**The Fractional Unit Root Distribution**

Fallaw Sowell

*Econometrica*, Vol. 58, No. 2 (Mar., 1990), 495-505.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199003%2958%3A2%3C495%3ATFURD%3E2.0.CO%3B2-F

*Econometrica* is currently published by The Econometric Society.

*Econometrica*, Vol. 58, No. 2 (March, 1990), 495–505

# THE FRACTIONAL UNIT ROOT DISTRIBUTION

BY FALLAW SOWELL[1]

Asymptotic distributions are derived for the ordinary least squares (OLS) estimate of a first order autoregression when the series is fractionally integrated of order $1+d$, for $-1/2 < d < 1/2$. The fractional unit root distribution is introduced to describe the limiting distribution. The unit root distribution ($d=0$) is seen to be an atypical member of this family because its density is nonzero over the entire real line. For $-1/2 < d < 0$ the fractional unit root distribution has nonpositive support, while if $0 < d < 1/2$ the fractional unit root distribution has nonnegative support. Any misspecification of the order of differencing leads to drastically different limiting distributions. Testing for unit roots is further complicated by the result that the $t$ statistic in this model only converges when $d=0$. Results are proven by means of functional limit theorems.

KEYWORDS: Unit root distribution, fractional differencing, functional limit theorem.

## 1. INTRODUCTION

THE NEED TO TEST economic theories which imply random walks has stimulated a large literature involving the unit root distribution (see Dickey and Fuller (1979, 1981), Evans and Savin (1981, 1984), Sargan and Bhargava (1983), Phillips (1987)). One facet of the unit root literature has concerned weakening the assumption of IID errors. In particular Phillips (1987) shows that the unit root distribution can be used to test for a random walk if the errors satisfy a strong mixing condition. Unfortunately, this condition may not be justified for some economic time series. For example, dependency greater than allowed in Phillips (1987), is permitted by fractionally integrated models which extend the ARIMA($p, d, q$) model to real values of $d$. Furthermore, studies that have looked for fractional integration (Granger and Joyeux (1981), Geweke and Porter-Hudak (1983)) have concluded that some economic time series possess fractional unit roots.

This paper generalizes the unit root distribution to fractionally integrated errors. It is shown that the limiting distributions of fractionally integrated series are radically different than for series integrated of order zero or one. The dissimilarities between the unit root distribution and other fractional unit root distributions underscore the importance of considering fractional models.

The approach used to obtain the limiting distributions is similar to that of Phillips (1987). Phillips uses functional limit theorems to obtain the unit root distribution when the underlying random variables are strong mixing. Fractionally integrated series have greater dependency than allowed in Phillips (1987) and a different functional limit theorem is required to obtain the limiting distributions.

[1] I thank John Geweke, Sastry Pantula, Peter Bloomfield, a co-editor, and two anonymous referees for helpful comments and suggestions. This is a generalization of Chapter 5 of my Ph.D. thesis, which was supported by the Alfred P. Sloan Foundation through a Doctoral Dissertation Fellowship.

Section 2 summarizes general properties of fractionally integrated time series and a related continuous stochastic process, fractional Brownian motion. In the third section, a new distribution is introduced, the fractional unit root distribution. It characterizes the limiting distribution of the OLS estimate of the parameter of an AR(1) model when the series is fractionally integrated. In the fourth section, the distribution is studied through simulations and directions for future research are discussed. Convergence in probability and in distribution will be denoted by $\to^p$ and $\Rightarrow$ respectively.

## 2. PRELIMINARY CONCEPTS

The fractional difference operator $(1-L)^d$ is defined by its Maclaurin series

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(-d+j)}{\Gamma(-d)\Gamma(j+1)} L^j$$

where

$$\Gamma(z) \equiv \begin{cases} \int_0^\infty s^{z-1} e^{-s}\, ds & \text{if } z > 0, \\ \infty & \text{if } z = 0. \end{cases}$$

If $z < 0$, $\Gamma(z)$ is defined in terms of the above expressions and the recurrence formula

$$z\Gamma(z) = \Gamma(z+1),$$

which holds for all values of $z$. Note that the recurrence formula and $\Gamma(0) = \infty$ imply that $\Gamma(z)$ has poles at the nonpositive integers.

