FRACTIONAL INTEGRATION
VERSUS TREND STATIONARY
IN TIME SERIES ANALYSIS

Francesc Marmol

97-02
Abstract

The objective of this paper is to study the effects of spurious detrending of a nonstationary fractionally integrated process (NFI(d), d ≥ 5) on the performance of the traditional least squares estimators and tests. We extend previous work on the subject undertaken by Durlauf and Phillips (1988) which considered only the leading difference stationary (d = 1) case. Moreover, we also consider the possibility of a double misspecification both in the stochastic and in the nonstochastic trends. Standard t-Student tests are shown to diverge in distribution invalidating any inference concerning the presence of time trends. On the other hand, we prove that, even under this double misspecification, the Durbin-Watson statistic remains to be a useful misspecification test.

Keywords:
Spurious detrending, nonstationary fractionally integrated processes, specification tests, misspecification.

*Marmol, Department of Statistics and Econometrics, Universidad Carlos III de Madrid; I am indebted to Niels Haldrup for helpful comments and discussions and to Juan C. Reboredo for computer assistance. I also thanks participants in the various seminars and meetings at which previous versions of this paper were presented for their remarks. Financial support from the European Commission through the Training and Mobility of Researchers Program is gratefully acknowledged. The usual disclaimers apply.
Fractional Integration Versus Trend Stationarity in Time Series Analysis

Francesc Marmol

GREMAQ-IDEI, Université des Sciences Sociales, Place Anatole France, 31042 Toulouse Cedex, France

and

Departamento de Estadística y Econometría, Universidad Carlos III, 28903, Getafe, Madrid, Spain.

The objective of this paper is to study the effects of spurious detrending of a nonstationary fractionally integrated process \( NFI(d), d \geq 0.5 \) on the performance of the traditional least squares estimators and tests. We extend previous work on the subject undertaken by Durlauf and Phillips (1988) which considered only the leading difference stationary \( (d = 1) \) case. Moreover, we also consider the possibility of a double misspecification both in the stochastic and in the nonstochastic trends. Standard t-Student tests are shown to diverge in distribution invalidating any inference concerning the presence of time trends. On the other hand, we prove that, even under this double misspecification, the Durbin-Watson statistic remains to be a useful misspecification test.

KEYWORDS: Spurious detrending, nonstationary fractionally integrated processes, specification tests, misspecification.

\[1\text{I am indebted to Niels Haldrup for helpful comments and discussions and to Juan C. Reboredo for computer assistance. I also thanks participants in the various seminars and meetings at which previous versions of this paper were presented for their remarks. Financial support from the European Commission through the Training and Mobility of Researchers Program is gratefully acknowledged. The usual disclaimers apply.}\]
1 INTRODUCTION

When a given time series \( Y_t \) becomes stationary after differencing \( d \) times, and the degree of differentiation or memory parameter, \( d \), is a real number, then the series is said to be fractionally integrated of order \( d \), denoted \( Y \sim FI(d) \). These processes have received an increasing attention because of their ability to provide a natural and flexible characterization of the nonstationary and persistent characteristics of economic time series.

A fractionally integrated process is both stationary and invertible if \( d \in (-\frac{1}{2}, \frac{1}{2}) \) and (weakly) nonstationary if \( d \geq \frac{1}{2} \). In spite of being nonstationary, the process is mean-reverting with transitory memory, i.e., with any random shock having only a temporary influence on the series, if \( d < 1 \), in contrast with the case when \( d \geq 1 \), where the process is both nonstationary and not mean-reverting with permanent memory, i.e., with any random shock having now a permanent effect on the present and future path of the series. On the other hand, a stationary fractionally integrated process has short-memory, with autocorrelations decaying at an exponential rate, if \( d < 0 \), whereas it has long-memory, with autocorrelations that die out at the slower hyperbolic rate, if \( 0 < d < 1/2 \) and as such may be expected to be useful in modelling long-term persistence. When \( d = 0 \), the process is white noise, with zero correlations and constant spectral density.

The implications of the fractionally integrated processes in economic data are, at least potentially, profound. They allow for more parsimonious models. Moreover, this class of processes is naturally introduced when we consider the aggregation of heterogeneous time series (Granger, 1980, Gonçalves and Gourieroux, 1987). Lastly, by allowing a rich range of spectral behavior near the origin, they can provide superior approximations to the Wold representations of many economic time series (Granger, 1966).

Hence, it seems quite reasonable to assume that the macroeconomic time series achieve stationarity after applying a fractional filter. However, it is not by large an easy task to discriminate whether a real economic time series is trend stationary or fractionally integrated. Empirically, macroeconomic variables appear to be fractionally integrated in the post-Second World War quarterly data, whilst as for the historical data covering
eighty or more years, it is difficult to offer unequivocal judgement as to whether many real economic variables are fractionally integrated or trend stationary. See Baillie (1996) for a recent survey and review of the major econometric work on fractionally integrated processes and their applications in economics and finance.

As an illustration of this difficulty, Chambers (1996) used fractional models to test trend and difference \((d = 1)\) stationarity in the logarithms of five U.K. macroeconomic time series, covering from 1955(2) to 1992(2), i.e., a total of 148 observations. The author concluded that three of these series (GNP, consumption and investment) were found to be consistent with the hypothesis of difference stationarity, while the remaining two series (exports and imports), however, rejected both trend and difference stationarity, suggesting, therefore, that these series were consistent with a fractional model.

Consequently, the empirical and theoretical debate of whether an economic series are best described as being trend stationary, difference stationary or fractionally integrated continues unresolved. The purpose of this paper is to analyze the effects of spuriously detrending a nonstationary fractionally integrated process (henceforth denoted \(NFJ\)). In doing this, the asymptotic distributions of Durlauf and Phillips (1988), which regressed a difference stationary process onto a constant and a linear time trend, are embedded in our results.

Moreover, we proof that the Durbin-Watson statistic is a valid diagnostic against misspecification even in the case where we ignore not only the stochastic trends but also misspecify the non-deterministic component of the underlying series of interest. Therefore, the results of this paper continue strongly supporting the importance of combining hypothesis testing with specification analysis as a powerful method of exposing spurious relationships.

The paper is organized as follows. In Section 2 we derive the asymptotic distributions of the traditional least squares statistics in a time trend regression when in fact the time series is in reality a nonstationary fractionally integrated process. Section 3 extends these results to the case of nonstationary fractionally integrated process with drift, whereas in
Section 4 we provide some experimental evidence on the power of the Durbin-Watson statistic in finite samples. Finally, some concluding remarks are collected in Section 5. Proofs are given in the Appendix.

With respect to the notation employed, the symbols "\( \Rightarrow \)", "\( \rightarrow_p \)" and "\( \equiv \)" denote weak convergence, convergence in probability and equality in distribution, respectively, while \([\cdot]\) denotes "integer part". Stochastic processes such that \( y_r(r) \) with \( r \in [0, 1] \) are frequently written as \( y_\infty \). Similarly, we write integrals with respect to Lebesgue measure such as \( \int y_\infty(r)dr \) as \( \int y_\infty \). The symbol \( \sum_{t=1}^T \) is denoted simply as \( \sum \) and \( \Gamma(\cdot) \) denotes the gamma function. Finally, all limits given in the paper are as the sample size \( T \to \infty \) unless otherwise stated.

### 2 TIME TRENDS VS. NFI PROCESSES

Let us initially consider the analysis of the following least squares regression:

\[
y_t = \hat{\alpha} + \hat{\beta} t + \hat{u}_t, \quad t = 1, 2, \ldots, T,
\]

where \( y_t \) is assumed to be a stationary time trend process.

In 1988, Durlauf and Phillips considered the estimation of this model when in fact the true data generating process (DGP) for \( y_t \) is a difference stationary process, \( \Delta y_t = \varepsilon_t \), with the \( \{\varepsilon_t\} \) sequence assumed to satisfy a functional central limit theorem (FCLT) of the type discusses in Phillips (1987), allowing for weak dependence and some heterogeneity over time.

