

Article

Fractional Unit Root Tests Allowing for a Structural Change in Trend under Both the Null and Alternative Hypotheses

Seong Yeon Chang ^{1,*} and Pierre Perron ²

¹ Wang Yanan Institute for Studies in Economics, Department of Statistics, MOE Key Laboratory of Econometrics and Fujian Key Laboratory of Statistical Science, Xiamen University, Xiamen 361005, China

² Department of Economics, Boston University, Boston, MA 02215, USA; perron@bu.edu

* Correspondence: sychang@xmu.edu.cn; Tel.: +86-156-5928-1231

Academic Editor: Steve Cook

Received: 17 November 2016; Accepted: 3 January 2017; Published: 8 January 2017

Abstract: This paper considers testing procedures for the null hypothesis of a unit root process against the alternative of a fractional process, called a fractional unit root test. We extend the Lagrange Multiplier (LM) tests of Robinson (1994) and Tanaka (1999), which are locally best invariant and uniformly most powerful, to allow for a slope change in trend with or without a concurrent level shift under both the null and alternative hypotheses. We show that the limit distribution of the proposed LM tests is standard normal. Finite sample simulation experiments show that the tests have good size and power. As an empirical analysis, we apply the tests to the Consumer Price Indices of the G7 countries.

Keywords: hypothesis testing; LM test; slope change; spurious break; trend function

JEL Classification: C22

1. Introduction

Non-stationarity in economic time series is a pervasive feature. In order to carry proper inference, it is important to find the exact features that lead to this non-stationarity. A unit root process is a well-known example of non-stationary processes, and testing for a unit root against stationarity has been a topic of substantial interest from both theoretical and empirical perspectives. Perron (1989) [1], however, showed that the Dickey and Fuller (1979) [2] type unit root test is biased in favor of a non-rejection of the unit root null hypothesis when the process is trend stationary with a structural change in slope. Perron (1989, 1990) [1,3] proposed testing procedures in which a structural break is allowed under both the null and alternative hypotheses. Later, Christiano (1992) [4] and Zivot and Andrews (1992) [5] criticized the assumption that the date of the structural break is known a priori. In succeeding research, Zivot and Andrews (1992) [5], Perron (1997) [6], and Vogelsang and Perron (1998) [7] treated the break date as unknown and proposed testing procedures for a unit root. In much work, especially that of Zivot and Andrews (1992) [5], it was common to allow for a structural break only under the alternative hypothesis, not under the null hypothesis of a unit root. This is very restrictive, and can lead to misleading results. Recent advances in testing for and estimating a structural break in a trend function have made possible the development of unit root tests that allow for a change in trend under both the null and alternative hypotheses. Perron and Zhu (2005) [8] established the consistency, rate of convergence, and limiting distribution of the parameter estimates when there is a break in a trend function with or without a concurrent level shift. Perron and Yabu (2009) [9] suggested a testing procedure for structural changes in the trend function of a

time series without any prior knowledge of whether the noise component is stationary or has an autoregressive unit root. Building on this work, Kim and Perron (2009) [10] proposed unit root testing procedures which allow for a structural change under both the null and alternative hypotheses; see also Carrion-i-Silvestre et al. (2009) [11] for an extension to the case with multiple changes.

Fractional processes with the order of integration $d \geq 0.5$ are also non-stationary. Standard unit root tests often reject the null hypothesis when the true process is fractionally integrated with $d \in (0.5, 1)$. This can lead to the misleading conclusion that the process of interest is stationary. This motivated researchers to introduce unit root tests which are powerful against the alternative hypothesis of a fractional process. Robinson (1991) [12] derived a Lagrange Multiplier (LM) test for fractional white noise disturbances in a linear regression, while Robinson (1994) [13] proposed tests for unit root, and actually any real values of d in both the frequency and the time domain. Tanaka (1999) [14] suggested an LM test in the time domain, and showed that it is locally best invariant and uniformly most powerful. Dolado et al. (2002) [15] introduced a Wald-type unit root test against the alternative of fractional integration. This test is based on the Dickey and Fuller (1979) [2] type test using an auxiliary regression with a consistent estimate of the integration order. Lobato and Velasco (2007) [16] established a Wald-type test which is more efficient and is asymptotically equivalent to the LM test. Recently, Cho et al. (2015) [17] suggested combining the test of Kwiatkowski et al. (1992) [18] and a unit root test to test the null of integer integration, i.e., $I(0)$ or $I(1)$ against the alternative of fractional integration, i.e., $I(d)$, $d \in (0, 1)$. In this line of work, the process of interest has been limited to either a random walk or a purely fractional process. Lobato and Velasco (2007) [16] considered short-run dynamics in the process. Dolado et al. (2008) [19] extended the work of Dolado et al. (2002) [15] and Lobato and Velasco (2007) [16] to incorporate some deterministic components; for instance, a constant and a linear trend function.

Our main contribution is to extend the LM test for a fractional unit root to allow for a structural change in a trend function under both the null and alternative hypotheses. This extension has some advantages, as follows: (i) it imposes a symmetric treatment of the nature of the deterministic trend under both the null and alternative hypotheses; (ii) it does not require long memory to be distinguished from structural change;¹ (iii) the power of fractional unit root tests can be substantially improved when a break is actually present. We consider linear trend models in which a structural change in slope occurs with or without a concurrent level shift.

The rest of this paper is organized as follows. In Section 2, we first introduce fractional processes and the Lagrange Multiplier test of Tanaka (1999) [14] along with preliminary results to be used subsequently. In Section 3, the LM tests are generalized to allow for a structural break in trend under both the null and alternative hypotheses. Extensions to processes with short-run dynamics are discussed in Section 4. The results of simulation experiments about the size and power of the tests are presented in Section 5. As an empirical application, we test for a fractional unit root in the Consumer Price Indices (CPIs) of the G7 countries in Section 6. Concluding remarks are provided in Section 7. All mathematical proofs are relegated to the Appendix A.

2. Lagrange Multiplier Test

For an integer $d = 1, 2, \dots$, the operator $\Delta^d = (1 - L)^d$ denotes the differencing operator with the usual lag operator L ; i.e., $LX_t \equiv X_{t-1}$, $\Delta X_t = X_t - X_{t-1}$, $\Delta^2 X_t = (1 - 2L + L^2)X_t = X_t - 2X_{t-1} + X_{t-2}$,

¹ Given that unit root and long memory processes share similar features, distinguishing between long memory processes and short memory processes with structural changes has been an important topic in econometrics and financial economics. Along the lines of Perron (1989) [1], it is well known that short memory processes with level shifts exhibit properties that lead standard tools to conclude that long memory is present (e.g., Diebold and Inoue (2001) [20], Granger and Hyung (2004) [21], Lu and Perron (2010) [22], Perron and Qu (2010) [23], Qu and Perron (2013) [24], Xu and Perron (2014) [25], and Varneskov and Perron (2016) [26], among many others). On the other hand, it has been also documented that long memory processes induce a rejection of the null hypothesis of no structural change when using conventional structural change tests (see Wright (1998) [27] and Krämer and Sibbertsen (2002) [28]).

and so on. For a non-integer real number $d > -1$, the difference operator $\Delta^d = (1 - L)^d$ is defined by means of the binomial expansion

$$\Delta^d = \sum_{k=0}^{\infty} \binom{d}{k} (-L)^k = \sum_{k=0}^{\infty} \pi_k(d) L^k,$$

where

$$\binom{d}{k} \equiv \frac{d!}{k!(d-k)!} = \frac{d \times (d-1) \times \cdots \times (d-k+1)}{k \times (k-1) \times \cdots \times 2 \times 1},$$

$$\pi_k(d) \equiv (-1)^k \binom{d}{k} = \prod_{s=1}^k \frac{s-1-d}{s} = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)},$$

with $\Gamma(\cdot)$ the gamma function, so that $\pi_k(d) = \binom{k-d-1}{k}$ and $\pi_0(d) = 1$. Recall that $x! \equiv \Gamma(x+1)$, $x = 0, 1, \dots$, and for $k = 1, 2, \dots$, $0 < x < 1$, $\Gamma(x-k)$ is defined as $\Gamma(x) = (x-1)\Gamma(x-1) = \cdots = (x-1) \cdots (x-k)\Gamma(x-k)$. To define a fractional process, we use the notation of Robinson (2005) [29]. Let $\{\eta_t, t = 0, \pm 1, \dots\}$ be a short-memory zero-mean covariance stationary process, with spectral density that is bounded and bounded away from zero. For $d \in (-0.5, 0.5)$,

$$\zeta_t = \Delta^{-d} \eta_t, \quad t = 0, \pm 1, \dots, \quad (1)$$

is covariance stationary and invertible. The truncated version of ζ_t is defined as

$$\zeta_t^\# = \zeta_t \mathbf{1}_{t \geq 1}, \quad t = 0, \pm 1, \dots, \quad (2)$$

where $\mathbf{1}_A$ is the indicator function for the event A . For an integer $m \geq 0$,

$$u_t = \Delta^{-m} \zeta_t^\#, \quad t = 0, \pm 1, \dots \quad (3)$$

is called a type I $I(m+d)$ process. Let $\mathcal{D}[0, 1]$ be the space of functions on $[0, 1]$ which are right continuous and have left limits, equipped with the Skorohod topology. Let \xrightarrow{p} denote convergence in probability and \xrightarrow{d} convergence in distribution. Denote by $[a]$ the integer part of $a \in \mathbb{R}$. The order of integration is $d_0 = d + m$, with $d \in (-0.5, 0.5)$ and $m \in \{0, 1\}$.²

Remark 1. Marinucci and Robinson (1999) [30] defined type I and type II fractional Brownian motions with $d \in (-0.5, 0.5)$ on $\mathcal{D}[0, 1]$, respectively, as follows:

$$B_I(t) = \frac{1}{\Gamma(d+1)} \left\{ \int_{-\infty}^0 [(t-s)^d - (-s)^d] dB(s) + \int_0^t (t-s) dB(s) \right\},$$

$$B_{II}(t) = \frac{1}{\Gamma(d+1)} \left\{ \int_0^t (t-s) dB(s) \right\},$$

where $B(\cdot)$ denotes the standard Brownian motion. Furthermore, Robinson (2005) [29] and Davidson and Hashimzade (2009) [31] pointed out that asymptotic results vary depending on the definition of fractional Brownian motions considered, which requires one to design simulation experiments in accordance with the particular type used.

Now we consider a fractional unit root test. Under the null hypothesis, $\{u_t\}$ is a unit root process; that is, $H_0 : d_0 = 1$ (i.e., $d = 0$ and $m = 1$). The alternative hypothesis can be either one-sided

² The restriction that $d_0 \neq 0.5$ is standard in the long memory literature. Tanaka (1999) [14] showed that the case with $d_0 = 0.5$ needs to be treated separately from the case with $d_0 \neq 0.5$.

($H_1 : d_0 > 1$ or $H_1 : d_0 < 1$) or two-sided ($H_1 : d_0 \neq 1$). Robinson (1994) [13] and Tanaka (1999) [14] considered the Lagrange Multiplier test in the frequency and time domain. It is well known that the LM test is locally best invariant. Further, Tanaka (1999) [14] showed that the LM test is locally uniformly most powerful invariant because it achieves the power envelope of all the invariant tests against local alternatives. The test statistic suggested in Tanaka (1999) [14] is

$$\mathcal{LM} = \sqrt{T} \sqrt{\frac{6}{\pi^2}} \sum_{k=1}^{T-1} \frac{1}{k} \rho_k,$$

where $\rho_k = (\sum_{j=k+1}^T \Delta u_{j-k} \Delta u_j) / \sum_{j=1}^T (\Delta u_j)^2$ is the k th order autocorrelation of the residuals Δu_t . Local alternatives to the null hypothesis are often considered in the literature, with the integration order defined as $d_0 = 1 + \delta T^{-1/2}$ with δ fixed, often referred to as Pitman drifts. We state the limiting distribution of the LM test under local alternatives, as it will be relevant for subsequent derivations.

Lemma 1 (Theorem 3.1 in Tanaka (1999) [14]). *Under the assumption that u_t is generated by (3) with $d_0 = 1 + \delta T^{-1/2}$ and δ fixed, it holds that, as $T \rightarrow \infty$, $\mathcal{LM} \xrightarrow{d} \mathcal{N}(\delta \sqrt{\pi^2/6}, 1)$.*

3. Deterministic Components Allowing for a Structural Change

In this section, we extend the LM test for a fractional unit root to allow for a structural change in trend with or without a concurrent level shift. We consider the time series of interest y_t as consisting of a deterministic component (f_t) and fractionally integrated errors. The data-generating process (DGP) is specified as

$$y_t = f_t + u_t.$$

For u_t , we impose $E(u_t) = 0$ and the following assumption.

Assumption 1. u_t is a type I $I(m+d)$ process which is defined in (1)–(3). Moreover, the short-memory zero-mean covariance stationary process η_t is assumed to be independent and identically distributed (i.i.d.) with zero mean and finite variance.

The i.i.d. assumption on the short-memory process η_t will be relaxed later to allow for short-run dynamics. The unit root null hypothesis corresponds to the case with $m = 1$ and $d = 0$, which implies that u_t is a weighted sum of η_t .

3.1. Change in Mean

We first consider the case where y_t experiences a level shift at an unknown time T_b . The DGP is specified as

$$y_t = \mu_1 + \mu_b C_t + u_t, \quad \mu_b \neq 0, \quad (4)$$

where C_t is a dummy variable for a level shift defined by:

$$C_t = \begin{cases} 0 & \text{if } t \leq T_b, \\ 1 & \text{if } t > T_b, \end{cases}$$

where $T_b = [T\lambda_b]$ is the true break date with the corresponding true break fraction $\lambda_b \in (0, 1)$.

