ACADEMIA ECONOMIC PAPERS
32 : 2 (June 2004), 273–312

Fractional Integration and the Phillips-Perron Test

Chingnun Lee *
Graduate Institute of Economics
National Sun Yat-sen University

Fu-Shuen Shie
Ph. D. Student
Department of Finance
National Taiwan University

Keywords: Unit root, Fractional integrated process, Power
JEL classification: C12, C22

* Correspondence: Chingnun Lee, Graduate Institute of Economics, National Sun Yat-sen University, Kaohsiung 804, Taiwan. Tel: (07) 525-2000 ext. 5618; Fax: (07) 525-5611; E-mail: lee_econ @mail.nsysu.edu.tw. We are grateful to the Professor Ching-Fan Chung, Professor Wen-Jen Tsay, and two anonymous referees for their helpful and constructive comments on an earlier version of this paper. We are especially grateful to the associate editor for the suggestion implemented in the proofs of Lemma 2 in this paper.

ABSTRACT

This paper derives the asymptotic distribution of the Phillips-Perron unit root tests statistics and some of their variants under a general non-stationary fractionally-integrated $\mathsf{I}(1+d)$ *process, for* $d \in (-0.5, 0.5)$*. By using the Newey-West estimator of long-run variance, we show that both the Phillips-Perron's* t *statistics and standardized coefficients estimator are consistent against a non-stationary but mean-reverting alternative, such as the* $\mathsf{I}(1+d)$ *process for* $d \in (-0.5, 0)$*. However, only the* t *statistic from a no-drift and no-time trend regression is consistent against a non-stationary and non-mean-reverting alternative, such as the* $\mathsf{I}(1+d)$ *process for* $d \in (0, 0.5)$*. Simulation results also confirm that the power of these test statistics in large samples will decrease as the lag number increases in the construction of a Newey-West estimator of the long-run variance.*

1. INTRODUCTION

It is very important that a statistical test be able to fully discriminate the null and the alternative hypothesis in large samples. This paper is concerned with this issue in testing the hypothesis of a unit root in the following simple autoregressive model:

$$y_t = \beta y_{t-1} + u_t, \tag{1}$$

where the disturbance u_t is a stationary but fractionally-integrated I(d) process with $-0.5 < d < 0.5$.

Sowell (1990) analyzes the asymptotic properties of the standard Dickey-Fuller (DF, Dickey and Fuller (1979)) unit root test statistics from the ordinary least square estimator (OLS) of β for the case in (1), where $\beta = 1$, $y_0 = 0$, and $(1-L)^d u_t = \varepsilon_t$, ε_t is independently and identically-distributed with mean zero and variance σ^2 ($i.i.d.$ $(0, \sigma^2)$) and $E(|\varepsilon_t|^m) < \infty$ for $m \geq \max[4, -8d/(1+2d)]$. Diebold and Rudebusch (1991) show via a Monte-Carlo simulation that these tests, although consistent, have little power in finite samples.

The results of Sowell (1990) are of limited value for real world applications, where one almost always has to allow under the null hypothesis, $d = 0$, for autocorrelation among u_t in (1). Therefore, the standard Dickey-Fuller test is rarely appropriate, and there is automatically some implied interest in the power of the Augmented Dickey-Fuller test (ADF, Dickey and Fuller (1979)) or the Phillips-Perron test (PP, Phillips (1987), Phillips and Perron (1988), and Perron and Ng (1996)), which allow for a more general process in u_t than as in Sowell (1990).

Under Monte Carlo results, Hassler and Wolters (1994) conclude that the ADF test is not consistent against a fractional alternative. They also say that the power of the PP (t-statistics) test is not very much influenced by the choice of the number of included residual autocovariances. However, the first conclusion of Hassler and Wolters (1994) is shown to be misleading by Krämer (1998), who shows that the ADF test is indeed consistent against a fractional alternative if the order of the autoregression does not tend to go to infinite too fast. This paper is concerned with the theoretical background underpinning the power of PP test statistics.

The objectives of our paper are twofold. The first extends the weak convergence

results in characterizing the fractional unit root distribution given by Sowell (1990) to general fractionally-integrated processes. Instead of assuming that the innovations u_t in model (1) are a fractional-$i.i.d.$ process, we allow for it after differencing d times to be a weakly dependent process as that in Phillips (1987) where the functional central limit theorem of Davidson and de Jong (2000) can apply.

Our second objective is to derive the asymptotic distribution of the PP test statistics under a non-stationary fractionally-integrated I(1 + d) process, $d \in (-0.5, 0.5)$. Our results show that both the Phillips-Perron's t statistics Z_t and standardized coefficients' estimator Z_β are consistent against a non-stationary but mean-reverting [1] alternative, such as the I($1+d$) process for $d \in (-0.5, 0)$. However, only the Z_t statistic from a no-drift and no-time trend regression is consistent against a non-stationary and non-mean-reverting alternative, such as the I(1 + d) process for $d \in (0, 0.5)$. The statistics Z_t from a regression with a drift, and with a drift and a time trend, and the Z_β test statistics are unable to distinguish consistently between the unit root and a non-stationary and non-mean-reverting fractionally-integrated I(1 + d) process, for $d \in (0, 0.5)$. We also show that in the cases when the test statistics are consistent, the power of these test statistics is in fact affected by the choice of the number included in the residual autocovariance in the construction of the Newey and West (1987) estimator of the long-run variance. Therefore, the results of Hassler and Wolters (1994) are again misleading in their conclusion on the power of the PP test against a fractionally-integrated process. According to our simulation, we find that the results of Hassler and Wolters (1994) may be due to the problem of a small sample size in their simulation.

This paper is organized as follows. Replacing a stronger condition with Davidson and De Jong's (2000), a functional central limit theory for a general fractionally-integrated process is given in section 2, where the generalized fractional unit root distribution is also presented. Section 3 analyzes the asymptotic properties of the Phillips-Perron test statistics under this general non-stationary fractionally-integrated process. Section 4 provides the simulation evidence of the power of the Phillips-Perron test. A conclusion completes the analysis. The main proofs are contained in the Appendix.

Throughout this paper, we use "$\Rightarrow$" to denote weak convergence of associated probability measures, "$\xrightarrow{p}$" denotes convergence in probability, [z] means the largest integer that is smaller than or equal to z, and we let $e_i \sim f_i$ denote that $e_i/f_i \to 1$ as $i \to \infty$.

[1] For a mean-reverting process, the impact of a unit innovation ε_t at time t on the process u_{t+k} is zero as $k \to \infty$.

2. ASYMPTOTICS OF A GENERAL FRACTIONALLY-INTEGRATED PROCESS

2.1 Functional central limit theory for a fractionally-integrated process

A general fractionally-integrated I(d) process u_t is customarily written in the form

$$(1-L)^d u_t = \varepsilon_t, \tag{2}$$

where the innovation process ε_t, the fractional difference of u_t, is assumed to be a stationary and weakly dependent process as specified below. By the obvious binomial expansion, they have the moving average (MA) representation:

$$u_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \tag{3}$$

where

$$b_j = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}.$$

Stirling's approximation formulation for the gamma function yields the well-known property that the MA coefficients decline hyperbolically to zero. We can now write

$$b_j \sim \frac{1}{\Gamma(d)} j^{d-1}.$$

The terms b_j are therefore square summable for $d < 0.5$, which is the condition necessary for the process to be stationary with a finite variance, whereas $d > -0.5$ is necessary for the process to have an invertible AR representation. Moreover, $d < 1$ implies that u_t is mean-reverting. The autocorrelations of u_t, ρ_j can be shown to be

$$\rho_j \sim \ cj^{2d-1},$$

where c is constant and has the same sign as d when ε_t is $i.i.d.$.

For $0<d<0.5$, the process is long memory in the sense that $\lim_{T\to\infty}\sum_{j=-T}^{T}|\rho_j|$ is non-finite. Its autocorrelations are all positive and decay at a hyperbolic rate. For $-0.5<d<0$, the sum of the absolute values of the processes autocorrelation tends to a constant, and all its autocorrelations, excluding the lag zero, are negative and decay hyperbolically to zero. Helson and Sarason (1967) show that an I(d) process with $d>-0.5$ and an autocorrelation given by ρ_j violate the strong mixing condition and hence is a long memory. Beran (1994) and Baillie (1996) are standard references to an I(d) process.

To begin we must be precise about the sequence ε_t of allowable innovations in (2) for the present paper. Following Phillips (1987), we assume that ε_t is a sequence of random variables that satisfy the following Assumption.

Assumption 1 The sequence ε_t, $-\infty<t<\infty$,
(a) has zero mean;
(b) satisfies $\sup_t E|\varepsilon_t|^\gamma<\infty$ for some $\gamma>2$;
(c) is stationary,[2] and $0<\sigma_\varepsilon^2<\infty$, where $\sigma_\varepsilon^2=\lim_{T\to\infty}T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}E(\varepsilon_t\varepsilon_s)$;
(d) is strong mixing with mixing coefficients α_m that satisfy $\sum_{m=1}^{\infty}\alpha_m^{1-2/\gamma}<\infty$.

Define the variance of the partial sums of the I(d) process u_t by $\sigma_T^2=E(\sum_{t=1}^{T}u_t)^2$. Davidson and De Jong (2000) have the following functional central limit theorem (FCLT) for this broad class I(d) process whose underlying shock variables may themselves exhibit quite general weak dependence.[3]

Lemma 1 Suppose $(1-L)^d u_t=\varepsilon_t$, $-0.5<d<0.5$ and ε_t satisfy Assumption 1; then as $T\to\infty$,
(a) $\sigma_T^2\sim\sigma_\varepsilon^2 V_d T^{1+2d}$, and
(b) $\sigma_T^{-1}\sum_{t=1}^{[Tr]}u_t\Rightarrow B_d(r)$, for $r\in[0,1]$.

Here, $B_d(r)$ is the normalized fractional Brownian motion that is defined by the fol-

[2] This is a stronger condition than Phillips (1987)'s. This condition excludes any possible heterogeneous innovation.

[3] This is a stronger assumption than as in Assumption 1 of Davidson and De Jong (2000). We impose the stationary condition in Assumption 1(c) that makes the u_t process be a linear function of a stationary process, ε_t. Because the stationary linear processes are ergodic, u_t under our assumption can apply the stationary ergodic Theorem (see Stout (1974), p.181)

lowing stochastic integral:[4]

$$B_d(r) \equiv \frac{1}{\Gamma(1+d)V_d^{\frac{1}{2}}}\left(\int_0^r (r-x)^d dB(x) + \int_{-\infty}^0 \left[(r-x)^d - (-x)^d\right] dB(x)\right), \quad (4)$$

with $V_d \equiv \Gamma(1+d)^{-2}\{(1+2d)^{-1} + \int_0^\infty [(1+\tau)^d - \tau^d] d\tau\} = \Gamma(1-2d)/[(1+2d)\Gamma(1+d)\Gamma(1-d)]$ and $B(r)$ is the standard Brownian motion.[5]

Lemma 1(a) shows that the variance of the partial sum of an I(d) process is $O(T^{1+2d})$. Sowell (1990) obtains the same result under the assumption that ε_t is $i.i.d.$. Lemma 1(b) is an FCLT for a general fractionally-integrated process that could apply to a large class of fractionally-integrated processes including the well-known Gaussian ARFIMA (p,d,q) process. These conditions allow for a wide variety of possibly weakly dependent generating mechanisms in the innovations process ε_t. The main novel feature of these results is that the innovation variables (fractional difference) are permitted to be a mixing process, a very general form of weak dependence allowing for various forms of non-linear dynamics.

Davydov (1970), Avram and Taqqu (1987), Mielniczuk (1997), and Chung (2002) give results on FCLT for a long memory process directly from an I(d) process written in a linear process as in (3) with square summable weights. Wang et al. (2003) derive the FCLT for a fractionally-integrated process from (2) by assuming that ε_t is a linear process of an $i.i.d.$ random variable, say η_t,

$$\varepsilon_t = \sum_{j=0}^{\infty} \varphi_j \eta_{t-j}, \; t = 1, 2, ... \quad (5)$$

with $\sum_{j=0}^{\infty} j^{0.5-d}|\varphi_j| < \infty$, $\sum_{j=0}^{\infty} \varphi_j \neq 0$, and $E|\eta_t|^{\max\{2,2/(1+2d)\}} < \infty$.