If $x_t \sim \text{IID}(0, \sigma^2)$, then its partial sum

$$y_t = \sum_{t=1}^{t} x_t$$

will be called integrated of order one, denoted $y_t \sim I(1)$, because after applying the differencing operator $(1-L)^1$ the series is $\text{IID}(0, \sigma^2)$. Similarly, a series $\varepsilon_t$ will be called integrated of order $d$, denoted $\varepsilon_t \sim I(d)$, if $(1-L)^d \varepsilon_t = u_t \sim \text{IID}(0, \sigma_u^2)$. When $d$ is not an integer the series is said to be fractionally integrated.

A fractionally integrated series is stationary and ergodic for $-1/2 < d < 1/2$. The unusual characteristic of fractionally integrated series is the dependence between distant observations. This can be seen in the autocovariance function or the spectral density. If $\varepsilon_t \sim I(d)$, then its autocovariance function is

$$\gamma_\varepsilon(s) = E[\varepsilon_t \varepsilon_{t-s}] = \frac{\Gamma(1-2d)\Gamma(d+s)}{\Gamma(d)\Gamma(1-d)\Gamma(1-d+s)} \sigma_u^2.$$

The autocovariances have the same sign as $d$ for $s \geq 1$ and the autocovariance

function satisfies

$$\lim_{s\to\infty} \gamma_\varepsilon(s)/s^{2d-1} = \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)}\sigma_u^2.$$

The dependence between observations is noted in the hyperbolic decay which is slower than the geometric decay of stationary ARMA models. The long run dependence can also be seen in the spectral density which satisfies

$$\lim_{\lambda\to 0} f_\varepsilon(\lambda)/\lambda^{-2d} = \sigma_u^2.$$

As the frequency approaches zero, the spectral density either goes to zero, if $d<0$, or infinity, if $d>0$. One appeal of fractionally integrated time series is their ability to capture a variety of long run behaviors with a single parameter.

A property of a fractionally integrated series is the dependence on $d$ of the growth of the partial sums' variance. This implies that the distribution theory developed in Phillips (1987) is not general enough to deal with fractionally integrated series. In Phillips (1987), Assumption 2.1(c) requires the variance of the partial sums to grow at a linear rate. The growth rates for fractionally integrated series are presented in the following theorem. All proofs are collected in the Appendix.

THEOREM 1: *If* $\varepsilon_t \sim I(d)$, $-1/2 < d < 1/2$, *and* $S_N = \sum_{t=1}^N \varepsilon_t$, *then*

$$\sigma_N^2 = \mathrm{Var}(S_N) = \frac{\sigma_u^2\Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)} \times\left[\frac{\Gamma(1+d+N)}{\Gamma(-d+N)} - \frac{\Gamma(1+d)}{\Gamma(-d)}\right]$$

*and*

$$\lim_{N\to\infty} \mathrm{Var}(S_N)/N^{1+2d} = \frac{\sigma_u^2\Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)}.$$

Theorem 1 shows $\mathrm{Var}(S_N) = O(N^{1+2d})$ which is an important characteristic of processes that are fractionally integrated. The variance of the partial sums of IID variables ($d=0$) grows at the linear rate $N$. Each random shock is uncorrelated with the others and only adds its own variance to the variance of the partial sum. When $-1/2 < d < 0$, each shock is negatively correlated with the others. Therefore, the variance of the partial sum grows less than the variance of the individual shock. When $d$ is near $-1/2$ the negative covariances almost totally offsets the added variance of the shock. For $0 < d < 1/2$ the shocks are positively correlated and the variance of the sum grows faster than the variance of a single shock. When $d$ is near $1/2$, the growth of the variance of the partial sum is almost quadratic. As shown below, this growth rate of the variance of partial sums is needed to apply functional limit theorems.

