Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Econometrics 123 (2004) 33–66
JOURNAL OF Econometrics
www.elsevier.com/locate/econbase

Tests of stationarity against a change in persistence

Fabio Busetti[a], A.M.Robert Taylor[b,*]

[a] *Banca d'Italia, Research Department, Via Nazionale 91, 00184 Rome, Italy*
[b] *Department of Economics, University of Birmingham, Edgbaston Park Road, Birmingham B15 2TT, UK*

Received 12 October 2001; accepted 22 October 2003

Abstract

This paper considers testing against a change in the order of integration of a time series, either from $I(0)$ to $I(1)$ or from $I(1)$ to $I(0)$, at some known or unknown point in the sample. The null hypothesis is that the series is stochastically stationary around a deterministic trend function. For the case of a known change-point the locally best invariant (LBI) tests against the above changes in the order of integration are derived under the assumption of Gaussianity. When the change-point is not known we construct our tests taking functions of the LBI statistics over all possible break-dates. Sub-sample implementations of existing stationarity tests are also considered. We demonstrate by a series of simulation experiments that, for a given direction of change, the LBI-based approach can deliver considerably more powerful tests than both the sub-sample stationarity tests and the ratio-based tests of Kim et al. (J. Econom. 109 (2002) 389) and Busetti and Taylor (Tests of stationarity against a change in persistence, University of Birmingham, Department of Economics, Discussion Paper 01-13, 2001). Moreover, the power losses from an unknown breakpoint do not appear to be large. We also find that standard stationarity tests have good power against both changes from $I(0)$ to $I(1)$ and vice versa, while the ratio-based tests are consistent only against a known direction of change. A further test constructed in terms of the LBI-based statistics for the two possible directions of change is shown to perform generally better than the standard stationarity tests when the direction of change under the alternative is not known. Finally, we apply the tests discussed in the paper to the US inflation rate and find evidence for a change in persistence from $I(1)$ to $I(0)$ behaviour although, significantly, the timing of this change varies according to whether or not a simultaneous change in the level of the series is allowed.
© 2003 Elsevier B.V. All rights reserved.

JEL classification: C22

Keywords: Persistence changes; LBI tests; Unknown direction of change; Trend breaks

* Corresponding author. Tel.: +44-121-414-6640; fax: +44-212-414-7377.
E-mail addresses: busetti.fabio@insedia.interbusiness.it (F. Busetti), r.taylor@bham.ac.uk (A.M.R. Taylor).

0304-4076/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2003.10.028

1. Introduction

In a recent paper, Kim (2000) develops residual-based ratio tests against changes in persistence in a time series, focusing on the case of a shift from stochastic stationarity, $I(0)$, to difference stationarity, $I(1)$, at some point in the sample. Kim (2000) also discusses the possibility of $I(1)$ to $I(0)$ shifts but does not provide tests against such alternatives. Like the stationarity tests of, *inter alia*, Nyblom and Mäkeläinen (1983) [NM] and Kwiatkowski et al. (1992) [KPSS], Kim (2000) tests the null that the series is $I(0)$ throughout its history, but differs from the latter which are designed as tests against a constant $I(1)$ alternative. Complementary tests have been proposed in Banerjee et al. (1992) [BLS] and Leybourne et al. (2000) [LKSN] who use Dickey-Fuller-type statistics to test the constant $I(1)$ null against the alternative of either a change from $I(0)$ to $I(1)$ or from $I(1)$ to $I(0)$.

Changes of this kind in macroeconomic variables are well documented; see the literature reviews in Kim (2000) and LKSN. The range of series for which such phenomena have been observed includes inflation rates, real output and short-term interest rates. The development of tests against such behaviour is therefore of considerable practical relevance. Although the ratio-based tests of Kim (2000) are inconsistent, simple modifications proposed independently in Busetti and Taylor (2001) [BT] and Kim et al. (2002), can provide consistent tests and breakpoint estimation under $I(0)$ to $I(1)$ changes. In this paper we propose new ratio-based tests and breakpoint estimators which are consistent under $I(1)$ to $I(0)$ changes, and demonstrate that the ratio-based tests which are consistent against changes from $I(1)$ to $I(0)$ are not consistent against changes from $I(0)$ to $I(1)$, and vice versa, with neither consistent against constant $I(1)$ processes. Consequently, acceptance of the null by the ratio-based tests cannot be taken to imply that the process is stationarity, even in large samples.

Assuming a known breakpoint and Gaussian innovations, we derive locally best invariant (LBI) tests of the constant $I(0)$ null against: (i) a shift from $I(0)$ to $I(1)$, and (ii) from $I(1)$ to $I(0)$. These tests are subsequently modified to allow for weakly dependent innovations and for an unknown breakpoint and are shown to contain the standard NM/KPSS tests as special cases. The consistency of the LBI-based tests is demonstrated and simulation evidence shows them to be considerably more powerful than the ratio-based tests. We show that LBI-based tests against $I(0)$ to $I(1)$ changes are also consistent against $I(1)$ to $I(0)$ changes, and vice versa, and against constant $I(1)$ processes. They therefore constitute useful Portmanteau statistics against any $I(1)$ behaviour in the process's history.

In practice the direction of the change may be unknown. In an attempt to maximise power for this situation, we also propose tests based on the pairwise maxima of the LBI-based and ratio-based statistics for $I(0)$ to $I(1)$ changes and vice versa. The NM/KPSS test is also found to display very good power properties when the change occurs near the beginning (end) of the sample when the switch is in the $I(0)$-$I(1)$ ($I(1)$-$I(0)$ respectively) direction, but appears to be dominated by our pairwise maximum LBI-based tests otherwise.

The plan of the paper is as follows. In Section 2 we discuss tests for the constant $I(0)$ null against the alternative of an $I(0)$ to $I(1)$ change. We first review the ratio-based

class of tests. We then develop the exact LBI tests under a known breakpoint. Modifications for an unknown breakpoint are also proposed. We also discuss sub-sample implementations of the NM/KPSS tests. In Section 3 we repeat the analysis of Section 2 for alternatives of a change from $I(1)$ to $I(0)$, while in Section 4 we discuss cases where the direction of change is not assumed known to the investigator. In Section 5 we use Monte Carlo methods to compare the finite sample size and power properties of the tests discussed in Sections 2–4 against processes which display either shifts from $I(0)$ to $I(1)$ or from $I(1)$ to $I(0)$. In Section 6 we provide some generalisations to allow for linear deterministic trends, deterministic structural breaks, and serially correlated innovations. Section 7 applies the statistics discussed in this paper to the US inflation rate. Section 8 concludes. Proofs are contained in an Appendix.

2. Tests against changes from $I(0)$ to $I(1)$

In this section we will focus attention on the Gaussian unobserved components model,

$$y_t = d_t + \mu_t + \varepsilon_t, \quad t = 1, \ldots, T, \tag{2.1}$$

$$\mu_t = \mu_{t-1} + 1(t > [\tau_0 T])\eta_t, \quad \tau_0 \in (0, 1), \tag{2.2}$$

where $1(\cdot)$ is the indicator function, and ε_t and η_t are mutually independent mean zero IID Gaussian processes with variances σ^2 and $\sigma_\eta^2\sigma^2$, respectively. Although Gaussianity has been assumed, the limiting results which follow hold under the weaker, martingale difference, conditions on $\{\varepsilon_t, \eta_t\}$ of Stock (1994, p. 2745). For the present, the deterministic component d_t is taken to be a constant; *viz*, $d_t = \beta_0$. Generalisations are given in Section 6.1. We may set μ_0 equal to zero without loss of generality. It can be seen that the data generating process [DGP] (2.1)–(2.2) yields a process which is stationary up to and including time $[\tau_0 T]$ but is $I(1)$ after the break, if and only if $\sigma_\eta^2 > 0$.

Consequently, a test for stationarity against a shift in persistence from stationarity to a unit root in the context of (2.1)–(2.2) can be framed in testing the null hypothesis

$$\mathsf{H}_0 : \sigma_\eta^2 = 0 \tag{2.3}$$

against the (fixed) alternative hypothesis

$$\mathsf{H}_1 : \sigma_\eta^2 > 0. \tag{2.4}$$

Following Tanaka (1996, p. 368) and Stock (1994, p. 2799), *inter alia*, it will also prove useful to consider the local alternative hypothesis,

$$\mathsf{H}_c : \sigma_\eta^2 = c^2/T^2, \quad c \geqslant 0. \tag{2.5}$$

In what follows we will derive representations for the limiting distributions of the statistics discussed in this paper under H_c. These representations are useful in that they can be used to delimit the asymptotic local power functions of the tests. Moreover, since H_c reduces to H_0 for $c = 0$, these representations will reduce to the limiting null distributions of the statistics on setting $c = 0$ in the expressions throughout.

2.1. Ratio tests

BT and Kim et al. (2002) have independently proposed the test which rejects H_0 for large values of the ratio statistic [1]

$$\mathscr{K}_M(\tau) = \frac{[(1-\tau)T]^{-2} \sum_{t=[\tau T]+1}^{T} (\sum_{i=[\tau T]+1}^{t} \hat{\varepsilon}_{1,i})^2}{[\tau T]^{-2} \sum_{t=1}^{[\tau T]} (\sum_{i=1}^{t} \hat{\varepsilon}_{0,i})^2}, \tag{2.6}$$

where $\hat{\varepsilon}_{0,t}$ are the OLS residuals from the regression of y_t on an intercept, $t = 1, \ldots [\tau T]$, and $\hat{\varepsilon}_{1,t}$ are the OLS residuals from the regression of y_t on an intercept, $t = [\tau T] + 1, \ldots, T$. Where the true breakpoint, τ_0, is known, $\mathscr{K}_M(\tau)$ of (2.6) is evaluated at $\tau = \tau_0$. Where τ_0 is unknown, they suggest computing the statistic $\mathscr{K}_M(\tau)$ of (2.6) for each value of $\tau \in \mathscr{T}$, where $\mathscr{T}$ is a given sub-interval of $[0,1]$ and taking an appropriate function of the resulting sequence of statistics. These authors investigate three such functions of the sequence $\{\mathscr{K}_M(\tau), \tau \in \mathscr{T}\}$. Firstly, after Andrews (1993), the maximum over the sequence of statistics, *viz.*,

$$H_1(\mathscr{K}_M(\cdot)) \equiv \max_{\tau \in \mathscr{T}} \mathscr{K}_M(\tau). \tag{2.7}$$

Secondly, Hansen's (1991) mean score statistic,

$$H_2(\mathscr{K}_M(\cdot)) \equiv \int_{\tau \in \mathscr{T}} \mathscr{K}_M(\tau)\, \mathrm{d}\tau. \tag{2.8}$$

Finally, after Andrews and Ploberger (1994), the mean-exponential statistic

$$H_3(\mathscr{K}_M(\cdot)) \equiv \log \left\{ \int_{\tau \in \mathscr{T}} \exp\left(\frac{1}{2}\, \mathscr{K}_M(\tau)\right) \mathrm{d}\tau \right\}. \tag{2.9}$$

In each case, H_0 of (2.3) is rejected for large values of the $H_j(\mathscr{K}_M(\cdot))$, $j = 1, \ldots, 3$, statistics. [2]

We now detail the limiting distributions of these statistics under H_c of (2.5).

Theorem 2.1. *Let y_t be generated by (2.1)–(2.2) under H_c of (2.5). Then, for $0 < \tau < 1$,*

$$\mathscr{K}_M(\cdot) \;\Rightarrow\; A_1(\cdot)/B_1(\cdot) \equiv \eta(\cdot) \tag{2.10}$$

$$H_j(\mathscr{K}_M(\cdot)) \;\Rightarrow\; H_j(\eta(\cdot)), \quad j = 1, \ldots, 3, \tag{2.11}$$

[1] Kim (2000) originally proposed a statistic of the form given in (2.6) but with $\hat{\varepsilon}_{0,i}$ and $\hat{\varepsilon}_{1,i}$ replaced by the full sample OLS residuals, $\hat{\varepsilon}_i$, from regressing y_i on an intercept, $i = 1, \ldots, T$. Representations for the limiting distributions of these statistics under H_c of (2.5) are provided in BT who also prove that they are of $\mathrm{O_p}(1)$, and hence yield *inconsistent* tests, under a class of DGPs which includes both (2.1)–(2.2) under H_1 of (2.4) and that considered in Theorem 3.3 of Kim (2000).

[2] In the actual computations of these statistics the integrals that appear in the functionals $H_j(\cdot)$, $j = 2, 3$, are replaced by averages: $H_2(\mathscr{K}_M(\cdot)) = T^{*-1} \sum_{\tau=\tau_*}^{\tau^*} \mathscr{K}_M(\tau)$ and $H_3(\mathscr{K}_M(\cdot)) = \log(T^{*-1} \sum_{\tau=\tau_*}^{\tau^*} \exp(\frac{1}{2}\mathscr{K}_M(\tau)))$, where $T^* = [\tau^* T] - [\tau_* T] + 1$ and $\mathscr{T} = [\tau_*, \tau^*] \subset [0,1]$; see also Hansen (1997).

where "⇒" denotes weak convergence of the associated probability measures, $A_1(\tau) = (1-\tau)^{-2}\int_\tau^1 [\mathbb{V}_1^{**}(r)]^2\,\mathrm{d}r$ *and* $B_1(\tau) = \tau^{-2}\int_0^\tau [\mathbb{V}_1^{***}(r)]^2\,\mathrm{d}r$, *with* $\mathbb{V}_1^{**}(r) \equiv \mathbb{V}_1(r) - \mathbb{V}_1(\tau) - (r-\tau)(1-\tau)^{-1}(\mathbb{V}_1(1) - \mathbb{V}_1(\tau))$ *and* $\mathbb{V}_1^{***}(r) \equiv \mathbb{V}_1(r) - r\tau^{-1}\mathbb{V}_1(\tau)$, *and where* $\mathbb{V}_1(r) \equiv \mathbb{W}_0(r) + c\int_{\tau_0}^r \mathbb{W}_c^*(s)\,\mathrm{d}s$, $\mathbb{W}_c^*(s) = \mathbb{W}_c(s) - \mathbb{W}_c(\tau_0)$, *with* $\mathbb{W}_0(r)$ *and* $\mathbb{W}_c(r)$ *independent standard Brownian motions on* $[0,1]$.

Remark 2.1. The representations given in Theorem 2.1 delimit the asymptotic local power functions of the ratio-based statistics under H_c of (2.5) and therefore generalise those presented in Kim et al. (2002) which apply only for $c=0$; that is, under H_0 of (2.3).

Remark 2.2. Although we have assumed that $\tau_0 \in (0,1)$, so that a break occurs under the alternative, the results of Theorem 2.1 also apply to $\tau_0 = 0$, the constant $I(1)$ process.

Remark 2.3. Following Harvey (2001) denote a first level Cramér-von Mises distribution with one degree of freedom by $CvM_1(1)$, then for $c=0$ and fixed τ, the (marginal) limiting distribution of $\mathscr{K}_M(\tau)$ is the ratio of two independent $CvM_1(1)$ distributions. That is, under H_0 of (2.3) $\mathscr{K}_M(\tau) \Rightarrow \left[\int_0^1 [\mathbb{B}_1(r)]^2\,\mathrm{d}r\right]\left[\int_0^1 [\mathbb{B}_2(r)]^2\,\mathrm{d}r\right]^{-1}$ where $\mathbb{B}_j(r) \equiv \mathbb{W}_j(r) - r\mathbb{W}_j(1)$, $j=1,2$, with $\mathbb{W}_j(r)$ two independent standard Brownian motions, $r \in [0,1]$. Asymptotic critical values for the $\mathscr{K}_M(\tau)$ and $H_j(\mathscr{K}_M(\cdot))$, $j=1,\ldots,3$, statistics are provided in Table 1.[3] The critical values for the $H_j(\mathscr{K}_M(\cdot))$, $j=1,\ldots,3$, statistics pertain to $\mathscr{T} = [0.2, 0.8]$.

We now show that the ratio-based statistics are of at most $\mathrm{O_p}(T^2)$ under a class of DGPs which includes both (2.1)–(2.2) under H_1 of (2.4) and that considered in Theorem 3.3 of Kim (2000). Remarks 2.4 and 2.5 then highlight two important practical issues.

Theorem 2.2. *Consider the process* y_t *generated by*

$$y_t = d_t + z_{t,1}, \quad t = 1,\ldots,[\tau_0 T], \quad \tau_0 \in (0,1) \tag{2.12}$$

$$y_t = d_t + z_{t,0}, \quad t = [\tau_0 T]+1,\ldots,T, \tag{2.13}$$

where $z_{t,0} = z_{t-1,0} + u_t$, *and* $z_{t,1}$ *and* u_t *are stationary processes satisfying Assumption* 1 *of Kim* (2000, *p.* 99). *Then* $\mathscr{K}_M(\tau)$, $0 < \tau \leqslant \tau_0$, *is of* $\mathrm{O_p}(T^2)$, *while for* $\tau_0 < \tau < 1$, $\mathscr{K}_M(\tau)$ *is of* $\mathrm{O_p}(1)$. *Consequently, if the intersection of the intervals* $[0,\tau_0]$ *and* $\mathscr{T}$ *is non-empty then the* $H_j(\mathscr{K}_M(\cdot))$, $j=1,\ldots,3$, *are each of* $\mathrm{O_p}(T^2)$, *they are otherwise of* $\mathrm{O_p}(1)$.