When our mixing processes ε_t are specified to be a summable linear process as in (5), it is interesting to compare the regularity conditions of Lemma 1 to those in Wang et al. First, the moment condition of ε_t in Assumption 1(b) is stronger than Wang et al's where only second moment is needed to exist. Second, the coefficient condition

[4] The original definition of a fractional Brownian motion shown in Sowell (1990) is $B_d(r) = 1/\Gamma(1+d)\int_0^r (r-x)^d dB(x)$. However, Marinucci and Robinson (1999) show that it requires a correction by replacing it with the definition of the fractional Brownian motion as in (4).

[5] This type of fractional Brownian motion is so defined as to make $EB_d(1)^2 = 1$. A fractional Brownian motion differs from a standard Brownian motion $B(r)$ by having correlated increments. Please refer to Marinucci and Robinson (1999) for additional details on the fractional Brownian motion.

on φ_j is weaker in our assumption due to mixing in a linear process, as we only need φ_j to be absolute summable (See Davidson (1994), p. 215)

2.2 Generalized fractional unit root distribution

From here on, we discuss the following non-stationary fractionally-integrated process y_t, defined by

$$y_t = \beta y_{t-1} + u_t, \tag{6}$$

$$\beta = 1, \tag{7}$$

where $y_0 = 0$, and u_t is the I(d) process satisfying the assumption in Lemma 1.

We consider the three least square regressions

$$y_t = \breve{\beta} y_{t-1} + \breve{u}_t, \tag{8}$$

$$y_t = \hat{\alpha} + \hat{\beta} y_{t-1} + \hat{u}_t, \tag{9}$$

and

$$y_t = \tilde{\alpha} + \tilde{\delta} t + \tilde{\beta} y_{t-1} + \tilde{u}_t, \tag{10}$$

where $\breve{\beta}, (\hat{\alpha}, \hat{\beta})$, and $(\tilde{\alpha}, \tilde{\delta}, \tilde{\beta})$ are the conventional least-squares regression coefficients. Following Sowell (1990), we are concerned with the limiting distribution of the regression in (8), (9), and (10) under the null hypothesis that the data are generated by (6) and (7). The limiting distribution of $\breve{\beta}, (\hat{\alpha}, \hat{\beta})$, and $(\tilde{\alpha}, \tilde{\delta}, \tilde{\beta})$ is described in the following Theorem.

Theorem 1 Let y_t satisfy (6) and (7); then as $T \to \infty$,
for the regression model (8),

(a) $T(\breve{\beta} - 1) \Rightarrow \dfrac{(1/2)[B_d(1)]^2}{\int_0^1 [B_d(r)]^2 dr}$, when $d > 0$;

(b) $T^{1+2d}(\breve{\beta} - 1) \Rightarrow -\dfrac{(1/2)\sigma_u^2}{\sigma_\varepsilon^2 V_d \int_0^1 [B_d(r)]^2 dr}$, when $d < 0$; and

(c) $T(\breve{\beta}-1) \Rightarrow \dfrac{(1/2)[B(1)^2-\sigma_u^2/\sigma_\varepsilon^2]}{\int_0^1 B(r)^2 dr}$, when $d=0$;[6]

for the regression model (9),

(d) $T(\hat{\beta}-1) \Rightarrow \dfrac{(1/2)[B_d(1)]^2 - B_d(1)\int_0^1 B_d(r)dr}{W_1}$, when $d>0$;

(e) $T^{1+2d}(\hat{\beta}-1) \Rightarrow -\dfrac{1}{2}\dfrac{\sigma_u^2}{V_d\sigma_\varepsilon^2 W_1}$, when $d<0$; and

(f) $T(\hat{\beta}-1) \Rightarrow \dfrac{(1/2)\{[B(1)]^2-(\sigma_u^2/\sigma_\varepsilon^2)\} - B(1)\int_0^1 B(r)dr}{\int_0^1[B(r)]^2dr - [\int_0^1 B(r)dr]^2}$, when $d=0$;

for the regression model (10),

(g) $T(\tilde{\beta}-1) \Rightarrow \dfrac{W_2}{W_3}$, when $d>0$;

(h) $T^{1+2d}(\tilde{\beta}-1) \Rightarrow -\dfrac{1}{2}\dfrac{\sigma_u^2}{V_d\sigma_\varepsilon^2 W_3}$, when $d<0$,

(i) $T(\tilde{\beta}-1)$

$$\Rightarrow \frac{(1/2)\{[B(1)-2\int_0^1 B(r)dr][B(1)+6\int_0^1 B(r)dr-12\int_0^1 rB(r)dr]-\sigma_u^2/\sigma_\varepsilon^2\}}{\int_0^1[B(r)]^2dr-4[\int_0^1 B(r)dr]^2+12\int_0^1 B(r)dr\int_0^1 rB(r)dr-12[\int_0^1 rB(r)dr]^2}, \quad \text{when } d=0;$$

where

$$\sigma_u^2 = E(u_t^2),$$

$$W_1 = \int_0^1[B_d(r)]^2dr - [\int_0^1 B_d(r)dr]^2,$$

$$W_2 = (1/2)[B_d(1)-2\int_0^1 B_d(r)dr][B_d(1)+6\int_0^1 B_d(r)dr - 12\int_0^1 rB_d(r)dr]$$

and

$$W_3 = \int_0^1[B_d(r)]^2dr - 4[\int_0^1 B_d(r)dr]^2 + 12\int_0^1 B_d(r)dr\int_0^1 rB_d(r)dr - 12[\int_0^1 rB_d(r)dr]^2.$$

[6] The limit distribution of the regression coefficient $T(\breve{\beta}-1)$ under the unit root hypothesis that $d=0$ and $\beta=1$ depends on unknown nuisance parameters, σ_u^2 and σ_ε^2; thus, the statistics $T(\breve{\beta}-1)$ cannot be used directly for unit root testing. However, σ_u^2 and σ_ε^2 can be consistently estimated, and there exists a simple transformation of the statistics $T(\breve{\beta}-1)$ which eliminates the nuisance parameters asymptotically. See Phillips (1987) and Xiao and Phillips (1998).

The convergence rates of $(\breve{\beta}-1)$ (and also $(\hat{\beta}-1)$ and $(\tilde{\beta}-1)$) depend intrinsically on the degree of the fractional integration in the u_t process. The distribution of $T^{\min[1,1+2d]}(\breve{\beta}-1)$ is therefore called a generalized fractional unit root distribution. Fractional unit root distribution was first derived by Sowell (1990) where ε_t in (2) is assumed to be $i.i.d.$ $(0,\sigma^2)$. Tanaka (1999) and Wang et al. (2003) extend the results of Sowell (1990) parametrically by assuming that ε_t is an infinite-order moving average process.[7]

Theorem 1 is a non-parametric extension of Sowell (1990) by assuming that ε_t satisfies Assumption 1. Under quite weak conditions, these results provide a unified treatment of most of the previously-cited results. For example, when the innovation process ε_t is $i.i.d.$ $(0,\sigma^2)$, we have $\sigma_u^2 = [\Gamma(1-2d)/\Gamma^2(1-d)]\sigma^2$, leading to the following simplification of parts (b) and (c) of Theorem 1 as the results of Sowell (1990): $T^{1+2d}(\breve{\beta}-1) \Rightarrow -\{[1/2+d][\Gamma(1+d)/\Gamma(1-d)]\}/\{\int_0^1[B_d(r)]^2dr\}$, for $d<0$ and $T(\breve{\beta}-1) \Rightarrow \{(1/2)[B(1)^2-1]\}/\{\int_0^1 B(r)^2dr\}$, for $d=0$.

It is interesting to note that when $d>0$, the assumption on ε_t does not play any role in determining the order of magnitude of the test statistic. The limit distribution of $T(\breve{\beta}-1)$ is free of the nuisance parameters of ε_t in (2). It converges to the same distribution as that of Sowell (1990), Tanaka (1999), and Wang et al. (2003). When $d<0$, the distribution of $T^{1+2d}(\breve{\beta}-1)$ has the same general form for a very wide class of the innovation process ε_t. It reduces to the distribution of Phillips (1987, Theorem 3.1, (c)) when $d=0$.

3. ASYMPTOTICS OF PHILLIPS-PERRON TEST WHEN u_t IS I(d)

3.1 Phillips and Perron test

Phillips (1987) proposes a test for the unit root hypothesis of the process y_t defined in (6) and (7) when $d=0$. Denoting $\ddot{\sigma}_u^2$=$\lim_{T\to\infty}T^{-1}E(\sum_{t=1}^T u_t^2)$ and $\ddot{\sigma}^2$=$\lim_{T\to\infty}T^{-1}$ $E(\sum_{t=1}^T u_t)^2$, Phillips's two statistics for testing the null of $\beta=1$ can be expressed as follows:

[7] In particular, Tanaka (1999) assumes that $\varepsilon_t = \sum_{j=0}^{\infty}\varphi_j\eta_{t-j}$ as in (5), but with the condition that $\sum_{j=0}^{\infty} j|\varphi_j| < \infty$.

$$Z_{\breve{\beta}} = T(\breve{\beta} - 1) - \frac{1}{2}\frac{(s^2 - s_u^2)}{T^{-2}\sum_{t=1}^{T} y_{t-1}^2}, \tag{11}$$

and

$$Z_{\breve{t}} = \left(\sum_{t=1}^{T} y_{t-1}^2\right)^{\frac{1}{2}} \frac{\breve{\beta} - 1}{s} - \frac{1}{2}\left(s^2 - s_u^2\right)\left[s\left(T^{-2}\sum_{t=1}^{T} y_{t-1}^2\right)^{\frac{1}{2}}\right]^{-1}, \tag{12}$$

where $\breve{\beta}$ is the OLS estimator of β in (8), and s_u^2 and s^2 are consistent estimates of $\ddot{\sigma}_u^2$ and $\ddot{\sigma}^2$, respectively.

Under the null hypothesis of the unit root, Phillips (1987) proves that these test statistics are distributed as $\{(1/2)[B(1)^2-1]\}/\{\int_0^1 B(r)^2 dr\}$ and $\{(1/2)[B(1)^2-1]\}/\{\int_0^1 B(r)^2 dr\}^{1/2}$, respectively. Phillips and Perron (1988) accommodate the regression model (8) with a drift as in (9) and a drift and time trend as in (10) and propose test statistics $Z_{\hat{\beta}}$, $Z_{\tilde{\beta}}$, $Z_{\hat{t}}$, and $Z_{\tilde{t}}$ corresponding to $Z_{\breve{\beta}}$ and $Z_{\breve{t}}$, respectively.

Among statistics in this class of unit root tests, the ADF and the PP are perhaps the most popular, as they are implemented in many statistical software packages. However, it is also a well documented fact that the PP test, as originally defined, suffers from severe size distortions when there are negative moving-average errors.[8] Perron and Ng (1996) suggest a modification of the PP tests to correct this problem. They use methods suggested by Stock (1990) to derive a modification of the Z_β and Z_t statistics. The modified Z_β and Z_t statistics, for example, in regression model (8) are

$$MZ_{\breve{\beta}} = Z_{\breve{\beta}} + \frac{T}{2}(\breve{\beta} - 1)^2, \tag{13}$$

and

$$MZ_{\breve{t}} = MSB \cdot MZ_{\breve{\beta}}, \tag{14}$$

where $MSB = (T^{-2}\sum_{t=1}^{T} y_{t-1}^2/s^2)^{1/2}$. Convergence of $\breve{\beta}$ at rate T when true $\beta = 1$ and u_t is I(0) ensures that MZ_β and Z_β are asymptotically equivalent. It is also true that MZ_t and Z_t are asymptotically equivalent.

[8] See Phillips and Perron (1988) and Schwert (1989), among others.