Fractional integration is a characteristic of discrete stochastic processes. Limiting functions of these discrete series, however, will be described by a continuous stochastic process, called fractional Brownian motion. Fractional Brownian motion was first introduced and studied in Mandelbrot and Van Ness (1968) and is defined for $d \in (-1/2, 1/2)$ by the stochastic integral

$$W_d(t) \equiv \frac{1}{\Gamma(d+1)} \int_0^t (t-x)^d \, dW(x).$$

When $d = 0$ this reduces the Brownian motion. Fractional Brownian motion is a stationary continuous Gaussian stochastic process, with mean zero and covariance function

$$E|W_d(t) - W_d(s)|^2 = |t-s|^{1+2d}.$$

The interested reader should see Jonas (1983) for a thorough presentation of the properties of fractional Brownian motion.

It is often possible to characterize limiting distributions of discrete stochastic processes as functions of continuous stochastic processes by applying functional limit theorems. The functional limit theorems that apply to fractionally integrated time series are presented in Davydov (1970) and Taqqu (1975) and the continuous stochastic process is fractional Brownian motion. The applicability of these functional limit theorems to fractionally integrated time series is noted in the following theorem.

THEOREM 2: *If* $\varepsilon_t \sim I(d)$ *for* $-1/2 < d < 1/2$ *and if* $(1-L)^d \varepsilon_t = u_t$ *have zero mean, are IID, and* $E|u_t|^r < \infty$ *for* $r \geq \max[4, -8d/(1+2d)]$, *then* $Z_N(t) = \sigma_N^{-1} S_{[Nt]} \Rightarrow W_d(t)$.

This is an invariance principle, i.e., the limiting distribution does not depend on the parameters (variance, skewness, kurtosis,...) of the error process. For fixed $d$, the *same limiting distribution* holds for all $u_t$ processes that satisfy the assumptions.

## 3. UNIT ROOT STATISTICS WITH FRACTIONALLY INTEGRATED SERIES

The limiting distribution of the OLS estimate of an AR(1) model for a random walk can be written as the ratio of functions of Brownian motions. Given the model

$$x_t = x_{t-1} + \varepsilon_t \quad \text{for} \quad t = 1, 2, \ldots, N,$$
$$x_0 = 0,$$

where $\varepsilon_t \sim \text{IID}(0, \sigma_\varepsilon^2 < \infty)$, the least squares slope estimate,

$$\hat{\beta} = \frac{\sum_{t=1}^{N} x_{t-1} x_t}{\sum_{t=1}^{N} x_{t-1}^2},$$

has the following asymptotic distribution

$$N(\hat{\beta}-1) \Rightarrow \frac{(1/2)\{W(1)^2-1\}}{\int_0^1 W(t)^2\,dt},$$

(see Phillips (1987) and the references cited therein). Because the error process is $I(0)$, this result can be considered the answer to the special case $d=0$ of the general question "What is the limiting distribution of $\hat{\beta}$ if $\varepsilon_t \sim I(d)$ for $-1/2 < d < 1/2$?" The answer to the general question is presented in the following theorem.

THEOREM 3:[2] *Given the model*

$$x_t = x_{t-1} + \varepsilon_t \quad \text{for} \quad t = 1,2,\ldots,N,$$
$$x_0 = 0,$$
$$\varepsilon_t \sim I(d) \qquad -1/2 < d < 1/2,$$

*with* $(1-L)^d \varepsilon_t = u_t$ *satisfying the assumptions of Theorem* 2 *and*

$$A_d = \frac{\frac{1}{2}x_N^2}{\sum_{t=1}^{N} x_{t-1}^2}, \qquad B_d = \frac{-\frac{1}{2}\sum_{t=1}^{N}\varepsilon_t^2}{\sum_{t=1}^{N} x_{t-1}^2},$$

*then* $(\hat{\beta}-1) = A_d + B_d$ *and*

$$NA_d \Rightarrow \frac{\frac{1}{2}[W_d(1)]^2}{\int_0^1 [W_d(s)]^2\,ds} \quad \text{and} \quad N^{1+2d}B_d \Rightarrow -\frac{\left[\frac{1}{2}+d\right]\frac{\Gamma(1+d)}{\Gamma(1-d)}}{\int_0^1 [W_d(s)]^2\,ds}.$$