Theorem 1 (Change in Mean). *Under Assumption 1, suppose that the process $\{y_t\}$ is generated under the null hypothesis of (4). Consider the Lagrange Multiplier test \mathcal{LM}_M defined by:*

$$\mathcal{LM}_M = \sqrt{T} \sqrt{\frac{6}{\pi^2}} \frac{\sum_{t=1}^T (-\log \Delta y_t) \Delta y_t}{\sum_{t=1}^T (\Delta y_t)^2}.$$

Under the null hypothesis $H_0 : d_0 = 1$, it holds that as $T \rightarrow \infty$, $\mathcal{LM}_M \xrightarrow{d} \mathcal{N}(0, 1)$.

Theorem 1 implies that Tanaka's (1999) [14] LM test is robust to the presence of a level shift. In the following subsection, we consider the LM test in the context of trending series.

3.2. Slope and Intercept Change in Trending Series

We now introduce a deterministic time trend in the models. We follow the notation in Kim and Perron (2009) [10] (henceforth KP) from which we will use some relevant results. The DGPs are specified as

1. Model A0: (Deterministic time trend without a structural change)

$$y_t = \mu_1 + \beta_1 t + u_t, \quad (5)$$

2. Model A1: (Level shift)

$$y_t = \mu_1 + \mu_b C_t + \beta_1 t + u_t, \quad (6)$$

3. Model A2: (Joint broken trend)

$$y_t = \mu_1 + \beta_1 t + \beta_b B_t + u_t, \quad (7)$$

where B_t is a dummy variable for a slope change in trend given by

$$B_t = \begin{cases} 0 & \text{if } t \leq T_b, \\ t - T_b & \text{if } t > T_b, \end{cases}$$

4. Model A3: (Locally disjoint broken trend)

$$y_t = \mu_1 + \mu_b C_t + \beta_1 t + \beta_b B_t + u_t.$$

Following KP, we can rewrite Models A1–A3 as follows:

$$y_t = z(T_b)'_t \phi + u_t = z'_{t,1} \phi_1 + z(T_b)'_{t,2} \phi_2 + u_t,$$

where $z_{t,1} = (1, t)'$, $\phi_1 = (\mu_1, \beta_1)'$,

$$z(T_b)_{t,2} = \begin{cases} C_t \\ B_t \\ (C_t, B_t)' \end{cases}, \quad \phi_2 = \begin{cases} \mu_b & \text{for Model A1} \\ \beta_b & \text{for Model A2} \\ (\mu_b, \beta_b)' & \text{for Model A3.} \end{cases}$$

In matrix notation, the models defined previously can be specified as $Y = Z_{T_b} \phi + U$, where $Y = [y_1, \dots, y_T]'$, $Z_{T_b} = [z(T_b)_1, \dots, z(T_b)_T]'$, $\phi = (\phi'_1, \phi'_2)'$, and $U = [u_1, \dots, u_T]'$.

Consider first Model A0, where no structural change is allowed. By taking first differences, we can rewrite (5) as follows:

$$\Delta y_t = \beta_1 + \Delta u_t = \beta_1 + \Delta^{1-d_0} \eta_t \mathbf{1}_{t \geq 1}. \quad (8)$$

The ordinary least squares (OLS) estimate of β_1 is $\hat{\beta}_1 = T^{-1} \sum_{t=1}^T \Delta y_t$, which is consistent under both H_0 and H_1 .³ We define $\widehat{\Delta y}_t = \Delta y_t - \hat{\beta}_1$, the OLS residuals from the regression model (8).

³ Under H_0 , $\hat{\beta}_1$ is a $T^{1/2}$ -consistent estimator of the slope coefficient β_1 . Hosking (1996) [32] considered a stationary ARFIMA (p, d, q) process $\{y_t\}$ and showed the weak convergence of the sample mean for $d \in (-0.5, 0.5)$. It is not difficult to generalize the result to the case where $d \in (0.5, 1)$, for which $\hat{\beta}_1$ is a $T^{3/2-d}$ -consistent estimator.

Theorem 2 (Linear Trend). Under Assumption 1, suppose that the process $\{y_t\}$ is generated under the null hypothesis of (5). Consider the Lagrange Multiplier test \mathcal{LM}_T defined by:

$$\mathcal{LM}_T = \sqrt{T} \sqrt{\frac{6}{\pi^2} \frac{\sum_{t=1}^T (-\log \Delta \widetilde{\Delta y_t}) \widetilde{\Delta y_t}}{\sum_{t=1}^T (\widetilde{\Delta y_t})^2}}.$$

Under the null hypothesis $H_0 : d_0 = 1$, it holds that as $T \rightarrow \infty$, $\mathcal{LM}_T \xrightarrow{d} \mathcal{N}(0, 1)$.

In what follows, the aim is to devise Lagrange Multiplier tests allowing for a slope change in trend with or without a concurrent level shift. The following assumption is essential to that effect.

Assumption 2. $\beta_b \neq 0$ and $\lambda_b \in (\pi, 1 - \pi)$ for some $\pi \in (0, 1/2)$.

Assumption 2 ensures that there is a single slope change in trend, and that the pre- and post-break samples are not asymptotically negligible, which is a standard assumption needed to derive useful asymptotic results. Model A1 (level shift only) will be revisited later.

The break date can be estimated by using a global least-squares criterion:

$$\hat{T}_b = \arg \min_{T_1 \in \Lambda} Y'(I - P_{T_1})Y,$$

where P_{T_1} is the matrix that projects on the range space of Z_{T_1} ; i.e., $P_{T_1} = Z_{T_1}(Z'_{T_1}Z_{T_1})^{-1}Z'_{T_1}$ and $\Lambda = [\pi T, (1 - \pi)T]$, $0 < \pi < 1/2$. Note that Z_{T_1} is the same as Z_{T_b} , except T_b is replaced with a generic break date T_1 . Perron and Zhu (2005) [8] (henceforth, PZ) established the consistency, rate of convergence, and limiting distribution of parameter estimates when the error is an $I(1)$ process. With $Z_{\hat{T}_b}$ constructed using the estimate \hat{T}_b , the OLS estimate of ϕ is $\hat{\phi} = (Z'_{\hat{T}_b}Z_{\hat{T}_b})^{-1}Z'_{\hat{T}_b}Y$, and the resulting sum of squared residuals is, for an estimated break fraction $\hat{\lambda}_s = \hat{T}_b/T$ (the subscript s refers to the fact that we consider a static regression; a dynamic regression with lagged dependent variables will be considered later):

$$S(\hat{\lambda}_s) = \sum_{t=1}^T \hat{u}_t^2 = \sum_{t=1}^T (y_t - z(\hat{T}_b)'_t \hat{\phi})^2 = Y'(I - P_{\hat{T}_b})Y,$$

where $P_{\hat{T}_b}$ is the projection matrix associated with $X_{\hat{T}_b}$. The rate of convergence of $\hat{\lambda}_s$ for Models A2 and A3 is $\hat{\lambda}_s - \lambda_b = O_p(T^{-1/2})$ with $I(1)$ errors (see Theorem 3 in PZ). Chang and Perron (2016) [33] derived the consistency and rate of convergence of $\hat{\lambda}_s$ when the noise component is a fractional process with the differencing parameter $d_0 \in (-0.5, 0.5) \cup (0.5, 1.5)$. Specifically, for Models A2 and A3, $\hat{\lambda}_s - \lambda_b = O_p(T^{-1/2+d})$ if $m = 1$ and $d \in (-0.5, 0.5)$. With the consistent estimates $(\hat{\lambda}_s, \hat{\mu}_1, \hat{\mu}_b, \hat{\beta}_1, \hat{\beta}_b)$, we can construct the detrended process $\{\tilde{y}_t\}$, and the Lagrange Multiplier test statistic $\mathcal{LM}_{T, \hat{\lambda}_s}$ is given by

$$\mathcal{LM}_{T, \hat{\lambda}_s} = \sqrt{T} \sqrt{\frac{6}{\pi^2} \frac{\sum_{t=1}^T (-\log \Delta \tilde{\Delta \tilde{y}_t}) \Delta \tilde{y}_t}{\sum_{t=1}^T (\Delta \tilde{y}_t)^2}}.$$

The convergence rate of the estimate $\hat{\lambda}_s$ is not fast enough to guarantee that $\mathcal{LM}_{T, \hat{\lambda}_s}$ has the standard normal limit under H_0 . KP faced a similar issue in dealing with unit root tests. They introduced a heuristic explanation of the issue involved, which we briefly review. Let $\hat{\lambda} = \hat{T}_b/T$ denote an estimate of the break fraction such that $\hat{\lambda} - \lambda_b = O_p(T^{-\varkappa})$ for some $\varkappa \geq 0$. The detrended series $\{\tilde{y}_t\}$ is given by

$$\tilde{Y} = \hat{M}_z Y = \hat{M}_z Z(T_b) \phi + \hat{M}_z U = \hat{M}_z Z(T_b)_2 \phi_2 + \hat{M}_z U, \quad (9)$$

where $Z(T_b)$ and $Z(T_b)_2$ are matrices stacking $\{z(T_b)_t'\}$ and $\{z(T_b)_{t,2}'\}$, respectively, and the idempotent matrix $\hat{M}_z = I - \hat{P}_z = I - Z(\hat{T}_b)[Z(\hat{T}_b)'Z(\hat{T}_b)]^{-1}Z(\hat{T}_b)'$. It is obvious that $\hat{M}_z Z(T_b)_2 \phi_2 = 0$ only if the true break date is used in \hat{M}_z . In finite samples, $\hat{\lambda} \neq \lambda_b$ in general; thereby, $\hat{M}_z Z(T_b)_2 \phi_2$ will not be zero. It turns out that a fast rate of convergence for the estimate of the break date is needed for the effect of $\hat{M}_z Z(T_b)_2 \phi_2$ on the Lagrange Multiplier test to become negligible asymptotically. The following proposition provides a sufficient condition under which $\mathcal{LM}_{T,\hat{\lambda}} \xrightarrow{d} \mathcal{N}(0,1)$ under H_0 .

Proposition 1. Suppose that the process $\{y_t\}$ is generated under the null hypothesis of Model A2 or A3, and that Assumptions 1 and 2 hold. Then, it holds that, as $T \rightarrow \infty$, $\mathcal{LM}_{T,\hat{\lambda}} \xrightarrow{d} \mathcal{N}(0,1)$ if $\hat{\lambda} - \lambda_b = o_p(T^{-1/2})$.

Proposition 1 implies that the estimate of the break fraction should converge at a rate faster than $T^{1/2}$. As shown above, $\hat{\lambda}_s$ does not satisfy this condition. Hence, we need to consider alternative ways to accelerate the rate of convergence of the estimate of the break fraction. KP suggested two possible approaches. The first is based on minimizing the sum of squared residuals (SSR) of a dynamic regression model. This method is similar to that in Hatanaka and Yamada (1999) [34]. The relevant dynamic regressions are specified as follows:

$$\begin{aligned} y_t &= \alpha y_{t-1} + \nu_1 D(T_b)_t + \nu_2 \mathbf{1}_{t \geq T_b} + z(T_b)_t' \phi + u_t, & (\text{Model A2}) \\ y_t &= \alpha y_{t-1} + \nu D(T_b)_t + z(T_b)_t' \phi + u_t, & (\text{Models A1 and A3}) \end{aligned} \quad (10)$$

where $D(T_b)_t = 1$ for $t = T_b + 1$ and 0 otherwise. Under the null hypothesis, we obtain an estimate of the break fraction $\hat{\lambda}_d$ which has a faster rate of convergence, such that $\hat{\lambda}_d - \lambda_b = O_p(T^{-1})$ for Models A2 and A3 (see Proposition 1 in KP). Let $\mathcal{LM}_{T,\hat{\lambda}_d}$ denote the Lagrange Multiplier test statistic with $\hat{\lambda}_d$ replacing $\hat{\lambda}_s$. It is worth noting that, as discussed by Hatanaka and Yamada (1999) [34] and KP, the estimate $\hat{\lambda}_d$ has a negative bias in finite samples, especially for Model A3. As we shall see, this will affect the finite sample properties of the tests.

The second approach is to use a trimmed data set using a window whose length depends on the sample size and which contains the estimated break date. The trimmed series then consists of the original one with the data points in the window excluded. KP showed that the rate of convergence of $\hat{\lambda}_s$ can be increased with the trimmed data set. Suppose that the estimate of the break fraction satisfies $\hat{\lambda} - \lambda_b = O_p(T^{-a})$ for some $0 < a < 1$, and the trimming window has length $2w(T)$ with $w(T) \equiv c_1 T^\delta$, $c_1 > 0$, and $-1 < -a < \delta < 0$. With this specification, the length of the window is negligible in the limit compared to the sample size T , but is still large enough to include the true break date asymptotically. Following KP's suggestion, one proceeds as follows:

- Estimate the break fraction $\hat{\lambda}_s$ from the original data set and form a window that ranges from $T_l \equiv T(\hat{\lambda}_s - w(T))$ to $T_h \equiv T(\hat{\lambda}_s + w(T))$.
- A trimmed data set is constructed by removing the original data from $T_l + 1$ to T_h and then shifting down the data after the window by $D(T) = y_{T_h} - y_{T_l}$. After the trimming and connecting procedures, we now have a new series $\{y_t^*\}$, for $t = 1, \dots, T^* (\equiv T - 2w(T)T)$, defined by:

$$y_t^* = \begin{cases} y_t & \text{if } t \leq T_l, \\ y_{t+T_h-T_l} - D(T) & \text{if } t > T_l. \end{cases}$$

- Test the null hypothesis $H_0 : d_0 = 1$ using T_l as the break date (i.e., $\hat{\lambda}_{tr} = T_l/T^*$). The Lagrange Multiplier test statistic is then given by

$$\mathcal{LM}_{T,\hat{\lambda}_{tr}} = \sqrt{T} \sqrt{\frac{6}{\pi^2} \frac{\sum_{t=1}^T (-\log \Delta \tilde{y}_t^*) \Delta \tilde{y}_t^*}{\sum_{t=1}^T (\Delta \tilde{y}_t^*)^2}},$$

where \tilde{y}_t^* is the detrended version of y_t^* using the estimate of the break date T_l (or break fraction $\hat{\lambda}_{tr}$).

