Distribution of the Durbin–Watson Statistic in Near Integrated Processes

Uwe Hassler and Mehdi Hosseinkouchack

Abstract This paper analyzes the Durbin–Watson (DW) statistic for near-integrated processes. Using the Fredholm approach the limiting characteristic function of DW is derived, in particular focusing on the effect of a "large initial condition" growing with the sample size. Random and deterministic initial conditions are distinguished. We document the asymptotic local power of DW when testing for integration.

1 Introduction

In a series of papers Durbin and Watson (1950, 1951, 1971) developed a celebrated test to detect serial correlation of order one. The corresponding Durbin–Watson (DW) statistic was proposed by Sargan and Bhargava (1983) in order to test the null hypothesis of a random walk,

$$y_t = \rho y_{t-1} + \varepsilon_t \,,\ t = 1, \ldots, T \,, \quad H_0 : \rho = 1 \,.$$

Bhargava (1986) established that the DW statistic for a random walk is uniformly most powerful against the alternative of a stationary AR(1) process. Local power of DW was investigated by Hisamatsu and Maekawa (1994) following the technique by White (1958). Hisamatsu and Maekawa (1994) worked under the following assumptions: (1) a model without intercept like above, (2) a zero (or at least negligible) starting value y_0, (3) serially independent innovations $\{\varepsilon_t\}$, and (4) homoskedastic innovations. Nabeya and Tanaka (1988, 1990a) and Tanaka (1990, 1996) introduced the so-called Fredholm approach to econometrics. Using this approach, Nabeya and Tanaka (1990b) investigated the local power of DW under a more realistic setup. They allowed for an intercept and also a linear trend in the model and for errors displaying serial correlation and heteroskedasticity of a certain

U. Hassler (✉) • M. Hosseinkouchack
Goethe-Universität Frankfurt, Grueneburgplatz 1, 60323 Frankfurt, Germany
e-mail: hassler@wiwi.uni-frankfurt.de; hosseinkouchack@wiwi.uni-frankfurt.de

degree. Here, we go one step beyond and relax the zero starting value assumption. To this end we adopt the Fredholm approach as well.[1]

In particular, we obtain the limiting characteristic function of the DW statistic for near-integrated processes driven by serially correlated and heteroskedastic processes with the primary focus to reveal the effect of a "large initial condition" growing with $\sqrt{T}$, where T is the sample size. This starting value assumption has been picked out as a central theme by Müller and Elliott (2003), see also Harvey et al. (2009) for a recent discussion.

The rest of the paper is organized as follows. Section 2 becomes precise on the notation and the assumptions. The underlying Fredholm approach is presented and discussed in Sect. 3. Section 4 contains the limiting results. Section 5 illustrates the power function of DW. A summary concludes the paper. Proofs are relegated to the Appendix.

2 Notation and Assumptions

Before becoming precise on our assumptions we fix some standard notation. Let $\mathbb{I}(\cdot)$ denote the usual indicator function, while I_T stands for the identity matrix of size T. All integrals are from 0 to 1 if not indicated otherwise, and $w(\cdot)$ indicates a Wiener process or standard Brownian motion.

We assume that the time series observations $\{y_t\}$ are generated from

$$y_t = x_t + \eta_t \,, \quad \eta_t = \rho\,\eta_{t-1} + u_t \,, \quad t = 1, \ldots, T, \tag{1}$$

where x_t is the deterministic component of y_t which we restrict to be a constant or a linear time trend. We maintain that the following conventional assumption governs the behavior of the stochastic process $\{u_t\}$.

Assumption 1 The sequence $\{u_t\}$ is generated by

$$u_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} \quad \text{with } \alpha_0 = 1 \,, \quad \sum_{j=0}^{\infty} |\alpha_j| < \infty \,, \quad a := \sum_{j=0}^{\infty} \alpha_j \neq 0 \,,$$

while $\{\varepsilon_t\}$ is a sequence of martingale differences with

$$\text{plim}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left(\varepsilon_t^2 | F_{t-1}\right) = \sigma^2 \,, \; \text{plim}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left(\varepsilon_t^2 \mathbb{I}\left(|\varepsilon_t| > \sqrt{T}\gamma\right) | F_{t-1}\right) = 0,$$

---

[1]Further applications of this approach in a similar context are by Nabeya (2000) to seasonal unit roots, and by Kurozumi (2002) and Presno and López (2003) to stationarity testing.

for any $\gamma > 0$, $0 < \sigma^2 < \infty$, and that F_t is the σ-algebra generated by the ε_s, $s \leq t$. Also we let $\sigma_u^2 := \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E\left(u_t^2\right)$ and $\omega_u^2 := \lim_{T \to \infty} T^{-1} E\left(\sum_{t=1}^{T} u_t\right)^2 = \sigma^2 a^2$.

Assumption 2 In model (1) we allow ρ to be time-dependent and set $\rho_T = 1 - \frac{c}{T}$ with $c > 0$, where the null distribution is covered as limiting case ($c \to 0$).

Assumption 3 For the starting value $\eta_0 = \xi$ we assume: a) $\xi = o_p(\sqrt{T})$ ("small starting value"), where ξ may be random or deterministic; b) $\xi = \delta \sqrt{\omega_u^2 / \left(1 - \rho_T^2\right)}$ where $\delta \sim N\left[\mu_\delta \mathbb{I}\left(\sigma_\delta^2 = 0\right), \sigma_\delta^2\right]$ and independent from $\{u_t\}$ ("large starting value").