The limiting distribution of $N^{\min[1,1+2d]}(\hat{\beta}-1)$ is called a fractional unit root distribution. This limiting distribution is achieved by normalizing, by $N$ if $d \geqslant 0$ and by $N^{1+2d}$ if $d \leqslant 0$. If $d>0$, then $N < N^{1+2d}$ and $NB_d$ converges in probability to zero leaving asymptotically the distribution of $NA_d$. Conversely when $d<0$, the limiting distribution is that of $N^{1+2d}B_d$ because $N > N^{1+2d}$ and $N^{1+2d}A_d$ converges in probability to zero. Because $A_d$ is a nonnegative random variable, if $d>0$ the fractional unit root distribution has nonnegative support. Similarly, if $d<0$ the support is nonpositive, because $B_d$ is a nonpositive random variable. If $d=0$, $NA_d$ and $N^{1+2d}B_d$ jointly converge and the support of the fractional unit root distribution is the entire real line.

[2] I thank Sastry Pantula for simplifying several steps of the proof and alleviating some unnecessary assumptions of an earlier version.

An immediate corollary is that the least squares estimate is consistent. The rate of convergence depends on the order of integration in a surprising way. It is well known that if $x_t$ is $I(0)$, the OLS estimate $\hat{\beta}$ converges at the rate $N^{1/2}$, and if $x_t$ is $I(1)$, convergence is at the rate $N$. This suggests that the rate of convergence increases with the order of integration. However, this is not the case. If a series is $I(1+d)$ for $0 \leqslant d < 1/2$, $\hat{\beta}$ converges at the rate $N$, the same for all $d$. If $-1/2 < d \leqslant 0$, $\hat{\beta}$ converges at the rate $N^{1+2d}$. So for $-1/2 < d < -1/4$ the rate of convergence is slower than $N^{1/2}$.

Fractional unit root distributions pose problems when testing for unit roots. The unit root distribution is not robust to any misspecification in the order of integration. An alternative approach to testing for a unit root is to consider the $t$ statistic, which for this model is

$$t = \frac{\sum_{t=1}^{N} x_{t-1}\varepsilon_t}{\left(\sum_{t=1}^{N} x_{t-1}^2\right)^{1/2} s},$$

where

$$s^2 = N^{-1}\sum_{t=1}^{N}\left(x_t - \hat{\beta}x_{t-1}\right)^2.$$

Unfortunately, a similar problem occurs with this statistic. As the following theorem notes, the statistic only converges when $d = 0$.

THEOREM 4: *Given the model*

$$x_t = x_{t-1} + \varepsilon_t \quad \textit{for} \quad t = 1,2,\ldots,N,$$
$$x_0 = 0,$$
$$\varepsilon_t \sim I(d), \qquad \textit{for} \quad -1/2 < d < 1/2,$$

*and* $(1-L)^d\varepsilon_t = u_t$ *satisfy the assumptions of Theorem* 2, *then*

$$t \to^p -\infty \quad \textit{if} \quad d < 0,$$
$$t \to^p \infty \qquad \textit{if} \quad d > 0.$$

## 4. DISCUSSION

The rate at which $N^{\min[1,1+2d]}(\hat{\beta}-1)$ converges to its limiting distribution is slow for values of $d$ near zero. The statistic is composed of the two random variables $N^{\max[1,1+2d]}A_d$ and $N^{\max[1,1+2d]}B_d$. Depending on the sign of $d$, one of these converges to zero. The random variable that converges to zero ($NB_d$ if $d > 0$ and $N^{1+2d}A_d$ if $d < 0$) does so at the rate $N^{-|2d|}$ which can be quite slow for $d$ near zero. This implies that $(\hat{\beta}-1)$ also converges very slowly to its limiting distribution when $d$ is near zero. This does not mean that the misspeci-

