LOCAL ASYMPTOTIC DISTRIBUTIONS RELATED TO THE AR(1) MODEL WITH DEPENDENT ERRORS

Seiji Nabeya
Tokyo International University

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

Econometric Research Program
Research Memorandum No. 362

May 1991 - Revised July 1991

This paper is a revised version of Sections 5 through 7 of Perron (1990a) and of Nabeya (1991). We are grateful to Bent E. Sørensen. This research was supported by grants from the Social Sciences and Humanities Research Council of Canada, the Natural and Engineering Council of Canada and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec. Some of the computations were carried out using the computer FACOM M-760/6 at the Computer Center, Hitotsubashi University which we thank for their support.

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

ABSTRACT

We consider the normalized least squares estimator of the parameter in a nearly integrated first–order autoregressive model with dependent errors. The dependence in the errors is modeled as either an MA(1) or an AR(1) process. As discussed in Perron (1991a), the usual asymptotic distribution is a poor guide to the finite sample distribution in the cases where i) the MA root approaches –1 ; and ii) the AR root approaches either 1 or –1. This occurs even for large sample sizes. The aim of this paper is to provide alternative asymptotic frameworks that treat the MA and AR roots as being local to their boundary. The appropriate limiting distributions are derived as well as the limiting characteristic functions allowing tabulation of distributional quantities via numerical integration. The results presented in this paper provide a better approximation to the finite sample distribution and helps explain many of the finite sample results discussed in Perron (1991a).

Key Words : Near–integrated model, functional weak convergence, simulation experiment, unit root process, nearly stationary model, nearly twice integrated model, nearly seasonally integrated model, Fredholm determinant, characteristic function.

Addresses:

Seiji Nabeya
Department of Economics
Tokyo International University
1–13–1, Matobakita
Kawagoe, Saitama, 350
Japan

Pierre Perron
Department of Economics
Princeton University
Princeton, NJ,
U.S.A., 08544

1. INTRODUCTION

In an attempt to cover more general time series structures, it has become popular in econometric methodology to consider models which permit that both the regressors and the errors have substantial heterogeneity and dependence over time. On a theoretical level, this advance has become possible due to a new class of central limit theorems (or functional central limit theorems) which provides asymptotic results allowing both substantial heterogeneity and dependence. This paper 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 :

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

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 given by :

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

Recently, a new class of models which specifically deal with the presence of a root close to, but not necessarily equal to one, has been studied. We consider a near–integrated process where the autoregressive parameter is defined by :

$$(1.3) \qquad \alpha = \exp(c/T).$$

Here, the constant c is a measure of the deviation from the unit root case. The model (1.1) and (1.3) may also be described as having a root local to unity : as the sample size increases, the autoregressive parameter converges to unity. When $c < 0$, the process $\{y_t\}$ is said to be (locally) stationary and when $c > 0$, it is said to be (locally) explosive. 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).

In this near–integrated context, with errors that are weakly dependent, Phillips (1987) showed that (under appropriate mixing conditions on the sequence $\{u_t\}$) :

$$(1.4) \qquad T(\hat{\alpha}-\alpha) \Rightarrow \left\{ \int_0^1 J_c(r)dW(r) + \lambda \right\} \left\{ \int_0^1 J_c(r)^2 dr \right\}^{-1},$$

where $\lambda = (\sigma^2 - \sigma_u^2)/(2\sigma^2)$, $\sigma^2 = \lim_{T\to\infty} E(T^{-1}S_T^2)$, $S_T = \Sigma_{j=1}^T u_j$, $\sigma_u^2 = \lim_{T\to\infty} T^{-1}E(\Sigma_{t=1}^T u_t^2)$, $J_c(r) = \int_0^r \exp((r-s)c)dW(r)$; and $W(r)$ is the unit Wiener process (or standard Brownian motion) on C[0,1], the space of real–valued continuous functions on the [0,1] interval. This type of asymptotic distribution provides a useful framework to analyze models with dependent errors.

Tabulations of the limiting distribution (1.4) with $\lambda = 0$ 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 (1991b,c) also considers a continuous–time approximation which performs well even in the case where the initial condition is non–zero. These asymptotic frameworks provide a substantial improvement 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 (1991a) presented an extensive simulation analysis to assess the adequacy of the limiting distribution (1.4) as an approximation to the finite sample distribution, concentrating on two leading cases, namely :

$$(1.5) \qquad \text{MA(1) errors :} \qquad u_t = e_t + \theta e_{t-1},$$

$$(1.6) \qquad \text{AR(1) errors :} \qquad u_t = \rho u_{t-1} + e_t,$$

where $\{e_t\}$ is a sequence of i.i.d. $N(0, \sigma_e^2)$ random variables. The results shown in Perron (1991a) 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 expressed in (1.4) ; 5) in the MA case the variance shows non–monotonic behavior as θ varies with a fixed sample size T.

This paper is devoted to providing an alternative asymptotic framework in each of the cases mentioned above where the usual asymptotic distribution fails to be a sensible guide to the finite sample distribution. The aim is twofold. First, to provide an asymptotic framework which is likely to provide a better approximation to the finite sample distribution. Secondly, our analysis will give theoretical explanations for the features of the finite sample distributions mentioned above.

In Section 2 we consider the limiting behavior of the normalized least–squares estimator allowing the MA parameter θ to approach -1 at a suitable rate. This provides an asymptotic framework which we label as "nearly white noise – nearly integrated process". We derive a characteristic function which allows the calculation of distributional quantities. The adequacy of this local framework is assessed.