Remark 2.4. Theorem 2.2 remains valid for the constant $I(1)$ process, (2.12)–(2.13) with $\tau_0 = 0$; i.e., $\mathscr{K}_M(\tau)$, $0 < \tau < 1$, and $H_j(\mathscr{K}_M(\cdot))$, $j=1,\ldots,3$, are all of $\mathrm{O_p}(1)$.

[3] All asymptotic critical values reported in this paper were obtained by direct simulation of the relevant limiting null functionals, obtained for $c=0$, for samples of size 1000 over 10000 replications. The random number generator of the matrix programming language Ox 2.0 of Doornik (1998) was used.

Table 1
Upper tail critical values for the tests: level and linear trend

Level	90%	95%	99%	Trend	90%	95%	99%
$\mathscr{K}_M(\tau)$	4.107	6.057	12.095	$\mathscr{K}_M(\tau)$	2.734	3.675	6.202
$\mathsf{H}_1(\mathscr{K}_M(.))$	13.630	18.183	29.890	$\mathsf{H}_1(\mathscr{K}_M(.))$	6.924	8.704	12.858
$\mathsf{H}_2(\mathscr{K}_M(.))$	3.486	4.611	7.507	$\mathsf{H}_2(\mathscr{K}_M(.))$	2.324	2.858	4.231
$\mathsf{H}_3(\mathscr{K}_M(.))$	3.328	5.128	10.526	$\mathsf{H}_3(\mathscr{K}_M(.))$	1.478	1.931	3.448
$\mathsf{H}_1(\mathscr{N}\mathscr{M}(.,1))$	0.783	0.933	1.265	$\mathsf{H}_1(\mathscr{N}\mathscr{M}(.,1))$	0.233	0.271	0.349
$\mathsf{H}_2(\mathscr{N}\mathscr{M}(.,1))$	0.301	0.375	0.541	$\mathsf{H}_2(\mathscr{N}\mathscr{M}(.,1))$	0.105	0.123	0.164
$\mathsf{H}_3(\mathscr{N}\mathscr{M}(.,1))$	0.154	0.191	0.279	$\mathsf{H}_3(\mathscr{N}\mathscr{M}(.,1))$	0.053	0.062	0.083
$\mathsf{H}_1(\mathscr{S}_1(.))$	1.224	1.586	2.529	$\mathsf{H}_1(\mathscr{S}_1(.))$	0.690	0.897	1.443
$\mathsf{H}_2(\mathscr{S}_1(.))$	0.729	0.987	1.590	$\mathsf{H}_2(\mathscr{S}_1(.))$	0.297	0.373	0.563
$\mathsf{H}_3(\mathscr{S}_1(.))$	0.374	0.505	0.822	$\mathsf{H}_3(\mathscr{S}_1(.))$	0.151	0.193	0.297
$\max \mathsf{H}_1\ (\mathscr{S})$	1.561	1.974	2.939	$\max \mathsf{H}_1\ (\mathscr{S})$	0.866	1.120	1.650
$\max \mathsf{H}_2\ (\mathscr{S})$	0.913	1.214	1.787	$\max \mathsf{H}_2\ (\mathscr{S})$	0.354	0.439	0.638
$\max \mathsf{H}_3\ (\mathscr{S})$	0.473	0.631	0.940	$\max \mathsf{H}_3\ (\mathscr{S})$	0.182	0.227	0.335
$\max \mathsf{H}_1\ (\mathscr{K})$	18.066	22.720	35.252	$\max \mathsf{H}_1\ (\mathscr{K})$	8.598	10.250	14.916
$\max \mathsf{H}_2\ (\mathscr{K})$	4.629	5.883	9.121	$\max \mathsf{H}_2\ (\mathscr{K})$	2.877	3.406	4.852
$\max \mathsf{H}_3\ (\mathscr{K})$	5.071	7.205	12.950	$\max \mathsf{H}_3\ (\mathscr{K})$	1.949	2.462	4.135
$\mathscr{S}_1(0.2)$	0.502	0.670	1.056	$\mathscr{S}_1(0.2)$	0.159	0.199	0.289
$\mathscr{S}_1(0.3)$	0.580	0.778	1.298	$\mathscr{S}_1(0.3)$	0.175	0.217	0.323
$\mathscr{S}_1(0.4)$	0.679	0.922	1.516	$\mathscr{S}_1(0.4)$	0.206	0.256	0.382
$\mathscr{S}_1(0.5)$	0.757	1.030	1.687	$\mathscr{S}_1(0.5)$	0.250	0.312	0.460
$\mathscr{S}_1(0.6)$	0.842	1.162	1.885	$\mathscr{S}_1(0.6)$	0.337	0.426	0.648
$\mathscr{S}_1(0.7)$	0.905	1.241	2.189	$\mathscr{S}_1(0.7)$	0.458	0.601	0.979
$\mathscr{S}_1(0.8)$	1.026	1.410	2.347	$\mathscr{S}_1(0.8)$	0.637	0.872	1.477

The ratio-based tests will therefore not display power which tends to unity as $T \to \infty$ against processes with constant persistence—either processes displaying constant $I(1)$ behaviour *or* constant $I(0)$ behaviour (the null model).

Remark 2.5. Where τ_0 is unknown, the $H_j(\mathscr{K}_M(\cdot))$, $j = 1, \ldots, 3$, statistics will *not* yield consistent inference if the intersection of the intervals $[0, \tau_0]$ and $\mathscr{T}$ is empty. This point has not been recognised by either Kim (2000) or Kim et al. (2002) and Theorem 2.2 therefore corrects this error in their stated results.

BT and Kim et al. (2002) independently propose $\hat{\tau}_M = \operatorname{argmax}_{\tau \in \mathscr{T}} \Lambda_M(\tau)$ as an estimator for the breakpoint τ_0, in the context of (2.12)–(2.13), where $\Lambda_M(\tau) = ([(1-\tau)T]^{-2} \sum_{t=[\tau T]+1}^{T} \hat{\varepsilon}_{1,t}^2)([\tau T]^{-2} \sum_{t=1}^{[\tau T]} \hat{\varepsilon}_{0,t}^2)^{-1}$. This estimator is T-consistent for τ_0, provided $\tau_0 \in \mathscr{T}$. Simulation evidence reported in Tables 5.5a–5.5b of BT shows that $\hat{\tau}_M$ performs well in practice.

2.2. LBI tests

Consider again (2.1)–(2.2) with τ_0 known. Using King and Hillier (1985, Eq. (6), p. 99), the LBI test of H_0 of (2.3) against H_1 of (2.4) is defined by the critical region

$$\mathscr{S}_1(\tau_0) = \hat{\sigma}^{-2}(T - [\tau_0 T])^{-2} \hat{\varepsilon}' \mathbf{A}(\tau_0) \hat{\varepsilon} > \ell, \tag{2.14}$$

where $\hat{\varepsilon} \equiv (\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_T)'$, $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \hat{\varepsilon}_t^2$, and ℓ a positive constant. The matrix $\mathbf{A}(\tau_0)$ in (2.14) is the variance covariance matrix of $\mu \equiv (\mu_1, \ldots, \mu_T)'$ which therefore has (i,j)th element equal to $\min\{i - [\tau_0 T], j - [\tau_0 T]\}$, $i, j = [\tau_0 T] + 1, \ldots, T$, and all other elements equal to zero. Straightforward algebra then demonstrates that $\mathscr{S}_1(\tau_0)$ of (2.14) may be written as

$$\mathscr{S}_1(\tau_0) = \hat{\sigma}^{-2}(T - [T\tau_0])^{-2} \sum_{t=[\tau_0 T]+1}^{T} \left(\sum_{j=t}^{T} \hat{\varepsilon}_j \right)^2. \tag{2.15}$$

Remark 2.6. Note that $\mathscr{S}_1(0)$ is precisely the stationary test proposed by NM; viz.,

$$\mathscr{NM} = T^{-2} \hat{\sigma}^{-2} \sum_{t=1}^{T} \left(\sum_{j=1}^{t} \hat{\varepsilon}_j \right)^2, \tag{2.16}$$

which is therefore LBI against the constant $I(1)$ alternative.

Where τ_0 is unknown there is no LBI test of H_0 against H_1. Here we again consider the functions $H_j(\mathscr{S}_1(\cdot))$, $j = 1, \ldots, 3$, of (2.7)–(2.9) applied to the sequence $\{\mathscr{S}_1(\tau), \tau \in \mathscr{T}\}$.

We now detail the limiting distributions of the above statistics under H_c of (2.5) and demonstrate that they are of $\mathrm{O_p}(T)$ under fixed alternatives.

Theorem 2.3. *Let y_t be generated by (2.1)–(2.2). Then, for $0 \leqslant \tau < 1$,*

$$\mathscr{S}_1(\cdot) \;\Rightarrow\; \xi_1(\cdot) \tag{2.17}$$

$$H_j(\mathscr{S}_1(\cdot)) \;\Rightarrow\; H_j(\xi_1(\,\cdot\,)), \quad j = 1, \ldots, 3, \tag{2.18}$$

where $\xi_1(\tau) = (1-\tau)^{-2} \int_\tau^1 [\mathbb{V}_1(r) - r\mathbb{V}_1(1)]^2 \,\mathrm{d}r$, $\mathbb{V}_1(r)$ as defined in Theorem 2.1.

Relevant asymptotic critical values are given in Table 1. The NM test, $\mathscr{NM}$ of (2.16), has a $CvM_1(1)$ limiting null distribution; critical values are given in Table 1 of KPSS, p. 166.

Theorem 2.4. *If y_t is generated by (2.12)–(2.13), then both $\mathscr{S}_1(\tau)$, $0 \leqslant \tau < 1$, and $H_j(\mathscr{S}_1(\cdot))$, $j = 1, \ldots, 3$, are of $\mathrm{O_p}(T)$. This result also holds if $\tau_0 = 0$ in (2.12)–(2.13); cf. Remark 2.4.*

Remark 2.7. This is an important result. It tells us that, contrary to current practice in applied work, a rejection by the NM/KPSS tests cannot be taken to imply that the process under test is a constant $I(1)$ process.

Where τ_0 is known, $\mathscr{NM}$ of (2.16) makes no use of that information. It therefore seems worthwhile applying the NM statistic to *only* the last $T - [\tau_0 T]$ observations; i.e.,

$$\mathscr{NM}(\tau_0, 1) = (T - [\tau_0 T])^{-2} \hat{\sigma}_1^{-2} \sum_{t=[\tau_0 T]+1}^{T} \left(\sum_{j=[\tau_0 T]+1}^{t} \hat{\varepsilon}_{1,j} \right)^2, \tag{2.19}$$

where $\hat{\varepsilon}_{1,t}$ are defined below (2.6), and $\hat{\sigma}_1^2 = (T - [\tau_0 T])^{-1} \sum_{t=[\tau_0 T]+1}^{T} \hat{\varepsilon}_{1,t}^2$. Notice that $\mathcal{NM}(\tau_0, 1)$ is the numerator of $\mathcal{K}_M(\tau_0)$ of (2.6), scaled by the variance estimator $\hat{\sigma}_1^2$. If τ_0 is not known the functions $H_j(\mathcal{NM}(\cdot, 1))$, $j = 1, \ldots, 3$, of (2.7)–(2.9) applied to the sequence $\{\mathcal{NM}(\tau, 1), \tau \in \mathcal{T}\}$, might be considered. Notice that $\{\mathcal{NM}(\tau, 1), \tau \in \mathcal{T}\}$ is the analogue of the sequence of *reverse recursive* Dickey-Fuller statistics of BLS.

We now detail the limiting distributions of the above statistics under H_c of (2.5) and demonstrate that they are of $\mathrm{O_p}(T)$ under fixed alternatives.

Theorem 2.5. *Let y_t be generated by* (2.1)–(2.2). *Then, under* H_c *of* (2.5), *for* $0 < \tau < 1$,

$$\mathcal{NM}(\cdot, 1) \;\Rightarrow\; A_1(\cdot) \tag{2.20}$$

$$H_j(\mathcal{NM}(\,\cdot\,, 1)) \;\Rightarrow\; H_j(A_1(\,\cdot\,)), \quad j = 1, \ldots, 3, \tag{2.21}$$

where $A_1(\tau)$ is as defined in Theorem 2.1.

Remark 2.8. For $c = 0$ and fixed τ, $\mathcal{NM}(\tau, 1) \;\Rightarrow\; CvM_1(1)$. Asymptotic critical values for the $H_j(\mathcal{NM}(\,\cdot\,, 1))$, $j = 1, \ldots, 3$, statistics are provided in Table 1.

Theorem 2.6. *If y_t is generated by* (2.12)–(2.13), *then both* $\mathcal{NM}(\tau, 1)$, $0 < \tau < 1$, *and* $H_j(\mathcal{NM}(\cdot, 1))$, $j = 1, \ldots, 3$, *are of* $\mathrm{O_p}(T)$. *These results also hold for the constant* $I(1)$ *process*, (2.12)–(2.13) *with* $\tau_0 = 0$.

3. Tests against changes from $I(1)$ to $I(0)$

Suppose now that μ_t in (2.1) is generated according to

$$\mu_t = \mu_{t-1} + 1(t \leqslant [\tau_0 T])\eta_t, \tag{3.1}$$

$t = 1, \ldots, T$, with μ_0 and η_t as defined in Section 2. Notice that $\tau_0 = 1$ now corresponds to the constant $I(1)$ model. The process y_t generated by (2.1)–(3.1) is $I(1)$ up to and including time $[\tau_0 T]$ if $\sigma_\eta^2 > 0$, but reverts to $I(0)$ behaviour after the break. Consequently, a test for stationarity against a shift in persistence from a unit root to stationarity can again be framed in testing H_0 of (2.3) against H_1 of (2.4), but now in the context of (2.1)–(3.1).

The level μ_t of (3.1) is observed to 'freeze' at time $[T\tau_0]$.[4] Interestingly, under this scheme for finite T, the process is stationary after the break but, asymptotically, the process is non-stationary after the breakpoint, since the variance of $\mu_{[\tau_0 T]}$ is a linear function of T.

[4] Other parameterisations can allow for different transitions between the two regimes under the $I(1)$ to $I(0)$ shift process. One possibility is $\mu_t = 1(t \leqslant [\tau_0 T])(\mu_{t-1} + \eta_t)$. Here μ_t vanishes from (2.1) after the break point. We believe that (3.1) is probably of most practical interest and so we focus on that case.

3.1. Ratio tests

When the data are generated by (3.4)–(3.5), of which (2.1)–(3.1) under H_1 of (2.4) is a special case, $\mathscr{K}_M(\tau)$ converges in probability to zero, at rate $\mathrm{O_p}(T^{-2})$, for $0 < \tau_0 \leqslant \tau < 1$, and is of $\mathrm{O_p}(1)$ if $\tau < \tau_0 \leqslant 1$. The tests of Section 2.1 which reject for large values of $\mathscr{K}_M(\tau)$, $0 < \tau < 1$, and $H_j(\mathscr{K}_M(\,\cdot\,))$, $j = 1, \ldots, 3$, will thus be inconsistent. Now, if τ_0 were *known* rejecting for *small* values of $\mathscr{K}_M(\tau_0)$ would clearly yield a consistent test against (3.4)–(3.5). However, where τ_0 is unknown the $H_j(\mathscr{K}_M(\cdot))$, $j = 1, \ldots, 3$, statistics are all of $\mathrm{O_p}(1)$, and so tests which reject for small values of these statistics will not yield consistent inference.

From the foregoing results it is seen that tests which reject for large values of statistics based on the *reciprocal* of $\mathscr{K}_M(\tau)$ can provide consistent inference. We now detail their limiting distributions under H_c of (2.5) and explore their behaviour against fixed alternatives.

Theorem 3.1. *Let y_t be generated by* (2.1)–(3.1) *under* H_c *of* (2.5). *Then, for* $0 < \tau < 1$,

$$(\mathscr{K}_M(\cdot))^{-1} \;\Rightarrow\; A_2(\cdot)/B_2(\cdot) \equiv \eta_1(\cdot) \tag{3.2}$$

$$H_j((\mathscr{K}_M(\,\cdot\,))^{-1}) \;\Rightarrow\; H_j(\eta_1(\,\cdot\,)), \quad j = 1, \ldots, 3, \tag{3.3}$$

where $A_2(\tau) = \tau^{-2} \int_0^\tau [\mathbb{V}_2^{***}(r)]^2 \,\mathrm{d}r$ *and* $B_2(\tau) = (1-\tau)^{-2} \int_\tau^1 [\mathbb{V}_2^{**}(r)]^2 \,\mathrm{d}r$ *with* $\mathbb{V}_2^{**}(r) \equiv \mathbb{V}_2(r) - \mathbb{V}_2(\tau) - (r-\tau)(1-\tau)^{-1}(\mathbb{V}_2(1) - \mathbb{V}_2(\tau))$, $\mathbb{V}_2^{***}(r) \equiv \mathbb{V}_2(r) - r\tau^{-1}\mathbb{V}_2(\tau)$, *and* $\mathbb{V}_2(r) \equiv \mathbb{W}_0(r) + c\left\{\int_0^{\min(r,\tau_0)} \mathbb{W}_c(s)\,\mathrm{d}s + 1(r > \tau_0)[(r-\tau_0)\mathbb{W}_c(\tau_0)]\right\}$.