Phillips and Perron (1988) and Perron and Ng (1996) analyze the asymptotic properties of PP test statistics under a near-integrated process. In this paper we are concerned with the alternative that y_t is a fractionally-integrated I($1+d$) process. Granger and Joyeux (1980), Geweke and Porter-Hudak (1983), and Baillie (1996) all conclude that some economic time series may posses unit roots, but with a fractionally-integrated error. The distinction of a series being I(1) or I($1+d$), for example, $d<0$, is also important when y_t is the OLS residual from a set of I(1) variables. The test of the order of the integration in the residual is called a residual-based test for fractional cointegration.[9]

3.2 Consistency of the Phillips-Perron test against I($1+d$) alternatives

In this section we derive the asymptotic distribution of Z_β, Z_t, MZ_β, and MZ_t test statistics under the alternative hypothesis of y_t being I($1+d$) as in (6) and (7) for $-0.5<d<0.5$. In deriving Theorem 2 below, we consider s_u^2 ($\breve{s}_u^2 = T^{-1}\sum_{t=1}^T \breve{u}_t^2$, $\hat{s}_u^2 = T^{-1}\sum_{t=1}^T \hat{u}_t^2$, and $\tilde{s}_u^2 = T^{-1}\sum_{t=1}^T \tilde{u}_t^2$) as the estimator of $\ddot{\sigma}_u^2$ for the regression models in (8), (9), and (10), respectively.[10] We use the Newey-West (1987) estimator s_{Tl}^2 ($\breve{s}_{Tl}^2 = T^{-1}\sum_{t=1}^T \breve{u}_t^2 + 2T^{-1}\sum_{\tau=1}^l w_{\tau l}\sum_{t=\tau+1}^T \breve{u}_t\breve{u}_{t-\tau}$, $\hat{s}_{Tl}^2 = T^{-1}\sum_{t=1}^T \hat{u}_t^2 + 2T^{-1}\sum_{\tau=1}^l w_{\tau l}\sum_{t=\tau+1}^T \hat{u}_t\hat{u}_{t-\tau}$, and $\tilde{s}_{Tl}^2 = T^{-1}\sum_{t=1}^T \tilde{u}_t^2 + 2T^{-1}\sum_{\tau=1}^l w_{\tau l}\sum_{t=\tau+1}^T \tilde{u}_t$ $\tilde{u}_{t-\tau}$)[11] as an estimator of the long-run variance $\ddot{\sigma}^2$, where $w_{\tau l} = 1-\tau/(l+1)$.

Phillips (1987) shows the consistency of s_u^2 and s_{Tl}^2 when u_t is a I(0) and strong mixing process. The following results establish the asymptotics of s_u^2 and s_{Tl}^2 when u_t is an I(d) process as in (2).

Lemma 2 Let y_t satisfy (6) and (7); but part (b) of Assumption 1 is replaced by the

[9] See, for example, Dittmann (2000).

[10] Phillips (1987) also recommends using the first difference of y_t, $u_t = y_t - y_{t-1}$, rather than the residual $\breve{u}$, $\hat{u}$, and $\tilde{u}_t$. However, it is noted by Phillips and Ouliaris (1990) that it would make the PP statistics inconsistent against the $I(0)$ process by using the first difference of y_t in the construction of the nuisance parameters estimator. Since we allow for a non-zero drift in the regression such as (10), the use of the first difference of y_t, is not preferred. See Phillips and Perron (1988).

[11] Instead of using the Newly-West estimator s_{Tl}^2 as in this paper, Perron and Ng (1996) suggest alternatively an autoregressive spectral density estimator defined as (for regression model (8)) $s_{AR}^2 = s_{el}^2/[(1-\breve{b}(1))^2]$, where $s_{el}^2 = T^{-1}\sum_{t=l+1}^T \breve{e}_{tl}^2$, $\breve{b}(1) = \sum_{j=1}^l \breve{b}_j$, with $\breve{b}_j$ and $\{\breve{e}_{tl}\}$ obtained from the autoregression: $\triangle y_t = \breve{b}_0 y_{t-1} + \sum_{j=1}^l \breve{b}_j \triangle y_{t-j} + \breve{e}_{tl}$. We find from simulation results that when using s_{AR}^2 as a long-run variance estimator, the power of MZ will decrease with an increase of l. In particular, the MZ turns out to perform very poorly as like any inconsistent test if the number of lagged differences l included in the regression increase too much. Overall, for a fixed T the power of MZ using s_{Tl}^2 clearly dominates the same tests statistics using s_{AR}^2 under an I($1+d$) alternative.

stronger moment condition: $\sup_t E|\varepsilon_t|^4 < \infty$; then providing that $l \to \infty$ as $T \to \infty$ such that $l = o(T^{1/4})$,

(a) $s_u^2 \xrightarrow{p} \sigma_u^2$, and

(b) $l^{-2d} s_{Tl}^2 \xrightarrow{p} \sigma_\varepsilon^2 V_d$.

Lemma 2(a) is the results of consistency in the OLS estimator and ergodic theorem. Lemma 2(b) shows that the exact order of magnitude of s_{Tl}^2 is equal to $O(l^{2d})$. That is, it has to multiply l^{-2d} on s_{Tl}^2 to achieve a "consistent" estimator of the long-run variance of u_t, as corresponding to the results of Lemma 1(a). We next give the asymptotics of the PP statistics under the alternative of y_t being I($1+d$) in the following Theorem.

Theorem 2 If the conditions of Lemma 2 are satisfied, then as $T \to \infty$, for the regression model (8):

(a) when $d < 0$, $\left(\frac{T}{l}\right)^{2d} Z_{\check{\beta}} \Rightarrow -\frac{1}{2}\frac{1}{\int_0^1 [B_d(r)]^2 dr}$, and

(b) when $d < 0$, $\left(\frac{T}{l}\right)^{d} Z_{\check{t}} \Rightarrow -\frac{1}{2}\frac{1}{\{\int_0^1 [B_d(r)]^2 dr\}^{1/2}}$;

(a') when $d > 0$, $Z_{\check{\beta}} \Rightarrow \frac{1}{2}\frac{[B_d(1)]^2}{\int_0^1 [B_d(r)]^2 dr}$, and

(b') when $d > 0$, $\left(\frac{T}{l}\right)^{-d} Z_{\check{t}} \Rightarrow \frac{1}{2}\frac{[B_d(1)]^2}{\{\int_0^1 [B_d(r)]^2 dr\}^{1/2}}$;

for the regression model (9):

(c) when $d < 0$, $\left(\frac{T}{l}\right)^{2d} Z_{\hat{\beta}} \Rightarrow -\frac{1}{2}\frac{1}{W_1}$, and

(d) when $d < 0$, $\left(\frac{T}{l}\right)^{d} Z_{\hat{t}} \Rightarrow -\frac{1}{2}\frac{1}{W_1^{1/2}}$;

(c') when $d > 0$, $Z_{\hat{\beta}} \Rightarrow \frac{(1/2)[B_d(1)]^2 - B_d(1)\int_0^1 B_d(r)dr}{W_1}$, and

(d') when $d > 0$, $\left(\frac{T}{l}\right)^{-d} Z_{\hat{t}} \Rightarrow \frac{(1/2)[B_d(1)]^2 - B_d(1)\int_0^1 B_d(r)dr}{W_1^{1/2}}$;

for the regression model (10):

(e) when $d < 0$, $\left(\frac{T}{l}\right)^{2d} Z_{\tilde{\beta}} \Rightarrow -\frac{1}{2}\frac{1}{W_3}$, and

(f) when $d < 0$, $\left(\frac{T}{l}\right)^{d} Z_{\tilde{t}} \Rightarrow -\frac{1}{2}\frac{1}{W_3^{1/2}}$;

(e') when $d > 0$, $Z_{\tilde{\beta}} \Rightarrow \frac{W_2}{W_3}$, and

(f') when $d > 0$, $\left(\frac{T}{l}\right)^{-d} Z_{\tilde{t}} \Rightarrow \frac{W_2}{W_3^{1/2}}$;

where

$$W_1 = \int_0^1 [B_d(r)]^2 dr - [\int_0^1 B_d(r)dr]^2,$$

$$W_2 = (1/2)[B_d(1) - 2\int_0^1 B_d(r)dr][B_d(1) + 6\int_0^1 B_d(r)dr - 12\int_0^1 rB_d(r)dr]$$

and

$$W_3 = \int_0^1 [B_d(r)]^2 dr - 4[\int_0^1 B_d(r)dr]^2 + 12\int_0^1 B_d(r)dr \int_0^1 rB_d(r)dr$$
$$-12[\int_0^1 rB_d(r)dr]^2.$$

When $d < 0$, since the limit of $(T/l)^{2d} Z_\beta$ and $(T/l)^d Z_t$ has a negative support, then Z_β and Z_t are unbounded and diverge to $-\infty$ as $(T/l) \to \infty$. Therefore, PP's lower tail Z_β and Z_t test statistics are consistent against a non-stationary but mean-reverting fractionally-integrated alternative.

When $d > 0$, the Z_β statistics have the same orders in probability under both the null of unit root and alternative of I($1 + d$), $d > 0$. In fact, for $d \in (0, 0.5)$, the orders in probability of the Z_β statistics are independent of the value of d, even though the form of their asymptotic distribution is affected by the value of d. This is in contrast to the cases when Z_t statistics for $d \in (-0.5, 0.5)$ and Z_β for $d \in (-0.5, 0)$, where both the order in probability and the form of the asymptotic distribution depend on d. Furthermore, although the Z_t test statistics have a different order in probability with the

null hypothesis of the unit root, only the limit of $(T/l)^{-d}Z_{\tilde{t}}$ has a positive support. The supports of the limit of $(T/l)^{-d}Z_{\hat{t}}$ and $(T/l)^{-d}Z_{\tilde{t}}$ are both possibly positive and negative. Therefore, only PP's upper $Z_{\tilde{t}}$ test statistic is consistent against a non-stationary and non mean-reverting fractionally-integrated alternative.

As a consequence of Theorems 1 and 2, we have the following corollary that shows the asymptotic equivalence of MZ_β and Z_β, MZ_t and Z_t.

Corollary 1 If the conditions of Lemma 2 are satisfied, then as $T \to \infty$, $MZ_\beta \Rightarrow Z_\beta$ and $MZ_t \Rightarrow Z_t$.

4. POWER IN FINITE SAMPLES

In this section we provide simulation evidence on the power of the MZ_β and MZ_t tests statistics against the I(1 + d) alternative. Following Schwert (1987), the number of lags l used in the MZ test statistics is chosen as $l_0 = 0$, $l_4 = int[4(T/100)^{1/4}]$, and $l_{12} = int[12(T/100)^{1/4}]$. We consider samples sizes T=100, 250, 500, and 1000 and the number of iterations is 5000. All of our tests are based on the 5% significance level. Observations on a fractional white noise I(d) process are generated using the Durbin-Levinson algorithm.

Tables 1, 3, and 5 give the percentage power of the 5% lower tail MZ_β and MZ_t unit root test against the alternatives I(1+d), $d = -0.1, -0.2, \ldots, -0.49$. Some results are quite clear and in accordance with our expectations. With other things being held constant: (i) Power increases as T increases. This is a reflection of the consistency of the test. (ii) When $T = 1000$, the power of both the MZ_β and MZ_t tests is lower when l is higher, which is in accordance with the asymptotics of the MZ_β and MZ_t test statistics under the I(1 + d), $d < 0$, alternative in the last section, which indicate that the power depends on (T/l) even asymptotically. However, when T is less than 1000, for example $T = 250$ is the sample size used by Hassler and Wolters (1994), then the conclusion that power will decrease as with an increase in l is ambiguous. Therefore, the results of Hassler and Wolters (1994) that the power of the PP test is not very much influenced by the choice of the number of included residual autocovariances may be due to small sample size results. (iii) Power is higher when d is larger in absolute value for a fixed T. This result is not surprising, as it is transparent from the relevant asymptotics for $-0.5 < d < 0$.

