

A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component

Mohitosh Kejriwal^{a,*†} and Pierre Perron^b

Perron and Yabu (2009a) consider the problem of testing for a break occurring at an unknown date in the trend function of a univariate time series when the noise component can be either stationary or integrated. This article extends their work by proposing a sequential test that allows one to test the null hypothesis of, say, I breaks versus the alternative hypothesis of $(I + 1)$ breaks. The test enables consistent estimation of the number of breaks. In both stationary and integrated cases, it is shown that asymptotic critical values can be obtained from the relevant quantiles of the limit distribution of the test for a single break. Monte Carlo simulations suggest that the procedure works well in finite samples.

Keywords: Structural change; sequential procedure; feasible GLS; unit root; structural breaks.

JEL classification: C22.

1. INTRODUCTION

This article considers the problem of selecting the number of breaks in the trend function of a univariate time series without any prior knowledge as to whether the noise component is stationary, $I(0)$, or contains an autoregressive unit root, $I(1)$. This is an important practical issue as typical macroeconomic series appear to be characterized by one or more breaks in trend. For instance, Lumsdaine and Papell (1997) find evidence of structural change for 9 out of 13 macroeconomic series when allowing for two breaks in the trend function. Ben-David and Papell (1997), using a data set consisting of 48 countries, show that most countries experienced statistically significant structural changes in the paths of their export-GDP and import-GDP ratios in light of the substantial movement towards trade liberalization during the postwar period. Ben-David and Papell (2000) find evidence of multiple breaks in per capita real GDP of the G7 countries over 1870–1989. Given the discontinuity of the growth process, they then provide a demarcation between different periods of growth along the development paths based on estimates of the break dates. In another interesting application, Loewy and Papell (1996) find that allowing for trend breaks permits more rejections of the unit root hypothesis in relative per-capita income among U.S. regions, an implication that follows from the notion of stochastic convergence among regions.

Testing whether a time series contains a broken trend is complicated by the fact that it is not known *a priori* whether the noise is $I(0)$ or $I(1)$. First, doing a structural change test based on the level of the data entails different limit distributions in both cases. Further, tests based on differenced data have very poor properties when the series contains an $I(0)$ component (Vogelsang, 1998). On the other hand, to conduct inference about the presence or absence of a unit root, it is useful to have information regarding the presence or absence of changes (see Carrion-i-Silvestre *et al.*, 2009; Kim and Perron, 2009). In particular, usual unit root tests based on search procedures, suggested by Zivot and Andrews (1992) and others, are not invariant to the magnitude of the trend break if one is present. The presence of a break in slope or level can adversely affect both the size and power properties of these tests. We thus have a circular testing problem between tests on the parameters of the trend function and unit root tests.

To deal with this circular problem, various approaches have been suggested to test for the stability of the trend function that are robust to whether the errors are $I(0)$ or $I(1)$. The first to provide such a solution is Vogelsang (2001), building on prior work related to hypothesis testing on the coefficients of a polynomial time trend reported in Vogelsang (1998). He shows that Wald tests for structural change in the coefficients of a linear trend function have non-degenerate limit distributions in both $I(0)$ and $I(1)$ cases. He weights the test statistic by a unit root test scaled by some parameters so that, for a given significance level, the value of the scaling parameter can be chosen to ensure that the asymptotic critical values will be the same.

^aPurdue University

^bBoston University

*Correspondence to: Mohitosh Kejriwal, Krannert School of Management, Purdue University, 403 West State Street, West Lafayette, IN 47907, USA.

†E-mail: mkejriwa@purdue.edu

[†]PY also consider a model that involves an intercept shift only. However, the focus in this article is on models that allow for breaks in the slope of the trend function. With a fixed break size, it is not possible to obtain a consistent estimate of the true number of pure level breaks with an $I(1)$ noise component. An alternative asymptotic framework in which the break size is modelled as an increasing function of the sample size in the $I(1)$ case is explored in Harvey *et al.* (2009b).

More recently, Harvey *et al.* (2009a) propose tests based on a weighted average of the regression t -statistics for a broken trend appropriate for the case of $I(0)$ and $I(1)$ noise. The weighting function they employ is based on the KPSS stationarity test applied to the levels and first-differenced data. In the unknown break date case, they use the supremum of the trend function t -statistics, calculated for all permissible break dates, for both the $I(0)$ and $I(1)$ cases. As in Vogelsang (2001), they use a correction to ensure that the weighted test has the same asymptotic critical value irrespective of whether the noise is $I(0)$ or $I(1)$.

Perron and Yabu (2009a, henceforth PY), propose an alternative approach based on a Feasible Generalized Least Squares procedure that uses a super-efficient estimate of the sum of the autoregressive parameters α when $\alpha = 1$. When the break date is known, they show that the standard Wald test from the feasible GLS regression has the chi-square limit distribution. When the break date is unknown, the limit distributions in the $I(0)$ and $I(1)$ cases are nearly the same when constructing the test using the Exp functional of the Wald test across all permissible break dates (Andrews and Ploberger, 1994). To improve the finite sample properties of their procedure, they also use a bias-corrected version of the OLS estimate of α (obtained from an autoregression based on the residuals from estimating the trend function parameters by OLS) as suggested by Roy and Fuller (2001). Based on Monte Carlo experiments, PY show their procedure to have a power function that is close to that attainable if one knew the true value of α in many cases. The advantage of their method over those of Vogelsang (2001) and Harvey *et al.* (2009a) is that it does not involve any random scaling so that the test used is more prone to have higher power and less size distortions (see, e.g. Perron and Yabu, 2009b in a related context).

Building on the work of PY, we propose a sequential procedure that allows one to test the null hypothesis of, say, I changes, against the alternative hypothesis of $(I + 1)$ changes. Such a sequential testing strategy has been developed by Bai and Perron (1998, 2003) in the context of stationary regression models. For the model with I breaks, the estimated break dates are obtained by a global minimization of the sum of squared residuals. The strategy proceeds by testing for the presence of an additional break in each of the $(I + 1)$ segments (obtained using the estimated partition). The test thus amounts to the application of $(I + 1)$ tests of the null hypothesis of no change versus the alternative hypothesis of a single change. We derive the asymptotic distribution of the sequential test and show that, in both $I(0)$ and $I(1)$ cases, asymptotic critical values can be obtained from the relevant quantiles of the limit distribution of the test for a single break. Monte Carlo experiments indicate that the procedure performs adequately in finite samples.

It is important to note that our results do not follow as a mere corollary to those in Bai and Perron (1998, 2003) since the latter studies are based on a stationary framework and do not allow for trending regressors or unit root errors. In fact, the result that the limit distribution of the sequential test is a function of that for a single break is somewhat surprising in the light of the analysis in Kejriwal and Perron (2008), who show that break date estimates are asymptotically dependent in the presence of unit root non-stationarities. Finally, unlike the case of the stationary regression model, there is a need to reconnect the data at the break dates with trending data in order to ensure that the initial conditions remain the same across different segments (see Section 3). For an application of the advocated methodology in the context of persistence of per-capita output for OECD countries, see Kejriwal and Lopez (2009).

The article is organized as follows. Section 2 presents the models allowing for a single break and reviews the PY testing procedure. In Section 3, we develop the sequential testing procedure and derive its asymptotic properties. Section 4 provides some Monte Carlo simulations, and Section 5 offers some concluding remarks. All technical derivations are included in a mathematical appendix.

2. THE MODELS AND TEST STATISTICS: THE SINGLE BREAK CASE

Consider the following data generating process (DGP) for a scalar random variable y_t :

$$\begin{aligned} y_t &= x_t' \Psi + u_t, \\ u_t &= \alpha u_{t-1} + e_t, \quad t = 2, \dots, T \\ u_1 &= e_1, \end{aligned} \quad (1)$$

where $e_t \sim \text{i.i.d.}(0, \sigma^2)$, x_t is an $(r \times 1)$ vector of deterministic components, and Ψ is an $(r \times 1)$ vector of unknown parameters. The parameter $\alpha \in (-1, 1]$ so that u_t can be stationary or have a unit root. For ease of exposition, we focus on the AR(1) case here and defer the case of a generalized error structure for u_t to Section 3. We will consider the following two models involving a break in the slope of the trend function.¹ We denote the true break date as $T_1^0 = [T\lambda_1^0]$ for some $\lambda_1^0 \in (0, 1)$, where $[\cdot]$ denotes the largest integer that is less than or equal to the argument. Also, $I(\cdot)$ is the indicator function.

- Model 1 (structural change in slope only). Here $x_t = (1, t, DT_t)'$, $\Psi = (\mu_0, \beta_0, \beta_1)'$, where $DT_t = I(t > T_1^0)(t - T_1^0)$.
- Model 2 (structural change in both intercept and slope). Here $x_t = (1, DU_t, t, DT_t)'$ and $\Psi = (\mu_0, \mu_1, \beta_0, \beta_1)'$ where $DU_t = I(t > T_1^0)$.

For Model 1, the null hypothesis of interest is $H_0: \beta_1 = 0$ while for Model 2 it is $H_0: \mu_1 = \beta_1 = 0$. Using the notation in (1), these hypotheses can be expressed in the form $R\Psi = \gamma$ where R is a $(q \times r)$ full rank matrix and γ is a $(q \times 1)$ vector, q being the number of restrictions. For Model 1, $R = (0, 0, 1)$, $\gamma = 0$ and for Model 2,

$$R = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We first discuss the testing procedure for some generic break date $T_1 = [T\lambda_1]$, where $\lambda_1 \in \Lambda_\epsilon$ with $\Lambda_\epsilon = \{\lambda: \epsilon \leq \lambda \leq 1 - \epsilon\}$ for some $\epsilon > 0$. If α were known, the GLS estimate of the parameters can be obtained by applying OLS to the regression

$$(1 - \alpha L)y_t = (1 - \alpha L)x_t'\Psi + (1 - \alpha L)u_t, \quad \text{for } t = 2, \dots, T$$

$$y_1 = x_1'\Psi + u_1. \quad (2)$$

In practice, however, α is unknown and must be replaced by an estimate. Perron and Yabu (2009b) proposed the use of the following super-efficient estimate of α :

$$\hat{\alpha}_S = \begin{cases} \hat{\alpha}, & \text{if } T^\delta |\hat{\alpha} - 1| > d \\ 1, & \text{if } T^\delta |\hat{\alpha} - 1| \leq d \end{cases} \quad (3)$$

for $\delta \in (0, 1)$ and $d > 0$ where

$$\hat{\alpha} = \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^T \hat{u}_{t-1}^2} \quad (4)$$

and $\{\hat{u}_t\}$ are the OLS residuals from the regression of y_t on x_t . Perron and Yabu (2009b) showed that: (i) $T^{1/2}(\hat{\alpha}_S - \alpha) \xrightarrow{d} N(0, 1 - \alpha^2)$ when $|\alpha| < 1$ and (ii) $T(\hat{\alpha}_S - 1) \xrightarrow{p} 0$ when $\alpha = 1$. These results remain valid both under the null of stability as well as under the alternative of a structural break.

For testing $H_0: R\Psi = \gamma$, PY propose using the Wald statistic based on the feasible GLS regression that uses $\hat{\alpha}_S$ as an estimate of α in (2):

$$W_{FS}(\lambda_1) = (R\hat{\Psi} - \gamma)'[s^2 R(X'X)^{-1}R']^{-1}(R\hat{\Psi} - \gamma),$$

where $X = (x_1, (1 - \hat{\alpha}_S)x_2, \dots, (1 - \hat{\alpha}_S)x_T)'$, $s^2 = T^{-1} \sum_{t=1}^T \hat{e}_t^2$ and $\hat{e}_t$ are the residuals associated with the feasible GLS regression. If $|\alpha| < 1$, PY show that

$$W_{FS}(\lambda_1) \Rightarrow \left[R \left(\int_0^1 F(s, \lambda_1) F(s, \lambda_1)' ds \right)^{-1} \int_0^1 F(s, \lambda_1) dW(s) \right]' \left[R \left(\int_0^1 F(s, \lambda_1) F(s, \lambda_1)' ds \right)^{-1} R' \right]^{-1} \\ \times \left[R \left(\int_0^1 F(s, \lambda_1) F(s, \lambda_1)' ds \right)^{-1} \int_0^1 F(s, \lambda_1) dW(s) \right] \equiv G_0(\lambda_1)$$

where $F(s, \lambda_1) = [1, s, I(s > \lambda_1)(s - \lambda_1)]'$ for Model 1 and $F(s, \lambda_1) = [1, I(s > \lambda_1), s, I(s > \lambda_1)(s - \lambda_1)]'$ for Model 2. Here $W(\cdot)$ represents a standard Brownian motion on $[0, 1]$. If $\alpha = 1$,

$$W_{FS}(\lambda_1) \Rightarrow \begin{cases} \frac{[\lambda_1 W(1) - W(\lambda_1)]^2}{[\lambda_1(1 - \lambda_1)]}, & \text{for Model 1} \\ \lim_{T \rightarrow \infty} \frac{e_{[T\lambda_1]+1}}{\sigma^2} + \frac{[\lambda_1 W(1) - W(\lambda_1)]^2}{\lambda_1(1 - \lambda_1)}, & \text{for Model 2} \end{cases} \\ \equiv G_1(\lambda_1).$$

In practice, since the break date is unknown, PY propose using the exp functional over the set of permissible break dates:

$$\exp-W_{FS} = \log \left[T^{-1} \sum_{[T\lambda_1]: \lambda_1 \in \Lambda_e} \exp \left(\frac{W_{FS}(\lambda_1)}{2} \right) \right] \Rightarrow \log \left[\int_{\lambda_1 \in \Lambda_e} \exp \left(\frac{g(\lambda_1)}{2} \right) d\lambda_1 \right]$$

where $g(\lambda_1)$ denotes $G_0(\lambda_1)$ and $G_1(\lambda_1)$ for the $i(0)$ and $i(1)$ cases respectively. They show that using the exp-functional, asymptotic critical values in the $I(1)$ and $I(0)$ cases are very close so that using the larger of the two can be expected to provide tests with the correct size for both stationary and integrated errors.