Remark 3.1. For $c = 0$, the asymptotic distribution of $(\mathscr{K}_M(\tau))^{-1}$ is equal to, but not independent of, that of the $\mathscr{K}_M(\tau)$ statistic. For the same reason, provided $\mathscr{T}$ is symmetric about 0.5, the asymptotic critical values for the $H_j((\mathscr{K}_M(\,\cdot\,))^{-1})$ statistics will coincide with those of the corresponding $H_j(\mathscr{K}_M(\,\cdot\,))$, $j = 1, \ldots, 3$, statistics. Asymptotic critical values for these tests may therefore be taken from Table 1.

In Theorem 3.2 we now demonstrate that the above statistics are of at most $\mathrm{O_p}(T^2)$ under a class of DGPs which includes (2.1)–(3.1) under H_1 of (2.4).

Theorem 3.2. *Consider the process y_t generated by*

$$y_t = d_t + z_{t,0}, \quad t = 1, \ldots, [\tau_0 T], \;\; \tau_0 \in (0,1) \tag{3.4}$$

$$y_t = d_t + z_{[\tau_0 T],0} + z_{t,1}, \quad t = [\tau_0 T] + 1, \ldots, T, \tag{3.5}$$

where $z_{t,j}$, $j = 0, 1$ and u_t are as defined in Theorem 2.2. *Then, for* $\tau \geqslant \tau_0$, $(\mathscr{K}_M(\tau))^{-1}$ *is of* $\mathrm{O_p}(T^2)$, *while for* $\tau < \tau_0$, $(\mathscr{K}_M(\tau))^{-1}$ *is of* $\mathrm{O_p}(1)$. *Consequently, if the intersection of the intervals* $[\tau_0, 1]$ *and* $\mathscr{T}$ *is non-empty then the* $H_j((\mathscr{K}_M(\cdot))^{-1}$, $j = 1, \ldots, 3$, *are each of* $\mathrm{O_p}(T^2)$, *but are otherwise of* $\mathrm{O_p}(1)$.

Remark 3.2. In the context of (3.4)–(3.5), $(\mathscr{K}_M(\tau))^{-1}$, $0 < \tau < 1$, and $H_j((\mathscr{K}_M(\cdot))^{-1})$, $j = 1, \ldots, 3$, are of $\mathrm{O_p}(1)$ and hence yield *inconsistent* tests if $\tau_0 = 1$,

the constant $I(1)$ model. Moreover, where τ_0 is unknown, the $H_j(\mathscr{K}_M(\cdot))$, $j=1,\ldots,3$, statistics will not yield consistent inference if the intersection of the intervals $[\tau_0,1]$ and $\mathscr{T}$ is empty; cf. Remarks 2.4 and 2.5.

Finally, and by analogy to results in Section 2.1, the estimator $\hat{\tau}_M = \text{argmin}_{\tau\in\mathscr{T}} \Lambda_M(\tau)$, where $\Lambda_M(\tau)$ is as defined in Section 2.1, is T-consistent, provided $\tau_0 \in \mathscr{T}$, for the breakpoint τ_0 in the context of (3.4)–(3.5).

3.2. LBI tests

Consider (2.1)–(3.1) with τ_0 assumed known. Using King and Hillier (1985, Eq. (6), p. 99), the LBI test of H_0 of (2.3) against H_1 of (2.4) is seen to be defined by the critical region

$$\mathscr{S}_0(\tau_0) = \hat{\sigma}^{-2}([\tau_0 T])^{-2}\hat{\varepsilon}'\mathbf{A}_1(\tau_0)\hat{\varepsilon} > \ell \tag{3.6}$$

where $\hat{\varepsilon}$ and $\hat{\sigma}^2$ are as defined in Section 2.3 and $\mathbf{A}_1(\tau_0)$, the variance covariance matrix of $\mu \equiv (\mu_1,\ldots,\mu_T)'$, has (i,j)th element equal[5] to $\min\{i,j,[\tau_0 T]\}$, $i,j=1,\ldots,T$, from which it follows that (3.6) may be re-written as

$$\mathscr{S}_0(\tau_0) = \hat{\sigma}^{-2}([T\tau_0])^{-2}\sum_{t=1}^{[\tau_0 T]}\left(\sum_{k=t}^{T}\hat{\varepsilon}_k\right)^2. \tag{3.7}$$

Remark 3.3. Notice that $\mathscr{S}_0(1)$ coincides with the NM statistic, $\mathscr{NM}$ of (2.16).

Where τ_0 is unknown there is no LBI test of H_0 against H_1. Here we again consider the functions $H_j(\mathscr{S}_0(\cdot))$, $j=1,\ldots,3$, of (2.7)–(2.9) applied to the sequence $\{\mathscr{S}_0(\tau), \tau\in\mathscr{T}\}$.

We now detail the limiting distributions of the above statistics under H_c of (2.5) and demonstrate the consistency of tests based upon these statistics under fixed alternatives.

Theorem 3.3. *Let y_t be generated by* (2.1)–(3.1) *under* H_c *of* (2.5). *Then, for* $0<\tau\leqslant 1$,

$$\mathscr{S}_0(\cdot) \;\Rightarrow\; \xi_0(\cdot), \tag{3.8}$$

$$H_j(\mathscr{S}_0(\cdot)) \;\Rightarrow\; H_j(\xi_0(\,\cdot\,)), \quad j=1,\ldots,3, \tag{3.9}$$

where $\xi_0(\tau) = \tau^{-2}\int_0^\tau [\mathbb{V}_2(r) - r\mathbb{V}_2(1)]^2\,\mathrm{d}r$, $\mathbb{V}_2(r)$ *as defined in Theorem* 3.1.

Remark 3.4. Asymptotic critical values for the $H_j(\mathscr{S}_0(\cdot))$, $j=1,\ldots,3$, statistics may again be obtained from Table 1 using those given for the corresponding $H_j(\mathscr{S}_1(\cdot))$ statistic. For the case of a known breakpoint, critical values for $\mathscr{S}_0(\tau_0)$, are as given for $\mathscr{S}_1(1-\tau_0)$.

[5] For the alternative formulation given for μ_t in footnote 3, $\mathbf{A}_1(\tau_0)$ has (i,j)th element equal to $\min\{i,j\}$, $i,j=1,\ldots,[\tau_0 T]$, and all other elements equal to zero, from which the resulting LBI test can be shown to reject for large values of the statistic $\mathscr{S}_0^*(\tau_0) = \hat{\sigma}^{-2}([T\tau_0])^{-2}\sum_{t=1}^{[\tau_0 T]}(\sum_{k=t}^{[\tau_0 T]}\hat{\varepsilon}_k)^2$.

Theorem 3.4. *If y_t is generated by* (3.4)–(3.5), *then both $\mathscr{S}_0(\tau)$, $0 < \tau \leqslant 1$, and $H_j(\mathscr{S}_0(\cdot))$, $j = 1, \ldots, 3$, are of* $\mathrm{O_p}(T)$. *These results also hold if $\tau_0 = 1$ in* (3.4)–(3.5); *cf. Remark* 3.2.

Remark 3.5. As demonstrated in the proof of Theorems 3.4 and 2.4, the OLS residuals $\hat{\varepsilon}_t$ are of $\mathrm{O_p}(T^{1/2})$ under either (2.12)–(2.13) or (3.4)–(3.5); i.e., under either direction of change. Consequently, the $\mathscr{S}_1(\tau)$, $0 \leqslant \tau < 1$, $H_j(\mathscr{S}_1(\cdot))$, $j = 1, \ldots, 3$, statistics and the $\mathscr{S}_0(\tau)$, $0 < \tau \leqslant 1$, and $H_j(\mathscr{S}_0(\cdot))$, $j = 1, \ldots, 3$, statistics will all be of $\mathrm{O_p}(T)$. That is, the LBI-based tests provide consistent inference even where neither the breakpoint, τ_0, nor the direction of change are known.

By analogy to the sub-sample NM tests of Section 2.2, if τ_0 is known it seems worthwhile applying the NM statistic to *only* the first $[\tau_0 T]$ observations; i.e.

$$\mathscr{NM}(0, \tau_0) = ([\tau_0 T])^{-2} \hat{\sigma}_0^{-2} \sum_{t=1}^{[\tau_0 T]} \left(\sum_{j=1}^{t} \hat{\varepsilon}_{0,j} \right)^2, \tag{3.10}$$

where $\hat{\varepsilon}_{0,t}$ are as defined below (2.6), and $\hat{\sigma}_0^2 = ([\tau_0 T])^{-1} \sum_{t=1}^{[\tau_0 T]} \hat{\varepsilon}_{0,t}^2$. Notice that $\mathscr{NM}(0, \tau_0)$ is the denominator of $\mathscr{K}_M(\tau_0)$ of (2.6), scaled by the variance estimator $\hat{\sigma}_0^2$. If τ_0 is unknown the functions $H_j(\mathscr{NM}(0, \cdot))$, $j = 1, \ldots, 3$, given in (2.7)–(2.9) applied to the sequence $\{\mathscr{NM}(0, \tau), \tau \in \mathscr{T}\}$, can be considered. Notice that the sequence $\{\mathscr{NM}(0, \tau), \tau \in \mathscr{T}\}$ is the analogue of the sequence of *recursive* Dickey-Fuller unit root statistics considered by BLS.

We now detail the limiting distributions of the above statistics under H_c of (2.5), and show that they are of $\mathrm{O_p}(T)$ under both fixed $I(1)$-$I(0)$ and constant $I(1)$ alternatives.

Theorem 3.5. *Let y_t be generated by* (2.1)–(3.1). *Then, under H_c of* (2.5), *for* $0 < \tau < 1$,

$$\mathscr{NM}(0, \cdot) \;\Rightarrow\; A_2(\cdot) \tag{3.11}$$

$$H_j(\mathscr{NM}(0, \cdot)) \;\Rightarrow\; H_j(A_2(\,\cdot\,)), \quad j = 1, \ldots, 3, \tag{3.12}$$

where $A_2(\tau)$ is as defined in Theorem 3.1.

Remark 3.6. For $c = 0$ and fixed τ, $\mathscr{NM}(0, \tau) \;\Rightarrow\; CvM_1(1)$. Asymptotic critical values for the $H_j(\mathscr{NM}(0, \,\cdot\,))$, $j = 1, \ldots, 3$, statistics for $\mathscr{T} = [0.2, 0.8]$ are as provided for the $H_j(\mathscr{NM}(\,\cdot\,, \tau))$, $j = 1, \ldots, 3$, statistics respectively in Table 1; cf. Remark 3.1.

Theorem 3.6. *If y_t is generated by* (3.4)–(3.5), *then both $\mathscr{NM}(0, \tau)$, $0 < \tau < 1$, and $H_j(\mathscr{NM}(0, \cdot))$, $j = 1, \ldots, 3$, are of* $\mathrm{O_p}(T)$. *These results also hold if $\tau_0 = 1$ in* (3.4)–(3.5).

Remark 3.7. As demonstrated in the proof of Theorem 3.2, the second sub-sample OLS residuals $\hat{\varepsilon}_{1,t}$, $t = [\tau T] + 1, \ldots, T$, are of $\mathrm{O_p}(T^{1/2})$, for all $\tau < \tau_0$, under the $I(1)$-$I(0)$ change DGP, (3.4)–(3.5). It therefore follows that $\mathscr{NM}(\tau, 1)$ will be of $\mathrm{O_p}(T)$, for all $\tau < \tau_0$, and, consequently, the $H_j(\mathscr{NM}(\,\cdot\,, 1))$, $j = 1, \ldots, 3$, statistics will all be

of $O_p(T)$, provided the intersection of $[0,\tau_0)$ and $\mathscr{T}$ is non-empty. Similarly, under the $I(0)$-$I(1)$ change DGP, (2.12)–(2.13), from the proof of Theorem 2.2, the first sub-sample OLS residuals $\hat{\varepsilon}_{0,t}$, $t=1,\ldots,[\tau T]$, are of $O_p(T^{1/2})$, and, hence, $\mathscr{N}\mathscr{M}(0,\tau)$ will be of $O_p(T)$, for all $\tau>\tau_0$. Consequently, the $H_j(\mathscr{N}\mathscr{M}(0,\cdot))$, $j=1,\ldots,3$, statistics will all be of $O_p(T)$, provided the intersection of $(\tau_0,1]$ and $\mathscr{T}$ is non-empty. Cf. Remark 3.5.

To conclude this Section we note that the limiting distributions for the tests of Section 2 under the DGPs of Section 3, and the tests of Section 3 under the DGPs of Section 2 can be obtained simply by replacing $\mathbb{V}_1(r)$ by $\mathbb{V}_2(r)$, and vice versa, throughout. To illustrate, if DGP (2.1)–(3.1) holds then under H_c of (2.5), $\mathscr{S}_1(\cdot) \Rightarrow \xi_3(\cdot)$, where $\xi_3(\tau)=(1-\tau)^{-2}\int_\tau^1[\mathbb{V}_2(r)-r\mathbb{V}_2(1)]^2\,dr$. Moreover, since for $c=0$, $\mathbb{V}_1(r)$ and $\mathbb{V}_2(r)$ are identically distributed it is clear that the asymptotic critical values for the tests considered in this Section can be obtained from Table 1; cf. Remarks 3.1, 3.4 and 3.6.

4. Testing when the direction of change is unknown

As shown above the ratio-based tests against $I(0)$-$I(1)$ alternatives are inconsistent against fixed $I(1)$-$I(0)$ alternatives, and vice versa. If the direction of change were unknown this might lead one to consider two-tailed tests, rejecting for either small or large values of the ratio-based statistics of Section 2.1. However, for the reasons outlined at the start of Section 3.1 this approach will only work if τ_0 is known. If not, such a procedure could only be consistent against changes from $I(0)$ to $I(1)$. In contrast, as noted in Remarks 3.5 and 3.7, the consistency results stated in Theorems 2.4 and those stated in Theorem 2.6 for the $H_j(\mathscr{N}\mathscr{M}(\cdot,1))$, $j=1,\ldots,3$, statistics remain valid in the case of fixed $I(1)$-$I(0)$ alternatives, as do those provided in Theorems 3.4 and those provided in Theorem 3.6 for the $H_j(\mathscr{N}\mathscr{M}(0,\cdot))$, $j=1,\ldots,3$, statistics, in the case of $I(0)$-$I(1)$ alternatives.

Monte Carlo results presented in Section 5 suggest that the finite sample power properties of the LBI-based tests of Sections 2.2 and 3.2 when directed against the 'wrong' alternative (by which we mean the application of tests for $I(0)$-$I(1)$ changes when the change is from $I(1)$-$I(0)$, and vice versa) are significantly diminished, relative to their power against the alternative for which they were designed. Since the LBI-based tests reject for large positive values, taking the maximum over the statistics upon which those tests are based seems worth exploring. We will also apply the same principle to the statistics based on the $\mathscr{K}_M(\cdot)$ and $(\mathscr{K}_M(\cdot))^{-1}$ sequences.[6] We therefore propose the following pairwise statistics

$$\max H_j(\mathscr{S}) \equiv \max\{H_j(\mathscr{S}_1(\cdot)), H_j(\mathscr{S}_0(\cdot))\}, \quad j=1,2,3, \tag{4.13}$$

$$\max H_j(\mathscr{K}) \equiv \max\{H_j(\mathscr{K}_M(\cdot)), H_j((\mathscr{K}_M(\cdot))^{-1})\}, \quad j=1,2,3, \tag{4.14}$$

[6] One might also consider adopting the same approach for the sub-sample NM-based tests of Sections 2.2 and 3.2. However, these turned out to have very low power and, hence, are omitted.

in each case rejecting for large values of the statistics. Since $\max(x, y)$ is continuous in both arguments, we may apply the continuous mapping theorem (CMT) directly to preceding results to establish the following theorem for the pairwise test statistics.

Theorem 4.1. *Let y_t be generated by* (2.1)–(3.1) *under* H_c *of* (2.5). *Then,*

$$\max H_j(\mathscr{S}) \;\Rightarrow\; \max\{H_j(\xi_1(\cdot)), H_j(\xi_0(\cdot))\}, \quad j = 1, 2, 3, \tag{4.15}$$

$$\max H_j(\mathscr{K}) \;\Rightarrow\; \max\{H_j(\eta(\cdot)), H_j(\eta_1(\cdot))\}, \quad j = 1, 2, 3. \tag{4.16}$$

The pairwise LBI-based test statistics are $\mathrm{O_p}(T)$ under fixed $I(1)$-$I(0)$, fixed $I(0)$-$I(1)$ and constant $I(1)$ alternatives, while the pairwise ratio-based tests are $\mathrm{O_p}(T^2)$ (provided $\tau_0 \in \mathscr{T}$) under fixed $I(1)$-$I(0)$ and fixed $I(0)$-$I(1)$ alternatives but are inconsistent against constant $I(1)$ alternatives. These results follow immediately from the properties of the statistics over which the maxima are taken; see Theorems 2.2, 2.4, 3.2 and 3.4.

Asymptotic critical values from the limiting null distributions of Theorem 4.1 are provided in Table 1. The finite sample size and power properties of the tests based on these pairwise maximum statistics are explored in Sections 5 and 6.