Assumption 1 allows for heteroskedasticity of $\{u_t\}$. If we assume homoskedasticity, $E\left(\varepsilon_t^2 | F_{t-1}\right) = \sigma^2$, then $\{u_t\}$ is stationary. Under Assumption 1 an invariance principle is guaranteed (see, for example, Phillips and Solo 1992). By Assumption 2, the process $\{\eta_t\}$ is near-integrated as defined by Phillips (1987). The initial condition under Assumption 3a) will be negligible as $T \to \infty$. The effect of initial condition under Assumption 3b) will not be negligible and the specification of δ, compactly, covers both random and fixed cases depending on the value of σ_δ.

We distinguish the model with demeaning from that with detrending using μ and τ for the corresponding cases. The test statistics $DW_{j,T}$ ($j = \mu, \tau$) are given by

$$DW_{j,T} = \frac{T}{\hat{\omega}^2} \frac{\sum_{t=2}^{T} \left(\hat{\eta}_t^j - \hat{\eta}_{t-1}^j\right)^2}{\sum_{t=1}^{T} \left(\hat{\eta}_t^j\right)^2}, \tag{2}$$

where $\hat{\eta}_t^j$ are OLS residuals calculated from (1) with $\hat{\omega}^2$ being a consistent estimator of σ_u^2 / ω_u^2 (see Hamilton 1994, Sect. 10.5, for further discussions).

$DW_{j,T}$ rejects a null hypothesis of $\rho = 1$ in favor of $\rho < 1$ for too large values. The critical values are typically taken from the limiting distributions, DW_j, which are characterized explicitly further down,

$$DW_{j,T} \xrightarrow{D} DW_j, \quad j = \mu, \tau,$$

where $\xrightarrow{D}$ denotes convergence in distribution as $T \to \infty$.

3 Fredholm Approach

The Fredholm approach relies on expressing limiting distributions as double Stieltjes integrals over a positive definite kernel $K(s, t)$ that is symmetric and continuous on $[0, 1] \times [0, 1]$.[2] Given a kernel, one defines a type I Fredholm integral equation as

$$f(t) = \lambda \int K(s, t) f(s) \, ds \,,$$

with eigenvalue λ and eigenfunction f. The corresponding Fredholm determinant (FD) of the kernel is defined as (see Tanaka 1990, Eq. (24))

$$D(\lambda) = \lim_{T \to \infty} \det \left(I_T - \frac{\lambda}{T} \left[K \left(\frac{j}{T}, \frac{k}{T} \right) \right]_{j,k=1,\ldots,T} \right). \tag{3}$$

Further, the so-called resolvent $\Gamma(s, t; \lambda)$ of the kernel (see Tanaka 1990, Eq. (25)) is

$$\Gamma(s, t; \lambda) = K(s, t) + \lambda \int \Gamma(s, u; \lambda) K(u, t) \, du \,. \tag{4}$$

Those are the ingredients used to determine limiting characteristic functions following Nabeya and Tanaka (1990a) and more generally Tanaka (1990).[3]

Let DW_j ($j = \mu, \tau$) represent the limit of $DW_{j,T}$. DW_j^{-1} can be written as $S_X = \int \{X(t) + n(t)\}^2 \, dt$ for some stochastic process $X(t)$ and an integrable function $n(t)$. Tanaka (1996, Theorem 5.9, p. 164) gives the characteristic function of random variables such as S_X summarized in the following lemma.

Lemma 1 *The characteristic function of*

$$S_X = \int [X(t) + n(t)]^2 \, dt \tag{5}$$

for a continuous function $n(t)$ is given by

$$E e^{i\theta S_X} = [D(2i\theta)]^{-1/2} \exp \left[i\theta \int n^2(t) \, dt - 2\theta^2 \int h(t) n(t) \, dt \right], \tag{6}$$

---

[2]Nabeya (2000, 2001) extend the Fredholm approach to cover a class of discontinuous kernels.

[3]For a more extensive discussion of integral equations of the Fredholm type, we refer to Hochstadt (1973) and Kondo (1991).

where h (t) is the solution of the following type II Fredholm integral equation

$$h\,(t) = m\,(t) + \lambda \int K\,(s,t)\,h\,(s)\,ds, \tag{7}$$

evaluated at $\lambda = 2i\theta$, $K\,(s,t)$ *is the covariance of* $X\,(t)$, *and* $m\,(t) = \int K\,(s,t)\,n\,(s)\,ds$.

Remark 1 Although Tanaka (1996, Theorem 5.9) presents this lemma for the covariance $K\,(s,t)$ of $X\,(t)$, his exposition generalizes for a more general case. Adopting his arguments one can see that Lemma 1 essentially relies on an orthogonal decomposition of $X\,(t)$, which does not necessarily have to be based on the covariance of $X\,(t)$. In particular, if there exists a symmetric and continuous function $C\,(s,t)$ such that

$$\int X^2\,(t)\,dt = \int \int C\,(s,t)\,dw\,(s)\,dw\,(t)\,,$$

then we may find $h\,(t)$ as in (7) by solving $h\,(t) = \int C\,(s,t)\,n\,(s)\,ds + \lambda \int C\,(s,t)\,h\,(s)\,ds$. This may in some cases shorten and simplify the derivations. As will be seen in the Appendix, we may find the characteristic function of DW_μ resorting to this remark.