Under this DGP, Durlauf and Phillips (1988) found that the estimated time trend coefficient in (1) was consistent, \( \hat{\beta} = O_p(T^{-1/2}) \), converging to the true structural coefficient of zero and that the constant term \( \hat{\alpha} \) diverged in distribution, \( \hat{\alpha} = O_p(T^{1/2}) \) as the significance \( (H_0: \alpha = 0, \beta = 0) \) t-Student tests, \( t_\alpha = O_p(T^{1/2}) \), \( t_\beta = O_p(T^{1/2}) \). In the same manner, the coefficient of determination \( R^2 \) was found to have a nondegenerate limiting distribution, whilst the Durbin-Watson (DW) statistic converged in probability to zero at a super consistent \( O_p(T^{-1}) \) rate, indicating the presence of some kind of misspecification in the time regression (1) proposed.
In this section, we will extend Durlauf and Phillips' findings from the assumption that the true DGP for \( y_t \) is a difference stationary process to the following generating mechanism:

**DGP-A:**

\[
\Delta^d y_t = \varepsilon_t, \quad d \geq 1/2, \quad \text{where } \varepsilon_t \sim i.i.d(0,\sigma^2), \quad E|\varepsilon_t|^q < \infty, \quad q \geq \max\{d-1/2, 2\}.
\]

Throughout this paper, we shall assume, without any loss of generality, that \( \varepsilon_t = 0 \) for \( t \leq 0 \). In order to derive the behavior of the OLS statistics in model (1), we need the following result:

**LEMMA 1:** Under DGP-A,

\[
(2) \quad T^{1/2-d} y_{[T]} \Rightarrow \frac{\sigma}{\Gamma(d)} \int_0^T (r-s)^{d-1} dW(s) \quad \Rightarrow y_\alpha^*(r), \text{ say},
\]

\[
(3) \quad T^{-1/2-d} \sum_{t=1}^T y_t \Rightarrow \int_0^t y_\alpha^*(r) dr,
\]

\[
(4) \quad T^{-3/2-d} \sum_{t=1}^T y_t^2 \Rightarrow \int_0^t y_\alpha^*(r) dr,
\]

\[
(5) \quad T^{-2d} \sum_{t=1}^T y_t^2 \Rightarrow \int_0^t y_\alpha^*(r) dr,
\]

where \( W(r) \sim N(0,r) \) is a standard Brownian motion associated with the \( \{\varepsilon_t\} \) sequence.

Result (2) is a FCLT for \( NF(\alpha) \) processes due to Akonom and Gourieroux (1988) and results (3)-(5) follow in a direct manner from (2) by applying the continuous mapping theorem (CMT). See also Marmol (1995).
THEOREM 1: Under DGP-A,

\[ T^{\frac{4}{3} - d} \hat{\alpha} \Rightarrow -6 \int_0^1 (r - 2/3) y_w(r) dr, \]

\[ T^{\frac{3}{2} - d} \hat{\beta} \Rightarrow 12 \int_0^1 (r - 1/2) y_w(r) dr, \]

\[ R^2 \Rightarrow \frac{12 \left[ \int (r - 1/2) y_w \right]^2}{\int y_w^2 - \left( \int y_w \right)^2}, \]

\[ T^{-\frac{1}{2}} \tilde{t}_\beta \Rightarrow \frac{12 \left[ \int y_w - 0.5 \int y_w \right]^2}{\int y_w^2 - \left( \int y_w \right)^2 - 12 \left[ \int (r - 1/2) y_w \right]^2} \]

\[ T^{-\frac{1}{2}} \tilde{t}_\alpha \Rightarrow \frac{9 \left[ \int (r - 2/3) y_w \right]^2}{\int y_w^2 - \left( \int y_w \right)^2 - 12 \left[ \int (r - 1/2) y_w \right]^2} \]

\[ T^{2d-1} Dw \Rightarrow \frac{\sigma^2 \Gamma(3 - 2d)}{\Gamma^2(2 - d) \left( \int y_w^2 - \left( \int y_w \right)^2 - 12 \left[ \int (r - 1/2) y_w \right]^2 \right)^{1/2}}. \]

if \( d \in \left[ \frac{1}{2}, \frac{3}{2} \right) \), and

\[ T^2 Dw \Rightarrow \frac{\int (\Delta y_w)^2 + 144 \left[ \int (r - 1/2) y_w \right]^2 - 24 \left( r - 1/2 \right) y_w \int \Delta y_w}{\int y_w^2 - \left( \int y_w \right)^2 - 12 \left[ \int (r - 1/2) y_w \right]^2}, \]

if \( d \geq 3/2 \), where \( \Delta y_w = \frac{\sigma}{\Gamma(d - 1)} \int_0^r (r - s)^{d-3} dW(s) \sim FI(d-1) \).

Thus, when the true DGP is assumed to be a \( NFI(d) \) process, we have that the constant term \( \hat{\alpha} \) in regression (1) diverges in distribution with a rate of convergence \( O_p \left( T^{d-1/2} \right) \) that depends on the memory parameter \( d \), except in the particular \( d = \frac{1}{2} \) case,
where \( \hat{\alpha} \) has a nondegenerate limiting distribution. With respect to the estimated time trend coefficient \( \hat{\beta} \), from expression (7) we can see that it has a well-defined limiting distribution upon suitable standardization given by \( T^{3/2-d} \). Consequently, this estimator will be consistent only if \( d < \frac{1}{2} \). When \( d > \frac{1}{2} \), \( \hat{\beta} \) actually diverges in distribution at a rate that depends on \( d \). Equally, in the particular case where \( d = \frac{1}{2} \), \( \hat{\beta} \) has a nondegenerate limiting distribution.

As regards the hypothesis testing, from expressions (9) and (10) we have that the distributions of both traditional \( t \)-Student tests diverge at a rate, \( O_p(T^{3/2}) \), that does not depend on \( d \). Consequently, the possible consistency of \( \hat{\beta} \) (if \( d < \frac{1}{2} \)) to the true structural coefficient of zero does not translate into desirable properties for these conventional significance tests. This latter result mirrors the asymptotic behavior of the \( t \)-Student statistic in the general spurious framework recently studied by Marmol (1996a), which allows for the presence of deterministic trends of general form in a multivariate set-up of nonstationary fractionally integrated processes with different memory parameters and possibly cointegrated in the Engle and Granger (1987) sense.

With respect to the \( R^2 \) statistic, we observe that it converges weakly to the nondegenerate random variable given in expression (8), independently, hence, of \( d \). This result is also obtained in the case where we regress a set of stochastically independent integrated processes with different memory parameters (Marmol, 1996b) with or without the presence of deterministic trends (Marmol, 1996a). When we allow for the presence of cointegrating relationships, then \( R^2 \to 1 \), as proved by Haldrup (1994) and Marmol (1996a).

Finally, as regards the \( DW \) statistic, it converges in probability to zero for all \( d \geq \frac{1}{2} \), even that at a different rates, depending on \( d \). For \( d \geq \frac{3}{2} \), we have from expression (12) that \( DW = O_p(T^{-2}) \), whereas in the case where \( d < \frac{3}{2} \), \( DW = O_p(T^{1-2d}) \) depending, consequently, on the memory parameter \( d \). This implies, in particular, that in the degenerate case \( d = \frac{1}{2} \), \( DW = O_p(1) \), having, therefore, a well-defined limiting
distribution. Indeed, even in this case the $DW$ statistic can be a valid method of exposing the spurious regression (see Section 4 below).

In this sense, Marmol (1995, 1996b) proved that when we regress two integrated processes (with or without the same memory parameters) with no deterministic components except for a constant term in the spurious regression, then $DW = O_p\left(T^{-1}\right)$ if $d < \frac{1}{2}$ whereas $DW = O_p\left(T^{-2}\right)$ otherwise. By contrast, Marmol (1996a) proved that if we regress a multivariate set of fractionally integrated processes with different memory parameters, possibly cointegrated, with deterministic components, then the results obtained in Theorem 1 for the $DW$ statistic reemerge and $DW = O_p\left(T^{1-2d^*}\right)$ for $d^* < \frac{1}{2}$ and $DW = O_p\left(T^{-2}\right)$ for $d^* \geq \frac{1}{2}$, with $d^*$ being the minimum order of differentiation included in the underlying spurious regression.

To close this section, it is worth noting that $\hat{\alpha}$ and $\hat{\beta}$, suitable standardized, have Gaussian limiting distributions, in spite of being functionals of the fractional Brownian motion $y_\omega$. To prove this claim, we need the following lemma, which is of interest in its own:

**LEMMA 2:** Under DGP-A,

$$(13)\quad y_\omega(r) \sim N(0, \sigma^2 r^{2d-1}),$$

$$(14)\quad \int y_\omega(r)dr \sim N\left(0, \frac{2\sigma^2}{(d+1)(2d+1)}\right),$$

$$(15)\quad \int y_\omega(r)dr \sim N\left(0, \frac{2\sigma^2}{(d+2)(2d+3)}\right),$$

and

$$(16)\quad \int (r-a)y_\omega(r)dr \sim N(0, \Theta_\omega),$$

where $a \in \mathbb{R}$ and
\[ \Theta_a = 2 \sigma^2 \left[ \frac{1}{(d+2)(2d+3)} + \frac{a^2}{(d+1)(2d+1)} - \frac{a(2d+3)}{(d+1)(d+2)(2d+2)} \right] \]

Now, with the help of the above lemma, it is straightforward to prove the following result:

**COROLLARY 1:** Under DGP-A, the OLS estimators of \( \alpha, \beta \) in (1) have the following limiting distributions:

\[ T^{\nu_2 - d} \hat{\alpha} = N(0, \Theta^a), \]
\[ T^{\nu_2 - d} \hat{\beta} = N(0, \Theta^b), \]

with \( \Theta^a = 36 \Theta_{2/3}, \Theta^b = 144 \Theta_{1/2} \) and with \( \Theta_a \) as defined in expression (16).

