Université de Montréal
Faculté des arts et des sciences
Département de sciences économiques

CAHIER 9420

APPROXIMATIONS TO SOME EXACT DISTRIBUTIONS IN THE FIRST ORDER AUTOREGRESSIVE MODEL WITH DEPENDENT ERRORS

Seiji NABEYA[1] and Pierre PERRON[2]

1 Department of Economics, Tokyo International University.

2 Département de sciences économiques and Centre de recherche et développement en économique (C.R.D.E.), Université de Montréal.

September 1994

Financial support from the National Science Foundation, the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec, and the Université de Montréal (CAFIR) is gratefully acknowledged. The computations were carried out using the computer FACOM M–760/6 at the Computer Center, Hitotsubashi University, which we thank for their support.

C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7

Télécopieur (FAX): (514) 343-5831
Courrier électronique (E-Mail): econo@tornade.ERE.Umontreal.CA

Ce cahier a également été publié au Centre de recherche et développement en économique (C.R.D.E.) (publication no 2494).

Dépôt légal - 1994
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISSN 0709-9231

RÉSUMÉ

On considère le modèle autorégressif quasi intégré $y_t = \alpha y_{t-1} + u_t$, où $\alpha = \exp(c/T)$ et où la séquence d'erreurs u_t suit soit un processus MA(1) ($u_t = e_t + \theta e_{t-1}$), soit un processus AR(1) ($u_t = \rho u_{t-1} + e_t$). On étudie la distribution de $\hat{\alpha}$, l'estimateur des moindres carrés de α. On suggère des modifications à la structure asymptotique locale telle qu'analysée par Nabeya et Perron (1992), lesquelles permettent d'excellentes approximations pour tous les cas où θ et ρ sont près de leur frontière respective. L'idée derrière cette nouvelle approximation découle de l'approche basée sur la théorie des déterminants de Fredholm. Plutôt que d'approximer, par exemple, la distribution échantillonnale finie $P[T(\hat{\alpha}-1) \leq x]$ par $P\left[\int_0^1 \int_0^1 K(s,t,x)\, dW(s)dW(t) \geq 0\right]$ où $K(s,t,x)$ est un noyau limite approprié (indépendant de la taille de l'échantillon T), on utilise un noyau $K_T(s,t,x)$ qui dépend de la taille de l'échantillon. Par un choix judicieux de la dépendance du noyau à T, il est possible d'obtenir des approximations qui sont non seulement excellentes mais aussi relativement faciles à évaluer en utilisant des techniques d'intégration numérique simples. Nos résultats sont très encourageants et manifestent le potentiel de recherche intéressant de cette approche et la nécessité de l'appliquer à des modèles plus complexes.

Mots-clés : approximations asymptotiques, fonction caractéristique, déterminant de Fredholm, modèles quasi intégrés, modèles quasi stationnaires.

ABSTRACT

We consider the near-integrated autogressive model $y_t = \alpha y_{t-1} + u_t$, where $\alpha = \exp(c/T)$ and the sequence of errors u_t is allowed to be an MA(1) process ($u_t = e_t + \theta e_{t-1}$) or an AR(1) process ($u_t = \rho u_{t-1} + e_t$). We study the distribution of $\hat{\alpha}$, the least-squares estimator of α. We suggest modifications to the local asymptotic framework analyzed by Nabeya and Perron (1992) which provide excellent approximations for all cases where θ or ρ are close to their relevant boundaries. The idea behind this new approximation is based on the "Fredholm determinant approach" where instead of approximating, for example, the finite sample distribution $P[T(\hat{\alpha} - 1) \leq x]$ by, say, $P\left[\int_0^1 \int_0^1 K(s,t,x)\, dW(s)dW(t) \geq 0\right]$ where $K(s,t,x)$ is a suitable limiting kernel (independent of the sample size T), we use a kernel, say, $K_T(s,t,x)$ that depends on the sample size. By a judicious choice of the dependence of the kernel on T, we are able to obtain approximations that are not only excellent but also relatively easy to evaluate using straightforward numerical integration techniques. Our results are very encouraging and show the approach to be worthy of further investigations and applications in more complex models.

Key words: asymptotic approximations; characteristic function; Fredholm determinant; near-integrated models; nearly stationary models.

1. INTRODUCTION.

This paper is concerned with the issue of approximating the exact distribution of estimators in dynamic linear models where substantial dependence over time is allowed in the errors. Models involving such features have become quite popular in econometrics mainly because of the possibility to apply a new class of central limit theorems (or functional central limit theorems) which provides asymptotic results allowing both substantial heterogeneity and dependence. Asymptotic inference in these models are now commonplace (see, e.g., White (1984) and Gallant and White (1988)), yet little is known about the adequacy of the asymptotic distributions as approximations to the finite sample distributions. In particular, when the asymptotic distribution is a poor guide to the exact distribution, little is known about what kind of alternative approximations provide useful improvements. This paper is an attempt to outline a class of approximations that we show is particularly successful in providing good approximations when existing asymptotic frameworks fail.

Our study considers the leading case of a dynamic first-order autoregressive model when the errors are allowed to be dependent. To be more precise, we consider the following first-order stochastic difference equation :

$$y_t = \alpha y_{t-1} + u_t, \qquad (t = 1,\ldots,T), \tag{1.1}$$

where y_0 is a fixed constant (or a random variable with a fixed distribution independent of T, the sample size) and $\{u_t\}$ is a sequence of weakly dependent random variables with mean zero. The least-squares estimator of α based on a sequence of observations $\{y_t\}_0^T$ is:

$$\hat{\alpha} = \Sigma_{t=1}^{T} y_t y_{t-1} (\Sigma_{t=1}^{T} y_{t-1}^2)^{-1}. \tag{1.2}$$

We consider the class of near-integrated processes with an autoregressive parameter defined by $\alpha = \exp(c/T)$ where the constant c is a measure of the deviation from the unit root case. Theoretical aspects of the limiting distribution of the least-squares estimator $\hat{\alpha}$ have been considered in Bobkoski (1983), Cavanagh (1986), Chan and Wei (1987) and Phillips (1987). Tabulations of the limiting distribution in the case where the errors $\{u_t\}$ are uncorrelated have been obtained by Chan (1988), Cavanagh (1986), Nabeya and Tanaka (1990) and Perron (1989) using different procedures. These studies also provide measures of the adequacy of this limiting distribution as an approximation to the finite sample distribution of $\hat{\alpha}$ when α is in the vicinity of 1. They show the approximation to be quite good in the case where $y_0 = 0$. Perron (1991a,b) and Tanaka (1990) also consider a continuous-time approximation which performs well even in the case where the initial condition is non-zero. These asymptotic frameworks provide substantial improvements over the traditional asymptotic distribution theory, when α is in the vicinity of one, essentially because the asymptotic distributions obtained are continuous with respect to the autoregressive parameter α.

Perron (1992) presented an extensive simulation analysis to assess the adequacy of this limiting distribution as an approximation to the finite sample distribution, concentrating on two leading cases, namely MA(1) errors ($u_t = e_t + \theta e_{t-1}$; $e_t \sim$ i.i.d. $N(0, \sigma^2)$) and AR(1) errors ($u_t = \rho u_{t-1} + e_t$). The results shown in Perron (1992) can be summarized as follows: 1) the asymptotic distribution is a very poor guide to the finite sample distribution, even for quite large sample sizes, when either θ (in the MA case) or ρ (in the AR case) are close to -1; 2) the inadequacy of the approximation is more severe in the MA case (for a given equal value of θ and ρ); 3) when ρ is close to + 1, the approximation is not as bad but the approach to the limiting distribution is quite slow; 4) the stochastic asymptotic expansion of the limiting distribution to order $O_p(T^{-1})$ provides a less accurate approximation in most cases than the standard limiting distribution.

In a recent paper, Nabeya and Perron (1992) considered an alternative framework where the parameter indexing the extent of the serial correlation in the errors (i.e. θ or ρ) is local to a boundary for which the asymptotic approximation fails. To be more precise, three cases were considered: a) MA root local to -1 specified such that $\theta_T = -1 + \delta/T^{1/2}$; b) AR root local to + 1 with $\rho_T = \exp(\phi/T)$; and c) AR root local to - 1 such that $\rho_T = -\exp(\phi/T)$. They derived the limiting distributions under these local frameworks as well as the appropriate limiting characteristic functions allowing the computation of distributional quantities of interest. It was shown that these alternative local asymptotic distributions not only give an explanation for the failures of the usual asymptotic distribution but also provide substantial improvements in approximating the exact distributions.

There are, nevertheless, some regions where none of the asymptotic frameworks available provide a satisfactory approximation. Examples of such regions include the case where the sample size is small and the autoregressive or moving-average parameters are close to -1. These cases illustrate the need for further refinements in the asymptotic approximations.

The purpose of this paper is to suggest modifications to the local asymptotic frameworks analyzed by Nabeya and Perron (1992) which provide excellent approximations for all cases where θ or ρ are close to their relevant boundaries. The idea behind this new approximation is based on the "Fredholm determinant approach" and can be explained as follows. The usual way to derive limiting distributions using the Fredholm determinant approach is to approximate the finite sample distribution $P[T(\hat{\alpha}-1) \leq x]$ by, say, $P[\int_0^1\int_0^1 K(s,t,x)dW(s)dW(t) \geq 0]$ where $K(s,t,x)$ is a suitable limiting kernel (independent of the sample size T). The characteristic function is then obtained as a transformation of the Fredholm determinant of this kernel. In the present work, instead of approximating the exact distribution using the limiting kernel, we use a kernel $K_T(s,t,x)$, say, that depends on the sample size. By a judicious choice of the dependence of the kernel on T we are able to obtain approximations that are not only excellent but also relatively easy to evaluate using straightforward numerical integration techniques.

We apply this technique to all cases mentioned above and emphasize i) an assessment of the quality of the approximations to the finite sample distributions and ii) their relative performances compared to other existing approximations. Our results are very encouraging and show the approach to be worthy of further applications in more complex models.

The plan of this paper is as follows. Section 2 provides preliminary material describing the models in more detail and covering the existing asymptotic approximations to the least-squares autoregressive estimator (1.2). Section 3 summarizes the so-called Fredholm determinant approach for deriving appropriate limiting characteristic functions that can be used to numerically evaluate the cumulative distribution functions. We also describe how our approach differs from the basic framework. Sections 4, 5, and 6 contain derivations of the proposed approximation to the limiting characteristic functions in each of the cases mentioned above for the structure of the errors, namely a) an AR root approaching 1, b) an AR root approaching -1, and c) an MA root approaching -1. In each case, the adequacy of the asymptotic approximation offered is assessed via simulations of the finite sample distribution. Section 7 offers concluding comments.

2. PRELIMINARIES.

It is useful at this point to review details about some existing asymptotic approximations in the nearly integrated first-order autoregressive model. As discussed in the introduction the data-generating process of interest is (1.1) and the estimator under consideration is (1.2). Limiting results about this OLS estimator are available under quite general conditions on the error sequence $\{u_t\}$. To be concrete we hereby state those originally considered by Herrndorf (1984) involving the concept of strong mixing.

ASSUMPTION 1: *(a)* $E(u_t) = 0$; *(b)* $\sup_t E|u_t|^{\beta+\epsilon} < \infty$ *for some* $\beta > 2$ *and* $\epsilon > 0$; *(c)* $\sigma^2 = \lim_{T\to\infty} T^{-1} E(S_T^2)$ *exists and* $\sigma^2 > 0$, *where* $S_t = \Sigma_1^t u_j$; *(d)* $\{u_t\}_1^\infty$ *is strong mixing with mixing numbers* α_m *that satisfy:* $\Sigma_1^\infty \alpha_m^{1-2/\beta} < \infty$.