4 Characteristic Functions

In Proposition 2 below, we will give expressions for the characteristic functions of DW_j^{-1} ($j = \mu, \tau$) employing Lemma 1. To that end we use that $DW_j^{-1} = \int \left[X_j\,(t) + n_j\,(t)\right]^2 dt$ for some integrable functions $n_j\,(t)$ and demeaned and detrended Ornstein–Uhlenbeck processes $X_\mu\,(t)$ and $X_\tau\,(t)$, respectively, see Lemma 2 in the Appendix. $n_\mu\,(t)$ and $n_\tau\,(t)$ capture the effect of the initial condition whose exact forms are given in the Appendix. Hence, we are left with deriving the covariance functions of $X_\mu\,(t)$ and $X_\tau\,(t)$. We provide these rather straightforward results in the following proposition.

Proposition 1 *The covariance functions of* $X_\mu\,(t)$ *and* $X_\tau\,(t)$ *from* $DW_j^{-1} = \int \left[X_j\,(t) + n_j\,(t)\right]^2 dt$ *are*

$$K_\mu\,(s,t) = K_1\,(s,t) - g\,(t) - g\,(s) + \omega_0,$$

$$K_\tau\,(s,t) = K_1\,(s,t) + \sum_{k=1}^{8} \phi_k\,(s)\,\psi_k\,(t)\,,$$

where $K_1\,(s,t) = \frac{1}{2c}\left[e^{-c|s-t|} - e^{-c(s+t)}\right]$ *and the functions* $\phi_k\,(s)$, $\psi_k\,(s)$, $k = 1, 2, \ldots, 8$, *and* $g\,(s)$ *and the constant* ω_0 *can be found in the Appendix.*

The problem dealt with in this paper technically translates into finding $h_j(t)$ as outlined in Lemma 1 for $K_j(s,t)$, $j = \mu, \tau$, i.e. finding $h_j(t)$ that solves a type II Fredholm integral equation of the form (7). Solving a Fredholm integral equation in general requires the knowledge of the FD and the resolvent of the associated kernel. The FD for $K_j(s,t)$ ($j = \mu, \tau$) are known (Nabeya and Tanaka 1990b),but not the resolvents. Finding the resolvent is in general tedious, let alone the difficulties one might face finding $h_j(t)$ once FD and the resolvent are known. To overcome these difficulties, we suggest a different approach to find $h_j(t)$ which follows.[4] As we see from Proposition 1 kernels of the integral equations considered here are of the following general form

$$K(s,t) = K_1(s,t) + \sum_{k=1}^{n} \phi_k(s)\,\psi_k(t).$$

Thus to solve for $h(t)$ in

$$h(t) = m(t) + \lambda \int K(s,t)\,h(s)\,ds, \tag{8}$$

we let $\upsilon = \sqrt{\lambda - c^2}$ and observe that (8) is equivalent to

$$h''(t) + \upsilon^2 h(t) = m''(t) - c^2 m(t) + \lambda \sum_{k=1}^{n} b_k \left[\psi_k''(t) - c^2 \psi_k(t)\right], \tag{9}$$

with the following boundary conditions

$$h(0) = m(0) + \lambda \sum_{k=1}^{n} b_k \psi_k(0), \tag{10}$$

$$h'(0) = m'(0) + \lambda \sum_{k=1}^{n} b_k \psi_k'(0) + \lambda b_{n+1}, \tag{11}$$

where

$$b_k = \int \phi_k(s)\,h(s)\,ds, \; k = 1, 2, \ldots, n \text{ and } b_{n+1} = \int e^{-cs} h(s)\,ds. \tag{12}$$

The solution to (9) can now be written as

$$h(t) = c_1 \cos \upsilon t + c_2 \sin \upsilon t + g_m(t) + \sum_{k=1}^{n} b_k g_k(t) \tag{13}$$

where $g_k(t)$, $k = 1, 2, \ldots, n$, are special solutions to the following differential equations

---

[4]Nabeya and Tanaka (1988) use a similar method to find the FD of kernel of the general form $K(s,t) = \min(s,t) + \sum_{k=1}^{n} \phi_k(s)\,\psi_k(t)$. See page 148 of Tanaka (1996) for some examples.

$$g_k''(t) + \upsilon^2 g_k(t) = \lambda \left[\psi_k''(t) - c^2 \psi_k(t) \right], k = 1, 2, \ldots, n,$$

and $g_m(t)$ is a special solution of

$$g_m''(t) + \upsilon^2 g_m(t) = m''(t) - c^2 m(t).$$

Using the boundary conditions (10) and (11) together with equations from (12) the unknowns c_1, c_2, b_1, b_2, ..., b_{n+1} are found giving an explicit form for (13). The solution for $h(t)$ can then be used for the purposes of Lemma 1. It is important to note that if we replace $K_1(s, t)$ with any other nondegenerate kernel, the boundary conditions (10) and (11) need to be modified accordingly.

Using the method described above we establish the following proposition containing our main results for DW_j ($j = \mu, \tau$).