Tables 2, 4 and 6 provide the percentage power of the 5% upper tails of the MZ_β

Table 1 Percentage Power of $MZ_{\check{\beta}}$ and $MZ_{\check{t}}$ Tests Against I($1+d$) Alternatives, $d < 0$

			Value of d						
Test	T	l	0.0	-0.1	-0.2	-0.3	-0.4	-0.45	-0.49
$MZ_{\check{\beta}}$	100	0	5.10	14.36	34.80	59.50	78.44	85.36	90.84
		1	5.40	12.50	25.76	47.44	67.30	74.76	82.06
		4	4.82	12.64	24.26	43.94	63.38	70.32	77.20
		12	5.92	12.94	27.96	49.48	69.64	77.44	85.04
	250	0	5.16	20.36	49.92	78.90	94.82	97.66	99.38
		1	4.80	18.58	41.80	70.06	88.40	94.22	96.58
		5	5.16	15.48	33.42	58.20	79.64	86.92	92.58
		15	5.88	14.38	35.78	62.28	84.26	91.10	95.00
	500	0	4.98	24.08	60.58	88.96	98.98	99.90	99.94
		1	4.84	22.58	53.02	82.02	96.92	98.88	99.76
		5	4.68	17.02	42.68	70.64	90.50	95.38	98.18
		17	5.76	15.94	40.26	69.92	90.94	96.02	98.66
	1000	0	5.02	30.24	71.26	95.12	99.94	100.00	100.00
		1	4.88	26.58	64.88	92.42	99.52	99.90	100.00
		7	4.52	20.06	49.32	80.22	96.36	98.72	99.79
		21	4.78	17.46	45.04	77.62	95.90	98.50	99.76
$MZ_{\check{t}}$	100	0	5.08	14.58	34.88	59.14	78.36	85.32	90.84
		1	5.08	12.02	25.22	46.44	66.60	74.02	81.54
		4	4.70	12.06	23.68	43.02	62.98	69.84	76.80
		12	5.60	12.34	27.10	48.58	69.28	77.14	84.76
	250	0	5.30	20.06	49.38	78.32	94.68	97.72	99.30
		1	4.84	17.84	41.24	69.06	87.92	94.10	96.36
		5	5.00	14.84	32.76	57.60	79.62	86.72	92.50
		15	5.50	14.08	35.26	61.80	84.06	90.86	94.80
	500	0	4.84	23.70	59.28	88.28	98.92	99.88	99.92
		1	4.86	21.76	51.64	81.08	96.76	98.84	99.78
		5	4.48	16.40	41.62	69.74	90.02	95.18	98.00
		17	5.68	15.58	39.60	69.34	90.54	95.92	98.64
	1000	0	4.74	29.70	70.70	95.00	99.92	100.00	100.00
		1	4.88	26.10	63.96	91.88	99.50	99.90	100.00
		7	4.38	19.52	48.84	79.76	96.28	98.70	99.78
		21	4.68	17.38	44.38	77.36	95.84	98.42	99.74

Note: 1. $MZ_{\check{\beta}}$ and $MZ_{\check{t}}$ are Perron and Ng's (1996) modified Phillips and Perron (1988) unit root test statistics.

2. DGP is $(1-L)^{1+d}y_t = \varepsilon_t, \varepsilon_t \overset{i.i.d.}{\sim} N(0,1)$.

3. The number of lag l other than 1 is chosen as $l_0 = 0$, $l_4 = int[4(T/100)^{1/4}]$ and $l_{12} = int[12(T/100)^{1/4}]$.

4. Number of iterations is 5000. Power of the tests is based on the 5% lower-tailed significance level.

Table 2 Percentage Power of $MZ_{\check{\beta}}$ and $MZ_{\check{t}}$ Tests Against I($1+d$) Alternatives, $d > 0$

			Value of d						
Test	T	l	0.0	0.1	0.2	0.3	0.4	0.45	0.49
$MZ_{\check{\beta}}$	100	0	5.44	12.76	24.74	35.22	47.72	54.80	73.20
		1	4.92	11.54	21.94	33.04	44.72	54.10	73.02
		4	5.14	11.60	19.64	28.88	41.84	52.36	69.58
		12	5.66	11.78	17.74	26.54	38.88	46.76	65.86
	250	0	5.46	16.08	27.68	39.32	51.00	59.82	76.98
		1	5.22	14.10	27.08	38.84	50.06	59.42	76.86
		5	5.60	12.66	22.16	33.50	47.16	57.90	74.98
		15	6.04	10.90	20.20	29.98	42.50	52.54	72.20
	500	0	5.20	16.34	31.54	40.60	51.94	60.50	77.18
		1	4.98	16.14	28.48	39.00	52.16	60.84	77.60
		5	5.28	13.72	26.10	36.58	49.14	57.42	76.76
		17	5.88	13.52	21.84	32.58	44.50	55.38	74.42
	1000	0	4.82	16.84	31.98	41.84	52.64	61.40	78.04
		1	4.84	17.52	30.38	40.52	52.28	59.98	78.90
		7	4.76	15.36	26.52	38.36	49.94	59.52	77.12
		21	5.38	13.34	23.46	33.90	46.72	56.52	75.54
$MZ_{\check{t}}$	100	0	5.58	18.82	41.18	60.40	77.88	84.90	93.90
		1	4.94	16.80	35.56	55.64	72.94	82.98	92.92
		4	5.90	16.42	30.86	47.76	66.90	78.36	91.02
		12	6.66	15.34	26.36	41.34	59.80	71.50	87.20
	250	0	5.02	23.94	47.12	66.92	81.96	89.00	94.78
		1	5.20	20.00	42.76	62.90	79.94	86.30	93.88
		5	5.32	17.02	35.22	54.52	72.48	82.76	92.46
		15	6.00	14.94	30.02	46.12	65.28	76.76	89.72
	500	0	5.28	25.84	52.38	71.22	85.42	90.38	96.08
		1	4.92	23.82	48.32	68.16	81.60	88.88	95.66
		5	5.34	19.48	40.38	60.98	76.96	85.58	93.86
		17	5.92	17.24	33.48	51.24	69.98	79.16	92.10
	1000	0	5.16	28.32	55.74	74.90	87.22	92.78	97.30
		1	5.12	26.60	53.46	71.38	85.16	91.06	96.58
		7	4.66	22.50	44.00	63.30	79.10	86.34	94.58
		21	5.40	17.52	36.02	55.22	73.36	82.24	92.40

For notation see footnote of Table 1.

Table 3 Percentage Power of $MZ_{\hat{\beta}}$ and $MZ_{\hat{t}}$ Tests Against I($1+d$) Alternatives, $d < 0$

			Value of d						
Test	T	l	0.0	-0.1	-0.2	-0.3	-0.4	-0.45	-0.49
$MZ_{\hat{\beta}}$	100	0	4.56	16.24	43.76	77.22	96.58	99.16	99.98
		1	4.14	13.10	32.78	60.82	88.40	95.60	99.08
		4	4.62	13.10	32.18	59.88	87.10	95.32	98.60
		12	3.80	13.80	38.12	70.00	94.32	98.54	99.82
	250	0	4.76	25.52	65.18	95.74	99.96	100.00	100.00
		1	4.06	20.34	56.12	89.28	99.44	99.98	100.00
		5	5.52	17.84	45.64	82.24	98.50	99.94	100.00
		15	5.26	20.00	51.72	88.48	99.74	99.98	100.00
	500	0	4.86	32.12	79.86	99.38	100.00	100.00	100.00
		1	4.90	27.16	72.18	97.58	100.00	100.00	100.00
		5	5.32	21.30	59.44	92.22	99.98	100.00	100.00
		17	5.66	22.18	61.06	94.46	100.00	100.00	100.00
	1000	0	5.32	39.88	89.66	99.96	100.00	100.00	100.00
		1	4.68	36.12	85.08	99.78	100.00	100.00	100.00
		7	5.14	24.82	68.78	97.72	100.00	100.00	100.00
		21	5.34	24.22	67.04	97.68	100.00	100.00	100.00
$MZ_{\hat{t}}$	100	0	4.92	12.80	35.28	68.02	93.46	97.94	99.74
		1	3.20	7.90	21.26	45.98	76.94	89.16	96.16
		4	3.44	7.92	21.86	46.54	76.92	90.00	96.24
		12	3.18	8.84	26.50	58.42	88.78	96.76	99.36
	250	0	5.02	19.98	56.44	92.56	99.76	100.00	100.00
		1	4.04	14.76	44.94	82.14	98.62	99.90	99.98
		5	5.06	13.22	36.94	74.38	96.88	99.70	99.98
		15	4.96	15.18	44.12	83.84	99.12	99.94	100.00
	500	0	5.32	26.36	72.56	98.58	100.00	100.00	100.00
		1	5.08	21.38	62.40	95.06	100.00	100.00	100.00
		5	5.14	16.76	50.18	87.80	99.76	99.98	100.00
		17	5.02	17.88	53.28	91.76	99.94	100.00	100.00
	1000	0	4.62	33.00	84.74	99.86	100.00	100.00	100.00
		1	5.16	29.68	78.54	99.38	100.00	100.00	100.00
		7	5.36	19.96	60.88	96.32	100.00	100.00	100.00
		21	5.08	19.72	60.72	96.20	100.00	100.00	100.00

For notation see footnote of Table 1.

Table 4 Percentage Power of $MZ_{\hat{\beta}}$ and $MZ_{\hat{t}}$ Tests Against I($1+d$) Alternatives, $d > 0$

			Value of d						
Test	T	l	0.0	0.1	0.2	0.3	0.4	0.45	0.49
$MZ_{\hat{\beta}}$	100	0	5.18	14.70	25.04	36.04	44.96	48.18	54.72
		1	4.30	12.32	23.60	31.42	41.96	46.92	54.60
		4	5.42	12.48	20.02	28.68	38.50	42.92	52.24
		12	5.42	10.78	19.02	25.40	34.04	39.86	51.28
	250	0	5.22	16.96	32.04	42.74	48.54	53.28	57.26
		1	4.74	16.08	28.50	38.80	46.04	49.48	56.62
		5	4.98	13.56	23.18	33.24	40.76	46.44	55.16
		15	5.72	11.88	20.66	28.56	38.32	43.72	53.60
	500	0	5.14	19.66	34.84	45.38	50.40	52.48	58.20
		1	5.02	17.30	32.70	41.26	48.12	53.16	57.86
		5	5.42	14.90	27.06	37.26	46.78	49.08	57.62
		17	5.22	13.04	21.60	31.86	41.78	46.26	56.02
	1000	0	4.94	21.46	37.72	46.76	51.42	52.80	60.88
		1	4.94	20.30	35.60	45.40	51.08	53.08	59.30
		7	5.12	16.60	28.64	39.38	46.28	50.98	56.10
		21	5.46	14.06	24.56	34.88	42.42	48.22	56.88
$MZ_{\hat{t}}$	100	0	5.16	14.40	24.10	34.70	43.02	45.00	48.20
		1	4.20	11.92	22.82	30.22	39.90	44.12	47.70
		4	5.38	12.28	19.28	27.58	36.48	40.44	45.84
		12	5.30	10.54	18.36	24.80	32.20	36.94	44.44
	250	0	5.08	16.46	30.82	40.98	45.42	49.42	48.26
		1	4.68	15.60	27.52	37.10	43.22	45.74	47.88
		5	4.82	13.26	22.52	31.92	38.62	42.78	46.84
		15	5.66	11.50	20.04	27.34	36.24	40.18	45.40
	500	0	5.04	19.20	33.42	43.20	46.82	48.42	48.84
		1	5.02	16.90	31.44	39.56	44.78	48.68	49.34
		5	5.38	14.72	26.28	35.38	43.92	44.82	48.46
		17	5.16	12.76	21.08	30.50	39.44	42.68	47.60
	1000	0	5.00	20.82	36.26	44.16	47.78	48.28	51.88
		1	5.02	20.04	34.36	43.32	47.88	48.72	49.18
		7	5.18	16.08	27.46	37.36	43.00	46.96	47.86
		21	5.50	13.78	23.84	33.68	40.20	44.42	48.36

For notation see footnote of Table 1.