Section 3 considers the case where ρ, the AR parameter, approaches 1. Our asymptotic analysis provides a limiting distribution for processes with nearly two unit roots. Section 4 considers the case where ρ approaches -1. Here the framework is shown to be related to a nearly integrated seasonal model of period 2. These asymptotic analyses help to understand the differing behavior of the normalized least squares estimator as ρ approaches plus or minus one. In each case, we derive the limiting distribution and an appropriate limiting characteristic function which allows tabulation of distributional quantities via numerical integration. The adequacy of these asymptotic distributions as approximations to the finite sample distributions is also assessed. Finally, Section 5 provides some concluding comments and an appendix contains mathematical derivations.

2. A NEARLY WHITE NOISE NEARLY INTEGRATED PROCESS

In this Section, we propose an alternative asymptotic framework that is intended, on the one hand, to provide an asymptotic distribution which better approximates the exact distribution of $T(\hat{\alpha} - \alpha)$ when the errors have an MA(1) structure with large negative correlation, i.e. when θ is close to -1. On the other hand, our intention, using this alternative approach, is also to explain some of the finite samples phenomena discussed in the introduction. Consider the following parameterization of the nearly integrated process with MA(1) errors :

(2.1) $$y_t = \exp(c/T)y_{t-1} + e_t + \theta_T e_{t-1} ,$$

where

(2.2) $$\theta_T = -1 + \delta/T^{1/2}.$$

For simplicity we assume that $e_t \sim$ i.i.d. $(0, \sigma_e^2)$. The process defined by (2.1) and (2.2) 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". A variant of this specification, with $c = 0.0$, has been considered by Pantula (1991) in a different context. Our aim, in this section, is to study the asymptotic distribution of $\hat{\alpha}$ under the specification (2.1) and (2.2). The next Theorem, proved in the Appendix, characterizes this asymptotic distribution.

THEOREM 1 : *Let $\{y_t\}$ be a sequence of random variables generated by (2.1) and (2.2) and assume that $y_0 = e_0 = 0$, then as $T \to \infty$:*

$$\hat{\alpha} \Rightarrow \left\{\delta^2 \int_0^1 J_c(r)^2 dr\right\} \left\{1 + \delta^2 \int_0^1 J_c(r)^2 dr\right\}^{-1} \equiv A(\delta, c);$$

where $J_c(r) = \int_0^r \exp((r-s)c)dW(s)$, and $W(s)$ is the unit Wiener process on $C[0,1]$.

Remark : i) The support of the limiting distribution of $\hat{\alpha}$ is the interval [0, 1], and since $\alpha \to 1$ as $T \to \infty$ the limiting distribution of $\hat{\alpha} - \alpha$ has negative support. Note also that the support of the limiting distribution of $\hat{\alpha} - \alpha$ is independent of c, though its distribution is not. ii) The condition $y_0 = e_0 = 0$ is not imposed for simplicity. The theorem is valid if $y_0 = e_0$ but otherwise the limiting distribution is not invariant to their values.

Theorem 1 shows that under this nearly white noise setting the asymptotic distribution of $\hat{\alpha}$ is degenerate unless $\delta = 0$, in the sense that $\hat{\alpha}$ converges to a random variable instead of fixed constant. Hence $\hat{\alpha}$ is not a consistent estimator of α. If $\delta = 0$, we have that $\hat{\alpha} \to 0$ in probability as expected ; and as $\delta \to \infty$ the limit of $\hat{\alpha}$ tends to 1.

This result helps to explain some of the findings in Perron (1991a). Note first that, under the present setting, $T(\hat{\alpha} - \alpha)$ is unbounded and converges to $-\infty$. Hence , on the one hand, we would expect the finite sample distribution of $T(\hat{\alpha} - \alpha)$ to shift leftward as θ decreases. On the other hand, we would also expect the usual asymptotic approximation to be inadequate for values of θ close to -1.

Theorem 1 presents an alternative distributional theory that could provide a more adequate approximation to the exact distribution of $T(\hat{\alpha} - \alpha)$ for values of T and θ where the usual asymptotic theory fails to provide a useful guide. To investigate this issue the next Theorem presents a limiting characteristic function that will allow computation of distributional quantities related to $\hat{\alpha}$.

THEOREM 2 : *Let $\{y_t\}$ be a stochastic process defined by (2.1) and (2.2) with $y_0 = e_0 = 0$. The limiting distribution of $\hat{\alpha}$, $\lim_{T\to\infty} P[\hat{\alpha} \leq x]$, is given by $P[Z_x^1 \geq 0]$ where Z_x^1 is a random variable with characteristic function $D_x^1(2iw)^{-1/2}$ where*

$$D_x^1(\lambda) = exp(c - \lambda x)(cosh(\mu) - c\ sinh(\mu)/\mu)$$

and $\mu = \{c^2 - \lambda\delta^2(x - 1)\}^{1/2}$.

Theorem 2 allows direct computation, by numerical integration, of the cumulative distribution and probability density functions as well as the moments of the asymptotic distribution. Denote the limiting distribution of $\hat{\alpha}$ by $F(x; c, \delta)$ and its associated density function by $G(x; c, \delta)$. These quantities can be numerically evaluated using Imhof's (1961) formulae :

$$F(x; c, \delta) = (1/2) + (1/\pi)\int_0^\infty (1/w)IM\{D_x^1(2iw)^{-1/2}\}dw\ , \tag{2.3}$$

$$G(x; c, \delta) = (1/\pi)\int_0^\infty (1/w)IM\{\partial D_x^1(2iw)^{-1/2}/\partial x\}dw\ , \tag{2.4}$$

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 $|D_x^1(2iw)^{-1/2}| < 10^{-10}$ holds.