Proposition 2 *For DW_j ($j = \mu, \tau$) we have under Assumptions 1, 2, and 3b)*

$$E\left[e^{i\theta DW_j^{-1}} \right] = \left[D_j(2i\theta) \right]^{-1/2} \left\{ 1 - \frac{\sigma_\delta^2}{c} \left[i\theta\Theta_j - 2\theta^2\Psi_j(\theta;c) \right] \right\}^{-1/2} \times$$

$$\exp\left\{ \frac{\mu_\delta^2 \left[i\theta\Theta_j - 2\theta^2\Psi_j(\theta;c) \right]}{2c - 2\sigma_\delta^2 \left[i\theta\Theta_j - 2\theta^2\Psi_j(\theta;c) \right]} \right\},$$

where $D_j(\lambda)$ is the FD of $K_j(s, t)$ with $\upsilon = \sqrt{\lambda - c^2}$,

$$D_\mu(\lambda) = \frac{e^{-c}}{\upsilon^4} \left[\upsilon\left(\lambda - c^3\right) \sin\upsilon - \left(c^2\upsilon^2 + 2c\lambda\right) \cos\upsilon + 2c\lambda \right],$$

$$D_\tau(\lambda) = e^{-c} \left[\left(\frac{c^5 - 4\lambda c^2}{\upsilon^4} - \frac{12\lambda(c+1)(c^2+\lambda)}{\upsilon^6} \right) \frac{\sin\upsilon}{\upsilon} \right.$$

$$\left. + \left(\frac{c^4}{\upsilon^4} + \frac{8\lambda c^3}{\upsilon^6} - \frac{48\lambda^2(c+1)}{\upsilon^8} \right) \cos\upsilon + \frac{4\lambda\upsilon^2 c^2(c+3) + 48\lambda^2(c+1)}{\upsilon^8} \right],$$

and

$$\Theta_\mu = \frac{e^{-2c}}{2c^2} \left(-1 + e^c\right)\left(c - 2e^c + ce^c + 2\right),$$

$$\Theta_\tau = \frac{e^{-c}}{c^4} \left[-4\left(-6 + c^2\right) - 8\left(3 + c^2\right)\cosh c + c\left(24 + c^2\right)\sinh c \right],$$

where $\Psi_j(\theta;c)$ for $j = \mu, \tau$ are given in the Appendix.

Remark 2 Under Assumption 3a) the limiting distributions of DW_j, $j = \mu, \tau$, are the same as the limiting distributions derived under a zero initial condition in Nabeya and Tanaka (1990b). These results are covered here when $\mu_\delta = \sigma_\delta^2 = 0$.

Remark 3 The Fredholm determinants, $D_j(\lambda)$, are taken from Nabeya and Tanaka (1990b).

Remark 4 It is possible to derive the characteristic functions using Girsanov's theorem (see Girsanov 1960) given, for example, in Chap. 4 of Tanaka (1996). Further, note that Girsanov's theorem has been tailored to statistics of the form of DW_j under Lemma 1 in Elliott and Müller (2006).

5 Power Calculation

To calculate the asymptotic power function of $DW_{j,T}$ $(j = \mu, \tau)$ we need the quantiles $c_{j,1-\alpha}$ as critical values where $(j = \mu, \tau)$

$$P(DW_j > c_{j,1-\alpha}) = P\left(DW_j^{-1} < c_{j,1-\alpha}^{-1}\right) = \alpha.$$

Our graphs rely on the $\alpha = 5\%$ level with critical values $c_{\mu,0.95} = 27.35230$ and $c_{\tau,0.95} = 42.71679$ taken, up to an inversion, from Table 1 of Nabeya and Tanaka (1990b). Let $\phi(\theta; c)$ stand for the characteristic functions obtained in Proposition 2 for large initial conditions of both deterministic and random cases in a local neighborhood to the null hypothesis characterized by c. With $x = c_{j,1-\alpha}^{-1}$ we hence can compute the local power by evaluating the distribution function of DW_j^{-1} where we employ the following inversion formula given in Imhof (1961)

$$F(x; c) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{1}{\theta} \text{Im}\left[e^{-i\theta x} \phi(\theta; c)\right] d\theta. \tag{14}$$

When it comes to practical computations, Imhof's formula (14) is evaluated using a simple Simpson's rule while correcting for the possible phase shifts which arise due to the square root map over the complex domain (see Tanaka 1996, Chap. 6, for a discussion) .[5]

Figure 1 shows the local power functions of DW_j $(j = \mu, \tau)$ for a deterministic large initial condition, that is for $\sigma_\delta^2 = 0$ in Assumption 3b). As is clear from Proposition 2, the power function is symmetric around $\mu_\delta = 0$ and decreasing in μ_δ^2 for any level of the local-to-unity parameter c.

For the random case we set $\mu_\delta = 0$. Figure 2 contains graphs for $\sigma_\delta \in \{0, 0.1, \ldots, 3\}$; to keep shape conformity with the case of a large deterministic initial condition the graphs are arranged symmetrically around 0. Figure 2 shows by how much the power decreases in the variance of the initial condition.

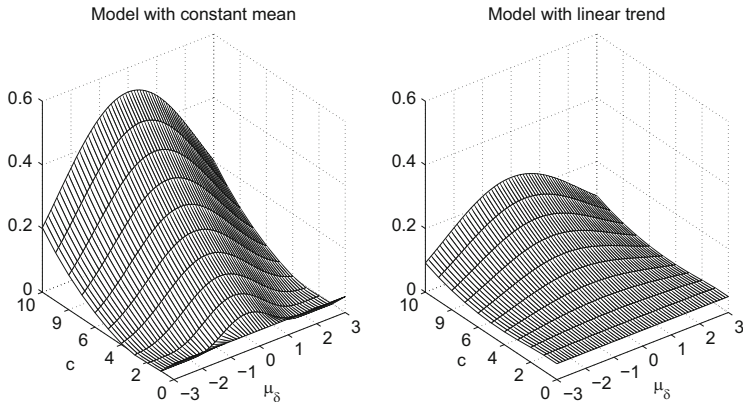[5]The inversion formula (14) is derived in Gurland (1948) and Gil-Pelaez (1951).