Remark 2. If the window contains either end of the data, then the process $\{y_t^*\}$ turns out to be Model A0 (no structural break), and the statistic in Theorem 2 should be applied to the trimmed data $\{y_t^*\}$.

The trimmed process $\{y_t^*\}$ will satisfy the properties of Model A2 regardless of the specification of the original data $\{y_t\}$, which implies that we can use a common limit distribution. The following proposition states the limiting distribution of the Lagrange Multiplier test based on the trimmed data, which is the same as would be obtained if the break date was known in Model A2.

Proposition 2. Suppose that the process $\{y_t\}$ is generated under the null hypothesis of Model A2 or A3, and that Assumptions 1 and 2 hold. Then, it holds that as $T \rightarrow \infty$, $\mathcal{LM}_{T, \hat{\lambda}_{tr}} \xrightarrow{d} \mathcal{N}(0, 1)$.

As shown in KP, under the null hypothesis of a unit root, the estimate of the break fraction $\hat{\lambda}_{tr} = T_l/T^*$ converges in probability to the true break fraction at some rate greater than T . Hence, the sufficient condition in Proposition 1 is satisfied, so that the proof of Proposition 2 is trivial and omitted.

In concluding this section, we consider the case where there is a change in mean; that is, $\mu_b \neq 0$, as in Models A1 and A3. In Model A1, we assume that there is a level shift only; that is, $\mu_b \neq 0$ and $\beta_b = 0$. Under the null hypothesis, a stochastic trend generated by the $I(1)$ error process tends to dominate a level shift. Hence, we cannot estimate the break fraction λ_b consistently, because the magnitude of the level shift is asymptotically negligible. In finite samples, we can ignore the level shift if the magnitude of the break is small. Then, Model A1 can be treated as Model A0, and we can follow the testing procedure pertaining to Theorem 2. However, a loss of power is inevitable if a large change in mean is ignored.

On the other hand, the level shift can be specified as an increasing function of the sample size; i.e., $\mu_b = c_2 T^{1/2+\alpha}$ for some $c_2 > 0$ and $\alpha > 0$. As addressed in Harvey et al. (2001) [35], PZ, and KP, this specification provides better approximations of the properties of the tests in finite samples when the level shifts are not very small. The models with $\mu_b = c_2 T^{1/2+\alpha}$ are labeled as Models A1b and A3b, respectively.

Proposition 3. Suppose that the process $\{y_t\}$ is generated under the null hypothesis of Model A1b or A3b. Then, $\mathcal{LM}_{T, \hat{\lambda}}$ diverges as $T \rightarrow \infty$.

Although the rate of convergence of the estimate of the break fraction is faster than in the case of a change in slope (see Proposition 7 in KP), Proposition 3 states that the LM tests cannot obtain the standard normal limiting distribution. Hence, the LM test $\mathcal{LM}_{T, \hat{\lambda}}$, using the critical values from the standard normal distribution, suffers from some liberal size distortions, even when $|\mu_b|$ is large.⁴

3.3. Using a Pre-Test for a Break in Slope

The results of Theorem 2 and Proposition 2 show that the limit distribution of the test is the same whether there is a break in slope introduced as a regressor or not, even when the DGP specifies that no break is present. Hence, unlike the case of testing for a unit root as in KP, theoretically there is no need to carry a pre-test to improve the power of the test. However, Chang and Perron (2016) [33] considered Models A2 and A3 with fractionally integrated errors and showed that the so-called spurious break issue occurs with the order of fractional integration $d_0 \in (0, 0.5) \cup (0.5, 1.5)$. This extended the results on Nunes et al. (1995) [36], who considered the unit root case. This means that under both the null

⁴ Simulation results related to this issue are available upon request.

and alternative hypotheses, if a break in slope is not present and one is allowed in the regression, the fitted model will with large probability suggests the presence of a break. This could have an effect on both the size and power of the test. On one hand, the slope change regressor may induce added liberal size distortions in finite samples because of the overfit. On the other hand, since when no break exists in the DGP it is a superfluous regressor, power may be reduced. Hence, it may be the case that in finite samples it is beneficial to use a pre-test for a change in slope and try to choose between models (5) and (7). Since a test for a change in mean will be inconsistent, there is no point in trying to distinguish between models (5) and (6) or between model (4) and the corresponding one without the change in mean.

Iacone et al. (2013) [37] suggested a sup-Wald type test (\mathcal{SW}) for Model A2. In particular, it is robust to any order of fractional integration d_0 located in an interval $[0, 1.5)$ excluding the boundary case 0.5. More precisely, given their recommended choice for the bandwidth when constructing the local Whittle estimate of d_0 , their test is consistent for values of d_0 in the interval $[0, 1.32]$, though we believe the proof can be modified to allow the interval $[0, 1.5)$. It follows the generalized least squares approach to construct the test statistic for a structural change in trend by taking d_0 -differences from the data. To make the test feasible, the fully extended local whittle (FELW) estimator \hat{d}_{FELW} of Abadir (2007) [38] is considered. While the FELW estimator is constructed under the null hypothesis of no structural change, Iacone et al. (2013) [37] showed that it also satisfies the necessary condition for consistency, even with a local break in trend. Since the true break date is unknown a priori, the final statistic \mathcal{SW} uses the *sup* functional of Andrews (1993) [39] across all admissible break dates. This test is asymptotically size controlled for all d_0 's in the prescribed range. Using this pre-test, we can then define the alternative estimate of the break fraction $\tilde{\lambda} = \hat{\lambda} \cdot \mathbf{1}_{\mathcal{SW} > \tau}$, where τ is the critical value for the \mathcal{SW} test with a nominal size $p\%$. Given that \mathcal{SW} is a consistent test, $\text{plim}_{T \rightarrow \infty} \tilde{\lambda} = \lambda_b$ if $\hat{\lambda}$ is a consistent estimate of λ_b . If there is no break in the DGP, we can expect that $p\%$ of the estimates $\tilde{\lambda}$'s are nonzero. In order to obtain a consistent estimate of λ_b under the null of no structural break, we assume that the critical value τ is a function of the sample size T . Since $\mathcal{SW} = O_p(T^q)$, $q > 0$ with a local break, let $\tau = cT^{q-\epsilon}$ for $0 < \epsilon < q$. This specification introduced in KP is useful because it does not have any effect on the consistency of the test \mathcal{SW} and does guarantee that $\text{plim}_{T \rightarrow \infty} \tilde{\lambda} = 0$ when no break is present. Hence, based on the consistency of $\tilde{\lambda}$, it is recommended to use \mathcal{LM}_T if $\tilde{\lambda} = 0$ and $\mathcal{LM}_{T, \tilde{\lambda}}$ if $\tilde{\lambda} \neq 0$. The LM test statistics with the pre-test are denoted by $\mathcal{LM}_{T, \hat{\lambda}_s}^p$, $\mathcal{LM}_{T, \hat{\lambda}_d}^p$, and $\mathcal{LM}_{T, \hat{\lambda}_{tr}}^p$. Whether using a pre-test is beneficial will be assessed later via simulation experiments about the size and power of the tests.

4. Short-Run Dynamics

We now relax Assumption 1 to introduce short-run dynamics in the noise component. A zero-mean short-memory covariance stationary process η_t can be represented as a one-sided moving average: $\eta_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, $t = 0, \pm 1, \dots$, where $\psi_0 = 1$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, $\psi(1) \equiv \sum_{j=0}^{\infty} \psi_j$, and $\{\epsilon_t, t = 0, \pm 1, \dots\}$ are *i.i.d.* random variables with mean zero. A special case of interest is an autoregressive moving average (ARMA(p, q)) process given by $\phi(L)\eta_t = \theta(L)\epsilon_t$. In order to implement the Lagrange Multiplier test, we first estimate the parameters $\Psi = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ consistently. Then, under the null hypothesis, we can construct $\hat{\epsilon}_t = \hat{\phi}(L)\hat{\theta}(L)^{-1}(1-L)^{d_0}\hat{u}_t$, where \hat{u}_t is the OLS residuals from the model considered, whereas $\hat{\phi}(L)$ and $\hat{\theta}(L)$ are estimated from $\phi(L)(1-L)^{d_0}\hat{u}_t = \theta(L)\epsilon_t$, using $d_0 = 1$. With short-run dynamics in the noise component, we consider the following test statistic:

$$\mathcal{LM} = \sqrt{T} \sum_{k=1}^{T-1} \frac{1}{k} \hat{\rho}_k,$$

where $\hat{\rho}_k$ is the k th order autocorrelation of residuals $\hat{\epsilon}_1, \dots, \hat{\epsilon}_T$. Tanaka (1999) [14] derived an important result related to this statistic when no break is present.

Lemma 2 (Theorem 3.3 in Tanaka (1999) [14]). Under local alternatives—that is, $d_0 = 1 + \delta T^{-1/2}$ with δ fixed—it holds that as $T \rightarrow \infty$, $\mathcal{LM} \xrightarrow{d} \mathcal{N}(\delta\omega^2, \omega^2)$, where

$$\omega^2 = \frac{\pi^2}{6} - (\kappa_1, \dots, \kappa_p, \xi_1, \dots, \xi_q) \Phi^{-1} (\kappa_1, \dots, \kappa_p, \xi_1, \dots, \xi_q)',$$

$$\kappa_i = \sum_{j=1}^{\infty} \frac{1}{j} g_{j-i}, \quad \xi_i = - \sum_{j=i}^{\infty} \frac{1}{j} h_{j-i},$$

with g_j and h_j the coefficients of L^j in the expansion of $1/\phi(L)$ and $1/\theta(L)$, respectively, and Φ the Fisher information matrix for ϕ and θ .

Remark 3. Note that $\omega^2 < \pi^2/6$; hence comparing Lemmas 1 and 2, the LM test has lower local asymptotic power in the presence of short-run dynamics of any kind. As will be shown via simulations, the loss in power can be substantial. It remains, nevertheless, inevitable.

With the maximum likelihood estimate $\hat{\omega}$, we show that $\mathcal{LM}/\hat{\omega} \xrightarrow{d} \mathcal{N}(0, 1)$ as $T \rightarrow \infty$ under the null hypothesis. In particular, when $p = 1$ or $q = 1$, it is easy to compute $\hat{\omega}$. Since $v_j = \varsigma v_{j-1} + \epsilon_j$ in both cases, $g_j = \varsigma^j$,

$$\kappa_1 = \sum_{j=1}^{\infty} \frac{1}{j} \varsigma^{j-1} = -\frac{1}{\varsigma} \log(1 - \varsigma), \quad \text{and} \quad \Phi^{-1} = 1 - \varsigma^2.$$

Hence, we have

$$\omega^2 = \frac{\pi^2}{6} - \frac{1 - \varsigma^2}{\varsigma^2} (\log(1 - \varsigma))^2,$$

and $\hat{\omega}$ can be computed using $\hat{\varsigma}$. All these results remain valid for all trending models with a break considered. The relevant correction needed is a simple scaling by $\hat{\omega}$ so that the test becomes $\mathcal{LM}_{T,\hat{\lambda}}^* \equiv \mathcal{LM}_{T,\hat{\lambda}}/\hat{\omega}$.

Proposition 4. Suppose that the process $\{y_t\}$ is generated under the null hypothesis of Model A2 or A3, and that Assumptions 1 and 2 hold with η_t being an ARMA(p, q) process. Then, it holds that as $T \rightarrow \infty$, $\mathcal{LM}_{T,\hat{\lambda}}^* \xrightarrow{d} \mathcal{N}(0, 1)$ if $\hat{\lambda} - \lambda_b = o_p(T^{-1/2})$ and $\hat{\omega} - \omega = o_p(1)$.

The sufficient conditions in Proposition 4 follow from Lemma 2 and Proposition 1, hence the proof is omitted. The finite sample performance of $\mathcal{LM}_{T,\hat{\lambda}}^*$ with $\hat{\lambda} \in \{\hat{\lambda}_s, \hat{\lambda}_d, \hat{\lambda}_{tr}\}$ allowing for a structural break under both the null and alternative hypotheses will be examined in the next section.

5. Simulation Experiments

In this section, we present results from simulation experiments to illustrate the various theoretical results. Throughout the simulations, the true break fraction is set to $\lambda_b = 0.5$.⁵ The DGP is specified as $y_t = f_t + u_t$ where

$$f_t = \begin{cases} \mu_1 + \beta_1 t & \text{for Model A0,} \\ \mu_b C_t & \text{for Model A1,} \\ \beta_b B_t & \text{for Model A2,} \\ \mu_b C_t + \beta_b B_t & \text{for Model A3,} \end{cases}$$

⁵ Unreported simulation results with $\lambda_b = \{0.3, 0.7\}$ are qualitatively similar to those with $\lambda_b = 0.5$.

and $u_t = \Delta^{-1}\zeta_t^\# = \Delta^{-1}\eta_t\mathbf{1}_{t \geq 1}$, $t = 0, \pm 1, \pm 2, \dots$, where η_t is a short-memory zero-mean covariance stationary process that will be specified below. We set some parameters as follows: $\mu_1 = 1.72$, $\beta_1 = 0.03$, $\mu_b = 1$, and $\beta_b = 1$. The configurations are the same as those in PZ, chosen to obtain distributions that easily reveal the main features of interest. In all cases, the results are obtained via 10,000 replications. Additionally, 5% nominal size tests are considered.

