite sample critical values (and the corresponding adjustment factors) of Table 4 unless sample sizes are reasonably large. We will therefore adopt these in all the remaining simulations.

Table 4 about here

5.1 Robustness to Non-Normal Errors

We next investigate the finite sample size of the recommended U procedure in the presence of non-normal errors. While the assumption of normality is not required for any of our asymptotic results, our finite sample critical values are nonetheless calculated using normally distributed errors. Since fat-tailed and/or skewed data is often encountered when modelling macroeconomic and financial time series, it is therefore important to assess the effects of such data on the finite sample size of our procedure.

Table 5 reports simulated empirical sizes of nominal 0.05-level U tests using the DGP (1)-(2) with $y_t = u_t$, $\rho = \{1, 0\}$, $u_1 = \varepsilon_1$, and $\varepsilon_t = \eta_t$ generated as $IIDN(0, 1)$, $IID\,t(5)$ (the smallest degrees of freedom we can permit since we require the error fourth moment to be finite; see Assumption LP), and $IID\,\chi^2(3)$ (centered). While still providing plausible

distributional assumptions for errors in macroeconomic and financial time series, the latter two distributions allow for a high degree of kurtosis in the innovations, and in the case of $\chi^2(3)$, the errors are also highly skewed. The simulations are conducted for the sample sizes $T = \{150, 300, 600, 1200\}$.

Table 5 about here

The U test sizes do not seem unduly sensitive to the normality assumption (under which the finite sample critical values are derived). For $T = 150$, our test has at most a size of around 0.16, occurring with $I(1)$ errors when the window width $m = 0.10$ and is generally much closer to the nominal level either under $I(0)$ errors or when we employ larger window widths. Of course, our procedure is asymptotically size controlled and therefore becomes robust to non-normality as we increase T, as Table 5 makes evident.

As a point of comparison, we also show the corresponding sizes of the KP procedure which allows for testing at least one simultaneous break in level and/or trend. Since the procedure essentially involves sequential application of the PY test, the KP test of the null of no breaks is simply the PY test for a single break in level/trend. This takes the form of a Wald test for a break, where the statistic is computed using either a first differenced regression (if PY's estimator of ρ is found to be within a $T^{-1/2}$ neighbourhood of 1), or a quasi-feasible GLS regression (otherwise).[4] We see that in the $I(1)$ case, when the errors are non-normal, PY suffers very significant upward size distortion, to the extent that in large samples the test can almost always reject the null of no breaks when the distribution of the innovations departs from the normal. This arises because here the PY statistic involves searching for a single outlier in first differenced data which precludes the application of an invariance principle (cf. section 3), so their critical values need to be based on a known error distribution, which they assume to be normal. Clearly then, such an approach will lack robustness to departures from this normality assumption, both in finite samples and asymptotically.

5.2 Robustness to Non-IID Errors

Here we analyze the finite sample size of the U procedure when the errors ε_t are autocorrelated. Table 6 reports simulated empirical sizes of nominal 0.05-level U tests using the DGP (1)-(2) with $y_t = u_t$ and ARIMA errors for u_t. Specifically, we consider $\rho = \{1.00, 0.95, 0.90, 0.70, 0.00\}$, with ε_t generated according to the $MA(1)$ scheme

$$\varepsilon_t = \eta_t - \theta\eta_{t-1},\ t = 1, ..., T,$$

with η_t generated as $IIDN(0, 1)$, $u_1 = \varepsilon_1 = \eta_1$, and $\theta = \{0.0, 0.5, -0.5\}$. Results are reported for $T = \{150, 300\}$ and these show the dynamics to have reassuringly little effect on finite sample size; in particular, they do not lead to much in the way of over-rejection, which is perhaps the greater concern. As would be expected from the asymptotic results, what over-sizing does occur generally diminishes as we increase sample size from $T = 150$ to $T = 300$.

Table 6 about here

[4]Here, and in Tables 7 and 8 below, due to the substantial computational requirements of PY, the KP entries are based on 1000 Monte Carlo replications, as in PY and KP.

5.3 Power

Lastly in this section, we consider the finite sample power of the tests, reporting results for nominal 0.05-level tests using a sample size of $T = 300$. We use the same four specifications for the number and timings of breaks as in the asymptotic power simulations, and again we consider $m = \{0.10, 0.15, 0.20, 0.25, 0.30\}$. In contrast to the settings for the asymptotic simulations, we here also report results for $m = 0.30$ in the case of three breaks, and $m = 0.25$ and $m = 0.30$ in the case of four breaks. Given our chosen dates for the breaks in the DGP, these settings violate our assumption that no more than one break can occur within any $\lfloor mT \rfloor$ observations, but we include these results to evaluate the behaviour of our proposed tests in such circumstances. The entries in the table for these cases are italicized, so as to distinguish them from the cases where our assumption is upheld.

The data are generated according to (1) and (2), with $\alpha = 0$, $u_1 = \varepsilon_1$ and $\varepsilon_t = \eta_t$ generated as $IIDN(0, 1)$. We consider $\rho = \{1.00, 0.95, 0.90, 0.70, 0.50\}$ and set a common break magnitude γ across all breaks and DGPs. Note that break magnitudes are held constant across both the $I(1)$ and $I(0)$ DGPs that we consider, rather than scaling the magnitudes according to the order of integration (as in our asymptotic analysis); this is done so as to provide some consistency across different values of ρ.

Tables 7 and 8 about here

In Table 7 we show the results for $\gamma = 10$. This value is selected so that power for S_1 is around 0.50 when $\rho = 1$ and $m = 0.10$ in the single break case, aiding comparison with our asymptotic results. In fact, across $\rho = 1$ the overall pattern of results for S_1, S_0 and U bears a close resemblance to the asymptotic case (see the $I(1)$, $c = 0$ rows of Table 3). In the cases where $\rho < 1$, S_0 behaves in the same manner as in the asymptotic case, i.e. power increases as the series becomes less persistent, and also broadly as m increases. However, in contrast to the asymptotic results, we now observe that while the power of S_1 generally falls when ρ changes from 1 to 0.95 and to 0.90, it then starts to increase as the level of persistence falls further. This arises because for the smaller values of ρ, the asymptotic results for $I(0)$ errors are not really applicable, since here the break magnitude is held common across $I(1)$ and $I(0)$ DGPs, unlike in the asymptotic model where the break magnitude is assumed to be an order T smaller in the $I(0)$ case. Although S_1 often has non-negligible power for the $I(0)$ DGPs considered, it is observed that S_0 is more powerful than S_1 when $\rho < 0.95$, in many cases quite substantially so.

As regards the performance of U, it is clear that, across all values of ρ, it displays power close to the maximum of S_1 and S_0. In fact, for $m = 0.10$ its power often exceeds the higher of both S_1 and S_0; although, for larger values of m, U displays power slightly lower than the better of S_1 and S_0, due to the effect of the larger critical values adjustment factors needed here. The robust performance of U seen in the asymptotic results therefore translates directly to finite samples, strengthening the case for its use in practical applications.

In the cases where our assumption that no more than one break can occur within any $\lfloor mT \rfloor$ observations does not hold (the italicized entries), we see that S_1, S_0 and therefore U have very low levels of power. This arises partly because $M_{t,\lfloor mT \rfloor}$ now often straddles two level breaks while only allowing for one, but chiefly because the long run variance estimators $\hat{\omega}_\varepsilon^2$ and $\hat{\omega}_u^2$ are now calculated from data for which not all of the level breaks have been purged. Predictably, these variance estimators become inflated, relative to their fully purged counterparts, and the values of S_1 and S_0 are consequently forced closer towards

zero. In an empirical context, however, if one suspects a non-rejection of the null hypothesis might be due to adopting too large a value m for the spacing of breaks, this can always be replaced by a smaller value, since the corresponding test would obviously be less susceptible to this potential problem.

Table 8 repeats the analysis using $\gamma = 5$. Not surprisingly, the powers are somewhat lower than in Table 8, particularly for the larger values of ρ. Otherwise, the same comments broadly apply, with the power of U again being fairly close to the maximum of S_1 and S_0, particularly for the smaller values of m. Tables 7 and 8 additionally report the power of the KP procedure. This is seen to be often substantially more powerful than U, most noticeably in those cases where ρ is close to one. High power when $\rho = 1$ is only to be expected, since, as noted above, KP is here detecting outliers in the first differenced data while (correctly) assuming normal errors. The price to be paid for these power gains is, of course, the very poor size control demonstrated when the errors are in fact non-normal.

Having established the finite sample properties of our proposed tests, we are now in a position to provide some comments as regards the choice of m. From an asymptotic perspective, the tests for all choices of m are never over-sized, while the results of Table 3 show that there is no setting for m that unambiguously provides the highest asymptotic power for U in all circumstances. However, additional considerations that may have some bearing on the decision are as follows. In the presence of non-normal innovations, upward finite sample size distortions are less pronounced for larger values of m, while smaller values of m allow more flexibility in terms of the maximum number of breaks that are assumed to be present in the data (i.e. $n_{\max}$), and also provide greater insurance against the possibility of more than one break occurring within any $\lfloor mT \rfloor$ observations. Using a smaller value of m also has some theoretical appeal in that it involves less adjustment to the critical values in the union of rejections (i.e. is closer to performing a raw unadjusted union of rejections). Overall, then, it would seem that a choice of $m = 0.10$ or $m = 0.15$ might be appealing in practice, provided the sample size is not too small.

6 Determining the Number and Timing of Breaks

A rejection by any of the test procedures outlined in section 4.2 informs us that at least one level break is present (subject to Type 1 error). We now discuss how one can proceed to detect and date possibly multiple breaks in level.