Denote by $\Psi_1(w;\ x)$ the moment-generating function associated with the characteristic function $D_x^1(2iw)^{-1/2}$. The moments of the asymptotic distribution can be obtained using Mehta and Swamy's (1978) result, which in our case implies :

$$E[A(\delta,c)]^r = \Gamma(r)^{-1}\int_0^\infty v^{r-1}\left\{\partial^r\Psi_1(-u;\ v/u)/\partial u^r\right\}_{u=0} dv\ . \tag{2.5}$$

To get an idea of the type of distribution involved, Figure 1 graphs the mean and standard deviation of the limiting distribution of $\hat{\alpha}$ as a function of δ, for the three cases c = 0.0, – 5.0 and 2.0. As expected for δ close to 0 the mean is close to 0, and as δ increases the mean approaches 1. The standard deviation of the process is close to 0 when either δ is very small or very large. As δ moves away from 0, both the mean and standard deviation increase more rapidly with the parameter c. From these considerations, we would expect : a) the approximation of the finite sample distribution to worsen as either δ gets large or close to zero (due to the implied zero variance) , and b) the approximation of the mean to be more adequate for small values of c (due to a less rapid approach of the mean of the asymptotic distribution towards 1) ; and c) the approximation of the variance to be more adequate for large values of c (due to a less flat asymptotic function).

Figure 2 presents the limiting density function of $\hat{\alpha}$ for several pairs of values for c and δ. As can be infered from this figure the limiting density shifts to the right as δ increases for a given value of c. The same behavior occurs, though to a lesser extent, as c increases with a given values of δ. Interestingly, the density can be bimodal as is the case for c = 2.0 and $\delta = 1.0$.

To use the asymptotic distribution of Theorem 1 as an approximation to the exact distribution of $T(\hat{\alpha} - \alpha)$, we specify the correspondence $\delta = T^{1/2}(1 + \theta)$. From the comments above one would expect a better approximation for combinations of T and θ such that δ is neither too small nor too large. Table I presents the percentage points of the distribution of $T(\hat{\alpha} - \alpha)$ calculated using this nearly white noise – nearly integrated asymptotic distribution. The cases considered are c = 0.0 ; $\theta = -0.9$, -0.7 and -0.5, with T = 25, 50, 100, 500, 1000 (except for $\theta = -0.5$), and T = 5000 (for $\theta = -0.9$). The finite sample values from Perron (1991a) are presented in parentheses. These are based on 10,000 replications using the model (2.1) and (2.2) with i.i.d. N(0, 1) innovations e_t. For $\theta = -0.9$, the approximation is excellent with $T \geq 500$, especially in the left tail. When T = 100 the approximation is still respectable but deteriorates as T reaches 50 or 25, especially in the right tail. Nevertheless, in all cases the approximation is much better than the standard asymptotic distribution (see Perron (1991a)). When $\theta = -0.7$, the approximation is best when T = 25 or 50 and deteriorates as T gets larger. Again the left tail is much better approximated than the right tail. When $\theta = -0.5$, the extreme right tail of the distribution is badly approximated due to the implied negativity of the asymptotic distribution of $T(\hat{\alpha} - \alpha)$ provided by the nearly white noise local framework.

Table II presents the approximation to the mean and variance of $T(\hat{\alpha} - \alpha)$ provided by the nearly white noise asymptotic framework for the three values of c considered (c = 0.0, –5.0, 2.0). With c = 0.0, the approximation is excellent for all sample sizes when $\theta = -0.9$. With $\theta = -0.7$, the approximation is adequate for samples of size less than 500. When θ is –0.50 the approximation is not as adequate, though it is highly superior to the standard asymptotic approximation when $T \leq 500$. The same qualitative features hold when c = 2.0 or – 5.0 but with a better approximation for c = – 5.0 and less adequate for c = 2.0.

Consider now the variance of the distribution of $T(\hat{\alpha} - \alpha)$. When $\theta = -0.9$, the variance is badly approximated unless $T \geq 500$. When θ is – 0.7 the approximation is reasonable for T = 100 and 500. When θ is – 0.5 it is reasonable for smaller sample sizes. For the cases c = – 5.0 and 2.0, the results show the same qualitative features but now, interestingly, the approximation is better when c = 2.0 and worse when c = – 5.0 (unlike what was found for the mean of the distribution). A feature of particular interest is the behavior of the variance as θ approaches – 1 with a given sample size. As mentioned in the introduction, the exact results shows a non–monotonic behavior. This feature is well explained by this local asymptotic theory. Indeed, this non–monotonic behavior is present in several of the cases presented in Table II.B. The reason for this behavior can be obtained by looking at Figure 1 where it is shown that the standard deviation of the local asymptotic distribution of $\hat{\alpha}$ is zero when $\delta = 0$ and eventually approaches zero again as δ increases. Given that $\delta = T^{1/2}(1 + \theta)$, a decrease in θ for a fixed T implies that δ approaches 0. The non–monotonic behavior occurs when the change in θ is such as to move δ over the hump in the standard deviation function presented in Figure 1.

Figure 3 considers a more detailed simulation experiment for the case c = 0 and $\delta = 1$. In that figure, the limiting distribution of $\hat{\alpha}$ is compared to the finite sample distributions for T = 50, 100 and 200. Note that this case corresponds to one where the adequacy of the approximation is rather poor compared to larger values of δ. This feature transpires mainly through the fact that the exact distribution of $\hat{\alpha}$ has a support that is not bounded below by 0 unlike its asymptotic counterpart. Hence, the left tail of the distribution is poorly approximated. The approximation is better in the right tail and improves rapidly as T increases.