Given that the OLS estimate of α may suffer from a serious downward bias especially when α is close to one, PY use a bias-corrected estimate based on the procedure proposed in Roy and Fuller (2001). The super-efficient estimate is then based on this bias-corrected estimate as opposed to the OLS estimate. The bias-corrected estimate is a function of a unit root test, namely the t -ratio $\hat{\tau} = (\hat{\alpha} - 1)/\hat{\sigma}_\alpha$, where $\hat{\alpha}$ is the OLS estimate and $\hat{\sigma}_\alpha$ is its standard deviation. The bias-corrected estimate is given by

$$\hat{\alpha}_M = \hat{\alpha} + C(\hat{\tau})\hat{\sigma}_\alpha, \\ C(\hat{\tau}) = \begin{cases} -\hat{\tau}, & \text{if } \hat{\tau} > \tau_{pct} \\ I_p T^{-1} \hat{\tau} - (1+r)[\hat{\tau} + c_2(\hat{\tau} + a)]^{-1}, & \text{if } -a < \hat{\tau} \leq \tau_{pct} \\ I_p T^{-1} \hat{\tau} - (1+r)\hat{\tau}^{-1}, & \text{if } -c_1^{1/2} < \hat{\tau} \leq -a \\ 0, & \text{if } \hat{\tau} \leq -c_1^{1/2} \end{cases}$$

where $c_1 = (1+r)T$ with r the dimension of Ψ in (2), $c_2 = [(1+r)T - \tau_{pct}^2(I_p + T)][\tau_{pct}(a + \tau_{pct})(I_p + T)]^{-1}$, a is a constant and τ_{pct} is a percentile of the limit distribution of $\hat{\tau}$ when $\alpha = 1$. Also, $I_p = (p+1)/2$, where p is the order of the autoregressive process considered for the noise component. Based on extensive simulation experiments, $a = 10$ is selected and $\tau_{0.01}$ is used as the choice for τ_{pct} .

3. THE SEQUENTIAL PROCEDURE

We consider a DGP for y_t that allows for the possibility of $(l + 1)$ breaks in the trend function. Following the notation in (1), it is:

- Model 1 ($(l + 1)$ breaks in slope only). Here $x_t = (1, t, DT_{1t}, \dots, DT_{(l+1)t})'$, $\Psi = (\mu_0, \beta_0, \beta_1, \dots, \beta_{l+1})'$ where $DT_{it} = I(t > T_i^0)(t - T_i^0)$.
- Model 2 ($(l + 1)$ breaks in both intercept and slope). Here $x_t = (1, DU_{1t}, \dots, DU_{(l+1)t}, t, DT_{1t}, \dots, DT_{(l+1)t})'$, $\Psi = (\mu_0, \mu_1, \dots, \mu_{l+1}, \beta_0, \beta_1, \dots, \beta_{l+1})'$ where $DU_{it} = I(t > T_i^0)$.

We are interested in testing the null hypothesis of l breaks against the alternative hypothesis that there are $(l + 1)$ breaks. For Model 1, this implies the null hypothesis $H_0: \beta_{l+1} = 0$ while for Model 2, the implication is $H_0: \beta_{l+1} = \mu_{l+1} = 0$. In Section 3.1, we present the sequential test and derive its limit distribution for the case where u_t is an AR(1) process. Section 3.2 subsequently extends the results to a general error structure for u_t .

3.1. The AR(1) case

Here we continue to assume that u_t is generated by (1). The sequential test of the null hypothesis of l breaks versus the alternative of $(l + 1)$ breaks is implemented as follows. First, we obtain the estimates of the break dates $\hat{T}_1, \dots, \hat{T}_l$ as global minimizers of the sum of squared residuals from the model with l breaks estimated by OLS:

$$(\hat{T}_1, \dots, \hat{T}_l) = \arg \min_{(T_1, \dots, T_l)} SSR(T_1, \dots, T_l).$$

This can be achieved using the dynamic programming algorithm proposed by Bai and Perron (2003). Next, we test for the presence of an additional break in each of the $(l + 1)$ segments obtained using the estimated partition $(\hat{T}_1, \dots, \hat{T}_l)$. In order to construct the test for the i th segment ($i = 1, \dots, l + 1$), we first estimate the following regression by OLS:

$$y_t = x_t^{(i)'} \Psi^{(i)} + u_t, \quad \text{for } t = \hat{T}_{i-1} + 1, \dots, \hat{T}_i, \quad (5)$$

where, for Model 1, we have $x_t^{(i)} = (1, t - \hat{T}_{i-1}, (t - k)I(t > k))'$, $\Psi^{(i)} = (\mu_0^{(i)}, \beta_0^{(i)}, \beta_1^{(i)})'$; while for Model 2, we have $x_t^{(i)} = (1, I(t > k), t - \hat{T}_{i-1}, (t - k)I(t > k))'$, $\Psi^{(i)} = (\mu_0^{(i)}, \mu_1^{(i)}, \beta_0^{(i)}, \beta_1^{(i)})'$. Here $k = [T\tau]$ where $\tau \in \Lambda_{i,\epsilon} = \{\tau: \hat{\lambda}_{i-1} + (\hat{\lambda}_i - \hat{\lambda}_{i-1})\epsilon \leq \tau \leq \hat{\lambda}_i - (\hat{\lambda}_i - \hat{\lambda}_{i-1})\epsilon\}$ with $\hat{\lambda}_i = \hat{T}_i/T$. We use the convention $\hat{T}_0 = 0$ and $\hat{T}_{l+1} = 0$. Note that the trend included in the i th segment is $(t - \hat{T}_{i-1})$ instead of t . This modification is needed to ensure that the initial conditions are the same across segments. The residuals from this regression denoted $\hat{u}_t^{(i)}$ are then used to compute the OLS estimate of α as in (4) for the i th segment. This estimate in turn is used to construct a super-efficient estimate of α , denoted $\hat{\alpha}_S^{(i)}$, as in (3).

Given the estimate $\hat{\alpha}_S^{(i)}$, the feasible GLS regression for the i th segment is

$$(1 - \hat{\alpha}_S^{(i)}L)y_t = (1 - \hat{\alpha}_S^{(i)}L)x_t^{(i)'} \Psi^{(i)} + (1 - \hat{\alpha}_S^{(i)}L)u_t \quad (6)$$

for $t \in [\hat{T}_{i-1} + 2, \dots, \hat{T}_i]$ together with $y_{\hat{T}_{i-1}+1} = x_{\hat{T}_{i-1}+1}^{(i)'} \Psi^{(i)} + u_{\hat{T}_{i-1}+1}$. Let $X^{(i)} = (x_{\hat{T}_{i-1}+1}^{(i)}, (1 - \hat{\alpha}_S^{(i)})x_{\hat{T}_{i-1}+2}^{(i)}, \dots, (1 - \hat{\alpha}_S^{(i)})x_{\hat{T}_i}^{(i)})'$ and the feasible GLS estimate of $\Psi^{(i)}$ be denoted by $\hat{\Psi}^{(i)}$. The Wald statistic for a given $\tau \in \Lambda_{i,\epsilon}$ is then given by

$$W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i) = (R\hat{\Psi}^{(i)} - \gamma)' [s_i^2 R(X^{(i)'} X^{(i)})^{-1} R']^{-1} (R\hat{\Psi}^{(i)} - \gamma), \quad (7)$$

where $s_i^2 = (\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} [\hat{e}_t^{(i)}]^2$ and $\hat{e}_t^{(i)}$ are the residuals from OLS estimation of (6). As in PY, we use the exp functional over all permissible break dates:

$$\exp -W_{FS}^{(i)} = \log \left[(\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{\tau \in \Lambda_{i,\epsilon}} \exp \left(\frac{W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i)}{2} \right) \right]$$

Given $\exp -W_{FS}^{(i)}$ for $i = 1, \dots, l + 1$, the sequential test is defined by

$$F_T(l + 1|l) = \max_{1 \leq i \leq l+1} \{\exp -W_{FS}^{(i)}\}.$$

We conclude in favour of a model with $(l+1)$ breaks if the maximum of the $\exp -W_{FS}^{(i)}$ tests is sufficiently large. The test thus amounts to the application of $(l + 1)$ tests of the null hypothesis of no change versus the alternative hypothesis of a single change. The following theorem states the limit distribution of the sequential test under the null hypothesis of l breaks.

THEOREM 1. Assume that $u_t(t = 1, \dots, T)$ is generated by (1). Under the null hypothesis that the true number of breaks is l , we have $\lim_{T \rightarrow \infty} P(F_T(l + 1|l) \leq x) = H_\epsilon(x)^{l+1}$ with $H_\epsilon(x)$ being the distribution function of $\log \int_{\lambda_1 \in A_\epsilon} \exp(g(\lambda_1)/2) d\lambda_1$ where $g(\lambda_1) = G_0(\lambda_1)$ if $|\alpha| < 1$ and $g(\lambda_1) = G_1(\lambda_1)$ if $\alpha = 1$.

The proof is in the appendix. The theorem states that, in both $l(0)$ and $l(1)$ cases, asymptotic critical values for the sequential test can be obtained from the relevant quantiles of the limit distribution for the single break test. A similar result was obtained by Bai and Perron (1998) in the context of stationary regression models. In our context, however, the result is somewhat surprising in view of the analysis in Kejriwal and Perron (2008), who show that break date estimates are asymptotically

dependent in the presence of unit root non-stationarities. We calculated the critical values by simulations using i.i.d. $N(0,1)$ random variables to approximate the Wiener process. The integrals are approximated by normalized sums with 2000 steps and the number of replications used is 10,000. Table 1 presents critical values for a wide range of values of the trimming parameter ϵ . As is evident from the tables, the quantiles in the $I(0)$ and $I(1)$ cases are quite close, and hence using the larger of the two can be expected to provide tests with the correct size in both cases. The argument for the consistency of the sequential test is the same as that in Bai and Perron (1998). Note also that the result continues to hold when using the bias-corrected estimates given that the bias correction procedure results in different finite sample properties but does not alter the rate of convergence of the estimates.

The test based on $F_T(l+1|l)$ can be used to provide an estimate of the number of breaks in the following way. First, apply the one break test $F_T(1|0)$ to determine if there is at least one break. Upon a rejection, use the test $F_T(2|1)$ to determine if there is more than one break. This process is repeated by increasing l sequentially until the test fails to reject the null hypothesis of no additional structural breaks. The estimated number of breaks is then obtained as the number of rejections. The sequential procedure can be made consistent by allowing the significance level of the test $F_T(l+1|l)$ to decrease to zero at a suitable rate as the sample size increases. It can be shown that if the true number of breaks is at least $l+1$, $F_T(l+1|l)$ diverges at rate $O_p(T^3)$ if $|\alpha| < 1$ and at rate $O_p(T)$ if $\alpha = 1$. Thus, if the critical value is allowed to increase at rate $O_p(T^{1-\epsilon})$, $0 < \epsilon < 1$, the size of the test converges to zero at T increases while ensuring the consistency of the test. This leads to the Theorem 2 whose proof is similar to that of Hosoya (1989) and is, therefore, omitted.

Table 1. Asymptotic critical values of the sequential test $F_T(l+1|l)$ for Models 1 and 2

		I(0)					I(1)				
		I					I				
ϵ	α	1	2	3	4	5	1	2	3	4	5
Model 1											
0.01	0.90	2.02	2.33	2.60	2.82	2.97	2.08	2.37	2.65	2.81	2.99
	0.95	2.61	2.98	3.24	3.41	3.59	2.66	3.03	3.25	3.47	3.62
	0.975	3.24	3.60	3.85	4.04	4.22	3.27	3.62	3.86	4.14	4.31
	0.99	4.04	4.40	4.75	5.09	5.30	4.14	4.50	4.79	4.95	5.07
0.05	0.90	1.90	2.25	2.52	2.72	2.88	1.93	2.23	2.46	2.62	2.82
	0.95	2.55	2.92	3.15	3.31	3.47	2.49	2.84	3.04	3.22	3.37
	0.975	3.15	3.49	3.63	3.86	4.04	3.05	3.38	3.68	3.80	3.98
	0.99	3.86	4.25	4.52	4.76	4.97	3.80	4.32	4.67	4.84	4.90
0.10	0.90	1.75	2.08	2.30	2.50	2.68	1.82	2.12	2.40	2.56	2.71
	0.95	2.32	2.72	3.00	3.23	3.38	2.41	2.73	2.90	3.07	3.25
	0.975	3.04	3.39	3.66	3.79	3.91	2.91	3.25	3.51	3.76	3.90
	0.99	3.79	4.11	4.53	4.76	4.86	3.76	4.18	4.51	4.72	4.82
0.15	0.90	1.67	1.94	2.18	2.36	2.53	1.66	1.97	2.20	2.37	2.54
	0.95	2.19	2.54	2.85	3.10	3.24	2.22	2.56	2.78	2.94	3.15
	0.975	2.88	3.25	3.41	3.64	3.79	2.78	3.15	3.44	3.66	3.82
	0.99	3.64	4.01	4.22	4.37	4.81	3.66	4.04	4.20	4.43	4.56
0.25	0.90	1.29	1.62	1.87	2.08	2.24	1.27	1.62	1.85	2.01	2.15
	0.95	1.89	2.27	2.45	2.61	2.73	1.88	2.19	2.41	2.61	2.75
	0.975	2.45	2.73	2.98	3.17	3.42	2.41	2.75	3.04	3.27	3.42
	0.99	3.17	3.61	4.01	4.13	4.34	3.27	3.59	3.93	4.07	4.25
Model 2											
0.01	0.90	3.34	3.70	3.97	4.19	4.38	3.52	3.86	4.11	4.34	4.52
	0.95	3.99	4.41	4.73	4.96	5.20	4.13	4.53	4.83	4.99	5.20
	0.975	4.74	5.21	5.39	5.53	5.88	4.84	5.20	5.42	5.59	5.72
	0.99	5.53	6.05	6.28	6.60	6.82	5.59	5.94	6.20	6.73	7.10
0.05	0.90	3.20	3.57	3.84	4.08	4.25	3.36	3.70	3.97	4.14	4.33
	0.95	3.87	4.27	4.56	4.74	4.94	4.02	4.37	4.67	4.87	5.02
	0.975	4.59	4.94	5.23	5.41	5.57	4.69	4.02	5.31	5.58	5.74
	0.99	5.41	5.81	6.13	6.34	6.76	5.58	5.97	6.16	6.30	6.52
0.10	0.90	2.96	3.37	3.64	3.87	4.13	3.26	3.60	3.83	3.99	4.10
	0.95	3.67	4.15	4.37	4.56	4.67	3.85	4.15	4.38	4.57	4.72
	0.975	4.39	4.67	5.06	5.21	5.45	4.39	4.74	5.00	5.15	5.34
	0.99	5.21	5.65	5.92	6.10	6.59	5.15	5.65	5.84	6.08	6.19
0.15	0.90	2.91	3.34	3.60	3.86	4.03	3.09	3.44	3.64	3.84	3.99
	0.95	3.63	4.06	4.34	4.59	4.79	3.66	4.00	4.28	4.61	4.73
	0.975	4.38	4.79	5.05	5.28	5.46	4.30	4.73	4.98	5.17	5.43
	0.99	5.28	5.70	5.83	5.98	6.22	5.17	5.57	5.93	6.07	6.15
0.25	0.90	2.54	2.88	3.16	3.36	3.57	2.72	3.05	3.31	3.52	3.69
	0.95	3.17	3.58	3.88	4.08	4.33	3.34	3.71	3.96	4.13	4.29
	0.975	3.90	4.35	4.57	4.77	5.00	3.99	4.30	4.55	4.80	4.95
	0.99	4.77	5.22	5.50	5.81	6.00	4.80	5.21	5.39	5.63	5.73

THEOREM 2. Let m be the true number of breaks and $\hat{m}$ be the number of breaks obtained using the sequential procedure based on the test statistic $F_T(l+1|l)$ applied with some size a_T . Consider a sequence of critical values $cv = cT^{1-\varepsilon}$, $0 < \varepsilon < 1$ so that a_T converges to zero while ensuring that $F_T(l+1|l)$ remains consistent. Then $P(\hat{m} = m) \rightarrow 1$ as $T \rightarrow \infty$.