Suppose first that it is the case that $S_1 := \max_{t \in \Lambda_T} S_{1,t,\lfloor mT \rfloor} > cv_\xi^1$, such that S_1 is significant at the ξ-level. We then assume that there is a (first) level break at $\tilde{t}_1 := \arg\max_{t \in \Lambda_T} S_{1,t,\lfloor mT \rfloor}$. Given that we exclude the possibility of more than one break in the interval $[\tilde{t}_1 - \lfloor mT \rfloor + 1, \tilde{t}_1 + \lfloor mT \rfloor - 1] =: \Lambda_{1,T}$, we next examine the possibility of a further break occurring in the remaining portion of Λ_T, not excluded by $\Lambda_{1,T}$, i.e. $\Lambda_T - \Lambda_{1,T}$. Then, if $\max_{t \in \Lambda_T - \Lambda_{1,T}} S_{1,t,\lfloor mT \rfloor} \leq cv_\xi^1$ we terminate the algorithm and conclude that only one level break is present. Otherwise, if $\max_{t \in \Lambda_T - \Lambda_{1,T}} S_{1,t,\lfloor mT \rfloor} > cv_\xi^1$, we then record a second level break at $\tilde{t}_2 := \arg\max_{t \in \Lambda_T - \Lambda_{1,T}} S_{1,t,\lfloor mT \rfloor}$. Next, we exclude the dates in the interval $[\tilde{t}_2 - \lfloor mT \rfloor + 1, \tilde{t}_2 + \lfloor mT \rfloor - 1] =: \Lambda_{2,T}$, and consider whether a further break can be detected in the range $\Lambda_T - \Lambda_{1,T} - \Lambda_{2,T}$, i.e. if $\max_{t \in \Lambda_T - \Lambda_{1,T} - \Lambda_{2,T}} S_{1,t,\lfloor mT \rfloor} \leq cv_\xi^1$, we terminate and conclude that two level breaks are present, while if $\max_{t \in \Lambda_T - \Lambda_{1,T} - \Lambda_{2,T}} S_{1,t,\lfloor mT \rfloor} > cv_\xi^1$ we then record a third level break at $\tilde{t}_3 := \arg\max_{t \in \Lambda_T - \Lambda_{1,T} - \Lambda_{2,T}} S_{1,t,\lfloor mT \rfloor}$. This sequential detection and dating procedure continues until there are no further significant level breaks

or the search set is null.

Entirely analogous procedures for detecting and dating multiple breaks using $S_{0,t,\lfloor mT \rfloor}$ can also be implemented, but must obviously be based on significance testing using the critical value, cv_ξ^0. Now let the number of significant breaks detected on the basis of $S_{1,t,\lfloor mT \rfloor}$ and $S_{0,t,\lfloor mT \rfloor}$ be denoted by n_1 and n_0, respectively. Since we can write $S_{0,t,\lfloor mT \rfloor} = \hat{\omega}_u^{-1}\hat{\omega}_\varepsilon T S_{1,t,\lfloor mT \rfloor}$, i.e. for a given T, $S_{0,t,\lfloor mT \rfloor}$ is simply a constant multiple of $S_{1,t,\lfloor mT \rfloor}$, it follows that if $n_1 \geq n_0$, then in terms of their locations, the n_0 breaks are simply a subset of the n_1 breaks, with both sets of break locations being identical if $n_1 = n_0$. Similarly, if $n_1 < n_0$, then the locations of the n_1 breaks are simply a subset of those of the n_0 breaks.

A corresponding approach based around the union of rejections decision rule, U, can also be developed in a straightforward way. Here, we first consider the detection and dating of breaks according to $S_{1,t,\lfloor mT \rfloor}$ and $S_{0,t,\lfloor mT \rfloor}$ separately, using the procedure outlined above, only with both critical values adjusted by the scaling factor κ_ξ, i.e. we replace cv_ξ^1 with $\kappa_\xi cv_\xi^1$ in the above procedure for detection based on $S_{1,t,\lfloor mT \rfloor}$, and replace cv_ξ^0 with $\kappa_\xi cv_\xi^0$ in for detection based on $S_{0,t,\lfloor mT \rfloor}$. Let the number of significant breaks detected according to these procedures be denoted by n_1' and n_0', respectively.[5] The number of breaks associated with the sequential procedure based on U is then simply $\max(n_1', n_0') := n_U$.

In Table 9 we present finite sample simulation results for the number of breaks detected by this sequential procedure. We focus on the case of three breaks in the DGP ($\tau_1, \tau_2, \tau_3 = 0.25, 0.50, 0.75$). The data are generated in exactly the same way as for the simulations reported in Tables 7-8, except that for the common break magnitude γ we now consider three different values: 5, 10 and 15. The table reports the frequency with which one, two, three or more than three breaks are detected by S_1, S_0 and U, when conducting these tests at the nominal 0.05-level using $m = \{0.10, 0.15, 0.20, 0.25\}$.[6] The sum of the frequencies is the total power of the test (for $\gamma = 10$ and $\gamma = 5$ these totals are the same as those powers given in Panel C of Tables 7 and 8, respectively, other than for rounding errors).

Table 9 about here

Other things equal, we see that the higher is the power of a given test, the more frequently the corresponding number of identified significant breaks is equal to the true value of three. Also, as the magnitude of the breaks increases, a migration towards detecting three breaks is clearly evident. Conversely, when a test has very low power, it is seen to detect three breaks only very infrequently. Given this relationship between the power of the test of the null hypothesis of no breaks and the frequency with which the correct number of breaks is identified, it is clear that break detection based on the U approach has inherent advantages in terms of identifying multiple breaks in level, in addition to the superior power properties of U when simply rejecting the null of no breaks.

7 Allowing for Linear Trends

In this section we briefly discuss how the procedures outlined thus far can be extended to accommodate a linear trend in the underlying DGP. In order to do so, we need to augment

[5] Note that due to the critical value adjustment factor κ_ξ, it must be true that $n_1' \leq n_1$ and $n_0' \leq n_0$.

[6] We omit results for $m = 0.30$ since, given our chosen dates for the breaks in the DGP, this setting of m would violate our assumption that no more than one break can occur within any $\lfloor mT \rfloor$ observations.

the observation equation in (1) with a linear trend term, *viz*,

$$y_t = \alpha + \beta t + \sum_{i=1}^{n} \gamma_{i,T}^{*} DU_t(\lfloor \tau_i T \rfloor) + u_t, \quad t = 1, ..., T. \tag{12}$$

In order to retain pivotal inference procedures in the presence of the additional linear trend term, βt, in (12) we must proceed as follows. First, let $\hat{\beta}$ denote the estimator of the trend coefficient, β, from the OLS regression of y_t on $(1, t)'$, $t = 1, ..., T$. We then define the corresponding de-trended fluctuation measure (which is exact invariant to both α and β)

$$M_{t,\lfloor mT \rfloor}^{*} := M_{t,\lfloor mT \rfloor} - \hat{\beta} \lfloor mT/2 \rfloor$$

and the corresponding maximum (in absolute value) of the sequence of such fluctuation measures, taken over all possible break points in Λ_T:

$$\mathcal{M}^{*} := \max_{t \in \Lambda_T} |M_{t,\lfloor mT \rfloor}^{*}|.$$

Next, in the computation of $\hat{\omega}_{\varepsilon}^2$, we re-define $\hat{t}_1 := (\arg\max_{t \in \Omega_T} |\Delta y_t - \tilde{\beta}|) - 1$ where $\tilde{\beta} := (T-1)^{-1} \sum_{t=2}^{T} \Delta y_t$ and similarly re-define $\hat{t}_2, ..., \hat{t}_{\bar{n}}$. We now calculate $\hat{\omega}_{\varepsilon}^2$ using the OLS residuals from (6) augmented to include a constant term in the regression. Similarly, we also calculate $\hat{\omega}_{u}^2$ from the OLS residuals from (8) now augmented to include a linear time trend in the regression (and using the re-defined $\hat{t}_1, ..., \hat{t}_{\bar{n}}$).

The time-trend adjusted analogues of the statistics in (10) and (11) are then given by

$$S_1^{*} := \hat{\omega}_{\varepsilon}^{-1} T^{-1/2} \mathcal{M}^{*} \tag{13}$$

and

$$S_0^{*} := \hat{\omega}_{u}^{-1} T^{1/2} \mathcal{M}^{*} \tag{14}$$

respectively. In Lemma 2 we detail the large sample behaviour of these statistics in both $I(1)$ and $I(0)$ environments.

Lemma 2 *Let the conditions of Theorem 1 hold. Then,*

(a) Under Assumption I(1),
(i) $S_1^{*} \xrightarrow{w} \sup_{r \in \Lambda} |L_1(r, m, c) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma}) + L_1^{*}(m, c) + K^{*}(m, \boldsymbol{\tau}, \boldsymbol{\gamma})|$;
(ii) $S_0^{*} \xrightarrow{w} \frac{\sup_{r \in \Lambda} \left| L_1(r,m,c) + K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma}) + L_1^{*}(m,c) + K^{*}(m,\boldsymbol{\tau},\boldsymbol{\gamma}) \right|}{Q^{*1/2}(c,d,\tilde{\boldsymbol{\tau}})}$.

(b) Under Assumption I(0),
(i) $S_1^{*} = o_p(1)$;
(ii) $S_0^{*} \xrightarrow{w} \sup_{r \in \Lambda} |L_0(r, m) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma}) + L_0^{*}(m, c) + K^{*}(m, \boldsymbol{\tau}, \boldsymbol{\gamma})|$.

where $L_1(r, m, c)$, $L_0(r, m)$ *and* $K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})$ *are as defined in Theorem 1, and*

$$\begin{aligned}
L_1^{*}(m, c) &:= -6m \left\{ \int_0^1 s W_c(s) ds - \frac{1}{2} \int_0^1 W_c(s) ds \right\}, \\
L_0^{*}(m, c) &:= -6m \left\{ \int_0^1 s dW(s) - \frac{1}{2} W(1) \right\}, \\
K^{*}(m, \boldsymbol{\tau}, \boldsymbol{\gamma}) &:= -3m \sum_{i=1}^{n} \gamma_i \tau_i (1 - \tau_i), \\
Q^{*}(c, d, \tilde{\boldsymbol{\tau}}) &:= \frac{\{\int_0^1 H^{*}(r, c, d, \tilde{\boldsymbol{\tau}})^2 dr\}^2}{\{\int_0^1 H^{*}(r, c, d, \tilde{\boldsymbol{\tau}}) dW_c(r)\}^2}
\end{aligned}$$

with $H^*(r, c, d, \tilde{\boldsymbol{\tau}})$ *is a continuous time residual from the projection of* $W_c(r)$ *onto the space spanned by* $\{1, r, 1(r > \tilde{\tau}_1), 1(r > \tilde{\tau}_2), ..., 1(r > \tilde{\tau}_{\bar{n}})\}$ *with* $\tilde{\tau}_i := \lim_{T\to\infty} T^{-1}\hat{t}_i$, $i = 1, 2, ..., \bar{n}$ *and* $\tilde{\boldsymbol{\tau}} := [\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_{\bar{n}}]$.

Asymptotic and finite sample null critical values for the tests based on the S_1^* and S_0^* statistics are given in Tables 10 and 11 respectively. These were obtained using the same settings as were outlined for Tables 1 and 4, respectively.

Tables 10 and 11 about here

Given the trend-modified statistics S_1^* and S_0^*, a union of rejections procedure, say U^*, that allows for trending behaviour can be formed as detailed in section 4.2, replacing the original statistics S_1 and S_0 with the trend-modified variants S_1^* and S_0^* of (13) and (14) respectively. An adjustment factor, κ_ξ^* say, is once more required to control the size of the U^* procedure. Asymptotic and finite sample values of κ_ξ^* are also reported in Tables 10 and 11; we see that these are again very modest, particularly for the smaller values of m.