The results of our experiments show the nearly white noise – nearly integrated asymptotic distribution to be a far better approximation to the finite sample distribution of $T(\hat{\alpha} - \alpha)$ when θ is close to – 1 than is the standard asymptotic distribution. However, the approximation still lacks some accuracy in an important range of cases. First when θ is away from – 1 and T is large. This case, however, is not of much consequences since in this region the usual asymptotic theory is adequate. Of more consequence is the fact that the approximation is inadequate when T is small and θ is close to – 1 (i.e., when δ is close to 0). Here none of the asymptotic distributions available so far provide a satisfactory approximation.

3. A NEARLY TWICE INTEGRATED MODEL

In this Section, the aim is to provide a local asymptotic framework that could explain the behavior of the distribution of $T(\hat{\alpha} - \alpha)$ when the errors have an AR(1) structure with (large) positive correlation. Our intention is also to assess whether this alternative asymptotic distribution provides a better approximation to the finite sample distribution. We start with the following parameterization of the process of interest :

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

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

where, for simplicity, we specify $e_t \sim$ i.i.d. $(0,\sigma_e^2)$. We can write (3.1) and (3.2) as :

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

As T converges to infinity $\{y_t\}$ becomes :

$$y_t = 2y_{t-1} - y_{t-2} + e_t \, .$$

Therefore, as T increases, $\{y_t\}$ converges to a process with two unit roots, hence the expression "nearly twice integrated". Our aim is to study the asymptotic behavior of $T(\hat{\alpha} - \alpha)$ under this specification.

We first need to define some new notation. Consider the following transformation of the random process $J_\phi(r)$:

(3.4) $$Q_c(J_\phi(r)) \equiv \int_0^r \exp((r - v)c)J_\phi(v)dv$$

where, as before, $J_\phi(v) = \int_0^v \exp((v - s)\phi)dW(s)$. Hence, $Q_c(J_\phi(r))$ is a weighted integral of the process $J_\phi(v)$ where the weight function depends upon the parameter c. If $c = 0$, we have $Q_0(J_\phi(r)) = \int_0^r J_\phi(v)dv$ and if $c = \phi = 0$, $Q_0(J_0(r)) = \int_0^r W(v)dv$. Using this notation, we characterize the asymptotic distribution of $T(\hat{\alpha} - 1)$ in the next Theorem .

THEOREM 3 : *Let $\{y_t\}$ be a stochastic process generated by (3.1) and (3.2), and let the function $Q_c(J_\phi(r))$ be as defined in (3.4), then as $T \to \infty$:*

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

Remarks : i) The conditions specified by (3.1) and (3.2) are overly stringent for the validity of this result. In fact Theorem 3 still holds if we assume that the sequence of innovations $\{e_t\}$ is a linear process of the form $e_t = \Sigma_{j=0}^{\infty} \nu_j \eta_{t-j}$ $(t = 1,2, ...)$ where $\{\eta_t\}$ is a martingale difference process satisfying some conditions (see Nabeya and Tanaka (1990)) These conditions are satisfied if $\{\eta_t\}$ is i.i.d. $(0, \sigma^2)$ with finite σ^2. Indeed, the proof in the appendix requires only some weak mixing conditions on the error sequence $\{e_t\}$ such that the partial sum of the errors $S_t = \Sigma_{j=1}^{t} e_j$ satisfies a functional central limit theorem; ii) The limiting distribution of $T(\hat{\alpha} - 1)$ has nonnegative support. iii) Let $\alpha = \exp(c/T)$ be the autoregressive parameter in (3.1). Since $T(\hat{\alpha} - \alpha) = T(\hat{\alpha} - 1) - T(\alpha - 1)$, and $T(\alpha - 1) \to c$ as $T \to \infty$, the limiting distribution of $T(\hat{\alpha} - \alpha)$ is given by

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

The support of the limiting distribution of $T(\hat{\alpha} - \alpha)$ is therefore bounded below by $-$ c.

There are several interesting features to note about Theorem 3. First, neither c nor ϕ is restricted to be negative; these parameters can take any real value. Hence the result applies to many cases of interest. In particular it can encompass a stationary process (c, ϕ < 0), a process with an explosive and a stationary root (either c or ϕ is negative and the other is positive), a process with two explosive roots (c, $\phi > 0$); or a process with two unit roots ($c = \phi = 0$). The latter is of particular interest. In that special case we have :

(3.6) $$T(\hat{\alpha} - 1) \Rightarrow (1/2)\left\{\int_0^1 W(r)dr\right\}^2 \left\{\int_0^1 \left(\int_0^r W(s)ds\right)^2 dr\right\}^{-1}.$$

This result is interesting in view of the simulation experiment reported in Dickey and Pantula (1987). They showed that with a sample of length 50, the standard Dickey–Fuller (1979) test rejects the null hypothesis of a single unit root in favor of a stationary process slightly more than 5% of the time when the series actually has two unit roots. Given our result in (3.6) this feature is due to the small sample size used in the simulations. Indeed, in large samples, the Dickey–Fuller criterion would never reject a unit root in favor of a stationary process when two unit roots are present as the limiting distribution in (3.6) has a positive support.

It is worth emphasizing about Theorem 3 that, contrary to the case analyzed in the previous section with an MA(1) root local to – 1, $T(\hat{\alpha} - \alpha)$ has a non–degenerate asymptotic distribution. This explains the relatively small discrepancies between the usual asymptotic approximation and the exact distribution when ρ is close to one (as opposed to the large ones when the MA or AR roots are close to – 1). For ρ close to one, the fact that the exact distribution approaches its asymptotic counterpart quite slowly is explained by the difference between the local asymptotic distribution described above and the asymptotic distribution described by (1.4).