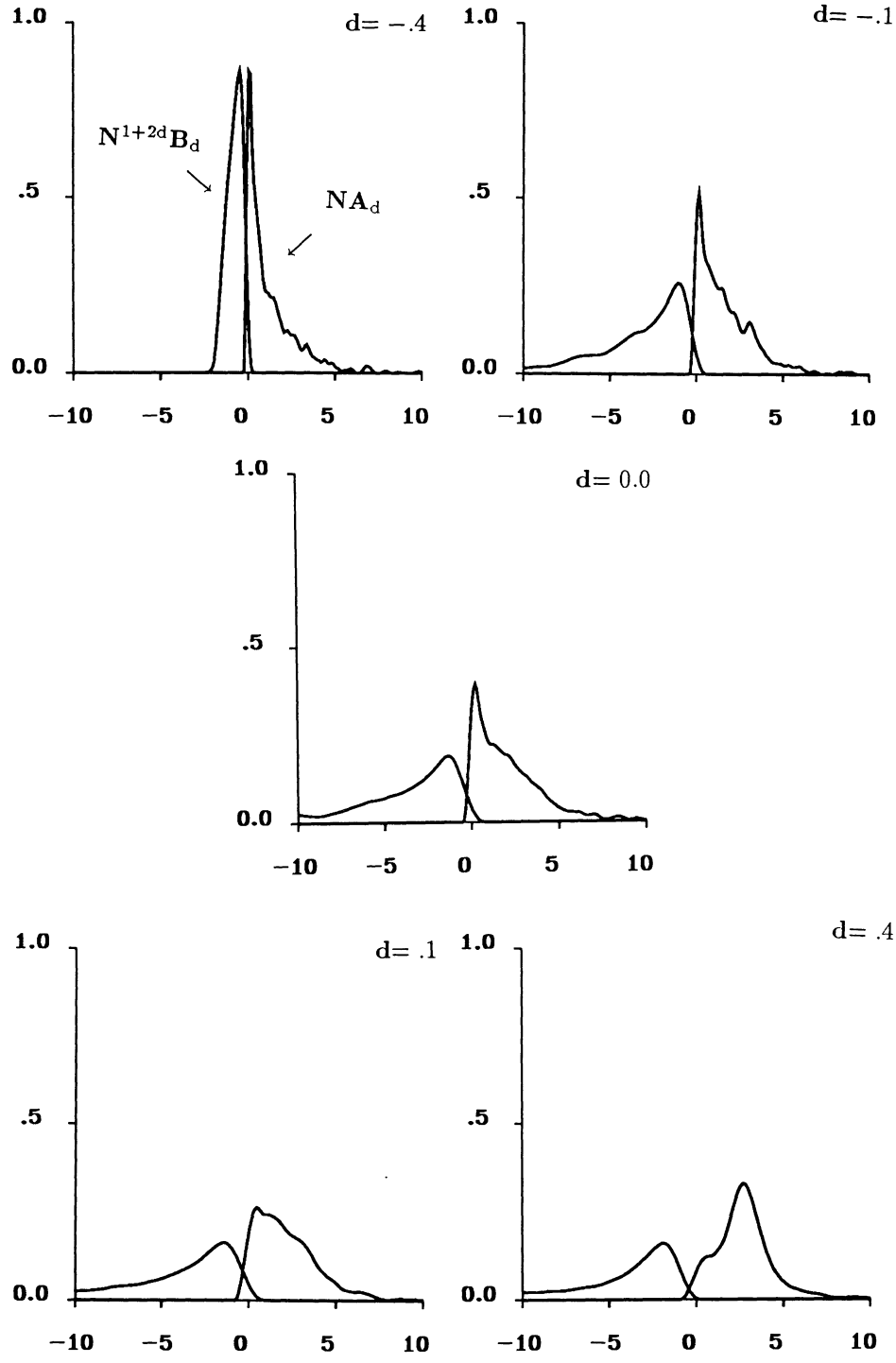


FIGURE 1.—Kernel estimates of the densities of $NA_d$ and $N^{1+2d}B_d$ using 1000 samples with $N = 900$.

fication is irrelevant for small values of $d$. Rather it implies that asymptotic theory may be inapplicable.

To illustrate the fractional unit root distribution, 1000 samples of $I(1+d)$ series were generated and estimates of $A_d$ and $B_d$ were calculated. Samples of 900 observations were generated for five different values of $d$: $-0.4$, $-0.1, 0.0, 0.1, 0.4$. The densities of $NA_d$ and $N^{1+2d}B_d$ were estimated by the kernel estimator

$$\hat{f}(x) = \frac{1}{(1000)h} \sum_{j=1}^{1000} \Psi\left(\frac{x-\hat{\theta}_j}{h}\right),$$

where $\hat{\theta}_j$ are the estimated values of $NA_d$ or $N^{1+2d}B_d$ for the 1000 samples and $\Psi(y)$ is the density defined by $(15/16)\ ((1-y^2)^2)$ for $-1 < y < 1$ and zero elsewhere. The value of $h$ was chosen to minimize the integrated mean square error (see Tapia and Thompson (1978, p. 67)). The estimated densities are presented in Figure 1.