First, to illustrate the effect of a structural break on the power of the fractional unit root test, we consider two different models when a structural change in slope is allowed in the DGP: (i) Model A0 (which ignores a relevant slope change); and (ii) Model A2 (which is well specified). The results are provided in Table 1. It is clear that the power of \mathcal{LM}_T is much lower than that of $\mathcal{LM}_{T,\hat{\lambda}_d}$, which supports the fact that a structural break in the DGP should be allowed when testing for a fractional unit root.

Table 1. The effect of a structural break in trend on the Lagrange multiplier (LM) tests.

$T = 150$											
d_0	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
\mathcal{LM}_T	0.767	0.617	0.458	0.290	0.187	0.117	0.070	0.044	0.029	0.020	0.004
$\mathcal{LM}_{T,\hat{\lambda}_d}$	1	1	0.997	0.986	0.949	0.858	0.703	0.485	0.303	0.169	0.036

Note: 5% nominal size tests are used. The data-generating process (DGP) is specified by $y_t = \beta_b B_t + u_t$ and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $\eta_t \sim i.i.d.\mathcal{N}(0,1)$ where $d_0 = m + d$, $d \in [-0.5, 0)$, and $m = 1$. The values of model parameters are set to $\lambda_0 = 0.5$, $\beta_b = 0.5$. The test \mathcal{LM}_T , designed for Model A1, ignores a structural break in the DGP, while the test $\mathcal{LM}_{T,\hat{\lambda}_d}$, designed for Model A2, is well specified with the estimate of the break date obtained from the dynamic regression (10).

Tables 2–5 present the rejection probabilities of the tests \mathcal{LM}_T and $\mathcal{LM}_{T,\hat{\lambda}}$ at the 5% significance level when $\eta_t \sim i.i.d.\mathcal{N}(0,1)$. In Table 2, no structural change is allowed in the DGP (Model A0); i.e., $y_t = \mu_1 + \beta_1 t + u_t$. The size of \mathcal{LM}_T is well controlled, which is 0.05 and 0.06 with sample sizes $T = 150, 500$, respectively. Table 3 reports the results for Model A1. The break fraction is not estimated consistently, because the level shift is negligible compared to the stochastic trend induced by the $I(1)$ errors. Hence, $\mathcal{LM}_{T,\hat{\lambda}_s}$ and $\mathcal{LM}_{T,\hat{\lambda}_d}$ suffer from severe size distortion, while \mathcal{LM}_T maintains size close to the nominal level 5%. Table 4 presents the results pertaining to Model A2. We also consider the test based on trimmed data, $\mathcal{LM}_{T,\hat{\lambda}_{tr}}$. The test $\mathcal{LM}_{T,\hat{\lambda}_d}$ is size-controlled, while the others show minor size distortion. However, the power of $\mathcal{LM}_{T,\hat{\lambda}_d}$ is always lower than that of the other two tests. Table 5 presents the results pertaining to Model A3. Here, we set $\mu_b = 0$ to consider the effect of an irrelevant level shift. Notice that $\mathcal{LM}_{T,\hat{\lambda}_s}$ exhibits liberal size distortion and $\mathcal{LM}_{T,\hat{\lambda}_d}$ also shows considerable size distortion. As noted by Chang and Perron (2016) [33], the estimate of the break date shows a pattern of bi-modality when an irrelevant level shift is introduced. This phenomenon is referred to the “contamination” effect, because the irrelevant level shift can make the estimate of the true break date less precise. By construction, the contamination effect is marginal on $\mathcal{LM}_{T,\hat{\lambda}_{tr}}$, whose exact size is 6.7% when $T = 500$.

Table 2. Rejection probabilities of the \mathcal{LM}_T test for Model A0.

d_0	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
$T = 150$											
1	1	1	0.997	0.983	0.924	0.778	0.544	0.307	0.154	0.051	
$T = 500$											
1	1	1	1	1	1	1	0.997	0.920	0.608	0.260	0.056

Note: 5% nominal size tests are used. The DGP is defined by $y_t = \mu_1 + \beta_1 t + u_t$, and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $\eta_t \sim i.i.d.\mathcal{N}(0,1)$ where $d_0 = m + d$, $d \in [-0.5, 0]$, and $m = 1$. The values of model parameters are set to $\mu_1 = 1.72$, $\beta_1 = 0.03$.

Table 3. Rejection probabilities of the $\mathcal{LM}_{T,\hat{\lambda}}$ and \mathcal{LM}_T tests for Model A1.

d_0	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
$T = 150$											
$\mathcal{LM}_{T,\hat{\lambda}_s}$	1	1	1	0.989	0.990	0.969	0.897	0.765	0.589	0.410	0.255
$\mathcal{LM}_{T,\hat{\lambda}_d}$	1	1	1	0.997	0.988	0.946	0.839	0.655	0.445	0.263	0.134
\mathcal{LM}_T	1	1	1	0.997	0.981	0.927	0.772	0.541	0.309	0.151	0.052
$T = 500$											
$\mathcal{LM}_{T,\hat{\lambda}_s}$	1	1	1	1	1	1	0.999	0.970	0.802	0.494	0.252
$\mathcal{LM}_{T,\hat{\lambda}_d}$	1	1	1	1	1	1	0.996	0.932	0.674	0.330	0.118
\mathcal{LM}_T	1	1	1	1	1	1	0.997	0.925	0.616	0.250	0.055

Note: 5% nominal size tests are used. The DGP is defined by $y_t = \mu_b C_t + u_t$, and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $\eta_t \sim i.i.d. \mathcal{N}(0,1)$, where $d_0 = m + d$, $d \in [-0.5, 0]$, and $m = 1$. The values of model parameters are set to $\lambda_0 = 0.5, \mu_b = 1$.

Table 4. Rejection probabilities of the $\mathcal{LM}_{T,\hat{\lambda}}$ tests for Model A2.

d_0	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
$T = 150$											
$\mathcal{LM}_{T,\hat{\lambda}_s}$	1	1	1	0.998	0.987	0.936	0.820	0.610	0.387	0.211	0.077
$\mathcal{LM}_{T,\hat{\lambda}_d}$	0.998	0.994	0.982	0.951	0.882	0.765	0.606	0.409	0.246	0.130	0.039
$\mathcal{LM}_{T,\hat{\lambda}_{tr}}$	1	0.999	0.999	0.997	0.981	0.922	0.794	0.587	0.369	0.205	0.072
$T = 500$											
$\mathcal{LM}_{T,\hat{\lambda}_s}$	1	1	1	1	1	1	0.995	0.904	0.567	0.229	0.063
$\mathcal{LM}_{T,\hat{\lambda}_d}$	1	1	1	1	0.997	0.974	0.893	0.717	0.426	0.165	0.048
$\mathcal{LM}_{T,\hat{\lambda}_{tr}}$	1	1	1	1	1	1	0.997	0.917	0.591	0.244	0.071

Note: 5% nominal size tests are used. The DGP is defined by $y_t = \beta_b B_t + u_t$, and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $\eta_t \sim i.i.d. \mathcal{N}(0,1)$, where $d_0 = m + d$, $d \in [-0.5, 0]$, and $m = 1$. The values of model parameters are set to $\lambda_0 = 0.5, \beta_b = 1$.

Table 5. Rejection probabilities of the $\mathcal{LM}_{T,\hat{\lambda}}$ tests for Model A3.

d_0	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
$T = 150$											
$\mathcal{LM}_{T,\hat{\lambda}_s}$	1	1	1	0.998	0.990	0.953	0.857	0.679	0.480	0.305	0.175
$\mathcal{LM}_{T,\hat{\lambda}_d}$	1	1	0.999	0.995	0.977	0.921	0.806	0.616	0.419	0.258	0.128
$\mathcal{LM}_{T,\hat{\lambda}_{tr}}$	1	1	1	0.996	0.981	0.926	0.798	0.587	0.380	0.218	0.083
$T = 500$											
$\mathcal{LM}_{T,\hat{\lambda}_s}$	1	1	1	1	1	1	0.995	0.924	0.659	0.350	0.156
$\mathcal{LM}_{T,\hat{\lambda}_d}$	1	1	1	1	0.999	0.996	0.974	0.851	0.561	0.259	0.101
$\mathcal{LM}_{T,\hat{\lambda}_{tr}}$	1	1	1	1	1	0.999	0.993	0.899	0.571	0.233	0.067

Note: 5% nominal size tests are used. The DGP is defined by $y_t = \beta_b B_t + u_t$, and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $\eta_t \sim i.i.d. \mathcal{N}(0,1)$, where $d_0 = m + d$, $d \in [-0.5, 0]$, and $m = 1$. The values of model parameters are set to $\lambda_0 = 0.5, \beta_b = 1$.

In Tables 6–8, we provide simulation results when the errors have short-run dynamics; i.e., $(1 - \rho L)\eta_t = \epsilon_t$, $\epsilon_t \sim i.i.d. \mathcal{N}(0,1)$. We set the value of the autoregressive (AR) parameter at $\rho \in \{-0.5, 0, 0.3, 0.6, 0.8\}$. When $\rho = 0$, we can compare the loss of power caused by allowing for dynamics when none is present. The other parameters remain unchanged. Table 6 reports the size and power of the Lagrange multiplier tests pertaining to Model A1, \mathcal{LM}_T . It is well size-controlled with less persistent AR parameters $\rho \in \{0, 0.3\}$, but it is very conservative with a higher AR coefficient $\rho \in \{0.6, 0.8\}$, while it shows liberal size distortions with $\rho = -0.5$. We find some interesting features in terms of power. First, power is higher when the AR parameter ρ is negative (in part due to the liberal size distortions). Second, as ρ becomes positive and large, power shrinks considerably. In particular, the loss of power is substantial when the AR parameter ρ increases from 0.6 to 0.8. This implies that

a sufficiently large time span is needed to distinguish fractional integration from weak dependence. Comparing Table 3 with Table 6, for the $\rho = 0$ case, it is obvious that power is substantially lower when an irrelevant AR parameter is introduced. This result suggests that selecting the number of lags in the noise component is crucial to obtain good power. Lastly, with a persistent AR parameter $\rho = 0.8$, the LM tests have non-monotonic power; that is, power does not increase when the order of integration d_0 moves away from the null of a unit root. As also discussed in Lobato and Velasco (2007) [16], it is difficult to distinguish fractional integration from a highly persistent stationary short-memory process. Table 7 reports the results pertaining to Model A2. They show similar patterns as for Model A1. It is noticeable that $\mathcal{LM}_{T,\hat{\lambda}_d}^*$ performs well in terms of size across all cases, while its power is always lower than that of the other tests. Table 8 presents the results for Model A3. $\mathcal{LM}_{T,\hat{\lambda}_d}^*$ has size distortion, even with negative and less persistent AR parameters. This happens because when using the dynamic regression to estimate the break fraction, the estimate $\hat{\lambda}_d$ is negatively biased for Model A3. Among the three LM statistics, $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$ based on the trimmed data performs well in finite samples. The size of $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$ is well controlled across various values of ρ , while $\mathcal{LM}_{T,\hat{\lambda}_s}^*$ and $\mathcal{LM}_{T,\hat{\lambda}_d}^*$ show liberal size distortions. Moreover, the power loss of $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$ is minor relative to the other tests. Hence, $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$ and $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$ are the recommended tests in practice.

Table 6. Rejection probabilities of $\mathcal{LM}_{T,\hat{\lambda}}^*$ and \mathcal{LM}_T for Model A1 with short-run dynamics.

		d_0										
ρ		0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
$T = 150$												
−0.5	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	0.999	0.997	0.987	0.965	0.894	0.776	0.618	0.449	0.305	0.187
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	0.999	0.997	0.986	0.962	0.891	0.769	0.598	0.429	0.277	0.168
	\mathcal{LM}_T	1	0.999	0.994	0.974	0.928	0.813	0.639	0.442	0.259	0.137	0.059
0	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.993	0.986	0.967	0.927	0.862	0.774	0.656	0.528	0.424	0.329	0.244
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.991	0.981	0.955	0.904	0.824	0.711	0.582	0.435	0.321	0.224	0.143
	\mathcal{LM}_T	0.986	0.969	0.930	0.851	0.728	0.572	0.411	0.267	0.163	0.102	0.046
0.3	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.946	0.899	0.839	0.767	0.671	0.553	0.450	0.365	0.272	0.215	0.158
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.922	0.863	0.788	0.683	0.573	0.441	0.337	0.245	0.168	0.115	0.069
	\mathcal{LM}_T	0.894	0.811	0.707	0.576	0.445	0.303	0.209	0.133	0.077	0.054	0.026
0.6	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.558	0.484	0.385	0.296	0.234	0.163	0.127	0.091	0.061	0.048	0.029
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.456	0.370	0.274	0.197	0.144	0.086	0.060	0.039	0.024	0.020	0.010
	\mathcal{LM}_T	0.361	0.277	0.184	0.121	0.079	0.042	0.029	0.018	0.013	0.013	0.007
0.8	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.096	0.074	0.049	0.035	0.025	0.017	0.014	0.010	0.012	0.011	0.008
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.044	0.033	0.019	0.013	0.014	0.012	0.015	0.022	0.024	0.027	0.036
	\mathcal{LM}_T	0.019	0.016	0.011	0.012	0.018	0.026	0.033	0.041	0.042	0.038	0.026
$T = 500$												
−0.5	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	1	1	1	0.999	0.979	0.856	0.581	0.305	0.129
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	1	1	1	0.999	0.974	0.829	0.537	0.262	0.100
	\mathcal{LM}_T	1	1	1	1	1	1	0.980	0.830	0.502	0.215	0.066
0	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	1	1	0.999	0.992	0.936	0.781	0.576	0.403	0.263
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	1	1	0.999	0.985	0.894	0.695	0.451	0.273	0.143
	\mathcal{LM}_T	1	1	1	1	0.999	0.979	0.858	0.595	0.330	0.168	0.058
0.3	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	1	0.998	0.979	0.920	0.777	0.575	0.394	0.257	0.155
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	1	0.996	0.967	0.869	0.684	0.447	0.261	0.162	0.074
	\mathcal{LM}_T	1	1	1	0.995	0.957	0.837	0.611	0.376	0.205	0.124	0.051
0.6	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.984	0.956	0.880	0.736	0.559	0.348	0.195	0.114	0.056	0.033	0.016
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.977	0.935	0.844	0.673	0.487	0.297	0.171	0.105	0.065	0.045	0.023
	\mathcal{LM}_T	0.974	0.927	0.833	0.670	0.497	0.316	0.199	0.135	0.083	0.058	0.028
0.8	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.249	0.148	0.075	0.040	0.024	0.015	0.008	0.007	0.005	0.006	0.004
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.202	0.133	0.084	0.067	0.066	0.069	0.084	0.091	0.101	0.107	0.103
	\mathcal{LM}_T	0.217	0.159	0.126	0.120	0.142	0.157	0.171	0.163	0.134	0.102	0.043

Note: 5% nominal size tests are used. The DGP is defined by $y_t = \mu_b C_t + u_t$ and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $(1 - \rho_L)\eta_t = \epsilon_t$ and $\epsilon_t \sim i.i.d. \mathcal{N}(0, 1)$, where $d_0 = m + d$, $d \in [-0.5, 0]$. We set $\mu_b = 1$.