Table 5 Percentage Power of $MZ_{\tilde{\beta}}$ and $MZ_{\tilde{t}}$ Tests Against I($1+d$) Alternatives, $d < 0$

			Value of d						
Test	T	l	0.0	-0.1	-0.2	-0.3	-0.4	-0.45	-0.49
$MZ_{\tilde{\beta}}$	100	0	5.16	28.94	72.32	97.28	100.00	100.00	100.00
		1	4.06	19.82	55.40	90.48	99.60	99.98	100.00
		4	5.18	19.08	49.88	86.22	99.32	99.92	100.00
		12	5.48	24.46	63.62	94.32	99.88	100.00	100.00
	250	0	5.42	36.86	87.04	99.84	100.00	100.00	100.00
		1	4.18	30.66	76.86	98.92	100.00	100.00	100.00
		5	5.16	23.98	65.60	96.04	99.98	100.00	100.00
		15	6.14	28.02	73.90	98.34	100.00	100.00	100.00
	500	0	4.36	45.54	95.46	100.00	100.00	100.00	100.00
		1	4.88	38.22	90.18	99.94	100.00	100.00	100.00
		5	4.48	29.22	76.24	98.98	100.00	100.00	100.00
		17	5.64	30.98	79.42	99.36	100.00	100.00	100.00
	1000	0	4.88	56.78	98.90	100.00	100.00	100.00	100.00
		1	4.82	50.48	97.14	100.00	100.00	100.00	100.00
		7	4.82	34.64	86.68	99.89	100.00	100.00	100.00
		21	5.70	32.78	86.40	99.85	100.00	100.00	100.00
$MZ_{\tilde{t}}$	100	0	5.34	26.06	67.72	96.00	99.94	100.00	100.00
		1	3.60	15.54	47.16	85.62	99.34	99.94	100.00
		4	4.26	15.02	42.34	81.40	98.54	99.88	100.00
		12	4.42	19.46	56.46	91.74	99.84	99.98	100.00
	250	0	5.18	33.04	83.42	99.76	100.00	100.00	100.00
		1	4.18	25.74	71.50	97.80	100.00	100.00	100.00
		5	4.66	20.10	59.12	94.20	99.96	100.00	100.00
		15	5.60	23.88	69.08	97.56	100.00	100.00	100.00
	500	0	4.54	42.00	93.94	100.00	100.00	100.00	100.00
		1	5.02	34.24	87.30	99.92	100.00	100.00	100.00
		5	4.32	26.30	72.36	98.52	100.00	100.00	100.00
		17	5.52	28.38	75.94	99.20	100.00	100.00	100.00
	1000	0	5.00	52.46	98.36	100.00	100.00	100.00	100.00
		1	4.94	46.02	95.94	100.00	100.00	100.00	100.00
		7	4.90	31.44	84.24	99.96	100.00	100.00	100.00
		21	5.10	30.02	84.02	99.84	100.00	100.00	100.00

For notation see footnote of Table 1.

Table 6 Percentage Power of $MZ_{\tilde{\beta}}$ and $MZ_{\tilde{t}}$ Tests Against I($1+d$) Alternatives, $d > 0$

			Value of d						
Test	T	l	0.0	0.1	0.2	0.3	0.4	0.45	0.49
$MZ_{\tilde{\beta}}$	100	0	5.10	16.32	31.48	48.36	61.66	65.38	69.04
		1	5.46	13.96	25.58	37.80	50.56	55.36	59.56
		4	4.42	11.34	18.90	26.78	36.10	40.60	42.62
		12	4.66	9.50	15.96	22.74	29.08	32.02	34.46
	250	0	5.42	20.96	44.74	64.54	75.84	78.76	80.12
		1	5.60	18.00	37.00	54.64	66.68	71.96	75.08
		5	5.24	13.14	25.14	36.30	49.80	53.62	56.06
		15	4.84	10.16	18.90	26.02	33.12	37.58	40.02
	500	0	4.80	26.26	56.10	74.44	83.00	83.28	83.66
		1	4.90	22.18	47.98	66.22	76.56	80.04	81.46
		5	5.36	18.30	33.38	50.46	61.92	66.66	68.98
		17	4.84	12.80	22.68	32.40	41.82	45.56	49.82
	1000	0	4.84	32.14	65.14	80.72	86.00	85.10	86.90
		1	4.80	27.62	58.72	76.02	83.40	84.32	84.88
		7	4.86	19.22	39.02	56.62	70.20	72.68	75.12
		21	5.30	14.86	27.30	39.18	51.70	55.76	59.00
$MZ_{\tilde{t}}$	100	0	4.82	13.78	23.98	35.52	44.54	46.48	48.62
		1	5.32	11.94	20.34	28.98	38.04	40.86	43.48
		4	4.24	9.80	16.48	22.50	28.44	31.34	32.58
		12	4.52	8.32	14.00	19.54	23.64	25.50	28.38
	250	0	5.16	17.28	33.20	45.74	54.26	56.42	56.16
		1	5.16	15.48	27.46	40.30	47.26	51.18	53.24
		5	5.04	11.42	20.62	29.00	38.10	40.24	41.42
		15	4.50	9.62	16.58	21.76	27.88	30.50	32.32
	500	0	4.74	20.76	41.46	53.10	57.86	57.92	58.12
		1	4.58	17.88	35.54	47.42	54.36	55.60	56.94
		5	5.06	15.62	26.52	37.28	44.40	47.66	48.16
		17	4.84	11.56	19.38	25.98	33.02	34.90	38.84
	1000	0	4.80	25.48	45.82	57.66	60.74	57.38	58.72
		1	4.72	21.46	42.38	54.94	58.96	58.38	59.08
		7	4.90	16.60	30.86	40.88	49.98	51.66	53.42
		21	5.66	13.52	21.58	30.66	38.38	41.16	43.16

For notation see footnote of Table 1.

Table 7 Percentage Power of $MZ_{\check{\beta}}$ and $MZ_{\check{t}}$ Tests Against I($1+d$)-GARCH(1, 1) Alternatives, $d < 0$

			Value of d						
Test	T	l	0.0	−0.1	−0.2	−0.3	−0.4	−0.45	−0.49
$MZ_{\check{\beta}}$	100	0	4.44	15.40	34.70	58.60	78.22	86.48	90.52
		1	4.32	12.46	28.24	47.32	69.56	76.66	82.20
		4	4.80	12.32	24.12	42.88	62.90	71.06	76.88
		12	5.58	12.84	28.20	48.70	70.82	77.10	85.30
	250	0	4.66	20.36	49.84	78.78	94.46	97.64	99.14
		1	4.78	17.76	41.24	69.28	88.64	93.22	96.64
		5	4.84	14.76	33.60	58.82	80.86	88.08	92.00
		15	5.86	15.72	34.86	64.18	83.76	90.56	94.98
	500	0	5.32	24.00	59.34	89.26	98.92	99.84	99.94
		1	4.78	21.96	53.46	82.58	96.64	98.90	99.62
		5	5.20	17.54	42.14	71.26	90.12	95.36	97.84
		17	5.50	15.30	38.58	70.94	90.50	96.00	98.60
	1000	0	5.24	28.76	70.56	95.62	99.90	100.00	100.00
		1	5.02	24.56	64.74	91.92	99.44	99.94	99.98
		7	4.74	19.30	49.44	79.94	96.16	98.90	99.48
		21	5.70	17.66	45.06	77.28	95.78	98.52	99.74
$MZ_{\check{t}}$	100	0	4.64	15.36	34.40	58.82	78.02	86.34	90.74
		1	4.18	12.10	27.56	46.38	68.72	76.10	82.10
		4	4.64	11.96	23.78	42.26	62.28	70.90	76.44
		12	5.34	12.14	27.34	47.84	70.16	76.82	85.22
	250	0	4.76	20.34	49.22	78.50	94.26	97.60	99.18
		1	4.62	17.38	40.44	68.24	88.24	93.10	96.54
		5	4.84	14.42	33.14	58.40	80.40	87.78	92.02
		15	5.76	15.44	34.36	63.66	83.56	90.32	94.82
	500	0	5.18	23.66	58.24	88.70	98.76	99.80	99.96
		1	4.52	21.28	52.42	82.44	96.46	98.82	99.60
		5	5.22	17.18	41.72	70.68	89.72	95.12	97.82
		17	5.28	15.00	38.02	70.62	90.40	95.86	98.62
	1000	0	5.14	28.40	69.96	95.32	99.90	100.00	100.00
		1	4.94	24.34	63.86	91.40	99.40	99.96	99.98
		7	4.68	19.26	48.92	79.48	96.06	98.90	99.50
		21	5.74	17.48	44.80	76.80	95.86	98.50	99.74

Note: 1. $MZ_{\check{\beta}}$ and $MZ_{\check{t}}$ are Perron and Ng's (1996) modified Phillips and Perron (1988) unit root test statistics.

2. DGP is $(1-L)^{1+d}y_t = \varepsilon_t$, $\varepsilon_t = \sqrt{h_t}v_t$, $h_t = 0.6 + 0.2h_{t-1} + 0.2\varepsilon_{t-1}^2$, $v_t \overset{i.i.d.}{\sim} N(0,1)$.

3. The number of lag l other than 1 is chosen as $l_0 = 0$, $l_4 = int[4(T/100)^{1/4}]$ and $l_{12} = int[12(T/100)^{1/4}]$.

4. Number of iterations is 5000. Power of the tests is based on the 5% lower-tailed significance level.

Table 8 Percentage Power of $MZ_{\check{\beta}}$ and $MZ_{\check{t}}$ Tests Against I($1+d$)-GARCH(1, 1) Alternatives, $d > 0$

			Value of d						
Test	T	l	0.0	0.1	0.2	0.3	0.4	0.45	0.49
$MZ_{\check{\beta}}$	100	0	4.64	12.38	23.76	34.22	46.26	55.66	73.60
		1	5.02	12.90	22.60	33.26	46.14	55.06	73.54
		4	5.46	11.04	19.64	30.54	41.96	53.18	68.68
		12	5.88	11.62	18.32	26.68	38.36	46.60	66.60
	250	0	5.08	15.84	28.92	39.60	50.92	59.78	76.74
		1	5.52	15.22	26.52	38.04	48.98	59.08	76.78
		5	5.50	12.50	22.76	34.66	47.20	57.82	74.84
		15	5.68	12.14	20.96	30.34	43.84	52.94	72.72
	500	0	5.30	17.08	31.44	40.38	51.86	60.74	77.80
		1	4.88	16.66	27.30	39.56	51.24	60.58	77.82
		5	4.78	13.66	26.40	37.52	48.42	58.32	76.92
		17	5.98	12.46	22.06	32.40	45.16	55.38	74.80
	1000	0	4.88	19.04	31.28	43.12	51.72	61.18	78.78
		1	4.94	17.88	31.34	41.86	52.74	61.24	78.78
		7	4.76	14.96	26.18	38.32	49.32	58.04	77.86
		21	5.36	13.48	23.48	34.26	48.58	57.26	75.70
$MZ_{\check{t}}$	100	0	4.70	19.18	38.96	59.08	76.66	84.90	93.96
		1	5.80	18.18	36.04	55.28	72.66	82.74	92.86
		4	5.70	15.68	31.34	47.60	66.82	78.46	91.94
		12	7.06	15.18	27.08	42.12	60.56	72.14	87.92
	250	0	5.22	23.82	47.94	67.28	82.12	88.68	94.94
		1	4.86	20.24	42.28	63.24	78.74	86.60	93.96
		5	5.32	17.72	35.80	55.96	72.08	81.56	92.94
		15	5.84	15.90	29.84	46.30	65.26	76.12	89.92
	500	0	4.96	26.92	51.74	71.46	84.84	90.20	96.96
		1	5.22	25.02	46.92	67.66	82.46	89.08	95.96
		5	4.56	19.94	41.74	60.92	76.22	85.14	93.98
		17	5.66	16.46	32.82	52.08	70.26	79.70	92.94
	1000	0	5.32	28.74	55.12	76.00	86.58	92.22	97.98
		1	5.12	27.16	53.56	72.06	86.02	90.32	96.96
		7	5.16	22.28	43.08	63.70	79.50	87.08	94.00
		21	5.68	18.44	37.16	55.44	73.98	81.58	92.90

For notation see footnote of Table 7.

and MZ_t unit root tests against the alternatives I($1+d$), $d = 0.1, 0.2, \ldots, 0.49$. The most important result is that, except for $MZ_{\check{t}}$ in Table 2, with d fixed, the power does not approach 100% as T increases. For example, for $d = 0.3$ and $l = l_4$, the power of $MZ_{\hat{t}}$ grows from 27.58% with $T = 100$ to only 37.36% with $T = 1000$. This is a reflection of the inconsistency of the $MZ_{\check{\beta}}$, $MZ_{\hat{\beta}}$, $MZ_{\hat{t}}$, $MZ_{\tilde{\beta}}$, and $MZ_{\tilde{t}}$ unit root tests against non-stationary and non-mean-reverting fractionally-integrated processes; the power is not expected to approach 100% even for an arbitrarily large value of T. The power of $MZ_{\check{t}}$ in Table 2 confirms that $MZ_{\check{t}}$ is consistent, but even with a rather large sample such as $T = 1000$, the power has not approached to 100%. We finally see that, for a fixed T, the power increases as d increases. This is not surprising, but not transparent from the relevant asymptotics for $0 < d < 0.5$.