5. Numerical results

In this Section we use Monte Carlo simulation methods to investigate the finite sample size and power properties of the tests developed in Sections 2–4 against data generated according to either (2.1)–(2.2) or (2.1)–(3.1), setting $\sigma^2 = 1$ throughout with no loss of generality. Both the tests which require the knowledge of the true breakpoint and those which do not are considered; in the former case we assume the breakpoint to be correctly specified. We investigate the impact of varying the signal-to-noise ratio among $\sigma_\eta = 0, 0.01, 0.025, 0.05, 0.10, 0.25, 05$, and the breakpoint among $\tau_0 = 0.3, 0.5, 0.7$. All experiments were programmed using the random number generator of Ox 2.20 and 10,000 replications. All results refer to tests run at the nominal 5% asymptotic level, for a sample of size $T = 100$. Other significance levels and sample sizes were considered but gave qualitatively similar results. As would be expected, size and power properties improved for all of the tests considered as sample sizes were increased. The corresponding results for $T = 200$ are reported in BT.

In Table 2 we report empirical rejection frequencies, size under $\mathsf{H}_0 : \sigma_\eta = 0$ and power under $\mathsf{H}_1 : \sigma_\eta > 0$, for the case of a change from $I(0)$ to $I(1)$. Both the ratio-based tests of Section 2.1 and the residual-based tests of Section 2.2 all display good size and power properties. The size of these tests all lie close the asymptotic 5% level, with impressive power properties where $\sigma_\eta \geqslant 0.1$. It is interesting to note that power is always higher the smaller is τ_0. This occurs because the smaller is τ_0, the greater the proportion of the sample containing a random walk component. When τ_0 is known, the highest power is achieved using the LBI statistic $\mathscr{S}_1(\tau_0)$, particularly near the null hypothesis, as one would expect. To illustrate, $\mathscr{S}_1(0.3)$ rejects H_0 55.56% of the time when $\sigma_\eta = 0.1$, as compared with 43.21% and 32.88% for the $\mathscr{NM}(0.3, 1)$ and $\mathscr{K}_M(0.3)$, respectively.

Table 2
Empirical rejection frequencies for the tests against $I(0)$-$I(1)$, $T = 100$

	$\sigma_\eta = 0$			$\sigma_\eta = 0.01$			$\sigma_\eta = 0.025$			$\sigma_\eta = 0.05$			$\sigma_\eta = 0.1$			$\sigma_\eta = 0.25$			$\sigma_\eta = 0.5$		
	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$
$\mathscr{K}_M(\tau_0)$	5.14	5.03	5.35	5.57	5.16	5.35	7.43	6.30	5.82	14.20	9.80	7.12	32.88	22.30	11.94	70.38	57.07	35.90	90.49	83.11	65.24
$\mathsf{H}_1(\mathscr{K}_M(.))$	4.35	4.35	4.35	4.48	4.63	4.49	5.46	5.66	5.05	9.97	9.91	7.32	27.29	24.76	16.18	67.00	65.43	49.70	89.34	89.05	78.54
$\mathsf{H}_2(\mathscr{K}_M(.))$	5.36	5.36	5.36	5.71	5.78	5.75	7.05	7.41	6.75	12.23	12.92	10.37	28.46	31.07	22.84	64.37	70.68	59.48	86.63	91.23	84.36
$\mathsf{H}_3(\mathscr{K}_M(.))$	5.45	5.45	5.45	5.64	5.67	5.53	6.87	6.97	6.36	12.02	11.96	8.96	30.49	28.58	19.32	69.82	69.10	54.31	90.62	90.64	81.26
$\mathscr{NM}$	5.28	5.28	5.28	6.34	5.89	5.44	11.70	9.10	6.56	27.53	19.89	10.40	54.30	44.40	23.91	84.80	76.49	57.76	95.19	90.40	77.01
$\mathscr{NM}(\tau_0, 1)$	5.08	5.29	5.22	5.77	5.44	5.38	8.46	7.15	5.81	18.68	12.62	7.70	43.21	29.64	15.59	77.65	65.07	43.26	91.77	84.43	67.80
$\mathsf{H}_1(\mathscr{NM}(.,1))$	2.89	2.89	2.89	3.16	3.16	3.04	4.94	4.61	3.81	12.66	11.48	7.25	38.58	33.62	20.17	81.48	76.43	57.86	95.99	93.27	81.92
$\mathsf{H}_2(\mathscr{NM}(.,1))$	4.66	4.66	4.66	4.87	4.86	4.72	7.47	7.26	6.18	16.29	15.56	11.58	41.76	39.28	27.67	82.61	79.90	65.60	96.32	94.81	85.75
$\mathsf{H}_3(\mathscr{NM}(.,1))$	4.68	4.68	4.68	5.04	4.93	4.83	7.66	7.42	6.33	16.73	15.86	11.64	42.79	39.97	27.97	83.43	80.28	65.80	96.47	94.89	85.80
$\mathscr{S}_1(\tau_0)$	5.47	5.60	4.20	6.71	6.06	4.62	12.31	10.28	6.54	28.41	22.20	12.20	55.56	47.79	30.40	85.02	79.34	65.28	95.43	92.52	83.56
$\mathsf{H}_1(\mathscr{S}_1(.))$	5.93	5.93	5.93	6.73	6.51	6.26	10.86	10.17	8.33	24.46	21.12	14.67	51.92	46.94	33.43	84.55	80.70	68.52	95.10	94.12	86.81
$\mathsf{H}_2(\mathscr{S}_1(.))$	5.42	5.42	5.42	6.24	5.87	5.67	11.31	10.23	7.84	25.94	22.09	13.58	52.48	47.52	31.75	83.11	79.77	65.88	94.29	93.29	84.00
$\mathsf{H}_3(\mathscr{S}_1(.))$	5.53	5.53	5.53	6.31	5.99	5.76	11.55	10.39	8.10	26.06	22.34	13.85	52.95	47.78	32.23	83.73	80.24	66.63	94.61	93.55	84.49
max H_1 ($\mathscr{S}$)	5.05	5.05	5.05	5.92	5.56	5.39	9.51	8.03	6.46	23.02	17.63	11.53	51.93	42.26	28.09	85.60	77.72	64.47	95.75	92.68	83.73
max H_2 ($\mathscr{S}$)	4.87	4.87	4.87	5.99	5.44	5.09	10.80	8.63	6.21	25.53	19.09	11.21	53.89	44.38	26.92	85.44	77.55	62.09	95.50	91.86	81.29
max H_3 ($\mathscr{S}$)	4.75	4.75	4.75	5.84	5.46	4.94	10.68	8.56	6.14	25.21	18.88	11.23	53.71	44.15	27.09	85.57	77.77	62.53	95.59	92.01	81.78
max H_1 ($\mathscr{K}$)	4.42	4.42	4.42	4.57	4.51	4.54	5.45	5.37	4.81	10.17	8.30	5.82	29.37	21.37	12.56	69.25	61.78	43.07	89.33	86.56	73.92
max H_2 ($\mathscr{K}$)	5.16	5.16	5.16	5.48	5.40	5.25	6.75	6.83	5.89	11.31	10.45	8.44	28.53	25.23	17.67	66.31	65.24	52.59	87.04	88.55	80.47
max H_3 ($\mathscr{K}$)	5.33	5.33	5.33	5.36	5.37	5.39	6.57	6.40	5.67	11.99	9.94	7.01	32.29	23.95	14.61	71.32	64.72	46.28	90.46	88.01	76.02

The standard NM test performs well: for small values of τ_0 its power is almost as high as the LBI test $\mathscr{S}_1(\tau_0)$, as one might expect given that $\mathscr{S}_1(0) \equiv \mathscr{NM}$. It is, however, very clearly dominated on power by the $\mathscr{S}_1(\tau_0)$ test for larger values of τ_0. The standard NM test also displays superior power to the sub-sample NM tests computed for known τ_0. This observation can be explained in terms of local power functions. The fixed alternative hypothesis $\mathsf{H}_1 : \sigma_\eta > 0$ of our simulations corresponds to the local alternative $\mathsf{H}_c : \sigma_\eta = c/T$ for the $\mathscr{NM}$ statistic and $\mathsf{H}_{c'} : \sigma_\eta = c'/(1-\tau_0)T$ for the $\mathscr{NM}(\tau_0, 1)$, since the latter is constructed using only $(1-\tau_0)T$ observations. Therefore the two local power functions are comparable when $c' = c(1-\tau_0)$, which corresponds to comparing the rejection frequencies for σ_η and $\sigma_\eta/(1-\tau_0)$ respectively. For example, when $\tau_0 = 0.5$ one should compare the observed rejection frequencies for, say, $\sigma_\eta = 0.025$ in the case of $\mathscr{NM}$ with those for $\sigma_\eta = 0.05$ in the case of $\mathscr{NM}(\tau_0, 1)$, namely 9.10% against 12.62%.

The performance of the max, mean, mean-exp tests based on the $H_i(.)$, $i = 1, 2, 3$, functions respectively, which do not assume knowledge of τ_0, is broadly comparable with the corresponding statistics for known τ_0. For example, if the functions $H_j(\mathscr{S}_1(.))$, $j = 1, 2, 3$, of the $\{\mathscr{S}_1(\tau), \tau \in \mathscr{T}\}$ sequence are adopted, the power loss incurred relative to the (theoretically) best test $\mathscr{S}_1(\tau_0)$, which requires knowledge of τ_0, is very small. Overall, the max-test, $H_1(\cdot)$, seems to be slightly less reliable in terms of size, particularly when applied to the $\{\mathscr{NM}(\tau, 1), \tau \in \mathscr{T}\}$ sequence and is also generally outperformed on power by the mean and mean-exp tests, with the latter apparently preferable for the tests of Section 2.2, especially so for small values of σ_η. The highest power is displayed by the tests obtained from the $\{\mathscr{S}_1(\tau), \tau \in \mathscr{T}\}$ sequence, followed by those obtained from the $\{\mathscr{NM}(\tau, 1), \tau \in \mathscr{T}\}$ and $\{\mathscr{K}_M(\tau), \tau \in \mathscr{T}\}$ sequences, in that order. In general the $H_j(\mathscr{S}_1(.))$, $j = 1, 2, 3$, tests are also significantly more powerful than the sub-sample NM test, $\mathscr{NM}(\tau_0, 1)$, and the ratio test, $\mathscr{K}_M(\tau_0)$, both of which require knowledge of τ_0, and largely comparable with the standard NM test for $\tau_0 = 0.3$, but higher otherwise, increasingly so as τ_0 increases. For example, with $\sigma_\eta = 0.1$ and $\tau_0 = 0.7$ the rejection probabilities of the $H_j(\mathscr{S}_1(.))$, $j = 2, 3$, tests lie around 32%, as opposed to 28% for the best of the subsample NM-type tests, 23.91% for the standard NM test, and 22.84% for the best of the ratio-based tests.

The final six rows of Table 2 report results for the tests based on the pairwise maximum statistics of (4.16). Recall from Section 4 that these are designed to be powerful against both the $I(0)$-$I(1)$ and $I(1)$-$I(0)$ alternatives. The power of the LBI-based statistics $\max H_j(\mathscr{S})$, $j = 1, \ldots, 3$, against $I(0)$-$I(1)$ turns out to be comparable to that of the standard NM test, $\mathscr{NM}$, when $\tau_0 \leqslant 0.5$ and generally higher when $\tau_0 = 0.7$. In general, tests based on the $\max H_j(\mathscr{S})$ statistics are less powerful than those based on the $H_j(\mathscr{S}_1(.))$, $j = 1, 2, 3$, statistics although the power of the latter is of course only attainable with *a priori* knowledge of the direction of change. Tests based on the $\max H_j(\mathscr{K})$, $j = 1, \ldots, 3$, statistics are significantly outperformed by those based on the $\max H_j(\mathscr{S})$, $j = 1, \ldots, 3$, and by the standard NM test.

Simulation results for the case of changes from $I(1)$ to $I(0)$ are reported in Table 3. It is striking that, for the tests proposed in this paper, the results in Table 3 are almost identical to those reported in Table 2 if one exchanges τ_0 with $1 - \tau_0$. The intuition behind this is that a process which switches from $I(1)$ to $I(0)$ behaviour at τ may

Table 3
Empirical rejection frequencies for the tests against $I(1)$-$I(0)$, $T = 100$

	$\sigma_\eta = 0$			$\sigma_\eta = 0.01$			$\sigma_\eta = 0.025$			$\sigma_\eta = 0.05$			$\sigma_\eta = 0.1$			$\sigma_\eta = 0.25$			$\sigma_\eta = 0.5$		
	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$
$1/\mathscr{K}_M(\tau_0)$	4.87	5.12	5.27	4.90	5.36	5.38	5.25	6.26	7.39	6.67	9.68	14.54	11.97	21.98	33.02	36.43	56.49	70.32	64.97	82.53	90.24
$\mathsf{H}_1(1/\mathscr{K}_M(.))$	4.22	4.22	4.22	4.34	4.44	4.55	5.02	5.57	5.69	7.21	9.70	10.30	16.33	25.40	28.03	48.71	64.85	68.37	77.75	88.63	90.16
$\mathsf{H}_2(1/\mathscr{K}_M(.))$	5.32	5.32	5.32	5.45	5.54	5.51	6.44	7.44	7.31	10.37	13.30	12.41	22.70	31.10	29.60	58.81	70.45	66.48	83.20	91.71	87.76
$\mathsf{H}_3(1/\mathscr{K}_M(.))$	5.11	5.11	5.11	5.33	5.48	5.45	6.12	6.90	6.87	8.98	11.77	12.64	19.61	28.86	31.37	53.04	68.34	71.03	80.29	90.65	91.27
$\mathscr{NM}$	5.28	5.28	5.28	5.42	5.96	6.32	6.50	9.04	11.56	9.89	19.28	27.25	22.65	43.08	54.41	56.19	74.91	84.57	75.99	89.62	95.13
$\mathscr{NM}(0, \tau_0)$	6.25	5.43	5.45	6.54	5.70	6.22	7.16	7.56	9.61	9.12	13.21	19.99	17.24	31.15	44.03	46.59	66.10	78.03	69.92	85.19	92.19
$\mathsf{H}_1(\mathscr{NM}(0,.))$	3.86	3.86	3.86	3.84	4.08	3.97	4.93	5.78	6.10	8.23	12.77	14.69	21.28	35.78	42.04	59.75	77.60	83.38	81.79	93.72	96.35
$\mathsf{H}_2(\mathscr{NM}(0,.))$	5.69	5.69	5.69	5.90	6.13	6.19	7.34	8.46	8.71	12.18	17.06	18.11	29.33	41.00	44.66	66.28	80.94	84.67	85.50	95.16	96.83
$\mathsf{H}_3(\mathscr{NM}(0,.))$	5.83	5.83	5.83	6.09	6.30	6.36	7.36	8.68	9.12	12.39	17.45	18.67	29.55	41.66	45.56	66.37	81.32	85.21	85.61	95.26	96.98
$\mathscr{S}_0(\tau_0)$	3.72	4.94	5.48	4.13	6.05	6.60	5.61	9.70	11.95	10.60	21.23	27.57	27.74	45.66	54.94	63.33	77.37	84.64	81.71	91.33	95.14
$\mathsf{H}_1(\mathscr{S}_0(.))$	4.89	4.89	4.89	5.26	5.68	5.74	6.92	8.79	9.31	12.03	18.80	22.59	29.67	43.95	50.39	66.03	78.55	83.32	84.65	92.82	94.72
$\mathsf{H}_2(\mathscr{S}_0(.))$	4.68	4.68	4.68	5.08	5.79	6.02	6.87	9.18	10.66	11.97	20.17	25.22	29.10	44.94	51.40	63.36	77.36	82.03	82.00	91.99	93.94
$\mathsf{H}_3(\mathscr{S}_0(.))$	4.70	4.70	4.70	5.18	5.82	6.09	6.95	9.24	10.63	12.14	20.27	25.35	29.37	45.28	51.81	64.03	77.67	82.65	82.46	92.25	94.19
max H_1 $(\mathscr{S})$	5.05	5.05	5.05	5.35	5.51	5.84	6.44	7.95	9.37	10.29	16.48	22.16	25.50	40.25	50.99	62.10	76.08	84.94	82.07	91.35	95.83
max H_2 $(\mathscr{S})$	4.87	4.87	4.87	5.25	5.66	5.94	6.20	8.70	10.20	10.45	18.24	25.31	25.01	41.84	54.05	59.95	75.81	84.90	79.38	90.71	95.54
max H_3 $(\mathscr{S})$	4.75	4.75	4.75	5.07	5.66	5.81	6.23	8.50	9.97	10.42	18.03	24.94	25.22	41.84	53.89	60.30	75.93	85.00	79.97	90.91	95.58
max H_1 $(\mathscr{K})$	4.42	4.42	4.42	4.61	4.61	4.72	4.88	5.28	5.83	6.21	8.04	10.16	12.34	21.51	28.60	42.26	60.89	68.95	72.83	86.30	89.45
max H_2 $(\mathscr{K})$	5.16	5.16	5.16	5.28	5.31	5.46	5.90	6.60	6.83	8.08	10.41	11.46	17.75	24.69	28.44	51.66	64.20	67.17	78.87	88.79	87.35
max H_3 $(\mathscr{K})$	5.33	5.33	5.33	5.43	5.47	5.53	5.72	6.30	6.83	7.26	9.64	11.81	14.59	23.87	31.45	45.46	63.63	71.47	75.02	87.69	90.38

also be viewed as a process with the opposite switch in behaviour at $(1-\tau)$ when the observations are taken in reverse order. Consequently, all of the comments made regarding the results in Tables 2 translate almost directly to the results from Tables 3. In particular, (i) the ranking of the tests is again LBI-based, NM-based, ratio-based, and (ii) power is higher the larger is τ_0.