When the sequence $\{u_t\}$ is strictly stationary condition (c) is implied by (a), (b) and (d) and $\sigma^2 = 2\pi f_u(0)$, where $f_u(0)$ is the non-normalized spectral density function of $\{u_t\}$ evaluated at frequency zero. The following notation is used throughout the paper: '$\Rightarrow$' denotes weak convergence in distribution; $\{e_t\}$ denotes a sequence of i.i.d. $(0, \sigma_e^2)$ random variables; W(r) is the unit Wiener process on C[0,1], the space of real-valued continuous functions on the [0,1] interval, and $J_c(r) = \int_0^r \exp((r-s)c)dW(r)$ is the Ornstein-Uhlenbeck process with unit variance and adjustment parameter c. Phillips (1987) proved the following Theorem under the conditions of Assumption 1 [1] (see also, Nabeya and Tanaka (1990, Theorem 4)).

LEMMA 1 *(Phillips (1987)): Let $\{y_t\}$ be generated by (1.1) with $\{u_t\}$ satisfying Assumption 1 and define $\tau = (\sigma^2 - \sigma_u^2)/(2\sigma^2)$, $\sigma_u^2 = \lim_{T\to\infty} T^{-1}E(\Sigma_{t=1}^T u_t^2)$. Then:*

$$T(\hat{\alpha} - \alpha) \Rightarrow \left\{ \int_0^1 J_c(r)dW(r) + \tau \right\} \left\{ \int_0^1 J_c(r)^2 dr \right\}^{-1}.$$

Also, $\lim_{T\to\infty} P[T(\hat{\alpha} - 1) \leq x]$ is given by $P[Z_x^0 \geq 0]$ with Z_x^0 is a random variable with characteristic function $D_x^0(2iw)^{-1/2}$ where, defining $\mu = (c^2 - 2\lambda x)^{1/2}$:

$$D_x^0(\lambda) = \exp(c - \lambda(1 - 2\tau))[\cosh(\mu) - (c - \lambda)\sinh(\mu)/\mu].$$

As discussed in the introduction, this limiting distribution provides an excellent approximation to the exact distribution when the errors are uncorrelated. However, as documented by Perron (1992), the quality of the approximation seriously deteriorates as the errors are more highly serially correlated, in particular when they are negatively correlated. To that effect, Nabeya and Perron (1992) considered alternative local asymptotic frameworks for restricted cases. Their results are summarized in the following lemmas.

LEMMA 2 *(Positive AR(1) errors): Let $\{y_t\}$ be generated by (1.1) and $\{u_t\}$ by $u_t = \exp(\phi/T)u_{t-1} + e_t$. Also define the function $Q_c(J_\phi(r)) \equiv \int_0^r \exp((r - v)c)J_\phi(v)dv$, then as $T \to \infty$:*

---

1 See also Perron (1992) for an extension considering $O_p(T^{-1})$ asymptotic expansions.

$$T(\hat{\alpha}-1) \Rightarrow (1/2)\, Q_c(J_\phi(1))^2 \left\{ \int_0^1 Q_c(J_\phi(r))^2 dr \right\}^{-1}.$$

Also, the limiting distribution of $T(\hat{\alpha}-1)$, $\lim_{T\to\infty} P[T(\hat{\alpha}-1) \le x]$, *is given by* $P(Z_x^1 \ge 0)$ *where* Z_x^1 *is a random variable with the characteristic function* $D_x^1(2iw)^{-1/2}$ *where:*

$$\begin{aligned} D_x^1(\lambda) = exp(c+\phi)\Big\{ & cosh(\mu_1)cosh(\mu_2) + c\phi\, sinh(\mu_1)sinh(\mu_2)/(\mu_1\mu_2) \\ & - (1/2)(c+\phi)[cosh(\mu_1)sinh(\mu_2)/\mu_2 + sinh(\mu_1)cosh(\mu_2)/\mu_1] \\ & + [\lambda + (c+\phi)(c-\phi)^2/2][cosh(\mu_1)sinh(\mu_2)/\mu_2 \\ & \qquad - sinh(\mu_1)cosh(\mu_2)/\mu_1]/(\mu_1^2-\mu_2^2) \\ & + 2\lambda(2x-c-\phi)[1-cosh(\mu_1)cosh(\mu_2) \\ & \qquad + (c^2+\phi^2)sinh(\mu_1)sinh(\mu_2)/(2\mu_1\mu_2)]/(\mu_1^2-\mu_2^2)^2 \Big\} \end{aligned}$$

with $\mu_1^2, \mu_2^2 = (1/2)\{c^2+\phi^2 \pm [(c^2-\phi^2)^2 + 8\lambda x]^{1/2}\}$.

LEMMA 3 *(Negative AR(1) errors): Let* $\{y_t\}$ *be generated by (1.1) and* $\{u_t\}$ *by* $u_t = -exp(\phi/T)u_{t-1} + e_t$. *Define the random functions* $J_{\phi,i}(s) \equiv \int_0^s exp((s-v)\phi)dW_i(v)$ *and* $Q_c(J_{\phi,i}(r)) \equiv \int_0^r exp((r-s)c)J_{\phi,i}(s)ds$ $(i=1,2)$*; where* $W_1(r)$ *and* $W_2(r)$ *are independent Wiener processes. Also let* $J_{c,1}(s) \equiv \int_0^s exp((s-v)c)dW_1(v)$. *Then as* $T\to\infty$:

$$\hat{\alpha} \Rightarrow 1 - 2\int_0^1 B(r)^2 dr \left\{ \int_0^1 \{[A(r)-B(r)]^2 + B(r)^2\} dr \right\}^{-1};$$

with $A(r) \equiv (\phi - c)[Q_c(J_{\phi,1}(r)) - Q_c(J_{\phi,2}(r))] + 2J_{c,1}(r)$ *and* $B(r) \equiv J_{\phi,1}(r) - J_{\phi,2}(r)$. *Also, the limiting distribution of* $\hat{\alpha}$, $\lim_{T\to\infty} P[\hat{\alpha} \le x]$, *is given by* $P(Z_x^2 \ge$

0) with Z_x^2 a random variable with characteristic function $D_x^2(2iw)^{-1/2}$ where:

$$D_x^2(\lambda) = exp(c+\phi)\Big\{[cosh(\{c^2 - \lambda(x-1)/2\}^{1/2})$$
$$- c\, sinh(\{c^2 - \lambda(x-1)/2\}^{1/2})/(c^2 - \lambda(x-1)/2)^{1/2}]$$
$$[cosh(\{\phi^2 - \lambda(x+1)/2\}^{1/2})$$
$$- \phi\, sinh(\{\phi^2 - \lambda(x+1)/2\}^{1/2})/(\phi^2 - \lambda(x+1)/2)^{1/2}]\Big\}.$$

LEMMA 4 *(Negative MA(1) errors): Let $\{y_t\}$ be generated by (1.1) with $y_0 = e_0 = 0$, and the sequence of errors generated by $u_t = e_t + \theta_T e_{t-1}$, $\theta_T = -1 + \delta/\sqrt{T}$, then as $T \to \infty$:*

$$\hat{\alpha} \Rightarrow \left\{\delta^2 \int_0^1 J_c(r)^2 dr\right\}\left\{1 + \delta^2 \int_0^1 J_c(r)^2 dr\right\}^{-1}.$$

Let $\mu = \{c^2 - \lambda\delta^2(x-1)\}^{1/2}$, the limiting distribution of $\hat{\alpha}$, $lim_{T\to\infty} P[\hat{\alpha} \leq x]$, is given by $P[Z_x^3 \geq 0]$ where Z_x^3 is a random variable with characteristic function $D_x^3(2iw)^{-1/2}$, where

$$D_x^3(\lambda) = exp(c - \lambda x)(cosh(\mu) - c\, sinh(\mu)/\mu).$$

The characteristic functions described in the above lemmas allow direct computation, by numerical integration, of the asymptotic cumulative distribution functions. The limiting distributions denoted by $F_j(x)$ $(j = 0, ..., 3)$ can be numerically evaluated using Imhof's (1961) formula:

$$F_j(x) = (1/2) + (1/\pi)\int_0^\infty (1/w) IM\{D_x^j(2iw)^{-1/2}\} dw\,, \qquad (2.1)$$

where $IM(\cdot)$ denotes the imaginary part of the argument. In computing the integrals the upper limit was set to a value w for which $|D_x^j(2iw)^{-1/2}| < 10^{-10}$ holds. Note that care must be exercised in evaluating these integrals since they involve the square root of complex valued quantities (see Perron (1989)).

3. OUTLINE OF THE FREDHOLM DETERMINANT APPROACH.

In this section, we briefly describe the so-called Fredholm determinant approach to obtain limiting characteristic functions. For a more detailed treatment oriented towards applications in econometrics, see Nabeya (1987) and Tanaka (1990). We suppose that we have under consideration a random variable ψ_T, say, with a non-degenerate limiting distribution and that we are interested in approximating its distribution, namely $P(\psi_T \leq x)$. What is assumed to hold is that $P(\psi_T \leq x)$ is equivalent to $P(S_T \geq 0)$ for some S_T of the form:

$$S_T = T^{-1}\Sigma_{j=1}^{T}\Sigma_{k=1}^{T}B_T(j,k)e_je_k, \tag{3.1}$$

where $\{B_T(j,k)\}$ is a deterministic sequence and $\{e_t\}$ is a sequence of i.i.d. random variables[2] with mean 0 and variance σ_e^2. The main ingredient of this approach is the existence of a symmetric and continuous function $K(s,t)$, called a kernel, defined on $[0,1] \times [0,1]$, that satisfies:

$$\lim_{T\to\infty}\max_{j,k}|B_T(j,k) - K(j/T,k/T)| = 0. \tag{3.2}$$

The kernel $K(s,t)$ is also assumed to be nearly definite in the sense that all but a finite number of eigenvalues of the following integral equation have the same sign:

$$f(t) = \lambda\int_0^1 K(s,t)f(s)ds. \tag{3.3}$$

The eigenvalues of the kernel $K(s,t)$ are those values of λ for which (3.3) has a non zero continuous solution $f(t)$. The solutions $f(t)$ are the eigenfunctions corresponding to λ. Denoting the sequences of eigenvalues and the associated

2 The method extends naturally to a sequence $\{e_t\}$ of random variables generated by a linear process (see Tanaka (1990)). Given the applications considered here, we only need $\{e_t\}$ to be i.i.d. and we restrict our attention to this special case.

orthonormal eigenfunctions by $\{\lambda_n, f_n(t);\ n = 1,\ldots, \infty\}$, Mercer's Theorem states that

$$K(s,t) = \Sigma_{n=1}^{\infty}\lambda_n^{-1}f_n(s)f_n(t). \tag{3.4}$$

Denoting by $D(\lambda)$ the Fredholm determinant associated with the kernel $K(s,t)$ we have:

$$D(\lambda) = \Pi_{n=1}^{\infty}[1-(\lambda/\lambda_n)], \tag{3.5}$$

and $D(\lambda)$ is an entire function of λ with $D(0) = 1$. In some applications, however, the Fredholm determinant $D(\lambda)$ is obtained using the integral equation (3.3). The procedure is described in Nabeya and Tanaka (1988).

The usefulness of this approach lies in the fact that the kernel $K(s,t)$ and the associated Fredholm determinant $D(\lambda)$ contain all the relevant information to characterize $\lim_{T\to\infty}P[S_T \geq 0]$ or, equivalently, the limiting distribution of the random variable of interest, namely $\lim_{T\to\infty}P[\psi_T \leq x]$. This characterization is summarized in the following lemma proved in Nabeya and Tanaka (1988).

