THE LIMITING DISTRIBUTION OF THE LEAST SQUARES ESTIMATOR IN NEARLY INTEGRATED SEASONAL MODELS

Pierre Perron
Princeton University and C.R.D.E.

Econometric Research Program
Research Memorandum No. 354

September 1990

Financial support is acknowledged from the Social Sciences and Humanities Research Coucil of Canada, the Natural Sciences and Engineering Council of Canada and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec.

Econometric Research Program
Princeton University
204 Fisher Hall
Princeton, NJ 08544-1021, USA

ABSTRACT

We consider the least squares estimator of the autoregressive parameter in a nearly integrated seasonal model. Building on the study by Chan (1989), who obtained the limiting distribution, we derive a closed form expression for the appropriate limiting joint moment-generating function. We use this function to tabulate percentage points of the asymptotic distribution for various seasonal periods via numerical integration. The results are extended by deriving an asymptotic expansion to order $O_p(T^{-1})$ whose percentage points are also obtained by numerically integrating the appropriate limiting joint moment-generating function. The adequacy of the approximation to the finite sample distribution is discussed.

Key words : Wiener process, unit roots, nonstationarity, numerical integration.

Address for Correspondence :

Pierre Perron
Department of Economics
Princeton University
Princeton, NJ, 08544–1021

1. INTRODUCTION

Consider a stochastic process $\{y_t\}$, of which a sample of size $(T + d)$ realizations is available, generated by the following seasonal autoregressive model :

$$y_t = \alpha_d y_{t-d} + u_t \ ; \qquad (t = 1, ..., T) \qquad (1)$$

where $\{u_t\}$ is a sequence of independent and identically distributed random variable with mean zero and finite variance σ^2. We shall also assume that y_i $(i = 0, ..., -d+1)$ are fixed constants. When $\alpha_d = 1$, (1) characterizes a seasonal random walk model. Models of this type play an important role in time series model building for business and economics data when a seasonal component is present (see Box and Jenkins (1976)). The presence of a seasonal root on the unit circle means that one should seasonally difference the data. Such a process has been considered by Latour and Roy (1987) in the more general context of a seasonal autoregressive integrated moving average model.

Denote by $\hat{\alpha}_d$ the least–squares estimator of α_d. The asymptotic distribution of $\hat{\alpha}_d$ has been studied by Dickey, Hasza and Fuller (1984) in the case where $\alpha_d = 1$. They show that the asymptotic distribution in this integrated case is not normal in contrast to the case where $|\alpha_d| < 1$. Additional results on the behavior of the sample autocovariances in that case can be found in Latour (1986) and Latour and Roy (1987). As documented in Dickey, Hasza and Fuller (1984), the asymptotic distribution of $\hat{\alpha}_d$ is an adequate approximation to the exact distribution of $\hat{\alpha}_d$ when $\alpha_d = 1$. However, when α_d is close to but not equal to one, the asymptotic normal distribution is a poor guide to the exact distribution even for quite large sample sizes. Such a feature is well documented in the non–seasonal case where $d = 1$ (see Evans and Savin (1981) and Perron (1989), among others). The basic reason for this poor approximation is the sharp discontinuity in the asymptotic distribution at the point $\alpha_d = 1$.

This feature has led to the development of an asymptotic framework that present no such discontinuities at the point $\alpha_d = 1$. Models proposed by Bobkoski (1983), Chan and Wei (1987) and Phillips (1987a) consider the presence of a root close to, but not necessarily equal, to unity. These are called nearly integrated processes and are characterized by an autoregressive parameter defined by :

$$\alpha_d = \exp(cd/T) \approx 1 + cd/T \,. \tag{2}$$

Here c, a real valued constant, is a measure of the deviation from the unit root case. When c is negative $\{y_t\}$ is said to be (locally) stationary and when c is positive it is (locally) explosive. When c = 0, we recover the standard seasonal random walk model with $\alpha_d = 1$. The asymptotic distribution of the least–squares estimator $\hat{\alpha}_d$ under model (1) and (2) has been derived by Chan (1989). The way α_d is specified in (2) is, however, different from that of Chan (1989) who specifies $\alpha_d = 1 + c/T$. We make this modification for two reasons. The first is that it is easier and more natural to make comparisons of processes with different values of d keeping a fixed α. To see this, invert (2) to obtain $c = (T/d)\ln(\alpha)$. For a fixed α and sample size T, c decreases as d increases. This is intuitively appealing since as d increases the data are sampled more frequently. For instance, a process sampled at monthly interval with coefficient $\alpha^* < 1$, say, exhibit stronger correlation than a process sampled at quarterly interval with the same coefficient α^*. Hence the former case should be associated with a process closer to the boundary of the nonstationary region, i.e. a value of c closer to zero. These features are present in the specification (2) as c decreases with d for given fixed α and T.

Secondly, we prefer the specification (2) since all the results of Chan (1989) hold under this specification, but not under the specification $\alpha = 1 + c/T$, as claimed in that paper. Upon correction for this error, the source of which we discuss below, the limiting distribution of $\hat{\alpha}_d$ when $\{y_t\}$ is generated by (1) and (2) can easily be obtained from Theorem 1 of Chan (1989). We state this result in the following proposition.