In practice it is likely that one will not know, a priori, whether the process has undergone a switch from $I(0)$ to $I(1)$ behaviour, or vice versa. It is therefore important to investigate the power properties of tests designed to detect a switch in one direction against a switch in the other direction. Table 4 reports simulated rejection frequencies for the various tests designed to pick up on changes from $I(0)$ to $I(1)$ when the true process switches from $I(1)$ to $I(0)$. We also considered the reverse scenario but this gave qualitatively similar results; results for this case are reported in BT.

Consider Table 4. The $\mathscr{K}_M(\tau_0)$ test displays power below the nominal level, while $\mathscr{NM}(\tau_0,1)$ has power roughly equal to size. Among the tests constructed in terms of the functionals $H_j(.)$, $j=1,2,3$, only the LBI-based statistics seem to provide reasonable power for the whole range of breakpoint locations $\tau_0=0.3,0.5,0.7$, while the sub-sample NM and ratio-based tests have very low power when τ_0 is small. Overall the ratio-based tests again perform worst. Notice that power is now higher the larger is τ_0. Indeed, for $\tau_0=0.7$ the LBI-based tests designed to detect a change from $I(0)$ to $I(1)$ are often more powerful against a change in the opposite direction. Moreover, a comparison with the results in Tables 2 and 3 shows that the standard NM test and the $\max\{H_j(\mathscr{S}_1(\cdot)),H_j(\mathscr{S}_0(\cdot))\}$, $j=1,2,3$, tests of (4.16), outperform the LBI-based tests designed for the 'wrong' alternative, as would be expected.

Overall, the simulation evidence is strongly in favour of using the LBI-based tests to detect a change in the order of integration of a time series when the direction of the change is known. If it is not, then the maximum LBI-based tests of (4.16) and the standard NM test, $\mathscr{NM}$, are preferred with the former displaying the higher power, in general. Significantly, knowledge of the time of the change does not appear to provide any notable advantage. As a final point, the inconsistency of the ratio-based tests against the 'wrong' alternative might be used constructively to help identify the direction of change, where it is unknown. Indeed, the ratio-based tests should also be useful in identifying if there is a change in persistence at all; cf. Remarks 2.4 and 3.2. We provide an illustration of this in Section 7.

6. Generalisations

In this Section we extend our model and test statistics in three directions. Firstly, we allow for a polynomial trend and/or structural breaks at known positions in the deterministic kernel d_t. Secondly, we discuss testing in the presence of structural breaks at unknown points in the sample and provide an asymptotically valid two-stage procedure for this case. Finally, as in KPSS, we extend our analysis to allow for weakly dependent errors.

Table 4
Empirical rejection frequencies for the tests against $I(0)$-$I(1)$ under the DGP $I(1)$-$I(0)$, $T = 100$

	$\sigma_\eta = 0$			$\sigma_\eta = 0.01$			$\sigma_\eta = 0.025$			$\sigma_\eta = 0.05$			$\sigma_\eta = 0.1$			$\sigma_\eta = 0.25$			$\sigma_\eta = 0.5$		
	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$
$\mathscr{K}_M(\tau_0)$	5.14	5.03	5.35	5.19	4.87	5.26	4.98	4.33	4.63	4.60	3.64	2.96	3.52	2.03	1.42	1.24	0.35	0.21	0.27	0.05	0.00
$\mathsf{H}_1(\mathscr{K}_M(.))$	4.35	4.35	4.35	4.18	4.26	4.33	4.22	4.17	4.86	3.77	4.11	6.13	2.66	5.30	13.71	1.03	13.23	31.06	0.43	18.92	38.07
$\mathsf{H}_2(\mathscr{K}_M(.))$	5.36	5.36	5.36	5.47	5.46	5.42	5.09	4.97	5.61	4.36	4.10	6.33	2.94	3.35	11.06	0.76	4.85	22.52	0.13	8.38	28.02
$\mathsf{H}_3(\mathscr{K}_M(.))$	5.45	5.45	5.45	5.45	5.43	5.45	5.12	5.16	5.87	4.53	4.90	7.54	3.20	6.19	15.91	1.22	14.11	32.90	0.44	19.73	39.57
$\mathscr{NM}(\tau_0, 1)$	5.08	5.29	5.22	5.08	5.29	5.22	5.08	5.29	5.22	5.08	5.29	5.22	5.08	5.29	5.22	5.08	5.29	5.22	5.08	5.29	5.22
$\mathsf{H}_1(\mathscr{NM}(., 1))$	2.89	2.89	2.89	2.89	2.96	2.95	2.91	3.08	3.73	2.93	3.93	8.06	2.96	8.99	26.47	3.49	40.11	69.30	6.51	68.28	89.78
$\mathsf{H}_2(\mathscr{NM}(., 1))$	4.66	4.66	4.66	4.64	4.61	4.72	4.65	4.66	5.59	4.64	5.09	9.48	4.68	7.30	24.51	4.72	25.00	63.58	4.99	52.39	86.37
$\mathsf{H}_3(\mathscr{NM}(., 1))$	4.68	4.68	4.68	4.69	4.66	4.83	4.70	4.74	5.69	4.70	5.22	9.88	4.74	7.78	25.83	4.80	28.32	65.49	5.13	56.26	87.62
$\mathscr{S}_1(\tau_0)$	5.47	5.60	4.20	5.76	5.73	4.66	6.56	7.59	6.51	8.78	12.87	13.49	17.44	30.23	32.32	46.73	61.71	62.24	68.15	76.08	74.29
$\mathsf{H}_1(\mathscr{S}_1(.))$	5.93	5.93	5.93	6.00	6.22	6.51	6.20	7.19	9.03	6.83	11.41	19.13	10.74	27.28	43.36	32.96	60.72	74.83	57.21	76.83	88.36
$\mathsf{H}_2(\mathscr{S}_1(.))$	5.42	5.42	5.42	5.41	5.56	5.93	5.65	7.24	9.39	7.14	12.59	21.91	11.89	30.30	45.99	34.08	61.63	74.13	57.64	76.28	86.09
$\mathsf{H}_3(\mathscr{S}_1(.))$	5.53	5.53	5.53	5.55	5.69	6.03	5.73	7.30	9.54	7.13	12.71	21.98	11.93	30.41	46.29	34.51	62.04	74.75	58.02	76.63	86.95

6.1. Extended deterministic specifications

Thus far we have assumed that $d_t = \beta_0$ in (2.1), but the results can be easily extended to allow $d_t = \mathbf{x}_t'\beta$, where $\mathbf{x}_t$ is a $(p \times 1)$, $p < T$, fixed sequence with associated parameter vector β. One simply needs to construct the various statistics with the OLS residuals from the regression of y_t on $\mathbf{x}_t$, $t = 1, \ldots, T$. The distributions and critical values will, however, change.

In the context of *NM* of (2.16), MacNeill (1978) provides the limiting distributions when $\mathbf{x}_t$ is the p-th order polynomial trend, $\mathbf{x}_t = (1, t, \ldots, t^{p-1}, t^p)'$, $0 \leqslant p < \infty$, while Busetti and Harvey (2001) [BH] consider the possibility of structural breaks in the trend, e.g. the broken level case, $\mathbf{x}_t = (1, h_t(\lambda_0))'$, and the broken trend case, $\mathbf{x}_t = (1, t, h_t(\lambda_0), th_t(\lambda_0))'$, where the indicator variable $h_t(\lambda_0) = 1(t > [T\lambda_0])$ and $\lambda_0 \in (0, 1)$ is a known deterministic breakpoint.

All of the test statistics developed in this paper may be extended in exactly the same way; that is, the OLS residuals $\{\hat{\varepsilon}_t\}_{t=1}^{T}$, $\{\hat{\varepsilon}_{0,t}\}_{t=1}^{[\tau T]}$ and $\{\hat{\varepsilon}_{1,t}\}_{t=[\tau T]+1}^{T}$ must now be obtained from the regression of y_t on $\mathbf{x}_t$ for $t=1, \ldots, T$, $t=1, \ldots [\tau T]$, and $t=[\tau T]+1, \ldots, T$, respectively.[7] Those tests of Sections 2 and 3 which were constructed to be LBI against a particular alternative model for $d_t = \beta_0$, when modified as above are correspondingly LBI for $d_t = \mathbf{x}_t'\beta$. Denote by $\mathbf{x}(r)$, $r \in [0, 1]$, the appropriately scaled limit on the unit interval of the $\mathbf{x}_t$; e.g. for the polynomial trend case $\mathbf{x}(r) = (1, r, \ldots, r^{p-1}, r^p)'$, $1 \leqslant p < \infty$, while for the broken trend case $\mathbf{x}(r) = (1, r, 1(r > \lambda_0), r.1(r > \lambda_0))'$. Then the limiting distributions of the statistics of Sections 2 and 3 are as detailed above, except that $\mathrm{d}\mathbb{V}_j^*(r)$, $\mathrm{d}\mathbb{V}_j^{**}(r)$ and $\mathrm{d}\mathbb{V}_j^{***}(r)$ now constitute the projection residuals of $\mathrm{d}\mathbb{V}_j(r)$ projected onto the subspace generated by $\mathbf{x}(r)$ in $L_2[0, 1]$, $L_2(\tau, 1]$ and $L_2[0, \tau]$, respectively. That is,

$$\mathbb{V}_j^*(r) \equiv \mathbb{V}_j(r) - \int_0^1 \mathbf{x}(r)'\,\mathrm{d}\mathbb{V}_j(r) \left(\int_0^1 \mathbf{x}(r)\mathbf{x}(r)'\,\mathrm{d}r \right)^{-1} \times \int_0^r \mathbf{x}(s)\,\mathrm{d}s, \quad r \in [0, 1], \tag{6.1}$$

$$\mathbb{V}_j^{**}(r) \equiv \mathbb{V}_j(r) - \mathbb{V}_j(\tau) - \int_\tau^1 \mathbf{x}(r)'\,\mathrm{d}\mathbb{V}_j(r) \left(\int_\tau^1 \mathbf{x}(r)\mathbf{x}(r)'\,\mathrm{d}r \right)^{-1} \times \int_\tau^r \mathbf{x}(s)\,\mathrm{d}s, \quad r \in (\tau, 1], \tag{6.2}$$

and

$$\mathbb{V}_j^{***}(r) \equiv \mathbb{V}_j(r) - \int_0^\tau \mathbf{x}(r)'\,\mathrm{d}\mathbb{V}_j(r) \left(\int_0^\tau \mathbf{x}(r)\mathbf{x}(r)'\,\mathrm{d}r \right)^{-1} \times \int_0^r \mathbf{x}(s)\,\mathrm{d}s, \quad r \in [0, \tau], \tag{6.3}$$

[7] When constructing the sub-sample residuals, any indicator variables should, of course, be omitted from $\mathbf{x}_t$ if $h_t(\lambda_0)$ assumes a fixed value throughout a given sub-sample. An obvious example occurs where $\tau = \lambda_0$.

in each case for $j = 1, 2$. Notice that these formulae reduce to those given in Sections 2 and 3 for the case of $\mathbf{x}_t = 1$. Critical values for the linear trend case, $\mathbf{x}_t = (1, t)'$, are provided in Table 1. The limiting null distribution of $\mathcal{NM}$ of (2.16) in this case is as a second level Cramér-von Mises distribution with one degree of freedom, denoted $CvM_2(1)$, critical values from which can be obtained from Table 1 of KPSS, p. 166. The case of a structural break in the trend is considered in more detail in the next section. The stated consistency results against fixed alternatives given in Sections 2 and 3 remain valid for the general $\mathbf{x}_t$ above, as do the corrections for weakly dependent errors outlined in Section 6.3.

We repeated the simulation experiments of Section 5 for the linear trend case, $\mathbf{x}_t = (1, t)'$. Qualitatively similar results, reported in BT, are obtained. In particular, the same ranking of the tests applies: LBI-based, NM-based, ratio-based. Interestingly, the size of the LBI-based tests seems somewhat more reliable for this case. As expected, power is somewhat lower, since one additional nuisance parameter has to be estimated; see also KPSS.

6.2. Dealing with structural breaks in the trend

We now consider further the two leading cases analysed by BH: (i) the broken level $\mathbf{x}_t = (1, h_t(\lambda_0))'$, and (ii) the broken trend $\mathbf{x}_t = (1, t, h_t(\lambda_0), th_t(\lambda_0))'$. In practice an interesting case is that of $\lambda_0 = \tau_0$; i.e. when a deterministic trend break and a persistence change occur simultaneously. We assume this to hold for the purposes of this section.

Suppose first that the breakpoint is known and consider the broken level case. Let $\hat{\varepsilon}_t$ be the OLS residuals from the regression of y_t on $\mathbf{x}_t = (1, h_t(\lambda_0))'$, $t = 1, \ldots, T$. It is then clear that $\hat{\varepsilon}_t = \hat{\varepsilon}_{0,t}$ for $t = 1, \ldots, [\lambda_0 T]$, while $\hat{\varepsilon}_t = \hat{\varepsilon}_{1,t}$ for $t = [\lambda_0 T] + 1, \ldots, T$, where $\hat{\varepsilon}_{j,t}$, $j = 0, 1$, are defined below (2.6). Consequently, the statistics $\mathcal{K}_M(\tau_0)$, $\mathcal{NM}(\tau_0, 1)$ and $\mathcal{NM}(0, \tau_0)$ are exactly as given in Sections 2 and 3, while $\mathcal{S}_1(\tau_0) = \mathcal{NM}(\tau_0, 1)\hat{\sigma}_1^2/\hat{\sigma}^2$ and $\mathcal{S}_0(\tau_0) = \mathcal{NM}(0, \tau_0)\hat{\sigma}_0^2/\hat{\sigma}^2$ where $\hat{\sigma}_1^2$ and $\hat{\sigma}_0^2$ are defined below (2.19) and (3.10) respectively and $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t^2$. Consequently, the limiting null distributions of the $\mathcal{NM}(\tau_0, 1)$, $\mathcal{NM}(0, \tau_0)$ $\mathcal{S}_1(\tau_0)$ and $\mathcal{S}_0(\tau_0)$ statistics are all $CvM_1(1)$, while that of $\mathcal{K}_M(\tau_0)$ is as given in (2.10), with critical values taken from Table 1, for the level case. Critical values for the full sample $\mathcal{NM}$ test in this case are obtained from Table I(a) of BH, p. 134, for $\lambda_0 = 0.1, 0.2, \ldots, 0.9$. A similar situation occurs for the broken trend case, except that the distribution of the sub-sample NM and LBI statistics is now $CvM_2(1)$, and the critical values for the ratio test and full sample NM test are obtained from the trend case of Table 1 and from Table I(b) of BH, p. 134, respectively.

A detailed analysis of the case $\lambda_0 \neq \tau_0$ and of the behaviour of the statistics constructed in terms of the functionals $H_i(.)$, $i = 1, 2, 3$, of (2.7)–(2.9) goes beyond the scope of this paper and it is left for future research. It is, however, worthwhile to note that for $\lambda_0 < \tau_0$ the limiting null distribution of $\mathcal{NM}(\tau_0, 1)$ remains either $CvM(1)$ [level case] or $CvM(2)$ [trend case], while critical values for $\mathcal{NM}(0, \tau_0)$ can be obtained from Tables I(a)–(b) of BH corresponding to a breakpoint location parameter of λ_0/τ_0. Similarly, for $\lambda_0 > \tau_0$ the critical values of $\mathcal{NM}(\tau_0, 1)$ can be taken from

the Tables I(a)–(b) of BH with breakpoint $(\lambda_0 - \tau_0)/(1 - \tau_0)$ and those of $\mathcal{NM}(0, \tau_0)$ from either the $CvM_1(1)$ or $CvM_2(1)$ distribution.

The case where the timing of the deterministic break is unknown is more complicated. The literature on stationarity tests has tackled this problem by estimating the breakpoint and using that estimate as if it were the true value; see, Kurozumi (2002) and Busetti and Harvey (2003). The same two-stage procedure can be used with our persistence change tests. Following Bai (1997), consider the breakpoint estimator

$$\hat{\lambda} = \arg\min_{\lambda \in \Lambda} \sum_{t=1}^{T} \hat{\varepsilon}_t(\lambda)^2, \tag{6.4}$$

where $\{\hat{\varepsilon}_t(\lambda)\}_{t=1}^{T}$ are the OLS residuals of a regression of y_t on $\mathbf{x}_t = (1, h_t(\lambda))'$ or $\mathbf{x}_t = (1, t, h_t(\lambda), th_t(\lambda))'$ for the broken level and broken trend cases respectively, and Λ is a sub-interval of $[0, 1]$ with $\lambda_0 \in \Lambda$. The estimator $\hat{\lambda}$ is T-consistent under H_0 and is then used as if it were λ_0: asymptotically valid critical values may thus be obtained from Table 1.