LEMMA 5: *As $T\to\infty$, we have $(S_T/\sigma_e^2) \Rightarrow \int_0^1\int_0^1 K(s,t)dW(s)dW(t)$, a random variable with characteristic function $D(2iw)^{-1/2}$.*

Our modification to this general procedure can be described as follows. Suppose we can decompose $B_T(j,k)$ as follows:

$$B_T(j,k) = \Sigma_{i=1}^{m}p_{i,T}B_{i,T}(j,k), \tag{3.6}$$

where $p_{i,T}$ $(i = 1, \ldots, m)$ are functions that depend on T but not on the indices j and k. In this framework, we have $p_{i,T} \to p_i$, say, as $T \to \infty$, and $\lim_{T\to\infty}\max_{j,k}|B_{i,T}(j,k) - K_i(j/T,k/T)| = 0$, say, such that $\Sigma_{i=1}^{m}p_iK_i(s,t) \equiv K(s,t)$, the original kernel. The modification proposed is then to use, in place of the limiting kernel $K(s,t)$, the function

$$K_T(s,t) \equiv \Sigma_{i=1}^{m} p_{i,T} K_i(s,t). \tag{3.7}$$

Note that, since $p_{i,T} \to p_i$ as $T \to \infty$ and $\Sigma_{i=1}^{m} p_i K_i(s,t) \equiv K(s,t)$, the modified kernel still satisfies a requirement similar to (3.2), namely:

$$\lim_{T\to\infty} \max_{j,k} | B_T(j,k) - K_T(j/T,k/T) | = 0. \tag{3.8}$$

Therefore, all the apparatus described before can be used. In the end, however, what is obtained is a characteristic function $D_T(2iw)^{-1/2}$ that depends on the sample size but which converges to the limiting characteristic function $D(2iw)^{-1/2}$ as $T \to \infty$.

It is important to note that the modified kernel $K_T(s,t)$ depends on sample size only through functions which do not depend on s or t. This allows the modified procedure to remain analytically tractable. Since we use the finite sample counterparts $p_{i,T}$ instead of their limit, we can expect the corresponding approximation to the characteristic function to be closer to the finite sample characteristic function. Hence, we can expect a better approximation to the finite sample distribution. Though we have not been able to show that the difference $|B_T(j,k) - K_T(j/T,k/T)|$ is uniformly smaller than the difference $|B_T(j,k) - K(j/T,k/T)|$, our results in the next sections provide strong evidence of a much improved approximation.

4. APPROXIMATIONS IN THE NEARLY TWICE INTEGRATED MODEL.

In this Section we consider an approximation to the distribution of $T(\hat{\alpha} - 1)$ when the errors have an AR(1) structure with (large) positive correlation. We have the following local parameterization:

$$y_t = \exp(c/T) y_{t-1} + u_t , \tag{4.1}$$

$$u_t = \exp(\phi/T)u_{t-1} + e_t , \tag{4.2}$$

where, for simplicity, we specify $e_t \sim$ i.i.d. $(0,\sigma_e^2)$ and $y_0 = y_{-1} = 0$. We can write $\{y_t\}$ as :

$$y_t = [\exp(c/T) + \exp(\phi/T)]y_{t-1} - \exp((c+\phi)/T)y_{t-2} + e_t . \tag{4.3}$$

As T converges to infinity $\{y_t\}$ becomes $y_t = 2y_{t-1} - y_{t-2} + e_t$; a process with two unit roots, hence the expression "nearly twice integrated". The inequality $T(\hat{\alpha}-1) \leq x$ is equivalent to the inequality $xV_T - U_T \geq 0$, where

$$U_T = 2T^{-3}(\Sigma_{t=2}^T y_t y_{t-1} - \Sigma_{t=2}^T y_{t-1}^2) , \tag{4.4}$$

$$V_T = 2T^{-4}\Sigma_{t=2}^T y_{t-1}^2 . \tag{4.5}$$

For ease of exposition, we let $\alpha_1 = \exp(c/T)$ and $\alpha_2 = \exp(\phi/T)$ and consider first the case

$$c \neq 0 , \phi \neq 0 \text{ and } c^2 \neq \phi^2 . \tag{4.6}$$

From equations (A.8) and (A.9) of Nabeya and Perron (1992) we have:

$$xV_T - U_T = \frac{1}{T}\Sigma_{j=1}^T \Sigma_{k=1}^T (1/T^2(\alpha_1 - \alpha_2)^2)\{p_T B_{1,T}(j,k) + q_T B_{2,T}(j,k) + r_T B_{3,T}(j,k)\} e_j e_k, \tag{4.7}$$

where

$$B_{1,T}(j,k) = (\alpha_1^{|j-k|} - \alpha_1^{2T-j-k+2}),$$

$$B_{2,T}(j,k) = (\alpha_2^{|j-k|} - \alpha_2^{2T-j-k+2}),$$

$$B_{3,T}(j,k) = -(\alpha_1^{T-j+1} - \alpha_2^{T-j+1})(\alpha_1^{T-k+1} - \alpha_2^{T-k+1}),$$

and

$$p_T = \frac{2x}{T}\left(\frac{1}{1-\alpha_1^2} - \frac{1}{1-\alpha_1\alpha_2}\right) + \frac{2}{1+\alpha_1} - \frac{2-\alpha_1-\alpha_2}{1-\alpha_1\alpha_2},$$

$$q_T = \frac{2x}{T}\left(\frac{1}{1-\alpha_2^2} - \frac{1}{1-\alpha_1\alpha_2}\right) + \frac{2}{1+\alpha_2} - \frac{2-\alpha_1-\alpha_2}{1-\alpha_1\alpha_2}, \qquad (4.8)$$

$$r_T = \frac{2x}{T}\frac{1}{1-\alpha_1\alpha_2} + \frac{2-\alpha_1-\alpha_2}{1-\alpha_1\alpha_2}.$$

Since $T^2(\alpha_1 - \alpha_2)^2/(c-\phi)^2 > 0$ (including the case of conjugate complex c and ϕ), we consider the inequality $S_T \geq 0$ instead of $xV_T - U_T \geq 0$, where

$$S_T = \{T^2(\alpha_1-\alpha_2)^2/(c-\phi)^2\}(xV_T - U_T)$$

$$= \frac{1}{T}\Sigma_{j=1}^T\Sigma_{k=1}^T\{\Sigma_{i=1}^3 p_{i,T}B_{i,T}(j,k)\}e_je_k,$$

with $p_{1,T} = p_T/(c-\phi)^2$, $p_{2,T} = q_T/(c-\phi)^2$, $p_{3,T} = r_T/(c-\phi)^2$. The inequality $S_T \geq 0$ is used since it allows simplifications in the statement of the approximate distribution that will be given in Theorem 1. We note the following limit relations:

$$p_{1,T} \to x\left(\frac{2}{c+\phi} - \frac{1}{c}\right)/(c-\phi)^2 \equiv p_1,$$

$$p_{2,T} \to x\left(\frac{2}{c+\phi} - \frac{1}{\phi}\right)/(c-\phi)^2 \equiv p_2,$$

$$p_{3,T} \to \left(1 - \frac{2x}{c+\phi}\right)/(c-\phi)^2 \equiv p_3.$$

Letting $j = Ts$ and $k = Tt$, we also have the limit relations [3]:

$$B_{1,T}(j,k) \to (e^{c|s-t|} - e^{c(2-s-t)}) \equiv K_1(s,t),$$

---

3 Note that $\Sigma_{i=1}^3 p_i K_i(s,t)$ corresponds to the limiting kernel used in Nabeya and Perron (1992), $xK_D(s,t) - K_N(s,t)$ in their notation.

$$B_{2,T}(j,k) \to (e^{\phi|s-t|} - e^{\phi(2-s-t)}) \equiv K_2(s,t),$$

$$B_{3,T}(j,k) \to -(e^{c(1-s)} - e^{\phi(1-s)})(e^{c(1-t)} - e^{\phi(1-t)}) \equiv K_3(s,t).$$

Given the strategy discussed in the introduction, we thereby approximate $P[T(\hat{\alpha}-1) \leq x]$ by $P[\int_0^1\int_0^1 K_T^1(s,t)dW(s)dW(t) \geq 0]$, where

$$K_T^1(s,t) \equiv \Sigma_{i=1}^3 p_{i,T} K_i(s,t). \tag{4.9}$$

The Fredholm determinant associated with the kernel $K_T^1(s,t)$ is expressed in Theorem 1.

THEOREM 1: *Under the conditions (4.6), the Fredholm determinant, $D_{T,x}^1(\lambda)$, associated with the kernel $K_T^1(s,t)$ defined by (4.9) is:*

$$\begin{aligned} D_{T,x}^1(\lambda) = exp(c+\phi)\Bigg[& cosh\,\mu_1 cosh\,\mu_2 + c\phi\frac{sinh\ \mu_1}{\mu_1}\frac{sinh\ \mu_2}{\mu_2} \\ & - \frac{c+\phi}{2}(cosh\,\mu_1\frac{sinh\ \mu_2}{\mu_2} + \frac{sinh\ \mu_1}{\mu_1}cosh\,\mu_2) \\ & + \frac{1}{\mu_1^2-\mu_2^2}[\lambda(r_T + (cp_T - \phi q_T)/(c-\phi)) + (c+\phi)(c-\phi)^2/2] \\ & \quad x\,[cosh\,\mu_1\frac{sinh\ \mu_2}{\mu_2} - \frac{sinh\ \mu_1}{\mu_1}cosh\,\mu_2] \\ & - 2\lambda\frac{(c+\phi)r_T}{(\mu_1^2-\mu_2^2)^2}\{1 - cosh\,\mu_1 cosh\,\mu_2 \\ & \quad + (\frac{c^2+\phi^2}{2} + \lambda\frac{cp_T+\phi q_T}{(c-\phi)^2})\frac{sinh\ \mu_1}{\mu_1}\frac{sinh\ \mu_2}{\mu_2}\}\Bigg], \end{aligned}$$

where μ_1^2 and μ_2^2 are the roots of the quadratic equation in z,

$$z^2 - 2z\left(\frac{c^2+\phi^2}{2} + \lambda\frac{cp_T + \phi q_T}{(c-\phi)^2}\right) + c^2\phi^2 + 2\lambda c\phi\frac{cq_T + \phi p_T}{(c-\phi)^2} = 0,$$

and, hence

$$\mu_1^2 - \mu_2^2 = \left\{(c^2-\phi^2)^2 + 4\lambda\frac{(c+\phi)(cp_T - \phi q_T)}{c-\phi} + 4\lambda^2\frac{(cp_T+\phi q_T)^2}{(c-\phi)^4}\right\}^{1/2}.$$

The distribution $P[T(\hat{\alpha}-1)\leq x]$ can be approximated by $P[\tilde{Z}_x^1 \geq 0]$ where $\tilde{Z}_x^1$ is a random variable with characteristic function $D_{T,x}^1(2iw)^{-1/2}$.

Remark (1): The above expression for $D_{T,x}^1(\lambda)$ presumes that conditions (4.6) are satisfied. If any of the conditions is violated, the corresponding Fredholm determinant can be evaluated by taking a suitable limiting process to $D_{T,x}^1(\lambda)$ as defined by Theorem 1, noting the definitions of p_T, q_T and r_T as stated in (4.8). For example, if we let $c \to 0$ and $\phi \to 0$, we have $(cp_T + \phi q_T)/(c-\phi)^2 \to -(x/6T^2 + 1/2T)$, $(c+\phi)(cp_T - \phi q_T)/(c-\phi) \to 2x$, $r_T + (cp_T - \phi q_T)/(c-\phi) \to x/T + 1$, and $(c+\phi)r_T \to -2x$. Therefore, when $c = \phi = 0$:

$$D_{T,x}^1(\lambda) = \cosh\mu_1\cosh\mu_2 + \frac{\lambda}{\mu_1^2-\mu_2^2}\left(\frac{x}{T}+1\right)\left(\cosh\mu_1\frac{\sinh\mu_2}{\mu_2} - \frac{\sinh\mu_1}{\mu_1}\cosh\mu_2\right)$$
$$+\frac{4\lambda x}{(\mu_1^2-\mu_2^2)^2}\left\{1-\cosh\mu_1\cosh\mu_2 - \lambda\left(\frac{x}{6T^2}+\frac{1}{2T}\right)\frac{\sinh\mu_1}{\mu_1}\frac{\sinh\mu_2}{\mu_2}\right\},$$

where

$$\mu_1^2, \mu_2^2 = -\lambda\left(\frac{x}{6T^2}+\frac{1}{2T}\right) \pm \left\{2\lambda x + \lambda^2\left(\frac{x}{6T^2}+\frac{1}{2T}\right)^2\right\}^{1/2}.$$

Remark (2): As $T \to \infty$, we have $(cp_T + \phi q_T)/(c-\phi)^2 \to 0$, $(c+\phi)(cp_T - \phi q_T)/(c-\phi) \to 2x$, $r_T + (cp_T - \phi q_T)/(c-\phi) \to 1$, and $(c+\phi)r_T \to c + \phi - 2x$. Hence,

$D^1_{T,x}(\lambda)$ converges to $D^1_x(\lambda)$ as defined by Lemma 2, and we recover the result of Nabeya and Perron (1992).