PROPOSITION 1 (Chan (1989)) : *Let $\{y_t\}$ be generated by (1) and (2) and assume that the sequence $\{u_t\}$ is a martingale sequence with respect to an increasing sequence of sigma fields $\{\mathcal{F}_t\}$ satisfying the following assumptions :*

(i) $T^{-1}\Sigma_1^T E(u_t^2 \mid \mathcal{F}_{t-1}) \to \sigma^2$ *in probability, and*

(ii) for all $\delta > 0$, $T^{-1}\Sigma_1^T E\left\{u_t^2 I(|u_t| > T^{1/2}\delta)|\mathcal{F}_{t-1}\right\} \to 0$

in probability, where $I(A)$ *denotes the indicator function of the set* A. *Then :*

$$T(\hat{\alpha}_d - \alpha_d) \Rightarrow Z(c,d) \equiv \mathrm{d}A(c,d)/B(c,d)\,, \tag{3}$$

where $\Rightarrow$ *denotes weak convergence in distribution, and*

$$A(c,d) \equiv \Sigma_{i=1}^d A_i(c) \equiv \Sigma_{i=1}^d \int_0^1 J_{i,c}(r)dW_i(r)\,, \tag{4}$$

$$B(c,d) \equiv \Sigma_{i=1}^d B_i(c) \equiv \Sigma_{i=1}^d \int_0^1 J_{i,c}(r)^2 dr\,, \tag{5}$$

with $J_{i,c}(r)$ *an Ornstein–Uhlenbeck process defined by :*

$$dJ_{i,c}(r) = cJ_{i,c}(r)dr + dW_i(r)\ ,\ J_{i,c}(0) = 0 \ \text{ for } i = 1, \ldots, d\,;$$

and $W_1(r), \ldots, W_d(r)$ *are independent Wiener processes.*

The error in the proof by Chan (1989) arises on p. 282 where it is claimed that Lemma 1(i) can be applied directly to the process Y_{i+td} to obtain the limiting distributions stated in Lemma 2. However, his Lemma 1 applies to a process for which a sample of size T (n in the notation of Chan (1989)) is available whereas there are m (=T/d) observations on Y_{i+td}. Hence, the noncentrality parameter γ (in Chan's notation) needs to be changed to γ/d in the statement of Lemma 2. The same error carries over to the statement of his Theorem 1. However, all of Chan's results are correct if α_d is specified as in (2) instead of $\alpha_d = 1 + c/T$ (except Lemma 2(ii) which should read as $n^{-1}Y_{n-i}^2 \Rightarrow d^{-1}W_i^2(B_\gamma(1))$ (or $T^{-1}Y_{T-i}^2 \Rightarrow d^{-1}J_c(1)^2$ in our notation). As stated above we prefer the specification (2) and shall use it in the rest of this paper.

As shown in Proposition 1, the limiting distribution of $T(\hat{\alpha}_d - \alpha_d)$ depends on the parameters c and d. A first feature of interest is that the limiting distribution is continuous with respect to the parameter c. Hence, this nearly integrated framework indeed provides a limiting representation that does not have the inherent discontinuity at $c = 0$ ($\alpha_d = 1$) present in the standard asymptotic framework with α_d fixed. One can therefore hope for a better approximation to the finite sample distribution of $\hat{\alpha}_d$ when α_d is in the vicinity of one. Also of interest is the fact that when $d = 1$, (3) reduces to the limiting distribution of the least–squares estimator in a nearly integrated first–order autoregressive process as studied by Bobkoski (1983), Chan and Wei (1987), Phillips (1987a), Perron (1989) and Chan (1988). If, in addition, $c = 0$, then (3) reduces to $(1/2)(W(1)^2 - 1)/\int_0^1 W(r)dr$, the limiting distribution of the least–squares estimator in a random walk model as studied by White (1958) and Evans and Savin (1981).

The aim of this paper is to provide percentage points of the limiting distribution given by (3) for various values of c and d. Chan (1988, 1989) suggests a method based on simulations of the Wiener processes (using an infinite series representation of the terms in (4) and (5)). We, however, adopt a different approach that provides exact values. We first derive the exact joint moment–generating function of $\{dA(c,d), B(c,d)\}$. Distributional quantities of interest can then be obtained using standard inversion formulae and numerical integration. The method is similar to that used in Perron (1989). This exercise is performed in Section 2. In the case where $c = 0$ (the seasonal random walk), our results are compared to those of Dickey, Hasza and Fuller (1984).

Section 3 considers a generalization of Chan's (1989) result by deriving the asymptotic expansion of the distribution of $T(\hat{\alpha}_d - \alpha_d)$ to order $O_p(T^{-1})$. Again an expression for the appropriate joint moment–generating function is derived and used to tabulate, via numerical integration, various percentage points. Section 4 contains concluding remarks and an appendix gives some technical derivations.

2. THE LIMITING DISTRIBUTION OF $T(\hat{\alpha}_d - \alpha_d)$

In this Section we analyze the joint moment–generating function of $\{dA(c,d), B(c,d)\}$ for arbitrary real values of c and positive integer values for d. Denote this joint moment–generating function by $M_{c,d}(v,u)$, which is defined by :

$$M_{c,d}(v,u) \equiv E[\exp(dvA(c,d) + uB(c,d))] \,. \tag{6}$$

Using (4) and (5) we have :

$$\begin{aligned} M_{c,d}(v,u) &= E[\exp(dv\Sigma_{i=1}^{d} A_i(c) + u\Sigma_{i=1}^{d} B_i(c))] \\ &= E[\exp(\Sigma_{i=1}^{d}\{dvA_i(c) + uB_i(c)\})] \,. \end{aligned} \tag{7}$$

Given the independence of the Wiener processes, $W_1(r)$, ..., $W_d(r)$, defining each pair $\{A_i(c), B_i(c)\}$, (7) reduces to :

$$M_{c,d}(v,u) = \Pi_{i=1}^{d} E[\exp(dvA_i(c) + uB_i(c))] \,. \tag{8}$$

Using the result of Phillips (1987a, eq. (A.1) corrected for a misprint) or Perron (1990a, Theorem 2) we have

$$E[\exp(dvA_i(c) + uB_i(c))] = \phi_c(dv,u)^{1/2} \,, \tag{9}$$

where

$$\phi_c(x,u) \equiv 2\lambda\exp(-(x + c))/[(\lambda + (x + c))\exp(-\lambda) + (\lambda - (x + c))\exp(\lambda)] \,, \tag{10}$$

with $\lambda = (c^2 + 2cx - 2u)^{1/2}$.

Using (8) through (10), it is easy to obtain the joint moment–generating function of $\{dA(c,d), B(c,d)\}$. Our result is summarized in the next theorem.