Table 5 provides empirical rejection frequencies for the $\mathcal{K}_M(\tau_0)$, $\mathcal{NM}(\tau_0, 1)$, $\mathcal{S}_1(\tau_0)$ and $\mathcal{NM}$ tests for data generated according to[8] (2.1)–(2.2) for $T = 100$ with $d_t = \beta_0 + \delta h_t(\lambda_0)$, $\sigma^2 = 1$, with $\lambda_0 = \tau_0$. The statistics were computed using the residuals from a regression of y_t on $\mathbf{x}_t = (1, h_t(\lambda))'$, with λ either the true value, λ_0, or the estimate from (6.4) with $\Lambda = [0.2, 0.8]$. We investigate the impact of varying the breakpoint and shift magnitude among $\lambda_0 = 0.3, 0.5, 0.7$ and $\delta = 0, 1, 2, 4$.

Where λ_0 is known all of the tests are exact invariant to δ and the results for $\mathcal{K}_M(\tau_0)$ and $\mathcal{NM}(\tau_0, 1)$ coincide with those in Table 2. All tests have empirical size close to the nominal asymptotic 5% level, with power highest for the exact LBI test although, as expected, its advantage over the sub-sample NM test is much reduced relative to the corresponding results in Table 2. As in Table 2, power is generally higher the lower is λ_0.

When the breakpoint is estimated, the results depend on δ. Table 5 shows that the tests do not seem to suffer from large size distortions, with the exception of the full sample NM test[9] for $\delta = 0$ and $\delta = 4$ and the ratio test for $\delta = 4$. In general, the larger is δ the easier it is to identify the breakpoint. Correspondingly, the power loss due to the estimation of λ_0 is low for $\delta \geqslant 2$. The ratio test is the most sensitive to breakpoint estimation: e.g., when $\delta = 4$, $\sigma_\eta = 0.25$, $\lambda_0 = 0.5$ its empirical power is only around 43%, as against 57% for the known breakpoint case. Interestingly, for the ratio test it no longer always holds that power is higher the lower is λ_0. Indeed for $\sigma_\eta = 0.5$ and $\delta \leqslant 1$ the reverse is true.

Finally, in unreported simulations (available upon request) we also considered testing without allowing for a break. It is well known that an unaccounted break induces spurious rejection, even asymptotically, for the stationarity tests; see e.g. Nyblom (1989). The same type of behaviour is expected for our tests. In fact, for $T = 100$, most tests

[8] We replicated these experiments for the corresponding $I(1)$-$I(0)$ change DGP. The results were qualitatively similar and available on request.

[9] Critical values for $\mathcal{NM}$ of (2.16) depend on the estimated breakpoint, and were obtained by linear interpolation of those reported in Table I(a) of BH.

Table 5
Empirical rejection frequencies for the case $I(0)$-$I(1)$ in the presence of a deterministic break, $T = 100$

		$\sigma_\eta = 0$			$\sigma_\eta = 0.01$			$\sigma_\eta = 0.025$			$\sigma_\eta = 0.05$			$\sigma_\eta = 0.1$			$\sigma_\eta = 0.25$			$\sigma_\eta = 0.5$		
		$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$	$\tau_0 = 0.3$	$\tau_0 = 0.5$	$\tau_0 = 0.7$
Breakpoint known																						
	$\mathscr{K}_M(\tau_0)$	5.14	5.03	5.35	5.57	5.16	5.35	7.43	6.30	5.82	14.20	9.80	7.12	32.88	22.30	11.94	70.38	57.07	35.90	90.49	83.11	65.24
	$\mathscr{NM}(\tau_0, 1)$	5.08	5.29	5.22	5.77	5.44	5.38	8.46	7.15	5.81	18.68	12.62	7.70	43.21	29.64	15.59	77.65	65.07	43.26	91.77	84.43	67.80
	$\mathscr{S}_1(\tau_0)$	5.27	5.38	5.25	5.81	5.59	5.47	8.70	7.33	6.22	19.41	13.03	8.21	44.03	30.70	16.64	79.22	68.39	47.18	93.53	88.53	75.60
	$\mathscr{NM}(\tau_0)$	4.96	4.72	5.07	5.58	4.94	5.05	8.26	6.12	5.11	18.68	10.52	5.24	43.26	25.30	5.84	78.68	61.73	13.26	93.12	83.42	34.47
Breakpoint estimated																						
$\delta = 0$	$\mathscr{K}_M(\tau_0)$	5.97	5.97	5.97	6.17	5.90	5.92	6.28	5.80	6.07	6.76	6.12	6.12	8.83	8.22	7.96	18.73	19.03	19.59	26.14	31.35	39.28
	$\mathscr{NM}(\tau_0, 1)$	3.75	3.75	3.75	3.68	3.75	3.79	3.81	3.93	3.75	5.02	4.38	4.40	10.96	7.85	6.65	40.14	31.07	22.65	67.24	57.88	46.60
	$\mathscr{S}_1(\tau_0)$	3.74	3.74	3.74	3.86	3.68	3.75	3.82	4.04	3.83	5.05	4.49	4.39	11.29	8.40	6.93	40.85	33.13	25.35	68.24	62.82	54.36
	$\mathscr{NM}$	1.93	1.93	1.93	1.82	1.94	1.87	2.22	1.97	1.92	3.05	2.45	2.29	10.30	4.73	2.97	47.83	24.70	5.97	78.78	55.61	14.59
$\delta = 1$	$\mathscr{K}_M(\tau_0)$	4.41	4.66	4.79	4.63	4.74	4.85	6.00	5.70	5.20	10.35	8.37	5.95	16.45	14.06	9.27	20.05	23.25	21.21	25.08	32.64	38.73
	$\mathscr{NM}(\tau_0, 1)$	3.86	4.25	4.34	4.44	4.62	4.49	6.48	5.65	4.74	13.53	8.99	5.86	24.61	19.11	10.73	45.74	38.90	27.29	68.89	60.79	49.04
	$\mathscr{S}_1(\tau_0)$	3.96	4.27	4.03	4.46	4.40	4.18	6.63	5.58	4.68	13.40	8.97	6.07	24.51	19.28	10.92	45.60	40.05	30.36	68.70	64.96	56.21
	$\mathscr{NM}$	3.56	3.17	3.39	4.06	3.36	3.37	6.32	4.14	3.47	13.87	6.54	3.59	31.20	16.77	4.25	58.70	40.16	9.41	81.47	61.72	19.65
$\delta = 2$	$\mathscr{K}_M(\tau_0)$	3.40	4.03	3.97	3.44	4.08	4.09	4.90	4.85	4.51	9.81	7.73	5.75	24.82	17.95	9.59	34.75	37.60	27.19	28.02	37.45	41.77
	$\mathscr{NM}(\tau_0, 1)$	4.93	5.13	5.19	5.61	5.44	5.31	8.33	6.98	5.70	17.76	11.84	7.21	40.80	27.67	14.09	60.67	54.75	36.77	73.12	66.65	55.31
	$\mathscr{S}_1(\tau_0)$	4.83	4.66	4.61	5.33	4.97	4.76	8.12	6.24	5.30	17.45	11.29	7.19	40.64	27.12	13.83	59.93	55.74	39.48	71.18	69.26	62.45
	$\mathscr{NM}$	4.38	5.05	5.75	4.82	5.22	5.67	7.52	6.09	5.69	16.53	9.66	5.90	40.32	22.96	6.37	72.59	55.93	12.02	86.34	72.21	28.15
$\delta = 4$	$\mathscr{K}_M(\tau_0)$	0.78	1.69	2.48	0.79	1.82	2.55	1.37	2.35	2.79	4.03	4.31	3.59	15.54	12.54	6.97	49.82	43.13	25.57	49.93	58.59	50.93
	$\mathscr{NM}(\tau_0, 1)$	5.01	5.12	5.41	5.54	5.54	5.46	8.45	7.19	5.93	18.27	12.46	7.70	42.33	29.20	15.23	76.11	63.85	41.43	83.88	79.39	64.82
	$\mathscr{S}_1(\tau_0)$	3.44	3.24	3.58	3.73	3.61	3.72	6.31	4.80	4.24	14.86	9.72	5.77	38.88	25.95	12.64	74.59	63.38	41.15	80.85	80.56	69.61
	$\mathscr{NM}$	3.73	7.65	8.85	3.95	7.76	8.84	6.55	8.66	8.94	14.90	12.37	9.05	39.16	25.46	9.47	75.86	59.72	14.60	91.24	81.26	33.01

have an empirical size of between 90% and 100% for a shift as small as twice the standard deviation of the noise. The exact distortions depend on the breakpoint, break magnitude and direction of change, as might be expected, and increased in magnitude as the sample size is increased.

6.3. Serial correlation

KPSS generalise (2.1)–(2.2) for $\tau_0 = 1$, to the case where the observation error process $\{\varepsilon_t\}$ satisfies the familiar α-mixing conditions of Phillips and Perron (1988, p. 336), with long run variance $\sigma_{\mathrm{L}}^2 = \lim_{T\to\infty} T^{-1}\mathrm{E}(\sum_{t=1}^{T} \varepsilon_t)^2$. In such cases, KPSS suggest replacing the OLS variance estimator $\hat{\sigma}^2$ in $\mathcal{NM}$ of (2.16) by the non-parametric estimator

$$\hat{\sigma}_{\mathrm{L}}^2 = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 + 2T^{-1}\sum_{i=1}^{m} w(i,m)\sum_{t=i+1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_{t-i}, \tag{6.5}$$

where $w(i,m) = 1 - i/(m+1)$, $i = 1,\ldots,m$, m the lag-truncation parameter. The rate conditions $m \to \infty$ and $m = \mathrm{o}(T^{1/2})$ as $T \to \infty$ are sufficient to ensure that $\hat{\sigma}_{\mathrm{L}}^2 \to^p \sigma_{\mathrm{L}}^2$ under both H_0 of (2.3) and the local alternative H_c that the long run variance of η_t is $\sigma_{\mathrm{L}}^2 c^2/T^2$; see Stock (1994, p. 2797–99).

If we generalise (2.1)–(2.2) and (2.1)–(3.1) to allow weak dependence in ε_t, no correction is needed for the ratio-based tests of Sections 2.1 and 3.1, while the LBI-based statistics of Sections 2.2 and 3.2 need to be modified in the manner suggested by KPSS. That is, by replacing $\hat{\sigma}^2$ in the right members of (2.16), (2.15) and (3.7) by $\hat{\sigma}_{\mathrm{L}}^2$ of (6.5). For the sub-sample NM-type statistics of Sections 2.2 and 3.2, we must replace the two sub-sample OLS variance estimators with the corresponding non-parametric long-run variance estimators. That is, we replace $\hat{\sigma}_{1,t}^2$ and $\hat{\sigma}_{0,t}^2$ in (2.19) and (3.10), and the functions $H_j(\cdot)$, $j = 1,\ldots,3$, thereof, by, respectively,

$$\hat{\sigma}_1^2 = (T - [\tau T])^{-1}\sum_{t=[\tau T]+1}^{T}\hat{\varepsilon}_{1,t}^2 + 2(T - [\tau T])^{-1} \times \sum_{i=1}^{m} w(i,m)\sum_{t=i+[\tau T]+1}^{T}\hat{\varepsilon}_{1,t}\hat{\varepsilon}_{1,t-i}, \tag{6.6}$$

and

$$\hat{\sigma}_0^2 = ([\tau T])^{-1}\sum_{t=1}^{[\tau T]}\hat{\varepsilon}_{0,t}^2 + 2([\tau T])^{-1}\sum_{i=1}^{m} w(i,m)\sum_{t=i+1}^{[\tau T]}\hat{\varepsilon}_{0,t}\hat{\varepsilon}_{0,t-i}. \tag{6.7}$$

The resulting non-parametrically modified statistics maintain the limiting null distributions given in Theorems 2.3, 2.5, 3.3 and 3.5 in the presence of weakly dependent errors. The consistency results stated in Sections 2 and 3 under fixed alternatives hold except that for those tests using long run variance estimators the rates are now $\mathrm{O_p}(T/m)$; cf KPSS.

We repeated the simulation experiments of Section 5 for cases where the errors display serial correlation. The results broadly confirm the findings reported in Section 5 and are therefore not presented in detail. As an example, we report in Table 6 the empirical rejection frequencies of the tests of Section 2.2 using the non-parametric correction for serial correlation outlined above with $m = 0, 4, 8$. We also report the corresponding quantities for the ratio-based tests of Section 2.1. The simulations are based on the DGP (2.1)–(2.2) for $T = 100$ and with $\tau_0 = 0.5$ but with the white noise process ε_t in (2.1) replaced by an AR(1) process with autoregressive coefficient $\rho = 0.5$.

Note first that the ratio-based tests, which do not require estimation of the long run variance, appear to be quite badly oversized, although unreported results demonstrate that empirical size does tend towards the nominal level as T is increased. For a given sample size, there is clearly nothing that the practitioner can do about the size properties of these tests. For the remaining tests, the usual trade-off between size and power in m seen in residual-based tests (see, *inter alia*, KPSS pp. 169–173) is apparent, with the exception of the $H_1(\mathcal{NM}(\cdot, 1)$ test which appears highly unreliable. Other simulation results, available on request, demonstrate that the closer is ρ to one the worse are the size distortions in the ratio tests and, for the remaining tests, the larger must m be to obtain tests with actual size close to the nominal level. As concerns power, the same ranking of the tests as noted in Section 5, i.e., LBI-based, NM-based, and ratio-based, in that order, seems to emerge from the results in Table 6, once m is chosen in each case to make the tests size-comparable.

7. Application: US inflation rate

We apply the tests discussed in this paper to the quarterly series of US inflation rate for the period 1960Q2–2000Q4. The series is calculated as first difference of the logarithm of the (seasonally adjusted) consumer price index, the latter obtained from the OECD Main Economic Indicators. Ratio-based tests apart, the statistics were computed using the non-parametric variance estimators, as outlined in Section 6.3, for $m = 0, 1, \ldots, 12$.

Consider first Table 7 which reports the tests of Sections 2 and 3 for the fixed level case, $\mathbf{x}_t = 1$. The standard NM statistic, $\mathcal{NM}$ and the $\max H_j(\mathcal{S})$, $j = 1, \ldots, 3$, statistics of (4.16) all reject H_0 at the 10% significance level when $m \leqslant 6$, at the 5% level when $m \leqslant 4$, and at the 1% level when $m \leqslant 2$. The null hypothesis that the US inflation rate follows a stationary process is clearly rejected. Notice that $\max H_j(\mathcal{K}) = H_j((\mathcal{K}_M(\cdot)^{-1})$, $j = 1, \ldots, 3$, and that in each case H_0 is rejected at the 1% level; cf Table 1.