Table 7. Rejection probabilities of $\mathcal{LM}_{T,\hat{\lambda}}^*$ for Model A2 with short-run dynamics.

		d_0										
ρ		0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
$T = 150$												
−0.5	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	0.999	0.996	0.985	0.949	0.861	0.713	0.515	0.354	0.215	0.100
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.993	0.982	0.955	0.908	0.816	0.680	0.521	0.350	0.233	0.136	0.054
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	0.996	0.990	0.969	0.926	0.822	0.671	0.478	0.335	0.203	0.098
0	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.991	0.978	0.943	0.885	0.795	0.652	0.507	0.364	0.268	0.189	0.097
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.912	0.851	0.779	0.668	0.562	0.425	0.311	0.205	0.153	0.104	0.046
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.988	0.965	0.931	0.868	0.774	0.645	0.496	0.373	0.280	0.205	0.113
0.3	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.907	0.848	0.761	0.662	0.533	0.429	0.319	0.239	0.170	0.129	0.071
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.691	0.602	0.502	0.409	0.311	0.230	0.163	0.115	0.076	0.056	0.027
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.868	0.792	0.696	0.594	0.467	0.369	0.271	0.198	0.141	0.109	0.055
0.6	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.436	0.360	0.282	0.215	0.157	0.119	0.083	0.061	0.050	0.038	0.027
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.230	0.176	0.130	0.092	0.058	0.044	0.028	0.020	0.019	0.016	0.012
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.421	0.344	0.269	0.211	0.156	0.118	0.082	0.062	0.054	0.043	0.030
0.8	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.057	0.040	0.038	0.038	0.044	0.053	0.061	0.063	0.058	0.044	0.028
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.022	0.012	0.013	0.017	0.025	0.032	0.044	0.047	0.044	0.034	0.023
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.040	0.026	0.024	0.025	0.034	0.039	0.049	0.053	0.051	0.038	0.024
$T = 500$												
−0.5	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	1	1	1	0.999	0.963	0.785	0.445	0.197	0.068
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	1	1	0.993	0.952	0.824	0.594	0.311	0.129	0.039
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	1	1	1	1	1	0.975	0.825	0.490	0.221	0.081
0	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	1	1	0.998	0.969	0.831	0.575	0.329	0.169	0.076
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	0.993	0.968	0.907	0.793	0.617	0.408	0.225	0.108	0.050
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	1	1	1	0.998	0.978	0.865	0.617	0.362	0.190	0.099
0.3	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	0.999	0.994	0.950	0.826	0.614	0.382	0.233	0.145	0.081
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.993	0.974	0.929	0.846	0.746	0.613	0.427	0.255	0.149	0.081	0.049
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	1	0.999	0.994	0.957	0.836	0.625	0.391	0.236	0.145	0.079
0.6	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.975	0.930	0.828	0.687	0.521	0.366	0.257	0.178	0.125	0.099	0.065
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.788	0.709	0.596	0.475	0.340	0.222	0.144	0.091	0.060	0.040	0.030
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.976	0.933	0.834	0.690	0.514	0.353	0.240	0.161	0.113	0.090	0.067
0.8	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.245	0.187	0.161	0.169	0.191	0.218	0.238	0.226	0.176	0.122	0.067
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.143	0.104	0.087	0.085	0.084	0.092	0.101	0.117	0.103	0.082	0.048
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.235	0.165	0.130	0.124	0.127	0.142	0.161	0.190	0.191	0.199	0.226

Note: 5% nominal size tests are used. The DGP is defined by $y_t = \beta_b B_t + u_t$ and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $(1 - \rho L)\eta_t = \epsilon_t$ and $\epsilon_t \sim i.i.d. \mathcal{N}(0, 1)$, where $d_0 = m + d$, $d \in [-0.5, 0]$. We set $\beta_b = 1$.

Table 8. Rejection probabilities of $\mathcal{LM}_{T,\hat{\lambda}}^*$ for Model A3 with short-run dynamics.

		d_0										
ρ		0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
$T = 150$												
−0.5	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	0.999	0.996	0.986	0.952	0.880	0.749	0.561	0.403	0.272	0.149
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	0.999	0.996	0.983	0.946	0.871	0.738	0.549	0.390	0.258	0.139
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	0.996	0.989	0.970	0.925	0.832	0.686	0.491	0.339	0.219	0.096
0	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.992	0.980	0.946	0.897	0.813	0.697	0.569	0.441	0.344	0.262	0.183
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.991	0.973	0.939	0.881	0.786	0.670	0.537	0.401	0.303	0.229	0.149
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.983	0.963	0.919	0.848	0.740	0.609	0.473	0.343	0.250	0.184	0.095
0.3	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.919	0.868	0.785	0.695	0.587	0.488	0.377	0.302	0.230	0.183	0.137
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.899	0.835	0.747	0.646	0.518	0.425	0.307	0.236	0.167	0.121	0.083
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.877	0.809	0.710	0.607	0.483	0.381	0.274	0.208	0.158	0.119	0.069
0.6	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.481	0.396	0.318	0.251	0.191	0.150	0.115	0.086	0.072	0.055	0.039
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.407	0.317	0.241	0.173	0.125	0.088	0.063	0.046	0.032	0.026	0.016
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.385	0.298	0.242	0.175	0.127	0.092	0.071	0.051	0.042	0.032	0.021
0.8	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.082	0.059	0.052	0.039	0.036	0.033	0.032	0.033	0.027	0.024	0.019
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.045	0.028	0.023	0.015	0.014	0.014	0.016	0.019	0.024	0.026	0.030
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.040	0.030	0.030	0.026	0.031	0.042	0.044	0.050	0.044	0.039	0.026
$T = 500$												
−0.5	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	1	1	1	0.999	0.970	0.807	0.501	0.243	0.103
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	1	1	1	0.993	0.936	0.741	0.434	0.199	0.079
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	1	1	1	1	0.999	0.972	0.801	0.476	0.212	0.073
0	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	1	1	0.999	0.979	0.876	0.674	0.439	0.286	0.164
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	1	1	0.991	0.953	0.829	0.622	0.379	0.226	0.124
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	1	1	1	0.998	0.970	0.835	0.591	0.326	0.177	0.077
0.3	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	1	1	0.999	0.995	0.962	0.859	0.673	0.468	0.307	0.192	0.123
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	1	1	0.997	0.978	0.925	0.792	0.580	0.372	0.215	0.124	0.076
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	1	1	0.999	0.992	0.945	0.812	0.590	0.371	0.222	0.131	0.078
0.6	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.978	0.935	0.843	0.697	0.521	0.348	0.220	0.140	0.091	0.070	0.050
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.942	0.864	0.739	0.572	0.393	0.240	0.138	0.080	0.050	0.035	0.025
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.971	0.917	0.818	0.667	0.493	0.343	0.231	0.162	0.108	0.087	0.063
0.8	$\mathcal{LM}_{T,\hat{\lambda}_s}^*$	0.232	0.153	0.114	0.097	0.095	0.099	0.105	0.093	0.071	0.052	0.052
	$\mathcal{LM}_{T,\hat{\lambda}_d}^*$	0.162	0.101	0.067	0.049	0.041	0.039	0.039	0.047	0.057	0.071	0.067
	$\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$	0.225	0.173	0.144	0.145	0.154	0.176	0.203	0.190	0.155	0.105	0.055

Note: 5% nominal size tests are used. The DGP is defined by $y_t = \beta_b B_t + u_t$ and $u_t = \Delta^{-m}(\Delta^{-d}\eta_t)\mathbf{1}_{t \geq 1}$ with $(1 - \rho L)\eta_t = \epsilon_t$ and $\epsilon_t \sim i.i.d.\mathcal{N}(0, 1)$, where $d_0 = m + d$, $d \in [-0.5, 0]$. We set $\beta_b = 1$.

The Size and Power When a Pre-Test Is Used

In Figures 1–4, we present the size and power of the LM tests as the slope change parameter (β_b) changes in Models A2 and A3 with and without the use of a pre-test. As a pre-test, we use the *SW* test of Iacone et al. (2013) [37] at the nominal 5% level. We only consider the version of the LM statistics based on the trimmed estimate of the break fraction, denoted $\mathcal{LM}_{T,\hat{\lambda}_{tr}}$ and $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^p$ when no short-run dynamics is allowed, and by $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$ and $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^{*p}$ when an AR(1) structure is allowed. To assess the extent of the differences in the size distortion and power, we also report the infeasible LM test based on the true value of break fraction, denoted $\mathcal{LM}_{T,\lambda_b}$ and $\mathcal{LM}_{T,\lambda_b}^*$. The results are presented in Figure 1 for Model A2 (no short-run dynamics), Figure 2 for Model A2, and in Figures 3 and 4 for Models A2 and A3 with short-run dynamics of the form $(1 - 0.3L)\eta_t = \epsilon_t$, $\epsilon_t \sim i.i.d.\mathcal{N}(0, 1)$. For $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^p$ and $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^{*p}$, the trimming window contains six observations (the simulation results are not sensitive to the length of the window). The magnitude of break in the slope

of the trend β_b varies from -4 to 4 in increments of 0.2 . We set $\mu_b = 0$ for Model A3. The sample sizes are $T = 150, 300$, and the number of replications is $10,000$ for each value of the parameters. One-sided tests against the alternative hypothesis $H_1 : d_0 < 1$ are constructed at the nominal 5% level. For power, d_0 is set to 0.8 .

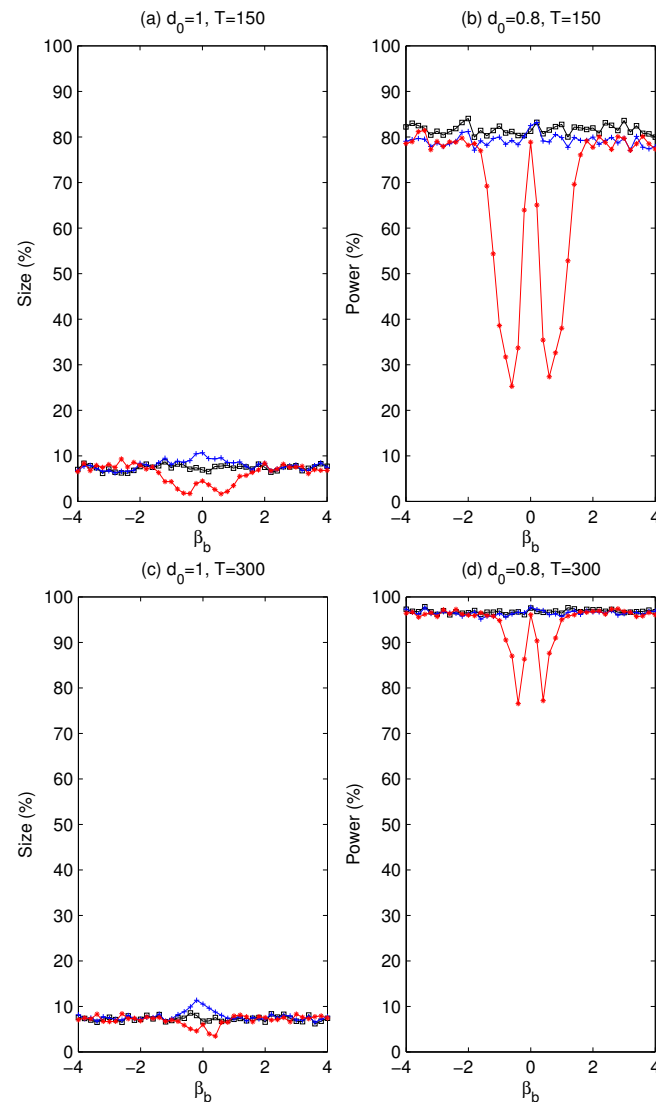


Figure 1. Null rejection probabilities in Model A2, $\lambda_b = 0.5$ [$- + -$: $\mathcal{LM}_{T, \hat{\lambda}_{tr}}$, $- * -$: $\mathcal{LM}_{T, \hat{\lambda}_{tr}}^p$, $- \square -$: $\mathcal{LM}_{T, \lambda_b}$]

The results for Model A2 (presented in Figure 1) show first that the version without the pre-test exhibits some liberal size distortions when β_b is near 0, which reduce when T increases, though remain noticeable even with $T = 300$. On the other hand, the version with the pre-test exhibits conservative size-distortions when β_b is near but not equal to 0, which again reduce but remain noticeable when $T = 300$. The most drastic differences occur when considering the power of the tests. The power of the version without the pre-test is slightly below but near to the power of the version with the true break fraction when $T = 150$ for all values of β_b . When $T = 300$, the power functions are nearly the same. Things are very different when the version with the pre-test is used. When β_b is near but not equal to zero, the power reduces drastically, creating pronounced power valleys. This reduction in power alleviates somewhat when $T = 300$, but remains important. This is due to the fact that for low values of β_b , the *SW* test of Iacone et al. (2013) [37] is not powerful enough, so a change in slope regressor is

not included. Yet, the magnitude of the change in slope is large enough to induce a considerable loss in power. This is akin to the problem faced by the Kim and Perron (2009) [10] test in the context of testing for a unit root.