Theorem 3 can be used to compute, for a given pair of values of c and ϕ, an alternative approximation to the exact distribution of $T(\hat{\alpha} - \alpha)$ or $T(\hat{\alpha} - 1)$ when the errors have an autoregressive root close to one. As in the previous section, we derive a closed form solution for an appropriate limiting characteristic function in this nearly twice integrated setting. The result is stated in the next Theorem.

THEOREM 4 : *Let $\{y_t\}$ be a stochastic process generated by (3.1) and (3.2), the limiting distribution of $T(\hat{\alpha} - 1)$, $\lim_{T\to\infty} P[T(\hat{\alpha} - 1) \le x]$, is given by $P(Z_x^2 \ge 0)$ where Z_x^2 is a random variable with the characteristic function $D_x^2(2iw)^{-1/2}$ where :*

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

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

Remarks : i) Again, the result of Theorem 4 holds under conditions more general than those stated, see Remark (i) of Theorem 3. ii) The parameters c and ϕ can be any two real numbers or any two complex conjugate pairs. Hence (3.3) is, in this general context, well defined in the realm of real numbers though (3.1) and (3.2) are not necessarily so. iii) The result with $c = \phi = 0$ was also obtained by Tanaka (1990). iv) The limiting charateristic function described in Theorem 4 can be used to numerically evaluate the limiting distribution and density functions of $T(\hat{\alpha} - 1)$ using (2.3) and (2.4). The specifications used are similar to those in Section 2.

To get an idea of the type of limiting distributions involved, Figure 4 presents limiting density functions of $T(\hat{\alpha} - 1)$ for some values of c and ϕ. Every curve in Figure 4 starts from ∞ at $x = 0$ and tends to 0 as $x \to \infty$.

Selected percentage points of the asymptotic distribution of $T(\hat{\alpha} - \alpha)$ with $\alpha = \exp(c/T)$ are presented in Tables III along with finite sample values in parentheses reproduced from Perron (1991a). The limiting values are obtained from the distribution (3.5) and the limiting characteristic function of Theorem 4 with $x - c$ substituted for x. Consider first the case where $c = 0$ presented in Table III.A. Here the left tail of the distribution is not well approximated for any value of ρ. This is due to the fact that when $c = 0.0$, the local asymptotic distribution considered implies a non–negative variable in the limit. However, the left tail of the finite sample distribution is in the negative part. On the other hand, the right tail of the distribution is much better approximated by the local asymptotic distribution than by the usual asymptotic distribution for a ρ value of 0.95 and to some extent 0.9. It provides, however, no improvement when ρ is 0.5. These facts are corroborated by the behavior of the mean and variance of the distribution. Both are better approximated by the local asymptotic distribution when $\rho = 0.95$ and to some extent when $\rho = 0.90$, but not when $\rho = 0.50$.

Consider now the case where $c = -5.0$ presented in Table III.B. The results are quite different from the case where $c = 0.0$. Here the "nearly doubly integrated" local asymptotic distribution seems to provide a worse approximation, in the left tail, than the usual asymptotic distribution for all values of ρ considered (even though the variance is better approximated). To get a closer view of the approximation, Figure 5 shows the results of simulations for the case $c = -5.0$ and $\phi = -2.5$ using N(0,1) random numbers for the innovation process $\{e_t\}$. Note that this case corresponds to one where the adequacy of the asymptotic approximation is rather poor compared to other cases considered in Table III. This fact is revealed by a substantial distance between the curves of the limiting and finite samples distributions in the left tail. This can be partly explained by the fact that the support of the empirical distribution of $T(\hat{\alpha} - 1)$ extends below zero whereas the support of its limiting distribution is bounded below by zero (see the Remark (ii) to Theorem 3).

The case with $c = 2.0$ is presented in Table III.C. For $\rho = 0.50$ the approximation is worse than the usual asymptotic distribution (see Perron (1991a)). On the other hand, when $\rho = 0.90$ or 0.95, the improvement is substantial. The median and the right tail of the distribution are very well approximated. For example, when $\rho = 0.9$ and $T = 25$, the values for the 95% point are 2.13 (exact) and 2.135 (local asymptotic). Unlike the case with $c = 0.0$ or -5.0 the left tail of the distribution is not as badly approximated especially for the 10% point.

Figures 6 and 7 present the results of a simulation exercise concerning the robustness of the limiting results to additional correlation in the errors $\{e_t\}$ as discussed in Remark (i) of Theorem 3. The results pertain to a comparison of the finite sample and asymptotic distributions of the statistic $T(\hat{\alpha} - 1)$. Figure 6 considers the case $c = \phi = 0$ with MA(1) errors $e_t = \eta_t - (3/4)\eta_{t-1}$. Figure 7 is for the case $c = -1 - i$ and $\phi = -1 + i$ using AR(1) errors $e_t = -(2/3)e_{t-1} + \eta_t$. In both cases, N(0,1) random numbers are used to generate the sequence $\{\eta_t\}$. Both figures illustrate the validity of the asymptotic result, though the convergence of the finite sample distribution to its asymptotic counterpart is rather slow. This is due to the fact that the moving–average and autoregressive parameters are large negative numbers. The convergence is more rapid with an error sequence that is less correlated or positively correlated.