Fig. 1 Power function of DW_j ($j = \mu, \tau$) for a large deterministic initial condition with $\mu_\delta \in \{-3, -2.9, \ldots, 3\}$, $c \in \{0, 1, \ldots, 10\}$ and $\sigma_\delta = 0$
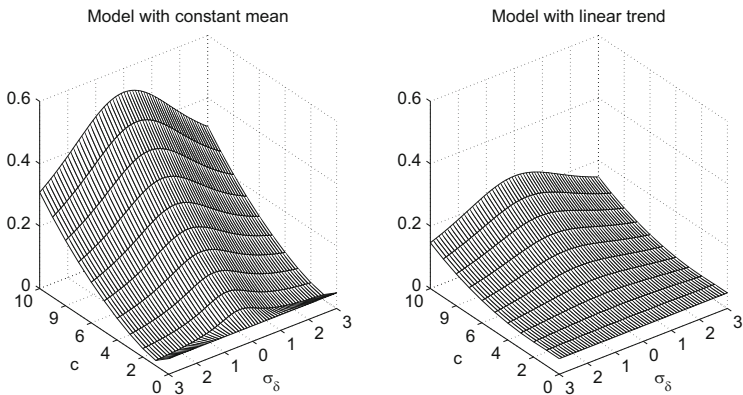

Fig. 2 Power function of DW_j ($j = \mu, \tau$) for a large random initial condition with $\sigma_\delta \in \{3, 2.9, \ldots, 0, 0.1, \ldots, 3\}$, $c \in \{0, 1, \ldots, 10\}$ and $\mu_\delta = 0$

6 Summary

We analyze the effect of a large initial condition, random or deterministic, on the local-to-unity power of the Durbin–Watson unit root test. Using the Fredholm approach the characteristic function of the limiting distributions are derived. We observe the following findings. First, the local power after detrending is considerably lower than in case of a constant mean. Second, a large initial value has a negative effect on power: the maximum power is achieved for $\mu_\delta = \sigma_\delta = 0$, which corresponds to a "small initial condition." Finally, comparing Figs. 1 and 2 one learns that deterministic and random initials values have a similar effect depending

only on the magnitude of the mean or the standard deviation, respectively, of the large initial condition.

Appendix

First we present a preliminary result. Lemma 2 contains the required limiting distributions in terms of Riemann integrals.

Lemma 2 *Let $\{y_t\}$ be generated according to (1) and satisfy Assumptions 1 and 2. It then holds for the test statistics from (2) asymptotically*

$$DW_{j,T} \xrightarrow{D} DW_j = \left(\int \{w_c^j (r)\}^2 \, dr \right)^{-1}, \ j = \mu, \ \tau,$$

where under Assumption 3b)

$$w_c^\mu (r) = w_c (r) - \int w_c (s) \, ds,$$

$$w_c^\tau (r) = w_c^\mu (r) - 12 \left(r - \frac{1}{2} \right) \int \left(s - \frac{1}{2} \right) w_c (s) \, ds,$$

with $w_c (r) = w (r)$ for $c = 0$ and

$$w_c (r) = \delta \, (e^{-cr} - 1) \, (2c)^{-1/2} + J_c (r) \ \text{ for } c > 0$$

and the standard Ornstein–Uhlenbeck process $J_c (r) = \int_0^r e^{-c(r-s)} dw (s)$.

Proof The proof is standard by using similar arguments as in Phillips (1987) and Müller and Elliott (2003).

Proof of Proposition 1

We set $\upsilon = \sqrt{\lambda - c^2}$. For DW_μ we have $w_c^\mu (s) = w_c (r) - \int w_c (s) \, ds$, thus

$$Cov \left[w_c^\mu (s) , w_c^\mu (t) \right] = Cov \left[J_c (s) - \int J_c (s) \, ds, J_c (t) - \int J_c (s) \, ds \right]$$

$$= K_1 (s, t) - \int Cov \left[J_c (s) , J_c (t) \right] ds - \int Cov \left[J_c (s) , J_c (t) \right] dt$$

$$+ \int \int Cov \left[J_c (s) , J_c (t) \right] ds$$

$$= K_1 (s, t) - g (t) - g (s) + \omega_0$$

where

$$g(t) = \frac{e^{-c(1+s)}(1 - e^{cs})(1 - 2e^c + e^{cs})}{2c^2},$$

$$\omega_0 = -\frac{3 - 2c + e^{-2c} - 4e^{-c}}{2c^3}.$$

For DW_τ we have $w_c^\tau(s) = w_c^\mu(s) - 12\left(s - \frac{1}{2}\right)\int\left(u - \frac{1}{2}\right)w_c(u)\,du$, thus

$$Cov\left[w_c^\tau(s), w_c^\tau(t)\right] = Cov\left[w_c^\mu(s), w_c^\mu(t)\right] - 12\left(t - \frac{1}{2}\right)\int\left(u - \frac{1}{2}\right)$$

$$Cov\left[J_c(s) - \int J_c(v)\,dv, J_c(u)\right]du$$

$$-12\left(s - \frac{1}{2}\right)\int\left(u - \frac{1}{2}\right)Cov\left[J_c(u), J_c(t)\right]$$