The results for Model A3 (presented in Figure 2) show a similar picture. This is also the case when considering the tests $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$ and $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^{*p}$ with serial correlation in the DGP (Figures 3 and 4), though here the size distortions of both tests are somewhat higher, liberal for $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$, and conservative for $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^{*p}$. The power losses of $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^{*p}$ is severe, especially when $T = 150$.

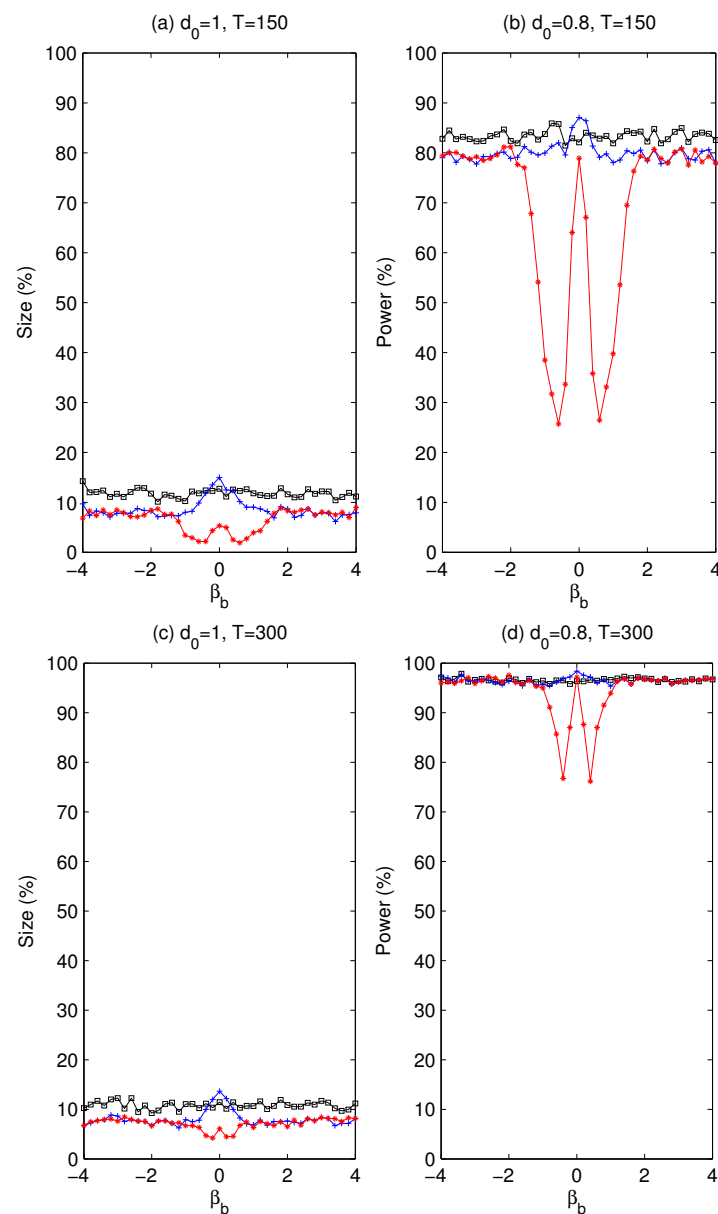


Figure 2. Null rejection probabilities in Model A3, $\lambda_b = 0.5$ [—+— : $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^*$, —*— : $\mathcal{LM}_{T,\hat{\lambda}_{tr}}^{*p}$, —□— : $\mathcal{LM}_{T,\lambda_b}$]

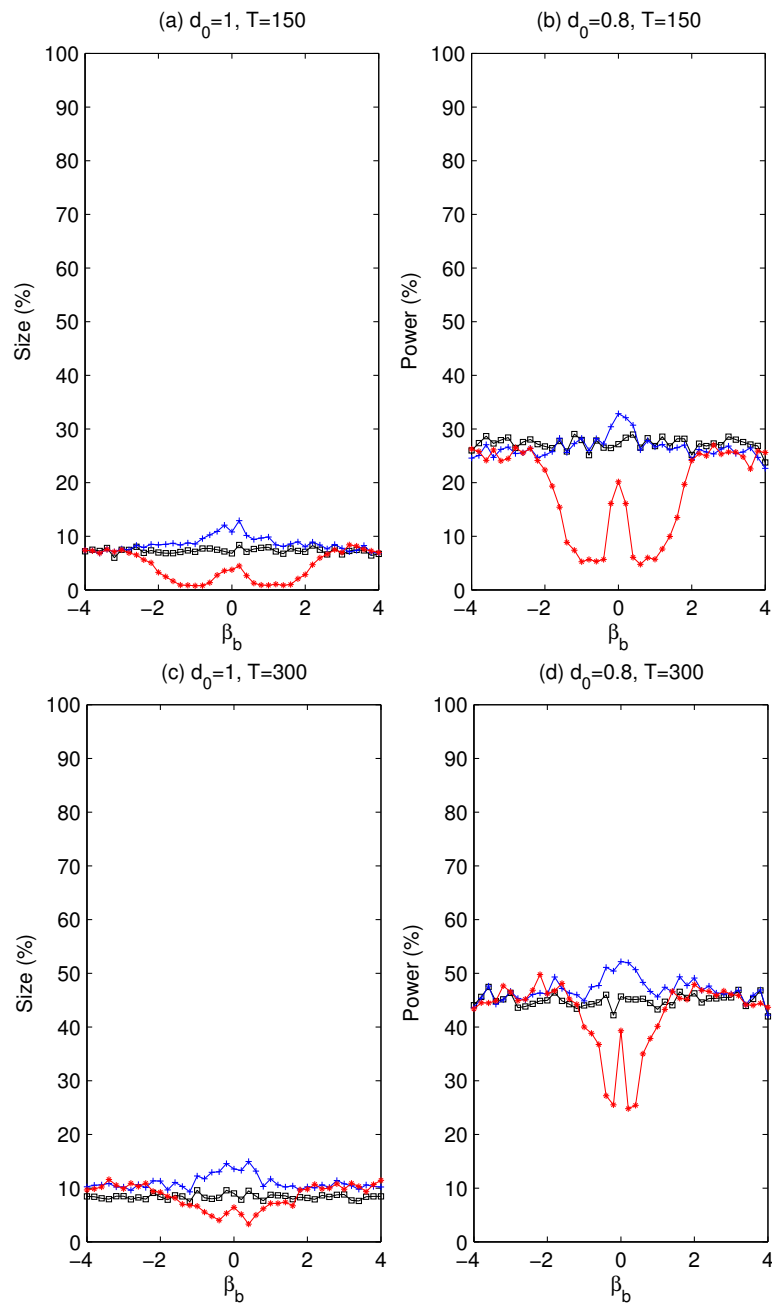


Figure 3. Null rejection probabilities in Model A2, where $(1 - 0.3L)\eta_t = \epsilon_t$, $\lambda_b = 0.5$ [$- + -$: $\mathcal{LM}_{T, \hat{\lambda}_{tr}}$, $- * -$: $\mathcal{LM}_{T, \hat{\lambda}_{tr}}^{*p}$, $- \square -$: $\mathcal{LM}_{T, \lambda_b}$].

Based on the simulation results, it is recommended to use the $\mathcal{LM}_{T, \hat{\lambda}_{tr}}$ or $\mathcal{LM}_{T, \hat{\lambda}_{tr}}^{*p}$ tests without the pre-test. In our view, the reduction in power when using a pre-test considerably outweighs the differences in size distortions. The *SW* test of Iacone et al. (2013) [37] is nevertheless still useful to assess the presence of large breaks.

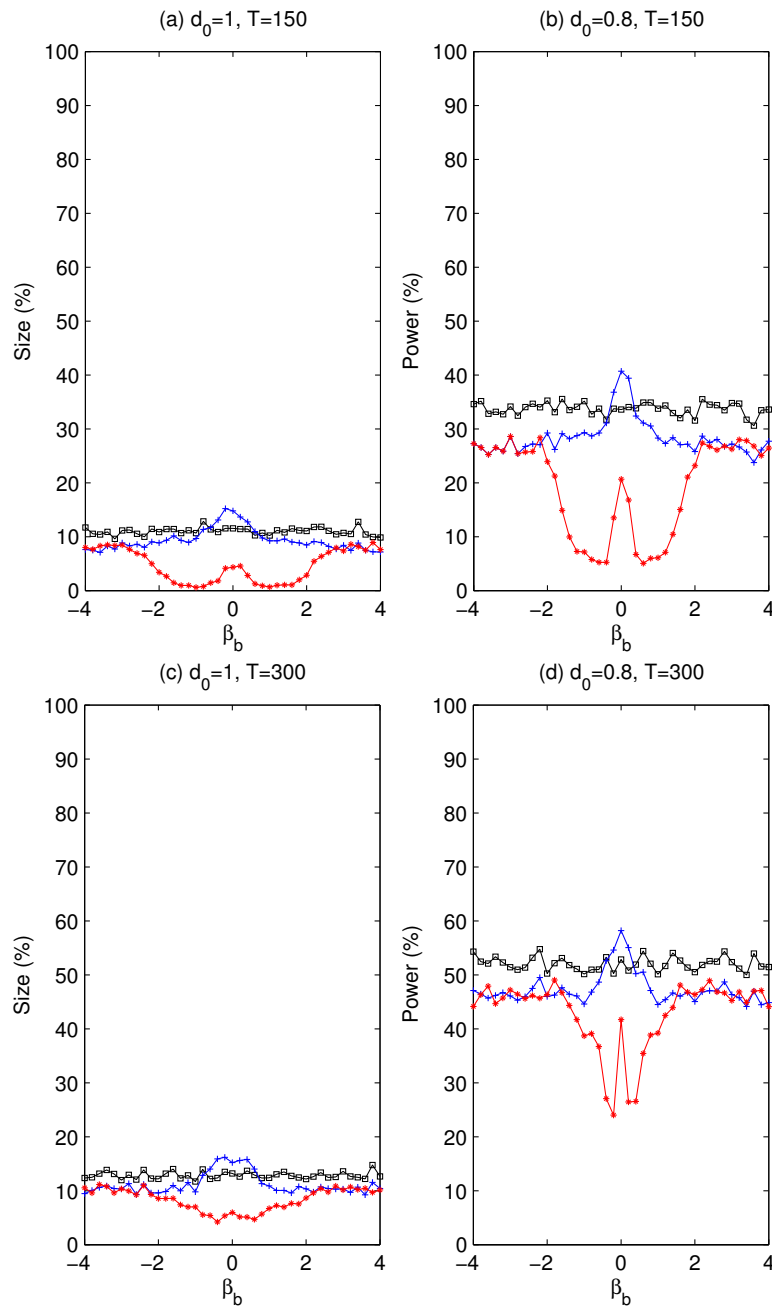
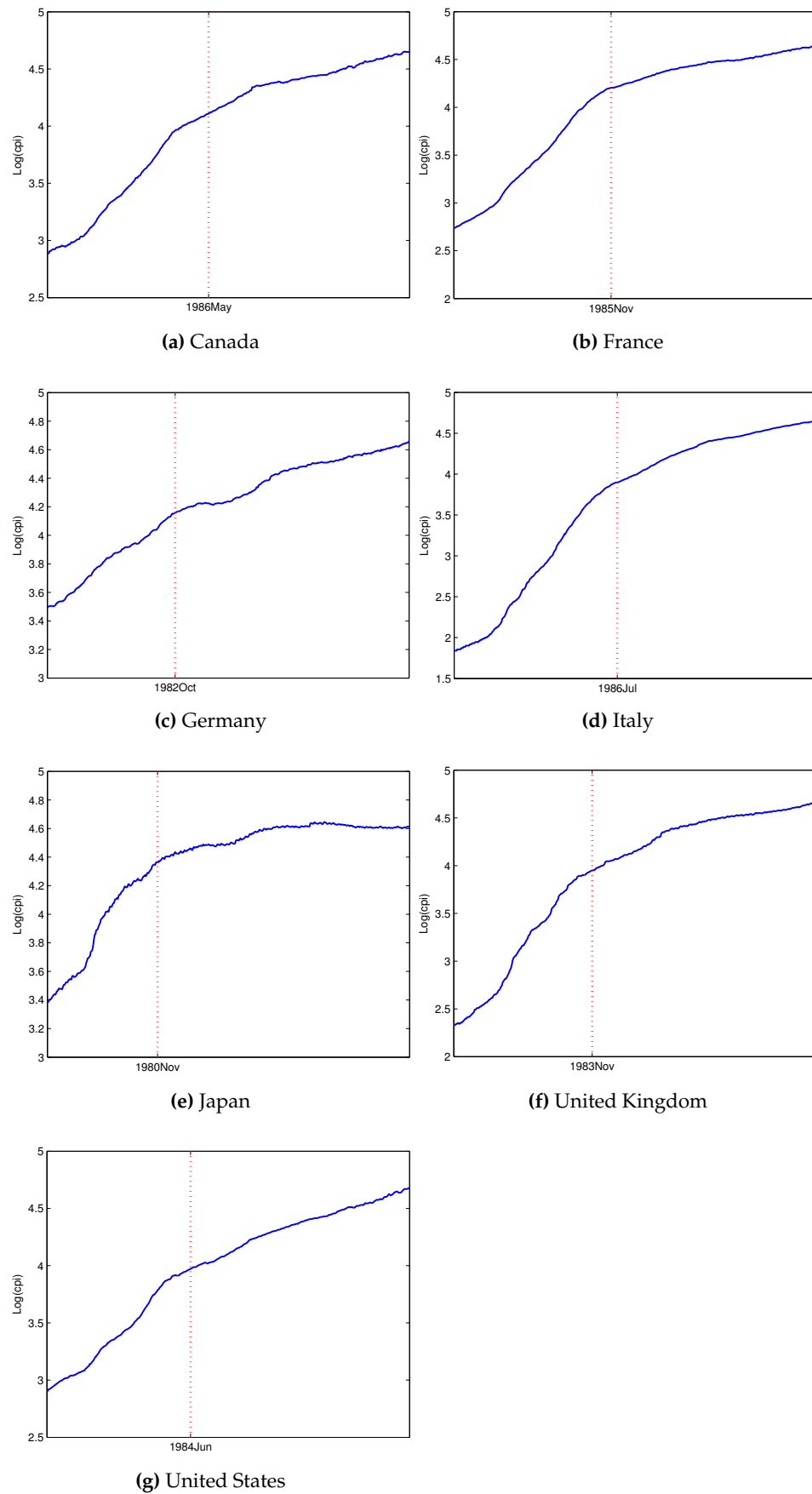