THEOREM 1 : *Let $\phi_c(x,u)$ be defined by (10), then $M_{c,d}(v,u)$, the joint moment–generating function of $\{dA(c,d), B(c,d)\}$, is given by :*

$$M_{c,d}(v,u) = \phi_c(dv,u)^{d/2}.$$

The moment–generating function stated in Theorem 1 allows us to compute the cumulative distribution function, as well as other distributional quantities, using numerical integration. Denote the joint characteristic function of {dA(c,d), B(c,d)} by $cf_{c,d}(v,u)$. With i denoting the imaginary number, we have :

$$cf_{c,d}(v,u) \equiv M_{c,d}(iv,iu) = \phi_c(idv,iu)^{d/2} .$$

The distribution function of Z(c,d) can be obtained as follows. Let $F_{c,d}(z) = P[Z(c,d) \leq z]$. Given that $P[B(c,d) \leq 0] = 0$, we have, from Theorem 1 of Gurland (1948) :

$$F_{c,d}(z) = (1/2) - (1/2\pi i) \lim_{\substack{\epsilon_1 \to 0 \\ \epsilon_2 \to \infty}} \int_{\epsilon_1 < |v| < \epsilon_2} [cf_{c,d}(v,-vz)/v]dv$$

$$= (1/2) - (1/2\pi) \int_0^\infty \text{AIMAG}[cf_{c,d}(v,-vz)/v]dv , \tag{11}$$

where AIMAG(·) denotes the imaginary part of the complex number. Further, the density function is given by :

$$f_{c,d}(z) \equiv \partial F_{c,d}(z)/\partial z = (1/2\pi i) \int_0^\infty \{\partial cf_{c,d}(v,u)/\partial u\}_{u=-vz} dv .$$

The moment–generating function described in Theorem 1 can also be used to calculate the moments of Z(c,d). Using Metha and Swamy's (1978) result, we have :

$$E[Z(c,d)]^r = \Gamma(r)^{-1} \int_0^\infty u^{r-1} \left\{\partial^r M_{c,d}(v,-u)/\partial v^r\right\}_{v=0} du ,$$

where Γ(·) denotes the Gamma function. The above expressions can be evaluated using numerical integration. In this section, we concentrate on calculating percentage points of the limiting distribution of $T(\hat{\alpha}_d - \alpha_d)$ using (11). We evaluate the integrals in the range $(0 + \epsilon, \bar{V})$ where $\bar{V}$ is an upper bound set such that the integrand evaluated at $\bar{V}$ is less than ϵ. ϵ was set at 1.0E–08 in each integration. For many cases, we also verified the sensitivity of the results to using a higher bound of integration. In all cases the results were identical at the precision reported.

A further comment about the nature of the numerical integration is warranted. The integrand in (11) involves the square root of complex quantities. It does so first through λ given that, in the relevant expression, $\lambda = (c^2 + 2idv(c + z))^{1/2}$. However, this causes no concern since the integrand in (11) is a function of λ only through λ^2. To see this, note that we can write $\phi_c(ix, iu)$ as (see Perron (1989)) :

$$\phi_c(ix,iu) = \exp(-(ix + c)/2)(\cos(\theta) - (ix + c)\sin(\theta)/\theta) ,$$

where $\theta = i\lambda$. Since $\cos(\theta)$ depends on θ only through θ^2 and $\sin(\theta)$ depends on θ only through odd powers of θ, it follows that $\phi_c(ix,iu)$ depends on λ only through λ^2. Hence, in this case, the choice of the branch of the square root is unimportant.

Of more importance is the fact that when d is an odd number the integrand involves again the square root of a complex valued quantity, namely $\phi_c(idv,-ivz)$. In that case special care must be applied to the integration as the use of the principal value of the square root may not ensure the continuity of the integrand. In that case, one must therefore integrate over the Reimann surface consisting here of two planes. This feature is present most notably in the non–seasonal model (d = 1) and the specifics of the procedure for the integration are discussed in Perron (1989). In the seasonal context, however, the values of d considered are most often even numbers (say 2, 4 and 12 as used below). Hence the numerical problems discussed above do not apply and standard integration formulae can be used. For the examples given below we have used the subroutine QDAG of the International Mathematical and Statistical Library (IMSL).

Table I presents various percentage points of the asymptotic distribution of $T(\hat{\alpha}_d - \alpha_d)$ for the following configuration of parameters : c = –10, –5, –2, –1, –.5, 0, .5, 1, 2 and 5. Three values of d are considered, d = 2, 4 and 12. The percentage points were obtained using (11) with a secant method such that the probability associated with a given critical value is within ± 1.0E–0.6 of the percentage point of interest. The results concerning the case where c = 0.0 are especially interesting since they allow comparisons with those of Dickey, Hasza and Fuller (1984) for the seasonal random walk case.

Several features are worth noting from these results. First, the spread of the distribution is much larger for negative values of c (the locally stationary case). It rapidly becomes more concentrated around zero as c increases. This is to be expected since the usual asymptotic theory with α_d fixed as the sample size increases suggests that $\hat{\alpha}_d$ converges at rate $T^{1/2}$ when $|\alpha_d| < 1$ instead of the rate T implied in the present setting. On the other hand, when $|\alpha_d| > 1$, $\hat{\alpha}_d$ converges to its true value at a rate faster than T. Secondly, when c is negative or zero the distribution shows higher variance as d increases. Such is not the case, however, when $c \geq 1.0$ where the variance diminishes as d increases (though a non–monotonic behavior is observed when c = 1.0).

When c = 0.0, our results can be compared to those of Dickey, Hasza and Fuller (1984), referred to as DHF. The values obtained are usually in close agreement, though some noticeable differences occur, especially in the left tail of the distribution. Consider, for instance, the first percentage point of the distribution when d = 4. Our value is –14.938, while that of DHF is –15.27. We believe our results to be more accurate given that those of DHF were obtained through simulation methods. Comparing our asymptotic critical value to the finite sample critical value reported in DHF shows the asymptotic distribution to match the finite sample critical value for a smaller sample size (near T = 200 instead of a T greater than 800 as implied by the results of DHF). A similar feature holds with d = 12 and to a lesser extent with d = 2. The differences are not as severe for other percentage points.