$$-\int J_c(v)\,dv\bigg]du$$

$$+144\left(s - \frac{1}{2}\right)\left(t - \frac{1}{2}\right)\int\int\left(u - \frac{1}{2}\right)\left(v - \frac{1}{2}\right)$$

$$\times Cov\left[J_c(u), J_c(v)\right]dudv$$

$$= Cov\left[w_c^\mu(s), w_c^\mu(t)\right]$$

$$-12\left(t - \frac{1}{2}\right)\int\left(u - \frac{1}{2}\right)Cov\left[J_c(s), J_c(u)\right]du$$

$$+12\left(t - \frac{1}{2}\right)\int\int\left(u - \frac{1}{2}\right)Cov\left[J_c(v), J_c(u)\right]dvdu$$

$$-12\left(s - \frac{1}{2}\right)\int\left(u - \frac{1}{2}\right)Cov\left[J_c(u), J_c(t)\right]du$$

$$+12\left(s - \frac{1}{2}\right)\int\int\left(u - \frac{1}{2}\right)Cov\left[J_c(u), J_c(v)\right]dvdu$$

$$+144\left(s - \frac{1}{2}\right)\left(t - \frac{1}{2}\right)\int\int\left(u - \frac{1}{2}\right)\left(v - \frac{1}{2}\right)$$

$$\times Cov\left[J_c(u), J_c(v)\right]dudv$$

With some calculus the desired result is obtained. In particular we have with $\phi_1(s) = -1$, $\phi_2(s) = -g(s)$, $\phi_3(s) = -3f_1(s)$, $\phi_4(s) = -3(s - 1/2)$, $\phi_5(s) = 3\omega_1$, $\phi_6(s) = 3\omega_1(s - 1/2)$, $\phi_7(s) = 6\omega_2(s - 1/2)$, $\phi_8(s) = \omega_0$, $\psi_1(t) = g(t)$,

$\psi_2(t) = \psi_6(t) = \psi_8(t) = 1$, $\psi_3(t) = \psi_5(t) = \psi_7(t) = t - 1/2$ and $\psi_4(t) = f_1(t)$ while

$$f_1(s) = \frac{e^{-c(1+s)}}{c^3} \times \left[2 + c + 2ce^c - (2+c)e^{2cs} + 2ce^{c+cs}(2s-1)\right],$$

$$\omega_1 = \frac{e^{-2c}(e^c - 1)}{c^4} \times [2 + c + (c-2)e^c],$$

$$\omega_2 = \frac{e^{-2c}}{c^5} \times \left[-3(c+2)^2 - 12c(2+c)e^c + (12c - 9c^2 + 2c^3 + 12)e^{2c}\right].$$

This completes the proof.

Proof of Proposition 2

Let $\mathcal{L}(X) = \mathcal{L}(Y)$ stand for equality in distribution of X and Y and set $A = \delta(2c)^{-1/2}$. To begin with, we do the proofs conditioning on δ. Consider first DW_μ. To shorten the proofs for DW_μ we work with the following representation for a demeaned Ornstein–Uhlenbeck process given under Theorem 3 of Nabeya and Tanaka (1990b), for their $R_1^{(2)}$ test statistic, i.e. we write

$$\mathcal{L}\left(\int \left\{J_c(r) - \int J_c(s)\,ds\right\}^2 dr\right) = \mathcal{L}\left(\int \int K_0(s,t)\,dw(t)\,dw(s)\right),$$

where $K_0(s,t) = \frac{1}{2c}\left[e^{-c|s-t|} - e^{-c(2-s-t)}\right] - \frac{1}{c^2}p(s)p(t)$ with $p(t) = 1 - e^{-c(1-t)}$. Using Lemma 2, we find that

$$n_\mu(t) = A\left(e^{-ct} - 1\right) - A\int \left(e^{-ct} - 1\right)dt.$$

For DW_μ we will be looking for $h_\mu(t)$ in

$$h_\mu(t) = m_\mu(t) + \lambda \int K_0(s,t)\,h_\mu(t)(s)\,ds, \tag{15}$$

where $m_\mu(t) = \int K_0(s,t)\,n_\mu(s)\,ds$. Equation (15) is equivalent to the following boundary condition differential equation

$$h_\mu''(t) + \upsilon^2 h_\mu(t) = m_\mu''(t) - c^2 m_\mu(t) + \lambda b_1, \tag{16}$$

with

$$h_\mu(1) = m_\mu(1) - \frac{1}{c^2}\lambda b_1 p(1), \tag{17}$$

$$h'_\mu(1) = m'_\mu(1) - \lambda e^{-c} b_2 - \frac{1}{c^2}\lambda p'(1) b_1, \tag{18}$$

where $b_1 = \int p(s) h_\mu(s) ds$ and $b_2 = \int e^{cs} h_\mu(s) ds$. Thus have

$$h_\mu(t) = c_1^\mu \cos \upsilon t + c_2^\mu \sin \upsilon t + g_\mu(t) + b_1 g_1(t),$$

where $g_\mu(t)$ is a special solution to $g''_\mu(t) + \upsilon^2 g_\mu(t) = m''_\mu(t) - c^2 m_\mu(t)$ and $g_1(t)$ is a special solution to $g''_1(t) + \upsilon^2 g_1(t) = \lambda$. Boundary conditions (17) and (18) together with $h_\mu(t)$ imply

$$c_1^\mu \cos \upsilon + c_1^\mu \sin \upsilon + \left[g_1(1) + \frac{1}{c^2}\lambda p(1)\right] b_1 = m_\mu(1) - g_\mu(1),$$