Figure 4. Null rejection probabilities in Model A3, where $(1 - 0.3L)\eta_t = \epsilon_t$, $\lambda_b = 0.5$ [$+-$: $\mathcal{LM}_{T, \hat{\lambda}_{tr}}$, $-*-$: $\mathcal{LM}_{T, \hat{\lambda}_{tr}}^{*p}$, $-\square-$: $\mathcal{LM}_{T, \lambda_b}$].

6. An Empirical Application

We analyze the aggregate price indices of the G7 countries. Monthly seasonally-adjusted CPI series were obtained from the OECD Main Economic Indicators. All series are analyzed with a logarithm transformation and are plotted in Figure 5, where the vertical line is the break date estimated by minimizing the sum of squared residuals from Model A2.

**Figure 5.** (log) CPI for the G7 Countries.

The results are presented in Table 9. We only consider Model A2, and use the test for a slope change of Iacone et al. (2013) [37]. Based on the simulation results, we recommend using the LM tests with $\hat{\lambda}_{tr}$. We present results with and without short-run dynamics. When dynamics is allowed, an AR(1) specification is used. The time span is from January 1969 to December 2007 ($T = 468$). First, the augmented Dickey-Fuller type test (ADF) cannot reject the null of a unit root against the alternative of trend stationarity for all G7 countries. On the other hand, the SW test of Iacone et al. (2013) [37] detects a change in the slope of the trend. Allowing for a structural change in trend, the fractional unit root tests $\mathcal{LM}_{T, \hat{\lambda}_{tr}}$ and $\mathcal{LM}_{T, \hat{\lambda}_{tr}}^*$ lead to a rejection of the unit root in favor of fractional integration. Specifically, the test results state that the order of fractional integration is greater than unity for all G7 countries. We apply the two-step feasible exact local Whittle estimator \hat{d}_{ELW} of Shimotsu (2010) [40] to the residuals from the fitted trend equipped with the estimate of the break date \hat{T}_b . This result is compatible with that in Gil-Alana (2008) [41], where he estimated the order of fractional integration for the U.S. CPI and found that the confidence intervals were located above unity. Hassler and Wolters (1995) [42] considered the inflation rates for various countries, and found that the order of fractional integration was located in an interval $(0, 0.5)$; that is, the inflation rate is a long-memory process.

Table 9. Empirical results for the logarithmic price indices of the G7 countries.

	ADF	SW	Break Date	\hat{d}_{ELW}	$\mathcal{LM}_{T, \hat{\lambda}_{tr}}$	$\mathcal{LM}_{T, \hat{\lambda}_{tr}}^*$
Canada	−0.16 (4)	19.40 ***	1986 May [0.45]	1.30	9.44 ***	9.37 *** (0.13)
France	−1.47 (6)	28.19 ***	1985 Nov [0.43]	1.69	34.65 ***	7.54 *** (0.62)
Germany	−1.59 (1)	10.78 **	1982 Oct [0.35]	1.32	18.98 ***	8.01 *** (0.41)
Italy	−0.61 (7)	16.37 ***	1986 Jul [0.45]	1.54	29.55 ***	11.64 *** (0.52)
Japan	−2.50 (11)	18.86 ***	1980 Nov [0.31]	1.65	18.93 ***	8.73 *** (0.42)
U.K.	−1.47 (7)	16.50 ***	1983 Nov [0.38]	1.51	9.46 ***	4.39 *** (0.5)
U.S.A.	−0.44 (2)	15.75 **	1984 Jun [0.40]	1.29	12.71 ***	3.48 *** (0.38)

Note: (1) The numbers in parentheses are the values of the autoregressive order selected by the Bayesian information criterion when constructing the ADF test; (2) the numbers in brackets $[\cdot]$ denote the estimated break fractions; (3) the numbers in $\langle \cdot \rangle$ are the estimates of the AR coefficient in the noise component. *, **, and *** denote a statistic significant at the 10%, 5%, and 1% level, respectively.

7. Conclusions

We established testing procedures for a fractional unit root, allowing for a structural change under both the null and alternative hypotheses. Following Robinson (1994) [13], Tanaka (1999) [14] derived a Lagrange multiplier test in the time domain, and Dolado et al. (2002, 2008) [15,19] and Lobato and Velasco (2007) [16] considered Wald-type tests for a unit root null hypothesis against fractional integration. Although Dolado et al. (2008) [19] introduced deterministic components, the case with a structural break in trend has not been considered in the literature. In contrast to the large amount of work related to testing the null hypothesis of long-memory against the alternative of stationarity with level shifts, and vice versa, work related to a fractional unit root test allowing for a structural break in trend is more scarce. To the best of our knowledge, this paper is the first that addresses testing for a fractional process allowing a structural break under both the null and alternative hypotheses.

Fractional unit root tests allowing for a structural break under both the null and alternative hypotheses have some desirable features: (i) given that economic variables are often subject to structural changes, our approach imposes a symmetric treatment of the change under both the null and alternative hypotheses; (ii) it is not required to distinguish long memory from structural change; (iii) the power of fractional unit root tests can be substantially improved when a break is actually present. Under some conditions, the proposed LM test statistics have the standard normal limit under the null hypothesis. Simulation experiments confirmed that the tests have good size and power. Hence, we believe that our procedures offer useful complements to existing tests and should be used in practical applications.

An extension of practical interest is to allow $I(d_0)$ processes under the null hypothesis, where $I(1)$ processes are included as a special case. The sufficient condition for the LM test statistic to have the standard normal limit may be different from that in Proposition 1. Recently, Chang and Perron

(2016) [33] extended PZ's analysis to cover the more general case of fractionally integrated errors for values of d_0 in the interval $(-0.5, 1.5)$ excluding the boundary case 0.5. In particular, they established the rate of the convergence of $\hat{\lambda}_s$ from the static regression [33] (Theorem 2). It is also important to examine the performance of $\hat{\lambda}_d$ and $\hat{\lambda}_{tr}$ under the null of $I(d_0)$ processes. Such investigations, and others, are the object of the ongoing subject.

Acknowledgments: We are grateful to academic editor Steve Cook, two anonymous referees, Tomoyoshi Yabu and participants at the 11th International Symposium on Econometric Theory and Applications (SETA 2015) for constructive comments, and Fabrizio Iacone, Dukpa Kim, and Katsumi Shimotsu for sharing their GAUSS and MATLAB codes. Chang's research was partly supported by the National Natural Science Foundation of China Grants No. 71131008 (Key Project).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Theorem A1. The DGP is specified by (4), that is, for $t = 1, \dots, T$,

$$y_t = \mu_1 + \mu_b C_t + u_t, \quad \mu_b \neq 0.$$

Under H_0 , take first differences of y_t and define $D(T_b)_t = 1$ if $t = T_b + 1$ and 0 otherwise. Then,

$$\begin{aligned} T^{-1} \sum_{t=1}^T (\Delta y_t)^2 &= T^{-1} \sum_{t=1}^T (\mu_b D(T_b)_t + \eta_t)^2 \\ &= T^{-1} \left(\sum_{t=1}^T \eta_t^2 + 2 \sum_{t=1}^T \mu_b D(T_b)_t \eta_t + \sum_{t=1}^T \mu_b^2 D(T_b)_t^2 \right) \\ &= T^{-1} \left(\sum_{t=1}^T \eta_t^2 + 2\mu_b \eta_{T_b+1} + \mu_b^2 \right) \\ &= T^{-1} \sum_{t=1}^T \eta_t^2 + o_p(1) \end{aligned}$$

and

$$\begin{aligned} T^{-1} \sum_{j=2}^T (-\log \Delta y_j) \Delta y_j &= T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} (\mu_b D(T_b)_{j-k} + \eta_{j-k}) \right) (\mu_b D(T_b)_j + \eta_j) \\ &= T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \mu_b D(T_b)_{j-k} + \sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) (\mu_b D(T_b)_j + \eta_j) \\ &= T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \mu_b^2 D(T_b)_{j-k} \right) D(T_b)_j + T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \mu_b D(T_b)_{j-k} \right) \eta_j \\ &\quad + T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \mu_b D(T_b)_j + T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \eta_j \\ &\equiv A_1 + A_2 + A_3 + A_4. \end{aligned}$$

It is easy to show that $A_1 = 0$. For A_2 , we have

$$\begin{aligned} (A_2)^2 &= \mu_b^2 \frac{1}{T^2} \left(\sum_{k=1}^{T-(T_b+1)} \frac{1}{k} \eta_{T_1+1+k} \right)^2 \leq \mu_b^2 \frac{1}{T^2} \left(\sum_{k=1}^{T-(T_b+1)} \frac{1}{k^2} \right) \left(\sum_{k=1}^{T-(T_b+1)} \eta_{T_1+1+k}^2 \right) \\ &\leq \mu_b^2 \frac{\pi^2}{6} \frac{1}{T^2} \left(\sum_{k=1}^{T-(T_b+1)} \eta_{T_1+1+k}^2 \right) = \mu_b^2 \frac{\pi^2}{6} \left(\frac{1}{T} \sum_{k=1}^{T-(T_b+1)} \eta_{T_1+1+k}^2 \right) \frac{1}{T} = O_p(T^{-1}) = o_p(1), \end{aligned}$$

where the first inequality follows from the Cauchy-Schwarz inequality, and the second inequality holds because $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$ (see Tanaka (1999) [14]). Because η_t is a short-memory zero-mean

covariance stationary process, it is straightforward to show that $T^{-1} \sum_{k=1}^{T-(T_b+1)} \eta_{T_1+1+k}^2 = O_p(1)$. By the continuous mapping theorem, $|A_2| = o_p(1)$, which implies that $A_2 = o_p(1)$. Similarly,

$$\begin{aligned} (A_3)^2 &= \mu_b^2 \frac{1}{T^2} \left(\sum_{k=1}^{T_b} \frac{1}{k} \eta_{T_b+1-k} \right)^2 \leq \mu_b^2 \frac{1}{T^2} \left(\sum_{k=1}^{T_b} \frac{1}{k^2} \right) \left(\sum_{k=1}^{T_b} \eta_{T_b+1-k}^2 \right) \\ &\leq \mu_b^2 \frac{\pi^2}{6} \left(\frac{1}{T} \sum_{k=1}^{T_b} \eta_{T_b+1-k}^2 \right) \frac{1}{T} = o_p(1). \end{aligned}$$

Given the results previously,

$$\mathcal{LM}_M = \sqrt{T} \sqrt{\frac{6}{\pi^2}} \frac{T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \eta_j + o_p(1)}{T^{-1} \sum_{j=2}^T \eta_j^2 + o_p(1)} = \sqrt{T} \sqrt{\frac{6}{\pi^2}} \sum_{k=1}^T \frac{1}{k} \rho_k + o_p(1),$$

hence following Lemma A1 and Slutsky's theorem, we have $\mathcal{LM}_M \xrightarrow{d} \mathcal{N}(0, 1)$. \square