The percentage points presented in Table I are useful for a variety of purposes. They can be used to construct asymptotic confidence intervals, to provide critical points on which test statistics can be based and also to analyze the limiting power function of tests of the null hypothesis of a seasonal random walk under a sequence of local alternatives. The reader is referred to Perron (1989) for more details. Percentage points other than those presented can be derived using (11). The numerical integration is simple and fast.

3. AN $O_p(T^{-1})$ ASYMPTOTIC EXPANSION

In this Section, we generalize the results of Section 2 by considering an $O_p(T^{-1})$ asymptotic expansion for the distribution of $T(\hat{\alpha}_d - \alpha_d)$. We also derive the appropriate limiting joint moment–generating function used to tabulate, via numerical integration, percentage points of the distribution for some values of c, d and sample size T.

We start with the following Lemma concerning the stochastic expansion for the moments of the data. Here and throughout the rest of this paper, $\overset{\mathcal{D}}{=}$ signifies equality in distribution.

LEMMA 1 : *Let $\{y_t\}$ be generated by (1) and (2) and suppose that $\{u_t\}$ satisfies the conditions of Proposition 1. Further assume that $\{u_t\}$ is Gaussian. Then :*

i) $T^{-1}\, y_{T-i+1}^2 \overset{\mathcal{D}}{=} d^{-1}\sigma^2\{J_{i,c}(1)^2 + 2\gamma_i exp(c) J_{i,c}(1)\} + O_p(T^{-1})\,;\qquad (i = 1, \ldots d)$

ii) $T^{-2}\Sigma_1^T y_{t-d}^2 \overset{\mathcal{D}}{=} \sigma^2 d^{-2} F(c,d) + O_p(T^{-1}) \equiv \sigma^2 d^{-2}\Sigma_{i=1}^d F_i(c,d) + O_p(T^{-1})$

where $F_i(c,d) \equiv \int_0^1 J_{i,c}^2(r)dr + 2\gamma_i \int_0^1 exp(cr) J_{i,c}(r)dr \quad (i = 1, \ldots, d);$

iii) $T^{-1}\Sigma_1^T u_t^2 \overset{\mathcal{D}}{=} \sigma^2 + \sigma^2 (2/T)^{1/2}\xi\ ;$

iv) $T^{-1}\Sigma_1^T y_{t-d} u_t \overset{\mathcal{D}}{=} \sigma^2 d^{-1} H(c,d) + O_p(T^{-1}) \equiv \sigma^2 d^{-1}\Sigma_{i=1}^d H_i(c,d) + O_p(T^{-1})$

where $H_i(c,d) \equiv \int_0^1 J_{i,c}(r) dW_i(r) + \gamma_i \int_0^1 exp(cr) dW_i(r) - d(2T)^{-1/2}\xi \quad (i = 1, \ldots, d);$

with $\gamma_i = (d/T)^{1/2} y_{-i+1}/\sigma$, ξ a N(0,1) random variable distributed independently of the Wiener processes $W_i(r)$ $(i = 1, \ldots, d)$, and with $J_{i,c}(r)$ as defined in Proposition 1.

The proof of Lemma 1 (based on derivations by Perron (1990b) and Phillips (1987b)) is presented in the Appendix. Noting that $T(\hat{\alpha}_d - \alpha_d) = T^{-1}\Sigma_1^T y_{t-d} u_t / T^{-2}\Sigma_1^T y_{t-d}^2$ it is easy to deduce, from parts (ii) and (iv), an asymptotic expansion for the distribution of $T(\hat{\alpha}_d - \alpha_d)$. The result is presented in the following Theorem.

THEOREM 2 : *Let $\{y_t\}$ be generated by (1) and (2) and suppose that the sequence $\{u_t\}$ satisfies the conditions of Proposition 1. Further assume that $\{u_t\}$ is Gaussian. Then :*

$$T(\hat{\alpha}_d - \alpha_d) \stackrel{\mathcal{D}}{=} Q(c,d) + O_p(T^{-1}),$$

where $Q(c,d) \equiv dH(c,d)/F(c,d)$ with $H(c,d)$ and $F(c,d)$ as defined in Lemma 1.

As can be seen from the result of Theorem 2, the $O_p(T^{-1})$ expansion for the distribution of $T(\hat{\alpha}_d - \alpha_d)$ involves the initial condition y_i ($i = 0, \ldots, -d+1$). When these initial conditions are zero, we have the following simplified result.

COROLLARY 1 : *Let the conditions of Theorem 2 be satisfied and further assume that $y_i = 0$ ($i = 0, \ldots, -d + 1$). Then :*

$$T(\hat{\alpha}_d - \alpha_d) \stackrel{\mathcal{D}}{=} d\left\{\Sigma_{i=1}^{d}\int_0^1 J_{i,c}(r)dW_i(r) - d(2T)^{-1/2}\xi\right\}\Big/\left\{\Sigma_{i=1}^{d}\int_0^1 J_{i,c}^2(r)dr\right\} + O_p(T^{-1}).$$

In the case where the initial conditions are zero, the $O_p(T^{-1/2})$ factor in the asymptotic expansion is quite simple. It only involves a standard N(0,1) variable in the numerator. Hence, the $O_p(T^{-1})$ expansion provides no adjustment to the location of the distribution (the correction factor having mean zero). The influence of the correction factor is to increase, for any finite sample size, the variance of the random variable describing the standard $O_p(1)$ asymptotic distribution. Hence, the expansion is likely to provide a better approximation to the exact distribution of $T(\hat{\alpha}_d - \alpha_d)$ in the cases where (1) the location of the distribution is fairly stable as T changes and (2) the variance of the exact distribution decreases as T increases.