In summary, combining the results of Perron (1991a) and those discussed above, the different types of asymptotic distributions considered for the case with positively correlated AR(1) errors appear to be complementary. None of them provides an approximation to the finite sample distribution that is satisfactory for all values of ρ and all percentage points. However, for a wide range of parameter configurations there is a particular asymptotic framework that seems appropriate. When ρ is small, say less than 0.5, the usual O(1) asymptotic performs quite well. When ρ is close to 0.5, the $O_p(T^{-1})$ stochastic expansion provides a substantial improvement in the left tail of the distribution (unless c = 2.0) but not in the right tail. This feature is not too troublesome, given that there are much less variations in the right tail of the distribution as T changes. When ρ approaches 1, the $O_p(T^{-1})$ stochastic expansion fails to provide much of an improvement. On the other hand, the "nearly twice integrated" model proposed here seems to provide a marked improvement in approximating the percentage points in the right tail of the distribution when ρ is close to one (especially when c = 0.0 or c = 2.0). The region where none of the asymptotic frameworks discussed provide an adequate approximation for sample sizes in the range from 25 to 100 is in the left tail of the distribution when ρ is close to one.

4. A NEARLY INTEGRATED SEASONAL MODEL

In this Section, our aim is to provide a local asymptotic framework for the case where the errors have an autoregressive root near minus unity. We consider the adequacy of such an asymptotic approximation and we also investigate how the theoretical results can shed light on the differing behavior of $T(\hat{\alpha} - \alpha)$ when the errors are negatively correlated with AR(1) or MA(1) structures as discussed in the introduction. Consider first the following parameterization of the process under study :

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

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

where we again specify $e_t \sim$ i.i.d. $(0,\sigma_e^2)$ and for simplicity $e_0 = u_0 = 0$. The model (4.1) and (4.2) can be written as :

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

As T increases to infinity $\{y_t\}$ approaches the process :

(4.4) $$y_t = y_{t-2} + e_t .$$

The equation (4.4) characterizes a seasonal model of period 2 with a root on the unit circle. We therefore label the process (4.1) and (4.2) as a "nearly integrated seasonal model". To get some insights into the result presented below, consider a special case where $c = \phi$. Then (4.3) reduces to :

(4.5) $$y_t = \exp(2c/T)y_{t-2} + e_t .$$

This is a special case of a class of nearly integrated seasonal models that have recently been studied by Chan (1989) and Perron (1990b). Chan (1989) derives the asymptotic distribution of $T(\hat{\alpha}_d - \alpha)$ where $\hat{\alpha}_d$ is the least-squares estimator of the coefficient on y_{t-2} in equation (4.5). Perron (1990b) tabulates the percentage points of this asymptotic distribution. The difference in focus here is that we wish to study the asymptotic distribution of the *first*-order autocorrelation coefficient when the process is a nearly integrated seasonal model of period 2.

Recall that $\hat{\alpha} = T^{-2}\Sigma_1^T y_t y_{t-1}/T^{-2}\Sigma_1^T y_{t-1}^2$. Under (4.5), it is easy to deduce from Chan (1989,Lemma 2.i) that :

(4.6) $$T^{-2}\Sigma_1^T y_{t-1}^2 \Rightarrow (\sigma_e^2/4)\Sigma_{i=1}^2 \int_0^1 [J_{c,i}(r)]^2\, dr\,;$$

where $J_{c,i}(r) = \int_0^r \exp((r-s)c)dW_i(s)$, i= 1, 2 ; and $W_1(r)$ and $W_2(r)$ and independent Wiener processes. Consider now the numerator of $\hat{\alpha}$. First note that we can write :

(4.7) $$y_t = \Sigma_{j=0}^{[t/2]}\exp(2cj/T)e_{t-2j}\ ,$$

and

(4.8) $$y_{t-1} = \Sigma_{j=0}^{[(t-1)/2]}\exp(2cj/T)e_{t-2j-1}\ ,$$

where $[\cdot]$ denotes the integer part of the number. Given that the errors $\{e_t\}$ are i.i.d., y_t and y_{t-1} are independent processes as they are functions of different subsets of the sequence $\{e_t\}$. Hence $y_t y_{t-1}$ is the product of two independent nearly integrated random processes having $[t/2]$ and $[(t-1)/2]$ elements respectively (assuming $e_0 = 0$). Following the results on the sums of products of two independent random walks, it is straightforward to show that :

(4.9) $$T^{-2}\Sigma_1^T y_t y_{t-1} \Rightarrow (\sigma_e^2/2) \int_0^1 J_{c,1}(r)J_{c,2}(r)dr\,.$$

Hence we have the following asymptotic result when $c = \phi$:

(4.10) $$\hat{\alpha} \Rightarrow \left\{2\int_0^1 J_{c,1}(r)J_{c,2}(r)dr\right\}\left\{\Sigma_{i=1}^2 \int_0^1 [J_{c,i}(r)]^2\, dr\right\}^{-1}.$$

Note that $\hat{\alpha}$ has a degenerate asymptotic distribution , in the sense that it converges to a random variable instead of a fixed constant as was the case with MA(1) errors having a root approaching – 1. Our result is consistent with that of Yajima (1985) who showed, among other things, that in seasonally integrated models of period k, the sample autocorrelations of orders other than kn (for any integer n) have a degenerate asymptotic distributions in the sense that they converge to random variables instead of fixed constants.

Given (4.10), $T(\hat{\alpha} - \alpha)$ is unbounded as T increases, which explains the large discrepancies between the exact and asymptotic distributions reported in Perron (1991a). Note, however, the different rate at which the root is permitted to approach – 1 as T increases to infinity. In the AR(1) case it does so at rate T, while in the MA(1) case the rate is $T^{1/2}$. This feature explains well the differences in the discrepancies between the finite sample and asymptotic distributions reported in Perron (1991a). As shown, the discrepancies are much larger in the MA(1) case for an equal value of θ and ρ. Such a feature can be theoretically interpreted by noting that for a given same value for θ and ρ, ϕ is further away than δ from the zero boundary because of the different normalizing power on T. Indeed, in the MA case, we have $\delta = T^{1/2}(1 + \theta)$ and, in the AR case, $\phi = T(1 + \rho)$. In this sense, for a given value of θ and ρ close to – 1 and a given T, the process (2.1) and (2.2) is closer to a white noise process than the process (4.1) and (4.2) is to a seasonal random walk. One can therefore expect the standard asymptotic distribution to provide a less accurate approximation in the former case.