Proof of Theorem A2. Under $H_0 : d_0 = 1$, $\Delta y_t = \beta_1 + \Delta u_t = \beta_1 + \eta_t \mathbf{1}_{t \geq 1}$. Define $\widetilde{\Delta y}_t = \Delta y_t - \hat{\beta}_1$. First, consider the denominator of \mathcal{LM}_T . Under H_0 , conditioning on $y_0 = 0$,

$$\begin{aligned} T^{-1} \sum_{t=1}^T \left(\widetilde{\Delta y}_t \right)^2 &= T^{-1} \sum_{t=1}^T (\Delta y_t - \hat{\beta}_1)^2 = T^{-1} \sum_{t=1}^T (\beta_1 - \hat{\beta}_1 + \eta_t)^2 \\ &= (\beta_1 - \hat{\beta}_1)^2 + 2(\beta_1 - \hat{\beta}_1) T^{-1} \sum_{t=1}^T \eta_t + T^{-1} \sum_{t=1}^T \eta_t^2 \\ &= O_p(T^{-1}) + o_p(1) + T^{-1} \sum_{t=1}^T \eta_t^2 = T^{-1} \sum_{t=1}^T \eta_t^2 + o_p(1), \end{aligned}$$

where we use the fact that $\hat{\beta}_1 - \beta_1 = O_p(T^{-1/2})$, and $T^{-1} \sum_{t=1}^T \eta_t \xrightarrow{p} 0$ by the weak law of large numbers. Second, the numerator of \mathcal{LM}_T is given by

$$\begin{aligned} T^{-1} \sum_{j=2}^T \left(-\log \Delta \widetilde{\Delta y}_j \right) \widetilde{\Delta y}_j &= T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \widetilde{\Delta y}_{j-k} \right) \widetilde{\Delta y}_j \\ &= T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} (\Delta y_{j-k} - \hat{\beta}_1) \right) (\Delta y_j - \hat{\beta}_1) \\ &= T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} (\beta_1 - \hat{\beta}_1 + \eta_{j-k}) \right) (\beta_1 - \hat{\beta}_1 + \eta_j) \\ &= (\beta_1 - \hat{\beta}_1)^2 T^{-1} \sum_{j=2}^T \sum_{k=1}^{j-1} \frac{1}{k} + (\beta_1 - \hat{\beta}_1) T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \right) \eta_j \\ &\quad + (\beta_1 - \hat{\beta}_1) T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) + T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \eta_j \\ &= B_1 + B_2 + B_3 + T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \eta_j, \end{aligned}$$

where we use the expansion $-\log \Delta = L + \frac{1}{2}L^2 + \frac{1}{3}L^3 + \dots$. We show that $B_i = o_p(1)$ for $i = 1, 2, 3$.

$$\begin{aligned} B_1 &\leq (\beta_1 - \hat{\beta}_1)^2 T^{-1} T (\log T + \gamma + \zeta_T) = O_p(T^{-1}) (\log T + \gamma + \zeta_T) = o_p(1), \\ B_2 &= (\beta_1 - \hat{\beta}_1) T^{-1} \sum_{k=1}^{T-1} \left(\frac{1}{k} \sum_{j=k+1}^T \eta_j \right) = O_p(T^{-1/2}) o_p(1) = o_p(1), \\ B_3 &= (\beta_1 - \hat{\beta}_1) T^{-1} \sum_{k=1}^{T-1} \left(\frac{1}{k} \sum_{j=1}^{T-k} \eta_j \right) = O_p(T^{-1/2}) o_p(1) = o_p(1), \end{aligned}$$

where γ is the Euler-Mascheroni constant and $\zeta_T \sim 1/(2T)$ which approaches 0 as $T \rightarrow \infty$. The results for B_2 and B_3 follow from the arguments used in the proof of Theorem 1 for the terms A_2 and A_3 . Hence,

$$T^{-1} \sum_{j=2}^T \left(-\log \Delta \widetilde{\Delta y_j} \right) \widetilde{\Delta y_j} = T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \eta_j + o_p(1).$$

Then, under H_0 ,

$$\begin{aligned} \mathcal{LM}_T &= \frac{T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \eta_j + o_p(1)}{T^{-1} \sum_{t=1}^T \eta_t^2 + o_p(1)} \\ &= \frac{T^{-1} \sum_{j=2}^T \left(\sum_{k=1}^{j-1} \frac{1}{k} \eta_{j-k} \right) \eta_j}{T^{-1} \sum_{t=1}^T \eta_t^2} + o_p(1). \end{aligned}$$

This completes the proof. \square

Proof of Propositions A1 and A3. When the estimate of the break date is consistent, the proof is trivial and omitted. We here focus on having a consistent estimate of the break fraction at some rate T^κ for $0 < \kappa \leq 1$. Specifically, suppose that the estimate of the break fraction $\hat{\lambda}$ satisfies that $\hat{\lambda} - \lambda_b = o_p(T^{-\kappa})$ for $0 < \kappa \leq 1$. For all models A1–A3, we have a detrended sequence $\{\tilde{y}_t\}$ based on the OLS method in PZ. The Lagrange Multiplier test statistic is given by:

$$\mathcal{LM}_{T,\hat{\lambda}} = \sqrt{T} \sqrt{\frac{6}{\pi^2} \frac{\sum_{t=2}^T (-\log \Delta \tilde{\Delta y}_t) \Delta \tilde{y}_t}{\sum_{t=1}^T (\Delta \tilde{y}_t)^2}} = \sqrt{T} \sqrt{\frac{6}{\pi^2} \frac{\sum_{t=2}^T \left(\sum_{k=1}^{t-1} \frac{1}{k} \Delta \tilde{y}_{t-k} \right) \Delta \tilde{y}_t}{\sum_{t=1}^T (\Delta \tilde{y}_t)^2}}.$$

We show that the stochastic orders of terms associated with the deterministic time trend are smaller than those of terms associated with the error process. We can write (9) as:

$$\tilde{Y} = \hat{M}_z Z(T_b) \phi + \hat{M}_z U = \hat{M}_z Z(T_b)_2 \phi_2 + \hat{M}_z U = \tilde{M} + \tilde{U},$$

and $\Delta \tilde{Y} = \Delta \tilde{M} + \Delta \tilde{U}$. Since $\mathcal{LM}_{T,\hat{\lambda}}$ is a functional of $\Delta \tilde{Y}$ and subvectors of $\Delta \tilde{Y}$, it suffices to consider the inner product of $\Delta \tilde{Y}' \Delta \tilde{Y}$, that is,

$$\Delta \tilde{Y}' \Delta \tilde{Y} = \Delta \tilde{M}' \Delta \tilde{M} + 2 \Delta \tilde{M}' \Delta \tilde{U} + \Delta \tilde{U}' \Delta \tilde{U}.$$

Note that we only need to check the stochastic order of $\Delta \tilde{Y}' \Delta \tilde{Y}$ because the lag of order k is controlled to be small relative to the sample size T . We want to show that the term $\Delta \tilde{U}' \Delta \tilde{U}$ dominates the others. It is straightforward to show that $\Delta \tilde{U}' \Delta \tilde{U} = O_p(T)$ uniformly over all admissible break dates $T_b \in \{\pi T, (1 - \pi)T\}$ for some $\pi \in (0, 1/2)$ in all models. When $\Delta \tilde{M}' \Delta \tilde{M}$ has a smaller order of magnitude compared to that of $\Delta \tilde{U}' \Delta \tilde{U}$, so does $\Delta \tilde{M} \Delta \tilde{U}$ by the Cauchy-Schwartz inequality. Further, the order of magnitude of $\Delta \tilde{M}' \Delta \tilde{M}$ cannot be greater than that of $\tilde{M}' \Delta \tilde{M}$. The order of magnitude of $\tilde{M}' \Delta \tilde{M}$ is $O_p(T^{2-2\kappa})$ for Models A2 and A3, which implies that the break fraction should be estimated consistently at some rate greater than $T^{1/2}$. On the other hand, for Model A1b, the

stochastic order of $\tilde{M}'\Delta\tilde{M}$ is $O_p(T^{1+2\alpha})$. Hence, the orders of magnitude of terms associated with the deterministic trend are greater than those of the terms associated with the error process, thereby the LM statistic diverges as $T \rightarrow \infty$. \square

References

1. Perron, P. The great crash, the oil price shock and the unit root hypothesis. *Econometrica* **1989**, *57*, 1361–1401.
2. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431.
3. Perron, P. Testing for a unit root in a time series with a changing mean. *J. Bus. Econ. Stat.* **1990**, *8*, 153–162.
4. Christiano, L.J. Searching for a break in GNP. *J. Bus. Econ. Stat.* **1992**, *10*, 237–250.
5. Zivot, E.; Andrews, D.W.K. Further evidence on the great crash, the oil price shock and the unit root hypothesis. *J. Bus. Econ. Stat.* **1992**, *10*, 251–270.
6. Perron, P. Further evidence on breaking trend functions in macroeconomic variables. *J. Econom.* **1997**, *80*, 355–385.
7. Vogelsang, T.J.; Perron, P. Additional tests for a unit root allowing for a break in the trend function at an unknown time. *Int. Econ. Rev.* **1998**, *39*, 1073–1100.
8. Perron, P.; Zhu, X. Structural breaks with deterministic and stochastic trends. *J. Econom.* **2005**, *129*, 65–119.
9. Perron, P.; Yabu, T. Testing for shifts in trend with an integrated or stationary noise component. *J. Bus. Econ. Stat.* **2009**, *27*, 369–396.
10. Kim, D.; Perron, P. Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. *J. Econom.* **2009**, *148*, 1–13.
11. Carrion-i-Silvestre, J.L.; Kim, D.; Perron, P. GLS-based unit root tests with multiple structural breaks under both the null and the alternative hypotheses. *Econom. Theory* **2009**, *25*, 1754–1792.
12. Robinson, P.M. Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *J. Econom.* **1991**, *47*, 67–84.
13. Robinson, P.M. Efficient tests of nonstationary hypothesis. *J. Am. Stat. Assoc.* **1994**, *89*, 1420–1437.
14. Tanaka, K. The nonstationary fractional unit root. *Econom. Theory* **1999**, *15*, 549–582.
15. Dolado, J.J.; Gonzalo, J.; Mayoral, L. A fractional Dickey-Fuller test for unit roots. *Econometrica* **2002**, *70*, 1963–2006.
16. Lobato, I.N.; Velasco, C. Efficient Wald tests for fractional unit roots. *Econometrica* **2007**, *75*, 575–589.
17. Cho, C-K.; Amsler, C.; Schmidt, P. A test of the null of integer integration against the alternative of fractional integration. *J. Econom.* **2015**, *187*, 217–237.
18. Kwiatkowski, D.; Phillips, P.C.B.; Schmidt, P.; Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root. *J. Econom.* **1992**, *54*, 154–178.
19. Dolado, J.J.; Gonzalo, J.; Mayoral, L. Wald tests of I(1) against I(d) alternatives: Some new properties and an extension to processes with trending components. *Stud. Nonlinear Dyn. E.* **2008**, *12*, 1–33.
20. Diebold, F.X.; Inoue, A. Long memory and regime switching. *J. Econom.* **2001**, *105*, 131–159.
21. Granger, C.W.J.; Hyung, N. Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *J. Empir. Finance* **2004**, *11*, 399–421.
22. Lu, Y.K.; Perron, P. Modeling and forecasting stock return volatility using a random level shift model. *J. Empir. Finance* **2010**, *17*, 138–156.
23. Perron, P.; Qu, Z. Long-memory and level shifts in the volatility of stock market return indices. *J. Bus. Econ. Stat.* **2010**, *28*, 275–290.
24. Qu, Z.; Perron, P. A stochastic volatility model with random level shifts and its application to S&P 500 and NASDAQ return indices. *Econom. J.* **2013**, *16*, 309–339.
25. Xu, J.; Perron, P. Forecasting return volatility: Level shifts with varying jump probability and mean reversion. *Int. J. Forecast.* **2014**, *30*, 449–463.
26. Varneskov, R.T.; Perron, P. *Combining Long Memory and Level Shifts in Modeling and Forecasting the Volatility of Asset Returns*; Department of Economics, Boston University, Boston, MA, USA; Unpublished manuscript, 2016.
27. Wright, J.H. Testing for a structural break at unknown date with long-memory disturbances. *J. Time Ser. Anal.* **1998**, *19*, 369–376.

28. Krämer, W.; Sibbertsen, P. Testing for structural changes in the presence of long memory. *Int. J. Bus. Econ.* **2002**, *1*, 235–242.
29. Robinson, P.M. The distance between rival nonstationary fractional processes. *J. Econom.* **2005**, *128*, 283–300.
30. Marinucci, D.; Robinson, P.M. Alternative forms of fractional Brownian motion. *J. Stat. Plan. Inference* **1999**, *80*, 111–122.
31. Davidson, J.; Hashimzade, N. Type I and type II Brownian motions: A reconsideration. *Comput. Stat. Data Anal.* **2009**, *53*, 2089–2106.
32. Hosking, J.R.M. Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *J. Econom.* **1996**, *73*, 261–284.
33. Chang, S.Y.; Perron, P. Inference on a structural change in trend with fractionally integrated errors. *J. Time Ser. Anal.* **2016**, *37*, 555–574.
34. Hatanaka, M.; Yamada, K. A unit root test in the presence of structural changes in I(1) and I(0) models. In *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger*; Engle, R.F., White, H., Eds.; Oxford University Press: Oxford, UK, 1999; pp. 256–282.
35. Harvey, D.I.; Leybourne, S.J.; Newbold, P. Innovational outlier unit root tests with an endogenously determined break in level. *Oxf. B. Econ. Stat.* **2001**, *63*, 559–575.
36. Nunes, L.C.; Kuan, C-M.; Newbold, P. Spurious break. *Econom. Theory* **1995**, *11*, 736–749.
37. Iacone, F.; Leybourne, S.J.; Taylor, A.M.R. Testing for a break in trend when the order of integration is unknown. *J. Econom.* **2013**, *176*, 30–45.
38. Abadir, K.M.; Distaso, W.; Giratis, L. Nonstationarity-extended local Whittle estimation. *J. Econom.* **2007**, *141*, 1353–1384.
39. Andrews, D.W.K. Tests for parameter instability and structural change with unknown change point. *Econometrica* **1993**, *61*, 821–856.
40. Shimotsu, K. Exact local Whittle estimation of fractional integration with unknown mean and time trend. *Econom. Theory* **2010**, *26*, 501–540.
41. Gil-Alana, L.A. Fractional integration and structural breaks at unknown periods of time. *J. Time Ser. Anal.* **2008**, *29*, 163–185.
42. Hassler, U.; Wolters, J. Long memory in inflation rates: International evidence. *J. Bus. Econ. Stat.* **1995**, *13*, 37–45.



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).