3 NFI PROCESSES WITH DRIFT

Consider now the following DGP for the \( y_i \) sequence, where we allow for the presence of a non-zero drift, acting as a nuisance parameter:

**DGP-B:**

(i) \( \Delta^d y_i = \mu + \varepsilon_i, \ d \geq 1/2, \varepsilon_i \sim i.i.d(0, \sigma^2), E|\varepsilon_i|^{q} < \infty, \ q \geq \max\{d - 1/2, 2\} \).

Moreover, throughout this section we shall assume the following truncated property for the \( \Delta^d \) operator:

**ASSUMPTION 1:**

\[ \Delta^d \sim \sum_{j=1}^{t} \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(1+j)}, \ t = 1, 2, \ldots. \]
Durflauf and Phillips (1988) were also concerned with the effect of estimating (1) when, in fact, the true DGP was a difference stationary process with drift, $\Delta y_t = \mu + \varepsilon_t$, showing that the regression theory for this DGP was identical to the driftless case.

The intuition behind this result is that a difference stationary process with drift can be converted into a time trend plus a driftless difference stationary process

$$\Delta y_t = \mu + \varepsilon_t \iff y_t = \Delta^{-1}\mu + \Delta^{-1}\varepsilon_t = \mu t + \tilde{y}_t,$$

assuming zero initial conditions. Consequently, in the difference stationary case, regressing this sum against a time trend will generate identical results to those obtained in Section 2 where $\mu = 0$.

The same results will be obtained in the fractional case if we assume that $y_t$ has been generated according to

$$y_t = \alpha + \beta t + y_t^0 = \gamma \zeta_t + y_t^0,$$

where $\gamma = (\alpha, \beta)$, $\zeta_t = (1, t)'$ and $y_t^0$ is a $NFI(d)$ process generated under DGP-A. In this case, least squares estimation of $\gamma$ in (1) yields

$$\hat{\gamma} - \gamma = (\sum \zeta_t \zeta_t')^{-1}(\sum \zeta_t y_t^0)$$

and by defining the diagonal matrix $\mathcal{S}_\tau = \text{diag}\{1, T\}$ jointly with the CMT and Lemma 1, we have that

$$T^{3/2-d} \mathcal{S}_\tau^{-1}(\hat{\gamma} - \gamma) \Rightarrow (\int \tau')^{-1}(\int \tau \nu_\omega)$$

where $\tau = (1, r)'$ so that

$$T^{3/2-d}(\hat{\alpha} - \alpha) \Rightarrow -6 \int (r - 2/3) \nu_\omega(r) dr$$

and

$$T^{3/2-d}(\hat{\beta} - \beta) \Rightarrow 12 \int (r - 1/2) \nu_\omega(r) dr$$

as in Theorem 1. Consequently, if we only erroneously ignore stochastic regressors in model (1) but we correctly specified the deterministic components, then the regression theory is identical to the case where $y_t$ is assume to be generated without these non-
stochastic trends. Moreover, this result can be extended in a straightforward manner to
general polynomial trends as noted by Haldrup (1991a, b) in the \( d = 2 \) case.

However, if we assume true DGP-B, then things change drastically. In effect, in this
case, a reparametrization of \( y_t \) allows us to write it as the sum of a driftless \( \text{NFI}(d) \)
process plus a deterministic function of time
\[
\Delta^d y_t = \mu + \varepsilon_t \Rightarrow y_t^d = \Delta^d \mu + \Delta^d \varepsilon_t \\
\Rightarrow y_t = \xi_t + \bar{y}_t,
\]
where \( \xi_t = \Delta^d \mu \) and \( \bar{y}_t = \Delta^d \varepsilon_t \). As regards the \( \xi_t \) term, under Assumption 1, it follows
that
\[
\xi_t = \Delta^d \mu = \frac{\mu \Gamma(d)}{\Gamma(d+1)} \sum_{j=1}^{d} j^{d-1} \equiv \frac{\mu \Gamma(d)}{\Gamma(d+1)} \xi_t,
\]
where the approximation follows from Sheppard's formula. See Granger (1988) for a
detailed justification of this truncated deterministic filter.

Hence, by assuming true DGP-B, if we implement a least squares testing procedure in
model (1) we are mistakenly ignore stochastic regressors and misspecifying the non-
stochastic component. In this section we will explore the consequences of this double
misspecification on the asymptotic behavior of the traditional least squares statistics. For
this, let us first announce the following results.

**Lemma 3:**

(17) \( T^{-d} \bar{\xi}_t \rightarrow \frac{r^d}{d} \),

(18) \( T^{-d-1} \sum \bar{\xi}_t \rightarrow \frac{1}{d(d+1)} \),

(19) \( T^{-d-2} \sum \bar{\xi}_t \rightarrow \frac{1}{d(d+2)} \),

(20) \( T^{-2d-1} \sum (\bar{\xi}_t)^2 \rightarrow \frac{1}{d^2(2d+1)} \).
LEMMA 4: Under DGP-B,

\[ T^{-2d-1/2} \sum \hat{z}_i \hat{y}_i \xrightarrow{p} \frac{1}{d} \int \frac{r^d y_0 (r) dr}{d-1} \]

(22) \[ T^{-d-1} \sum y_t \xrightarrow{p} \frac{\mu}{\Gamma(d + 2)} \]

(23) \[ T^{-1-2d} \sum y_t^2 \xrightarrow{p} \frac{\mu^2}{\Gamma^2(d + 1)(2d + 1)} \]

(24) \[ T^{-1-2d} \sum (y_t - \bar{y})^2 \xrightarrow{p} \frac{\mu^2}{\Gamma^2(d)(d + 1)^2(2d + 1)} \]

(25) \[ T^{-2-d} \sum t y_t \xrightarrow{p} \frac{\mu}{\Gamma(d + 2)} \]

(26) \[ T^{-2-d} \sum (y_t - \bar{y})(t - \bar{t}) \xrightarrow{p} \frac{\mu}{2\Gamma(d + 3)} \]

Using the above results and the methods employed in Theorem 1, it is now easy to find the relevant asymptotic theory for model (1).

THEOREM 2:

(i) Under DGP-B, if \( d \neq 1 \),

(27) \[ T^{-d} \hat{\alpha} \xrightarrow{p} \mu \frac{2(1-d)}{\Gamma(d+3)} \]

(28) \[ T^{-d} \hat{\beta} \xrightarrow{p} \mu \frac{6d}{\Gamma(d+3)} \]

(29) \[ R^2 \xrightarrow{p} \frac{6d - 3}{(d + 2)^2} (\geq 0) \]

(30) \[ T^{-1/2} \hat{\rho} \xrightarrow{p} \frac{d(6d + 3)^{1/2}}{d-1} \]
(31) \( T^{-1/2} t_a \xrightarrow{p} \frac{(2d+1)^{1/2}}{d} \),

(32) if \( 1/2 \leq d < 1 \),

\( T^{2d} DW \xrightarrow{p} \frac{\sigma^2 \Gamma(3-2d) \Gamma^2(d+3)(2d+1)}{\mu^2 \Gamma(2-d)(d-1)^2} \),

(33) if \( d > 1 \),

\( T^2 DW \xrightarrow{p} \frac{(2d+1) \left[ \frac{\Gamma^2(d+3) + 36 \Gamma^2(d+1)(2d-1)}{\Gamma^2(d)(2d-1)} \right]}{(d-1)^2} \)

\( \cdot \frac{-12 \Gamma(d) \Gamma(d+3)(2d-1)}{\Gamma^2(d)(2d-1)} \).

(ii) Under DGP-B, in the case where \( d = 1 \),

(34) \( T^{-1/2} \hat{\alpha} = N \left( 0, \frac{2\sigma^2}{15} \right) \),

(35) \( T^{d/2}(\hat{\beta} - \mu) = N \left( 0, \frac{6\sigma^2}{15} \right) \),

whilst \( R^2, t_a, t_{\hat{\beta}} \) and \( DW \) have the same limiting distributions as in Theorem 1.

First at all, notice that, when \( d = 1 \), in model (1) we only misspecify the stochastic trends. Consequently, and according to aforementioned comments, the asymptotic behavior of the \( OLS \) statistics obtained in the second part of the theorem is the same than in the driftless case.