8 Conclusion

In this paper we have discussed procedures, based on generalized fluctuation measures, for testing for the presence and location of multiple levels breaks (possibly around a constant linear trend) in autocorrelated time series processes. In contrast to the extant literature, our proposed union of rejections based procedure is robust as to whether the data are $I(1)$ or $I(0)$ and to distributional assumptions on the underlying errors. We have provided representations for and critical values from the asymptotic distributions of our proposed statistics (those appropriate for the $I(0)$ and $I(1)$ environments and the union of rejections approach) under the null hypothesis of no level breaks, together with representations for and numerical evaluation of their asymptotic local power functions under both $I(0)$ and $I(1)$ environments. Associated estimators of the level break fractions, based on the statistics appropriate for the $I(0)$ and $I(1)$ environments, were also provided and evaluated numerically. Monte Carlo simulations were also reported which suggested that our proposed methods perform well in small samples, regardless of the (unknown) order of integration of the data. Our proposed tests and level break detection procedure require a choice of window width which in turn impacts on the maximum number of breaks allowable. A detailed numerical examination of the impact of the choice of window width on both the asymptotic and finite sample performance of our approach was reported, with a choice of $m = 0.10$ or $m = 0.15$ appearing to deliver the greatest flexibility, obviously provided the sample size involved is not too small. Overall, the robust tests that we propose should prove useful in practical applications, particularly when dealing with long spans of macroeconomic or financial data where multiple level breaks are an important consideration, and where uncertainty exists as to the order of integration properties of the data.

Appendix

Proof of Theorem 1

(a) Under Assumption I(1)

$$T^{-1/2}M_{\lfloor rT\rfloor,\lfloor mT\rfloor} = T^{-1/2}\frac{\sum_{i=1}^{\lfloor \frac{m}{2}T\rfloor} u_{\lfloor rT\rfloor+i} - \sum_{i=1}^{\lfloor \frac{m}{2}T\rfloor} u_{\lfloor rT\rfloor-i+1}}{\lfloor \frac{m}{2}T\rfloor} + \omega_\varepsilon K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma})$$

noting that $M_{\lfloor rT\rfloor,\lfloor mT\rfloor}$ is invariant to the intercept term α and to any level breaks that occur prior to the earliest time period spanned by $M_{\lfloor rT\rfloor,\lfloor mT\rfloor}$, and where $K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma})$ is as given in the statement of Theorem 1. Then, since we have the FCLT, $T^{-1/2}u_{\lfloor rT\rfloor} \xrightarrow{w} \omega_\varepsilon W_c(r)$, it follows from the Continuous Mapping Theorem [CMT] that

$$T^{-1/2}M_{\lfloor rT\rfloor,\lfloor mT\rfloor} \xrightarrow{w} 2m^{-1}\omega_\varepsilon\left\{\int_r^{r+m/2} W_c(s)ds - \int_{r-m/2}^{r} W_c(s)ds\right\} + \omega_\varepsilon K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma}).$$

This result combined with the CMT delivers the stated result in (a).

(b) Under Assumption I(0)

$$T^{1/2}M_{\lfloor rT\rfloor,\lfloor mT\rfloor} = T^{1/2}\frac{\sum_{i=1}^{\lfloor \frac{m}{2}T\rfloor} u_{\lfloor rT\rfloor+i} - \sum_{i=1}^{\lfloor \frac{m}{2}T\rfloor} u_{\lfloor rT\rfloor-i+1}}{\lfloor \frac{m}{2}T\rfloor} + \omega_u K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma}).$$

Then, since we now have the FCLT, $T^{-1/2}\sum_{t=1}^{\lfloor rT\rfloor} u_t \xrightarrow{w} \omega_u W(r)$, it follows from the CMT that

$$T^{1/2}M_{\lfloor rT\rfloor,\lfloor mT\rfloor} \xrightarrow{w} 2m^{-1}\omega_u\left\{W(r+m/2) - 2W(r) + W(r-m/2)\right\} + \omega_u K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma}).$$

This result combined with the CMT delivers the stated result in (b).

Proof of Theorem 2

(a) (i) This part of the proof follows the approach of Perron and Zhu (2005, Theorem 3.1). Writing (1) in first differences we have that

$$\Delta y_t = \sum_{i=1}^{n} \gamma_{i,T}^{*} D_t(\lfloor \tau_i T\rfloor) + \Delta u_t. \tag{A.1}$$

Now, it should be obvious that $\hat{t}_1 := (\arg\max_{t\in\Omega_T} |\Delta y_t|) - 1$ is identical to the estimator

$$\left\{\arg\min_{j\in\Omega_T} T^{-1}\sum_{t=1}^{T} f_t(j)^2\right\} - 1$$

where $f_t(j)$ are the OLS residuals from a regression of Δy_t on $D_t(\lfloor jT\rfloor)$. As a consequence, denoting the sum of squared residuals from a regression of Δy_t on $D_t(\lfloor .\rfloor)$ as $SSR(.)$, we find

$$SSR(\hat{t}_1) - SSR(\lfloor \tau_1 T\rfloor) \leq 0. \tag{A.2}$$

Next, let $\tilde{t} \in \Omega_T - 1$. Then,

$$T^{-1}\{SSR(\tilde{t}) - SSR(\lfloor \tau_1 T \rfloor)\} = \omega_\varepsilon^2 \begin{cases} 0 & \tilde{t} = \lfloor \tau_1 T \rfloor \\ \gamma_1^2 - \gamma_i^2 + o_p(1) & \tilde{t} = \lfloor \tau_i T \rfloor,\ i = 2, ..., n \\ \gamma_1^2 + o_p(1) & \text{otherwise.} \end{cases} \quad \text{(A.3)}$$

This follows since, from (A.1),

$$\begin{aligned} T^{-1}SSR(\tilde{t}) &= T^{-1} \sum_{t \neq \tilde{t}+1}^{T} \left\{ \sum_{i=1}^{n} \gamma_{i,T}^* D_t(\lfloor \tau_i T \rfloor) + \Delta u_t \right\}^2 \\ &= T^{-1} \sum_{t=1}^{T} \Delta u_t^2 + o_p(1) + \omega_\varepsilon^2 \begin{cases} \sum_{i=1}^{n} \gamma_i^2 & \tilde{t} \neq \lfloor \tau_i T \rfloor,\ i = 1, ..., n \\ \sum_{j=1, j \neq i}^{n} \gamma_j^2 & \tilde{t} = \lfloor \tau_i T \rfloor,\ i = 1, ..., n \end{cases} \quad \text{(A.4)} \end{aligned}$$

from which we have that

$$T^{-1}SSR(\lfloor \tau_1 T \rfloor) = T^{-1} \sum_{t=1}^{T} \Delta u_t^2 + o_p(1) + \omega_\varepsilon^2 \sum_{i=2}^{n} \gamma_i^2. \quad \text{(A.5)}$$

Subtracting (A.5) from (A.4) establishes (A.3).

Now define $\Pi := \{\tilde{t} : |\tilde{t} - \lfloor \tau_1 T \rfloor| > 0\}$; we find that $\lim_{T \to \infty} \min_{\tilde{t} \in \Pi} T^{-1}\{SSR(\tilde{t}) - SSR(\lfloor \tau_1 T \rfloor)\} > 0$, as $\gamma_1^2 > 0$ and $\gamma_1^2 - \gamma_i^2 > 0$ for $i = 2, ..., n$. Hence, asymptotically, $\hat{t}_1 \notin \Pi$ as this leads to a contradiction of (A.2). Therefore $\lim_{T \to \infty} \Pr(\hat{t}_1 - \lfloor \tau_1 T \rfloor = 0) = 1$; that is, $\hat{t}_1 - \lfloor \tau_1 T \rfloor \xrightarrow{a.s.} 0$, which implies $\hat{t}_1 - \lfloor \tau_1 T \rfloor \xrightarrow{p} 0$.

To show that $\hat{t}_2 - \lfloor \tau_2 T \rfloor \xrightarrow{p} 0$, we simply repeat the above argument, replacing Ω_T with $\Omega_T - \Omega_{1,T}$, and noting that the interval $\Omega_{1,T} = [\hat{t}_1 - \lfloor dT \rfloor + 2, \hat{t}_1 + \lfloor dT \rfloor]$ contains $\lfloor \tau_2 T \rfloor$, but excludes $\lfloor \tau_1 T \rfloor$ almost surely. The results for $\hat{t}_3, ..., \hat{t}_n$ follow in entirely analogous fashion.

Finally, if $\Omega_{n,T} \neq \varnothing$, the procedure will continue to determine break points until $\Omega_{\bar{n}+1,T} = \varnothing$; it follows, therefore, that $\bar{n} \geq n$ as $T \to \infty$.

(a) (ii) Using (A.1), if the n break points $\lfloor \tau_i T \rfloor$, $i = 1, ..., n$, were known, we would regress Δy_t on the $D_t(\lfloor \tau_i T \rfloor)$ and obtain the residuals

$$\begin{aligned} g_t := \quad & \Delta y_t - \sum_{i=1}^{n} \Delta y_t D_t(\lfloor \tau_i T \rfloor) \\ = \quad & \Delta u_t - \sum_{i=1}^{n} \Delta u_t D_t(\lfloor \tau_i T \rfloor) \end{aligned}$$

and then construct the long run variance based on g_t rather than $\hat{\varepsilon}_t$ of equation (7). Denoting such an estimator $\hat{\omega}_\varepsilon^2(g)$, this would consistently estimate ω_ε^2 since we have simply removed a finite number n of the Δu_t, replacing each with zero, which has no asymptotic impact.