Finally, to find the robustness of our results from the assumption of strong stationarity, we consider the shock variable ε_t to be a GARCH(1,1) process. In particular we assume that $\varepsilon_t = \sqrt{h_t} v_t$, $h_t = 0.6 + 0.2h_{t-1} + 0.2\varepsilon_{t-1}^2$,[12] $v_t \overset{i.i.d.}{\sim} N(0,1)$. Tables 7 and 8 provide the percentage power of 5% lower and upper tail of $MZ_{\check{\beta}}$ and $MZ_{\check{t}}$, respectively, and they are not much different from their counterparts as in Tables 1 and 2.[13] Therefore, the simulation shows that our results are robust to autoregressive conditional heteroskedasticity.

5. CONCLUSION

In this paper we have generalized the fractional unit root distribution of Sowell (1990) to a general fractionally-integrated process. Our characterization of this generalized fractional unit root distribution also extends those parametric representation in Tanaka (1999) and Wang et al. (2003) to a general semiparametric fractionally-integrated process. We also show that the Phillips and Perron's unit root test statistics Z_β and Z_t can be used to distinguish a unit root non-stationary process from a non-stationary but mean-reverting I($1+d$), $d < 0$ process, although it has low power. Only the t statistic from a regression without drift and the time trend model $Z_{\check{t}}$, is consistent against the non-stationary and non-mean-reverting I($1+d$), $d > 0$ process. Moreover, we have provided Monte-Carlo evidence on their power in finite samples and show the robustness of our results to autoregressive conditional heteroskedasticity.

[12] This assumption would satisfy the requirement of nonnegativity and stationarity.

[13] We also perform $MZ_{\hat{\beta}}$, $MZ_{\hat{t}}$, $MZ_{\tilde{\beta}}$ and $MZ_{\tilde{t}}$ under this assumption. This conclusion does not change.

Appendix: Proofs of the Theorems

We first provide the following asymptotic results of the sample moments which are useful to derive the asymptotics of the OLS estimator.

Lemma A.1 Let $(1-L)^d u_t = \varepsilon_t$, where $-0.5 < d < 0.5$ and ε_t satisfies Assumption 1. Define

$$\gamma_j = E(u_t u_{t-j}) \text{ for } j = 0, 1, 2, ...,$$

and

$$y_t = u_1 + u_2 + \cdots + u_t \text{ for } t = 1, 2, ..., T, \tag{A1}$$

with $y_0 = 0$. Therefore,

(a) $T^{-\frac{1}{2}-d} \sum_{t=1}^{T} u_t \Rightarrow V_d^{\frac{1}{2}} \sigma_\varepsilon B_d(1)$,

(b) $T^{-2-2d} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow V_d \sigma_\varepsilon^2 \int_0^1 [B_d(r)]^2 dr$,

(c) $T^{-1-2d} y_T^2 \Rightarrow V_d \sigma_\varepsilon^2 [B_d(1)]^2$,

(d) $T^{-\frac{3}{2}-d} \sum_{t=1}^{T} y_{t-1} \Rightarrow V_d^{\frac{1}{2}} \sigma_\varepsilon \int_0^1 B_d(r) dr$,

(e) $T^{-1} \sum_{t=1}^{T} u_t u_{t-j} \xrightarrow{p} \gamma_j$,

(f) If $d > 0$,

$$T^{-1-2d} \sum_{t=j+1}^{T} y_{t-1} u_{t-j} \Rightarrow \frac{1}{2} V_d \sigma_\varepsilon^2 [B_d(1)]^2 \text{ for } j = 0, 1, 2, ...,$$

(g) If $d < 0$,

$$T^{-1} \sum_{t=j+1}^{T} y_{t-1} u_{t-j} \xrightarrow{p} \begin{cases} -\frac{1}{2}\sigma_u^2 & \text{for } j = 0; \\ -\frac{1}{2}\sigma_u^2 + \gamma_0 + \gamma_1 + \gamma_2 + \cdots + \gamma_{j-1} & \text{for } j = 1, 2, ..., \end{cases}$$

(h) If $d > 0$,

$$T^{-1-2d}\sum_{t=j+1}^{T} u_t y_{t-1-j} \Rightarrow \frac{1}{2}V_d\sigma_\varepsilon^2[B_d(1)]^2 \text{ for } j=0,1,2,...,$$

(i) If $d<0$,

$$T^{-1}\sum_{t=j+1}^{T} u_t y_{t-1-j} \xrightarrow{p} \begin{cases} -\frac{1}{2}\sigma_u^2 & \text{for } j=0; \\ -\frac{1}{2}\sigma_u^2-\gamma_1-\gamma_2-\gamma_3-....-\gamma_j & \text{for } j=1,2,...,\end{cases}$$

(j) $T^{-\frac{3}{2}-d}\sum_{t=1}^{T} tu_t \Rightarrow V_d^{\frac{1}{2}}\sigma_\varepsilon[B_d(1)-\int_0^1 B_d(r)dr]$,

(k) $T^{-\frac{5}{2}-d}\sum_{t=1}^{T} ty_{t-1} \Rightarrow V_d^{\frac{1}{2}}\sigma_\varepsilon\int_0^1 rB_d(r)dr$,

(l) $T^{-3-2d}\sum_{t=1}^{T} ty_{t-1}^2 \Rightarrow V_d\sigma_\varepsilon^2\int_0^1 r[B_d(r)]^2dr$,

(m) $T^{-2-2d}\sum_{t=j+1}^{T} y_{t-1}y_{t-1-j} \Rightarrow V_d\sigma_\varepsilon^2\int_0^1 [B_d(r)]^2dr$.

A joint weak convergence for the sample moments given above to their respective limits is easily established and will be used below.

Proof of Lemma A.1

The proofs of items (a) to (d) are a straightforward application of the continuous mapping theorem from Lemma 1(b)'s results. Item (e) is due to ergodicity of the stationary linear process u_t.

To prove items (f) and (g) for $j=0$, recall that $y_0=0$, and thus it is convenient to write $\sum_{t=1}^T y_{t-1}u_t = 1/2y_T^2 - 1/2\sum_{t=1}^T u_t^2$. From items (c) and (e), we know that y_T^2 is $O_p(T^{1+2d})$ and $\sum_{t=1}^T u_t^2$ is $O_p(T)$; then, $\sum_{t=1}^T y_{t-1}u_t$ would be $O_p(T^\kappa)$, where $\kappa=\max(1+2d,1)$. Therefore, for $d>0$,

$$T^{-1-2d}\sum_{t=1}^{T} y_{t-1}u_t \Rightarrow T^{-1-2d}\frac{1}{2}y_T^2 \Rightarrow \frac{1}{2}V_d\sigma_\varepsilon^2[B_d(1)]^2, \tag{A2}$$

and for $d<0$,

$$T^{-1}\sum_{t=1}^{T} y_{t-1}u_t \Rightarrow -\frac{1}{2}\sum_{t=1}^{T} u_t^2 \xrightarrow{p} -\frac{1}{2}\sigma_u^2, \tag{A3}$$

which establish results (f) and (g) for $j=0$.

For $j > 0$, observe that

$$y_{t-1} = y_{t-j-1} + u_{t-j} + u_{t-j+1} + \cdots + u_{t-1},$$

which implies that

$$\begin{aligned}\sum_{t=j+1}^{T} y_{t-1}u_{t-j} &= \sum_{t=j+1}^{T} (y_{t-j-1} + u_{t-j} + u_{t-j+1} + \cdots + u_{t-1})u_{t-j} \\ &= \sum_{t=j+1}^{T} y_{t-j-1}u_{t-j} + \sum_{t=j+1}^{T} (u_{t-j} + u_{t-j+1} + \cdots + u_{t-1})u_{t-j}.\end{aligned}$$

If $d > 0$, then

$$\begin{aligned}T^{-1-2d} \sum_{t=j+1}^{T} y_{t-j-1}u_{t-j} &= \left(\frac{T-j}{T}\right)^{1+2d} (T-j)^{-1-2d} \sum_{t=1}^{T-j} y_{t-1}u_t \\ &\Rightarrow \frac{1}{2} V_d \sigma_\varepsilon^2 [B_d(1)]^2\end{aligned}$$

as in (A2). Moreover,

$$T^{-1-2d} \sum_{t=j+1}^{T} (u_{t-j} + u_{t-j+1} + \cdots + u_{t-1})u_{t-j} \xrightarrow{p} 0$$

from result (e). Thus,

$$T^{-1-2d} \sum_{t=j+1}^{T} y_{t-1}u_{t-j} \Rightarrow \frac{1}{2} V_d \sigma_\varepsilon^2 [B_d(1)]^2.$$

The proof for $d < 0$ in item (g) is analogous.

The proof of items (h) and (i) is analogous with items (f) and (g).

To prove item (j), we first observe that $\sum_{t=1}^{T} y_{t-1} = \sum_{t=1}^{T} Tu_t - \sum_{t=1}^{T} tu_t$, or $\sum_{t=1}^{T} tu_t = T\sum_{t=1}^{T} u_t - \sum_{t=1}^{T} y_{t-1}$. Therefore, $T^{-3/2-d}\sum_{t=1}^{T} tu_t = T^{-1/2-d}\sum_{t=1}^{T} u_t - T^{-3/2-d}\sum_{t=1}^{T} y_{t-1}$. By applying the continuous mapping theorem to the joint convergence of items (a) and (d), we have

$$T^{-\frac{3}{2}-d}\sum_{t=1}^{T} tu_t \Rightarrow V_d^{\frac{1}{2}}\sigma_\varepsilon B_d(1) - V_d^{\frac{1}{2}}\sigma_\varepsilon \int_0^1 B_d(r)dr.$$

To prove items (k) and (l), recall the notation that $X_T(r) = \sum_{t=1}^{[Tr]} u_t$, for $r \in [0,1]$. As $T \to \infty$, the following results then hold.

$$T^{-1}\sum_{t=1}^{T} \frac{t}{T} y_{t-1} = \int_0^1 rX_T(r)dr, \tag{A4}$$

and

$$T^{-1}\sum_{t=1}^{T} \frac{t}{T} y_{t-1}^2 = \int_0^1 rX_T^2(r)dr. \tag{A5}$$

The result of item (k) follows immediately from (A4) and item (d). Similarly, the result of item (l) follows immediately from (A5) and item (b).

For item (m), we finally observe that

$$\begin{aligned}
&T^{-2-2d}\sum_{t=j+1}^{T} y_{t-1}y_{t-1-j} \\
&\quad = T^{-2-2d}\sum_{t=j+1}^{T} (y_{t-1-j} + u_{t-j} + u_{t-j+1} + \cdots + u_{t-1})y_{t-1-j} \\
&\quad = T^{-2-2d}\sum_{t=j+1}^{T} (y_{t-1-j}^2 + u_{t-j}y_{t-1-j} + u_{t-j+1}y_{t-1-j} + \cdots + u_{t-1}y_{t-1-j}),
\end{aligned}$$

which converge to $V_d\sigma_\varepsilon^2 \int_0^1 [B_d(r)]^2 dr$ by virtue of items (b), (f), and (g). This completes the proofs of Lemma A.1.