In order to compute percentage points associated with the distribution of the random variable Q(c,d) in the asymptotic expansion, we proceed as in Section 2 with the derivation of the joint moment–generating function of {dH(c,d), F(c,d)} which we denote by

$$MT_{c,d}(v,u) \equiv E[\exp(dvH(c,d) + uF(c,d))] .$$

We start with the following Lemma concerning the joint moment–generating function of the individual elements $\{dH_i(c,d), F_i(c,d)\}$ $(i = 1,..., d)$.

LEMMA 2 : *Let $\phi_c(x, u)$ be as defined in (10). Then :*

$$E[\exp\{dH_i(c,d) + uF_i(c,d)\}] = \exp(-u\delta_i + (dv)^2/4T)\phi_c(dv, u)^{1/2}\Upsilon_c(\gamma_i, dv, u) ,$$

where $\Upsilon_c(\gamma_i, x, u) \equiv \exp\{-(\gamma_i^2/2)(x + c + \lambda)[1 - \exp(x + c + \lambda)\psi_c(x, u)]\}$;

with $\lambda = (c^2 + 2cx - 2u)^{1/2}$, $\gamma_i = (d/T)^{1/2}y_{-i+1}/\sigma$ *and* $\delta_i = \gamma_i^2(\exp(2c) - 1)/2c$.

The proof of Lemma 2 can be obtained as a special case of Theorem 1 of Perron (1990b). With this result, it is straightforward to deduce the joint moment–generating function of $\{dH(c,d), F(c,d)\}$ which we state in the following Theorem.

THEOREM 3 : *Let $\phi_c(x, u)$ be defined by (10) and $\Upsilon_c(\gamma_i, x, u)$ be as defined in Lemma 2, then the joint moment–generating function of $\{dH(c,d), F(c,d)\}$ is given by :*

$$MT_{c,d}(v, u) = \phi_c(dv,u)^{d/2}\exp(v^2d^3/4T)\ \Pi_{i=1}^{d}\exp(-u\delta_i)\Upsilon_c(\gamma_i, dv, u)$$

where $\gamma_i = (d/T)^{1/2}y_{-i+1}/\sigma$ *and* $\delta_i = \gamma_i^2(\exp(2c) - 1)/2c$.

Proof : We have $MT_{c,d}(v,u) = E[\exp(dv\Sigma_{i=1}^{d}H_i(c,d) + u\Sigma_{i=1}^{d}F_i(c,d))] = \Pi_{i=1}\ E[\exp(dvH_i(c,d) + uF_i(c,d)]$, given the independence of the Wiener processes involved in the pairs $\{H_i(c,d), F_i(c,d)\}$ $(i = 1, ..., d)$, as well as the independence of the variable ξ. Hence, using Lemma 2, $MT_{c,d}(v,u) = \Pi_{i=1}^{d}[\phi_c(dv,u)^{1/2}\exp(-u\delta_i + (dv)^2/4T)\Upsilon_c(\gamma_i, dv, u)] = \phi_c(dv,u)^{d/2}\exp(d^3v^2/4T)\ \Pi_{i=1}^{d}\exp(-u\delta_i)\Upsilon_c(\gamma_i, dv, u)$. □

When the initial conditions are 0, we have the following special case.

COROLLARY 2 : *If* $y_0 = \ldots = y_{-d+1} = 0$, *then :*

$$MT_{c,d}(v,u) = exp(v^2 d^3/4T)\phi_c(dv,u)^{d/2}.$$

The results of Theorem 2 and Corollary 2 can be used to obtain percentage points of Q(c,d), the $O_p(T^{-1})$ asymptotic distribution of $T(\hat{\alpha}_d - \alpha_d)$, for various values of c, d, T and normalized initial conditions (y_i/σ) (i = 0, ..., –d + 1). The method is similar to that used in Section 2 and is based on numerically integrating expression (11) with appropriate modification for the joint moment–generating function used.

We have calculated the percentage points of the distribution of Q(c,d) in the case where the initial conditions are set to 0. Tables II through IV present the results for c = –5, 0 and 2 respectively. The values of d chosen are again 2, 4 and 12. The sample sizes used vary with d and were chosen to allow comparisons (in the case c = 0) with the finite sample simulated values presented in Dickey, Hasza and Fuller (1984).

As discussed before, the $O_p(T^{-1/2})$ correction increases the spread of the distribution in all cases. The corrections are quite minor when c = 2.0. They become more important when c = 0 and –5, especially when d is small. Considering the case c = 0 and comparing the results with the percentage points from the usual $O_p(1)$ asymptotic distribution and the exact values reported in Dickey, Hasza and Fuller (1984) we draw the following conclusions. The critical values from the $O_p(T^{-1})$ asymptotic expansion are closer to the exact value of the distribution of $T(\hat{\alpha}_d - \alpha_d)$ in the right tail of the distribution compared to the critical values from the standard $O_p(1)$ asymptotic distribution. Hence, it provides an improvement, though minor, in the right tail of the distribution. Concerning the median and the left tail of the distribution, the critical values from the $O_p(T^{-1})$ expansion are actually further away from the exact values compared to those of the standard $O_p(1)$ asymptotic distribution (except for the 1% point with d = 12). Hence, in the left tail, the asymptotic expansion fails to provide an improvement.

4. CONCLUSION

This study offers a simple way to calculate percentage points and other distributional quantities of interest for the limiting distribution of the least–squares estimator in a nearly integrated seasonal model. The method involves only a one–dimensional numerical integration based on the exact joint moment–generating function of the variables characterizing the limiting distribution. Hence, it is faster and more accurate than the simulation method proposed by Chan (1989).

The method proposed allow easy calculation of asymptotically valid confidence intervals which do not suffer from the discontinuity problem inherent in the standard fixed α_d asymptotic theory. It also allows an easily computable approximation to the exact distribution of $T(\hat{\alpha}_d - \alpha_d)$ which is quite accurate for values of α_d in a neighborhood of one, where the usual asymptotic theory fails to provide a useful guide. To be more precise, the approximation is adequate for value of T, d and α_d such that, from (2), $c = T\ln(\alpha_d)/d$ is not "too far" from 0. Finally, given that the model studied here can be interpreted as a sequence of models that are local to the seasonal random walk model as T increases, the method can easily deliver values for the local power of tests for the null hypothesis of a seasonal random walk.