3.2. The general case

We now provide an extension of the previous analysis to the case where u_t is allowed to have the following more general structure:

$$\begin{aligned} u_t &= \alpha u_{t-1} + v_t \\ v_t &= d(L)e_t \end{aligned} \quad (8)$$

for $t = 1, \dots, T$ with $d(L) = \sum_{i=0}^{\infty} d_i L^i$, $e_t \sim \text{i.i.d. } (0, \sigma^2)$, e_t being defined on the set of integers. We make the following assumption regarding the polynomial $d(L)$:

ASSUMPTION 1. We assume that $d(z) \neq 0$ for all $|z| \leq 1$ and $\sum_{i=0}^{\infty} i|d_i| < \infty$.

These restrictions on the polynomial $d(L)$ ensure the invertibility and weak stationarity of v_t (see Chang and Park, 2002 for a detailed discussion of these assumptions). Under these conditions, u_t has an autoregressive representation, say $A(L)u_t = e_t$, where $A(L) = 1 - \sum_{i=1}^{\infty} a_i L^i$. In (8), we wish to have α represent the sum of the autoregressive coefficients, $\sum_{i=1}^{\infty} a_i$. Accordingly, we consider the representation

$$u_t = \alpha u_{t-1} + \sum_{j=1}^{\infty} a_j^* \Delta u_{t-j} + e_t, \quad (9)$$

where $a_j^* = -\sum_{i=j+1}^{\infty} a_i$. The infinite order autoregression in Δu_{t-j} is approximated by a truncated version whose order k_T is a function of the number of observations, T :

$$u_t = \alpha u_{t-1} + \sum_{j=1}^{k_T} a_j^* \Delta u_{t-j} + e_{tk},$$

where $e_{tk} = \sum_{j>k_T} a_j^* \Delta u_{t-j} + e_t$. Let $\hat{u}_t^{(i)}$ be the residuals from estimating (5) by OLS. Then the estimate of α considered is the OLS estimate $\hat{\alpha}^{(i)}$ obtained from the regression

$$\hat{u}_t^{(i)} = \alpha^{(i)} \hat{u}_{t-1}^{(i)} + \sum_{j=1}^{k_T} \hat{\varsigma}_j^{(i)} \Delta \hat{u}_{t-j}^{(i)} + e_{tk}^{(i)}, \quad t = \hat{T}_{i-1} + k_T + 2, \dots, \hat{T}_i. \quad (10)$$

The order k_T of the autoregression (10) is assumed to satisfy both an upper bound and a lower bound condition. This is stated in Assumption 2:

ASSUMPTION 2. The order k_T is assumed to satisfy (as $T \rightarrow \infty$) (a) (upper bound condition) $k_T^2/T \rightarrow 0$ and (b) (lower bound condition) $k_T \sum_{i>k_T} |a_i^*| \rightarrow 0$.

The upper bound condition is the same as that used by Chang and Park (2002) in the context of unit root tests and is weaker than the assumption $k_T = o_p(T^{1/3})$ made in Ng and Perron (1995). Note that the lower bound condition allows a logarithmic rate of increase of k_T thereby permitting the use of data-dependent rules such as information criteria for selecting k_T in practice. PY recommend using the Bayesian information criterion (BIC) for choosing k_T . Again a bias-correction is applied and the super-efficient estimate $\hat{\alpha}^{(i)}$ is constructed as in (3) and used in the feasible GLS regression (6).

The specific form of the Wald test depends on the nature of the errors, $I(0)$ or $I(1)$, and the model. Consider first the $I(0)$ case. For both models, we need to simply replace s_f^2 in (7) by $\hat{h}_v^{(i)}$, an estimate of $(2\pi \text{ times})$ the spectral density function at frequency zero of $v_t = (1 - \alpha L)u_t$. PY propose using an estimator based on a weighted sum of autocovariances using the quadratic spectral kernel and the bandwidth selected according to the plug-in method advocated by Andrews (1991) using an AR(1) approximation. The estimator is

$$\hat{h}_v^{(i)} = (\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} \{\hat{v}_t^{(i)}\}^2 + (\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{j=1}^{\hat{T}_i - \hat{T}_{i-1} - 1} w(j/l_T) \sum_{t=\hat{T}_{i-1}+j+1}^{\hat{T}_i} \hat{v}_t^{(i)} \hat{v}_{t-j}^{(i)}$$

where $\hat{v}_t^{(i)}$ are the residuals from the regression analogous to (6) for the general case. The kernel function $w(\cdot)$ is assumed to satisfy the following:

ASSUMPTION 3. The kernel function $w(\cdot)$ is a continuous and even function with $|w(\cdot)| \leq 1$, $w(0) = 1$, and $\int_{-\infty}^{\infty} w^2(x) < \infty$; and the bandwidth l_T is such that $l_T \rightarrow \infty$ and $l_T = o(T^{1/2})$ as $T \rightarrow \infty$.

Consider now the case where the errors are $I(1)$. For Model 1, the form of the test statistic is the same as in (7), except that we replace s_i^2 by an autoregressive spectral density estimate with the lag length of the autoregression again selected by BIC. The autoregression is

$$\hat{v}_t^{(i)} = \sum_{j=1}^{k_T} \hat{\zeta}_j^{(i)} \hat{v}_{t-j}^{(i)} + \hat{e}_{tk}^{(i)}.$$

Denote the estimate by

$$\hat{\zeta}^{(i)}(L) = (1 - \hat{\zeta}_1^{(i)}L - \dots - \hat{\zeta}_{k_T}^{(i)}L^{k_T}) \quad \text{and} \quad \{\hat{\sigma}_{ek}^{(i)}\}^2 = (\hat{T}_i - \hat{T}_{i-1} - k_T)^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \hat{e}_{tk}^2.$$

Then,

$$\hat{h}_V^{(i)} = \frac{\{\hat{\sigma}_{ek}^{(i)}\}^2}{\{\hat{\zeta}^{(i)}(1)\}^2}.$$

For Model 2, PY propose a modified test statistic, which ensures that the limit distribution is the same as that in the AR(1) case. See PY for details on the modification used.

Given Assumptions 1–3, we are now in a position to state the result for the general case. This is presented in Theorem 3 whose proof can be found in the Appendix:

THEOREM 3. *Suppose that u_t is generated by (8). Then, under Assumptions 1–3, Theorems 1 and 2 remain valid.*

4. SIMULATION EXPERIMENTS

We conduct simulation experiments to assess the finite sample performance of the proposed sequential procedure. We consider cases where the DGP involve either one or two breaks. The sample sizes used are $T = 120, 240, 360$. The level of trimming is set at $\epsilon = 0.15$. The maximum number of allowable breaks is set at three. In all experiments, $\{e_t\}_{t=0}^T$ denotes a sequence of i.i.d. standard normal random variables. The sequence $\{u_t\}_{t=1}^T$ is generated by the ARMA(1,1) model

$$u_t = \alpha u_{t-1} + e_t + \theta e_{t-1}, \quad t = 1, \dots, T \\ u_0 = 0.$$

We consider six values for the autoregressive parameter: $\alpha = 0.5, 0.8, 0.85, 0.9, 0.95, 1$ and five values for the moving average parameter: $\theta = -0.5, -0.3, 0, 0.3, 0.5$. We report results for $\theta = 0$ and $\theta = 0.5$. The results for the other values of θ are qualitatively similar to those for $\theta = 0.5$ and are available upon request. All experiments are based on 1000 replications.

The estimate of the autoregressive parameter is obtained from an autoregression where the lag length is selected using BIC. Regarding the choice of parameters δ and d , we experimented with $\delta = 0.3, 0.4, 0.5, 0.6$ and $d = 0.5, 1, 2$ as in Perron and Yabu (2009b). The results indicated that the choice $\delta = 0.5, d = 1$ works relatively well in finite samples, and therefore we use these values for the construction of the super-efficient estimate. We construct the bias-corrected estimate of the autoregressive parameter using the method of Roy and Fuller (2001) as used in PY. We present our results in terms of the probabilities of selecting a given number of breaks, i.e. $p(\hat{m} = m^*)$ for $m^* = 0, 1, 2, 3$.

4.1. The case with one break

We consider two models, the first involving a break in the slope of the trend only and the second involving a break in both level and slope. The data are generated by

- Model 1 (a single break in slope only): $y_t = \eta DT_t + u_t$,
- Model 2 (a Single break in level and slope): $y_t = \eta(2DU_t + DT_t) + u_t$

where the break date is set to $T_1^0 = T/2$, at mid-sample.

Table 2 presents the probability of break selection corresponding to different values of η for Model 1 when $\theta = 0$. First, when $\alpha = 0.5$ so that the process exhibits only moderate persistence, the procedure selects one break with probability at least 90% irrespective of the magnitude of break and the sample size. When the degree of persistence increases, the probability of underestimation increases, at least for small break sizes. This is expected, given that power of the one break test declines as α approaches 1. The performance of the procedure generally improves as the magnitude of the break increases, mirroring corresponding increases in the power of the single break test. As expected, the probabilities of selecting a single break increase when the sample size is increased. The corresponding probabilities when $\theta = 0.5$ are reported in Table 3. Here the results are broadly similar to those for $\theta = 0$ except that there is a larger probability of underestimation, especially for small break sizes with $T = 120$. This probability, however, diminishes relatively rapidly with increasing break size and/or increasing sample size.

Table 2. Probability of break selection for Model 1 [$m = 1, \theta = 0$]

η	$\alpha = 1$				$\alpha = 0.95$				$\alpha = 0.90$				$\alpha = 0.85$				$\alpha = 0.80$				$\alpha = 0.50$			
	m^*				m^*				m^*				m^*				m^*				m^*			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$T = 120$																								
0.5	0.317	0.593	0.045	0.045	0.265	0.650	0.044	0.041	0.231	0.714	0.032	0.023	0.213	0.732	0.034	0.021	0.161	0.776	0.044	0.019	0.000	0.920	0.060	0.020
0.6	0.141	0.741	0.071	0.047	0.058	0.839	0.054	0.049	0.019	0.888	0.053	0.040	0.005	0.914	0.050	0.031	0.005	0.914	0.053	0.028	0.000	0.914	0.052	0.034
0.7	0.067	0.809	0.058	0.066	0.012	0.885	0.056	0.047	0.000	0.914	0.050	0.036	0.000	0.917	0.048	0.035	0.000	0.911	0.052	0.037	0.000	0.903	0.064	0.033
0.8	0.022	0.838	0.068	0.072	0.000	0.897	0.055	0.048	0.000	0.924	0.036	0.040	0.000	0.923	0.040	0.037	0.000	0.927	0.038	0.035	0.000	0.910	0.063	0.027
0.9	0.008	0.863	0.060	0.069	0.000	0.912	0.050	0.038	0.000	0.931	0.043	0.026	0.000	0.931	0.038	0.031	0.000	0.937	0.035	0.028	0.000	0.910	0.062	0.028
1.0	0.000	0.865	0.065	0.070	0.000	0.913	0.039	0.048	0.000	0.919	0.050	0.031	0.000	0.920	0.043	0.037	0.000	0.926	0.043	0.031	0.000	0.899	0.059	0.042
$T = 240$																								
0.5	0.064	0.874	0.044	0.018	0.000	0.957	0.033	0.010	0.000	0.955	0.042	0.003	0.000	0.942	0.051	0.007	0.000	0.933	0.056	0.011	0.000	0.916	0.083	0.001
0.6	0.009	0.927	0.054	0.010	0.000	0.963	0.031	0.006	0.000	0.963	0.026	0.011	0.000	0.949	0.047	0.004	0.000	0.934	0.060	0.006	0.000	0.920	0.077	0.003
0.7	0.000	0.923	0.070	0.007	0.000	0.959	0.033	0.008	0.000	0.958	0.037	0.005	0.000	0.952	0.042	0.006	0.000	0.936	0.059	0.005	0.000	0.912	0.086	0.002
0.8	0.001	0.915	0.067	0.017	0.000	0.962	0.030	0.008	0.000	0.945	0.047	0.008	0.000	0.940	0.057	0.003	0.000	0.932	0.064	0.004	0.000	0.919	0.077	0.004
0.9	0.000	0.920	0.058	0.022	0.000	0.966	0.024	0.010	0.000	0.970	0.023	0.007	0.000	0.953	0.037	0.010	0.000	0.938	0.051	0.011	0.000	0.915	0.080	0.005
1.0	0.000	0.907	0.076	0.017	0.000	0.964	0.025	0.011	0.000	0.965	0.027	0.008	0.000	0.958	0.034	0.008	0.000	0.945	0.043	0.012	0.000	0.917	0.075	0.008
$T = 360$																								
0.5	0.012	0.942	0.034	0.012	0.000	0.967	0.021	0.012	0.000	0.954	0.040	0.006	0.000	0.934	0.057	0.004	0.000	0.919	0.070	0.011	0.000	0.918	0.075	0.007
0.6	0.002	0.930	0.054	0.014	0.000	0.968	0.025	0.007	0.000	0.962	0.033	0.005	0.000	0.939	0.057	0.004	0.000	0.928	0.062	0.010	0.000	0.934	0.060	0.006
0.7	0.000	0.939	0.046	0.015	0.000	0.969	0.029	0.002	0.000	0.950	0.045	0.005	0.000	0.937	0.054	0.009	0.000	0.919	0.069	0.012	0.000	0.925	0.066	0.009
0.8	0.000	0.926	0.062	0.018	0.000	0.960	0.031	0.009	0.000	0.951	0.036	0.013	0.000	0.941	0.052	0.007	0.000	0.927	0.061	0.012	0.000	0.929	0.063	0.008
0.9	0.000	0.927	0.058	0.015	0.000	0.972	0.020	0.008	0.000	0.963	0.025	0.012	0.000	0.946	0.044	0.010	0.000	0.928	0.063	0.009	0.000	0.932	0.062	0.006
1.0	0.000	0.947	0.049	0.004	0.000	0.975	0.021	0.004	0.000	0.963	0.029	0.008	0.000	0.947	0.038	0.015	0.000	0.930	0.055	0.015	0.000	0.929	0.062	0.009