Note also that (4.10) does not presume that c is negative; it can also accommodate locally explosive processes as well as a seasonal random walk. In the latter case we have an interesting result, namely the asymptotic behavior of the first–order autocorrelation coefficient when the true model is a seasonal random walk of period 2. In this case :

$$(4.11) \qquad \hat{\alpha} \Rightarrow \left\{2\int_0^1 W_1(r)W_2(r)dr\right\}\left\{\Sigma_{i=1}^2 \int_0^1 W_i(r)^2dr\right\}^{-1},$$

where $W_1(r)$ and $W_2(r)$ are independent Wiener processes. The general case where $c \neq \phi$ is more complex but yields qualitatively similar results. The following Theorem, proved in the Appendix, provides the formal asymptotic distribution.

THEOREM 5 : *Let $\{y_t\}$ be a stochastic process generated by (4.1) and (4.2). Define the random functions $J_{\phi,i}(s) = \int_0^s exp((s - v)\phi)dW_i(v)$ and $Q_c(J_{\phi,i}(r)) = \int_0^r exp((r - s)c)$ $J_{\phi,i}(s)ds$ $(i = 1,2)$; where $W_1(r)$ and $W_2(r)$ are independent Wiener processes. Also let $J_{c,1}(s) = \int_0^s exp((s - v)c)dW_1(v)$. Then as $T \to \infty$:*

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

where $A(r) = (\phi - c)[Q_c(J_{\phi,1}(r)) - Q_c(J_{\phi,2}(r))] + 2J_{c,1}(r)$,

and $B(r) = J_{\phi,1}(r) - J_{\phi,2}(r)$.

Remark : The limiting distribution of $\hat{\alpha}$ is of the form $(Z_1 - Z_2)/(Z_1 + Z_2)$ where Z_1 ($\equiv \int_0^1 [A(r) - B(r)]^2 dr$) and Z_2 ($\equiv \int_0^1 B(r)^2 dr$) are nonnegative random variables. Hence, the support of the limiting distribution is restricted to the interval [–1, 1].

Note that the result in Theorem 5 reduces to (4.10) when $c = \phi$. Note also that, again in this general context, c and ϕ can take any real value. This result shows that $\hat{\alpha}$ has a degenerate asymptotic distribution even in the general case where $c \neq \phi$. Hence $\hat{\alpha}$ is not a consistent estimator and $T(\hat{\alpha} - \alpha)$ is unbounded as T increases. However unlike the MA(1) case and similar to the AR(1) case with a positive root on the unit circle, the asymptotic distribution in Theorem 5 has a non–zero variance even on the boundary $\phi = 0$. The next theorem characterizes the appropriate limiting characteristic function that can be used to numerically evaluate the limiting distribution of $\hat{\alpha}$.

THEOREM 6 : *Let $\{y_t\}$ be a stochastic process generated by (4.1) and (4.2). The limiting distribution of the first–order least–squares estimator $\hat{\alpha}$, $\lim_{T\to\infty} P[\hat{\alpha} \leq x]$, 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 $D_x^3(\lambda)$ is defined by :*

$$\begin{aligned} D_x^3(\lambda) = \exp(c + \phi) \Big\{ & [\cosh(\{c^2 - \lambda(x - 1)/4\}^{1/2}) \\ & - c \sinh(\{c^2 - \lambda(x - 1)/4\}^{1/2})/(c^2 - \lambda(x - 1)/4)^{1/2}] \\ & [\cosh(\{\phi^2 - \lambda(x + 1)/4\}^{1/2}) \\ & - \phi \sinh(\{\phi^2 - \lambda(x + 1)/4\}^{1/2})/(\phi^2 - \lambda(x + 1)/4)^{1/2}] \Big\}. \end{aligned}$$

Remark : The limiting characteristic function described in Theorem 6 can be used to numerically evaluate the limiting distribution and density functions of $\hat{\alpha}$ using (2.3) and (2.4). The specifications used are similar to those in Section 2.

To get an idea of the type of limiting distributions involved, Figure 8 presents the limiting density functions for some values of the parameters c and ϕ. Note that as c increases the limiting density function exhibits a high peak near 1 though the maximum height is finite.

Table IV presents the distribution of $T(\hat{\alpha} - \alpha)$ based upon the local asymptotic framework described in Theorem 5 using the joint characteristic function of Theorem 6 for $\rho = -0.9$. The cases considered are again c = 0.0, – 5.0 and 2.0 with T = 25, 50, 100 and 500. The finite sample results, reproduced in parentheses from Perron (1991a), are based on normally distributed innovations. In general, the approximation is satisfactory and certainly represent a major improvement over the standard asymptotic distribution (see Perron (1991a)). The approximation is best, and indeed very good, when c = – 5.0 (most notably with T = 50 and 100). It deteriorates as c increases. Also, for a fixed value of c, the approximation is better when T is small; it deteriorates as T increases to 500. This last feature is to be expected given that our asymptotic framework is local to the boundary $\rho = -1$; when T increases the noncentrality parameter ϕ is correspondingly higher. Finally, it is to be noted that the approximation is better in the left tail of the distribution.