On the other hand, when \( d \neq 1 \) all the statistics converge in probability to some constant after suitable standardization. This constitutes a kind of return to more classical asymptotic results with stationary variables, and, of course, this fact is due to the asymptotically dominant \( \xi_t \) term. The influence of this deterministic component on the relevant asymptotic theory for the \( OLS \) statistics that we are examining only cancels exactly when \( d = 1 \).
Otherwise, in the case where \( d \neq 1 \), we can see in the first part of the theorem that the constant term is \( o_p(T^d) \), i.e., it diverges and a rate that depends on \( d \). By contrast, the estimated time trend coefficient \( \hat{\beta} \) diverges when \( d > 1 \) but converges to the true structural coefficient of zero when \( d < 1 \). Further, the coefficient of determination \( R^2 \) converges in probability to a non-negative constant and the \( t \)-Student tests diverge at the rate \( o_p(T^{1/2}) \), as always. Finally, as in the driftless case, \( DW \to 0 \) for all \( d \geq \frac{1}{2} \), and, hence, it remains to be a valid tool against misspecification.

### 4 SOME MONTE CARLO EVIDENCE

In this section we perform some Monte Carlo experiments to examine the power of the \( DW \) statistic as a misspecification test against spurious detrending in small to moderate samples. The parameter space considered in this study is the following:

\[
\{ T = 50, 100, 200 \} \otimes \{ d = 0.5, 0.6, 0.8, 1, 1.2, 1.5, 1.8, 2 \} \otimes \{ \mu = 0, 1 \} \otimes \{ \vartheta = 0, 0.4, 0.8 \}
\]

Observations on the \( NFI(d) \) process were generated in the following manner: First, we simulate a stationary \( \eta_{\delta} \sim SF_{\delta}(\delta) \) process for \( \delta \in \left[ -\frac{1}{2}, \frac{1}{2} \right] \). In this case we have that \( \eta_{\delta} = \Delta^{-\delta} u_{\delta} \), where the perturbation term \( u_{\delta} \) is generated as \( u_{\delta} = \varepsilon_{\delta} - \vartheta \varepsilon_{\delta-1} \), where \( \varepsilon_{\delta} \) is a sequence of identically and independently distributed \( N(0,1) \) variables. The fractional difference operator \( \Delta^{-\delta} \) is defined as an infinite lag polynomial, so that we must truncate it in some point \( m \). Herein we choose \( m = 10,000 \). Now to mimic the sample path of the \( NFI(d) \) process \( y_{\delta} \), for \( d \in \left[ \frac{1}{2}, 2 \right] \), we simply take partial sums of \( \eta_{\delta} \) with initial condition \( y_0 = 0 \). Thus, as a matter of definition, \( y_{\delta} \) is \( NFI(1.4) \) if \( \Delta y_{\delta} = \eta_{\delta} \sim SF_{\delta}(0.4) \).

Using this procedure, for each sample size \( T \), we generate \( T + 100 \) observations. Then, the first 100 observations are discarded in order to eliminate the influence of the initial conditions.

The results of our experiments are given in Tables 1-3. These results were generated by a simulation using 20,000 replications. As is well-known, a serious shortcoming of the \( DW \) test is that its exact distribution depends on the data matrix. Notwithstanding, the true distribution of the \( DW \) statistic lies between that of two other statistics, \( d_{i} \) (the
lower bound) and \( d_* \) (the upper bound), which only depend on \( T \) and the number of regressors. The null hypothesis of no autocorrelation is rejected against the alternative of positive autocorrelation if \( DW < d_* \) (Region 1) against the alternative of negative autocorrelation if \( DW > 4 - d_* \) (Region 5) and not rejected if \( d_* < DW < 4 - d_* \) (Region 3), where asterisks indicate tabulated values at appropriate significance levels (e.g., Savin and White, 1977). If \( d_* < DW < d_* \) (Region 2) or if \( 4 - d_* < DW < 4 - d_* \) (Region 4) the test is inconclusive. For each point of the parameter space, tables give percent of times each of these five region contains the DW statistic.

Table 1 gives the power of the DW statistic in the most favorable situation, namely, where the true model is a pure NFI with white noise innovations. In Tables 2 and 3 we compare power in the presence of a NFI with moderate to large MA(1) innovations, respectively. Some results are clear and in accord with the asymptotics of the previous sections. With other things held constant, (i) power increases as \( T \) increases This is a reflection of the consistency of the test. (ii) Power is higher when \( d \) is larger, i.e., as the alternative hypothesis becomes further from the null.

Other important practical conclusions that can be drawn from our simulations are the following. First, the presence or absence of the drift term \( \mu \) does not change the power properties of the DW statistic. Durlauf and Phillips (1988) and our Theorem 2 prove that this must be the case for \( d = 1 \). However, Tables 1-3 show that this result remains true for all \( d \geq \frac{1}{2} \). Consequently, it seems that the power properties of the DW statistic are independent of misspecifications in the deterministic part. Second, the percent of times the DW statistic lies in Regions 2 and 4 (the inconclusive regions) are almost zero in all cases (perhaps to a surprising degree, in fact).

On the other hand, the DW statistic has good power properties for all \( d \geq \frac{1}{2} \) in the presence of moderate moving average innovations (Table 2). Indeed, this power can be very low for \( d < 1 \) against NFI processes with large moving average innovations (Table 3), pointing out the presence of severe identification problems. For instance, when \( d = 0.8 \) the null hypothesis of correct specification is accepted in about 70% for \( T = 200 \).
Finally, note that, for this sample size, for \( d = 0.5 \) or \( d = 0.6 \), the \( DW \) statistic lies in on 5 in almost all the occasions.

5 CONCLUSIONS

The aim of this paper has been to provide a generalization of the available results on the behavior of difference stationary processes which are misspecified as trend stationary time series to the more general framework of misspecified nonstationary fractionally integrated processes.

Three conclusions can be drawn from our study. First, depending on the memory parameter \( d \), the estimated time trend coefficient can either diverge or converge to the structural parameter of zero. Second, when we assume that the true generating mechanism includes a non-zero drift, then the asymptotic behavior of the standard \( OLS \) statistics changes with respect to the driftless case, with the only exception of the case where \( d = 1 \), i.e., in the difference stationary case. This is due to the fact that, except in the unit root case, for most of values of \( d \), a least squares regression in model (1) implies to commit two different misspecifications, namely, to ignore the stochastic trends and to misspecify the deterministic component of the underlying series of interest. Third, in any case, the conventional \( t \)-Student statistics diverge at the rate \( O_p(T^{1/2}) \). Consequently, they will, with probability one, reject the null hypothesis of no significance. Hence, standard \( OLS \) inference remains invalid, as in the difference stationary case. Finally, we showed that, in any case, \( DW \xrightarrow{p} 0 \) for all \( d \geq \frac{1}{2} \), and, therefore, this statistic will, with probability one, reject the hypothesis of correct model specification, as \( T \to \infty \). This, in turn, constitutes an useful diagnostic against misspecification. However, from our Monte Carlo experiments we conclude that, in moderate samples and for \( d < 1 \), the \( DW \) statistic has important identification problems except in the simplest case of a pure \( NFI \) process with white noise innovations.
<table>
<thead>
<tr>
<th>T</th>
<th>μ</th>
<th>Region</th>
<th>0.5</th>
<th>0.6</th>
<th>0.8</th>
<th>1</th>
<th>1.2</th>
<th>1.5</th>
<th>1.8</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0</td>
<td>1</td>
<td>0.8770</td>
<td>0.9560</td>
<td>0.9970</td>
<td>1.0000</td>
<td>0.9990</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>2</td>
<td>0.0370</td>
<td>0.0160</td>
<td>0.0020</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>3</td>
<td>0.0860</td>
<td>0.0280</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>1</td>
<td>0.8670</td>
<td>0.9550</td>
<td>0.9970</td>
<td>0.9990</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9990</td>
<td>1.0000</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>2</td>
<td>0.0350</td>
<td>0.0170</td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>3</td>
<td>0.0980</td>
<td>0.0280</td>
<td>0.0030</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>1</td>
<td>0.9970</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>2</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>3</td>
<td>0.0020</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>1</td>
<td>0.9950</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>2</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>3</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>1</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>2</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>3</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>1</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>2</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>3</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**True model:** \( \Delta^T y_t = \mu + \varepsilon_t, \quad \varepsilon_t \sim \text{NIID}(0,1) \).