Remark (3): The support of the limiting distribution of $T(\hat{\alpha} - 1)$ considered in Nabeya and Perron (1992) and described by Lemma 2 is limited to the interval $[0,\infty)$, and its density approaches ∞ as x approaches 0 from above. On the other hand, the support of both the finite sample distribution of $T(\hat{\alpha}-1)$ and the approximation obtained using Theorem 1 are not limited to the interval $[0,\infty)$. We can, accordingly, expect our modified approximation to be more accurate, especially for values of x near 0.

Table I.a presents results assessing the adequacy of various approximations to the finite sample distribution. It shows the maximum of the absolute distance between the empirical distribution (based on 10,000 replications) and three asymptotic approximations, namely: 1) our modified approximation obtained using the result of Theorem 1 (denoted NP1), the local asymptotic distribution of Nabeya and Perron (1992) described in Lemma 2 (denoted NP2), and 3) the standard asymptotic distribution described by Lemma 1 (denoted PHL). In conducting the simulations, we used uniform random numbers [4] for the sequence $\{e_t\}$ and the sample sizes are T = 25, 50 and 100. We considered 18 cases with the various combinations of (c, α_2) obtained using c = 0, -5, 2, and $\alpha_2 = \exp(\phi/T) = 0.1, 0.3, 0.5, 0.7, 0.9, 0.95$. The results show that NP1 is far better than the other two approximations. Though not presented here, the two curves showing the cumulative distribution function from the empirical and the NP1-approximation would be almost indistinguishable even for T = 25.

Table I.b considers in more detail the quality of the NP1-approximation in the tails of the distribution. The entries provide the signed maximum distance between the finite sample distribution and the NP1 approximation. The left tail is defined to include percentage point lower than 5% and the right tail percentage points greater than 95%. Two facts are worth noting from these

---

[4] The use of uniform random variables allows us to assess the validity of the approximation with errors that are clearly different from normality. The use of normal variables yields, however, qualitatively similar results.

TABLE I.a

Maximum distance between the empirical distributions and its asymptotic approximations.

Nearly Twice Integrated Model; $y_t = \exp(c/T)y_{t-1} + u_t$, $u_t = \alpha_2 u_{t-1} + e_t$; $\alpha_2 = \exp(\phi/T)$.

		T = 25			T = 50			T = 100		
c	α_2	NP1	NP2	PHL	NP1	NP2	PHL	NP1	NP2	PHL
0	.1	.021	.617	.021	...	...	.016	...	...	.008
0	.3	.013	.527	.018	.008	.538	.010	...	...	.009
0	.5	.010	.422	.034	.011	.435	.017	.003	.435	.011
0	.7	.010	.303	.086	.005	.312	.048	.009	.321	.021
0	.9	.015	.162	.222	.005	.162	.155	.012	.173	.076
0	.95	.021	.128	.303	.007	.115	.244	.012	.118	.151
–5	.1	.024	.992	.049	...	...	.027	...	...	.015
–5	.3	.016	.969	.036	.009	.973	.014	...	...	.010
–5	.5	.014	.915	.043	.008	.924	.032	.005	.922	.017
–5	.7	.018	.778	.083	.010	.794	.052	.006	.806	.033
–5	.9	.011	.486	.168	.012	.497	.104	.005	.507	.057
–5	.95	.014	.364	.210	.006	.342	.138	.004	.353	.097
2	.1	.009	.198	.060	...	...	.030	...	...	.017
2	.3	.013	.156	.066	.010	.154	.037	...	...	.017
2	.5	.015	.124	.087	.006	.124	.054	.005	.127	.025
2	.7	.013	.094	.136	.005	.096	.071	.007	.091	.031
2	.9	.024	.055	.282	.011	.048	.172	.012	.051	.081
2	.95	.014	.056	.368	.018	.041	.263	.012	.040	.159

Note: Some entries are left blanked due to the fact that the Fortran functions cosh(·) and sinh(·) are ill behaved when the absolute value of the real part of the argument is very large.

TABLE I.b

Adequacy of the NP1 Approximation in the Tails.

Nearly Twice Integrated Model; $y_t = \exp(c/T)y_{t-1} + u_t$, $u_t = \alpha_2 u_{t-1} + e_t$; $\alpha_2 = \exp(\phi/T)$.

		T = 25		T = 50		T = 100	
c	α_2	Left	Right	Left	Right	Left	Right
0	.1	−.012	.004	...	...	...	...
0	.3	−.002	.005	−.002	.003	...	...
0	.5	−.001	.004	−.003	.004	.003	.001
0	.7	−.006	.005	−.003	−.001	−.002	−.003
0	.9	−.004	.006	.002	−.002	−.001	.003
0	.95	.003	.006	.002	.004	.002	.001
−5	.1	−.010	.003	...	...	...	...
−5	.3	−.002	−.002	−.003	−.003	...	...
−5	.5	−.003	.004	.003	.003	−.003	−.002
−5	.7	−.006	.003	−.003	.003	.004	.002
−5	.9	−.002	.004	.004	.006	.001	−.003
−5	.95	−.004	.005	−.001	.005	.001	−.002
2	.1	−.003	.004	...	...	...	...
2	.3	−.007	.004	−.004	.004	...	...
2	.5	−.004	.005	−.003	−.002	.001	.004
2	.7	−.001	.005	.001	.001	−.004	−.001
2	.9	−.003	.010	−.004	.005	.003	.005
2	.95	−.003	.004	.002	.006	.004	.003

Notes: Let $F_T(x)$ denote the finite sample distribution and $F_{NP1}(x)$ its NP1 approximation. The entries in the columns labelled "Left" report the value of $F_T(x) - F_{NP1}(x)$ whose absolute value is largest in the range $\min\{F_T(x), F_{NP1}(x)\} \leq .05$. Similarly, the entries in the columns labelled "Right" report the values of $F_T(x) - F_{NP1}(x)$ whose absolute value is largest in the range $\max\{F_T(x), F_{NP1}(x)\} \geq .95$. Some entries are left blanked because the Fortran functions cosh(·) and sinh(·) are ill behaved when the absolute value of the real part of the argument is very large.

results. First, the approximations in the tails are better than the approximations in the center of the distribution, namely the range (.05, .95). Indeed, the distances reported in Table I.b are generally smaller (in absolute value) than those reported in Table I.a. This feature is particularly encouraging since oftentimes the tails are of more interest, for example when hypothesis testing is involved. The second feature of interest is that for small sample sizes the approximation to the c.d.f. lies above the exact c.d.f. in the left tail and below it in the right tail. Hence, critical values taken from the approximation would yield slightly conservative tests on either side. Note, however, that this discrepancy is very small even for $T = 25$.

5. THE NEARLY INTEGRATED SEASONAL MODEL.

In this Section we consider an approximation to the distribution of $\hat{\alpha}$ when the errors have an AR(1) structure with (large) negative correlation. We have the following local parameterization:

$$y_t = \exp(c/T)y_{t-1} + u_t, \tag{5.1}$$

$$u_t = -\exp(\phi/T)u_{t-1} + e_t, \tag{5.2}$$

where $e_t \sim$ i.i.d. $(0,\sigma_e^2)$ and for simplicity $e_0 = u_0 = 0$. Combining (5.1) and (5.2):

$$y_t = [\exp(c/T) - \exp(\phi/T)]y_{t-1} + \exp((c+\phi)/T)y_{t-2} + e_t. \tag{5.3}$$

As T increases to infinity $\{y_t\}$ approaches the process $y_t = y_{t-2} + e_t$ which characterizes a seasonal model of period 2 with a root on the unit circle. We therefore label the process (5.1) and (5.2) as a "nearly integrated seasonal model". We note first that the inequality $\hat{\alpha} \leq x$ is equivalent to the inequality $xV_T^* - U_T^* \geq 0$ where $U_T^* = 2T^{-2}\Sigma_{t=2}^T y_t y_{t-1}$ and $V_T^* = 2T^{-2}\Sigma_{t=2}^T y_{t-1}^2$. For simplicity of notation, we specify $\alpha_1 = \exp(c/T)$ and $\alpha_2 = -\exp(\phi/T)$. Straightforward algebra yields the following expression:

$$xV_T^* - U_T^* = \frac{1}{(\alpha_1-\alpha_2)^2}\Sigma_{j=1}^{T}\Sigma_{k=1}^{T}\frac{1}{T}\Big\{p_T(\alpha_1^{|j-k|} - \alpha_1^{2T-j-k+2})$$

$$+ q_T(\alpha_2^{|j-k|} - \alpha_2^{2T-j-k+2}) - r_T(\alpha_1^{2T-j-k+2} + \alpha_2^{2T-j-k+2})\Big\}e_j e_k$$

$$+ \frac{2r_T}{T(\alpha_1-\alpha_2)^2}(\Sigma_{j=1}^{T}\alpha_1^{T-j+1}e_j)(\Sigma_{j=1}^{T}\alpha_2^{T-j+1}e_j) \qquad (5.4)$$

where

$$p_T = \frac{1}{T}\Big(\frac{2(x-\alpha_1)}{1-\alpha_1^2} - \frac{2x-\alpha_1-\alpha_2}{1-\alpha_1\alpha_2}\Big),$$

$$q_T = \frac{1}{T}\Big(\frac{2(x-\alpha_2)}{1-\alpha_2^2} - \frac{2x-\alpha_1-\alpha_2}{1-\alpha_1\alpha_2}\Big),$$

$$r_T = \frac{1}{T}\frac{2x-\alpha_1-\alpha_2}{1-\alpha_1\alpha_2}.$$

Note that $P[xV_T^* - U_T^* \geq 0] = P[(\alpha_1 - \alpha_2)^2(xV_T^* - U_T^*) \geq 0]$ since $(\alpha_1 - \alpha_2)^2 > 0$. Hence, we can, without loss of generality, consider approximating $P[\hat{\alpha} \leq x]$ by $P[S_T \geq 0]$ where, using (5.4):

$$S_T = \frac{1}{T}\Sigma_{j=1}^{T}\Sigma_{k=1}^{T}\Big\{p_T(\alpha_1^{|j-k|} - \alpha_1^{2T-j-k+2}) + q_T(\alpha_2^{|j-k|} - \alpha_2^{2T-j-k+2})$$

$$- r_T(\alpha_1^{2T-j-k+2} + \alpha_2^{2T-j-k+2})\Big\}e_j e_k + O_p(T^{-1}). \qquad (5.5)$$

The contributions of terms with $j = T$ or $k = T$ in (5.5) can be shown to be of order $O_p(T^{-3/2})$ and can therefore be neglected if needed. We use this fact to consider from now on that T is an even number ($T = 2L$, say) since it allows us to state that omitting the elements corresponding to $j = T$ or $k = T$ when T is odd introduces negligible errors [5].