η	$\alpha = 1$				$\alpha = 0.95$				$\alpha = 0.90$				$\alpha = 0.85$				$\alpha = 0.80$				$\alpha = 0.50$			
	m^*				m^*				m^*				m^*				m^*				m^*			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$T = 120$																								
0.5	0.482	0.427	0.032	0.059	0.455	0.459	0.034	0.052	0.422	0.521	0.028	0.029	0.349	0.610	0.027	0.014	0.228	0.733	0.022	0.017	0.011	0.952	0.016	0.021
0.6	0.314	0.560	0.040	0.086	0.310	0.614	0.031	0.045	0.233	0.708	0.029	0.030	0.155	0.807	0.017	0.021	0.078	0.873	0.025	0.024	0.001	0.961	0.019	0.019
0.7	0.190	0.648	0.050	0.112	0.133	0.761	0.032	0.074	0.080	0.859	0.027	0.034	0.051	0.898	0.023	0.028	0.022	0.930	0.021	0.027	0.000	0.963	0.021	0.016
0.8	0.137	0.678	0.072	0.113	0.053	0.808	0.062	0.077	0.026	0.903	0.022	0.049	0.012	0.922	0.038	0.028	0.006	0.952	0.019	0.023	0.000	0.953	0.023	0.024
0.9	0.079	0.730	0.075	0.116	0.019	0.854	0.047	0.080	0.003	0.930	0.016	0.051	0.001	0.923	0.040	0.036	0.001	0.945	0.033	0.021	0.000	0.963	0.019	0.018
1.0	0.030	0.777	0.084	0.109	0.005	0.853	0.055	0.087	0.000	0.911	0.051	0.038	0.000	0.938	0.032	0.030	0.000	0.953	0.028	0.019	0.000	0.955	0.021	0.024
$T = 240$																								
0.5	0.268	0.651	0.052	0.029	0.114	0.845	0.035	0.006	0.033	0.948	0.016	0.003	0.005	0.969	0.022	0.004	0.001	0.973	0.021	0.005	0.000	0.970	0.030	0.000
0.6	0.125	0.782	0.060	0.033	0.007	0.946	0.040	0.007	0.001	0.975	0.019	0.005	0.000	0.970	0.027	0.003	0.000	0.974	0.025	0.001	0.000	0.971	0.028	0.001
0.7	0.039	0.857	0.068	0.036	0.000	0.959	0.033	0.008	0.000	0.976	0.016	0.008	0.000	0.963	0.033	0.004	0.000	0.967	0.027	0.006	0.000	0.963	0.035	0.002
0.8	0.008	0.895	0.073	0.024	0.000	0.967	0.020	0.013	0.000	0.972	0.024	0.004	0.000	0.974	0.023	0.003	0.000	0.975	0.024	0.001	0.000	0.970	0.027	0.003
0.9	0.001	0.889	0.073	0.037	0.000	0.949	0.035	0.016	0.000	0.973	0.023	0.004	0.000	0.975	0.022	0.003	0.000	0.975	0.022	0.003	0.000	0.977	0.022	0.001
1.0	0.000	0.858	0.094	0.048	0.000	0.960	0.031	0.009	0.000	0.975	0.020	0.005												

η	$\alpha = 1$				$\alpha = 0.95$				$\alpha = 0.90$				$\alpha = 0.85$				$\alpha = 0.80$				$\alpha = 0.50$			
	m^*				m^*				m^*				m^*				m^*				m^*			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$T = 120$																								
0.5	0.317	0.493	0.054	0.136	0.324	0.496	0.057	0.123	0.330	0.530	0.041	0.099	0.319	0.536	0.041	0.104	0.274	0.576	0.048	0.102	0.000	0.798	0.076	0.126
0.6	0.174	0.577	0.065	0.184	0.112	0.668	0.053	0.167	0.070	0.720	0.073	0.137	0.045	0.765	0.061	0.129	0.031	0.7666	0.063	0.140	0.000	0.788	0.059	0.153
0.7	0.077	0.696	0.075	0.152	0.020	0.761	0.065	0.154	0.006	0.782	0.057	0.155	0.001	0.776	0.062	0.161	0.000	0.782	0.065	0.153	0.000	0.787	0.079	0.134
0.8	0.030	0.736	0.062	0.172	0.001	0.765	0.072	0.162	0.000	0.805	0.058	0.137	0.000	0.819	0.065	0.116	0.000	0.826	0.060	0.114	0.000	0.777	0.065	0.158
0.9	0.007	0.745	0.083	0.165	0.000	0.774	0.074	0.152	0.000	0.805	0.049	0.146	0.000	0.788	0.052	0.160	0.000	0.811	0.050	0.139	0.000	0.806	0.060	0.134
1.0	0.004	0.724	0.079	0.193	0.000	0.769	0.074	0.157	0.000	0.794	0.060	0.146	0.000	0.777	0.063	0.160	0.000	0.786	0.051	0.163	0.000	0.784	0.050	0.166
$T = 240$																								
0.5	0.096	0.770	0.089	0.045	0.000	0.879	0.080	0.041	0.000	0.904	0.066	0.030	0.000	0.886	0.076	0.038	0.000	0.880	0.078	0.042	0.000	0.875	0.088	0.037
0.6	0.020	0.830	0.094	0.056	0.000	0.911	0.054	0.035	0.000	0.896	0.062	0.042	0.000	0.901	0.066	0.033	0.000	0.878	0.082	0.040	0.000	0.880	0.087	0.033
0.7	0.005	0.842	0.097	0.056	0.000	0.893	0.068	0.039	0.000	0.916	0.056	0.028	0.000	0.886	0.072	0.042	0.000	0.884	0.075	0.041	0.000	0.873	0.095	0.032
0.8	0.001	0.0854	0.094	0.051	0.000	0.909	0.061	0.030	0.000	0.908	0.058	0.034	0.000	0.883	0.082	0.035	0.000	0.885	0.077	0.038	0.000	0.884	0.086	0.030
0.9	0.000	0.850	0.091	0.059	0.000	0.895	0.066	0.039	0.000	0.907	0.054	0.039	0.000	0.911	0.061	0.028	0.000	0.898	0.067	0.035	0.000	0.862	0.106	0.032
1.0	0.000	0.841	0.100	0.059	0.000	0.904	0.053	0.043	0.000	0.906	0.056	0.0												

η	$\alpha = 1$				$\alpha = 0.95$				$\alpha = 0.90$				$\alpha = 0.85$				$\alpha = 0.80$				$\alpha = 0.50$			
	m^*				m^*				m^*				m^*				m^*				m^*			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$T = 120$																								
0.5	0.478	0.356	0.058	0.108	0.466	0.381	0.049	0.104	0.453	0.436	0.033	0.078	0.391	0.498	0.027	0.084	0.260	0.617	0.034	0.089	0.015	0.830	0.056	0.099
0.6	0.314	0.465	0.075	0.146	0.326	0.502	0.050	0.122	0.258	0.607	0.050	0.085	0.179	0.690	0.035	0.096	0.109	0.761	0.040	0.090	0.006	0.847	0.040	0.107
0.7	0.191	0.563	0.065	0.181	0.143	0.639	0.063	0.155	0.108	0.714	0.050	0.128	0.081	0.776	0.039	0.104	0.034	0.795	0.053	0.118	0.000	0.849	0.046	0.105
0.8	0.130	0.606	0.076	0.188	0.065	0.711	0.092	0.132	0.043	0.802	0.054	0.101	0.017	0.817	0.056	0.110	0.012	0.838	0.052	0.098	0.000	0.828	0.053	0.119
0.9	0.075	0.633	0.095	0.197	0.029	0.727	0.055	0.189	0.004	0.820	0.051	0.125	0.005	0.844	0.043	0.108	0.002	0.845	0.042	0.111	0.000	0.860	0.044	0.096
1.0	0.032	0.661	0.095	0.212	0.010	0.733	0.072	0.185	0.001	0.826	0.046	0.127	0.000	0.855	0.045	0.100	0.000	0.868	0.049	0.083	0.000	0.848	0.044	0.108
$T = 240$																								
0.5	0.328	0.558	0.069	0.045	0.237	0.695	0.051	0.017	0.115	0.843	0.026	0.016	0.035	0.903	0.046	0.016	0.006	0.933	0.043	0.018	0.000	0.941	0.044	0.015
0.6	0.172	0.684	0.106	0.038	0.048	0.867	0.056	0.029	0.015	0.932	0.039	0.014	0.002	0.952	0.030	0.016	0.000	0.935	0.041	0.024	0.000	0.924	0.050	0.026
0.7	0.058	0.803	0.092	0.047	0.000	0.906	0.058	0.036	0.000	0.948	0.036	0.016	0.000	0.945	0.039	0.016	0.000	0.922	0.058	0.020	0.000	0.935	0.045	0.020
0.8	0.011	0.815	0.122	0.052	0.000	0.927	0.051	0.022	0.000	0.950	0.030	0.020	0.000	0.948	0.037	0.015	0.000	0.931	0.048	0.021	0.000	0.935	0.044	0.021
0.9	0.002	0.827	0.112	0.059	0.000	0.919	0.051	0.030	0.000	0.949	0.034	0.017	0.000	0.945	0.042	0.013	0.000	0.940	0.046	0.014	0.000	0.942	0.041	0.017
1.0	0.002	0.792	0.133	0.073	0.000	0.933	0.041	0.026	0.000	0.939	0.041	0.020												

η	$\alpha = 1$				$\alpha = 0.95$				$\alpha = 0.90$				$\alpha = 0.85$				$\alpha = 0.80$				$\alpha = 0.50$			
	m^*				m^*				m^*				m^*				m^*				m^*			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$T = 120$																								
0.5	0.000	0.589	0.292	0.119	0.000	0.604	0.288	0.108	0.000	0.625	0.295	0.080	0.000	0.616	0.329	0.055	0.000	0.575	0.371	0.054	0.000	0.028	0.884	0.088
0.6	0.000	0.433	0.430	0.137	0.000	0.444	0.431	0.125	0.000	0.437	0.457	0.106	0.000	0.418	0.495	0.087	0.000	0.397	0.519	0.084	0.000	0.023	0.871	0.106
0.7	0.000	0.332	0.516	0.152	0.000	0.295	0.555	0.150	0.000	0.274	0.578	0.148	0.000	0.257	0.616	0.127	0.000	0.231	0.651	0.118	0.000	0.012	0.875	0.113
0.8	0.000	0.233	0.599	0.168	0.000	0.191	0.654	0.155	0.000	0.157	0.706	0.137	0.000	0.119	0.752	0.129	0.000	0.095	0.787	0.118	0.000	0.009	0.898	0.093
0.9	0.000	0.164	0.667	0.169	0.000	0.128	0.716	0.156	0.000	0.066	0.789	0.145	0.000	0.061	0.805	0.134	0.000	0.047	0.838	0.115	0.000	0.004	0.900	0.096
1.0	0.000	0.098	0.720	0.182	0.000	0.070	0.759	0.171	0.000	0.046	0.802	0.152	0.000	0.027	0.843	0.130	0.000	0.021	0.862	0.117	0.000	0.000	0.899	0.101
$T = 240$																								
0.5	0.000	0.297	0.639	0.064	0.000	0.229	0.720	0.051	0.000	0.180	0.777	0.043	0.000	0.125	0.826	0.049	0.000	0.070	0.876	0.054	0.000	0.000	0.950	0.050
0.6	0.000	0.145	0.781	0.074	0.000	0.056	0.893	0.051	0.000	0.021	0.925	0.054	0.000	0.012	0.936	0.052	0.000	0.003	0.939	0.058	0.000	0.000	0.944	0.056
0.7	0.000	0.060	0.864	0.076	0.000	0.011	0.927	0.062	0.000	0.005	0.934	0.061	0.000	0.000	0.951	0.049	0.000	0.000	0.950	0.050	0.000	0.000	0.950	0.050
0.8	0.000	0.014	0.905	0.081	0.000	0.001	0.935	0.064	0.000	0.000	0.939	0.061	0.000	0.000	0.948	0.052	0.000	0.000	0.951	0.049	0.000	0.000	0.957	0.043
0.9	0.000	0.003	0.906	0.091	0.000	0.000	0.948	0.052	0.000	0.000	0.949	0.051	0.000	0.000	0.943	0.057	0.000	0.000	0.946	0.054	0.000	0.000	0.954	0.046
1.0	0.000	0.001	0.913	0.086	0.000	0.000	0.944	0.056	0.000	0.000	0.953	0.047												

Table 7. Probability of break selection for Model 2 [$m = 2, \theta = 0$]