A striking feature of Figure 1 is the general agreement in the estimated densities for different values of $d$. The densities of $N^{1+2d}B_d$ are all similar except as $d$ approaches $-1/2$ the density appears to be converging to unit mass at zero. This is the case because the numerator of $N^{1+2d}B_d$ equals zero at $d = -1/2$. The similarity in the densities of $NA_d$ is due in part to its numerator which asymptotically has a chi-square distribution with one degree of freedom and hence is asymptotically independent of $d$.

The similarity between these densities with different values of $d$ suggest low power for tests of the order of integration based on the $N^{1+2d}B_d$ and $NA_d$. Alternative approaches (i.e., statistics) are required to test for unit roots when the data may be fractionally integrated. Tests that would be appropriate in this situation would involve estimating the differencing parameter $d$. One approach is the unconditional maximum likelihood estimation of a univariate ARIMA$(p, d, q)$ model. The univariate maximum likelihood procedure is a special case of the general results in Sowell (1988).

*Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, PA 15213-3890, U.S.A.*

*Manuscript received July, 1987; final revision received March, 1989.*

## APPENDIX

The proof of Theorem 1 is simplified by the following.

LEMMA: *For $N = 1, 2, \ldots$ with $a$ and $b \neq -1, -2, -3, \ldots$,*

$$\sum_{k=1}^{N} \frac{\Gamma(a+k)}{\Gamma(b+k)} = \frac{1}{1+a-b}\left[\frac{\Gamma(1+a+N)}{\Gamma(b+N)} - \frac{\Gamma(1+a)}{\Gamma(b)}\right];$$

*when $b = 0$ this reduces to $(1/(1+a))[\Gamma(1+a+N)/\Gamma(N)]$.*

PROOF: The lemma follows by induction and the relation $z\Gamma(z) = \Gamma(z+1)$. *Q.E.D.*

PROOF OF THEOREM 1: Write $\gamma_\varepsilon(s) = K\Gamma(d+s)/\Gamma(1-d+s)$ where $K \equiv \sigma_u^2\Gamma(1-2d)/(\Gamma(d)\Gamma(1-d))$. The variance of the partial sum can be written

$$\begin{aligned}
\operatorname{Var}\left[\sum_{j=1}^{N}\varepsilon_j\right] &= -N\gamma_\varepsilon(0) + 2\sum_{i=0}^{N-1}(N-i)\gamma_\varepsilon(i) \\
&= -\sum_{k=1}^{N}\gamma_\varepsilon(0) + 2\sum_{i=0}^{N-1}\sum_{k=1}^{N-i}\gamma_\varepsilon(i) \\
&= -\sum_{k=1}^{N}\gamma_\varepsilon(0) + 2\sum_{k=1}^{N}\sum_{i=0}^{N-k}\gamma_\varepsilon(i) \\
&= K\sum_{k=1}^{N}\left[-\frac{\Gamma(d)}{\Gamma(1-d)} + 2\sum_{i=0}^{N-k}\frac{\Gamma(d+i)}{\Gamma(1-d+i)}\right] \\
&= K\sum_{k=1}^{N}\left[-\frac{\Gamma(d)}{\Gamma(1-d)} + 2\sum_{i=1}^{N-k+1}\frac{\Gamma(d-1+i)}{\Gamma(-d+i)}\right].
\end{aligned}$$

Using the above lemma twice this reduces to

$$\frac{K}{d(1+2d)}\left[\frac{\Gamma(1+d+N)}{\Gamma(-d+N)} - \frac{\Gamma(1+d)}{\Gamma(-d)}\right].$$