---

[5] In fact, the contribution of the terms with $j = T$ or $k = T$ in (5.5) is

We now introduce the sequence of random variables $\{\xi_j, \eta_j; j = 1, \ldots, L \equiv T/2\}$. We define $\xi_j = (1/\sqrt{2})(e_{2j-1} + e_{2j})$ and $\eta_j = (1/\sqrt{2})(e_{2j-1} - e_{2j})$. Note that ξ_j and η_j are uncorrelated and i.i.d. sequences with mean 0 and variance σ_e^2 just as the original sequence $\{e_t\}$. This specification implies that:

$$e_{2j-1} = \frac{1}{\sqrt{2}}(\xi_j + \eta_j) \text{ and } e_{2j} = \frac{1}{\sqrt{2}}(\xi_j - \eta_j), \quad (j = 1, \ldots, L),$$

and upon rearrangements:

$$S_T = T^{-1}\Sigma_{m=1}^{L}\Sigma_{n=1}^{L}(a_{mn}\xi_m\xi_n + 2b_{mn}\xi_m\eta_n + c_{mn}\eta_m\eta_n) + O_p(T^{-1}), \tag{5.6}$$

where [6]

$T^{-1}\{p_T(1 - \alpha_1^2) - r_T\alpha_1^2\}(2e_T\Sigma_{j=1}^{T-1}\alpha_1^{T-j}e_j + e_T^2) + T^{-1}\{q_T(1 - \alpha_2^2) - r_T\alpha_2^2\}(2e_T\Sigma_{j=1}^{T-1}\alpha_2^{T-j}e_j + e_T^2)$. Since $p_T(1 - \alpha_1^2) - r_T\alpha_1^2$ and $q_T(1 - \alpha_2^2) - r_T\alpha_2^2$ are $O(T^{-1})$ as $T \to \infty$, the whole expression is $O_p(T^{-3/2})$ and can be neglected.

6 The exact values are as follows. Let $A_{mn} = p_T(2 + 2\alpha_1) + q_T(2 + 2\alpha_2)$, if $m = n$, $A_{mn} = p_T(1 + \alpha_1)^2\alpha_1^{2|m-n|-1} + q_T(1 + \alpha_2)^2\alpha_2^{2|m-n|-1}$, if $m \neq n$; $B_{mn} = \text{sgn}(n - m)\{p_T(1 - \alpha_1^2)\alpha_1^{2|m-n|-1} + q_T(1 - \alpha_2^2)\alpha_2^{2|m-n|-1}\}$; $C_{mn} = p_T(2 - 2\alpha_1) + q_T(2 - 2\alpha_2)$, if $m = n$, and $C_{mn} = -p_T(1 - \alpha_1)^2\alpha_1^{2|m-n|-1} - q_T(1 - \alpha_2)^2\alpha_2^{2|m-n|-1}$, if $m \neq n$. Then $a_{mn} = (1/2)\{A_{mn} - (p_T + r_T)(1 + \alpha_1)^2\alpha_1^{4L-2m-2n+2} - (q_T + r_T)(1 + \alpha_2)^2\alpha_2^{4L-2m-2n+2}\}$; $b_{mn} = (1/2)\{B_{mn} + (p_T + r_T)(1 - \alpha_1^2)\alpha_1^{4L-2m-2n+2} + (q_T + r_T)(1 - \alpha_2^2)\alpha_2^{4L-2m-2n+2}\}$; and $c_{mn} = (1/2)\{C_{mn} - (p_T + r_T)(1 - \alpha_1)^2\alpha_1^{4L-2m-2n+2} - (q_T + r_T)(1 - \alpha_2)^2\alpha_2^{4L-2m-2n+2}\}$.

$$a_{mn} = a_{nm} = 2p_T\alpha_1^{2|m-n|} - 2(p_T + r_T)\alpha_1^{4L-2m-2n+2} + O(T^{-1}),$$
$$b_{mn} = O(T^{-1}),$$
$$c_{mn} = c_{nm} = 2q_T\alpha_2^{2|m-n|} - 2(q_T + r_T)\alpha_2^{4L-2m-2n+2} + O(T^{-1}).$$

Using the above relations, we can express (5.6) as:

$$S_T = T^{-1}\Sigma_{m=1}^{L}\Sigma_{n=1}^{L}\Big[\Big\{2p_T\alpha_1^{2|m-n|} - 2(p_T + r_T)\alpha_1^{4L-2m-2n+2}\Big\}\xi_m\xi_n + \Big\{2q_T\alpha_2^{2|m-n|} - 2(q_T + r_T)\alpha_2^{4L-2m-2n+2}\Big\}\eta_m\eta_n\Big] + O_p(T^{-1}). \quad (5.7)$$

Note that (5.7) is the sum of two terms of the form (3.6), viz.:

$$S_T = S_T^A + S_T^B + O_p(T^{-1})$$
$$= T^{-1}\Sigma_{m=1}^{L}\Sigma_{n=1}^{L}\Big\{\Sigma_{i=1}^{2}p_{i,T}^{A}B_{i,T}^{A}(m,n)\Big\}\xi_m\xi_n + T^{-1}\Sigma_{m=1}^{L}\Sigma_{n=1}^{L}\Big\{\Sigma_{i=1}^{2}p_{i,T}^{B}B_{i,T}^{B}(m,n)\Big\}\eta_m\eta_n + O_p(T^{-1}), \quad (5.8)$$

where $p_{1,T}^A \equiv p_T$, $p_{2,T}^A \equiv p_T + r_T$, $p_{1,T}^B \equiv q_T$, $p_{2,T}^B \equiv q_T + r_T$, $B_{1,T}^A(m,n) \equiv 2\alpha_1^{2|m-n|}$, $B_{2,T}^A(m,n) \equiv -2\alpha_1^{4L-2m-2n+2}$, $B_{1,T}^B(m,n) \equiv 2\alpha_2^{2|m-n|}$, $B_{2,T}^B(m,n) \equiv -2\alpha_2^{4L-2m-2n+2}$.

Consistently with our strategy to find better approximations, we consider the limit of terms in (5.8) which do not involve the $p_{i,T}$'s. Using the fact that $\alpha_1 = \exp(c/T) = \exp(c/2L)$ and $\alpha_2 = -\exp(\phi/T) = -\exp(\phi/2L)$ and denoting $m = sL$ and $n = tL$, we have the approximate kernels corresponding to S_T^A and S_T^B:

$$K_{A,T}(s,t) = p_T\exp(c|s-t|) - (p_T + r_T)\exp(c(2-s-t)), \quad (5.9)$$

$$K_{B,T}(s,t) = q_T\exp(\phi|s-t|) - (q_T + r_T)\exp(\phi(2-s-t)). \quad (5.10)$$

Our approximation for $P[\hat{\alpha} \leq x]$ is therefore

$$P[\int_0^1\int_0^1 K_{A,T}(s,t)dW_1(s)dW_1(t) + \int_0^1\int_0^1 K_{B,T}(s,t)dW_2(s)dW_2(t) \geq 0], \tag{5.11}$$

where $W_1(\cdot)$ and $W_2(\cdot)$ are independent Wiener processes, constructed from the partial sums of $\{\xi_n\}$ and $\{\eta_n\}$, respectively. The Fredholm determinant for the kernels $K_{A,T}(s,t)$ and $K_{B,T}(s,t)$ are stated in the following Theorem.

THEOREM 2: *Under the conditions that $c \neq 0$ and $\phi \neq 0$, the Fredholm determinants, $D^A_{T,x}(\lambda)$ and $D^B_{T,x}(\lambda)$, associated with the kernels (5.9) and (5.10) are (denoting $a_A \equiv c$ and $a_B \equiv \phi$):*

$$D^j_{T,x}(\lambda) = exp(a_j)\{cosh\,\mu_j - (a_j - \lambda r_T)\frac{sinh\,\mu_j}{\mu_j}\}, \qquad (j = A, B),$$

where $\mu_A^2 = c^2 + 2\lambda c p_T$ and $\mu_B^2 = \phi^2 + 2\lambda\phi q_T$. The distribution $P[\hat{\alpha} \leq x]$ can be approximated by $P[\tilde{Z}_x^2 \geq 0]$ where $\tilde{Z}_x^2$ is a random variable with characteristic function $D^2_{T,x}(2iw)^{-1/2} \equiv [D^A_{T,x}(2iw)D^B_{T,x}(2iw)]^{-1/2}$.

Remark (1): If $c = 0$, Theorem 2 applies with the modification $\mu_A^2 = 2\lambda(1 - r_T - x)$. Similarly if $\phi = 0$, the modification is $\mu_B^2 = -2\lambda(1 + x + r_T)$.

Remark (2): If we let $T \to \infty$, we have $p_T \to (1-x)/c$, $q_T \to -(1+x)/\phi$, $r_T \to 0$ and $D^j_{T,x}(\lambda) \to \exp(a_j)[\cosh(\mu_j) - a_j \sinh(\mu_j)/\mu_j]$, for $j = A, B$ where $\mu_A^2 = c^2 - 2\lambda(x - 1)$ and $\mu_B^2 = \phi^2 - 2\lambda(x + 1)$. This result corresponds to that stated in Theorem 6 of Nabeya and Perron (1992) (see Lemma 3 of Section 2).

Table II.a presents results showing the maximum of the absolute distance between the empirical distribution (based on 10,000 replications) and three asymptotic approximations, namely: 1) our modified approximation obtained

TABLE II.a

Maximum distance between the empirical distributions and its asymptotic approximations.

Nearly Integrated Seasonal Model; $y_t = \exp(c/T)y_{t-1} + u_t$, $u_t = \alpha_2 u_{t-1} + e_t$; $\alpha_2 = -\exp(\phi/T)$.

		T = 25			T = 50			T = 100		
c	α_2	NP1	NP2	PHL	NP1	NP2	PHL	NP1	NP2	PHL
0	–.1	.022	.306	.023	...	...	.015	...	...	.004
0	–.3	.020	.224	.038	.004	.220	.016	...	...	.013
0	–.5	.018	.157	.070	.005	.149	.034	.006	.150	.021
0	–.7	.018	.103	.145	.012	.090	.077	.007	.081	.037
0	–.9	.007	.040	.389	.006	.030	.228	.007	.027	.124
0	–.95	.008	.027	.602	.010	.022	.402	.010	.016	.237
-5	–.1	.041	.180	.267	...	...	.268	...	...	.268
–5	–.3	.024	.056	.173	.014	.056	.173	...	...	.173
–5	–.5	.020	.019	.294	.013	.017	.154	.009	.012	.086
–5	–.7	.012	.028	.513	.012	.014	.297	.005	.013	.151
–5	–.9	.007	.021	.947	.010	.009	.739	.006	.007	.465
–5	–.95	.007	.023	.997	.010	.007	.954	.005	.007	.749
2	–.1	.010	.764	.495	...	...	.500	...	...	.489
2	–.3	.018	.718	.545	.014	.719	.546	...	...	.537
2	–.5	.014	.650	.567	.007	.638	.555	.007	.641	.558
2	–.7	.020	.550	.527	.009	.533	.509	.007	.526	.508
2	–.9	.016	.354	.345	.012	.317	.301	.014	.296	.273
2	–.95	.022	.290	.278	.009	.217	.201	.011	.177	.147

See note to Table I.a.

TABLE II.b

Adequacy of the NP1 Approximation in the Tails.

Nearly Integrated Seasonal Model; $y_t = \exp(c/T)y_{t-1} + u_t$, $u_t = \alpha_2 u_{t-1} + e_t$; $\alpha_2 = -\exp(\phi/T)$.