η	$\alpha = 1$						$\alpha = 0.95$						$\alpha = 0.90$						$\alpha = 0.85$						$\alpha = 0.80$						$\alpha = 0.50$					
	m^*						m^*						m^*						m^*						m^*						m^*					
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3				
$T = 120$																																				
0.5	0.000	0.000	0.487	0.513	0.000	0.000	0.507	0.492	0.000	0.000	0.568	0.432	0.000	0.000	0.598	0.402	0.000	0.000	0.630	0.370	0.000	0.000	0.653	0.347	0.000	0.000	0.630	0.370	0.000	0.000	0.653	0.347				
0.6	0.000	0.003	0.484	0.513	0.000	0.000	0.545	0.455	0.000	0.000	0.587	0.413	0.000	0.000	0.613	0.387	0.000	0.000	0.626	0.374	0.000	0.000	0.630	0.370	0.000	0.000	0.630	0.374	0.000	0.000	0.630	0.370				
0.7	0.000	0.001	0.511	0.488	0.000	0.001	0.503	0.496	0.000	0.000	0.545	0.455	0.000	0.001	0.580	0.419	0.000	0.001	0.596	0.403	0.000	0.001	0.616	0.384	0.000	0.000	0.616	0.403	0.000	0.000	0.616	0.384				
0.8	0.000	0.000	0.521	0.479	0.000	0.001	0.541	0.458	0.000	0.001	0.580	0.419	0.000	0.000	0.609	0.391	0.000	0.000	0.640	0.360	0.000	0.000	0.649	0.351	0.000	0.000	0.649	0.360	0.000	0.000	0.649	0.351				
0.9	0.000	0.000	0.493	0.507	0.000	0.002	0.522	0.476	0.000	0.000	0.572	0.428	0.000	0.000	0.581	0.417	0.000	0.002	0.608	0.392	0.000	0.000	0.656	0.344	0.000	0.000	0.656	0.392	0.000	0.000	0.656	0.344				
1.0	0.000	0.000	0.507	0.493	0.000	0.001	0.546	0.453	0.000	0.001	0.583	0.416	0.000	0.000	0.608	0.391	0.000	0.001	0.619	0.380	0.000	0.001	0.666	0.334	0.000	0.000	0.666	0.380	0.000	0.000	0.666	0.334				
$T = 240$																																				
0.5	0.000	0.000	0.613	0.387	0.000	0.000	0.711	0.289	0.000	0.000	0.806	0.194	0.000	0.000	0.822	0.178	0.000	0.000	0.840	0.160	0.000	0.000	0.873	0.127	0.000	0.000	0.873	0.160	0.000	0.000	0.873	0.127				
0.6	0.000	0.000	0.674	0.326	0.000	0.000	0.744	0.256	0.000	0.001	0.800	0.199	0.000	0.000	0.851	0.149	0.000	0.000	0.856	0.144	0.000	0.000	0.872	0.128	0.000	0.000	0.872	0.144	0.000	0.000	0.872	0.128				
0.7	0.000	0.000	0.671	0.329	0.000	0.000	0.769	0.231	0.000	0.000	0.819	0.181	0.000	0.000	0.833	0.167	0.000	0.000	0.859	0.141	0.000	0.000	0.859	0.141	0.000	0.000	0.859	0.141	0.000	0.000	0.859	0.141				
0.8	0.000	0.000	0.703	0.297	0.000	0.000	0.787	0.213	0.000	0.000	0.841	0.159	0.000	0.000	0.861	0.139	0.000	0.000	0.872	0.128	0.000	0.000	0.873	0.127	0.000	0.000	0.873	0.128	0.000	0.000	0.873	0.127				
0.9	0.000	0.000	0.749	0.251	0.000	0.000	0.789	0.211	0.000	0.000	0.839	0.161	0.000	0.000	0.857	0.143	0.000	0.000	0.870	0.130	0.000	0.000	0.854	0.146	0.000	0.000	0.854	0.130	0.000	0.000	0.854	0.146				
1.0	0.000	0.000	0.727	0.273	0.000	0.000	0.806	0.194	0.000	0.000	0.856	0.144	0.000	0.000	0.872	0.128	0.000	0.000	0.875	0.125	0.000	0.000	0.877	0.123	0.000	0.000	0.877	0.125	0.000	0.000	0.877	0.123				
$T = 360$																																				
0.5	0.000	0.000	0.969	0.304	0.000	0.000	0.842	0.158	0.000	0.000	0.901	0.099	0.000	0.000	0.901	0.099	0.000	0.000	0.897	0.103	0.000	0.000	0.905	0.095	0.000	0.000	0.905	0.103	0.000	0.000	0.905	0.095				
0.6	0.000	0.000	0.743	0.257	0.000	0.000	0.860	0.140	0.000	0.000	0.885	0.115	0.000	0.000	0.891	0.109	0.000	0.000	0.896	0.104	0.000	0.000	0.915	0.085	0.000	0.000	0.915	0.104	0.000	0.000	0.915	0.085				
0.7	0.000	0.000	0.803	0.197	0.000	0.000	0.860	0.140	0.000	0.000	0.921	0.079	0.000	0.000	0.918	0.082	0.000	0.000	0.911	0.089	0.000	0.000	0.903	0.097	0.000	0.000	0.903	0.089	0.000	0.000	0.903	0.097				
0.8	0.000	0.000	0.815	0.185	0.000	0.000	0.865	0.135	0.000	0.000	0.921	0.079	0.000	0.000	0.921	0.079	0.000	0.000	0.901	0.099	0.000	0.000	0.905	0.095	0.000	0.000	0.905	0.099	0.000	0.000	0.905	0.095				
0.9	0.000	0.000	0.808	0.192	0.000	0.000	0.898	0.102	0.000	0.000	0.904	0.096	0.000	0.000	0.895	0.105	0.000	0.000	0.904	0.096	0.000	0.000	0.905	0.095	0.000	0.000	0.905	0.096	0.000	0.000	0.905	0.095				
1.0	0.000	0.000	0.822	0.178	0.000	0.000	0.883	0.117	0.000	0.000	0.897	0.103	0.000	0.000	0.894	0.106	0.000	0.000	0.883	0.117	0.000	0.000	0.902	0.098	0.000	0.000	0.902	0.106	0.000	0.000	0.902	0.098				

Table 8. Probability of break selection for Model 1 [$m = 2$, $\theta = 0.5$]

η	$\alpha = 1$				$\alpha = 0.95$				$\alpha = 0.90$				$\alpha = 0.85$				$\alpha = 0.80$				$\alpha = 0.50$			
	m^*				m^*				m^*				m^*				m^*				m^*			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$T = 120$																								
0.5	0.002	0.592	0.240	0.166	0.000	0.600	0.245	0.155	0.000	0.608	0.275	0.117	0.000	0.611	0.310	0.079	0.000	0.551	0.364	0.085	0.000	0.193	0.741	0.066
0.6	0.004	0.500	0.301	0.195	0.000	0.516	0.296	0.188	0.000	0.507	0.354	0.139	0.000	0.479	0.422	0.099	0.000	0.447	0.447	0.106	0.000	0.145	0.781	0.074
0.7	0.000	0.430	0.351	0.219	0.000	0.443	0.378	0.179	0.000	0.389	0.449	0.162	0.000	0.329	0.530	0.141	0.000	0.260	0.623	0.117	0.000	0.073	0.833	0.094
0.8	0.001	0.351	0.430	0.218	0.000	0.345	0.446	0.209	0.000	0.293	0.531	0.176	0.000	0.222	0.640	0.138	0.000	0.175	0.693	0.132	0.000	0.046	0.856	0.098
0.9	0.000	0.283	0.469	0.248	0.000	0.283	0.508	0.209	0.000	0.188	0.621	0.191	0.000	0.150	0.702	0.148	0.000	0.086	0.778	0.136	0.000	0.014	0.894	0.092
1.0	0.000	0.253	0.498	0.249	0.000	0.218	0.557	0.225	0.000	0.138	0.687	0.175	0.000	0.084	0.758	0.158	0.000	0.057	0.799	0.144	0.000	0.004	0.892	0.104
$T = 240$																								
0.5	0.000	0.493	0.396	0.111	0.000	0.467	0.473	0.060	0.000	0.350	0.606	0.044	0.000	0.233	0.727	0.040	0.000	0.127	0.836	0.037	0.000	0.000	0.956	0.044
0.6	0.000	0.351	0.540	0.109	0.000	0.292	0.646	0.062	0.000	0.181	0.768	0.051	0.000	0.071	0.894	0.035	0.000	0.025	0.930	0.045	0.000	0.000	0.960	0.040
0.7	0.000	0.253	0.616	0.131	0.000	0.125	0.781	0.094	0.000	0.064	0.883	0.053	0.000	0.024	0.924	0.052	0.000	0.007	0.936	0.057	0.000	0.000	0.956	0.044
0.8	0.000	0.136	0.740	0.124	0.000	0.073	0.848	0.079	0.000	0.017	0.932	0.051	0.000	0.003	0.951	0.046	0.000	0.001	0.959	0.040	0.000	0.000	0.949	0.051
0.9	0.000	0.110	0.749	0.141	0.000	0.024	0.882	0.094	0.000	0.005	0.924	0.071	0.000	0.000	0.952	0.048	0.000	0.000	0.968	0.032	0.000	0.000	0.959	0.041
1.0	0.000	0.046	0.796	0.158	0.000	0.014	0.854	0.132	0.000	0.000	0.938	0.062	0.000	0.001	0.954	0.045	0.000	0.000	0.966	0.034	0.000	0.000	0.961	0.039
$T = 360$																								
0.5	0.000	0.324	0.603	0.073	0.000	0.255	0.717	0.028	0.000	0.121	0.853	0.026	0.000	0.031	0.940	0.029	0.000	0.000	0.970	0.030	0.000	0.000	0.966	0.034
0.6	0.000	0.220	0.681	0.099	0.000	0.074	0.885	0.041	0.000	0.026	0.942	0.032	0.000	0.000	0.964	0.036	0.000	0.000	0.964	0.036	0.000	0.000	0.966	0.034
0.7	0.000	0.105	0.780	0.115	0.000	0.011	0.934	0.055	0.000	0.000	0.961	0.039	0.000	0.000	0.965	0.035	0.000	0.000	0.965	0.035	0.000	0.000	0.959	0.041
0.8	0.000	0.045	0.828	0.127	0.000	0.000	0.943	0.057	0.000	0.000	0.964	0.036	0.000	0.000	0.959	0.041	0.000	0.000	0.953	0.047	0.000	0.000	0.966	0.034
0.9	0.000	0.020	0.859	0.121	0.000	0.000	0.946	0.054	0.000	0.000	0.961	0.039	0.000	0.000	0.967	0.037	0.000	0.000	0.981	0.019	0.000	0.000	0.959	0.041
1.0	0.000	0.008	0.867	0.125	0.000	0.000	0.956	0.044	0.000	0.000	0.958	0.042	0.000	0.000	0.963	0.037	0.000	0.000	0.961	0.019	0.000	0.000	0.965	0.035

η	$\alpha = 1$				$\alpha = 0.95$				$\alpha = 0.90$				$\alpha = 0.85$				$\alpha = 0.80$				$\alpha = 0.50$			
	m^*				m^*				m^*				m^*				m^*				m^*			
	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$T = 120$																								
0.5	0.000	0.030	0.388	0.582	0.000	0.039	0.417	0.544	0.000	0.029	0.527	0.444	0.000	0.023	0.569	0.408	0.000	0.019	0.614	0.367	0.000	0.118	0.576	0.306
0.6	0.000	0.054	0.384	0.562	0.000	0.049	0.428	0.523	0.000	0.049	0.518	0.433	0.000	0.041	0.560	0.399	0.000	0.020	0.625	0.355	0.000	0.082	0.655	0.263
0.7	0.000	0.057	0.402	0.541	0.000	0.056	0.427	0.517	0.000	0.063	0.492	0.445	0.000	0.046	0.537	0.417	0.000	0.033	0.587	0.380	0.000	0.030	0.656	0.314
0.8	0.000	0.051	0.404	0.545	0.000	0.060	0.450	0.490	0.000	0.063	0.510	0.427	0.000	0.057	0.571	0.372	0.000	0.040	0.611	0.349	0.000	0.025	0.681	0.294
0.9	0.000	0.055	0.420	0.525	0.000	0.059	0.444	0.497	0.000	0.050	0.522	0.428	0.000	0.055	0.578	0.367	0.000	0.040	0.611	0.349	0.000	0.006	0.700	0.294
1.0	0.000	0.053	0.406	0.541	0.000	0.050	0.470	0.480	0.000	0.038	0.541	0.421	0.000	0.050	0.600	0.350	0.000	0.032	0.623	0.345	0.000	0.001	0.677	0.322
$T = 240$																								
0.5	0.000	0.071	0.476	0.453	0.000	0.068	0.623	0.309	0.000	0.032	0.716	0.252	0.000	0.020	0.786	0.194	0.000	0.005	0.870	0.125	0.000	0.000	0.920	0.080
0.6	0.000	0.063	0.517	0.420	0.000	0.058	0.668	0.274	0.000	0.027	0.779	0.194	0.000	0.013	0.838	0.149	0.000	0.002	0.885	0.113	0.000	0.000	0.923	0.077
0.7	0.000	0.043	0.589	0.368	0.000	0.029	0.707	0.264	0.000	0.008	0.822	0.170	0.000	0.002	0.855	0.143	0.000	0.000	0.880	0.120	0.000	0.000	0.933	0.067
0.8	0.000	0.028	0.637	0.335	0.000	0.000	0.016	0.729	0.000	0.001	0.866	0.133	0.000	0.001	0.892	0.107	0.000	0.001	0.903	0.096	0.000	0.000	0.914	0.086
0.9	0.000	0.016	0.655	0.329	0.000	0.004	0.775	0.221	0.000	0.001	0.863	0.136	0.000	0.000	0.910	0.090	0.000	0.000	0.914	0.086	0.000	0.000	0.919	0.081
1.0	0.000	0.009	0.662	0.329	0.000	0.001	0.788	0.211	0.000	0.000	0.866	0.134												

Table 4 presents results for Model 2 with $\theta = 0$. Again, the procedure performs relatively better when $\alpha = 0.5$, although now there is a non-negligible probability of overestimation which increases as α increases. This is due to the fact that the tests suffer from size distortions which become more severe with increases in α and the number of breaks assumed under the null hypothesis. These size inaccuracies persist for $\alpha = 1$ and small break sizes even with the two larger sample sizes. However, as with Model 1, the probabilities of selecting one break are higher relative to those with the smallest sample size. The selection probabilities in Table 5 for $\theta = 0.5$ again demonstrate the higher underestimation probability in the presence of a general error structure.

4.2. The case with two breaks

With two breaks the DGPs considered are the following:

- Model 1 (two breaks in slope only): $y_t = \eta_1 DT_{1t} + \eta_2 DT_{2t} + u_t$,
- Model 2 (two breaks in level and slope):

$$y_t = \eta_1 (DT_{1t} + 2DU_{1t}) + \eta_2 (DT_{2t} + 2DU_{2t}) + u_t.$$

We set $\eta_1 = 1$ and report results for a range of values of η . The dates of the breaks are set at $T_1^0 = T/3$ and $T_2^0 = 2T/3$.

The results for Model 1 with $\theta=0$ are reported in Table 6. With $\alpha = 0.5$, the probabilities of selecting the true number of breaks is close to 90% even with $T = 120$ and small break sizes. However, in contrast to the one break case, these probabilities are reduced as α approached 1. This suggests that the power of the one-versus-two breaks test is low relative to that of the zero-versus-one break test. For $\alpha > 0.5$, there is also a non-negligible probability of overestimation reflecting the size distortions of the two-versus-three breaks test. As for the one break case, the selection probabilities for two breaks increase with the sample size.

Table 7 reports corresponding results for Model 2. As in the one break case, there is a substantial probability of overestimation, especially for values of α close to 1. Noticeably, the probability of underestimation is negligible even for small break sizes and $T = 120$. When the sample size increases to $T = 240$, the probabilities of selecting two breaks increase to about 80–85% for moderate break sizes when $\alpha \leq 0.95$. These probabilities further increase when T is increased to 360. For $\alpha = 1$, the size distortions are still in play even for large magnitudes of the breaks and large sample sizes. For both models, the probabilities reported in Tables 8 and 9 for $\theta = 0.5$ can be interpreted in a manner similar to the one break case.

In summary, the performance of the proposed sequential procedure is qualitatively different for Models 1 and 2. For Model 1, there is a tendency to underestimate the true number of breaks while for Model 2, there is a probability of overestimation. This difference can be traced to the finite sample properties of the tests for these models with low power being more of an issue for Model 1 and size distortions being the dominant factor for Model 2. The power problem is alleviated to a considerable extent for large magnitudes of the breaks while the size distortions in Model 2 remain somewhat of a concern, especially in the presence of strong persistence in the error component, though these concerns are mitigated for larger sample sizes. The simulation results point to the importance of the choice of the maximal value of the number of breaks in relation to the size of the sample available. For example, when testing for two breaks in a sample of size 120, one ends up with fewer than 40 observations per segments. It is then not surprising to see low power and/or size distortions. Hence, practitioners must exercise caution to allow a sufficient number of observations in each segment and chose the maximum number of breaks permissible accordingly.

5. CONCLUSION

Testing whether a time series contains a broken trend is complicated by the fact that we do not have *a priori* knowledge of whether the noise is stationary or integrated. This has motivated the development of tests that are robust to the extent of persistence in the error component. These are designed to evaluate the null hypothesis of no structural change versus the alternative of a single change in trend but do not allow researchers to select the number of changes. Given that selecting the number of breaks is an important practical issue, we attempted to fill a gap in the literature by proposing a sequential procedure that enables consistent estimation of the number of breaks. Monte Carlo evidence demonstrated that the procedure works well in samples sizes that are common in applied work.