**Estimated model:** \( y_t = \hat{\alpha} + \hat{\beta} t + \text{res} \).

*Regions 1 and 5:* regions of rejection of the null hypothesis of correct specification. *Region 3:* region of no rejection of the null hypothesis of correct specification. *Regions 2 and 4:* inconclusive regions. 5% significance level.
### TABLE 2

Power of the $DW$ statistic against spurious detrending

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\mu$</th>
<th>Region</th>
<th>0.5</th>
<th>0.6</th>
<th>0.8</th>
<th>1</th>
<th>1.2</th>
<th>1.5</th>
<th>1.8</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0</td>
<td>1</td>
<td>0.0960</td>
<td>0.2570</td>
<td>0.7540</td>
<td>0.9600</td>
<td>0.9680</td>
<td>0.9520</td>
<td>0.9600</td>
<td>0.9490</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>2</td>
<td>0.0340</td>
<td>0.0880</td>
<td>0.0640</td>
<td>0.0140</td>
<td>0.0130</td>
<td>0.0180</td>
<td>0.0090</td>
<td>0.0200</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>3</td>
<td>0.8050</td>
<td>0.6430</td>
<td>0.1790</td>
<td>0.0260</td>
<td>0.0190</td>
<td>0.0300</td>
<td>0.0310</td>
<td>0.0310</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>4</td>
<td>0.0340</td>
<td>0.0660</td>
<td>0.0030</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>5</td>
<td>0.0310</td>
<td>0.0080</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>1</td>
<td>0.1010</td>
<td>0.2490</td>
<td>0.7380</td>
<td>0.9610</td>
<td>0.9590</td>
<td>0.9680</td>
<td>0.9560</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>2</td>
<td>0.0390</td>
<td>0.0810</td>
<td>0.0580</td>
<td>0.0200</td>
<td>0.0160</td>
<td>0.0140</td>
<td>0.0110</td>
<td>0.0120</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>3</td>
<td>0.7870</td>
<td>0.6500</td>
<td>0.2040</td>
<td>0.0340</td>
<td>0.0230</td>
<td>0.0260</td>
<td>0.0210</td>
<td>0.0320</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>4</td>
<td>0.0290</td>
<td>0.0100</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>5</td>
<td>0.0440</td>
<td>0.0100</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>1</td>
<td>0.2790</td>
<td>0.6530</td>
<td>0.9890</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>2</td>
<td>0.0550</td>
<td>0.0470</td>
<td>0.0030</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>3</td>
<td>0.6360</td>
<td>0.2970</td>
<td>0.0080</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>4</td>
<td>0.0110</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>5</td>
<td>0.0190</td>
<td>0.0020</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>1</td>
<td>0.2790</td>
<td>0.6680</td>
<td>0.9910</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>2</td>
<td>0.0500</td>
<td>0.0490</td>
<td>0.0020</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>3</td>
<td>0.6630</td>
<td>0.2800</td>
<td>0.0070</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>4</td>
<td>0.0070</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>5</td>
<td>0.0100</td>
<td>0.0020</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>1</td>
<td>0.6410</td>
<td>0.9590</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>2</td>
<td>0.0250</td>
<td>0.0040</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>3</td>
<td>0.3330</td>
<td>0.0370</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>4</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>1</td>
<td>0.6510</td>
<td>0.9670</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>2</td>
<td>0.0380</td>
<td>0.0080</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>3</td>
<td>0.1080</td>
<td>0.0250</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>5</td>
<td>0.0030</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**True model:** $\Delta y_t = \mu + u_t$, $u_t = \varepsilon_t - 0.4 \varepsilon_{t-1}$, $\varepsilon_t \sim \text{NID}(0,1)$.

**Estimated model:** $y_t = \hat{\alpha} + \hat{\beta} t + \text{res}$.

Regions 1 and 5: regions of rejection of the null hypothesis of correct specification. Region 3: region of no rejection of the null hypothesis of correct specification. Regions 2 and 4: inconclusive regions. 5% significance level.
## TABLE 3

Power of the $DW$ statistic against spurious detrending

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\mu$</th>
<th>Region</th>
<th>0.5</th>
<th>0.6</th>
<th>0.8</th>
<th>1</th>
<th>1.2</th>
<th>1.5</th>
<th>1.8</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0130</td>
<td>0.1580</td>
<td>0.2040</td>
<td>0.1850</td>
<td>0.1860</td>
<td>0.1810</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>2</td>
<td>0.0000</td>
<td>0.0020</td>
<td>0.0150</td>
<td>0.0730</td>
<td>0.0640</td>
<td>0.0580</td>
<td>0.0660</td>
<td>0.0730</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>3</td>
<td>0.2880</td>
<td>0.4650</td>
<td>0.8110</td>
<td>0.7200</td>
<td>0.7060</td>
<td>0.7280</td>
<td>0.7220</td>
<td>0.7160</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>4</td>
<td>0.1300</td>
<td>0.1350</td>
<td>0.0580</td>
<td>0.0220</td>
<td>0.0140</td>
<td>0.0150</td>
<td>0.0090</td>
<td>0.0160</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>5</td>
<td>0.5820</td>
<td>0.3980</td>
<td>0.1030</td>
<td>0.0270</td>
<td>0.0120</td>
<td>0.0140</td>
<td>0.0170</td>
<td>0.0140</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0170</td>
<td>0.1860</td>
<td>0.1920</td>
<td>0.1760</td>
<td>0.1630</td>
<td>0.1940</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>2</td>
<td>0.0000</td>
<td>0.0120</td>
<td>0.0680</td>
<td>0.0650</td>
<td>0.0600</td>
<td>0.0890</td>
<td>0.0550</td>
<td>0.0660</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>3</td>
<td>0.3010</td>
<td>0.4680</td>
<td>0.7960</td>
<td>0.7160</td>
<td>0.7210</td>
<td>0.7440</td>
<td>0.7250</td>
<td>0.7280</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>4</td>
<td>0.1150</td>
<td>0.1320</td>
<td>0.0790</td>
<td>0.0080</td>
<td>0.0150</td>
<td>0.0100</td>
<td>0.0130</td>
<td>0.0110</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>5</td>
<td>0.5840</td>
<td>0.4000</td>
<td>0.0960</td>
<td>0.0220</td>
<td>0.0070</td>
<td>0.0100</td>
<td>0.0100</td>
<td>0.0120</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0530</td>
<td>0.5730</td>
<td>0.6130</td>
<td>0.5790</td>
<td>0.5990</td>
<td>0.5760</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>2</td>
<td>0.0000</td>
<td>0.0140</td>
<td>0.0440</td>
<td>0.0490</td>
<td>0.0530</td>
<td>0.0550</td>
<td>0.0550</td>
<td>0.0460</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>3</td>
<td>0.0860</td>
<td>0.2670</td>
<td>0.8030</td>
<td>0.3820</td>
<td>0.3380</td>
<td>0.3640</td>
<td>0.3450</td>
<td>0.3770</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>4</td>
<td>0.0380</td>
<td>0.0780</td>
<td>0.0300</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0020</td>
<td>0.0010</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>5</td>
<td>0.8760</td>
<td>0.6540</td>
<td>0.1000</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0020</td>
<td>0.0000</td>
<td>0.0010</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0520</td>
<td>0.5860</td>
<td>0.6430</td>
<td>0.5890</td>
<td>0.6050</td>
<td>0.5600</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>2</td>
<td>0.0000</td>
<td>0.0300</td>
<td>0.0300</td>
<td>0.0540</td>
<td>0.0450</td>
<td>0.0550</td>
<td>0.0490</td>
<td>0.0500</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>3</td>
<td>0.0760</td>
<td>0.2640</td>
<td>0.7840</td>
<td>0.3560</td>
<td>0.3090</td>
<td>0.3520</td>
<td>0.3440</td>
<td>0.3890</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>4</td>
<td>0.0430</td>
<td>0.0530</td>
<td>0.0320</td>
<td>0.0020</td>
<td>0.0010</td>
<td>0.0020</td>
<td>0.0010</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>5</td>
<td>0.8810</td>
<td>0.5830</td>
<td>0.1020</td>
<td>0.0020</td>
<td>0.0020</td>
<td>0.0020</td>
<td>0.0010</td>
<td>0.0010</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.2000</td>
<td>0.9610</td>
<td>0.9540</td>
<td>0.9550</td>
<td>0.9480</td>
<td>0.9530</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>2</td>
<td>0.0000</td>
<td>0.0310</td>
<td>0.0090</td>
<td>0.0100</td>
<td>0.0060</td>
<td>0.0130</td>
<td>0.0110</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>3</td>
<td>0.0090</td>
<td>0.1190</td>
<td>0.7190</td>
<td>0.0300</td>
<td>0.0360</td>
<td>0.0390</td>
<td>0.0390</td>
<td>0.0360</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>4</td>
<td>0.0010</td>
<td>0.0300</td>
<td>0.0060</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>5</td>
<td>0.9900</td>
<td>0.6810</td>
<td>0.8510</td>
<td>0.0440</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>1</td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.2270</td>
<td>0.9580</td>
<td>0.9670</td>
<td>0.9530</td>
<td>0.9410</td>
<td>0.9480</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>2</td>
<td>0.0000</td>
<td>0.0240</td>
<td>0.0070</td>
<td>0.0070</td>
<td>0.0080</td>
<td>0.0150</td>
<td>0.0080</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>3</td>
<td>0.0060</td>
<td>0.1060</td>
<td>0.6940</td>
<td>0.0340</td>
<td>0.0260</td>
<td>0.0390</td>
<td>0.0430</td>
<td>0.0440</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>4</td>
<td>0.0010</td>
<td>0.0320</td>
<td>0.0070</td>
<td>0.0010</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>5</td>
<td>0.9930</td>
<td>0.6810</td>
<td>0.8610</td>
<td>0.0480</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