Proof of Theorem 1

We prove items (g) and (h) from regression model (10). The proof of models (8) and (9) is analogous. Let the data generating process be

$$y_t = \alpha + y_{t-1} + u_t,$$

and the regression model be

$$y_t = \alpha + \beta y_{t-1} + \delta t + u_t. \tag{A6}$$

Note that the regression model of (A6) can be equivalently rewritten as

$$\begin{aligned} y_t &= (1-\beta)\alpha + \beta(y_{t-1} - \alpha(t-1)) + (\delta + \beta\alpha)t + u_t, \\ &\equiv \alpha^* + \beta^* \xi_{t-1} + \delta^* t + u_t, \end{aligned} \tag{A7}$$

where $\alpha^* = (1-\beta)\alpha$, $\beta^* = \beta$, $\delta^* = \delta + \beta\alpha$, and $\xi_{t-1} = y_{t-1} - \alpha(t-1)$. Moreover, under the null hypothesis that $\beta = 1$ and $\delta = 0$,

$$\xi_t = y_0 + u_1 + u_2 + \cdots + u_t;$$

that is, ξ_t is the random walk as described in (A1). Under the maintained hypothesis, $\alpha = \alpha_0$, $\beta = 1$, and $\delta = 0$, which in (A7) means that $\alpha^* = 0$, $\beta^* = 1$ and $\delta^* = \alpha_0$. The deviation of the OLS estimate from these true values is given by

$$\begin{bmatrix} \tilde{\alpha}^* \\ \tilde{\beta} - 1 \\ \tilde{\delta}^* - \alpha_0 \end{bmatrix} = \begin{bmatrix} T & \sum_{t=1}^{T} \xi_{t-1} & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} \xi_{t-1} & \sum_{t=1}^{T} \xi_{t-1}^2 & \sum_{t=1}^{T} t\xi_{t-1} \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t\xi_{t-1} & \sum_{t=1}^{T} t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} u_t \\ \sum_{t=1}^{T} \xi_{t-1} u_t \\ \sum_{t=1}^{T} t u_t \end{bmatrix}, \tag{A8}$$

or in shorthand as

$$\mathbf{C} = \mathbf{A}^{-1}\mathbf{f}.$$

From Lemma A.1, the order of probability of the individual terms in (A8) is as follows,

$$\begin{bmatrix} \tilde{\alpha}^* \\ \tilde{\beta} - 1 \\ \tilde{\delta}^* - \alpha_0 \end{bmatrix} = \begin{bmatrix} O_p(T) & O_p(T^{\frac{3}{2}+d}) & O_p(T^2) \\ O_p(T^{\frac{3}{2}+d}) & O_p(T^{2+2d}) & O_p(T^{\frac{5}{2}+d}) \\ O_p(T^2) & O_p(T^{\frac{5}{2}+d}) & O_p(T^3) \end{bmatrix}^{-1} \begin{bmatrix} O_p(T^{\frac{1}{2}+d}) \\ O_p(T^{\max[1+2d,1]}) \\ O_p(T^{\frac{3}{2}+d}) \end{bmatrix}.$$

To prove item (g), we note that when $d > 0$, $\sum_{t=1}^{T} y_{t-1}u_t$ is $O_p(T^{1+2d})$ as in item (f) of Lemma A1. We define two rescaling matrices,

$$\Upsilon_T = \begin{bmatrix} T^{\frac{1}{2}} & 0 & 0 \\ 0 & T^{1+d} & 0 \\ 0 & 0 & T^{\frac{3}{2}} \end{bmatrix} \text{ and } \aleph_T = \begin{bmatrix} T^{-\frac{1}{2}-d} & 0 & 0 \\ 0 & T^{-1-2d} & 0 \\ 0 & 0 & T^{-\frac{3}{2}-d} \end{bmatrix}.$$

Multiplying the rescaling matrices on (A8), we get

$$\Upsilon_T \mathbf{C} = \Upsilon_T \mathbf{A}^{-1} \Upsilon_T \Upsilon_T^{-1} \aleph_T^{-1} \aleph_T \mathbf{f}. \tag{A9}$$

Substituting the results of Lemma A.1 to (A9), we establish that

$$\tilde{\mathbf{b}}_{\mathbf{1}} \Rightarrow \mathbf{Q}^{-\mathbf{1}} \mathbf{h}_{\mathbf{1}}, \tag{A10}$$

where

$$\tilde{\mathbf{b}}_{\mathbf{1}} \equiv \begin{bmatrix} T^{\frac{1}{2}-d}\tilde{\alpha^*} \\ T(\tilde{\beta}-1) \\ T^{\frac{3}{2}-d}(\tilde{\delta^*}-\alpha_0) \end{bmatrix},$$

$$\mathbf{Q} \equiv \begin{bmatrix} 1 & V_d^{\frac{1}{2}}\sigma_\varepsilon \int_0^1 B_d(r)dr & \frac{1}{2} \\ V_d^{\frac{1}{2}}\sigma_\varepsilon \int_0^1 B_d(r)dr & V_d\sigma_\varepsilon^2 \int_0^1 [B_d(r)]^2 dr & V_d^{\frac{1}{2}}\sigma_\varepsilon \int_0^1 rB_d(r)dr \\ \frac{1}{2} & V_d^{\frac{1}{2}}\sigma_\varepsilon \int_0^1 rB_d(r)dr & \frac{1}{3} \end{bmatrix},$$

$$\mathbf{h}_{\mathbf{1}} \equiv \begin{bmatrix} V_d^{\frac{1}{2}}\sigma_\varepsilon B_d(1) \\ \frac{1}{2}V_d\sigma_\varepsilon^2[B_d(1)]^2 \\ V_d^{\frac{1}{2}}\sigma_\varepsilon[B_d(1) - \int_0^1 B_d(r)dr] \end{bmatrix}.$$

Thus, the asymptotic distribution of $T(\tilde{\beta}-1)$ is given by the middle row of (A10), which is

$$T(\tilde{\beta}-1) \Rightarrow \frac{W_2}{W_3}.$$

Note that this distribution does not depend on α.

To prove item (k), we notice that when $d<0$, $\sum_{t=1}^{T} y_{t-1}u_t$ is $O_p(T)$. We define another rescaling matrix

$$\Re_T = \begin{bmatrix} T^{-\frac{1}{2}+d} & 0 & 0 \\ 0 & T^{-1} & 0 \\ 0 & 0 & T^{-\frac{3}{2}+d} \end{bmatrix}.$$

Multiplying Υ_T and $\Re_T$ on (A8) to get

$$\Upsilon_T \mathbf{C} = \Upsilon_T \mathbf{A}^{-1} \Upsilon_T \Upsilon_T^{-1} \Re_T^{-1} \Re_T \mathbf{f}. \tag{A11}$$

Substitute the results of lemma A.1 to (A11), we establish that

$$\tilde{\mathbf{b}}_{\mathbf{2}} \Rightarrow \mathbf{Q}^{-\mathbf{1}} \mathbf{h}_{\mathbf{2}}, \tag{A12}$$

where

$$\tilde{\mathbf{b}}_{\mathbf{2}} \equiv \begin{bmatrix} T^{\frac{1}{2}+d}\tilde{\alpha}^* \\ T^{1+2d}(\tilde{\beta}-1) \\ T^{\frac{3}{2}+d}(\tilde{\delta}^*-\alpha_0) \end{bmatrix} \text{ and } \mathbf{h}_{\mathbf{2}} \equiv \begin{bmatrix} 0 \\ -\frac{1}{2}\sigma_u^2 \\ 0 \end{bmatrix}.$$

Thus, the asymptotic distribution of $T^{1+2d}(\tilde{\beta}-1)$ is given by the middle row of (A12), which is

$$T^{1+2d}(\tilde{\beta}-1) \Rightarrow -\frac{1}{2}\frac{\sigma_u^2}{V_d \sigma_\varepsilon^2 W_3}.$$

This completes the proofs of Theorem 1.

Proof of Lemma 2

We prove the case where $s_u^2 = \breve{s}_u^2$ and $s_{Tl}^2 = \breve{s}_{Tl}^2$ from regression model (8). The proofs of models (9) and (10) are analogous.

To prove item (a), since $\breve{\beta}$ is consistent as shown in Theorem 1 and u_t is ergodic, then applying Theorem 4 from Sowell (1990), we have

$$s_u^2 = T^{-1} \sum_{t=1}^{T} \breve{u}_t^2 \xrightarrow{p} \sigma_u^2.$$

To prove item (b), we start by defining the population counterpart of $\breve{s}_{Tl}^2$, i.e., $\sigma^2(l)$ as

$$\sigma^2(l) = T^{-1} \sum_{t=1}^{T} E(u_t^2) + 2T^{-1} \sum_{\tau=1}^{l} w_{\tau l} \sum_{t=\tau+1}^{T} E(u_t u_{t-\tau}).$$

Following Lee and Schmidt (1996) we have

$$l^{-2d}\sigma^2(l) \xrightarrow{p} \sigma_\varepsilon^2 V_d.$$

Given $l = o(T^{1/4})$, Tsay (2001, Lemma 1 and Theorem 1) has shown that

$$\breve{s}_{Tl}^2 - \sigma^2(l) \xrightarrow{p} 0,$$

that is, the difference between $\sigma^2(l)$ and $\breve{s}_{Tl}^2$ will converge in probability to zero. The desired result then is obtained. This completes the proofs of Lemma 2.

Proof of Theorem 2

To prove item (a), we first rewrite $Z_{\breve{\beta}}$ as

$$Z_{\breve{\beta}} = T(\breve{\beta} - 1) - \frac{1}{2} \frac{T^2 \sigma_{\breve{\beta}}^2}{\breve{s}_u^2} (\breve{s}_{Tl}^2 - \breve{s}_u^2), \tag{A13}$$

where as a matter of notation, $\sigma_{\breve{\beta}}^2 = \breve{s}_u^2 / \sum_{t=1}^{T} y_{t-1}^2$.

By item (b) of Lemma A.1, we have

$$\frac{T^{2+2d}\sigma_{\check{\beta}}^2}{\check{s}_u^2} \Rightarrow \frac{1}{V_d\sigma_\varepsilon^2 \int_0^1 [B_d(r)]^2 dr}. \tag{A14}$$

Multiplying T^{2d} on $Z_{\check{\beta}}$, and collecting the results of Theorem 1(b), (A14), and Lemma 2, we obtain

$$\begin{aligned} T^{2d}Z_{\check{\beta}} &= T^{1+2d}(\check{\beta}-1) - \frac{1}{2}\left(\frac{T^{2+2d}\sigma_{\check{\beta}}^2}{\check{s}_u^2}\right)(\check{s}_{Tl}^2 - \check{s}_u^2) \\ &\Rightarrow -\frac{1}{2}\frac{V_d\sigma_\epsilon^2 l^{2d}}{V_d\sigma_\epsilon^2 \int_0^1 [B_d(r)]^2 dr}; \end{aligned}$$

therefore,

$$\left(\frac{T}{l}\right)^{2d} Z_{\check{\beta}} \Rightarrow -\frac{1}{2}\frac{1}{\int_0^1 [B_d(r)]^2 dr}.$$

To prove item (b), we note that

$$\begin{aligned} Z_{\check{t}} &= \frac{(\check{\beta}-1)(\sum_{t=1}^T y_{t-1}^2)^{\frac{1}{2}}}{\check{s}_{Tl}} - \frac{1}{2}\frac{\check{s}_{Tl}^2 - \check{s}_u^2}{\check{s}_{Tl}(T^{-2}\sum_{t=1}^T y_{t-1}^2)^{1/2}} \\ &= \frac{(T^{-2}\sum_{t=1}^T y_{t-1}^2)^{\frac{1}{2}}}{\check{s}_{Tl}}\left[T(\check{\beta}-1) - \frac{1}{2}\frac{\check{s}_{Tl}^2 - \check{s}_u^2}{(T^{-2}\sum_{t=1}^T y_{t-1}^2)}\right] \\ &= \frac{1}{\check{s}_{Tl}}\left(\frac{T\sigma_{\check{\beta}}}{\check{s}_u}\right)^{-1} Z_{\check{\beta}} \\ &\Rightarrow \frac{1}{(V_d\sigma_\epsilon^2 l^{2d})^{\frac{1}{2}}} \times \left\{T^{2d}V_d\sigma_\epsilon^2 \int_0^1 [B_d(r)]^2 dr\right\}^{\frac{1}{2}} \times \\ &\qquad \left[-\frac{1}{2}\left(\frac{T}{l}\right)^{-2d}\frac{1}{\int_0^1 [B_d(r)]^2 dr}\right] \\ &= -\frac{1}{2}\left(\frac{T}{l}\right)^{-d}\frac{1}{(\int_0^1 [B_d(r)]^2 dr)^{\frac{1}{2}}}. \end{aligned} \tag{A15}$$

Therefore,

$$\left(\frac{T}{l}\right)^d Z_{\check{t}} \Rightarrow -\frac{1}{2}\frac{1}{\{\int_0^1 [B_d(r)]^2 dr\}^{\frac{1}{2}}}.$$

To prove items (a') and (b'), just use the results of Theorem 1(a) to replace all steps in the proofs of items (a) and (b). For example, Theorem 1(a) implies that the first term of (A13) is $O_p(1)$ and (A14) implies that the second term of (A13) is $o_p(T^{-2d})$. Therefore, we have the results that $Z_{\check{\beta}} = O_p(1) - o_p(T^{-2d}) = O_p(1) \Rightarrow (1/2)([B_d(1)]^2/\int_0^1 [B_d(r)]^2 dr)$.