Possible extensions to this line of research include the derivation of the appropriate joint–moment generating function associated with the limiting distribution of the least–squares estimators when a constant and/or a trend are included in the regression. Also of interest are similar derivations concerning the t–statistic on $\hat{\alpha}_d$ in such models. These topics involve non–trivial extensions to the present study.

APPENDIX : PROOF OF LEMMA 1

Throughout we assume, without loss of generality, that T = md for some integer m. For simplicity of notation we let $\alpha = \alpha_d = \exp(cd/T)$. First note that we can write y_t as

$$y_t = \sum_{k=0}^{[(t-1)/d]} \alpha^k u_{t-kd} + \alpha^{[(t+1)/d]} y_{t-d\{[(t-1)/d]+1\}} ,$$

where $[\cdot]$ denotes the integer part of the argument. Note that y_t depends only on a subset of length $[(t-1)/d] + 1$ of the sequence $\{u_t\}$ and an initial condition. As in Chan (1989), we can write :

$$y_{td-i+1} = \alpha y_{(t-1)d-i+1} + u_{td-i+1} \qquad (t = 1, ..., m) \quad (i = 1, ..., d)$$

where the sequences $\{u_{td-i+1}\}$ are independent for $i = 1, ..., d$. Hence, y_{td-i+1} can be represented by a near–integrated process of length m with non–centrality parameter c, given that $\alpha = \exp(cd/T) = \exp(c/m)$. Hence $y_{td-i+1} \overset{\mathcal{D}}{=} x_{t,i}$, say, where the process $\{x_{t,i}$; $t = 1, ..., m\}$ satisfies, for a given fixed i, the difference equation :

$$x_{t,i} = \exp(c/m)\, x_{(t-1),i} + \epsilon_{t,i} \qquad (t = 1, ..., m) \quad (i = 1, ..., d) \qquad (A.1)$$

with initial condition $x_{0,i} = y_{-i+1}$ and where the sequence $\{\epsilon_{t,i}$; $t = 1, ..., m\}$ is a subset of the sequence $\{u_t$; $t = 1, ..., md\}$. To prove part (i), we apply Lemma A.1 (a) of Perron (1990b) to obtain (for all $i = 1, ..., d$) :

$$T^{-1/2}\, y_{T-i+1} \overset{\mathcal{D}}{=} d^{-1/2} m^{-1/2} x_{T,i}$$

$$\overset{\mathcal{D}}{=} \sigma J_c(1) + T^{-1/2}\exp(c) y_{-i+1} + O_p(T^{-1}) .$$

Part (i) follows easily. To prove part (ii), we proceed similarly applying Lemma A.1 (c) of Perron (1990b) to $\Sigma_{t=0}^{m-1} x_{t,i}^2$ in the following way :

$$T^{-2}\Sigma_1^T y_{t-d}^2 = T^{-2}\Sigma_{i=1}^d \Sigma_{t=0}^{m-1} y_{td-i+1}^2 \overset{\mathcal{D}}{=} d^{-2}\Sigma_{i=1}^d m^{-2}\Sigma_{t=0}^{m-1} x_{t,i}^2$$

$$\overset{\mathcal{D}}{=} d^{-2}\Sigma_{i=1}^d\left\{\sigma^2\int_0^1 J_{i,c}(r)^2 dr + 2m^{-1/2}\sigma y_{-i+1}\int_0^1 \exp(cr)J_{i,c}(r)dr\right\} + O_p(T^{-1}) .$$

The result of part (ii) follows using the fact that T = md. Part (iii) follows from Phillips (1987b) who shows that $T^{-1}\Sigma_1^T u_t^2 \overset{\mathcal{D}}{=} \sigma^2 + T^{-1/2}\eta + O_p(T^{-1})$ where $\eta \sim N(0, \nu^2)$ with $\nu^2 = 2\pi f_2(0)$ where $f_2(\lambda)$ is the spectral density of $\{u_t^2 - Eu_t^2\}$. In the case of i.i.d. Gaussian errors, $\nu^2 = 2\sigma^4$. Hence, $T^{-1}\Sigma_1^T u_t^2 \overset{\mathcal{D}}{=} \sigma^2 + \sigma^2(2/T)^{1/2}\xi + O_p(T^{-1})$ where $\xi \sim N(0, 1)$.

To prove part (iv), first note that, after some manipulations using (1), we obtain :

$$T^{-1}\Sigma_1^T y_{t-d}u_t = (1/2\alpha)\left\{\Sigma_{i=1}^d y_{T-d+i}^2 - \Sigma_{i=1}^d y_{-i+1}^2 - T(\alpha^2-1)T^{-2}\Sigma_1^T y_{t-d}^2 - T^{-1}\Sigma_1^T u_t^2\right\} .$$

Using parts (i) through (iii) with the facts that $\alpha = \exp(cd/T) = 1 + O_p(T^{-1})$ and $T(\alpha^2 - 1) = 2cd + O_p(T^{-1})$, we obtain :

$$\begin{aligned}
T^{-1}\Sigma_1^T y_{t-d}u_t \overset{\mathcal{D}}{=} (1/2)\Big\{&\Sigma_{i=1}^d d^{-1}\sigma^2[J_{i,c}(1)^2 + 2\gamma_i \exp(c)J_{i,c}(1)] \\
&- 2cd^{-1}\sigma^2\Sigma_{i=1}^d[\textstyle\int_0^1 J_{ic}(r)^2 dr + 2\gamma_i\int_0^1 \exp(cr)J_{i,c}(r)dr] \\
&- \sigma^2 - \sigma^2(2/T)^{1/2}\xi\Big\} + O_p(T^{-1}) \\
= (\sigma^2/d)\Big\{&(1/2)\,\Sigma_{i=1}^d[J_{i,c}(1)^2 - 2c\textstyle\int_0^1 J_{i,c}(r)^2 dr - 1] \\
&+ \Sigma_{i=1}^d \gamma_i[\exp(c)J_{i,c}(1) - 2c\textstyle\int_0^1 \exp(cr)J_{i,c}(r)dr] \\
&- d(2T)^{-1/2}\xi\Big\} + O_p(T^{-1}) . \qquad \text{(A.2)}
\end{aligned}$$