From (6), we find

$$\begin{aligned} \hat{\varepsilon}_t &= \Delta y_t - \sum_{i=1}^{\bar{n}} \Delta y_t D_t(\hat{t}_i) \\ &= g_t + \sum_{i=1}^{n} \Delta y_t \{D_t(\lfloor \tau_i T \rfloor) - D_t(\hat{t}_i)\} - \sum_{i=n+1}^{\bar{n}} \Delta y_t D_t(\hat{t}_i). \end{aligned}$$

Now, for $i = 1, ..., n$

$$\begin{aligned}\left|\Delta y_t\{D_t(\lfloor\tau_i T\rfloor) - D_t(\hat{t}_i)\}\right| &\leq \max_{t\in\Omega_T}|\Delta y_t| \max_{t\in\Omega_T}\left|D_t(\lfloor\tau_i T\rfloor) - D_t(\hat{t}_i)\right| \\ &= \{\omega_\varepsilon T^{1/2}|\gamma_1| + O_p(1)\}o_p(1) \\ &= o_p(T^{1/2})\end{aligned}$$

where the $o_p(1)$ term arises since $\max_{t\in\Omega_T}\left|D_t(\lfloor\tau_i T\rfloor) - D_t(\hat{t}_i)\right| \xrightarrow{p} 0$ uniformly in i, which in turn arises because $\hat{t}_i - \lfloor\tau_i T\rfloor \xrightarrow{a.s.} 0$ uniformly in i. Also, for $i > n$, $\Delta y_t D_t(\hat{t}_i) = \Delta y_{\hat{t}_i} = \Delta u_{\hat{t}_i}$ which is of $O_p(1)$. As a consequence, $\hat{\varepsilon}_t$ is identical to g_t apart from at a finite number $\bar{n}$ of indices i. At such points the difference between $\hat{\varepsilon}_t$ and g_t is bounded by an $o_p(T^{1/2})$ random variable. Since $\bar{n}$ is not increasing in T, it then follows that $\hat{\omega}_\varepsilon^2 - \hat{\omega}_\varepsilon^2(g) \xrightarrow{p} 0$, and therefore $\hat{\omega}_\varepsilon^2 \xrightarrow{p} \omega_\varepsilon^2$. This consistency result also holds if $n = 0$, since here $\hat{\varepsilon}_t$ is identical to Δu_t apart from at a finite number of points, $\bar{n}$, which are stochastically distributed on Ω_T; at these $\bar{n}$ points, Δu_t is simply replaced by zero, which has no impact on the asymptotic behaviour of $\hat{\omega}_\varepsilon^2$.

(b) Consider first a simple example, where we define the series $z_t := \Delta\eta_t$, with η_t as in Assumption LP, i.e. $IID(0, \sigma_\eta^2)$. Suppose we estimate the OLS regression

$$z_t = \hat{\boldsymbol{\lambda}}'\mathbf{z}_{t,k} + h_t \tag{A.6}$$

where $\mathbf{z}_{t,k} := [z_{t-1}, z_{t-2}, .., z_{t-k}]'$. Then, defining the $k \times 1$ vector $\mathbf{i} := [1, 1, ..., 1]'$ we construct

$$\begin{aligned}k^{-1}\mathbf{i}'\hat{\boldsymbol{\lambda}} &= k^{-1}\mathbf{i}'\left(T^{-1}\sum_{t=k+1}^{T}\mathbf{z}_{t,k}\mathbf{z}_{t,k}'\right)^{-1}T^{-1}\sum_{t=k+1}^{T}\mathbf{z}_{t,k}z_t \\ &= k^{-1}\mathbf{i}'\left[\sigma_\eta^2\mathbf{A} + \mathbf{o}_p(1)\right]^{-1}\left[\sigma_\eta^2\mathbf{b} + \mathbf{o}_p(1)\right]\end{aligned}$$

where $\mathbf{A}$ is the $k \times k$ matrix

$$\mathbf{A} := \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & ... & 0 \\ -1 & 2 & -1 & 0 & 0 & ... & 0 \\ 0 & -1 & 2 & -1 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & ... & 0 & 0 & -1 & 2 & -1 \\ 0 & ... & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

and $\mathbf{b}$ is the $k \times 1$ vector $[-1, 0, ..., 0]'$. Next, observe that

$$\begin{aligned}k^{-1}\mathbf{i}'\left[\sigma_\eta^2\mathbf{A} + \mathbf{o}_p(1)\right]^{-1}\left[\sigma_\eta^2\mathbf{b} + \mathbf{o}_p(1)\right] &= k^{-1}\mathbf{i}'\left[\sigma_\eta^{-2}\mathbf{A}^{-1} + \mathbf{o}_p(1)\right]\left[\sigma_\eta^2\mathbf{b} + \mathbf{o}_p(1)\right] \\ &= k^{-1}\mathbf{i}'\mathbf{A}^{-1}\mathbf{b} + k^{-1}\mathbf{i}'\mathbf{o}_p(1) \\ &= k^{-1}\mathbf{i}'\mathbf{A}^{-1}\mathbf{b} + o_p(1)\end{aligned}$$

where the inverse term in the first line follows from a matrix version of a Taylor series expansion of the form $(a + o)^{-1} = a^{-1} - a^{-2}o + a^{-3}o^2 - ....$

Now, it can be shown that the first column of $\mathbf{A}^{-1}$ is given by $[\frac{k}{k+1}, \frac{k-1}{k+1}, ..., \frac{1}{k+1}]'$ and, hence, $\mathbf{A}^{-1}\mathbf{b} = -[\frac{k}{k+1}, \frac{k-1}{k+1}, ..., \frac{1}{k+1}]'$, so that

$$k^{-1}\mathbf{i}'\mathbf{A}^{-1}\mathbf{b} = -\frac{1}{k(k+1)}\sum_{i=1}^{k} i = -\frac{1}{2}.$$

Consequently, $k^{-1}\mathbf{i}'\hat{\boldsymbol{\lambda}} = -\frac{1}{2} + o_p(1)$. Next,

$$\begin{aligned}\hat{\sigma}_h^2 := T^{-1}\sum_{t=k+1}^{T} h_t^2 &= T^{-1}\sum_{t=k+1}^{T} z_t^2 - T^{-1}\sum_{t=k+1}^{T} z_t\mathbf{z}_{t,k}'\left(T^{-1}\sum_{t=k+1}^{T}\mathbf{z}_{t,k}\mathbf{z}_{t,k}'\right)^{-1} T^{-1}\sum_{t=k+1}^{T}\mathbf{z}_{t,k}z_t \\ &= 2\sigma_\eta^2 - \sigma_\eta^2\mathbf{b}'\mathbf{A}^{-1}\mathbf{b} + o_p(1) \\ &= 2\sigma_\eta^2 - \sigma_\eta^2\frac{k}{k+1} + o_p(1).\end{aligned}$$

Then, given that $k \to \infty$, as $T \to \infty$ it follows that $T^{-1}\sum_{t=k+1}^{T} h_t^2 \xrightarrow{p} \sigma_\eta^2$. An estimate of the long run variance of z_t based on (A.6) is then given by

$$\hat{\omega}_z^2 := \frac{\hat{\sigma}_h^2}{\left(1 - \sum_{j=1}^{k}\hat{\lambda}_i\right)^2} = \frac{\sigma_\eta^2 + o_p(1)}{O_p(k^2)} = O_p(k^{-2}).$$

Note that an identical expression for $\hat{\omega}_z^2$ is obtained as

$$\hat{\omega}_z^2 := \frac{\hat{\sigma}_h^2}{\hat{\pi}^2}$$

where $\hat{\pi}$ is estimated from the OLS regression

$$\Delta z_t = \hat{\pi} z_{t-1} + \sum_{j=1}^{k-1}\psi_j \Delta z_{t-j} + h_t.$$

It follows that $\hat{\omega}_\varepsilon^2 = O_p(k^{-2})$ on noting that replacing the over-differenced noise $z_t = \Delta\eta_t$ with a more general $I(-1)$ process, such as $\Delta\varepsilon_t$ where ε_t is as in Assumption LP, will not affect the rate of convergence to zero. Similarly, using $\Delta\hat{\varepsilon}_t$ in place of z_t, where $\hat{\varepsilon}_t$ are the residuals from (6), will also leave this rate of convergence unchanged.

Proof of Theorem 3

(a) For $r \in \Lambda$, straightforward extensions of results in Perron and Vogelsang (1992) yield the result that $T^{-1/2}\hat{u}_{\lfloor rT \rfloor} \xrightarrow{w} \omega_\varepsilon H(r, d, c, \tilde{\boldsymbol{\tau}})$, where $H(r, d, c, \hat{\boldsymbol{\tau}})$ is a continuous time residual from the projection of $W_c(r)$ onto the space spanned by $\{1, 1(r > \tilde{\tau}_1), 1(r > \tilde{\tau}_2), ..., 1(r > \tilde{\tau}_{\bar{n}})\}$ with $\tilde{\tau}_i := \lim_{T\to\infty} T^{-1}\hat{t}_i$, $i = 1, 2, ..., \bar{n}$ and $\tilde{\boldsymbol{\tau}} = [\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_{\bar{n}}]$. Then, from (9) we obtain, using the CMT, that

$$T\hat{\pi} \xrightarrow{w} \frac{\sigma_\eta}{\omega_\varepsilon}\frac{\int_0^1 H(r, c, d, \tilde{\boldsymbol{\tau}})dW_c(r)}{\int_0^1 H(r, c, d, \tilde{\boldsymbol{\tau}})^2 dr}$$

and, since $\hat{\sigma}^2 \xrightarrow{p} \sigma_\eta^2$, we therefore obtain that

$$T^{-2}\hat{\omega}_u^2 := \frac{\hat{\sigma}^2}{(T\hat{\pi})^2} \xrightarrow{w} \omega_\varepsilon^2 \frac{\{\int_0^1 H(r,c,d,\tilde{\boldsymbol{\tau}})^2 dr\}^2}{\{\int_0^1 H(r,c,d,\tilde{\boldsymbol{\tau}}) dW_c(r)\}^2} =: \omega_\varepsilon^2 Q(c,d,\tilde{\boldsymbol{\tau}})$$

as required.

Now, from Theorem 2 (a) (i), we have that as $T \to \infty$, $\hat{t}_i - \lfloor \tau_i T \rfloor \xrightarrow{p} 0$, $i = 1, ..., n$, and $\bar{n} \geq n$. Thus if $n = \bar{n}$, $\tilde{\boldsymbol{\tau}} = [\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_n] = [\tau_1, \tau_2, ..., \tau_n] + o_p(T^{-1})$, confirming case (i).