		T = 25		T = 50		T = 100	
c	α_2	Left	Right	Left	Right	Left	Right
0	–.1	–.020	.010	...	...	...	...
0	–.3	–.008	.007	–.003	.002	...	...
0	–.5	–.004	.004	–.002	–.003	–.002	–.002
0	–.7	–.009	–.004	–.005	–.007	.001	–.001
0	–.9	–.002	–.003	.002	–.001	.002	.002
0	–.95	.003	–.002	–.001	–.002	.002	–.002
–5	–.1	–.033	.027	...	...	...	...
–5	–.3	–.020	.013	–.011	.009	...	...
–5	–.5	–.015	.009	–.009	.005	.001	.003
–5	–.7	–.006	.002	–.005	.002	–.002	–.002
–5	–.9	–.003	.003	–.003	.003	–.006	.004
–5	–.95	.004	.001	–.002	–.002	.003	.001
2	–.1	–.005	.009	...	...	...	...
2	–.3	–.007	.003	–.004	.001	...	...
2	–.5	–.007	–.003	–.002	–.002	.002	–.002
2	–.7	–.002	–.018	–.002	–.008	–.006	–.001
2	–.9	–.004	–.010	–.003	–.009	.003	–.002
2	–.95	–.002	–.012	.002	–.007	.002	–.005

See notes to Table I.b.

using the result of Theorem 2 (denoted NP1), the local asymptotic distribution of Nabeya and Perron (1992) described in Lemma 3 (denoted by NP2), and 3) the standard asymptotic distribution described by Lemma 1 (denoted by PHL). In conducting the simulations, we used uniform random numbers for the sequence $\{e_t\}$ and the sample sizes are T = 25, 50 and 100. We considered 18 cases with the various combinations of (c, α_2) obtained using c = 0, -5, 2, and $\alpha_2 = -\exp(\phi/T) = -.1, -.3, -.5, -.7, -.9, -.95$. The superiority of NP1 is again evident although in three cases with c = -5 NP1 provides a less accurate approximation than NP2 (the differences are, however, only in the third decimals). Again, the two curves showing the cumulative distribution function from the empirical and the NP1-approximation would be almost indistinguishable even for T = 25.

Table II.b considers in more detail the quality of the NP1-approximation in the tails of the distribution. The analysis is similar to that reported in Section 4 for the nearly twice integrated case. The general results are also similar. First, the approximations in the tails are better than the approximations in the center of the distribution, and indeed very accurate. Secondly, for a small sample size the approximation to the c.d.f. again lies, in general, above the exact c.d.f. in the left tail and below it in the right tail. Hence, critical values taken from the approximation would, also in this case, yield slightly conservative tests on either side. The discrepancies are, however, very small even for T = 25.

6. APPROXIMATIONS IN THE NEARLY WHITE NOISE MODEL.

In this Section we consider an approximation to the distribution of $\hat{\alpha}$ when the errors have an MA(1) structure with (large) negative correlation. We have the following local parameterization:

$$y_t = \exp(c/T)y_{t-1} + u_t , \tag{6.1}$$

$$u_t = e_t + \theta_T e_{t-1}. \tag{6.2}$$

We assume $e_t \sim$ i.i.d. $(0, \sigma_e^2)$ and $y_0 = e_0 = 0$. The process defined by (6.1) and (6.2) is an ARMA(1,1) where the autoregressive root approaches 1 and the moving average root approaches - 1 as T converges to infinity. In the limit, the roots cancel and the process $\{y_t\}$ is white noise provided the sequence $\{e_t\}$ is white noise. However, in any finite sample, $\{y_t\}$ is nearly integrated, hence the expression "nearly white noise - nearly integrated model". For simplicity of notation, we specify $\alpha_T = \exp(c/T)$.

Define $U_T = 2T^{-1}\Sigma_{t=2}^{T} y_{t-1}(y_t - y_{t-1})$ and $V_T = 2T^{-1}\Sigma_{t=2}^{T} y_{t-1}^2$, then the inequality $\hat{\alpha} \leq x$ is equivalent to the inequality:

$$[U_T - (1-\alpha_T)^2 V_T/2\alpha_T]/V_T \quad = \hat{\alpha} - (1+\alpha_T^2)/2\alpha_T$$
$$\leq x - (1+\alpha_T^2)/2\alpha_T \equiv z_T. \qquad (6.3)$$

Expanding U_T and V_T as:

$$U_T = -2[T(1+\alpha_T)]^{-1}\Sigma_{j=1}^{T}\Sigma_{k=1}^{T}(\alpha_T^{|j-k|} - \alpha_T^{2T-j-k})u_j u_k$$
$$+ (T\alpha_T)^{-1}\Sigma_{j=1}^{T}\Sigma_{k=1}^{T}\alpha_T^{|j-k|}u_j u_k - (T\alpha_T)^{-1}\Sigma_{j=1}^{T}u_j^2,$$

and

$$V_T = 2[T(1-\alpha_T^2)]^{-1}\Sigma_{j=1}^{T-1}\Sigma_{k=1}^{T-1}(\alpha_T^{|j-k|} - \alpha_T^{2T-j-k})u_j u_k,$$

we have

$$U_T - (1-\alpha_T)^2 V_T/2\alpha_T = (T\alpha_T)^{-1}(\Sigma_{j=1}^{T}\alpha_T^{T-j}u_j)^2 - (T\alpha_T)^{-1}\Sigma_{j=1}^{T}u_j^2.$$

Using (6.1) and (6.2), we also deduce that:

$$\alpha_T[U_T - (1-\alpha_T)^2 V_T/2\alpha_T] = T^{-1}\Big[\{e_T + (\alpha_T + \theta_T)\Sigma_{j=1}^{T-1}\alpha_T^{T-j-1}e_j\}^2$$
$$-e_T^2 - (1+\theta_T^2)\Sigma_{j=1}^{T-1}e_j^2 - 2\theta_T\Sigma_{j=2}^{T}e_j e_{j-1}\Big] \qquad (6.4)$$

and

$$\alpha_T V_T = [2(\alpha_T + \theta_T)/(1-\alpha_T^2)]$$

$$\left\{T^{-1}\Sigma_{j=1}^{T-1}\Sigma_{k=1}^{T-1}\{(1+\alpha_T\theta_T)\alpha_T^{|j-k|}-(\alpha_T+\theta_T)\alpha_T^{2T-j-k-1}\}e_je_k\right\}$$

$$-2\theta_T T^{-1}\Sigma_{j=1}^{T-1} e_j^2. \qquad (6.5)$$

Defining $W_T = \alpha_T\{z_T V_T - (U_T - (1-\alpha_T)^2 V_T/2\alpha_T)\}$, the inequality (6.3) is equivalent to $W_T \geq 0$, with W_T expressed as:

$$W_T = W_{1,T} + W_{2,T}, \qquad (6.6)$$

where

$$W_{1,T} \equiv T^{-1}\Sigma_{j=1}^{T-1}\Sigma_{k=1}^{T-1} B_T(j,k)e_je_k + 2T^{-1/2}b_Te_T\,\Sigma_{j=1}^{T-1} B_T(j,T)e_j, \qquad (6.7)$$

$$W_{2,T} \equiv T^{-1}(1+\theta_T^2 - 2\theta_T z_T)\Sigma_{j=1}^{T-1} e_j^2 + 2\theta_T T^{-1}\Sigma_{j=2}^{T} e_je_{j-1}, \qquad (6.8)$$

with

$$B_T(j,k) = p_T\alpha_T^{|j-k|} + q_T\alpha_T^{2T-j-k}, (j,k=1,\dots,T), \qquad (6.9)$$

and

$$b_T = -T^{-1/2}(\alpha_T+\theta_T)/[\alpha_T(p_T+q_T)],$$

$$p_T = 2(\alpha_T+\theta_T)(1+\alpha_T\theta_T)z_T/(1-\alpha_T^2),$$

$$q_T = -(\alpha_T+\theta_T)^2(2z_T/\alpha_T(1-\alpha_T^2)+\alpha_T^{-2}).$$

Note that the first component $W_{1,T}$ defined by (6.7) involves $B_T(j,k)$ defined by (6.9). Following our strategy to obtain improved approximations, we consider the following associated kernel

$$K_T^3(s,t) = p_T e^{c|s-t|} + q_T e^{c(2-s-t)}. \tag{6.10}$$

The component $(W_{1,T}/\sigma_e^2)$ can then be be approximated by the random variable Q_T, where:

$$Q_T \equiv \int_0^1 \int_0^1 K_T^3(s,t)\, dW(s) dW(t) + 2b_T(e_T/\sigma_e) \int_0^1 K_T^3(s,1) dW(s). \tag{6.11}$$

Under the assumption that $e_T \sim N(0, \sigma_e^2)$, the characteristic function of Q_T can be obtained using Theorem 3 in Tanaka (1990) [7]. The Fredholm determinant $D_T^3(\lambda)$ for $K_T^3(s,t)$ and the value of the resolvent $R_T^3(s,t;\lambda)$ for $K_T^3(s,t)$ evaluated at $s = t = 1$ are given by

$$D_T^3(\lambda) = \exp(c)[\cosh \mu - \{c + \lambda(p_T + q_T)\} \frac{\sinh \mu}{\mu}], \tag{6.12}$$

and

$$R_T^3(1,1;\lambda) = \exp(c)(p_T + q_T)(\cosh \mu - c \frac{\sinh \mu}{\mu})/D_T^3(\lambda), \tag{6.13}$$

where $\mu^2 = c^2 + 2cp_T\lambda$. The characteristic function of Q_T is then given by:

$$[D_T^3(2iw) \{1 - 2ib_T^2 w(R_T(1,1;2iw) - K_T(1,1))\}]^{-1/2} = \exp(-c/2)[\cosh \mu - (c + 2iw(p_T + q_T) + (2iw)^2(\alpha_T + \theta_T)^2/T\alpha_T^2) \frac{\sinh \mu}{\mu}]^{-1/2},$$

where $\mu^2 = c^2 + 4icp_T w$.

Consider now the distribution of the second component $W_{2,T}$ defined by

---

7 For the following derivations to hold, the common distribution of $\{e_1, \ldots, e_{T-1}\}$ need not be the same as that of e_T, though it is assumed that they have the same mean and variance.

(6.8). Using the fact that $T^{-1}\Sigma_{j=1}^{T-1} e_j^2 \to \sigma_e^2$ and assuming that e_t has finite fourth moment, we have the following approximation:

$$(W_{2,T}/\sigma_e^2) \simeq N(1 + \theta_T^2 - 2\theta_T z_T,\ T^{-1}[(1 + \theta_T^2 - 2\theta_T z_T)^2\gamma + 4\theta_T^2]), \quad (6.14)$$

where γ is the variance of (e_t^2/σ_e^2) [8]. To obtain a manageable expression, we proceed under the assumption that $W_{1,T}$ and $W_{2,T}$ are asymptotically independent. The error introduced by this approximation will be small, for large samples, if α_T is close to 1 and θ_T is close to -1. Under this approximation we are now able to describe the approximate characteristic function of (W_T/σ_e^2) which we state in the following Theorem [9].

THEOREM 3: *Consider the case where $c \neq 0$. Let $\mu^2 = c^2 + 2\lambda c p_T$ and p_T, q_T be defined by (6.9). Under the approximation that $W_{1,T}$ and $W_{2,T}$ are asymptotically independent and under the assumption that $e_T \sim N(0, \sigma_e^2)$ and that e_t has finite fourth moment, the distribution function $P[\hat{\alpha} \leq x]$ can be approximated by $P[\tilde{Z}_x^3 \geq 0]$, with $\tilde{Z}_x^3$ a random variable with characteristic function $D_{T,x}^3(2iw)^{-1/2}$, where:*

$$D_{T,x}^3(\lambda) = exp\Big\{c - \lambda(1 + \theta_T^2 - 2\theta_T z_T)$$

$$- (\lambda^2/4T)[(1 + \theta_T^2 - 2\theta_T z_T)^2\gamma + 4\theta_T^2]\Big\}$$

$$\times \left\{\cosh\mu - (c + \lambda(p_T + q_T) + \lambda^2(\alpha_T + \theta_T)^2/T\alpha_T^2)\frac{\sinh\mu}{\mu}\right\}$$

---

8 If $e_t \sim N(0,\sigma_e^2)$ we have $\gamma = 2$, while if e_t has the uniform distribution over the interval $[-\sqrt{3}\sigma_e, \sqrt{3}\sigma_e]$, $\gamma = 4/5$.