Given the rejection of H_0, we now investigate in which direction (and at what date) the order of integration changes, and indeed if there does appear to have been a change or whether the series is a constant $I(1)$ process. To that end, notice first that, although both the statistics designed for detecting $I(0)$-$I(1)$ and those for $I(1)$-$I(0)$ point towards a rejection of H_0, the outcome looks firmer in the latter case: for example, $H_1(\mathcal{S}_0(.))$ rejects at 10% when $m \leqslant 9$, while $H_1(\mathcal{S}_1(.))$ only rejects at this significance level when $m \leqslant 3$. Recall that, with the exception of the ratio-based statistics, each statistic also yields consistent tests against persistence changes in the opposite direction. It was also remarked, in Section 5, that this feature of the ratio-based statistics might

Table 6
Empirical rejection frequencies for the tests $I(0)$-$I(1)$ modified for serial correlation: case of AR(1) disturbance with parameter 0.5, $T = 100$, $\tau_0 = 0.5$

	$\sigma_\eta = 0$			$\sigma_\eta = 0.01$			$\sigma_\eta = 0.025$			$\sigma_\eta = 0.05$			$\sigma_\eta = 0.1$			$\sigma_\eta = 0.25$			$\sigma_\eta = 0.5$		
	$m = 0$	$m = 4$	$m = 8$	$m = 0$	$m = 4$	$m = 8$	$m = 0$	$m = 4$	$m = 8$	$m = 0$	$m = 4$	$m = 8$	$m = 0$	$m = 4$	$m = 8$	$m = 0$	$m = 4$	$m = 8$	$m = 0$	$m = 4$	$m = 8$
$\mathscr{K}_M(\tau_0)$	7.56	7.56	7.56	7.46	7.46	7.46	7.96	7.96	7.96	9.16	9.16	9.16	13.26	13.26	13.26	34.42	34.42	34.42	61.64	61.64	61.64
$\mathsf{H}_1(\mathscr{K}_M(.))$	9.66	9.66	9.66	9.80	9.80	9.80	10.18	10.18	10.18	11.80	11.80	11.80	18.34	18.34	18.34	45.66	45.66	45.66	75.20	75.20	75.20
$\mathsf{H}_2(\mathscr{K}_M(.))$	10.04	10.04	10.04	10.10	10.10	10.10	10.64	10.64	10.64	12.88	12.88	12.88	19.96	19.96	19.96	49.22	49.22	49.22	77.46	77.46	77.46
$\mathsf{H}_3(\mathscr{K}_M(.))$	11.32	11.32	11.32	11.64	11.64	11.64	11.92	11.92	11.92	13.80	13.80	13.80	20.98	20.98	20.98	49.12	49.12	49.12	78.00	78.00	78.00
$\mathscr{NM}$	35.68	8.54	5.18	35.72	8.72	5.50	37.10	9.80	6.46	42.70	13.82	9.22	56.30	25.26	18.26	78.90	55.02	45.02	90.78	71.28	61.96
$\mathscr{NM}(\tau_0, 1)$	34.12	8.12	3.04	34.62	8.02	3.08	35.16	8.54	3.32	37.96	10.08	4.10	46.22	15.54	7.48	70.20	36.00	20.38	85.54	52.74	33.12
$\mathsf{H}_1(\mathscr{NM}(., 1))$	44.98	0.44	2.46	45.52	0.44	2.54	46.58	0.54	2.38	50.98	1.02	2.50	61.52	4.18	2.72	83.78	23.64	1.76	94.26	42.30	1.10
$\mathsf{H}_2(\mathscr{NM}(., 1))$	49.48	7.24	3.84	49.50	7.36	3.82	50.88	7.70	4.20	55.06	10.06	5.78	66.16	18.32	10.92	86.32	47.24	31.48	95.64	68.12	46.72
$\mathsf{H}_3(\mathscr{NM}(., 1))$	50.06	7.10	3.84	50.10	7.28	4.00	51.54	7.62	4.22	55.84	10.06	5.70	66.62	18.36	11.00	86.48	47.58	31.30	95.74	68.12	46.24
$\mathscr{S}_1(\tau_0)$	30.68	9.44	6.16	30.76	9.86	6.38	32.60	11.02	7.86	39.08	15.10	11.30	53.28	29.06	23.12	78.92	58.34	51.80	91.78	74.16	67.72
$\mathsf{H}_1(\mathscr{S}_1(.))$	35.72	9.42	6.16	35.90	9.80	6.52	38.48	10.94	7.14	43.42	14.88	10.12	57.22	27.30	19.26	82.30	57.98	47.84	93.40	76.06	66.02
$\mathsf{H}_2(\mathscr{S}_1(.))$	32.40	9.12	5.94	32.52	9.62	6.34	34.88	10.96	7.80	40.44	14.76	11.16	54.52	28.60	22.44	80.40	57.94	50.44	92.36	73.98	66.48
$\mathsf{H}_3(\mathscr{S}_1(.))$	33.10	9.30	6.06	33.26	9.70	6.60	35.54	11.10	7.80	40.92	15.12	11.18	55.36	28.98	22.68	80.98	58.34	50.64	92.66	74.68	67.06
$\max \mathsf{H}_1\ (\mathscr{S})$	41.26	8.02	4.08	41.60	8.16	4.06	43.80	9.06	4.50	48.26	12.10	5.92	60.56	22.28	12.06	82.30	52.40	36.60	92.90	71.58	56.00
$\max \mathsf{H}_2\ (\mathscr{S})$	36.98	8.36	4.60	37.18	8.68	4.72	39.30	9.68	5.26	44.56	13.06	7.76	57.36	25.16	17.02	79.98	55.02	44.10	91.94	72.04	61.38
$\max \mathsf{H}_3\ (\mathscr{S})$	37.68	8.48	4.52	38.02	8.52	4.60	39.86	9.70	4.96	45.26	13.00	7.32	57.68	25.02	16.50	80.38	54.84	43.66	92.16	72.32	61.36
$\max \mathsf{H}_1\ (\mathscr{K})$	10.92	10.92	10.92	10.94	10.94	10.94	11.24	11.24	11.24	13.30	13.30	13.30	18.80	18.80	18.80	43.50	43.50	43.50	72.86	72.86	72.86
$\max \mathsf{H}_2\ (\mathscr{K})$	11.22	11.22	11.22	10.88	10.88	10.88	11.16	11.16	11.16	12.62	12.62	12.62	18.46	18.46	18.46	44.00	44.00	44.00	73.82	73.82	73.82
$\max \mathsf{H}_3\ (\mathscr{K})$	12.90	12.90	12.90	12.86	12.86	12.86	13.00	13.00	13.00	15.00	15.00	15.00	21.22	21.22	21.22	46.28	46.28	46.28	75.24	75.24	75.24

be usefully exploited to help identify the direction of the change, and, indeed, if there is a change in persistence. For the US inflation rate series, the ratio-based statistics provide a massive rejection of H_0 in the direction $I(1)$-$I(0)$ and no rejection, or at best a borderline rejection, in the direction $I(0)$-$I(1)$. Given the characteristic of these tests to be over-sized in finite samples in the presence of serially correlated innovations (cf. Table 6), the foregoing outcomes are all consonant with a change in persistence from $I(1)$-$I(0)$ in the US inflation rate series.

The change-point estimators of Sections 2 and 3 are then computed for our data. Optimizing over the set $\mathscr{T} = [0.2, 0.8]$, as above, we obtained $\text{argmax}_{\tau\in\mathscr{T}} \Lambda_M(\tau) = 0.2$, which corresponds to 1968Q2, and $\text{argmin}_{\tau\in\mathscr{T}} \Lambda_M(\tau) = 0.76$; i.e., 1990Q4. The corner maximum at 0.2 becomes 0.15 if we stretch the feasible set to $[0.1, 0.9]$, while the minimum again suggests a change from $I(1)$ to $I(0)$, with the estimated breakpoint unchanged: $\hat{\tau}_M = 0.76$. It is interesting to observe that this is located at the time of the US recession of 1990–1991: the growth rate of US real GDP was negative for three consecutive quarters, from 1990Q3 to 1991Q1, with a quarterly inflation rate not significantly declining before 1991Q1. That the argmax estimator points towards the beginning of the sample makes it less convincing. It should be clear that for a series with an $I(1)$-$I(0)$ switch the statistic $\Lambda_M(\tau)$ takes high values for small τ, regardless of the true changepoint τ_0; that is, it is inconsistent against $I(1)$-$I(0)$ changepoints. We next computed, at the estimated breakpoint $\hat{\tau}_M = 0.76$, the statistics for a known breakpoint. In particular, as expected, $\mathscr{K}_M(\hat{\tau}_M)$ and $\mathscr{NM}(\hat{\tau}_M, 1)$ do not reject the null hypothesis, whereas $1/\mathscr{K}_M(\hat{\tau}_M)$, $\mathscr{NM}(0, \hat{\tau}_M)$ and $\mathscr{S}_0(\hat{\tau}_M)$ all provide clear rejections. The divergence in the inference between the $\mathscr{K}_M(\hat{\tau}_M)$ and $1/\mathscr{K}_M(\hat{\tau}_M)$ statistics again suggests a change in persistence, rather than a constant $I(1)$ process.

The results reported in Table 7 assume that there are no breaks in level in the inflation series. We now apply the tests of Section 6.2 to test for the possibility of a simultaneous change in level and persistence. Because the breakpoint is unknown we use the two-stage procedure outlined in Section 6.2, first estimating the breakpoint according to (6.4). This yields $\hat{\lambda} = 0.55$, corresponding to a break in level in 1982. The two-stage statistics of Section 6.2 computed at this estimated date are reported in Table 8. The level-break NM test rejects H_0 at the 1% level for all m. Of the tests designed for detecting $I(1)$ to $I(0)$ changes the $(\mathscr{K}_M(\hat{\lambda}))^{-1}$ and $\mathscr{S}_0(\hat{\lambda})$ statistics both reject H_0 at the 1% level, the latter for all m, while $\mathscr{NM}(0, \hat{\lambda})$ rejects at the 1% level for all $m \leqslant 9$ and at the 5% level for $m \leqslant 12$. In contrast, the tests designed to detect $I(0)$ to $I(1)$ changes provide much less evidence against the null. Only $\mathscr{NM}(\hat{\lambda}, 1)$ can reject H_0 at the 10% level, and then only for $m \leqslant 4$. Overall, these results are consonant with a simultaneous level break and change in persistence from $I(1)$ to $I(0)$ at $\tau = 0.55$.

As a final note, the tests developed in this paper to be consistent against changes from $I(1)$ to $I(0)$ are also consistent against changes from $I(2)$ to $I(0)$; for example, the ratio-based tests of Section 2.1 which are of $\mathrm{O_p}(T^k)$, $k \in \{0, 2\}$, against the former will be of $\mathrm{O_p}(T^{2k})$ against the latter. A referee has suggested that the US inflation rate undergoes a change from $I(2)$ to $I(0)$ in the early to mid 1980s. Allowing for a simultaneous level change, the above results appear supportive of that view.

Table 7
Results of the tests for the US inflation rate

US	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	$m=10$	$m=11$	$m=12$	10%	5%	1%
$\mathcal{NM}$	2.094	1.129	0.788	0.607	0.499	0.427	0.376	0.338	0.310	0.287	0.269	0.254	0.241	0.347	0.461	0.743
$\max \mathsf{H}_1\ (\mathcal{S})$	9.207	4.963	3.465	2.669	2.192	1.878	1.653	1.486	1.361	1.262	1.181	1.114	1.059	1.561	1.974	2.939
$\max \mathsf{H}_2\ (\mathcal{S})$	5.183	2.794	1.951	1.503	1.234	1.057	0.930	0.837	0.766	0.710	0.665	0.627	0.596	0.913	1.214	1.787
$\max \mathsf{H}_3\ (\mathcal{S})$	3.200	1.573	1.059	0.800	0.650	0.552	0.484	0.433	0.395	0.366	0.342	0.322	0.305	0.473	0.631	0.940
$\mathcal{K}_M(0.76)$	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	4.107	6.057	12.095
$\mathsf{H}_1(\mathcal{K}_M(.))$	16.841	16.841	16.841	16.841	16.841	16.841	16.841	16.841	16.841	16.841	16.841	16.841	16.841	13.630	18.163	29.890
$\mathsf{H}_2(\mathcal{K}_M(.))$	1.814	1.814	1.814	1.814	1.814	1.814	1.814	1.814	1.814	1.814	1.814	1.814	1.814	3.486	4.611	7.507
$\mathsf{H}_3(\mathcal{K}_M(.))$	4.090	4.090	4.090	4.090	4.090	4.090	4.090	4.090	4.090	4.090	4.090	4.090	4.090	3.328	5.128	10.526
$\mathcal{NM}(0.76,1)$	0.340	0.246	0.198	0.161	0.143	0.136	0.132	0.132	0.137	0.144	0.154	0.168	0.185	0.347	0.461	0.743
$\mathsf{H}_1(\mathcal{NM}(.,1))$	5.651	3.064	2.150	1.657	1.361	1.167	1.027	0.923	0.845	0.784	0.733	0.692	0.658	0.783	0.933	1.265
$\mathsf{H}_2(\mathcal{NM}(.,1))$	2.812	1.609	1.169	0.919	0.771	0.674	0.602	0.548	0.508	0.477	0.453	0.434	0.419	0.301	0.375	0.541
$\mathsf{H}_3(\mathcal{NM}(.,1))$	1.753	0.899	0.627	0.484	0.401	0.348	0.309	0.280	0.259	0.243	0.230	0.220	0.212	0.154	0.191	0.279
$\mathcal{S}_1(0.76)$	4.735	2.553	1.782	1.373	1.128	0.966	0.850	0.764	0.700	0.649	0.607	0.573	0.545	1.026	1.410	2.347
$\mathsf{H}_1(\mathcal{S}_1(.))$	4.913	2.648	1.849	1.424	1.170	1.002	0.882	0.793	0.726	0.673	0.630	0.595	0.565	1.224	1.586	2.529
$\mathsf{H}_2(\mathcal{S}_1(.))$	3.943	2.126	1.484	1.143	0.939	0.804	0.708	0.637	0.583	0.540	0.506	0.477	0.453	0.729	0.987	1.590
$\mathsf{H}_3(\mathcal{S}_1(.))$	2.057	1.089	0.755	0.579	0.475	0.406	0.357	0.321	0.293	0.272	0.254	0.240	0.228	0.374	0.505	0.822
$1/\mathcal{K}_M(0.76)$	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	4.107	6.057	12.095
$\mathsf{H}_1(1/\mathcal{K}_M(.))$	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	151.91	13.630	18.163	29.890
$\mathsf{H}_2(1/\mathcal{K}_M(.))$	20.049	20.049	20.049	20.049	20.049	20.049	20.049	20.049	20.049	20.049	20.049	20.049	20.049	3.486	4.611	7.507
$\mathsf{H}_3(1/\mathcal{K}_M(.))$	72.18	72.18	72.18	72.18	72.18	72.18	72.18	72.18	72.18	72.18	72.18	72.18	72.18	3.328	5.128	10.526
$\mathcal{NM}(0,0.76)$	2.616	1.417	0.993	0.766	0.630	0.541	0.477	0.430	0.395	0.368	0.346	0.327	0.312	0.347	0.461	0.743
$\mathsf{H}_1(\mathcal{NM}(0,.))$	6.174	3.254	2.241	1.718	1.407	1.202	1.057	0.952	0.871	0.809	0.758	0.715	0.680	0.783	0.933	1.265
$\mathsf{H}_2(\mathcal{NM}(0,.))$	3.842	2.068	1.441	1.114	0.919	0.791	0.700	0.633	0.583	0.543	0.511	0.484	0.462	0.301	0.375	0.541
$\mathsf{H}_3(\mathcal{NM}(0,.))$	2.087	1.075	0.739	0.568	0.467	0.401	0.354	0.320	0.294	0.274	0.258	0.244	0.233	0.154	0.191	0.279
$\mathcal{S}_0(0.76)$	3.151	1.698	1.186	0.913	0.750	0.643	0.566	0.509	0.466	0.432	0.404	0.381	0.362	0.502	0.670	1.056
$\mathsf{H}_1(\mathcal{S}_0(.))$	9.207	4.963	3.465	2.669	2.192	1.878	1.653	1.486	1.361	1.262	1.181	1.114	1.059	1.224	1.586	2.529
$\mathsf{H}_2(\mathcal{S}_0(.))$	5.183	2.794	1.951	1.503	1.234	1.057	0.930	0.837	0.766	0.710	0.665	0.627	0.596	0.729	0.987	1.590
$\mathsf{H}_3(\mathcal{S}_0(.))$	3.200	1.573	1.059	0.800	0.650	0.552	0.484	0.433	0.395	0.366	0.342	0.322	0.305	0.374	0.505	0.822

Table 8
Results of the tests for the US inflation rate with estimated level shift at 1982

US	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	$m=10$	$m=11$	$m=12$	10%	5%	1%
$\mathscr{NM}$	2.895	1.581	1.113	0.861	0.710	0.610	0.539	0.486	0.447	0.416	0.390	0.369	0.352	0.157	0.197	0.289
$\mathscr{K}_M(0.55)$	0.023	0.023	0.023	0.023	0.023	0.023	0.023	0.023	0.023	0.023	0.023	0.023	0.023	4.107	6.057	12.095
$\mathscr{NM}(0.55,1)$	0.870	0.592	0.480	0.403	0.362	0.334	0.309	0.287	0.272	0.261	0.252	0.247	0.242	0.347	0.461	0.743
$\mathscr{S}_1(0.55)$	0.262	0.143	0.101	0.078	0.064	0.055	0.049	0.044	0.040	0.038	0.035	0.033	0.032	0.347	0.461	0.743
$1/\mathscr{K}_M(0.55)$	43.540	43.540	43.540	43.540	43.540	43.540	43.540	43.540	43.540	43.540	43.540	43.540	43.540	4.107	6.057	12.095
$\mathscr{NM}(0,0.55)$	5.947	3.138	2.168	1.663	1.362	1.164	1.023	0.920	0.843	0.782	0.732	0.691	0.656	0.347	0.461	0.743
$\mathscr{S}_0(0.55)$	9.324	5.092	3.586	2.774	2.286	1.965	1.736	1.566	1.439	1.339	1.257	1.189	1.132	0.347	0.461	0.743

8. Conclusions

We have considered the problem of testing the null hypothesis that a series is stochastically stationary, around a (possibly broken) deterministic trend function, against the alternative that the series displays a change in persistence, either from $I(0)$ to $I(1)$ or from $I(1)$ to $I(0)$.

BT and Kim et al. (2002) develop ratio-based tests against $I(0)$ to $I(1)$ changes. We have shown these to be inconsistent against constant $I(1)$ processes and against $I(1)$ to $I(0)$ changes, and have developed consistent ratio-based tests and breakpoint estimators against $I(1)$ to $I(0)$ changes, and ratio-based tests which are consistent against either direction of change. Under Gaussianity and for a known direction of change at a known point, we have also derived LBI tests against changes in persistence. Where the change-point is unknown we proposed taking functions of the LBI statistics over all possible break-dates. Sub-sample implementations of the stationarity tests of KPSS and NM were also considered.

Numerical results suggested that, for a given direction of change, the LBI-based tests are considerably more powerful than either the ratio-based or sub-sample stationarity tests. Significant power did not appear to be gained from knowledge of the breakpoint. With the exception of the ratio-based tests, all the tests considered were shown to have good power against both changes from $I(0)$ to $I(1)$, and vice versa, and against constant $I(1)$ alternatives. Moreover, a test based on the pairwise maximum of the LBI-based statistics for the two possible directions of change generally outperformed standard stationarity tests.