APPENDIX

In what follows, $W^{(i)}(i = 1, \dots, l+1)$ denotes a set of $(l+1)$ independent standard Brownian motions on $[0,1]$ and W denotes a standard Brownian motion on $[0,1]$ that is independent of $W^{(i)}$ for all i . Also, $\Lambda_{i,\epsilon} = \{\tau : \hat{\lambda}_{i-1} + (\hat{\lambda}_i - \hat{\lambda}_{i-1})\epsilon \leq \tau \leq \hat{\lambda}_i - (\hat{\lambda}_i - \hat{\lambda}_{i-1})\epsilon\}$ and $\Lambda_\epsilon = \{r : 1 - \epsilon \leq r \leq \epsilon\}$ for some $\epsilon > 0$. We use $\|\cdot\|$ to denote the Euclidean norm, i.e. for a $p \times q$ matrix A , $\|A\| = (\sum_{i=1}^p \sum_{j=1}^q x_{ij}^2)^{1/2}$. Also, we use $\|A\|_1$ to denote the vector-induced norm, i.e. $\|A\|_1 = \sup_{x \neq 0} \|Ax\|/\|x\|$. We shall also use the fact that the estimates of the break fractions are consistent for the true break fractions. As shown in Perron and Zhu (2005): $T^{3/2}(\hat{\lambda}_i - \lambda_i^0) = O_p(1)$ for Model 1 with $I(0)$ errors, $T(\hat{\lambda}_i - \lambda_i^0) = O_p(1)$ for Model 2 with $I(0)$ errors, while $T^{1/2}(\hat{\lambda}_i - \lambda_i^0) = O_p(1)$ for Models 1 and 2 with $I(1)$ errors. Though their proof is for the single-break case, the results continue to hold with multiple breaks (see Oka and Perron, 2009 for the case of $I(0)$ errors).

PROOF OF THEOREM 1. Consider first Model 1. Let $X^{(i)} = \{x_{\hat{T}_{i-1}+1}^{(i)}, (1 - \hat{\alpha}_S^{(i)})x_{\hat{T}_{i-1}+2}^{(i)}, \dots, (1 - \hat{\alpha}_S^{(i)})x_{\hat{T}_i}^{(i)}\}'$ with $x_j^{(i)} (j = 1, 2, 3)$ being the j th column of $X^{(i)}$. For a given $\tau \in A_{i,\tau}$ the Wald test of $\beta_1^{(i)} = 0$ can be expressed as

$$W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i) = \frac{[\hat{\beta}_1^{(i)}]^2 (X_3^{(i)'} M^{(i)} X_3^{(i)})}{S_i^2} \quad (11)$$

with $M^{(i)} = I^{(i)} - Z^{(i)}(Z^{(i)'}Z^{(i)})^{-1}Z^{(i)'}$, $Z^{(i)} = (X_1^{(i)}, X_2^{(i)})$ and S_i^2 the residual error variance from the feasible GLS regression. Denoting $\tilde{U}^{(i)} = \{u_t - \hat{\alpha}_S^{(i)}u_{t-1}\}_{t=\hat{T}_{i-1}+1}^{\hat{T}_i}$, we have, under the null hypothesis of l breaks,

$$\begin{aligned} \hat{\beta}_1^{(i)} &= \frac{X_3^{(i)'} M^{(i)} \tilde{U}^{(i)}}{X_3^{(i)'} M^{(i)} X_3^{(i)}} = \frac{X_3^{(i)'} \tilde{U}^{(i)} - X_3^{(i)'} Z^{(i)} (Z^{(i)'} Z^{(i)})^{-1} Z^{(i)'} \tilde{U}^{(i)}}{X_3^{(i)'} X_3^{(i)} - X_3^{(i)'} Z^{(i)} (Z^{(i)'} Z^{(i)})^{-1} Z^{(i)'} X_3^{(i)}} \\ &= \frac{\begin{bmatrix} r_3^{(i)} - \begin{pmatrix} q_{13}^{(i)} & q_{23}^{(i)} \end{pmatrix} \begin{pmatrix} q_{11}^{(i)} & q_{12}^{(i)} \\ q_{12}^{(i)} & q_{22}^{(i)} \end{pmatrix}^{-1} \begin{pmatrix} r_1^{(i)} \\ r_2^{(i)} \end{pmatrix} \end{bmatrix}}{\begin{bmatrix} q_{33}^{(i)} - \begin{pmatrix} q_{13}^{(i)} & q_{23}^{(i)} \end{pmatrix} \begin{pmatrix} q_{11}^{(i)} & q_{12}^{(i)} \\ q_{12}^{(i)} & q_{22}^{(i)} \end{pmatrix}^{-1} \begin{pmatrix} q_{13}^{(i)} \\ q_{23}^{(i)} \end{pmatrix} \end{bmatrix}} \end{aligned} \quad (12)$$

where

$$\begin{aligned} q_{11}^{(i)} &= 1 + (\hat{T}_i - \hat{T}_{i-1} - 1)(1 - \hat{\alpha}_S^{(i)})^2 \\ q_{12}^{(i)} &= 1 + (1 - \hat{\alpha}_S^{(i)})^2 \sum_{t=\hat{T}_{i-1}+2}^{\hat{T}_i} (t - \hat{T}_{i-1}) + \hat{\alpha}_S^{(i)}(1 - \hat{\alpha}_S^{(i)})(\hat{T}_i - \hat{T}_{i-1} - 1) \\ q_{13}^{(i)} &= (1 - \hat{\alpha}_S^{(i)})^2 \sum_{t=k+1}^{\hat{T}_i} (t - k) + \hat{\alpha}_S^{(i)}(1 - \hat{\alpha}_S^{(i)})(\hat{T}_i - k) \\ q_{22}^{(i)} &= 1 + (1 - \hat{\alpha}_S^{(i)})^2 \sum_{t=\hat{T}_{i-1}+2}^{\hat{T}_i} (t - \hat{T}_{i-1})^2 + [\hat{\alpha}_S^{(i)}]^2 (\hat{T}_i - \hat{T}_{i-1} - 1) + 2\hat{\alpha}_S^{(i)}(1 - \hat{\alpha}_S^{(i)}) \sum_{t=\hat{T}_{i-1}+2}^{\hat{T}_i} (t - \hat{T}_{i-1}) \\ q_{23}^{(i)} &= (1 - \hat{\alpha}_S^{(i)})^2 \sum_{t=k+1}^{\hat{T}_i} (t - \hat{T}_{i-1})(t - k) + \hat{\alpha}_S^{(i)}(1 - \hat{\alpha}_S^{(i)}) \sum_{t=k+1}^{\hat{T}_i} (t - \hat{T}_{i-1}) + \hat{\alpha}_S^{(i)}(1 - \hat{\alpha}_S^{(i)}) \sum_{t=k+1}^{\hat{T}_i} (t - k) + [\hat{\alpha}_S^{(i)}]^2 (\hat{T}_i - k) \\ q_{33}^{(i)} &= (1 - \hat{\alpha}_S^{(i)})^2 \sum_{t=k+1}^{\hat{T}_i} (t - k)^2 + [\hat{\alpha}_S^{(i)}]^2 (\hat{T}_i - k) + 2\hat{\alpha}_S^{(i)}(1 - \hat{\alpha}_S^{(i)}) \sum_{t=k+1}^{\hat{T}_i} (t - k) \\ r_1^{(i)} &= u_{\hat{T}_{i-1}+1} - \hat{\alpha}_S^{(i)}u_{\hat{T}_{i-1}} + (1 - \hat{\alpha}_S^{(i)}) \sum_{t=\hat{T}_{i-1}+2}^{\hat{T}_i} (u_t - \hat{\alpha}_S^{(i)}u_{t-1}) \\ r_2^{(i)} &= u_{\hat{T}_{i-1}+1} - \hat{\alpha}_S^{(i)}u_{\hat{T}_{i-1}} + (1 - \hat{\alpha}_S^{(i)}) \sum_{t=\hat{T}_{i-1}+2}^{\hat{T}_i} (t - \hat{T}_{i-1})(u_t - \hat{\alpha}_S^{(i)}u_{t-1}) + \hat{\alpha}_S^{(i)} \sum_{t=\hat{T}_{i-1}+2}^{\hat{T}_i} (u_t - \hat{\alpha}_S^{(i)}u_{t-1}) \\ r_3^{(i)} &= (1 - \hat{\alpha}_S^{(i)}) \sum_{t=k+1}^{\hat{T}_i} (t - k)(u_t - \hat{\alpha}_S^{(i)}u_{t-1}) + \hat{\alpha}_S^{(i)} \sum_{t=k+1}^{\hat{T}_i} (u_t - \hat{\alpha}_S^{(i)}u_{t-1}) \end{aligned}$$

Next, we derive the limit of each term separately for $|\alpha| < 1$ and $\alpha = 1$.

STATIONARY CASE ($|\alpha| < 1$). We use the fact that

$$\begin{aligned} T^{-1/2} \sum_{t=1}^{[Ts]} (u_t - \hat{\alpha}_S^{(i)}u_{t-1}) &= T^{-1/2} \sum_{t=1}^{[Ts]} [(\alpha - \hat{\alpha}_S^{(i)})u_{t-1} + e_t] \\ &= T^{-1/2} \sum_{t=1}^{[Ts]} e_t - T^{-1/2} [T^{1/2}(\hat{\alpha}_S^{(i)} - \alpha)] T^{-1/2} \sum_{t=1}^{[Ts]} u_{t-1} \\ &= T^{-1/2} \sum_{t=1}^{[Ts]} e_t + o_p(1) \Rightarrow \sigma W(s) \end{aligned}$$

The convergence results for each of the components are then as follows: $T^{-1}q_{11}^{(i)} \xrightarrow{P} (\lambda_i^0 - \lambda_{i-1}^0)(1 - \alpha)^2$, $T^{-2}q_{12}^{(i)} \xrightarrow{P} (1 - \alpha)^2 \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0) ds = (1 - \alpha)^2 (\lambda_i^0 - \lambda_{i-1}^0)^2 / 2$, $T^{-2}q_{13}^{(i)} \xrightarrow{P} (1 - \alpha)^2 (\lambda_i^0 - \tau)^2$, $T^{-3}q_{22}^{(i)} \xrightarrow{P} (1 - \alpha)^2 (\lambda_i^0 - \lambda_{i-1}^0)^3 / 3$, $T^{-3}q_{23}^{(i)} \xrightarrow{P} (1 - \alpha)^2 \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0)(s - \tau) ds$, $T^{-3}q_{33}^{(i)} \xrightarrow{P} (1 - \alpha)^2 \int_{\tau}^{\lambda_i^0} (s - \tau)^2 ds$, $T^{-1/2}r_1^{(i)} \Rightarrow (1 - \alpha)\sigma[W(\lambda_i^0) - W(\lambda_{i-1}^0)]$, $T^{-1/2}r_2^{(i)} \Rightarrow (1 - \alpha)$

$\sigma \int_{\lambda_{i-1}^0}^{\lambda_i^0} [W(\lambda_i^0) - W(s)] ds$ and $T^{-3/2} r_3^{(i)} \Rightarrow (1 - \alpha) \sigma \int_{\tau}^{\lambda_i^0} [W(\lambda_i^0) - W(s)] ds$. Using these results in (11) together with the fact that $s_i^2 \xrightarrow{P} \sigma^2$ for each i , we obtain

$$W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i) \Rightarrow \frac{A_i^2}{B_i} \equiv \zeta_i(\tau)$$

where

$$\begin{aligned} A_i &= \int_{\tau}^{\lambda_i^0} [W(\lambda_i^0) - W(s)] ds - ((\lambda_i^0 - \tau)^2 \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0)(s - \tau) ds) \\ &\quad \times \left(\begin{pmatrix} (\lambda_i^0 - \lambda_{i-1}^0) & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} \\ \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^3}{3} \end{pmatrix}^{-1} \begin{pmatrix} [W(\lambda_i^0) - W(\lambda_{i-1}^0)] \\ \int_{\lambda_{i-1}^0}^{\lambda_i^0} [W(\lambda_i^0) - W(s)] ds \end{pmatrix} \right) \\ B_i &= \int_{\tau}^{\lambda_i^0} (s - \tau)^2 ds - ((\lambda_i^0 - \tau)^2 \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0)(s - \tau) ds) \\ &\quad \times \left(\begin{pmatrix} (\lambda_i^0 - \lambda_{i-1}^0) & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} \\ \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^3}{3} \end{pmatrix}^{-1} \begin{pmatrix} (\lambda_i^0 - \tau)^2 \\ \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0)(s - \tau) ds \end{pmatrix} \right) \end{aligned}$$

Note that the random variables $\xi_1, \dots, \xi_{l+1}$ are independent. This follows since, for $s_i \in [\lambda_{i-1}^0, \lambda_i^0]$, the processes $W(\lambda_1^0) - W(s_1), W(\lambda_2^0) - W(s_2), \dots, W(\lambda_{l+1}^0) - W(s_{l+1})$ are independent. Next, we use the fact that $W(\lambda_i^0) - W(s)$ has the same distribution as

$$\sqrt{\lambda_i^0 - \lambda_{i-1}^0} \left[W^{(i)}(1) - W^{(i)}\left(\frac{s - \lambda_{i-1}^0}{\lambda_i^0 - \lambda_{i-1}^0}\right) \right]$$

for $s \in [\lambda_{i-1}^0, \lambda_i^0]$. Then with the change of variable $r = (\tau - \lambda_{i-1}^0)/(\lambda_i^0 - \lambda_{i-1}^0)$, $\zeta_i(\tau)$ has the same distribution as $\zeta_i^*(r) = [A_i^*]^2/B_i^*$ where

$$\begin{aligned} A_i^* &= \int_r^1 [W^{(i)}(1) - W^{(i)}(s)] ds - \left((1-r)^2 \int_r^1 s(s-r) ds \right) \\ &\quad \times \left(\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} \begin{pmatrix} W^{(i)}(1) \\ \int_0^1 [W^{(i)}(1) - W^{(i)}(s)] ds \end{pmatrix} \right) \\ B_i^* &= \int_r^1 (s-r)^2 ds - \left((1-r)^2 \int_r^1 s(s-r) ds \right) \\ &\quad \times \left(\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} \begin{pmatrix} (1-r)^2 \\ \int_r^1 s(s-r) ds \end{pmatrix} \right) \end{aligned}$$

We then obtain

$$\log \left[(\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{\tau \in \Lambda_{i,\epsilon}} \exp \left(\frac{W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i)}{2} \right) \right] \Rightarrow \log \left[\int_{r \in \Lambda_\epsilon} \exp(\zeta_i^*(r)) \right]$$

which is the limit distribution of the zero-versus-one break test. Using the independence of $\zeta_1^*, \zeta_2^*, \dots, \zeta_{l+1}^*$, the result follows.