9 Note that we use the fact that $P[W_T \geq 0] = P[(W_T/\sigma_e^2) \geq 0]$ since $\sigma_e^2 > 0$.

Remark (1): As $c \to 0$, $\alpha_T \to 1$, $z_T \to x - 1$, $cp_T \to T(1 + \theta_T)^2(1 - x)$ and $p_T + q_T \to -(1 + \theta_T)(1 - \theta_T + 2\theta_T x)$, hence the characteristic function for the case $c = 0$ is obtained from Theorem 3 using these limits and $\mu^2 = 2\lambda T(1 + \theta_T)^2(1 - x)$.

Remark (2): The result of Theorem 3 allows existing approximations as special cases. Consider first treating θ_T fixed at θ as T increases. The limit of $\mathbf{P}[T(\hat{\alpha} - 1) \leq x]$ as $T \to \infty$ is then equivalent to $P[\tilde{Z}_x^3 \geq 0]$ where the characteristic function of $\tilde{Z}_x^3$ is given by Theorem 3 replacing z_T by x/T and taking the limit as $T \to \infty$, which then reduces to:

$$\exp\{(-c/2) + iw(1 + \theta^2)\}\,[\cosh \mu - \{c - 2iw(1 + \theta)^2\} \sinh(\mu)/\mu]^{-1/2},$$

where $\mu^2 = c^2 - 4i(1 + \theta)^2 xw$. This characteristic function corresponds to that in Phillips (1987) and in Nabeya and Tanaka (1990, Theorem 4) (see also Lemma 1 of Section 2).

Consider now the case where $\theta_T = -1 + \delta/\sqrt{T}$. The limit, as $T \to \infty$, of the characteristic function in Theorem 3 (noting the inequality (6.3)) is:

$$\exp((-c/2) + 2ixw)(\cosh \mu - c \sinh(\mu)/\mu)^{-1/2},$$

where $\mu^2 = c^2 + 4i\delta^2(1 - x)w$, which is the result derived by Nabeya and Perron (1992, Theorem 2) (see also Lemma 4 of Section 2).

Remark (3): The characteristic function of Theorem 3 can also be applied to the framework considered by Pantula (1991). He analyzed the null distribution (i.e. with $c = 0$) of several unit root tests specifying a sequence of local values of the form $\theta_T = -1 + \delta T^{-d}$ with $d > 0$. Among the tests he considered is the normalized bias $T(\hat{\alpha} - 1)$. The results above permit the analysis of the limit of $T(\hat{\alpha} - 1)$, in a framework slightly more general than Pantula's, i.e. under a sequence of local alternative indexed by c. The quantities of interest are $\alpha_T[U_T - (1-\alpha_T)^2 V_T/2\alpha_T]$ and $\alpha_T V_T$ as defined by (6.4) and (6.5). If d = 0 (Phillips' (1987) approximation), then using (6.4):

$$\alpha_T[U_T - (1-\alpha_T)^2 V_T/2\alpha_T] \Rightarrow \sigma_e^2(1+\theta)^2\{\int_0^1 \exp(c(1-s))dW(s)\}^2 - \sigma_e^2(1+\theta^2)$$

whereas if d > 0, the first term in (6.4) vanishes, as $T \to \infty$, and

$$\alpha_T[U_T - (1-\alpha_T)^2 V_T/2\alpha_T] \Rightarrow -2\sigma_e^2.$$

If d = 1/2 (Nabeya and Perron's (1992) approximation), then

$$\alpha_T V_T \Rightarrow \delta^2\sigma_e^2\int_0^1\int_0^1 [(\exp(c(2-s-t) - \exp(c|s-t|))/c]dW(s)dW(t) + 2\sigma_e^2.$$

If d < 1/2, the first term in the right side of (6.5) is dominant and

$$T^{-1+2d}\alpha_T V_T \Rightarrow \delta^2\sigma_e^2\int_0^1\int_0^1 [(\exp(c(2-s-t)) - \exp(c|s-t|))/c]dW(s)dW(t).$$

On the other hand, if d > 1/2, the second term is dominant and $\alpha_T V_T \Rightarrow 2\sigma_e^2$. It is important to note that in all the cases discussed above, some terms in (6.4) or (6.5) vanish as $T \to \infty$ and, hence, do not make a contribution to the asymptotic distribution. The superiority of the approximation offered in this paper can be understood by noting that many terms are taken into account as completely as possible so that they contribute to a better approximation.

Table III.a presents results showing the maximum of the absolute distance between the empirical distribution (based on 10,000 replications) and three asymptotic approximations, namely: 1) our modified approximation obtained using the result of Theorem 3 (denoted NP1), the local asymptotic distribution of Nabeya and Perron (1992) described in Lemma 4 (denoted by NP2), and 3) the standard asymptotic distribution described by Lemma 1 (denoted by PHL). As before, we used uniform random numbers for the sequence $\{e_t\}$ and the sample sizes are T = 25, 50 and 100. We considered 18 cases with the various

TABLE III.a

Maximum distance between the empirical distributions and its asymptotic approximations.

Nearly White Noise Model; $y_t = \exp(c/T)y_{t-1} + u_t$, $u_t = e_t + \theta_T e_{t-1}$; $\theta_T = -1 + \delta/\sqrt{T}$.

		T = 25			T = 50			T = 100		
c	θ_T	NP1	NP2	PHL	NP1	NP2	PHL	NP1	NP2	PHL
0	–.1	.017	.380	.025	.009	.375	.017	.008	.361	.005
0	–.3	.009	.303	.061	.008	.285	.029	.005	.287	.020
0	–.5	.013	.233	.175	.010	.207	.096	.007	.199	.055
0	–.7	.025	.192	.479	.014	.156	.321	.010	.126	.187
0	–.9	.042	.281	.980	.031	.189	.921	.027	.107	.804
0	–.95	.035	.393	1.000	.027	.310	1.000	.029	.208	.992
–5	–.1	.051	.363	.108	.023	.409	.053	.018	.447	.031
–5	–.3	.024	.295	.258	.016	.348	.142	.008	.381	.073
–5	–.5	.030	.227	.604	.010	.258	.383	.009	.296	.215
–5	–.7	.053	.342	.968	.038	.176	.862	.019	.182	.641
–5	–.9	.063	.625	1.000	.011	.505	1.000	.023	.314	1.000
–5	–.95	.079	.664	1.000	.032	.593	1.000	.008	.496	1.000
2	–.1	.012	.762	.058	.006	.766	.027	.009	.755	.016
2	–.3	.007	.704	.053	.006	.697	.027	.005	.685	.015
2	–.5	.013	.583	.068	.018	.558	.037	.005	.564	.020
2	–.7	.016	.408	.186	.012	.359	.123	.009	.337	.075
2	–.9	.023	.334	.635	.012	.216	.499	.010	.146	.372
2	–.95	.022	.476	.930	.022	.320	.827	.019	.190	.689

TABLE III.b

Adequacy of the NP1 Approximation in the Tails.

Nearly White Noise Model; $y_t = \exp(c/T)y_{t-1} + u_t$, $u_t = e_t + \theta_T e_{t-1}$; $\theta_T = -1 + \delta/\sqrt{T}$.

		T = 25		T = 50		T = 100	
c	θ_T	Left	Right	Left	Right	Left	Right
0	–.1	–.016	.008	–.009	.003	–.004	.004
0	–.3	.002	.009	–.001	.002	–.002	.001
0	–.5	.013	.008	.006	–.001	.002	–.001
0	–.7	.019	.005	.011	–.005	.006	.002
0	–.9	.017	.006	.022	–.001	.020	–.002
0	–.95	–.007	.013	.010	–.005	.015	–.006
–5	–.1	–.033	.042	–.019	.015	–.012	.010
–5	–.3	–.003	.023	–.006	.011	–.003	.005
–5	–.5	.028	.014	.009	.006	.007	.002
–5	–.7	.026	.011	.030	–.004	.019	–.005
–5	–.9	–.063	.029	–.011	.009	.013	–.008
–5	–.95	–.078	.033	–.027	.018	–.005	.002
2	–.1	–.002	.004	–.002	.002	–.002	.003
2	–.3	–.003	.004	–.002	.002	–.002	.003
2	–.5	.003	.003	.003	.003	.003	.001
2	–.7	.015	.005	.008	.002	–.003	.004
2	–.9	.012	.013	.009	.004	.010	.002
2	–.95	.005	.009	.012	.003	.011	.003

See notes to Table I.b.

combinations of (c, θ_T) obtained using c = 0, -5, 2, and $\theta_T = -1 + \delta/\sqrt{T}$ = -.1, -.3, -.5, -.7, -.9, -.95. In computing the NP1-approximation, we specified $\gamma = 2$, the value corresponding to the normal distribution even though uniform random numbers are used to simulate the empirical distributions [10]. The results show the NP1 approximation to be clearly the best [11]. The maximum distances are, however, higher than those reported in the previous two sections especially for c = -5.

Table III.b considers in more detail the quality of the NP1-approximation in the tails of the distribution. Here the qualitative results are somewhat different from the previous two cases. First, comparing Table III.a and III.b, it is seen that, in general, most of the errors in approximating the exact distribution comes from errors in approximating the left tail. Indeed, the entries in the column "left" are, in general, larger in absolute value than those under the column "right" and are close to the maximal distances reported in Table III.a. As in the previous two cases, the approximation to the c.d.f. in the right tail is, for a small sample size, below the exact distribution. Hence, critical values taken from the approximation would yield slightly conservative right sided tests. However, contrary to previous cases, there is no such clear pattern in the left tail.

7. CONCLUSIONS.

This paper has considered alternative asymptotic approximations in a specialized dynamic model with dependent errors. The method is based on substituting a suitable finite sample version for the limiting kernel in the so-called Fredholm determinant approach often used to derive approximating characteristic functions. Our results are very encouraging since they show the modified procedure to yield excellent approximations to the finite sample

---

10 We performed sensitivity analyses by using different values of γ to generate the approximate distributions. The results showed the same qualitative features. In particular, the adequacy of the NP1-approximation is similar to that reported here.

11 The only case where NP1 is not the best is c = 0, θ_T = -.1 and T = 100, but the difference is in the third decimal place.

distribution when other existing approximations fail to be adequate.

At the heart of the method is the idea to use the finite sample counterparts of components in the coefficients $B_T(j,k)$ that do not depend on the indices j or k. This allows analytical tractability in deriving the associated Fredholm determinant, yet allows much improved accuracy as our experiments show. Two remarks about this procedure are of immediate interest.

First, when considering the standard asymptotic approximation (denoted PHL in the Tables), the kernels considered are limits of coefficients associated with the products $u_j u_k$ in quadratic forms in the sequence of variables $\{u_j\}$ which are, in the general case considered, substantially correlated. In the approximation considered by Nabeya and Perron (1992), denoted NP2, the kernels used pertain to limits of coefficients associated with products $e_j e_k$ where $\{e_j\}$ is a sequence of i.i.d. random variables. The superiority of NP2 over PHL is thereby likely to be due to the inadequacy of asymptotic approximations provided by limiting results involving substantially correlated variables (see also Perron (1992)). In the alternative asymptotic approximations considered in this paper (denoted NP1), the kernels used again involve coefficients associated with products $e_j e_k$ with $\{e_j\}$ i.i.d.. The approximate kernels used are, however, not the limiting counterparts but are rather suitable functions of the sample size that are expected to approximate more closely the finite sample coefficients associated with $e_j e_k$.

The second remark concerns the theoretical justification, or lack thereof, of the proposed modification. Indeed, as stated in Section 3, we have not been able to show that the difference $|B_T(j,k) - K_T(j/T,k/T)|$ is uniformly smaller than the difference $|B_T(j,k) - K(j/T,k/T)|$, or that the approximate Fredholm determinants yield characteristic functions closer to their finite sample counterparts than those implied using their limiting form. This issue is clearly an important topic for further research.

REFERENCES

Bobkoski, M.J. (1983): Hypothesis Testing in Nonstationary Time Series. Unpublished PhD Dissertation, Univeristy of Wisconsin, Madison.