To show case (ii), recognize that since $n = 0$, $\hat{t}_1 = \arg\max_{t \in \Omega_T} |\Delta u_t| - 1$. Now $\hat{t}_1$ is (marginally) distributed uniformly on Ω_T *whatever the distribution of* Δu_t *or its covariance structure* $E(\Delta u_t, \Delta u_{t-s})$, $s \neq 0$; this is because no position on Ω_T is any more likely to yield the maximum of $|\Delta u_t|$ than is any other.[7] Next, the distribution of $\hat{t}_2 = (\arg\max_{t \in \Omega_T - \Omega_{1,T}} |\Delta u_t|) - 1$ where $\Omega_{1,T} := [\hat{t}_1 - \lfloor dT \rfloor + 2, \hat{t}_1 + \lfloor dT \rfloor]$, while clearly dependent on $\hat{t}_1$, still does not depend on the distribution of Δu_t, nor on its covariance structure as $\max_{t \in \Omega_T} |\Delta u_t|$ and $\max_{t \in \Omega_T - \Omega_{1,T}} |\Delta u_t|$ are asymptotically independent under Assumption LP for ε_t, since they are separated by at least $\lfloor mT \rfloor$ observations.[8,9] Continuing on like this, we find that $\hat{t}_3 = (\arg\max_{t \in \Omega_T - \Omega_{1,T} - \Omega_{2,T}} |\Delta u_t|) - 1$ where $\Omega_{2,T} := [\hat{t}_2 - \lfloor dT \rfloor + 2, \hat{t}_2 + \lfloor dT \rfloor]$, is dependent on $\hat{t}_1$ and $\hat{t}_2$, but not on the distribution of Δu_t, nor on its covariance structure, and so on. We therefore find that $\tilde{\boldsymbol{\tau}} := [\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_{\bar{n}}]$ is a vector of $\bar{n}$ dependent random variables, whose length is also stochastic but satisfies (5), whose distribution is the same for all ε_t satisfying Assumption LP. Thus, $\tilde{\boldsymbol{\tau}}$ is a stochastic argument of H. Notice also that the distribution of $\tilde{\boldsymbol{\tau}}$ is independent of $W_c(r)$. This is because each of the (finite) $\bar{n}$ random variables $\max_{t \in \Omega_T} |\Delta u_t|$, $\max_{t \in \Omega_T - \Omega_{1,T}} |\Delta u_t|$, ... becomes independent of $T^{-1/2}(u_{\lfloor rT \rfloor} - u_1) = T^{-1/2} \sum_{t=2}^{\lfloor rT \rfloor} \Delta u_t$, as $T \to \infty$.

Finally, case (iii) is a hybrid of those in (i) and (ii) above. We have $[\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_n] = [\tau_1, \tau_2, ..., \tau_n] + o_p(T^{-1})$. So $\tilde{\boldsymbol{\tau}} = [\tau_1, \tau_2, ..., \tau_n, \tilde{\tau}_{n+1}, ..., \tilde{\tau}_{\bar{n}}]$ has n non-stochastic elements and $\bar{n} - n$ stochastic terms ($\bar{n}$ again being stochastic, but satisfying (5)). The stochastic terms, while dependent, have a distribution that does not depend on the distribution of Δu_t, nor on its covariance structure, and are independent of $W_c(r)$.

(b) As regards $\hat{\omega}_u^2$, the level breaks $\gamma_{i,T}^* = \omega_u T^{-1/2} \gamma_i$, $i = 1, 2, .., n$, have no asymptotic effect under Assumption I(0). Thus in (8) we are simply introducing a finite number, $\bar{n}$, of asymptotically irrelevant level break regressors. Furthermore, in (9) an additional $\bar{n}k$ one-time dummy variable regressors are introduced. Although $\bar{n}k \to \infty$ as $T \to \infty$, since $\bar{n}k = o(T^{1/2})$ the effect of these dummy variables is again asymptotically negligible. Hence $\hat{\omega}_u^2$ behaves asymptotically as if calculated directly from u_t, and therefore $\hat{\omega}_u^2 \xrightarrow{p} \omega_u^2$.

Proof of Lemma 1

(a) The result in (i) follows directly from Theorem 1 (a) and Theorem 2 (a) (ii). In (ii) the result follows from Theorem 1 (a) and Theorem 3 (a), along with application of the CMT.

(b) The result in (i) follows directly from Theorem 1 (b) and Theorem 2 (b). In (ii) the result follows from Theorem 1 (b) and Theorem 3 (b).

[7] This is not saying that $\max_{t \in \Omega_T} |\Delta u_t|$ is uniform, only its positioning on Ω_T.

[8] This holds regardless of whether $c = 0$ or $c > 0$.

[9] We might consider $\hat{t}_2$ as being *conditionally* (marginally) distributed across the interval $\Omega_T - \Omega_{1,T}$.

Proof of Lemma 2

(a) First note that $M^*_{t,\lfloor mT \rfloor}$ is invariant to β, so we may set β equal to zero without loss of generality in what follows. Then

$$T^{-1/2} M^*_{t,\lfloor mT \rfloor} = T^{-1/2} M_{t,\lfloor mT \rfloor} - (m/2) T^{1/2} \hat{\beta} \tag{A.7}$$

and so the limit of the first term in the right member of (A.7) is $\omega_\varepsilon \{L_1(r,m,c) + K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma})\}$, cf. Theorem 1 (a). As regards the second term in the right member of (A.7), straightforward but tedious algebra shows that

$$T^{1/2}\hat{\beta} \xrightarrow{w} \omega_\varepsilon \left\{ 12 \int_0^1 s W_c(s) ds - 6 \int_0^1 W_c(s) ds + 6 \sum_{i=1}^{n} \gamma_i \tau_i (1 - \tau_i) \right\}$$

and, hence, by the CMT

$$T^{-1/2} M^*_{t,\lfloor mT \rfloor} \xrightarrow{w} \omega_\varepsilon \{L_1(r,m,c) + K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma}) + L^*_1(m,c) + K^*(m,\boldsymbol{\tau},\boldsymbol{\gamma})\}.$$

The other results required to establish the results in (a), which concern the large sample behaviour of the long run variance estimators $\hat{\omega}^2_\varepsilon$ and $\hat{\omega}^2_u$, are all entirely routine generalizations of the corresponding results in the non-trend case.

(b) The results follow along similar lines to those in (a) and in Lemma 1.

References

Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821-856.

Berk, K.N. (1974). Consistent autoregressive spectral estimates. Annals of Statistics 2, 489–502.

Bunzel, H. and Vogelsang, T.J. (2005). Powerful trend function tests that are robust to strong serial correlation, with an application to the Prebisch-Singer hypothesis. Journal of Business and Economic Statistics 23, 381-94.

Burridge, P. and Taylor, A.M.R. (2006). Additive outlier detection via extreme-value theory. Journal of Time Series Analysis 27, 685-701.

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2009a). Simple, robust and powerful tests of the breaking trend hypothesis. Econometric Theory 25, 995-1029.

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2009b). Unit root testing in practice: dealing with uncertainty over the trend and initial condition (with commentaries and rejoinder). Econometric Theory 25, 587-667.

Keifer, N.M. and Vogelsang, T.J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory 21, 1130-1164.

Kejriwal, M. and Perron, P. (2009). A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component. Manuscript, Department of Economics, Boston University.

Kuan, C.-M. and Hornik, K. (1995). The generalized fluctuation test: a unifying view. Econometric Reviews 14, 135-161.

Leisch, F. Hornik, K. and Kuan, C.-M. (2000). Monitoring structural changes with the generalized fluctuation test. Econometric Theory 16, 835-854.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57, 1361-1401.

Perrron, P. (2006). Dealing with structural breaks. In Patterson, K. and Mills, T.C. (eds.), Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, Palgrave Macmillan, 278-352.

Perron, P. and Rodríguez, G. (2003). Searching for additive outliers in non-stationary time series. Journal of Time Series Analysis 24, 193-220.

Perron, P. and Vogelsang, T. (1992). Testing for a unit root in a time series with a changing mean: corrections and extensions. Journal of Business and Economic Statistics 10, 467-470.

Perron, P. and Yabu, T. (2009a). Estimating deterministic trends with an integrated or stationary noise component. Journal of Econometrics 151, 56-69.

Perron, P. and Yabu, T. (2009b). Testing for shifts in trend with an integrated or stationary noise component. Journal of Business and Economic Statistics 27, 369-396.

Perron, P. and Zhu, X. (2005). Structural breaks with deterministic and stochastic trends. Journal of Econometrics 129, 65-119.

Saygınsoy, O. and Vogelsang, T. (2010). Testing for a shift in trend at an unknown date: a fixed-b analysis of heteroskedasticity autocorrelation robust OLS based tests. Econometric Theory, forthcoming.

Stock, J.H. (1994). Unit roots, structural breaks and trends. In Engle, R.F. and McFadden, D.L. (eds.), Handbook of Econometrics 4, Elsevier Science, 2739–2840.

Stock, J.H. and Watson, M.W. (1996). Evidence on structural instability in macroeconomic time series relations. Journal of Business and Economic Statistics 14, 11-30.

Stock, J.H. and Watson, M.W. (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In Engle, R.F. and White, H. (eds.), Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W.J. Granger, Oxford University Press, 1–44.

Vogelsang, T.J. (1998). Trend function hypothesis testing in the presence of serial correlation. Econometrica 66, 123-148.

Table 1. Asymptotic critical values for nominal ξ-level S_1 and S_0 tests, and asymptotic κ_ξ values for U

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	Critical values			Critical values			Critical values			Critical values			Critical values		
	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ
$\xi = 0.10$	0.536	22.063	1.003	0.626	17.366	1.011	0.697	14.660	1.024	0.758	12.816	1.044	0.811	11.492	1.061
$\xi = 0.05$	0.579	23.400	1.003	0.679	18.550	1.018	0.763	15.683	1.031	0.833	13.764	1.055	0.893	12.392	1.076
$\xi = 0.01$	0.670	26.132	1.006	0.799	20.929	1.028	0.899	17.826	1.059	0.987	15.786	1.094	1.059	14.242	1.138

Note: The critical values for S_1 and S_0 are for the $I(1)$ and $I(0)$ cases, respectively.

Table 2. Asymptotic sizes of nominal ξ-level tests under $I(1)$ and $I(0)$ errors

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
Panel A. $\xi = 0.10$															
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U
$I(1)$, $c = 0$	0.100	0.004	0.100	0.100	0.013	0.100	0.100	0.025	0.100	0.100	0.038	0.100	0.100	0.051	0.100
$I(1)$, $c = 10$	0.018	0.005	0.022	0.009	0.014	0.020	0.004	0.024	0.023	0.002	0.034	0.027	0.001	0.039	0.028
$I(1)$, $c = 20$	0.002	0.005	0.006	0.000	0.012	0.011	0.000	0.018	0.015	0.000	0.024	0.018	0.000	0.030	0.019
$I(1)$, $c = 40$	0.000	0.005	0.005	0.000	0.011	0.010	0.000	0.015	0.012	0.000	0.021	0.013	0.000	0.024	0.013
$I(0)$	0.000	0.100	0.097	0.000	0.100	0.090	0.000	0.100	0.079	0.000	0.100	0.067	0.000	0.100	0.058
Panel B. $\xi = 0.05$															
	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U
$I(1)$, $c = 0$	0.050	0.003	0.050	0.050	0.008	0.050	0.050	0.016	0.050	0.050	0.028	0.050	0.050	0.036	0.050
$I(1)$, $c = 10$	0.007	0.003	0.009	0.003	0.008	0.009	0.001	0.015	0.012	0.000	0.022	0.016	0.000	0.026	0.017
$I(1)$, $c = 20$	0.000	0.003	0.003	0.000	0.007	0.006	0.000	0.011	0.008	0.000	0.015	0.009	0.000	0.017	0.009
$I(1)$, $c = 40$	0.000	0.003	0.003	0.000	0.005	0.004	0.000	0.007	0.005	0.000	0.010	0.006	0.000	0.011	0.005
$I(0)$	0.000	0.050	0.048	0.000	0.050	0.040	0.000	0.050	0.036	0.000	0.050	0.028	0.000	0.050	0.022
Panel C. $\xi = 0.01$															
	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U
$I(1)$, $c = 0$	0.010	0.001	0.010	0.010	0.004	0.010	0.010	0.008	0.010	0.010	0.013	0.010	0.010	0.018	0.010
$I(1)$, $c = 10$	0.001	0.001	0.001	0.000	0.003	0.002	0.000	0.006	0.003	0.000	0.008	0.004	0.000	0.011	0.004
$I(1)$, $c = 20$	0.000	0.001	0.001	0.000	0.002	0.002	0.000	0.003	0.002	0.000	0.004	0.002	0.000	0.006	0.001
$I(1)$, $c = 40$	0.000	0.001	0.001	0.000	0.001	0.001	0.000	0.002	0.001	0.000	0.002	0.001	0.000	0.002	0.001
$I(0)$	0.000	0.010	0.009	0.000	0.010	0.007	0.000	0.010	0.004	0.000	0.010	0.002	0.000	0.010	0.001

Note: The rejections for S_1 are computed using critical values for S_1 under $I(1)$, $c = 0$ errors; the rejections for S_0 are computed using critical values for S_0 under $I(0)$ errors.