Using the definition of $K$, the fact that $z\Gamma(z) = \Gamma(1+z)$, and the result $\Gamma(a+N)/\Gamma(b+N) \sim N^{a-b}$ as $N \to \infty$, the theorem is proven. *Q.E.D.*

PROOF OF THEOREM 2: Apply Theorem 2 from Davydov (1970) with Theorem 1 above. Note Davydov (1970) uses $\gamma$ to denote $1+2d$. *Q.E.D.*

PROOF OF THEOREM 3: If the variance of $x_N$ is denoted by $\sigma_N^2$, then

$$(\hat{\beta}-1) = \frac{\dfrac{1}{N\sigma_N^2}\displaystyle\sum_{t=1}^{N}x_{t-1}\varepsilon_t}{\dfrac{1}{N\sigma_N^2}\displaystyle\sum_{t=1}^{N}x_{t-1}^2}.$$

First consider the denominator,

$$\begin{aligned}
\frac{1}{N\sigma_N^2}\sum_{t=1}^{N}x_t^2 &= \frac{1}{N}\sum_{t=1}^{N}\left[\sigma_N^{-1}x_t\right]^2 \\
&= \sum_{t=1}^{N}\int_{(t-1)/N}^{t/N}\left[Z_N(s)\right]^2 ds \\
&= \int_0^1\left[Z_N(s)\right]^2 ds \Rightarrow \int_0^1\left[W_d(s)\right]^2 ds
\end{aligned}$$

which holds for $-1/2 < d < 1/2$. The last result follows from the Continuous Mapping Theorem and Theorem 2.

Now consider the numerator

$$\frac{1}{N\sigma_N^2}\sum_{t=1}^{N} x_{t-1}\varepsilon_t = \frac{1}{2N\sigma_N^2}\sum_{t=1}^{N}\left[x_t^2 - x_{t-1}^2\right] - \frac{1}{2N\sigma_N^2}\sum_{t=1}^{N}\varepsilon_t^2$$

$$= \frac{1}{2N}\left[\sigma_N^{-1}x_N\right]^2 - \frac{1}{2N\sigma_N^2}\sum_{t=1}^{N}\varepsilon_t^2$$

$$= \frac{1}{2N}\left[Z_N(1)\right]^2 - \frac{1}{2N\sigma_N^2}\sum_{t=1}^{N}\varepsilon_t^2.$$

The first term when multiplied by $N$ converges in distribution to $(1/2)[W_d(1)]^2$ for $-1/2 < d < 1/2$ again because of the Continuous Mapping Theorem and Theorem 2. This proves the limiting distribution of the $A_d$ term. The limiting distribution of the second term follows by first noting

$$\left[\frac{1}{N}\sum_{t=1}^{N}\varepsilon_t^2\right] \to^p E\varepsilon_t^2 = \gamma_\varepsilon(0) = \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(1-d)}\,\sigma_u^2$$

by the Ergodic Theorem. Finally, using Theorem 1,

$$N^{1+2d}/\sigma_N^2 \to \frac{(1+2d)\Gamma(1-d)\Gamma(1+d)}{\Gamma(1-2d)}\sigma_u^{-2}$$

which shows the limiting distribution of $B_d$ and proves the theorem. *Q.E.D.*

PROOF OF THEOREM 4: First show that $s^2 \to^p \gamma_\varepsilon(0)$:

$$s^2 = N^{-1}\sum_{t=1}^{N}\left(x_t - x_{t-1} - (\hat{\beta}-1)x_{t-1}\right)^2$$

$$= N^{-1}\sum_{t=1}^{N}\left(\varepsilon_t - (\hat{\beta}-1)x_{t-1}\right)^2$$

$$= N^{-1}\sum_{t=1}^{N}\varepsilon_t^2 - 2(\hat{\beta}-1)N^{-1}\sum_{t=1}^{N}x_{t-1}\varepsilon_t + (\hat{\beta}-1)^2N^{-1}\sum_{t=1}^{N}x_{t-1}^2.$$