$$-c_1^\mu \upsilon \sin \upsilon + c_1^\mu \upsilon \cos \upsilon + \left[\frac{1}{c^2}\lambda p'(1) + g'_1(1)\right] b_1 + \lambda e^{-c} b_2 = m'_\mu(1) - g'_\mu(1),$$

while expressions for b_1 and b_2 imply that

$$c_1^\mu \int p(s) \cos \upsilon s ds + c_1^\mu \int p(s) \sin \upsilon s ds + b_1 \left(\int p(s) g_1(s) ds - 1\right)$$

$$= -\int p(s) g_\mu(s) ds$$

$$c_1^\mu \int e^{cs} \cos \upsilon s ds + c_1^\mu \int e^{cs} \sin \upsilon s ds + b_1 \int e^{cs} g_1(s) ds - b_2$$

$$= -\int e^{cs} g_\mu(s) ds$$

These form a system of linear equations in c_1^μ, c_2^μ, b_1, and b_2, which in turn identifies them. With some calculus we write

$$\int n_\mu(t) h_\mu(t) dt = \frac{Ae^{-2c}}{2c^2\lambda\upsilon} \times$$

$$[-A(-1 + e^c)(2 + c + (-2 + c)e^c)$$
$$+2e^c(cc_2^\mu \lambda - ce^c(c^2 c_2^\mu - c^2 c_1^\mu - (-1 + c)c_2^\mu \upsilon^2))$$
$$+2ce^c(c^3 c_2^\mu - cc_2^\mu \lambda + c_2^\mu(-1 + e^c)\lambda - c^2 c_1^\mu \upsilon)\cos \upsilon$$
$$-2ce^c(c^3 c_1^\mu - cc_1^\mu \lambda + c_1^\mu(-1 + e^c)\lambda + c^2 c_2^\mu \upsilon)\sin \upsilon].$$

Solving for c_1^μ and c_2^μ we find that they are both a multiple of A, hence

$$\Psi_\mu (\theta; c) = \frac{1}{A^2} \int n_\mu (t) \, h_\mu (t) \, dt,$$

is free of A. Now with $\Theta_\mu = \int n_\mu^2 (t) \, dt$, an application of Lemma 1 results in

$$E\left[e^{i\theta \int \{w_c^\mu (r)\}^2 dr} | \delta \right] = \left[D_\mu (2i\theta) \right]^{-1/2} \exp \left[i\theta A^2 \Theta_\mu - 2\theta^2 A^2 \Psi_\mu (\theta; c) \right].$$

As $\sqrt{2c} A = \delta \sim N\left(\mu_\delta, \sigma_\delta^2 \right)$, standard manipulations complete the proof for $j = \mu$.

Next we turn to DW_τ. Using Lemma 2 we find that

$$n_\tau (t) = A \left[\left(e^{-ct} - 1 \right) - \int \left(e^{-ct} - 1 \right) dt - 12 \, (t - 1/2) \int (t - 1/2) \left(e^{-ct} - 1 \right) dt \right].$$

Here we will be looking for $h_\tau (t)$ in the

$$h_\tau (t) = m_\tau (t) + \lambda \int K_\tau (s, t) \, h_\tau (t) \, (s) \, ds, \tag{19}$$

where $m_\tau (t) = \int K_\tau (s, t) \, n_\tau (s) \, ds$ and $K_\tau (s, t)$ is from Proposition 1. Equation (19) can be written as

$$h_\tau'' (t) + \upsilon^2 h_\tau (t) = m_\tau'' (t) - c^2 m_\tau (t) + \lambda \sum_{k=1}^{8} b_k \left[\psi_k'' (t) - c^2 \psi_k (t) \right], \tag{20}$$

with the following boundary conditions

$$h_\tau (0) = m_\tau (0) + \lambda \sum_{k=1}^{8} b_k \psi_k (0), \tag{21}$$

$$h_\tau' (0) = m_\tau' (0) + \lambda \sum_{k=1}^{8} b_k \psi_k' (0) + \lambda b_9, \tag{22}$$

where

$$b_k = \int \phi_k (s) \, h_\tau (s) \, ds, \ k = 1, \ldots, 8 \text{ and } b_9 = \int e^{-cs} h (s) \, ds. \tag{23}$$

The solution to (20) is

$$h_\tau (t) = c_1^\tau \cos \upsilon t + c_2^\tau \sin \upsilon t + g_\tau (t) + \sum_{k=1}^{8} b_k g_k (t) \tag{24}$$

where $g_k(t)$, $k = 1, 2, \ldots, 8$, are special solutions to the following differential equations

$$g_k''(t) + \upsilon^2 g_k(t) = \lambda \left[\psi_k''(t) - c^2 \psi_k(t) \right], k = 1, 2, \ldots, 8,$$

and $g_\tau(t)$ is a special solution of $g_\tau''(t) + \upsilon^2 g_\tau(t) = m_\tau''(t) - c^2 m_\tau(t)$. The solution given in (24) can be written as

$$h_\tau(t) = c_1^\tau \cos \upsilon t + c_2^\tau \sin \upsilon t + g_\tau(t) - b_1 \frac{\lambda}{\upsilon^2} - \frac{\lambda c^2}{\upsilon^2} (b_2 + b_6 + b_8) \qquad (25)$$

$$+ (b_3 + b_5 + b_7) \lambda c^2 \frac{1 - 2t}{2\upsilon^2} + b_4 \lambda \frac{1 - 2t}{2\upsilon^2}.$$