Table 3. Asymptotic powers of nominal 0.05-level tests under $I(1)$ and $I(0)$ errors

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
Panel A. 1 break; $\tau_1 = 0.50$															
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U
$I(1)$, $c = 0$	0.500	0.068	0.501	0.342	0.092	0.344	0.265	0.102	0.269	0.220	0.164	0.255	0.195	0.220	0.265
$I(1)$, $c = 10$	0.478	0.200	0.507	0.270	0.265	0.370	0.152	0.308	0.332	0.089	0.415	0.386	0.050	0.496	0.440
$I(1)$, $c = 20$	0.467	0.409	0.579	0.215	0.529	0.543	0.082	0.621	0.597	0.028	0.737	0.693	0.007	0.807	0.754
$I(1)$, $c = 40$	0.451	0.860	0.868	0.121	0.941	0.934	0.016	0.975	0.969	0.001	0.992	0.987	0.000	0.997	0.994
$I(0)$	0.000	0.500	0.495	0.000	0.779	0.757	0.000	0.929	0.909	0.000	0.980	0.966	0.000	0.995	0.988
Panel B. 2 breaks; $\tau_1 = 0.33$, $\tau_2 = 0.67$															
	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U
$I(1)$, $c = 0$	0.736	0.105	0.738	0.553	0.229	0.571	0.431	0.089	0.408	0.363	0.146	0.350	0.324	0.205	0.333
$I(1)$, $c = 10$	0.727	0.286	0.748	0.462	0.494	0.622	0.281	0.320	0.420	0.165	0.445	0.441	0.097	0.554	0.496
$I(1)$, $c = 20$	0.712	0.535	0.785	0.373	0.757	0.777	0.155	0.677	0.664	0.052	0.795	0.753	0.015	0.869	0.823
$I(1)$, $c = 40$	0.694	0.941	0.955	0.218	0.989	0.987	0.025	0.991	0.988	0.002	0.998	0.996	0.000	0.999	0.999
$I(0)$	0.000	0.725	0.721	0.000	0.950	0.938	0.000	0.995	0.992	0.000	1.000	0.999	0.000	1.000	1.000
Panel C. 3 breaks; $\tau_1 = 0.25$, $\tau_2 = 0.50$, $\tau_3 = 0.75$															
	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)					
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U			
$I(1)$, $c = 0$	0.864	0.242	0.867	0.689	0.106	0.673	0.564	0.208	0.562	0.482	0.316	0.520			
$I(1)$, $c = 10$	0.866	0.496	0.890	0.619	0.343	0.663	0.399	0.527	0.617	0.245	0.664	0.654			
$I(1)$, $c = 20$	0.850	0.735	0.916	0.511	0.662	0.754	0.227	0.826	0.818	0.079	0.913	0.887			
$I(1)$, $c = 40$	0.835	0.982	0.989	0.315	0.987	0.985	0.044	0.997	0.996	0.002	0.999	0.999			
$I(0)$	0.000	0.858	0.854	0.000	0.988	0.984	0.000	1.000	0.999	0.000	1.000	1.000			
Panel D. 4 breaks; $\tau_1 = 0.20$, $\tau_2 = 0.40$, $\tau_3 = 0.60$, $\tau_4 = 0.80$															
	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)								
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U						
$I(1)$, $c = 0$	0.925	0.244	0.926	0.784	0.222	0.779	0.667	0.378	0.703						
$I(1)$, $c = 10$	0.932	0.509	0.943	0.729	0.507	0.792	0.498	0.698	0.774						
$I(1)$, $c = 20$	0.920	0.755	0.955	0.620	0.793	0.862	0.289	0.913	0.910						
$I(1)$, $c = 40$	0.902	0.987	0.994	0.387	0.995	0.994	0.055	1.000	0.999						
$I(0)$	0.000	0.923	0.920	0.000	0.998	0.997	0.000	1.000	1.000						

Table 4. Finite sample critical values for nominal ξ-level S_1 and S_0 tests, and finite sample κ_ξ values for U

Panel A. $T = 150$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	Critical values			Critical values			Critical values			Critical values			Critical values		
	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ
$\xi = 0.10$	0.569	21.745	1.015	0.659	16.677	1.044	0.730	13.876	1.071	0.774	12.454	1.090	0.828	11.050	1.117
$\xi = 0.05$	0.610	23.315	1.022	0.712	17.851	1.063	0.792	14.899	1.099	0.844	13.409	1.126	0.905	11.896	1.171
$\xi = 0.01$	0.699	26.600	1.038	0.821	20.256	1.141	0.913	16.897	1.225	0.979	15.302	1.279	1.057	13.647	1.375

Panel B. $T = 300$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	Critical values			Critical values			Critical values			Critical values			Critical values		
	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ
$\xi = 0.10$	0.562	20.919	1.012	0.642	16.781	1.024	0.716	14.043	1.047	0.771	12.429	1.069	0.827	11.061	1.098
$\xi = 0.05$	0.605	22.260	1.015	0.697	17.930	1.031	0.782	15.045	1.064	0.843	13.342	1.095	0.909	11.927	1.131
$\xi = 0.01$	0.693	25.034	1.028	0.806	20.159	1.070	0.913	17.035	1.132	0.997	15.092	1.193	1.076	13.566	1.277

Panel C. $T = 600$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	Critical values			Critical values			Critical values			Critical values			Critical values		
	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ
$\xi = 0.10$	0.551	21.225	1.006	0.639	16.801	1.020	0.710	14.220	1.035	0.770	12.489	1.057	0.821	11.239	1.075
$\xi = 0.05$	0.594	22.570	1.009	0.692	17.937	1.025	0.774	15.198	1.046	0.841	13.415	1.075	0.902	12.096	1.101
$\xi = 0.01$	0.680	25.178	1.013	0.805	20.284	1.048	0.902	17.255	1.092	0.988	15.243	1.162	1.063	13.790	1.212

Panel D. $T = 1200$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	Critical values			Critical values			Critical values			Critical values			Critical values		
	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ	S_1	S_0	κ_ξ
$\xi = 0.10$	0.545	21.540	1.004	0.635	17.076	1.015	0.705	14.447	1.030	0.764	12.666	1.051	0.816	11.364	1.072
$\xi = 0.05$	0.589	22.912	1.004	0.687	18.210	1.017	0.769	15.456	1.039	0.836	13.583	1.068	0.897	12.256	1.089
$\xi = 0.01$	0.675	25.706	1.007	0.795	20.568	1.034	0.895	17.557	1.075	0.983	15.435	1.128	1.061	13.972	1.181

Table 5. Finite sample sizes of nominal 0.05-level tests: normal and non-normal innovations

Panel A. $T = 150$							
		U					
		$m = 0.10$	$m = 0.15$	$m = 0.20$	$m = 0.25$	$m = 0.30$	KP
$N(0,1)$	$\rho = 1.0$	0.050	0.050	0.050	0.050	0.050	0.089
	$\rho = 0.0$	0.040	0.025	0.016	0.012	0.008	0.076
$t(5)$	$\rho = 1.0$	0.157	0.112	0.096	0.087	0.080	0.482
	$\rho = 0.0$	0.074	0.041	0.026	0.018	0.012	0.085
$\chi^2(3)$	$\rho = 1.0$	0.148	0.092	0.074	0.065	0.060	0.664
	$\rho = 0.0$	0.080	0.048	0.029	0.021	0.013	0.066
Panel B. $T = 300$							
		U					
		$m = 0.10$	$m = 0.15$	$m = 0.20$	$m = 0.25$	$m = 0.30$	KP
$N(0,1)$	$\rho = 1.0$	0.050	0.050	0.050	0.050	0.050	0.059
	$\rho = 0.0$	0.041	0.035	0.024	0.016	0.011	0.056
$t(5)$	$\rho = 1.0$	0.114	0.091	0.080	0.074	0.070	0.611
	$\rho = 0.0$	0.066	0.048	0.031	0.021	0.014	0.058
$\chi^2(3)$	$\rho = 1.0$	0.103	0.078	0.066	0.061	0.057	0.800
	$\rho = 0.0$	0.073	0.053	0.034	0.024	0.015	0.051
Panel C. $T = 600$							
		U					
		$m = 0.10$	$m = 0.15$	$m = 0.20$	$m = 0.25$	$m = 0.30$	KP
$N(0,1)$	$\rho = 1.0$	0.050	0.050	0.050	0.050	0.050	0.063
	$\rho = 0.0$	0.045	0.038	0.030	0.021	0.017	0.055
$t(5)$	$\rho = 1.0$	0.092	0.079	0.075	0.070	0.067	0.803
	$\rho = 0.0$	0.063	0.047	0.036	0.025	0.018	0.046
$\chi^2(3)$	$\rho = 1.0$	0.082	0.068	0.061	0.058	0.057	0.910
	$\rho = 0.0$	0.064	0.050	0.038	0.027	0.019	0.053
Panel D. $T = 1200$							
		U					
		$m = 0.10$	$m = 0.15$	$m = 0.20$	$m = 0.25$	$m = 0.30$	KP
$N(0,1)$	$\rho = 1.0$	0.050	0.050	0.050	0.050	0.050	0.056
	$\rho = 0.0$	0.048	0.041	0.032	0.024	0.018	0.052
$t(5)$	$\rho = 1.0$	0.075	0.071	0.066	0.062	0.064	0.936
	$\rho = 0.0$	0.060	0.046	0.034	0.024	0.019	0.047
$\chi^2(3)$	$\rho = 1.0$	0.069	0.063	0.058	0.056	0.056	0.975
	$\rho = 0.0$	0.059	0.048	0.037	0.026	0.021	0.031