It is shown in Phillips (1987a, eq. (8)) that (for i = 1, ..., d) :

$$(1/2)[J_{i,c}(1)^2 - 2c\int_0^1 J_{i,c}(r)^2 dr - 1] \overset{\mathcal{D}}{=} \int_0^1 J_{i,c}(r)dW_i(r) , \qquad \text{(A.3)}$$

and in Perron (1990a, Lemma A.1) that (for i = 1, ..., d) :

$$[\exp(c)J_{i,c}(1) - 2c\int_0^1 \exp(cr)J_{i,c}(r)dr] \overset{\mathcal{D}}{=} \int_0^1 \exp(cr)dW_i(r) . \qquad \text{(A.4)}$$

Part (iv) follows upon substituting (A.3) and (A.4) into (A.2).

REFERENCES

Bobkoski, M.J. (1983). Hypothesis testing in nonstationary time series. PhD Thesis, University of Wisconsin–Madison.

Box, G.E.P., and Jenkins, G.M. (1976). *Time Series Analysis : Forecasting and Control.* Holden–Day, San Francisco.

Chan, N.H. (1988). The parameter inference for nearly nonstationary time series. *Journal of the American Statistical Association*, 83, 857–862.

Chan, N.H. (1989). On the nearly nonstationary seasonal time series. *The Canadian Journal of Statistics*, 17, 279–284.

Chan, N.H. and Wei, C.Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics*, 15, 1050–1063.

Dickey, D.A.; Hasza, D.P. and Fuller, W.A. (1984). Testing for unit roots in seasonal time series. *Journal of the American Statistical Association*, 79, 355–367.

Evans, G.B.A., and Savin, N.E. (1981). The calculation of the limiting distribution of the least squares estimator of the parameter in a random walk model. *Annals of Statistics*, 9, 1114–1118.

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

Latour, A. (1986). Études des autocovariances et des autocorrélations échantillonnales de modèles saisonniers. PhD Thesis, Université de Montréal.

Latour, A., and Roy, R. (1987). Some exact results on the sample autocovariances of seasonal ARIMA model. *The Canadian Journal of Statistics*, 27, 143–161.

Metha, J.S., and Swamy, P.A.V.B. (1978). The existence of moments of some simple bayes estimators of coefficients in a simultaneous equation model. *Journal of Econometrics*, 7, 1–14.

Perron, P. (1989). The calculation of the limiting distribution of the least–squares estimator in a near integrated model. *Econometric Theory*, 5, 241–255.

Perron, P. (1990a). A continuous time approximation to the unstable first–order autoregressive process : the case without an intercept. Forthcoming in *Econometrica.*

Perron, P. (1990b). The adequacy of limiting distributions in the AR(1) model with dependent errors. Econometric Research Program Memorandum No. 349, Princeton University.

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

Phillips. P.C.B. (1987b). Asymptotic expansions in nonstationary vector autoregressions. *Econometric Theory*, 3, 45–68.

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

TABLE I : Percentage Points of the Asymptotic Distribution of $T(\hat{\alpha}_d - \alpha_d)$; $\alpha_d = \exp(cd/T)$

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
d=2									
c=−10.0	−22.973	−18.570	−15.086	−11.404	−0.989	6.128	7.679	8.889	10.158
c=−5.0	−19.312	−15.419	−12.399	−9.275	−0.977	4.030	5.037	5.809	6.623
c=−2.0	−16.509	−12.970	−10.285	−7.578	−0.944	2.490	3.182	3.726	4.282
c=−1.0	−15.389	−11.976	−9.413	−6.866	−0.903	1.989	2.580	3.068	3.638
c=−0.5	−14.779	−11.426	−8.925	−6.461	−0.857	1.759	2.298	2.753	3.293
c=0.0	−14.115	−10.827	−8.389	−6.010	−0.775	1.539	2.028	2.449	2.956
c=0.5	−13.400	−10.165	−7.788	−5.492	−0.652	1.332	1.770	2.155	2.630
c=1.0	−12.604	−9.417	−7.091	−4.867	−0.509	1.137	1.525	1.875	2.315
c=2.0	−10.645	−7.447	−5.121	−3.095	−0.247	0.790	1.080	1.359	1.728
c=5.0	−1.058	−0.561	−0.350	−0.213	−0.007	0.153	0.225	0.311	0.446
d=4									
c=−10.0	−28.811	−23.551	−19.306	−14.721	−0.993	9.448	11.906	13.883	16.010
c=−5.0	−23.325	−18.871	−15.343	−11.606	−0.982	6.373	8.002	9.287	10.652
c=−2.0	−18.950	−15.107	−12.131	−9.062	−0.941	4.042	5.106	5.955	6.873
c=−1.0	−17.126	−13.512	−10.757	−7.960	−0.880	3.228	4.112	4.830	5.625
c=−0.5	−16.090	−12.603	−9.967	−7.319	−0.817	2.837	3.630	4.282	5.018
c=0.0	−14.938	−11.584	−9.074	−6.589	−0.721	2.460	3.166	3.753	4.427
c=0.5	−13.636	−10.413	−8.038	−5.738	−0.598	2.103	2.721	3.245	3.855
c=1.0	−12.106	−9.023	−6.812	−4.749	−0.461	1.768	2.301	2.761	3.309
c=2.0	−8.025	−5.548	−3.992	−2.688	−0.217	1.176	1.551	1.890	2.311
c=5.0	−0.565	−0.413	−0.313	−0.222	−0.005	0.190	0.260	0.331	0.428
d=12									
c=−10.0	−43.994	−36.410	−30.139	−23.193	−0.996	17.924	22.746	26.751	31.201
c=−5.0	−34.082	−28.019	−23.077	−17.675	−0.986	12.450	15.749	18.447	21.410
c=−2.0	−25.913	−21.088	−17.227	−13.092	−0.933	8.144	10.292	12.036	13.943
c=−1.0	−22.375	−18.073	−14.678	−11.089	−0.855	6.526	8.264	9.680	11.241
c=−0.5	−20.325	−16.332	−13.207	−9.936	−0.783	5.714	7.249	8.505	9.898
c=0.0	−18.039	−14.396	−11.579	−8.665	−0.685	4.917	6.252	7.353	8.580
c=0.5	−15.481	−12.252	−9.795	−7.290	−0.563	4.151	5.292	6.241	7.310
c=1.0	−12.683	−9.959	−7.921	−5.870	−0.430	3.430	4.386	5.189	6.104
c=2.0	−7.210	−5.632	−4.473	−3.314	−0.197	2.177	2.803	3.343	3.975
c=5.0	−0.617	−0.513	−0.423	−0.323	−0.004	0.301	0.391	0.470	0.559