As in the proof of Theorem 3, the first term converges in probability to $\gamma_\varepsilon(0)$. The other two terms both converge to zero using the following reasoning for each:

$$(\hat{\beta}-1)N^{-1}\sum_{t=1}^{N}x_{t-1} = \sigma_N(\hat{\beta}-1)\frac{1}{N}\sum_{t=1}^{N}\sigma_N^{-1}x_{t-1} \to^p 0$$

because

$$\frac{1}{N}\sum_{t=1}^{N}\sigma_N^{-1}x_{t-1} \Rightarrow \int_0^1 W_d(s)\,ds$$

as in the proof of Theorem 3 and

$$\sigma_N(\hat{\beta}-1) \to^p 0$$

by Theorem 1 and Theorem 3. Similar reasoning shows

$$(\hat{\beta}-1)^2N^{-1}\sum_{t=1}^{N}x_{t-1}^2 \to^p 0.$$

Now, the $t$ statistic is

$$\frac{\sum_{t=1}^{N}x_{t-1}\varepsilon_t}{\left(\sum_{t=1}^{N}x_{t-1}^2\right)^{1/2}s} = \frac{\frac{1}{\sqrt{N}\sigma_N}\sum_{t=1}^{N}x_{t-1}\varepsilon_t}{\frac{1}{\sqrt{N}\sigma_N}\left(\sum_{t=1}^{N}x_{t-1}^2\right)^{1/2}s}.$$

The same reasoning used in the proof of Theorem 3 shows that the denominator converges:

$$\frac{1}{\sqrt{N}\,\sigma_N}\left(\sum_{t=1}^{N} x_{t-1}^2\right)^{1/2} = \left(\frac{1}{N\sigma_N^2}\sum_{t=1}^{N} x_{t-1}^2\right)^{1/2} \Rightarrow \left(\int_0^1 W_d(s)^2\,ds\right)^{1/2}.$$

Finally, consider the numerator. From the proof of Theorem 3,

$$\sum_{t=1}^{N} x_{t-1}\varepsilon_t = x_N^2 - \sum_{t=1}^{N} \varepsilon_t^2;$$

hence, the numerator can be written

$$\frac{\sigma_N}{\sqrt{N}}\left[\sigma_N^{-1}x_N\right]^2 - \frac{\sqrt{N}}{\sigma_N}\left[N^{-1}\sum_{t=1}^{N}\varepsilon_t^2\right]$$

and, because $\sigma_N$ is $O(N^{1/2+d})$, if $d > 0$, $(\sigma_N/\sqrt{N}) \to \infty$ and $(\sqrt{N}/\sigma_N) \to 0$, while if $d < 0$, $(\sigma_N/\sqrt{N}) \to 0$ and $(\sqrt{N}/\sigma_N) \to \infty$ which proves the theorem. *Q.E.D.*

## REFERENCES


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

DICKEY, DAVID A., AND WAYNE A. FULLER (1979): "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

——— (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057–1072.

EVANS, G. B. A., AND N. E. SAVIN (1981): "Testing for Unit Roots: 1," *Econometrica*, 49, 753–779.

——— (1984): "Testing for Unit Roots: 2," *Econometrica*, 52, 1241–1269.

GEWEKE, JOHN F., AND SUSAN PORTER-HUDAK (1982): "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

GRANGER, C. W. J., AND ROSELYNE JOYEUX (1980): "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, 1, 1529.

JONAS, ANDREW BARRANDE (1983): *Persistent Memory Random Processes*. Ph.D. Thesis, Department of Statistics, Harvard University.

MANDELBROT, BENOIT B., AND JOHN W. VAN NESS (1968): "Fractional Brownian Motions, Fractional Brownian Noises and Applications," *SIAM Review*, 10, 422–437.

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

SARGAN, J. D., AND A. BHARGAVA (1983): "Testing Residuals from Least Squares Regression for Being Generated By a Gaussian Random Walk," *Econometrica* 51, 153–174.

SOWELL, FALLAW B. (1988): "Maximum Likelihood Estimation of Fractionally Integrated Time Series Models," GSIA, Carnegie Mellon University, Working Paper.

TAPIA, RICHARD A., AND JAMES R. THOMPSON (1978): *Nonparametric Probability Density Estimation*. Baltimore: The Johns Hopkins University Press.

TAQQU, MURAD S. (1975): "Weak Convergence to Fractional Brownian Motion and to the Rosenblatt Process," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 31, 287–302.