Table 6. Finite sample sizes of nominal 0.05-level U tests: ARMA errors

Panel A. $T = 150$						
ρ	θ	$m = 0.10$	$m = 0.15$	$m = 0.20$	$m = 0.25$	$m = 0.30$
1.00	0.00	0.050	0.050	0.050	0.050	0.050
	0.50	0.048	0.084	0.095	0.095	0.096
	−0.50	0.062	0.078	0.080	0.079	0.081
0.95	0.00	0.032	0.033	0.032	0.031	0.028
	0.50	0.060	0.100	0.109	0.106	0.097
	−0.50	0.077	0.088	0.082	0.070	0.063
0.90	0.00	0.024	0.024	0.021	0.019	0.016
	0.50	0.060	0.099	0.106	0.100	0.087
	−0.50	0.083	0.083	0.065	0.055	0.044
0.70	0.00	0.019	0.012	0.010	0.007	0.005
	0.50	0.052	0.058	0.050	0.043	0.032
	−0.50	0.090	0.059	0.039	0.028	0.020
0.00	0.00	0.040	0.025	0.016	0.012	0.008
	0.50	0.131	0.061	0.031	0.020	0.010
	−0.50	0.064	0.037	0.023	0.016	0.010
Panel B. $T = 300$						
ρ	θ	$m = 0.10$	$m = 0.15$	$m = 0.20$	$m = 0.25$	$m = 0.30$
1.00	0.00	0.050	0.050	0.050	0.050	0.050
	0.50	0.031	0.044	0.059	0.066	0.072
	−0.50	0.060	0.063	0.066	0.064	0.064
0.95	0.00	0.013	0.014	0.016	0.015	0.013
	0.50	0.057	0.059	0.063	0.059	0.053
	−0.50	0.043	0.037	0.035	0.029	0.025
0.90	0.00	0.010	0.010	0.009	0.008	0.007
	0.50	0.083	0.079	0.072	0.059	0.048
	−0.50	0.040	0.031	0.025	0.020	0.015
0.70	0.00	0.011	0.009	0.008	0.006	0.004
	0.50	0.120	0.116	0.093	0.074	0.055
	−0.50	0.048	0.032	0.022	0.016	0.011
0.00	0.00	0.041	0.035	0.024	0.016	0.011
	0.50	0.094	0.052	0.027	0.015	0.008
	−0.50	0.060	0.046	0.032	0.022	0.015

Table 7. Finite sample powers of nominal 0.05-level tests: $\gamma = 10$, $T = 300$

Panel A. 1 break; $\tau_1 = 0.50$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.530	0.125	0.520	0.370	0.134	0.364	0.281	0.143	0.268	0.242	0.202	0.254	0.210	0.261	0.257	1.000
$\rho = 0.95$	0.526	0.382	0.581	0.299	0.442	0.481	0.157	0.504	0.465	0.088	0.612	0.535	0.040	0.705	0.596	1.000
$\rho = 0.90$	0.549	0.723	0.762	0.267	0.816	0.800	0.095	0.893	0.859	0.033	0.945	0.911	0.009	0.975	0.946	1.000
$\rho = 0.70$	0.789	1.000	1.000	0.459	1.000	1.000	0.251	1.000	1.000	0.166	1.000	1.000	0.106	1.000	1.000	1.000
$\rho = 0.50$	0.977	1.000	1.000	0.902	1.000	1.000	0.800	1.000	1.000	0.695	1.000	1.000	0.595	1.000	1.000	1.000

Panel B. 2 breaks; $\tau_1 = 0.33$, $\tau_2 = 0.67$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.747	0.170	0.733	0.569	0.293	0.579	0.426	0.111	0.362	0.372	0.173	0.319	0.324	0.243	0.295	1.000
$\rho = 0.95$	0.753	0.501	0.784	0.487	0.687	0.726	0.248	0.557	0.533	0.140	0.679	0.598	0.064	0.775	0.673	1.000
$\rho = 0.90$	0.768	0.846	0.901	0.431	0.951	0.944	0.142	0.945	0.921	0.051	0.975	0.956	0.014	0.990	0.975	1.000
$\rho = 0.70$	0.912	1.000	1.000	0.590	1.000	1.000	0.350	1.000	1.000	0.234	1.000	1.000	0.151	1.000	1.000	1.000
$\rho = 0.50$	0.995	1.000	1.000	0.945	1.000	1.000	0.915	1.000	1.000	0.832	1.000	1.000	0.729	1.000	1.000	1.000

Panel C. 3 breaks; $\tau_1 = 0.25$, $\tau_2 = 0.50$, $\tau_3 = 0.75$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.851	0.353	0.849	0.680	0.138	0.646	0.559	0.263	0.522	0.492	0.364	0.493	*0.189*	*0.125*	*0.146*	1.000
$\rho = 0.95$	0.855	0.719	0.901	0.588	0.558	0.700	0.354	0.746	0.722	0.209	0.845	0.786	*0.009*	*0.121*	*0.074*	1.000
$\rho = 0.90$	0.858	0.946	0.968	0.504	0.932	0.928	0.199	0.981	0.970	0.074	0.994	0.986	*0.001*	*0.069*	*0.032*	1.000
$\rho = 0.70$	0.947	1.000	1.000	0.646	1.000	1.000	0.380	1.000	1.000	0.254	1.000	1.000	*0.001*	*0.030*	*0.007*	1.000
$\rho = 0.50$	0.998	1.000	1.000	0.975	1.000	1.000	0.921	1.000	1.000	0.838	1.000	1.000	*0.008*	*0.034*	*0.008*	1.000

Panel D. 4 breaks; $\tau_1 = 0.20$, $\tau_2 = 0.40$, $\tau_3 = 0.60$, $\tau_4 = 0.80$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.908	0.343	0.903	0.778	0.269	0.758	0.660	0.446	0.672	*0.180*	*0.098*	*0.141*	*0.148*	*0.085*	*0.098*	1.000
$\rho = 0.95$	0.910	0.741	0.936	0.701	0.720	0.828	0.442	0.871	0.855	*0.009*	*0.076*	*0.049*	*0.002*	*0.048*	*0.024*	1.000
$\rho = 0.90$	0.905	0.959	0.981	0.603	0.970	0.969	0.261	0.994	0.990	*0.000*	*0.030*	*0.014*	*0.000*	*0.016*	*0.006*	1.000
$\rho = 0.70$	0.967	1.000	1.000	0.689	1.000	1.000	0.402	1.000	1.000	*0.000*	*0.001*	*0.000*	*0.000*	*0.000*	*0.000*	1.000
$\rho = 0.50$	0.999	1.000	1.000	0.977	1.000	1.000	0.923	1.000	1.000	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	1.000

Table 8. Finite sample powers of nominal 0.05-level tests: $\gamma = 5$, $T = 300$

Panel A. 1 break; $\tau_1 = 0.50$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.104	0.021	0.102	0.086	0.035	0.087	0.078	0.046	0.074	0.076	0.084	0.087	0.072	0.133	0.106	0.835
$\rho = 0.95$	0.033	0.054	0.066	0.010	0.081	0.072	0.003	0.111	0.083	0.001	0.175	0.120	0.000	0.240	0.154	0.775
$\rho = 0.90$	0.010	0.127	0.118	0.001	0.201	0.176	0.000	0.289	0.234	0.000	0.401	0.304	0.000	0.510	0.375	0.783
$\rho = 0.70$	0.014	0.846	0.834	0.003	0.948	0.938	0.000	0.977	0.968	0.000	0.979	0.970	0.000	0.975	0.967	1.000
$\rho = 0.50$	0.073	0.995	0.994	0.024	0.992	0.991	0.005	0.988	0.984	0.001	0.978	0.971	0.000	0.970	0.959	1.000

Panel B. 2 breaks; $\tau_1 = 0.33$, $\tau_2 = 0.67$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.151	0.027	0.146	0.121	0.079	0.141	0.097	0.028	0.076	0.095	0.054	0.078	0.090	0.090	0.087	0.968
$\rho = 0.95$	0.058	0.084	0.105	0.019	0.188	0.171	0.004	0.120	0.089	0.002	0.200	0.137	0.001	0.291	0.188	0.951
$\rho = 0.90$	0.023	0.197	0.188	0.003	0.391	0.352	0.000	0.383	0.307	0.000	0.526	0.408	0.000	0.654	0.500	0.945
$\rho = 0.70$	0.029	0.946	0.940	0.007	0.972	0.968	0.001	0.985	0.976	0.000	0.984	0.977	0.000	0.985	0.978	0.991
$\rho = 0.50$	0.148	0.993	0.992	0.057	0.984	0.981	0.026	0.988	0.985	0.006	0.977	0.971	0.001	0.975	0.966	0.998

Panel C. 3 breaks; $\tau_1 = 0.25$, $\tau_2 = 0.50$, $\tau_3 = 0.75$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.188	0.073	0.200	0.138	0.029	0.124	0.128	0.079	0.125	0.124	0.144	0.150	*0.085*	*0.090*	*0.083*	0.991
$\rho = 0.95$	0.073	0.189	0.204	0.022	0.110	0.103	0.007	0.254	0.200	0.002	0.382	0.285	*0.000*	*0.172*	*0.113*	0.985
$\rho = 0.90$	0.028	0.369	0.354	0.003	0.340	0.304	0.000	0.569	0.480	0.000	0.717	0.599	*0.000*	*0.194*	*0.117*	0.978
$\rho = 0.70$	0.042	0.971	0.969	0.011	0.987	0.984	0.002	0.992	0.989	0.000	0.994	0.991	*0.000*	*0.384*	*0.193*	0.982
$\rho = 0.50$	0.227	0.991	0.990	0.133	0.993	0.992	0.033	0.991	0.988	0.007	0.989	0.987	*0.000*	*0.367*	*0.198*	0.990

Panel D. 4 breaks; $\tau_1 = 0.20$, $\tau_2 = 0.40$, $\tau_3 = 0.60$, $\tau_4 = 0.80$

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)			
	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	KP
$\rho = 1.00$	0.214	0.073	0.222	0.172	0.068	0.175	0.160	0.175	0.210	*0.095*	*0.084*	*0.093*	*0.090*	*0.083*	*0.080*	0.996
$\rho = 0.95$	0.091	0.210	0.233	0.030	0.221	0.202	0.009	0.427	0.355	*0.001*	*0.149*	*0.105*	*0.000*	*0.138*	*0.086*	0.994
$\rho = 0.90$	0.035	0.433	0.414	0.004	0.504	0.459	0.000	0.737	0.658	*0.000*	*0.152*	*0.100*	*0.000*	*0.125*	*0.072*	0.987
$\rho = 0.70$	0.059	0.979	0.977	0.015	0.993	0.991	0.002	0.994	0.993	*0.000*	*0.165*	*0.080*	*0.000*	*0.113*	*0.043*	0.984
$\rho = 0.50$	0.320	0.993	0.992	0.146	0.993	0.993	0.036	0.985	0.984	*0.000*	*0.141*	*0.053*	*0.000*	*0.094*	*0.035*	0.982