UNIT ROOT CASE ($\alpha = 1$). Here, we use the fact that $T(\hat{\alpha}_S^{(i)} - 1) \xrightarrow{P} 0$. Also, we have

$$T^{-1/2} \sum_{t=1}^{[Ts]} (u_t - \hat{\alpha}_S^{(i)} u_{t-1}) = T^{-1/2} \sum_{t=1}^{[Ts]} e_t - T^{-1} [T(\hat{\alpha}_S^{(i)} - 1)] T^{-1/2} \sum_{t=1}^{[Ts]} u_{t-1} \Rightarrow \sigma W(s).$$

Then the convergence results for the components in (12) are: $q_{11}^{(i)} \xrightarrow{P} 1$, $q_{12}^{(i)} \xrightarrow{P} 1$, $q_{13}^{(i)} \xrightarrow{P} 0$, $T^{-1} q_{22}^{(i)} \xrightarrow{P} \lambda_i^0 - \lambda_{i-1}^0$, $T^{-1} q_{23}^{(i)} \rightarrow \lambda_i^0 - \tau$, $T^{-1} q_{33}^{(i)} \xrightarrow{P} \lambda_i^0 - \tau$, $r_1^{(i)} \Rightarrow \lim_{T \rightarrow \infty} e_{[T\lambda_{i-1}^0]+1}$, $T^{-1/2} r_2^{(i)} \Rightarrow \sigma [W(\lambda_i^0) - W(\lambda_{i-1}^0)]$ and $T^{-1/2} r_3^{(i)} \Rightarrow \sigma [W(\lambda_i^0) - W(\tau)]$. Again, using these results in (11) together with the fact that $s_i^2 \xrightarrow{P} \sigma^2$ for each i , we obtain

$$W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i) \Rightarrow \frac{C_i^2}{D_i} \equiv \eta_i(\tau)$$

where

$$\begin{aligned} C_i &= W(\lambda_i^0) - W(\tau) - \frac{(\lambda_i^0 - \tau)[W(\lambda_i^0) - W(\lambda_{i-1}^0)]}{\lambda_i^0 - \lambda_{i-1}^0} \\ D_i &= \frac{(\lambda_i^0 - \tau)(\tau - \lambda_{i-1}^0)}{\lambda_i^0 - \lambda_{i-1}^0} \end{aligned}$$

Again, it is straightforward to verify that $\eta_1, \dots, \eta_{t+1}$ are independent. Then, as in Model 1, we use the fact that $W(\lambda_i^0) - W(s)$ has the same distribution as

$$\sqrt{\lambda_i^0 - \lambda_{i-1}^0} \left[W^{(i)}(1) - W^{(i)}\left(\frac{s - \lambda_{i-1}^0}{\lambda_i^0 - \lambda_{i-1}^0}\right) \right]$$

for $s \in [\lambda_{i-1}^0, \lambda_i^0]$. With the change of variable $r = (\tau - \lambda_{i-1}^0)/(\lambda_i^0 - \lambda_{i-1}^0)$, $\eta_i(\tau)$ has the same distribution as

$$\eta_i^*(r) = \frac{[W^{(i)}(r) - rW^{(i)}(1)]^2}{r(1-r)}.$$

We thus have

$$\log \left[(\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{\tau \in \Lambda_{i,\epsilon}} \exp \left(\frac{W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i)}{2} \right) \right] \Rightarrow \log \left[\int_{r \in \Lambda_i} \exp(\eta_i^*(r)) \right]$$

which is the limit distribution of the zero-versus-one break test. Again, the result follows from the independence of $\eta_1^*, \eta_2^*, \dots, \eta_{t+1}^*$.

Consider now the proof for Model 2. Again, let $X^{(i)} = \{x_{t_{i-1}+1}^{(i)}, (1 - \hat{\alpha}_s^{(i)})x_{t_{i-1}+2}^{(i)}, \dots, (1 - \hat{\alpha}_s^{(i)})x_{t_i}^{(i)}\}'$ with $x_j^{(i)}$ ($j = 1, 2, 3, 4$) being the j th column of $X^{(i)}$. For a given $\tau \in \Lambda_{i,\epsilon}$, the Wald test for testing $\mu_1^{(i)} = \beta_1^{(i)} = 0$ can be expressed as

$$W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i) = [\hat{y}^{(i)}]' \frac{(Z_1^{(i)'} M_2^{(i)} Z_1^{(i)}) [\hat{y}^{(i)}]}{s_i^2} \quad (13)$$

with $\hat{y}^{(i)} = (\hat{\mu}_1^{(i)}, \hat{\beta}_1^{(i)})'$, $Z_1^{(i)} = (X_2^{(i)}, X_4^{(i)})$, $M_2^{(i)} = I^{(i)} - Z_2^{(i)}(Z_2^{(i)'} Z_2^{(i)})^{-1} Z_2^{(i)'}$, $Z_2^{(i)} = (X_1^{(i)}, X_3^{(i)})$ and s_i^2 the residual error variance from the feasible GLS regression.

STATIONARY CASE ($|\alpha| < 1$). Using arguments similar to those for Model 1, we have

$$W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i) \Rightarrow F_i'(E_i)^{-1} F_i \equiv \Phi_i(\tau)$$

where

$$\begin{aligned} E_i &= \begin{pmatrix} \lambda_i^0 - \tau & \int_{\tau}^{\lambda_i^0} (s - \tau) ds \\ \int_{\tau}^{\lambda_i^0} (s - \tau) ds & \int_{\tau}^{\lambda_i^0} (s - \tau)^2 ds \end{pmatrix} - \begin{pmatrix} \lambda_i^0 - \tau & \int_{\tau}^{\lambda_i^0} (s - \tau) ds \\ (\lambda_i^0 - \tau)^2 & \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0)(s - \tau) ds \end{pmatrix} \\ &\quad \times \begin{pmatrix} (\lambda_i^0 - \lambda_{i-1}^0) & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} \\ \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^3}{3} \end{pmatrix}^{-1} \begin{pmatrix} \lambda_i^0 - \tau & (\lambda_i^0 - \tau)^2 \\ \int_{\tau}^{\lambda_i^0} (s - \tau) ds & \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0)(s - \tau) ds \end{pmatrix} \\ F_i &= \begin{pmatrix} [W(\lambda_i^0) - W(\tau)] \\ \int_{\tau}^{\lambda_i^0} [W(\lambda_i^0) - W(s)] ds \end{pmatrix} - \begin{pmatrix} \lambda_i^0 - \tau & \int_{\tau}^{\lambda_i^0} (s - \tau) ds \\ (\lambda_i^0 - \tau)^2 & \int_{\tau}^{\lambda_i^0} (s - \lambda_{i-1}^0)(s - \tau) ds \end{pmatrix} \\ &\quad \times \begin{pmatrix} (\lambda_i^0 - \lambda_{i-1}^0) & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} \\ \frac{(\lambda_i^0 - \lambda_{i-1}^0)^2}{2} & \frac{(\lambda_i^0 - \lambda_{i-1}^0)^3}{3} \end{pmatrix}^{-1} \begin{pmatrix} [W(\lambda_i^0) - W(\lambda_{i-1}^0)] \\ \int_{\lambda_{i-1}^0}^{\lambda_i^0} [W(\lambda_i^0) - W(s)] ds \end{pmatrix} \end{aligned}$$

Applying the same transformations as in Model 1, $\Phi_i(\tau)$ has the same distribution as $\Phi_i^*(r) = F_i^{*'}(E_i^*)^{-1} F_i^*$ where

$$\begin{aligned} E_i^* &= \begin{pmatrix} 1 - r & \int_r^1 (s - r) ds \\ \int_r^1 (s - r) ds & \int_r^1 (s - r)^2 ds \end{pmatrix} - \begin{pmatrix} 1 - r & \int_r^1 (s - r) ds \\ (1 - r)^2 & \int_r^1 s(s - r) ds \end{pmatrix} \\ &\quad \times \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} \begin{pmatrix} 1 - r & (1 - r)^2 \\ \int_r^1 (s - r) ds & \int_r^1 s(s - r) ds \end{pmatrix} \\ F_i^* &= \begin{pmatrix} [W^{(i)}(1) - W^{(i)}(r)] \\ \int_r^1 [W^{(i)}(1) - W^{(i)}(s)] ds \end{pmatrix} - \begin{pmatrix} 1 - r & \int_r^1 (s - r) ds \\ (1 - r)^2 & \int_r^1 s(s - r) ds \end{pmatrix} \\ &\quad \times \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} \begin{pmatrix} W^{(i)}(1) \\ \int_0^1 [W^{(i)}(1) - W^{(i)}(s)] ds \end{pmatrix} \end{aligned}$$

which is the limit distribution of the zero-versus-one break test in a model that allows for a break in intercept as well as the slope. Using the independence of $\Phi_1^*, \dots, \Phi_{t+1}^*$, the result follows.

Unit Root Case ($\alpha = 1$). Derivations similar to those used for Model 1 yield

$$W_{FS}(\hat{\lambda}_{i-1}, \tau, \hat{\lambda}_i) \Rightarrow \frac{[\lim_{T \rightarrow \infty} e_{[T\tau]+1}]^2}{\sigma^2} + \frac{(\lambda_i^0 - \lambda_{i-1}^0)}{(\lambda_i^0 - \tau)(\tau - \lambda_{i-1}^0)} \times \left[W(\lambda_i^0) - W(\tau) - \frac{(\lambda_i^0 - \tau)[W(\lambda_i^0) - W(\lambda_{i-1}^0)]}{(\lambda_i^0 - \lambda_{i-1}^0)} \right]^2. \quad (14)$$

Under the assumption that e_t is i.i.d. normal, for $\tau \in \Lambda_{i,\epsilon}$, $e_{[T\tau]+1}$ is asymptotically independent of both $T^{-1/2} \sum_{t=\tau_{i-1}^0}^{\tau_i^0} e_t$ and $T^{-1/2} \sum_{t=\tau_{i-1}^0}^{\tau_i^0} e_t$ so that the first and second terms in (14) are independent. Then using the same variable and distribution transformations as in Model 1 and the fact that $\lim_{T \rightarrow \infty} e_{[T\tau]+1}$ has the same distribution as

$$\lim_{T \rightarrow \infty} e_{\left\lceil \frac{\tau(\tau - \lambda_{i-1}^0)}{\lambda_i^0 - \lambda_{i-1}^0} \right\rceil + 1},$$

the result of the theorem follows from the independence of the tests over the $(l+1)$ regimes.

PROOF OF THEOREM 3. We will prove the theorem for Model 1. The proof for Model 2 is similar and hence omitted. As in the proof of Theorem 1, we analyze the stationary and unit root cases separately.

STATIONARY CASE ($|\alpha| < 1$). Let $z_t = (\Delta u_{t-1}, \Delta u_{t-2}, \dots, \Delta u_{t-k_T})'$, $b_t = \sum_{j>k_T} \Delta u_{t-j} a_j^*$ and $\varsigma = (a_1^*, a_2^*, \dots, a_{k_T}^*)'$. Here we have

$$\begin{aligned} T^{-1/2} \sum_{t=1}^{[Ts]} (u_t - \hat{\alpha}_S^{(i)} u_{t-1}) &= T^{-1/2} \sum_{t=1}^{[Ts]} \{(\alpha - \hat{\alpha}_S^{(i)}) u_{t-1} + z_t' \varsigma + b_t + e_t\} \\ &= T^{-1/2} \sum_{t=1}^{[Ts]} e_t - T^{-1/2} [T^{1/2} (\hat{\alpha}_S^{(i)} - \alpha)] T^{-1/2} \sum_{t=1}^{[Ts]} u_{t-1} \\ &\quad + \left\{ T^{-1/2} \sum_{t=1}^{[Ts]} z_t' \right\} \varsigma + T^{-1/2} \sum_{t=1}^{[Ts]} b_t. \end{aligned} \quad (15)$$

As in the AR(1) case, the second term in (15) is $o_p(1)$. For the third term, note that

$$\left\| T^{-1/2} \sum_{t=1}^{[Ts]} z_t' \right\| = O_p(k_T^{1/2} T^{-1/2}) = o_p(1).$$

The last term in (15) is

$$\begin{aligned} T^{-1/2} \sum_{t=1}^{[Ts]} \sum_{j>k_T} \Delta u_{t-j} a_j^* &= T^{-1/2} \sum_{j>k_T} a_j^* \sum_{t=1}^{[Ts]} \Delta u_{t-j} \leq T^{-1/2} \sum_{j>k_T} |a_j^*| \left\| \sum_{t=1}^{[Ts]} \Delta u_{t-j} \right\| \\ &= o_p(k_T^{-1} T^{-1/2}) = o_p(1), \end{aligned}$$

where we use the fact that $\left\| \sum_{t=1}^{[Ts]} \Delta u_{t-j} \right\| = O_p(1)$ uniformly over j and s . We thus have

$$\begin{aligned} T^{-1/2} \sum_{t=1}^{[Ts]} (u_t - \hat{\alpha}_S^{(i)} u_{t-1}) &= T^{-1/2} \sum_{t=1}^{[Ts]} e_t + o_p(1) \\ &\Rightarrow \sigma W(s). \end{aligned}$$

The result of the theorem then follows if $\hat{h}_v^{(i)} \xrightarrow{p} \sigma^2$ for each i . Now, let $f_t = z_t' \varsigma + b_t + e_t = \sum_{j=1}^{\infty} \Delta u_{t-j} a_j^* + e_t = F(L)e_t$, where $F(L) = 1 + (1 - \alpha L)^{-1} (1 - L) d(L) \sum_{j=1}^{\infty} a_j^* L^j$. Then $(2\pi \text{ times})$ the spectral density function of f_t at the zero frequency is

$$h_f = F(1)^2 \sigma^2 = \sigma^2, \quad (16)$$

since $F(1) = 1$. Thus, we have that

$$(\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} f_t^2 + (\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{j=1}^{\hat{T}_i - \hat{T}_{i-1} - 1} w(j/l_T) \sum_{t=\hat{T}_{i-1}+j+1}^{\hat{T}_i} f_t f_{t-j} = \sigma^2 + o_p(1). \quad (17)$$

Next, we write

$$\begin{aligned} \hat{v}_t^{(i)} &= y_t - \hat{\alpha}_S^{(i)} y_{t-1} - (x_t^{(i)} - \hat{\alpha}_S^{(i)} x_{t-1}^{(i)})' \hat{\Psi}^{(i)} \\ &= \left\{ D_{1T}^{-1} (x_t^{(i)} - \hat{\alpha}_S^{(i)} x_{t-1}^{(i)}) \right\}' \left\{ D_{1T} (\Psi^{(i)} - \hat{\Psi}^{(i)}) \right\} + (\alpha - \hat{\alpha}_S^{(i)}) u_{t-1} + f_t \end{aligned}$$

where $D_{1T} = \text{diag}(T^{1/2}, T^{3/2}, T^{3/2})$. We then have, for $s \geq 0$,

$$T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} \hat{v}_t^{(i)} \hat{v}_{t-s}^{(i)} = T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} f_t f_{t-s} + \sum_{g=1}^5 S_g, \text{ (say)}$$

We now consider each of the terms $S_1, \dots, S_5$ in turn.