To prove item (c), using the fact as in (A14), we have

$$\frac{T^{2+2d}\sigma^2_{\hat{\beta}}}{\hat{s}^2_u} \Rightarrow \frac{1}{V_d \sigma^2_{\varepsilon} W_1},$$

where $\sigma^2_{\hat{\beta}} = \hat{s}^2_u \mathbf{e}'_{\mathbf{1}} \mathbf{A}^{-\mathbf{1}}_{\mathbf{1}} \mathbf{e}_{\mathbf{1}}$, $\mathbf{e}_{\mathbf{1}} = [0, 1]'$, and $\mathbf{A}_{\mathbf{1}}$=

$$\begin{bmatrix} T & \sum_{t=1}^{T} y_{t-1} \\ \sum_{t=1}^{T} y_{t-1} & \sum_{t=1}^{T} y^2_{t-1} \end{bmatrix}.$$

Multiplying T^{2d} on $Z_{\hat{\beta}}$, we have

$$\begin{aligned} T^{2d} Z_{\hat{\beta}} &= T^{1+2d}(\hat{\beta} - 1) - \frac{1}{2}\left(\frac{T^{2+2d}\sigma^2_{\hat{\beta}}}{\hat{s}^2_u}\right)(\hat{s}^2_{Tl} - \hat{s}^2_u) \\ &\Rightarrow -\frac{1}{2}\frac{V_d \sigma^2_{\epsilon} l^{2d}}{V_d \sigma^2_{\epsilon} W_1}; \end{aligned}$$

therefore,

$$\left(\frac{T}{l}\right)^{2d} Z_{\hat{\beta}} \Rightarrow -\frac{1}{2}\frac{1}{W_1}. \tag{A16}$$

To prove item (d) of Theorem 2, we note that

$$\begin{aligned} Z_{\hat{t}} &= \frac{1}{\hat{s}_{Tl}} \left(\frac{T\sigma_{\hat{\beta}}}{\hat{s}_u} \right)^{-1} Z_{\hat{\beta}} \\ &\Rightarrow \frac{1}{(V_d \sigma_\epsilon^2 l^{2d})^{\frac{1}{2}}} \times (T^{2d} V_d \sigma_\epsilon^2 W_1)^{\frac{1}{2}} \times \left[-\frac{1}{2} \left(\frac{T}{l} \right)^{-2d} \frac{1}{W_1} \right] \\ &= -\frac{1}{2} \left(\frac{T}{l} \right)^{-d} \frac{1}{W_1^{\frac{1}{2}}}; \end{aligned}$$

thus,

$$\left(\frac{T}{l} \right)^{d} Z_{\hat{t}} \Rightarrow -\frac{1}{2} \frac{1}{W_1^{\frac{1}{2}}}. \tag{A17}$$

Similarly, to prove items (c') and (d'), just use the results of Theorem 1(d) to replace all the steps in the proofs of items (c) and (d).

To prove items (e) and (f), using the fact as in (A14), we have

$$\frac{T^{2+2d} \sigma_{\tilde{\beta}}^2}{\tilde{s}_u^2} \Rightarrow \frac{1}{V_d \sigma_{\tilde{\varepsilon}}^2 W_3},$$

where $\sigma_{\tilde{\beta}}^2 = \tilde{s}_u^2 \mathbf{e}_2' \mathbf{A}_2^{-1} \mathbf{e}_2$, $\mathbf{e}_2 = [0, 1, 0]'$, and $\mathbf{A}_2$=

$$\begin{bmatrix} T & \sum_{t=1}^{T} y_{t-1} & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} y_{t-1} & \sum_{t=1}^{T} y_{t-1}^2 & \sum_{t=1}^{T} t y_{t-1} \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t y_{t-1} & \sum_{t=1}^{T} t^2 \end{bmatrix}.$$

We obtain results of items (e) and (f) that simply replace W_1 in (A16) and (A17) by W_3.

Similarly, to prove items (e') and (f'), just use the result of Theorem 1(g). This completes the proofs of Theorem 2.

Proof of Corollary 1

By definition,

$$MZ_{\check{\beta}} = Z_{\check{\beta}} + \frac{T}{2}(\check{\beta} - 1)^2. \tag{A18}$$

For $d < 0$, the first term in the right hand-side of (A18) is $O_p(T^{-2d})$ and the second term is $O_p(T^{-1-3d})$ from Theorem 2(a) and Theorem 1(b), respectively. Given $-0.5 < d < 0$, $-2d$ is never to be less than $-1 - 3d$. Therefore, we have the results that $MZ_{\check{\beta}}$ and $Z_{\check{\beta}}$ are asymptotically equivalent. Consequently, for $d > 0$, the first term in the right-hand side of (A18) is $O_p(1)$ and the second term is $O_p(T^{-1})$ from Theorem 2(b) and Theorem 1(a), respectively. It follows that $MZ_{\check{\beta}}$ and $Z_{\check{\beta}}$ are asymptotically equivalent.

We next note that

$$MZ_{\check{t}} = MSB \cdot MZ_{\check{\beta}}.$$

It is easy to show that

$$MSB \Rightarrow \left(\frac{T}{l}\right)^d \left(\int_0^1 [B_d(r)]^2 dr\right)^{\frac{1}{2}}$$

as in (A15). Therefore, $MZ_{\check{t}}$ has the asymptotics as in $Z_{\check{t}}$. The proofs for $MZ_{\hat{\beta}}$, $MZ_{\tilde{\beta}}$, $MZ_{\hat{t}}$, and $MZ_{\tilde{t}}$ are analogous. This completes the proofs of Corollary 1.

REFERENCES

Avram, F. and M. S. Taqqu (1987), "Noncentral Limit Theorems and Appell Polynomials," *Annals of Probability*, 15, 767–775.

Baillie, R. T. (1996), "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5–59.

Beran, J. (1994), *Statistics for Long-Memory Processes*, New York: Chapman & Hall.

Chung, C. F. (2002), "Sample Means, Sample Autocovariances, and Linear Regression of Stationary Multivariate Long Memory Processes," *Econometric Theory*, 18, 51–78.

Davidson, J. (1994), *Stochastic Limit Theory*, Oxford: Oxford University Press.

Davidson, J. and R. M. De Jong (2000), "The Functional Central Limit Theorem and Weak Convergence to Stochastic Integrals II: Fractionally Integrated Processes," *Econometric Theory*, 16, 643–666.

Davydov, Y. A. (1970), "The Invariance Principle for Stationary Processes," *Theory of Probability and Its Applications*, 15, 487–489.

Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

Diebold, F. X. and G. D. Rudebusch (1991), "On the Power of Dickey-Fuller Tests Against Fractional Alternatives," *Economics Letters*, 35, 155–160.

Dittmann, I. (2000), "Residual-Based Tests for Fractional Cointegration: A Monte Carlo Study," *Journal of Time Series Analysis*, 21(6), 615–647.

Geweke, J. F. and S. Porter-Hudak (1983), "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

Granger, C. W. J. and R. Joyeux (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, 1, 15–39.

Hassler, U. and J. Wolters (1994), "On the Power of Unit Root Tests Against Fractional Alternatives," *Economics Letters*, 45, 1–5.

Helson, J. and Y. Sarason (1967), "Past and Future," *Mathematica Scandanavia*, 21, 5–16.

Krämer, W. (1998), "Fractional Integration and the Augmented Dickey-Fuller Test," *Economics Letters*, 61, 269–272.

Lee, D., and P. Schmidt (1996), "On the Power of the KPSS Test of Stationarity Against Fractionally-Integrated Alternatives," *Journal of Econometrics*, 73, 285–302.

Marinucci, D. and P. M. Robinson (1999), "Alternative Forms of Fractional Brownian Motion," *Journal of Statistical Planning and Inference*, 80, 111–122.

Mielniczuk, J. (1997), "Long and Short-Range Dependent Sums of Infinite-Order Moving Averages and Regression Estimation," *Acta Sci Math (Szeged)*, 63, 301–316.

Newey, W. and K. West (1987), "A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Perron, P. and S. Ng (1996), "Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties," *Reviews of Economic Studies*, 63, 435–463.

Phillips, P. C. B. (1987), "Time Series Regression with a Unit Root," *Econometrica*, 55, 277–301.

Phillips, P. C. B. and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 58(1), 165–193.

Phillips, P. C. B. and P. Perron (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75(2), 335–346.

Schwert, G. W. (1987), "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data," *Journal of Monetary Economics*, 20, 73–103.

Schwert, G. W. (1989), "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics*, 7, 147–159.

Sowell, F. B. (1990), "The Fractional Unit Root Distribution," *Econometrica*, 58, 2, 495–505.

Stock, J. H. (1990), "A Class of Tests for Integration and Cointegration," Unpublished manuscript, Kennedy School of Government, Harvard University.

Stout, W. F. (1974), *Almost Sure Convergence*, New York: Academic Press.

Tanaka, K. (1999), "The Nonstationary Fractional Unit Root," *Econometric Theory*, 15, 549–582.

Tsay, W. J. (2001), "Alternative Proof for the Consistency of the KPSS Tests Against Fractional Alternatives," *Journal of Social Sciences and Philosophy*, 13(4), 401–416.

Wang, Q., Y. Lin, and C. Gulati (2003), "Asymptotics for General Fractionally Integrated Processes with Applications to Unit Root Tests," *Econometric Theory*, 19, 143–164.

Xiao, Z. and P. C. B. Phillips (1998), "An ADF Coefficient Test for a Unit Root in ARMA Models of Unknown Order with Empirical Applications to the US Economy," *Econometrics Journal*, 1, 27–43.

分數整合與 Phillips-Perron 檢定

李慶男 *
國立中山大學經濟學研究所

謝富順
博士班研究生
國立台灣大學財務金融學系

關鍵詞: 單根、分數整合過程、檢定力
JEL **分類代號**: C12, C22

* 聯繫作者: 李慶男, 國立中山大學經濟學研究所, 高雄市 804 鼓山區西子灣蓮海路 70 號。 電話: (07) 525-2000 分機 5618; 傳真: (07) 525-5611: E-mail: lee_econ@mail.nsysu.edu.tw。 作者感謝鍾經樊教授、蔡文禎教授與兩位匿名審稿者精闢的建議與指正, 特別感謝責任編輯對於本文輔助定理 2 證明部份的指正與建議。

摘　　要

本文在非定態分數整合的過程下, 亦即 I(1+d) 的 $d \in (-0.5, 0.5)$ 下, 推導出 Phillips-Perron 單根檢定統計量的漸進分配。藉由使用 Newey-West 的長期變異數估計式, 我們證明出 Phillips-Perron 的 t 統計量和標準係數估計式, 在非定態但均數復歸的分數整合過程, 如 I$(1+d)$ 的 $d \in (-0.5, 0)$ 下是具一致性的。然而, 在非定態且非均數復歸的分數整合過程, 如I$(1+d)$ 的 $d \in (0, 0.5)$ 下, 只有從沒有截距項和沒有時間趨勢項迴歸下的 t 統計量是具一致性的。模擬的結果亦支持我們的發現, 即檢定統計量的檢定力, 在大樣本下將隨著 Newey-West 的長期變異數估計式所選取的落後期數增加而遞減。