Cavanagh, C. (1986): "Roots Local to Unity," mimeo, Harvard University.

Chan, N. H. (1988): "The Parameter Inference for Nearly Nonstationary Time Series," *Journal of the American Statistical Association*, 83, 857-862.

Chan, N.H., and C.Z. Wei (1987): "Asymptotic Inference for Nearly Nonstationary AR(1) Processes," *Annals of Statistics*, 15, 1050-1063.

Gallant, A.R., and H. White (1988): *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*, Oxford: Basil Blackwell.

Herrndorf, N. (1984): "A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables," *Annals of Probability*, 12, 141-153.

Imhof, P.J. (1961): "Computing the Distribution of Quadratic Forms in Normal Variables," *Biometrika*, 48, 419-426.

Nabeya, S. (1987): "Application of the Theory of Integral Equations to the Derivation of Asymptotic Distributions in Time Series Analysis," Unpublished PhD Dissertation, Hitotsubashi University, Tokyo.

Nabeya, S., and P. Perron (1992): "Local Asymptotic Distributions Related to the AR(1) Model with Dependent Errors," forthcoming in *Journal of Econometrics*.

Nabeya, S., and K. Tanaka (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., and K. Tanaka (1990): "A General Approach to the Limiting Distribution for Estimators in Time Series Regression with Nonstable Autoregressive Errors," *Econometrica*, 58, 145-163.

Pantula, S.G. (1991): "Asymptotic Distributions of Unit Root Tests when the Process is Nearly Stationary," *Journal of Business and Economic Statistics*, 9, 63-72.

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. (1991a): "A Continuous Time Approximation to the Unstable First-Order Autoregressive Process: The Case without an Intercept," *Econometrica*, 59, 211-236.

Perron, P. (1991b): "A Continuous-Time Approximation to the Stationary First-order Autoregressive Model," *Econometric Theory*, 7, 236-252.

Perron, P. (1992): "The Adequacy of Asymptotic Approximations in the Near-integrated Model with Dependent Errors," manuscript, Université de Montréal.

Phillips, P.C.B. (1987): "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74, 535-47.

Tanaka, K. (1990): "The Fredholm Approach to Asymptotic Inference on Nonstationary and Noninvertible Time Series Models," *Econometric Theory*, 6, 411-432.

White, H. (1984): *Asymptotic Theory for Econometricians*, Academic Press.

Université de Montréal
Département de sciences économiques
Centre de documentation
C.P. 6128, succursale Centre-ville
Montréal (Québec)
H3C 3J7

Cahiers de recherche (Discussion papers)
1993 à aujourd'hui (1993 to date)

Si vous désirez obtenir un exemplaire, vous n'avez qu'à faire parvenir votre demande et votre paiement (5 $ l'unité) à l'adresse ci-haut mentionnée. / To obtain a copy ($ 5 each), please send your request and prepayment to the above-mentioned address.

9301 : Mercenier, Jean, "Nonuniqueness of Solutions in Applied General-Equilibrium Models with Scale Economies and Imperfect Competition : A Theoretical *Curiosum*?", 26 pages.
9302 : Lemieux, Thomas, "Unions and Wage Inequality in Canada and in the United States", 66 pages.
9303 : Lemieux, Thomas, "Estimating the Effects of Unions on Wage Inequality in a Two-Sector Model with Comparative Advantage and Non-Random Selection", 54 pages.
9304 : Harchaoui, Tarek H., "Time-Varying Risks and Returns : Evidence from Mining Industries Data", 20 pages.
9305 : Lévy-Garboua, Louis et Claude Montmarquette, "Une étude économétrique de la demande de théâtre sur données individuelles", 40 pages.
9306 : Montmarquette, Claude, Rachel Houle et Sophie Mahseredjian, "The Determinants of University Dropouts : A Longitudinal Analysis", 17 pages.
9307 : Gaudry, Marc, Benedikt Mandel et Werner Rothengatter, "A Disagregate Box-Cox Logit Mode Choice Model of Intercity Passenger Travel in Germany", 17 pages.
9308 : Fortin, Nicole M., "Borrowing Constraints and Female Labor Supply : Nonparametric and Parametric Evidence of the Impact of Mortgage Lending Rules", 38 pages.
9309 : Dionne, Georges, Robert Gagné, François Gagnon et Charles Vanasse, "Debt, Moral Hazard and Airline Safety : an Empirical Evidence", 34 pages.
9310 : Dionne, Georges, Anne Gibbens et Pierre St-Michel, "An Economic Analysis of Insurance Fraud", 40 pages.
9311 : Gaudry, Marc, "Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes", 26 pages.
9312 : Laferrière, Richard et Marc Gaudry, "Testing the Linear Inverse Power Transformation Logit Mode Choice Model", 29 pages.

9313 : Kollmann, Robert, "Fiscal Policy, Technology Shocks and the US Trade Balance Deficit", 38 pages.

9314 : Ghysels, Eric, "A Time Series Model With Periodic Stochastic Regime Switching", 54 pages.

9315 : Allard, Marie, Camille Bronsard et Lise Salvas-Bronsard, "C^*-Conjugate Expectations and Duality", 22 pages.

9316 : Dudley, Leonard et Claude Montmarquette, "Government Size and Economic Convergence", 28 pages.

9317 : Bronsard, Camille, "L'histoire de l'économie mathématique racontée à Juliette", 17 pages.

9318 : Tremblay, Rodrigue, "The Quest for Competitiveness and Export-Led Growth", 16 pages.

9319 : Proulx, Pierre-Paul, "L'ALÉNA", 12 pages.

9320 : Proulx, Pierre-Paul, "Le Québec dans l'ALÉNA", 28 pages.

9321 : Dionne, Georges, Denise Desjardins, Claire Laberge-Nadeau et Urs Magg, "Medical Conditions, Risk Exposure and Truck Drivers' Accidents : an Analysis with Count Data Regression Models, 20 pages.

9322 : Ghysels, Eric, "Seasonal Adjustment and Other Data Transformations", 28 pages.

9323 : Dufour, Jean-Marie et David Tessier, "On the Relationship Between Impulse Response Analysis, Innovation Accounting and Granger Causality", 12 pages.

9324 : Dufour, Jean-Marie et Eric Renault, "Causalités à court et à long terme dans les modèles VAR et ARIMA multivariés", 68 pages.

9325 : Ghysels, Eric et Alastair Hall, "On Periodic Time Series and Testing the Unit Root Hypothesis", 36 pages.

9326 : Campbell, Bryan et Jean-Marie Dufour, "Exact Nonparametric Orthogonality and Random Walk Tests", 28 pages.

9327 : Proulx, Pierre-Paul, "Quebec in North America : from a Borderlands to a Borderless Economy : an Examination of its Trade Flows with the U.S.A. at the National and Regional Levels", 24 pages.

9328 : Proulx, Pierre-Paul, "L'ALÉNA, le Québec et la mutation de son espace économique", 36 pages.

9329 : Sprumont, Yves, "Strategyproof Collective Choice in Economic and Political Environments", 48 pages.

9330 : Cardia, Emanuela et Steve Ambler, "The Cyclical Behaviour of Wages and Profits Under Imperfect Competition", 24 pages.

9331 : Arcand, Jean-Louis L. et Elise S. Brezis, "Disequilibrium Dynamics During the Great Depression", 64 pages.

9332 : Beaudry, Paul et Michel Poitevin, "Contract Renegotiation : a Simple Framework and Implications for Organization Theory", 48 pages.

9333 : Ghysels, Eric et Alastair Hall, "On Periodic Time Series and Testing the Unit Root Hypothesis", 36 pages. (Ce cahier portait déjà le no 9325.)

9333 : Dagenais, Marcel G. et Jean-Marie Dufour, "Pitfalls of Rescaling Regression Models with Box-Cox Transformations", 16 pages.

9334 : Bonomo, Marco et René Garcia, "Disappointment Aversion as a Solution to the Equity Premium and the Risk-Free Rate Puzzles", 40 pages.

9335 : Ghysels, Eric et Offer Lieberman, "Dynamic Regression and Filtered Dat Series : A Laplace Approximation to the Effects of Filtering in Small Samples' 24 pages.

9336 : Kollmann, Robert, "The Duration of Unemployment as A Signal : Implication for Labor Market Equilibrium", 19 pages.

9337 : Kollmann, Robert, "Fertility, Consumption and Bequests in a Model with Non Dynastic Parental Altruism", 22 pages.

9401 : Mercenier, Jean et Bernardin Akitoby, "On Intertemporal General-Equilibrium Reallocation Effects of Europe's move to a Single Market", janvier 1994, 41 pages.

9402 : Gauthier, Céline et Michel Poitevin, "Using Ex Ante Payments in Self-Enforcing Risk-Sharing Contracts, février 1994, 38 pages.

9403 : Ghysels, Eric et Joanna Jasiak, "Stochastic Volatility and Time Deformation : an Application of Trading Volume and Leverage Effects, février 1994, 37 pages.

9404 : Dagenais, Marcel G. et Denyse L. Dagenais, "GMM Estimators for Linear Regression Models with Errors in the Variables", avril 1994, 33 pages.

9405 : Bronsard, C., F. Rosenwald et L. Salvas-Bronsard, "Evidence on Corporate Private Debt Finance and the Term Structure of Interest Rates", avril 1994, 42 pages.

9406 : Dinardo, John, Nicole M. Fortin et Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992 : A Semiparametric Approach", avril 1994, 73 pages.

9407 : Campbell, Bryan et Jean-Marie Dufour, "Exact Nonparametric Tests of Orthogonality and Random Walk in the Presence of a Drift Parameter", avril 1994, 32 pages.

9408 : Bollerslev, Tim et Eric Ghysels, "Periodic Autoregressive Conditional Heteroskedasticity", mai 1994, 29 pages.

9409 : Cardia, Emanuela, "The Effects of Government Financial Policies : Can We Assume Ricardian Equivalence?", mai 1994, 42 pages.

9410 : Kollmann, Robert, "Hidden Unemployment : A Search Theoretic Interpretation", mai 1994, 9 pages.

9411 : Kollmann, Robert, "The Correlation of Productivity Growth Across Regions and Industries in the US", juin 1994, 14 pages.

9412 : Gaudry, Marc, Benedikt Mandel et Werner Rothengatter, "Introducing Spatial Competition through an Autoregressive Contiguous Distributed (AR-C-D) Process in Intercity Generation-Distribution Models within a Quasi-Direct Format (QDF)", juin 1994, 61 pages.

9413 : Gaudry, Marc et Alexandre Le Leyzour, "Improving a Fragile Linear Logit Model Specified for High Speed Rail Demand Analysis in the Quebec-Windsor Corridor of Canada, août 1994, 39 pages.

9414 : Lewis, Tracy et Michel Poitevin, "Disclosure of Information in Regulatory Proceedings", juillet 1994, 38 pages.

9415 : Ambler, Steve, Emanuela Cardia et Jeannine Farazli, "Export Promotion and Growth", août 1994, 41 pages.

9416 : Ghysels, Eric et Haldun Sarlan, "On the Analysis of Business Cycles Through the Spectrum of Chronologies", août 1994, 37 pages.

9417 : Martel, Jocelyn et Timothy C.G. Fisher, "The Creditors' Financial Reorganization Decision : New Evidence from Canadian Data", août 1994, 21 pages.

9418 : Cannings, Kathy, Claude Montmarquette et Sophie Mahseredjian, "Entrance Quotas and Admission to Medical Schools : A Sequential Probit Model", septembre 1994, 26 pages.

9419 : Cannings, Kathy, Claude Montmarquette et Sophie Mahseredjian, "Major Choices : Undergraduate Concentrations and the Probability of Graduation", septembre 1994, 26 pages.

9420 : Nabeya, Seiji et Pierre Perron, Approximations to Some Exact Distributions in the First Order Autoregressive Model with Dependent Errors", septembre 1994, 40 pages.