The boundary conditions in (21) and (22) imply

$$m_\tau(0) - g_\tau(0) = c_1^\tau + \frac{\lambda}{2\upsilon^2} (-2b_1 + b_4)$$

$$+ \frac{\lambda}{2} \left(\frac{c^2}{\upsilon^2} + 1 \right) (-2b_2 + b_3 + b_5 - 2b_6 + b_7 - 2b_8)$$

$$m_\tau'(0) - g_\tau'(0) = c_2^\tau \mu - \lambda b_1 g'(0) - \lambda \left(\frac{c^2}{\upsilon^2} + 1 \right) (b_3 + b_5 + b_7)$$

$$- \left(\frac{1}{\upsilon^2} + f'(0) \right) \lambda b_4 - \lambda b_5,$$

while expressions given under (23) characterize nine more equations. These equations form a system of linear equations in unknowns c_1^τ, c_2^τ, b_1, $\ldots$, b_9, which can be simply solved to fully identify (25). Let $\Theta_\tau = \int n_\tau(t)^2 dt$. Also as for the constant case we set

$$\Psi_\tau(\theta; c) = \frac{1}{A^2} \int n_\tau(t) h_\tau(t) dt,$$

whose expression is long and we do not report here. When solving this integral we see that $\Psi_\tau(\theta; c)$ is free of A. As before we apply Lemma 1 to establish the following

$$E \left[e^{i\theta \int \{w_c^\tau(r)\}^2 dr} | \delta \right] = [D_\tau(2i\theta)]^{-1/2} \exp \left[i\theta A^2 \Theta_\tau - 2\theta^2 A^2 \Psi_\tau(\theta; c) \right].$$

Now, using $E \left[e^{i\theta \int \{w_c^\tau(r)\}^2 dr} \right] = EE \left[e^{i\theta \int \{w_c^\tau(r)\}^2 dr} | \delta \right]$, standard manipulations complete the proof.

References

Bhargava, A. (1986). On the theory of testing for unit roots in observed time series. *Review of Economic Studies*, *53*, 369–384.

Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression. *I. Biometrika*, *37*, 409–428.

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. *II. Biometrika*, *38*, 159–178.

Durbin, J., & Watson, G. S. (1971). Testing for serial correlation in least squares regression. *III. Biometrika*, *58*, 1–19.

Elliott, G., & Müller, U. K. (2006). Minimizing the impact of the initial condition on testing for unit roots. *Journal of Econometrics*, *135*, 285–310.

Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika*, *38*, 481–482.

Girsanov, I. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and Its Applications*, *5*, 285–301.

Gurland, J. (1948). Inversion formulae for the distribution of ratios. *Annals of Mathematical Statistics*, *19*, 228–237.

Hamilton, J. D. (1994). *Time series analysis*. Cambridge: Cambridge University Press.

Harvey, D. I., Leybourne, S. J., & Taylor, A. M. R. (2009). Unit root testing in practice: Dealing with uncertainty over the trend and initial condition. *Econometric Theory*, *25*, 587–636.

Hisamatsu, H., & Maekawa, K. (1994). The distribution of the Durbin-Watson statistic in integrated and near-integrated models. *Journal of Econometrics*, *61*, 367–382.

Hochstadt, H. (1973). *Integral equations*. New York: Wiley.

Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, *4*, 419–426.

Kondo, J. (1991). *Integral equations*. Oxford: Clarendon Press.

Kurozumi, E. (2002). Testing for stationarity with a break. *Journal of Econometrics*, *108*, 63–99.

Müller, U. K., & Elliott, G. (2003). Tests for unit roots and the initial condition. *Econometrica*, *71*, 1269–1286.

Nabeya, S. (2000). Asymptotic distributions for unit root test statistics in nearly integrated seasonal autoregressive models. *Econometric Theory*, *16*, 200–230.

Nabeya, S. (2001). Approximation to the limiting distribution of t- and F-statistics in testing for seasonal unit roots. *Econometric Theory*, *17*, 711–737.

Nabeya, S., & Tanaka, K. (1988). Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. *Annals of Statistics*, *16*, 218–235.

Nabeya, S., & Tanaka, K. (1990a). A general approach to the limiting distribution for estimators in time series regression with nonstable autoregressive errors. *Econometrica*, *58*, 145–163.

Nabeya, S., & Tanaka, K. (1990b). Limiting power of unit-root tests in time-series regression. *Journal of Econometrics*, *46*, 247–271.

Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregressions. *Biometrika*, *74*, 535–547.

Phillips, P. C. B., & Solo, V. (1992). Asymptotics for linear processes. *The Annals of Statistics*, *20*, 971–1001.

Presno, M. J., & López, A. J. (2003). Testing for stationarity in series with a shift in the mean. A Fredholm approach. *Test*, *12*, 195–213.

Sargan, J. D., & Bhargava, A. (1983). Testing residuals from least squares regression for being generated by the gaussian random walk. *Econometrica*, *51*, 153–174.

Tanaka, K. (1990). The Fredholm approach to asymptotic inference on nonstationary and noninvertible time series models. *Econometric Theory*, *6*, 411–432.

Tanaka, K. (1996). *Time series analysis: Nonstationary and noninvertible distribution theory*. New York: Wiley.

White, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Annals of Mathematical Statistics*, *29*, 1188–1197.