S_1 :

$$S_1 = T^{-1} \{D_{1T}(\Psi^{(i)} - \hat{\Psi}^{(i)})\}' \left\{ D_{1T}^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} (x_t^{(i)} - \hat{\alpha}_S^{(i)} x_{t-1}^{(i)}) (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\} \{D_{1T}(\Psi^{(i)} - \hat{\Psi}^{(i)})\}$$

Note that $\|D_{1T}(\Psi^{(i)} - \hat{\Psi}^{(i)})\| = O_p(1)$, $\|D_{1T}^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} (x_t^{(i)} - \hat{\alpha}_S^{(i)} x_{t-1}^{(i)}) (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1}\| = O_p(1)$. Thus, $S_1 = O_p(T^{-1})$.

S_2 :

$$S_2 = T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} f_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \{D_{1T}(\Psi^{(i)} - \hat{\Psi}^{(i)})\} \\ + T^{-1} \{D_{1T}(\Psi^{(i)} - \hat{\Psi}^{(i)})\}' D_{1T}^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} (x_t^{(i)} - \hat{\alpha}_S^{(i)} x_{t-1}^{(i)}) f_{t-s}.$$

Note that

$$\left\| \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} f_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| \leq \left\| \zeta' \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} z_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| \\ + \left\| \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} b_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| + \left\| \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} e_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\|,$$

where

$$\left\| \zeta' \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} z_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| = O_p(k_T^{1/2} T^{-1/2}) \\ \left\| \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} b_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| \leq \sum_{j>k_T} |a_j^*| \left\| \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} \Delta u_{t-j} (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| \\ = o_p(k_T^{-1}) O_p(T^{-1/2}) \\ \left\| \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} e_t (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| = O_p(1).$$

The first term in S_2 is thus $O_p(T^{-1})$. Similar arguments can be used to show the second term in S_2 is also $O_p(T^{-1})$.

S_3 :

$$S_3 = (\alpha - \hat{\alpha}_S^{(i)}) \left\{ T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-1} (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\} D_{1T}(\Psi^{(i)} - \hat{\Psi}^{(i)}) \\ + (\alpha - \hat{\alpha}_S^{(i)}) \left\{ T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-s-1} (x_t^{(i)} - \hat{\alpha}_S^{(i)} x_{t-1}^{(i)})' D_{1T}^{-1} \right\} D_{1T}(\Psi^{(i)} - \hat{\Psi}^{(i)}).$$

Note that $\left\| \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-1} (x_{t-s}^{(i)} - \hat{\alpha}_S^{(i)} x_{t-s-1}^{(i)})' D_{1T}^{-1} \right\| = O_p(1)$ and $\alpha - \hat{\alpha}_S^{(i)} = O_p(T^{-1/2})$. Thus, the first term in S_3 is $O_p(T^{-3/2})$. Similarly, the second term in S_3 is also $O_p(T^{-3/2})$.

S_4 :

$$S_4 = (\alpha - \hat{\alpha}_S^{(i)})^2 T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-1} u_{t-s-1}.$$

Since $\|T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-1} u_{t-s-1}\| = O_p(1)$, it follows that $S4 = O_p(T^{-1})$.

S5:

$$S5 = (\alpha - \hat{\alpha}_S^{(i)})T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-s-1} f_t + (\alpha - \hat{\alpha}_S^{(i)})T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-1} f_{t-s}.$$

Here, we have

$$\begin{aligned} \left\| T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-s-1} f_t \right\| &\leq \sum_{j=1}^{\infty} \|a_j^*\| \left\| T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-s-1} \Delta u_{t-j} \right\| + \left\| T^{-1} \sum_{t=\hat{T}_{i-1}+s+1}^{\hat{T}_i} u_{t-s-1} e_t \right\| \\ &= O_p(1) + O_p(T^{-1/2}) = O_p(1). \end{aligned}$$

so that the first term in S5 is $O_p(T^{-1/2})$. That the second term in S5 is also $O_p(T^{-1/2})$ follows in an analogous manner.

Hence, combining the results for terms S1–S5, using (17) and the fact that $T^{-1} \sum_{j=1}^{T-1} |w(j/l_T)| \rightarrow \int_0^1 |w(s)| ds$, we have

$$\begin{aligned} \hat{h}_v^{(i)} &= (\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} \left\{ \hat{v}_t^{(i)} \right\}^2 + (\hat{T}_i - \hat{T}_{i-1})^{-1} \sum_{j=1}^{\hat{T}_i - \hat{T}_{i-1} - 1} w(j/l_T) \sum_{t=\hat{T}_{i-1}+j+1}^{\hat{T}_i} \hat{v}_t^{(i)} \hat{v}_{t-j}^{(i)} \\ &= \sigma^2 + o_p(1). \end{aligned}$$

This proves the theorem for the $I(0)$ case.

UNIT ROOT CASE ($\alpha = 1$). Note that in this case

$$\begin{aligned} T^{-1/2} \sum_{t=1}^{[Ts]} (u_t - \hat{\alpha}_S^{(i)} u_{t-1}) &= (1 - \hat{\alpha}_S^{(i)})T^{-1/2} \sum_{t=1}^{[Ts]} u_{t-1} + T^{-1/2} \sum_{t=1}^{[Ts]} v_t \\ &= [T(1 - \hat{\alpha}_S^{(i)})][T^{-3/2} \sum_{t=1}^{[Ts]} u_{t-1}] + T^{-1/2} \sum_{t=1}^{[Ts]} v_t \end{aligned} \quad (18)$$

with $v_t = \Delta u_t$. Now since $1 - \hat{\alpha}_S^{(i)} = o_p(T^{-1})$, $T^{-3/2} \sum_{t=1}^{[Ts]} u_{t-1} = O_p(1)$, the first term in (18) is $o_p(1)$. For the second term, note that $\varsigma(L)v_t = e_t$, where $\varsigma(L) = (1 - \sum_{j=1}^{\infty} a_j^* L^j)$. We thus have $T^{-1/2} \sum_{t=1}^{[Ts]} v_t \Rightarrow h_v W(s)$, where $h_v = \sigma^2/\varsigma(1)^2$. Again, it suffices to show that $\hat{h}_v^{(i)} = \{\hat{\sigma}_{ek}^{(i)}\}^2 / \{\hat{\varsigma}^{(i)}(1)\}^2$ is a consistent estimate of h_v .

Let $D_{2T} = \text{diag}(1, T^{1/2}, T^{1/2})$, $\eta_t = (v_{t-1}, \dots, v_{t-k_T})'$, $\hat{\eta}_t^{(i)} = (\hat{v}_{t-1}^{(i)}, \dots, \hat{v}_{t-k_T}^{(i)})'$ and $\hat{\varsigma}^{(i)} = (\hat{\varsigma}_1^{(i)}, \dots, \hat{\varsigma}_{k_T}^{(i)})'$. Then we can write $\hat{v}_t^{(i)} = \hat{\eta}_t^{(i)'} \hat{\varsigma}^{(i)} + \tilde{e}_{tk}^{(i)}$. Note that

$$\hat{v}_t^{(i)} = \{D_{2T}^{-1}(x_t^{(i)} - \hat{\alpha}_S^{(i)} x_{t-1}^{(i)})\}' \{\Psi^{(i)} - \hat{\Psi}^{(i)}\} + (1 - \hat{\alpha}_S^{(i)})u_{t-1} + v_t.$$

Next, it is straightforward to verify that, uniformly in $j \in [0, k_T]$,

$$T^{-1} \sum_{t=\hat{T}_{i-1}+j+1}^{\hat{T}_i} \hat{v}_t^{(i)} \hat{v}_{t-j}^{(i)} = T^{-1} \sum_{t=\hat{T}_{i-1}+j+1}^{\hat{T}_i} v_t v_{t-j} + O_p(T^{-1}).$$

This implies that

$$T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \hat{\eta}_t^{(i)} \hat{v}_t^{(i)} = T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \eta_t v_t + O_p(k_T^{1/2} T^{-1})$$

and

$$T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \hat{\eta}_t^{(i)} \hat{\eta}_t^{(i)'} = T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \eta_t \eta_t' + O_p(k_T T^{-1}).$$

Now, let

$$n = \left\| \left[T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \hat{\eta}_t^{(i)} \hat{\eta}_t^{(i)'} \right]^{-1} - \left[T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \eta_t \eta_t' \right]^{-1} \right\|$$

and $n_1 = \left\| \left[T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \eta_t \eta_t' \right]^{-1} \right\|$. Then, observe that

$$n \leq \{n + n_1\} \left\| T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \hat{\eta}_t^{(i)} \hat{\eta}_t^{(i)'} - T^{-1} \sum_{t=\hat{T}_{i-1}+k_T+1}^{\hat{T}_i} \eta_t \eta_t' \right\| n_1.$$

From Lemma 3.2(a) in Chang and Park (2002), it follows that $n_1 = O_p(1)$. Thus, we get

$$n \leq \frac{O_p(1)O_p(k_T T^{-1})}{1 - O_p(k_T T^{-1})} = O_p(k_T T^{-1})$$

Next, we have

$$\hat{\zeta}^{(i)} = \left[\sum_{t=\hat{\tau}_{i-1}+k_T+1}^{\hat{\tau}_i} \hat{\eta}_t^{(i)} \hat{\eta}_t^{(i)'} \right]^{-1} \sum_{t=\hat{\tau}_{i-1}+k_T+1}^{\hat{\tau}_i} \hat{\eta}_t^{(i)} \hat{v}_t^{(i)}$$

so that

$$\begin{aligned} \|\hat{\zeta}^{(i)} - \zeta\| &\leq (n) \left\| T^{-1} \sum_{t=\hat{\tau}_{i-1}+k_T+1}^{\hat{\tau}_i} \eta_t \eta_t' \right\|_1 \|\zeta\| \\ &\quad + (n + n_1) \left\| T^{-1} \sum_{t=\hat{\tau}_{i-1}+k_T+1}^{\hat{\tau}_i} \eta_t \sum_{j>k_T} v_{t-j} a_j^* \right\| + \left\| T^{-1} \sum_{t=\hat{\tau}_{i-1}+k_T+1}^{\hat{\tau}_i} \eta_t e_t \right\| + O_p(k_T^{1/2} T^{-1}) \\ &= O_p(k_T T^{-1}) O_p(1) + [O_p(k_T T^{-1}) + O_p(1)] [o_p(k_T^{-1/2}) + O_p(k_T^{1/2} T^{-1/2})] + O_p(k_T^{1/2} T^{-1}) \\ &= o_p(k_T^{-1/2}). \end{aligned} \quad (19)$$

Next, denoting ι as the $(k_T \times 1)$ vector of ones and using the result in (19), we have

$$\iota'(\hat{\zeta}^{(i)} - \zeta) \leq \|\iota'\| \|\hat{\zeta}^{(i)} - \zeta\| = O(k_T^{1/2}) o_p(k_T^{-1/2}) = o_p(1).$$

We therefore get $\hat{\zeta}^{(i)}(1)^2 \xrightarrow{P} \zeta(1)^2$.

Finally, the fact that $\{\hat{\sigma}_{ek}^{(i)}\}^2$ is a consistent estimate of σ^2 follows using steps similar to those employed above for showing the consistency of $\hat{\zeta}^{(i)}(1)^2$ for $\zeta(1)^2$. The details are omitted to conserve space.

Acknowledgement

The authors are grateful to an anonymous referee for detailed comments and suggestions which helped improve the article. Any errors are our own responsibility.

REFERENCES

- Andrews, D. W. K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**, 817–88.
- Andrews, D. W. K. and Ploberger, W. (1994) Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* **62**, 1383–414.
- Bai, J. and Perron, P. (1998) Estimating and testing linear models with multiple structural changes. *Econometrica* **66**, 47–78.
- Bai, J. and Perron, P. (2003) Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* **18**, 1–22.
- Ben-David, D. and Papell, D. H. (1997) International trade and structural change. *Journal of International Economics* **43**, 513–23.
- Ben-David, D. and Papell, D. H. (2000) Some evidence on the continuity of the growth process among the G7 countries. *Economic Inquiry* **38**, 320–30.
- Carrion-i-Silvestre, J. L., Kim, D. and Perron, P. (2009) GLS-based unit root tests with multiple structural breaks both under the null and the alternative hypotheses. *Econometric Theory* **25**, 1754–92.
- Chang, Y. and Park, J. Y. (2002) On the asymptotics of ADF tests for unit roots. *Econometric Reviews* **21**, 431–47.
- Harvey, D. I., Leybourne, S. J. and Taylor, A. M. R. (2009a) Simple, robust and powerful tests of the breaking trend hypothesis. *Econometric Theory* **25**, 995–1029.
- Harvey, D. I., Leybourne, S. J. and Taylor, A. M. R. (2009b) *Robust methods for detecting multiple level breaks in autocorrelated time series*. Granger Centre Discussion Paper No. 09/01.
- Hosoya, Y. (1989) Hierarchical statistical models and a generalized likelihood ratio test. *Journal of the Royal Statistical Society, Series B* **51**, 435–47.
- Kejriwal, M. and Lopez, C. (2009) *Unit roots, level shifts and trend breaks in per capita output: a robust evaluation*. Krannert Working Paper No. 1227, Purdue University.
- Kejriwal, M., and Perron, P. (2008) The limit distribution of the estimates in cointegrated regression models with multiple structural changes. *Journal of Econometrics* **146**, 59–73.
- Kim, D., and Perron, P. (2009) Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. *Journal of Econometrics* **148**, 1–13.
- Loewy, M. B. and Papell, D. H. (1996) Are US regional incomes converging? Some further evidence. *Journal of Monetary Economics* **38**, 587–98.
- Lumsdaine, R. L. and Papell, D. H. (1997) Multiple trend breaks and the unit-root hypothesis. *Review of Economics and Statistics* **79**, 212–18.
- Ng, S. and Perron, P. (1995) Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association* **90**, 268–81.
- Oka, T. and Perron, P. (2009) *Testing for common breaks in a multiple equations system*. Manuscript, Department of Economics, Boston University.
- Perron, P. and Yabu, T. (2009a) Tests for shifts in trend with an integrated or stationary noise component. *Journal of Business & Economic Statistics* **27**, 369–96.
- Perron, P. and Yabu, T. (2009b) Estimating deterministic trends with an integrated or stationary noise component. *Journal of Econometrics* **151**, 56–69.
- Perron, P. and Zhu, X. (2005) Structural breaks with deterministic and stochastic trends. *Journal of Econometrics* **129**, 65–119.
- Roy, A. and Fuller, W. A. (2001) Estimation for autoregressive processes with a root near one. *Journal of Business & Economic Statistics* **19**, 482–93.

- Vogelsang, T. J. (1998) Trend function hypothesis testing in the presence of serial correlation. *Econometrica* **66**, 123–48.
- Vogelsang, T. J. (2001) *Tests for a shift in trend when serial correlation is of unknown form*. Unpublished Manuscript, Department of Economics, Cornell University.
- Zivot, E. and Andrews, D. W. K. (1992) Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics* **10**, 251–70.