TABLE II

Percentage Points of the Distribution of $T(\hat{\alpha}_d - \alpha_d)$; $O(T^{-1})$ Asymptotic Expansion ; $\alpha_d = \exp(cd/T)$, c = 0.0

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
d=2									
T=20	–14.995	–11.307	–8.640	–6.095	–0.724	1.603	2.106	2.546	3.084
T=30	–14.717	–11.147	–8.553	–6.063	–0.741	1.582	2.080	2.512	3.040
T=40	–14.570	–11.068	–8.511	–6.049	–0.750	1.571	2.066	2.496	3.018
T=100	–14.297	–10.924	–8.438	–6.025	–0.765	1.552	2.043	2.468	2.979
d=4									
T=40	–15.259	–11.774	–9.187	–6.644	–0.701	2.502	3.215	3.810	4.494
T=60	–15.151	–11.710	–9.149	–6.625	–0.708	2.488	3.198	3.791	4.472
T=80	–15.098	–11.678	–9.130	–6.616	–0.711	2.481	3.190	3.781	4.461
T=200	–15.001	–11.622	–9.097	–6.600	–0.717	2.469	3.176	3.765	4.440
d=12									
T=120	–18.132	–14.461	–11.626	–8.696	–0.679	4.941	6.281	7.384	8.615
T=180	–18.103	–14.439	–11.611	–8.685	–0.681	4.933	6.271	7.374	8.605
T=240	–18.088	–14.430	–11.603	–8.681	–0.682	4.929	6.266	7.368	8.597
T=600	–18.059	–14.408	–11.589	–8.671	–0.684	4.922	6.258	7.359	8.586

TABLE III

Percentage Points of the Distribution of $T(\hat{\alpha}_d - \alpha_d)$; $O(T^{-1})$ Asymptotic Expansion ; $\alpha_d = \exp(cd/T)$, $c = -5.0$

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
d=2									
T=20	–22.622	–17.755	–14.059	–10.311	–0.735	4.778	5.878	6.730	7.645
T=30	–21.623	–17.037	–13.538	–9.976	–0.802	4.526	5.586	6.402	7.270
T=40	–21.093	–16.657	–13.266	–9.804	–0.841	4.401	5.444	6.246	7.096
T=100	–20.063	–15.932	–12.754	–9.488	–0.919	4.176	5.196	5.978	6.805
d=4									
T=40	–25.230	–20.272	–16.380	–12.290	–0.854	6.901	8.602	9.941	11.365
T=60	–24.627	–19.821	–16.045	–12.067	–0.893	6.726	8.401	9.721	11.121
T=80	–24.312	–19.591	–15.873	–11.953	–0.914	6.638	8.301	9.611	11.001
T=200	–23.727	–19.165	–15.559	–11.746	–0.954	6.479	8.122	9.415	10.791
d=12									
T=120	–34.954	–28.696	–23.602	–18.045	–0.941	12.766	16.123	18.867	21.879
T=180	–34.668	–28.470	–23.427	–17.924	–0.956	12.661	15.999	18.728	21.723
T=240	–34.517	–28.360	–23.341	–17.862	–0.963	12.609	15.936	18.658	21.643
T=600	–34.255	–28.158	–23.182	–17.749	–0.976	12.514	15.824	18.533	21.504

TABLE IV

Percentage Points of the Distribution of $T(\hat{\alpha}_d - \alpha_d)$; $O(T^{-1})$ Asymptotic Expansion ; $\alpha_d = \exp(cd/T)$, $c = 2.0$

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
d=2									
T=20	–10.799	–7.464	–5.115	–3.090	–0.243	0.799	1.094	1.379	1.760
T=30	–10.740	–7.456	–5.117	–3.092	–0.245	0.796	1.089	1.372	1.749
T=40	–10.718	–7.453	–5.118	–3.093	–0.245	0.795	1.087	1.369	1.743
T=100	–10.674	–7.449	–5.119	–3.095	–0.247	0.792	1.083	1.363	1.734
d=4									
T=40	–8.059	–5.562	–4.000	–2.691	–0.216	1.181	1.557	1.898	2.323
T=60	–8.048	–5.557	–3.997	–2.690	–0.217	1.179	1.555	1.895	2.319
T=80	–8.042	–5.554	–3.996	–2.689	–0.217	1.179	1.554	1.894	2.316
T=200	–8.031	–5.551	–3.993	–2.688	–0.217	1.177	1.552	1.892	2.314
d=12									
T=120	–7.217	–5.636	–4.476	–3.316	–0.197	2.179	2.806	3.346	3.978
T=180	–7.214	–5.635	–4.475	–3.315	–0.197	2.179	2.805	3.345	3.977
T=240	–7.214	–5.635	–4.474	–3.315	–0.197	2.178	2.804	3.345	3.977
T=600	–7.210	–5.633	–4.474	–3.314	–0.197	2.178	2.804	3.344	3.976