*True model: $\Delta y_t = \mu + u_t, \quad u_t = \epsilon_t - 0.8 \epsilon_{t-1}, \quad \epsilon_t \sim \text{NIID}(0,1)$

*Estimated model: $y_t = \tilde{\alpha} + \tilde{\beta} t + \text{res.}$

*Regions 1 and 5: regions of rejection of the null hypothesis of correct specification. Region 3: region of no rejection of the null hypothesis of correct specification. Regions 2 and 4: inconclusive regions. 5% significance level.*
APPENDIX

PROOF OF THEOREM 1: First, consider the OLS estimators of $\alpha$ and $\beta$ in model (1)

$$\hat{\alpha} = \theta^{-1}\left[\sum i^2 \sum y_i - \sum ty_i \sum t\right],$$

and

$$\hat{\beta} = \theta^{-1}\left[T \sum ty_i - \sum t \sum y_i\right],$$

where $\theta = T \sum t^2 - \left(\sum t\right)^2$. Given the general result

(A1) \[ T^{-m} \sum i^{m-1} = T^{-1} \sum (t/T)^{m-1} \to \int_0^1 r^{m-1} dr = \frac{r^m}{m} \bigg|_0^1 = \frac{1}{m}. \]

it follows that

(A2) \[ T^{-4} \theta = T^{-3} \sum i^2 - (T^{-2} \sum t)^2 \to \frac{1}{12}. \]

Therefore, using Lemma 1, (A2) and the CMT, we have that

$$T^{1/2-d} \hat{\alpha} = \left(T^{-4} \theta\right)^{-1} \left[ T^{-3} \sum i^2 T^{-1/2-d} \sum y_i - T^{-3/2-d} \sum ty_i T^{-d} \sum t \right]$$

$$\Rightarrow 12 \left[ \frac{1}{3} \int y_\infty - \frac{1}{2} \int r y_\infty \right] = -6 \int \left(r - 2/3\right) y_\infty.$$

and

$$T^{3/2-d} \hat{\beta} = \left(T^{-4} \theta\right)^{-1} \left[ T^{-3/2-d} \sum ty_i - T^{-2} \sum t T^{-1/2-d} \sum y_i \right]$$

$$\Rightarrow 12 \left[ \int r y_\infty - \frac{1}{2} \int y_\infty \right] = 12 \int \left(r - 1/2\right) y_\infty.$$

With respect to the $R^2$ statistic,

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{\sum \hat{u}_i^2}{\sum (y_i - \bar{y})^2},$$

using Lemma 1, (A1) and the CMT, we get the following results

(A3) \[ T^{-d} \sum (y_i - \bar{y})^2 = T^{-d} \sum y_i^2 - \left(T^{-1/2-d} \sum y_i\right)^2 \Rightarrow \int y_\infty^2 - \left[ \int y_\infty \right]^2, \]

(A4) \[ T^{-3/2-d} \sum (y_i - \bar{y})(t - \bar{t}) \Rightarrow \int \left(r - 1/2\right) y_\infty, \]

and

(A5) \[ T^{-d} \sum \hat{u}_i^2 = T^{-d} \sum (y_i - \bar{y})^2 - T^{3/2-d} \hat{\beta} T^{-3/2-d} \sum (y_i - \bar{y})(t - \bar{t}). \]
where $\bar{y} = T^{-1}\sum y_i$ and $\bar{t} = T^{-1}\sum t$. Now, (A3)-(A5) and the CMT together imply

\[ R^2 \Rightarrow 1 - \frac{\sum y_i^2 - \left( \sum y_i \right)^2 - 12\left( \sum (r-1/2)y_i \right)^2}{\sum y_i^2 - \left( \sum y_i \right)^2} = \frac{12\left( \sum (r-1/2)y_i \right)^2}{\sum y_i^2 - \left( \sum y_i \right)^2}. \]

Consider now the $t$-Student statistics,

\[ t_\beta^2 = \frac{\hat{\beta}^2}{T^{-1}\sum \hat{u}^2}, \]

and

\[ t_\alpha^2 = \frac{\hat{\alpha}^2}{T^{-1}\sum \hat{u}^2 T^{-3}\sum t^2}. \]

In this case, from (6), (7), (A1), (A5) and the CMT, it is straightforward to show that

\[ T^{-1}t_\beta^2 = \frac{T^{-3}d\hat{\beta}^2}{T-2d\sum \hat{u}^2 \sum t^2} \Rightarrow \frac{12\left( \sum ry_i - 0.5\sum y_i \right)^2}{\sum y_i^2 - \left( \sum y_i \right)^2 - 12\left( \sum (r-1/2)y_i \right)^2}. \]

and

\[ T^{-1}t_\alpha^2 = \frac{T^{-1}d\hat{\alpha}^2}{T^{-3}d\sum \hat{u}^2 T^{-3}\sum t^2} \Rightarrow \frac{9\left( \sum (r-2/3)y_i \right)^2}{\sum y_i^2 - \left( \sum y_i \right)^2 - 12\left( \sum (r-1/2)y_i \right)^2}. \]

Finally, consider the asymptotic behavior of the $DW$ statistic, defined as

\[ DW = \frac{\sum (\Delta\hat{u}_i)^2}{\sum \hat{u}_i^2}. \]
Manipulating the numerator, we get
\[ \sum (\Delta u_i)^2 = \sum (\Delta y_i)^2 + T\beta - 2\beta \sum \Delta y_i. \]

Let us first assume that \( d \geq 3/2 \). In this case, \( \Delta y_i \) is a \( NFI(d-1) \) process and hence, using Lemma 1, it follows that
\[
\text{(A6)} \quad T^{2-d} \sum (\Delta y_i)^2 \Rightarrow \int (\Delta y_\omega)^2,
\]
and
\[
\text{(A7)} \quad T^{1/2-d} \sum \Delta y_i \Rightarrow \int \Delta y_\omega.
\]

Consequently, using (7), (A5)-(A7) and the CMT, it can be deduced that
\[
T^2 DW \Rightarrow \frac{\int (\Delta y_\omega)^2 + 144 \left[ \int (r-1/2)y_\omega \right]^2 - 24 \int (r-1/2)y_\omega \int \Delta y_\omega}{\int y_\omega^2 - \left[ \int y_\omega \right]^2 - 12 \left[ \int (r-1/2)y_\omega \right]^2}.
\]

In the same manner, consider the case where \( d < 3/2 \). Now, observe that \( \Delta y_i \) is a \( SFI(d-1) \) process, and standard asymptotic results can be applied to its sample moments, yielding
\[
\text{(A8)} \quad T^{-1} \sum \Delta y_i \rightarrow^p 0
\]
and
\[
\text{(A9)} \quad T^{-1} \sum (\Delta y_i)^2 \rightarrow^p \text{var}(\Delta y_i) = \frac{\sigma^2 \Gamma(3-2d)}{\Gamma(2-d)}.
\]

Hence, by virtue of expressions (7), (A5), (A8), (A9) and the CMT, it is direct to show that \( T^{-1} \sum (\Delta u_i)^2 \approx T^{-1} \sum (\Delta y_i)^2 \) which, in turn, implies that
\[
T^{2d-1} DW \Rightarrow \frac{\sigma^2 \Gamma(3-2d)}{\Gamma^2(2-d)} \left\{ \frac{\int y_\omega^2 - \left[ \int y_\omega \right]^2}{\int (r-1/2)y_\omega^2} \right\},
\]
as required.