A feature of the ratio-based tests is that in the presence of serially correlated innovations they do not require the arguably arbitrary decisions over the lag truncation parameter that must be taken in the context of the other statistics discussed in this paper. However, the simulation results presented in this paper have suggested that the finite sample size properties of the ratio-based tests are not satisfactory in practical situations.

Finally, we applied the tests to the US inflation rate. When no level break was allowed, the outcomes were consistent with a change in persistence from $I(1)$-$I(0)$ around the time of the 1990/1991 US recession. Where a simultaneous level change was allowed the results were consistent with a change in persistence from $I(1)$-$I(0)$ in the early to mid 1980s. Overall, the evidence presented in this paper suggests that the ratio-based and LBI-based tests are useful complements and that their use in tandem forms a useful synergy for the applied researcher.

Acknowledgements

Both authors are grateful to Max King, Steve Leybourne, Cheng Hsiao and two anonymous referees for their helpful comments on earlier versions of this paper. The second author also wishes to thank Michael McAleer and the University of Western Australia for their hospitality whilst he worked on certain aspects of the first draft of this paper.

Appendix A. Proof of theorems

Since all of the tests discussed in this paper are exact invariant to β_0 we set $\beta_0 = 0$ throughout this Appendix purely to simplify notation.

Proof of Theorem 2.1. Define the (independent) partial sum processes $\mathscr{S}^{\varepsilon}_{[Tr]} = \sum_{k=1}^{[Tr]} \varepsilon_k$, $r \in [0,1)$, where $[Tr]$ denotes the integer part of Tr, with the convention $\mathscr{S}^{\varepsilon}_{T} = \sum_{k=1}^{T} \varepsilon_k$, and $\mathscr{S}^{\eta^*}_{[Tr]} = \sum_{k=1}^{[Tr]} \eta^*_k$, $r \in [0,1)$, again with the convention that $\mathscr{S}^{\eta^*}_{T} = \sum_{k=1}^{T} \eta^*_k$, and where $\eta^*_t \equiv \sigma_\eta^{-1}\eta_t$. Under the conditions stated on (ε_t, η_t) in Section 2, cf. Chan and Wei (1988, Theorem 2.2, p. 372), $\mathscr{S}^{\varepsilon}_{[Tr]}$ and $\mathscr{S}^{\eta^*}_{[Tr]}$ satisfy a multivariate invariance principle such that

$$T^{-1/2}(\mathscr{S}^{\varepsilon}_{[Tr]}, \mathscr{S}^{\eta^*}_{[Tr]}) \;\Rightarrow\; \sigma(\mathbb{W}_0(r), \mathbb{W}_c(r)), \quad r \in [0,1], \tag{A.1}$$

jointly, where $\mathbb{W}_0(r)$ and $\mathbb{W}_c(r)$ are independent standard Brownian motions.

Consider the process y_t generated according to (2.1)–(2.2). This may be written as,

$$y_t = \varepsilon_t + \sum_{j=1}^{t} 1(j > [\tau_0 T])\eta_j, \quad t = 1, \ldots, T. \tag{A.2}$$

Consequently, from (A.2), (A.1) and an application of the CMT, under H_c of (2.5),

$$\begin{aligned} T^{-1/2}\sum_{t=1}^{[Tr]} y_t &= T^{-1/2}\sum_{t=1}^{[Tr]} \varepsilon_t + T^{-1/2}\sum_{t=1}^{[Tr]}\left(\sum_{j=1}^{t} 1(j > [\tau_0 T])\eta_j\right) \\ &\Rightarrow \sigma\left(\mathbb{W}_0(r) + c\int_{\tau_0}^{r} [\mathbb{W}_c(s) - \mathbb{W}_c(\tau_0)]\,\mathrm{d}s\right) \equiv \sigma\mathbb{V}_1(r), \end{aligned} \tag{A.3}$$

$r \in [0,1]$, where the integral is understood to exist only where $r > \tau_0$.

Consider first the case where τ is fixed. Then from (A.2), the OLS residuals from a regression of y_t on an intercept, $t = 1, \ldots, [\tau T]$, satisfy $\hat{\varepsilon}_{0,t} = y_t - [\tau T]^{-1}\sum_{s=1}^{[\tau T]} y_s$, hence

$$\begin{aligned} \hat{\varepsilon}_{0,t} &= \varepsilon_t - [\tau T]^{-1}\sum_{s=1}^{[\tau T]} \varepsilon_s + \sum_{j=1}^{t} 1(j > [\tau_0 T])\eta_j \\ &\quad - [\tau T]^{-1}\sum_{s=1}^{[\tau T]}\left(\sum_{j=1}^{s} 1(j > [\tau_0 T])\eta_j\right) \end{aligned} \tag{A.4}$$

and therefore from (A.3), (A.4) and an application of the CMT

$$T^{-1/2}\sum_{t=1}^{[Tr]} \hat{\varepsilon}_{0,t} \Rightarrow \sigma(\mathbb{V}_1(r) - r\tau^{-1}\mathbb{V}_1(\tau)), \quad r \in [0,\tau]. \tag{A.5}$$

Similarly, from (A.2), the OLS residuals from a regression of y_t on an intercept, $t=[\tau T]+1,\ldots,T$, satisfy $\hat{\varepsilon}_{1,t}=y_t-(T-[\tau T])^{-1}\sum_{s=[\tau T]+1}^{T} y_s$, hence

$$\hat{\varepsilon}_{1,t}=\varepsilon_t-(T-[\tau T])^{-1}\sum_{s=[\tau T]+1}^{T}\varepsilon_s+\sum_{j=[\tau T]}^{t}1(j>[\tau_0 T])\eta_j$$

$$-(T-[\tau T])^{-1}\sum_{s=[\tau T]+1}^{T}\left(\sum_{j=[\tau T]+1}^{s}1(j>[\tau_0 T])\eta_j\right). \tag{A.6}$$

Observe that $T^{-1/2}\sum_{t=[\tau T]+1}^{[Tr]}\hat{\varepsilon}_t\equiv T^{-1/2}(\sum_{t=1}^{[Tr]}\hat{\varepsilon}_t-\sum_{t=1}^{[\tau T]}\hat{\varepsilon}_t)$. Therefore from (A.3), (A.6) and an application of the CMT,

$$T^{-1/2}\sum_{t=[\tau T]+1}^{[Tr]}\hat{\varepsilon}_{1,t}\Rightarrow\sigma\{\mathbb{V}_1(r)-\mathbb{V}_1(\tau)$$

$$-(r-\tau)(1-\tau)^{-1}(\mathbb{V}_1(1)-\mathbb{V}_1(\tau))\}, \tag{A.7}$$

$r\in(\tau,1]$. It then follows directly from applications of the CMT that $\mathscr{K}_M(\tau)$ of (2.6) weakly converges to $\eta(\tau)$. This result holds formally only for fixed τ. The joint convergence result (2.10) stated for the sequence of statistics $\{\mathscr{K}_M(\tau),0<\tau<1\}$, then follows immediately from the fixed representation using arguments proved in Zivot and Andrews (1992). The result in (2.11) then follows directly from the main result using applications of the CMT, noting that the $H_j(\cdot)$, $j=1,\ldots,3$, are continuous functions. □

Proof of Theorem 2.2. Under the DGP (2.12)–(2.13) it is clearly seen that the first sub-sample OLS residuals $\hat{\varepsilon}_{0,t}$, $t=1,\ldots,[\tau T]$, are of $\mathrm{O_p}(1)$ provided $\tau\leqslant\tau_0$, since $\hat{\varepsilon}_{0,t}=y_t-[\tau T]^{-1}\sum_{s=1}^{[\tau T]}y_s$. However, the second sub-sample OLS residuals, $\hat{\varepsilon}_{1,t}$, are seen to be of $\mathrm{O_p}(T^{1/2})$, $t=[\tau T]+1,\ldots,T$. Therefore, for $\tau\leqslant\tau_0$, $\mathscr{K}(\tau)$ of (2.6) is of $\mathrm{O_p}(T^2)$, while for $\tau>\tau_0$, $\mathscr{K}(\tau)$ is of $\mathrm{O_p}(1)$. Consequently, the H_j are of $\mathrm{O_p}(T^2)$, provided the intersection of the intervals $[0,\tau_0]$ and $\mathscr{T}$ is non-empty. □

Proof of Theorem 2.3. From (A.2), the OLS residuals from a regression of y_t on an intercept, $t=1,\ldots,T$, satisfy $\hat{\varepsilon}_t=y_t-T^{-1}\sum_{s=1}^{T}y_s$, and hence

$$\hat{\varepsilon}_t=\varepsilon_t-T^{-1}\sum_{s=1}^{T}\varepsilon_s+\sum_{j=1}^{t}1(j>[\tau_0 T])\eta_j-T^{-1}\sum_{s=1}^{T}\left(\sum_{j=1}^{s}1(j>[\tau_0 T])\eta_j\right). \tag{A.8}$$

Therefore from (A.3), (A.8) and an application of the CMT

$$T^{-1/2}\sum_{t=1}^{[Tr]}\hat{\varepsilon}_t\ \Rightarrow\ \sigma(\mathbb{V}_1(r)-r\mathbb{V}_1(1)),\qquad r\in[0,1]. \tag{A.9}$$

The weak convergence of $\mathscr{S}_1(\tau)$ to $\xi_1(\tau)$ for fixed τ follows directly from (A.9), the CMT and the consistency of $\hat{\sigma}^2$ for σ^2 under H_c of (2.5); see Elliott and Stock (1994,

Theorem 2). The joint convergence result (2.17) stated for the sequence of statistics $\{\mathscr{S}_1(\tau), 0 < \tau < 1\}$, again follows from Zivot and Andrews (1992). The results in (2.18) then follow using the CMT. □

Proof of Theorem 2.4. Under the DGP (2.12)–(2.13), for $t = 1, \ldots, [\tau_0 T]$, y_t is of $O_p(1)$, while for $t = [\tau_0 T] + 1, \ldots, T$, y_t is of $O_p(T^{1/2})$. Since $\hat{\varepsilon}_t = y_t - T^{-1} \sum_{s=1}^{T} y_s$, these are of $O_p(T^{1/2})$ for all t. Consequently, $\sum_{t=[\tau T]+1}^{T} (\sum_{j=t}^{T} \hat{\varepsilon}_j)^2$ is of $O_p(T^4)$ for $0 \leqslant \tau < 1$. Noting from KPSS that under (2.12)–(2.13) $\hat{\sigma}^2$ is of $O_p(T)$ the stated result follows immediately. □

Proof of Theorem 2.5. The results follow from Theorem 2.1 and the consistency of $\hat{\sigma}_1^2$ for σ^2 under H_c of (2.5). □

Proof of Theorem 2.6. Since, as shown in proof of Theorem 2.2, the $\hat{\varepsilon}_{1,t}$ are of $O_p(T^{1/2})$, $t = [\tau T] + 1, \ldots, T$, the results follow immediately noting that $\hat{\sigma}_1^2$ is of $O_p(T)$. □

Proof of Theorem 3.1. Under (2.1)–(3.1), $y_t = \varepsilon_t + \sum_{j=1}^{t} 1(j \leqslant [\tau_0 T])\eta_j$, $t = 1, \ldots, T$, from which using (A.1) and the CMT it follows that, under H_c of (2.5),

$$T^{-1/2} \sum_{t=1}^{[Tr]} y_t = T^{-1/2} \sum_{t=1}^{[Tr]} \varepsilon_t + T^{-1/2} \sum_{t=1}^{[Tr]} \left(\sum_{j=1}^{t} 1(j \leqslant [\tau_0 T])\eta_j \right)$$

$$\Rightarrow \sigma \left(\mathbb{W}_0(r) + c \left\{ \int_0^{\min(r,\tau_0)} \mathbb{W}_c(s)\, \mathrm{d}s + 1(r > \tau_0)[(r - \tau_0)\mathbb{W}_c(\tau_0)] \right\} \right)$$

$$\equiv \sigma \mathbb{V}_2(r), \quad r \in [0, 1]. \tag{A.10}$$

The proof of (3.2) and (3.3) then follow along exactly the same lines as the proof of Theorem 2.1, replacing $\mathbb{V}_1(r)$ by $\mathbb{V}_2(r)$ throughout. □

Proof of Theorem 3.2. Under (3.4)–(3.5), y_t is $O_p(T^{1/2})$ for $t = 1, \ldots, [\tau_0 T]$. The first sub-sample residuals, $\hat{\varepsilon}_{0,t}$, $t = 1, \ldots, [\tau T]$, are thus also of $O_p(T^{1/2})$. Although y_t is of $O_p(T^{1/2})$ for $t = [\tau_0 T] + 1, \ldots, T$, due to $z_{[\tau_0 T],0}$ in the right member of (3.5), OLS residuals from regressing y_t on a constant for any data within the set of observations $\{y_s\}_{s=[\tau_0 T]+1}^{T}$ will be purged of $z_{[\tau_0 T],0}$, and hence of $O_p(1)$. The second sub-sample residuals, $\hat{\varepsilon}_{1,t}$, are thus of $O_p(1)$ provided $\tau \geqslant \tau_0$, and of $O_p(T^{1/2})$ otherwise. Consequently, $(\mathscr{K}_M)^{-1}$ is of $O_p(T^2)$ for all $\tau \geqslant \tau_0$, but $O_p(1)$ otherwise. The $H_j(\cdot)$, $j = 1, \ldots, 3$, are thus of $O_p(T^2)$, provided the intersection of the intervals $[\tau_0, 1]$ and $\mathscr{T}$ is non-empty. □

Proof of Theorem 3.3. The weak convergence of $\mathscr{S}_0(\tau)$ to $\xi_0(\tau)$ for fixed τ follows directly from (A.10), applications of the CMT and the consistency of $\hat{\sigma}^2$ for σ^2 under H_c of (2.5). The joint convergence result (3.8) again follows from Zivot and Andrews (1992). The results in (3.9) then follow from (3.8) using the CMT. □

Proof of Theorem 3.4. From Theorem 3.2, the OLS residuals $\hat{\varepsilon}_t$ are of $O_p(T^{1/2})$, $t = 1, \ldots, T$ under (3.4)–(3.5). The proof therefore follows exactly as for Theorem 2.4. □

Proof of Theorem 3.5. The results follow from Theorem 3.1 and the consistency of $\hat{\sigma}_1^2$ for σ^2 under H_c. □

Proof of Theorem 3.6. The stated results follow immediately from the results given in Theorem 3.2 and noting that $\hat{\sigma}_0^2$ is of $O_p(T)$. □

References

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

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.

Banerjee, A., Lumsdaine, R., Stock, J., 1992. Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. Journal of Business and Economics Statistics 10, 271–288.

Bai, J., 1997. Estimation of a change point in multiple regression models. Review of Economics and Statistics 79, 551–563.

Busetti, F., Harvey, A.C., 2001. Testing for the presence of a random walk in series with structural breaks. Journal of Time Series Analysis 22, 127–150.

Busetti, F., Harvey, A.C., 2003. Further comments on stationarity tests in series with structural breaks at unknown points. Journal of Time Series Analysis 29, 137–140.

Busetti, F., Taylor, A.M.R., 2001. Tests of stationarity against a change in persistence, University of Birmingham, Department of Economics, Discussion Paper 01-13.

Chan, N.H., Wei, C.Z., 1988. Limiting distributions of least squares estimates of unstable autoregressive processes. Annals of Statistics 16, 367–401.

Doornik, J.A., 1998. Object-oriented matrix programming using Ox 2.0. Timberlake Consultants Press, London.

Elliott, G., Stock, J.H., 1994. Inference in time series regression when the order of integration of a regressor is unknown. Econometric Theory 10, 672–700.

Hansen, B.E., 1991. Testing for structural change of unknown form in models with nonstationary regressors, Department of Economics, University of Rochester, Mimeo.

Hansen, B.E., 1997. Approximate asymptotic p-values for structural-change tests. Journal of Business and Economic Statistics 15, 60–67.

Harvey, A.C., 2001. Testing in unobserved components models. Journal of Forecasting 20, 1–19.

Kim, J.Y., 2000. Detection of change in persistence of a linear times series. Journal of Econometrics 95, 97–116.

Kim, J.Y., Belaire Franch, J., Badilli Amador, R., 2002. Corrigendum to "Detection of change in persistence of a linear times series". Journal of Econometrics 109, 389–392.

King, M.L., Hillier, G.H., 1985. Locally best invariant tests of the error covariance matrix of the linear regression model. Journal of the Royal Statistical Society, (Series B) 47, 98–102.

Kurozumi, E., 2002. Testing for stationarity with a break. Journal of Econometrics 108, 63–99.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?. Journal of Econometrics 54, 159–178.

Leybourne, S.J., Kim, T.-H., Smith, V., Newbold, P., 2000. Tests for a change in persistence against the null of difference-stationarity, Department of Economics, University of Nottingham, UK, Mimeo.

MacNeill, I., 1978. Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. Annals of Statistics 6, 422–433.

Nyblom, J., 1989. Testing for the constancy of parameters over time. Journal of the American Statistical Association 84, 223–230.

Nyblom, J., Mäkeläinen, T., 1983. Comparisons of tests for the presence of random walk coefficients in a simple linear model. Journal of the American Statistical Association 78, 856–864.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346.

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

Zivot, E., Andrews, D.W.K., 1992. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business and Economic Statistics 10, 251–270.