Table 9. Number of breaks detected at the nominal 0.05-level: 3 breaks; $\tau_1 = 0.25$, $\tau_2 = 0.50$, $\tau_3 = 0.75$; $T = 300$

			$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)		
ρ	γ		S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U
1.00	5	# = 1:	0.175	0.060	0.179	0.132	0.025	0.118	0.124	0.065	0.114	0.120	0.117	0.132
		# = 2:	0.013	0.011	0.018	0.006	0.003	0.006	0.004	0.013	0.010	0.004	0.025	0.017
		# = 3:	0.000	0.002	0.002	0.000	0.000	0.000	0.000	0.002	0.001	0.000	0.002	0.001
		# > 3:	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	10	# = 1:	0.366	0.181	0.359	0.432	0.088	0.422	0.409	0.151	0.366	0.390	0.198	0.337
		# = 2:	0.356	0.113	0.347	0.211	0.037	0.189	0.135	0.082	0.129	0.094	0.122	0.123
		# = 3:	0.127	0.057	0.140	0.037	0.013	0.035	0.015	0.031	0.028	0.007	0.045	0.033
		# > 3:	0.001	0.003	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	15	# = 1:	0.017	0.153	0.018	0.116	0.127	0.135	0.242	0.165	0.265	0.327	0.181	0.320
		# = 2:	0.181	0.188	0.181	0.380	0.103	0.388	0.430	0.154	0.390	0.408	0.196	0.336
		# = 3:	0.794	0.328	0.789	0.489	0.098	0.457	0.274	0.163	0.267	0.166	0.209	0.213
		# > 3:	0.007	0.006	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.95	5	# = 1:	0.071	0.145	0.163	0.021	0.093	0.088	0.007	0.196	0.161	0.002	0.277	0.223
		# = 2:	0.003	0.037	0.034	0.001	0.016	0.013	0.000	0.051	0.035	0.000	0.095	0.057
		# = 3:	0.000	0.007	0.006	0.000	0.002	0.001	0.000	0.007	0.004	0.000	0.010	0.005
		# > 3:	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	10	# = 1:	0.367	0.212	0.267	0.412	0.218	0.334	0.301	0.221	0.258	0.193	0.196	0.233
		# = 2:	0.352	0.242	0.342	0.152	0.191	0.229	0.049	0.250	0.241	0.015	0.309	0.286
		# = 3:	0.136	0.258	0.286	0.024	0.149	0.138	0.004	0.275	0.223	0.000	0.341	0.268
		# > 3:	0.000	0.007	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	15	# = 1:	0.004	0.054	0.003	0.051	0.109	0.046	0.173	0.068	0.077	0.299	0.042	0.070
		# = 2:	0.085	0.124	0.061	0.282	0.170	0.214	0.413	0.140	0.192	0.424	0.126	0.159
		# = 3:	0.911	0.762	0.928	0.663	0.588	0.735	0.390	0.738	0.715	0.208	0.805	0.751
		# > 3:	0.000	0.009	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.90	5	# = 1:	0.027	0.246	0.240	0.003	0.239	0.221	0.000	0.317	0.293	0.000	0.346	0.336
		# = 2:	0.001	0.096	0.090	0.000	0.085	0.070	0.000	0.191	0.149	0.000	0.284	0.211
		# = 3:	0.000	0.026	0.023	0.000	0.017	0.013	0.000	0.061	0.039	0.000	0.087	0.052
		# > 3:	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	10	# = 1:	0.342	0.099	0.108	0.353	0.112	0.133	0.168	0.049	0.066	0.066	0.024	0.040
		# = 2:	0.347	0.213	0.235	0.126	0.214	0.227	0.027	0.139	0.167	0.006	0.130	0.163
		# = 3:	0.169	0.626	0.617	0.025	0.605	0.568	0.003	0.794	0.737	0.001	0.841	0.784
		# > 3:	0.000	0.009	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	15	# = 1:	0.001	0.004	0.000	0.018	0.006	0.004	0.102	0.001	0.002	0.220	0.000	0.001
		# = 2:	0.031	0.019	0.009	0.151	0.025	0.027	0.310	0.008	0.012	0.367	0.007	0.010
		# = 3:	0.968	0.967	0.982	0.830	0.966	0.969	0.570	0.991	0.986	0.346	0.992	0.989
		# > 3:	0.000	0.009	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 9. Continued

			$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)		
ρ	γ		S_1	S_0	U	S_1	S_0	U	S_1	S_0	U	S_1	S_0	U
0.70	5	# = 1:	0.032	0.061	0.066	0.009	0.026	0.031	0.002	0.007	0.009	0.000	0.005	0.006
		# = 2:	0.008	0.207	0.218	0.002	0.097	0.112	0.000	0.024	0.035	0.000	0.082	0.087
		# = 3:	0.002	0.695	0.678	0.000	0.865	0.841	0.000	0.961	0.945	0.000	0.907	0.898
		# > 3:	0.000	0.008	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	10	# = 1:	0.134	0.000	0.000	0.171	0.000	0.000	0.103	0.000	0.000	0.071	0.000	0.000
		# = 2:	0.232	0.000	0.000	0.141	0.000	0.000	0.084	0.000	0.000	0.056	0.002	0.002
		# = 3:	0.581	0.991	0.992	0.334	1.000	1.000	0.194	1.000	1.000	0.128	0.998	0.998
		# > 3:	0.000	0.009	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	15	# = 1:	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.027	0.000	0.000
		# = 2:	0.000	0.000	0.000	0.001	0.000	0.000	0.018	0.000	0.000	0.063	0.000	0.000
		# = 3:	1.000	0.991	0.992	0.999	1.000	1.000	0.978	1.000	1.000	0.900	1.000	1.000
		# > 3:	0.000	0.009	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.50	5	# = 1:	0.116	0.008	0.009	0.078	0.004	0.005	0.024	0.003	0.004	0.006	0.002	0.003
		# = 2:	0.069	0.019	0.020	0.039	0.009	0.010	0.007	0.005	0.006	0.001	0.035	0.037
		# = 3:	0.042	0.955	0.953	0.015	0.981	0.978	0.001	0.983	0.978	0.000	0.951	0.947
		# > 3:	0.000	0.009	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	10	# = 1:	0.007	0.000	0.000	0.017	0.000	0.000	0.037	0.000	0.000	0.050	0.000	0.000
		# = 2:	0.024	0.000	0.000	0.028	0.000	0.000	0.056	0.000	0.000	0.055	0.000	0.000
		# = 3:	0.967	0.990	0.991	0.931	1.000	1.000	0.828	1.000	1.000	0.734	1.000	1.000
		# > 3:	0.000	0.010	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	15	# = 1:	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		# = 2:	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		# = 3:	1.000	0.990	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		# > 3:	0.000	0.010	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 10. Asymptotic critical values for nominal ξ-level S_1^* and S_0^* tests, and asymptotic κ_ξ^* values for U^*; trend case

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	Critical values			Critical values			Critical values			Critical values			Critical values		
	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*
$\xi = 0.10$	0.513	22.048	1.004	0.585	17.354	1.015	0.636	14.619	1.033	0.677	12.753	1.054	0.710	11.380	1.081
$\xi = 0.05$	0.554	23.394	1.004	0.637	18.526	1.018	0.698	15.629	1.041	0.744	13.698	1.069	0.787	12.247	1.097
$\xi = 0.01$	0.638	26.135	1.006	0.743	20.852	1.031	0.824	17.785	1.068	0.887	15.648	1.119	0.935	14.085	1.173

Table 11. Finite sample critical values for nominal ξ-level S_1^* and S_0^* tests, and finite sample κ_ξ^* values for U^*; trend case

	$m = 0.10$ ($n_{\max} = 8$)			$m = 0.15$ ($n_{\max} = 5$)			$m = 0.20$ ($n_{\max} = 4$)			$m = 0.25$ ($n_{\max} = 3$)			$m = 0.30$ ($n_{\max} = 3$)		
	Critical values			Critical values			Critical values			Critical values			Critical values		
	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*	S_1^*	S_0^*	κ_ξ^*
Panel A. $T = 150$															
$\xi = 0.10$	0.552	22.123	1.023	0.625	16.934	1.056	0.675	14.082	1.088	0.703	12.599	1.112	0.736	11.130	1.151
$\xi = 0.05$	0.593	23.767	1.029	0.676	18.172	1.078	0.734	15.134	1.123	0.768	13.586	1.156	0.807	12.032	1.201
$\xi = 0.01$	0.678	27.337	1.065	0.780	20.712	1.170	0.852	17.250	1.244	0.895	15.632	1.308	0.948	13.928	1.404
Panel B. $T = 300$															
$\xi = 0.10$	0.540	21.067	1.016	0.604	16.896	1.033	0.659	14.112	1.059	0.694	12.455	1.086	0.727	11.045	1.121
$\xi = 0.05$	0.581	22.426	1.018	0.654	18.085	1.042	0.719	15.131	1.079	0.762	13.381	1.112	0.803	11.909	1.150
$\xi = 0.01$	0.664	25.216	1.040	0.755	20.324	1.085	0.834	17.187	1.161	0.893	15.204	1.238	0.946	13.602	1.301
Panel C. $T = 600$															
$\xi = 0.10$	0.528	21.288	1.008	0.599	16.836	1.026	0.651	14.234	1.046	0.691	12.451	1.074	0.722	11.166	1.101
$\xi = 0.05$	0.570	22.629	1.011	0.651	17.959	1.034	0.711	15.239	1.061	0.758	13.386	1.093	0.795	12.029	1.131
$\xi = 0.01$	0.655	25.226	1.023	0.754	20.332	1.067	0.832	17.249	1.112	0.892	15.210	1.180	0.942	13.721	1.246
Panel D. $T = 1200$															
$\xi = 0.10$	0.522	21.559	1.005	0.593	17.070	1.020	0.645	14.423	1.038	0.684	12.606	1.062	0.715	11.262	1.086
$\xi = 0.05$	0.563	22.948	1.006	0.644	18.230	1.024	0.703	15.451	1.050	0.751	13.529	1.080	0.789	12.128	1.108
$\xi = 0.01$	0.646	25.712	1.012	0.744	20.629	1.042	0.822	17.556	1.090	0.885	15.461	1.151	0.937	13.881	1.209