\textit{Q.E.D.}

**PROOF OF LEMMA 2:** Given that \( W(r) \) is a Brownian motion, we know, by linearity with respect to \( W(r) \), that \( y_\omega, \int y_\omega, \int ry_\omega \) and \( \int (r-a)y_\omega \) are also Gaussian processes.
Indeed, expression (13) follows from Jonas (1983). To prove expression (14), notice from (13) that

\[(A10) \quad \int_{0}^{1} y_{\omega}(r)dr = \sigma \int_{0}^{1} r^{d-1} W(r)dr. \]

Now, given that

\[E \int_{0}^{1} r^{d-1} W(r)dr = \int_{0}^{1} r^{d-1} E(W(r))dr = 0\]

and

\[\text{var} \int_{0}^{1} r^{d-1} W(r)dr = E \int_{0}^{1} r^{d-1} s^{d-1} W(r)W(s)drds\]

\[= \int_{0}^{1} \int_{0}^{1} r^{d-1} s^{d-1} E(W(r)W(s))drds\]

\[= \int_{0}^{1} \int_{0}^{1} s^{d-1} \min\{r, s\}drds = \int_{0}^{1} \left[ \int_{0}^{s} ds + r \int_{r}^{1} s^{d-1} ds \right] dr\]

\[= \int_{0}^{1} r^{d-1} \left[ \frac{s^{d+1}}{d+1} + r \frac{s^{d}}{d} \right] dr = \int_{0}^{1} r^{d-1} \left[ \frac{r^{d+1}}{d+1} + \frac{r - r^{d}}{d} \right] dr = \frac{2}{(d+1)(2d+1)},\]

we get \(\int_{0}^{1} r^{d-1} W(r)dr \sim N(0, 2/(d+1)(2d+1))\), meaning that

\[\int_{0}^{1} y_{\omega}(r)dr \sim N\left(0, \frac{2\sigma^2}{(d+1)(2d+1)}\right).\]

Similarly, it can be proved that \(\int_{0}^{1} r^{d} W(r)dr \sim N\left(0, 2/(d+2)(2d+3)\right)\), implying

\[\int_{0}^{1} r y_{\omega}(r)dr \sim N\left(0, \frac{2\sigma^2}{(d+2)(2d+3)}\right).\]

Finally, to prove expression (16), notice that

\[\int_{0}^{1} (r-a) y_{\omega}(r)dr = \int_{0}^{1} r y_{\omega}(r)dr - a \int_{0}^{1} y_{\omega}(r)dr.\]
Hence,

$$E \int_0^1 (r-a) y_\omega(r) dr = \int_0^1 rE(y_\omega(r)) dr - a \int_0^1 E(y_\omega(r)) dr = 0$$

and

$$\text{var} \int_0^1 (r-a) y_\omega(r) dr = \Theta_a = \text{var} \int_0^1 r y_\omega(r) dr + a^2 \text{var} \int_0^1 y_\omega(r) dr$$

$$-2a \text{cov}\left( \int_0^1 r y_\omega(r) dr, \int_0^1 y_\omega(r) dr \right).$$

Thus, given that

$$\int_0^1 \int_0^1 \int_0^1 E(y_\omega(r) y_\omega(s)) dr ds = \int_0^1 \int_0^1 E(W_\omega(r) W_\omega(s)) dr ds,$$

we obtain after some manipulations

$$\text{cov}\left( \int_0^1 r y_\omega(r) dr, \int_0^1 y_\omega(r) dr \right) = \sigma^2 \int_0^1 s \text{min}\{r, s\} dr ds$$

$$= \frac{\sigma^2 (2d+3)}{(2d+2)(d+1)(d+2)}.$$
\[ T^{-d-1} \sum \tilde{z}_i = T^{-1} \sum \left( T^{-d} \tilde{z}_i \right) \to \frac{1}{d} \int_0^1 r^d \, dr = \frac{1}{d(d+1)}, \]

\[ T^{-d-2} \sum \tilde{t}_i = T^{-1} \sum \left( T^{-d} \tilde{t}_i \right) \to \frac{1}{d^2} \int_0^1 r^{d+1} \, dr = \frac{1}{d(d+2)}, \]

\[ T^{-2d-1} \sum \left( \tilde{z}_i \right)^2 = T^{-1} \sum \left( T^{-d} \tilde{z}_i \right)^2 \to \frac{1}{d^2} \int_0^1 r^{2d} \, dr = \frac{1}{d^2(2d+1)}, \]

\[ T^{-2d-1/2} \sum \tilde{z}_i \tilde{y}_i = T^{-1} \sum \left( T^{-d} \tilde{z}_i \right) \left( T^{-1/2} \tilde{y}_i \right) \to \frac{1}{d} \int_0^1 r^d \, y_i(r) \, dr. \]

Q.E.D.

PROOF OF LEMMA 4. Using Lemma 1, Lemma 3 and the CMT, expressions (22)-(26) can be proved in the following manner:

\[ T^{-d-1} \sum y_i = \frac{\mu}{\Gamma(d)} T^{-d-1} \sum \tilde{z}_i + T^{-1/2} T^{-1/2-d} \sum \tilde{y}_i \to \frac{\mu}{\Gamma(d)d(d+1)}, \]

\[ T^{-1-2d} \sum y_i^2 = \frac{\mu^2}{\Gamma^2(d)} T^{-1-2d} \sum \tilde{z}_i^2 + T^{-1} T^{-2d} \sum \tilde{y}_i^2 \]

\[ +2 \frac{\mu}{\Gamma(d)} T^{-1/2} T^{-1/2-2d} \sum \tilde{z}_i \tilde{y}_i \to \frac{\mu^2}{\Gamma^2(d+1)(2d+1)}, \]

\[ T^{-1-2d} \sum (y_i - \bar{y})^2 = T^{-1-2d} \sum y_i^2 - \left( T^{-d} \bar{y} \right)^2 \]

\[ \to \frac{\mu^2}{\Gamma^2(d+1)(2d+1)} - \frac{\mu^2}{\Gamma^2(d+1)(2d+1)} = \frac{\mu^2}{\Gamma^2(d)(d+1)^2(2d+1)}, \]

\[ T^{-2d} \sum t_y = \frac{\mu}{\Gamma(d)} T^{-2-d} \sum \tilde{t}_i + T^{-1/2} T^{-3/2-d} \sum \tilde{t}_i \tilde{y}_i \to \frac{\mu}{\Gamma(d)d(d+2)}, \]

and

\[ T^{-2d} \sum (y_i - \bar{y})(t - t_i) = T^{-2-d} \sum t_y(t - t_i) - T^{-2} \sum t T^{-1-d} \sum y_i \]

\[ \to \frac{\mu}{\Gamma(d)d(d+2)} - \frac{\mu}{2\Gamma(d)d(d+1)} = \frac{\mu}{2\Gamma(d+3)}. \]

Q.E.D.
PROOF OF THEOREM 2. Under $DGP-B$, 
\[ \hat{\alpha} = \theta^{-1} \left[ \sum i^2 \sum y_i - \sum i \bar{y}, \sum t \right] \]
\[ = \theta^{-1} \left[ \sum i^2 \sum \tilde{y}_i - \sum i \bar{\tilde{y}}_i, \sum t \right] + \theta^{-1} \frac{\mu}{\Gamma(d)} \left[ \sum i^2 \sum \tilde{\xi}_i - \sum i \bar{\tilde{\xi}}_i, \sum t \right]. \]

If $d \neq 1$, then $\sum i^2 \sum \tilde{\xi}_i \neq \sum i \bar{\tilde{\xi}}_i, \sum t$. Hence, using (A1), Lemma 1, Lemma 2 and the CMT
\[ T^{-d} \hat{\alpha} = (T^{-d} \theta)^{-1} \left[ T^{-1/2} \left[ T^{-1} \sum i^2 T^{-1/2-d} \sum \tilde{y}_i - T^{-3/2-d} \sum i \tilde{y}_i, T^{-2} \sum t \right] \right. \]
\[ + (T^{-d} \theta)^{-1} \frac{\mu}{\Gamma(d)} \left[ T^{-3} \sum i^2 T^{-1-d} \sum \tilde{\xi}_i - T^{-2-d} \sum i \tilde{\xi}_i, T^{-2} \sum t \right] \]
\[ \left. \quad \rightarrow \frac{\mu}{\Gamma(d)} \left[ \frac{4}{d(d+1)} - \frac{6}{d(d+2)} \right] = \frac{2(1-d)}{\Gamma(d)} \frac{2(1-d)}{\Gamma(d+3)} \right]. \]

where in the last equality we have employed the well-known recursion formula
\[ \Gamma(z+1) = z \Gamma(z), \quad z \in \mathbb{R}^+. \]

With respect to $\hat{\beta}$, we have that, under $DGP-B$, 
\[ \hat{\beta} = \theta^{-1} \left[ T \sum \tilde{y}_i - \sum t \sum y_i \right] \]
\[ = \theta^{-1} \left[ T \sum \tilde{y}_i - \sum t \sum \bar{y}_i \right] + \theta^{-1} \frac{\mu}{\Gamma(d)} \left[ T \sum i \tilde{\xi}_i - \sum i \bar{\tilde{\xi}}_i \right]. \]

When $d \neq 1$, $T \sum i \tilde{\xi}_i - \sum i \bar{\tilde{\xi}}_i \neq 0$. Therefore, using again Lemma 1, Lemma 2 and the CMT we get
\[ T^{-d} \hat{\beta} = (T^{-d} \theta)^{-1} \left[ T^{-1/2} \left[ T^{-3/2-d} \sum \tilde{y}_i - T^{-2-d} \sum i \tilde{y}_i, T^{-2} \sum t \right] \right. \]
\[ + (T^{-d} \theta)^{-1} \frac{\mu}{\Gamma(d)} \left[ T^{-2-d} \sum i \tilde{\xi}_i - T^{-2-d} \sum i \bar{\tilde{\xi}}_i \right] \]
\[ \left. \quad \rightarrow \frac{\mu}{\Gamma(d)} \frac{6}{(d+1)(d+2)} = \frac{2(1-d)}{\Gamma(d)} \frac{2(1-d)}{\Gamma(d+3)} \right]. \]

As regards the coefficient of determination given that
\[ \sum \tilde{\alpha}^2 = \sum (y_i - \bar{y})^2 - \hat{\beta} \sum (y_i - \bar{y})(t - \bar{t}), \]

then, under $DGP-B$, if $d \neq 1$, from (24), (26) and the CMT it follows that

26
\[ T^{-1-2d} \sum \hat{u}_i^2 = T^{-1-2d} \sum (y_t - \bar{y})^2 - T^{-1-d} \hat{\beta} T^{-2-d} \sum (y_t - \bar{y})(t - \bar{t}) \]

\[ \frac{\mu}{\Gamma^2(d)(d+1)^3(2d+1)} - \frac{3\mu^2}{\Gamma^2(d)(d+1)^3(d+2)^2} \]

\[ = \frac{\mu^2(d-1)^2}{\Gamma^2(d+3)(2d+1)}. \]

Hence, using (24), (A11) and the CMT yields

\[ R^2 = 1 - \frac{T^{-1-2d} \sum \hat{u}_i^2}{T^{-1-2d} \sum (y_t - \bar{y})^2} \xrightarrow{P} \frac{6d - 3}{(d+2)^2}. \]

With respect to the t-ratios, using (A1), (A2), (28), (A11) and the CMT, it is direct to show that

\[ T^{-1} t_{\beta}^2 = \frac{T^{-2-2d} \hat{\beta}^2}{T^{-2-2d-1} \sum \hat{u}_i^2} \xrightarrow{P} \frac{d(6d + 3)^{1/2}}{d - 1} \]

and

\[ T^{-1} t_{\alpha}^2 = \frac{T^{-2d} \hat{\alpha}^2}{T^{-2d-1} \sum \hat{u}_i^2 T^{-1} \sum t^2} \xrightarrow{P} \frac{(2d + 1)^{1/2}}{d}. \]

Finally, with respect to the DW statistic, under DGP-B, when \( d \geq 3/2 \),

\[ \Delta y_i = \Delta \xi_i + \Delta \tilde{y}_i = \frac{\mu}{\Gamma(d)} t^{d-1} + \Delta \tilde{y}_i. \]

Consequently, using (A1), Lemma 1 and the CMT, it follows that

\[ T^{-d} \sum \Delta y_i = \frac{\mu}{\Gamma(d)} T^{-d} \sum t^{d-1} + T^{-1/2-d} \sum \Delta \tilde{y}_i \xrightarrow{P} \frac{\mu}{\Gamma(d+1)} \]

and

\[ T^{-1-2d} (\Delta y_i)^2 = \frac{\mu^2}{\Gamma^2(d)} T^{-1-2d} \sum t^{2d-2} + T^{-1} T^{-2d+2} \sum (\Delta \tilde{y}_i)^2 \]

27
\[+2 \mu \frac{2}{\Gamma(d)} T^{-1/2} T^{v_{2-2d}} \sum t^{d-1} \Delta \tilde{y}_i \sim \frac{\mu^2}{\Gamma^2(d)(2d-1)} ,\]

which implies, using (A11) and the CMT, that

\[T^2 DW \sim \frac{(2d+1)}{\Gamma^2(d)(2d-1)(d-1)^2} \left[ \Gamma^2(d+3) \right.\]
\[\left. + 36 \Gamma^2(d+1)(2d-1) - 12 \Gamma(d)(d+3)(2d-1) \right].\]

When \(d < 1\),

\[T^{-1} \sum \Delta y_i = \frac{\mu}{\Gamma(d)} T^{d-1} T^{-d} \sum t^{d-1} + T^{-1} \sum \Delta \tilde{y}_i \sim \epsilon \to 0\]

and

\[T^{-1} \sum (\Delta y_i)^2 = \frac{\mu^2}{\Gamma^2(d)} T^{2d-2} T^{-2d} \sum t^{2d-2} + T^{-1} \sum (\Delta \tilde{y}_i)^2\]

\[+2 \frac{\mu}{\Gamma(d)} T^{-1} \sum t^{d-1} \Delta \tilde{y}_i \sim \text{var} (\Delta \tilde{y}_i) = \frac{\sigma^2 \Gamma(3-2d)}{\Gamma(2-d)}.\]

Hence,

\[T^{2d} DW = \frac{T^{-1} \sum (\Delta u_i)^2}{T^{-2d-1} \sum u_i^2} = \frac{T^{-1} \sum (\Delta \tilde{y}_i)^2}{T^{-2d-1} \sum u_i^2} + o(1)\]

\[\sim \frac{\sigma^2 \Gamma(3-2d) \Gamma^2(d+3)(2d+1)}{\mu^2 \Gamma(2-d)(d-1)^2}.\]

Similarly, when \(1 < d < 3/2\),

\[T^{-d} \sum \Delta y_i = \frac{\mu}{\Gamma(d)} T^{-d} \sum t^{d-1} + T^{-1} \sum \Delta \tilde{y}_i \sim \frac{\mu}{\Gamma(d+1)}\]

and

\[T^{1-2d} \sum (\Delta y_i)^2 = \frac{\mu^2}{\Gamma^2(d)} T^{1-2d} \sum t^{2d-2} + T^{2-2d} T^{-1} \sum (\Delta \tilde{y}_i)^2\]

\[+2 \frac{\mu}{\Gamma(d)} T^{2-2d} T^{-1} \sum t^{d-1} \Delta \tilde{y}_i \sim \frac{\mu^2}{\Gamma^2(d)(2d-1)},\]
so that

\[ T^{1-2d} \sum (\Delta u_i)^2 = T^{1-2d} \sum (\Delta \tilde{y}_i)^2 + T^{2-2d} \tilde{\beta}^2 - 2T^{1-d} \tilde{\beta} T^{-d} \sum \Delta \tilde{y}_i. \]

Consequently, collecting the above results we obtain

\[
\lim_{n \to \infty} T^2 D W \to \frac{(2d + 1)^2}{(d - 1)^2} \left[ \frac{\Gamma^2(d + 3) + 36 \Gamma^2(d + 1)(2d - 1) - 12 \Gamma(d + 3) \Gamma(d)(2d - 1)}{\Gamma^2(d)(2d - 1)} \right].
\]

Finally, it is straightforward to prove that, if \( d = 1 \), the OLS statistics have the same limiting distributions as in Theorem 3 under DGP-A. For instance, when \( d = 1 \), \( \tilde{\xi}_i = t \) and \( \Gamma(d) = \Gamma(1) = 1 \). Hence,

\[
\sum t^3 \sum \tilde{\xi}_i - \sum t \tilde{\xi}_i \sum t = \sum t^3 \sum t - \sum t^2 \sum t = 0
\]

and

\[
\frac{T \sum t \tilde{\xi}_i - \sum t \sum \tilde{\xi}_i}{\theta \Gamma(d)} = \frac{T \sum t^3 - (\sum t)^2}{T \sum t^2 - (\sum t)^2} = 1,
\]

so that

\[
\hat{\alpha} = \theta^{-1} \left[ \sum t^3 \sum \tilde{y}_i - \sum t \sum \tilde{y}_i \sum t \right]
\]

and

\[
\hat{\beta} - \mu = \theta^{-1} \left[ T \sum \tilde{y}_i - \sum t \sum \tilde{y}_i \right],
\]

as in the driftless case, i.e., under DGP-A. Consequently, when \( d = 1 \), Corollary 1 applies and, then, using expression (16) in Lemma 2, we obtain \( T^{-1/2} \hat{\alpha} \equiv N(0, \Theta^\alpha) \) and \( T^{1/2}(\hat{\beta} - \mu) \equiv N(0, \Theta^\theta) \), where \( \Theta^\alpha = 2 \sigma^2 / 15 \) and \( \Theta^\theta = 6 \sigma^2 / 15 \).

Q.E.D.
REFERENCES