Figure 9 presents an extended simulation for the case c = 2.0 and $\phi = -2.5$ whose asymptotic density is graphed in Figure 8. When c is positive, the finite sample distribution of $\hat{\alpha}$ has a relatively heavy weight in the region x > 1, in spite of the fact that the support of the limiting distribution of $\hat{\alpha}$ is restricted to the interval [–1, 1] (see the remark to Theorem 5). It is seen that the convergence of the finite sample distribution to its asymptotic counterpart is accordingly slow in the right tail of the distribution (near 1) whereas it is fast in the left tail. This reflects the fact mentioned above concerning the high peak of the limiting density near one with the fact that the support is bounded above by one. For other combinations of c and ϕ considered in Figure 8, the approach of the finite sample distribution to its asymptotic counterpart is faster.

5. CONCLUDING COMMENTS

This study presented alternative frameworks that could provide better approximations to the finite sample distributions of the statistic of interest. Our results are encouraging in that our local asymptotic distributions provide substantial improvements in approximating the finite sample distributions in the region of the parameter space where the traditional asymptotic framework provides severe inaccuracies. These local asymptotic distributions still depend upon nuisance parameters, namely those indexing the extent of correlation in the residuals. In practice one would need to have an estimate of these parameters in order to use our distributional results. These could be obtained by a preliminary investigation of the nature of the correlation structure of the residuals. Consider, for example, the case of testing for a unit root. A preliminary estimate of the correlation structure under the null hypothesis can be obtained by analyzing the sample correlation of the first–difference of the data. Suppose, for illustration, that a large negative MA(1) component is estimated. The test can then be carried using the local asymptotic distribution described in Section 2 with δ chosen according to the estimated value of the MA parameter θ. Of course, similar procedures can be followed in the case where the residuals have an autoregressive structure.

On a theoretical side, our study along with that of Perron (1991a) show how different asymptotic frameworks can be complementary in several respects. First, each framework provides a better approximation to the finite sample distribution where the other shows great inaccuracies. Secondly, the asymptotic results in the local asymptotic frameworks were shown to be useful in explaining why and when the usual asymptotic theory may fail. Nevertheless, our results also show the need for a unified asymptotic theory that could provide a sensible guide to the finite sample distribution over most of the relevant parameter space, albeit with possibly the need to estimate nuisance parameters. Such a topic is of interest for future research.

REFERENCES

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

Cavanagh, C. (1986): "Roots Local to Unity," manuscript, 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. (1989): "On the Nearly Nonstationary Seasonal Time Series," *The Canadian Journal of Statistics*, 17, 279–284.

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

Dickey, D.A., and W.A. Fuller (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

Dickey, D.A., and S.G. Pantula (1987): "Determining the Order of Differencing in Autoregressive Processes," *Journal of Business and Economic Statistics*, 5, 455–461.

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

Metha, J.S., and P.A.V.B. Swamy (1978): "The Existence of Moments of Some Simple Bayes Estimators of Coefficients in a Simultaneous Equation Model," *Journal of Econometrics*, 7, 1–14.

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. (1991): "Characteristic Functions of the Limiting Distributions Related to the AR(1) Estimator in the AR(2) Model," manuscript, Tokyo International University.

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. (1990a): "The Adequacy of Limiting Distributions in the AR(1) Model with Dependent Errors," Econometric Research Program Memorandum No. 349, Princeton University.

Perron, P. (1990b): "The Limiting Distribution of the Least Squares Estimator in Nearly Integrated Seasonal Models," Econometric Research Program Memorandum No. 354, Princeton University.

Perron, P. (1991a): "The Adequacy of Asymptotic Approximations in the Near–integrated Model with Dependent Errors," manuscript, Princeton University.

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

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

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

Phillips, P.C.B., and P. Perron (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335–346.

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

Yajima, Y. (1985): "Asymptotic Properties of the Sample Autocorrelations and Partial Autocorrelations of a Multiplicative Process," *Journal of Time Series Analysis*, 6, 187–201.

MATHEMATICAL APPENDIX

Proof of Theorem 1: Assuming that $y_0 = e_0 = 0$, we can write $y_t = \Sigma_{j=1}^{t}\exp((t-j)c/T)u_j$. Given that $u_t = e_t - e_{t-1} + \delta T^{-1/2}e_{t-1}$, simple manipulations show that :

$$\text{(A.1)} \qquad y_t = a_T e_t + b_T X_t \, ,$$

where

$$\text{(A.2)} \qquad a_T = (1 - \delta T^{-1/2})\exp(-c/T) \, ,$$

$$\text{(A.3)} \qquad b_T = 1 - \exp(-c/T)(1 - \delta T^{-1/2}) \, ,$$

and $X_t = \Sigma_{j=1}^{t}\exp((t-j)c/T)e_j$ is a near–integrated process given by $X_t = \exp(c/T)X_{t-1} + e_t$, with $e_t \sim$ i.i.d. $(0, \sigma_e^2)$ and $X_0 = 0$. Note that $a_T \to 1$ and $T^{1/2}b_T \to \delta$ as $T \to \infty$. Consider first the second sample moment of y_t. We have :

$$T^{-1}\Sigma_1^T y_t^2 = T^{-1}\Sigma_1^T (a_T e_t + b_T X_t)^2$$

$$= a_T^2 T^{-1}\Sigma_1^T e_t^2 + T b_T^2 T^{-2}\Sigma_1^T X_t^2 + 2a_T T^{1/2} b_T T^{-3/2}\Sigma_1^T X_t e_t \, .$$

Note that $T^{-1}\Sigma_1^T e_t^2 \to \sigma_e^2$ (in probability) and $T^{-2}\Sigma_1^T X_t^2 \Rightarrow \sigma_e^2 \int_0^1 J_c(r)^2 dr$ as $T \to \infty$. Furthermore, in a manner similar to Theorem 2.4 of Chan and Wei (1987), it can be shown that $\Sigma_1^T X_t e_t = O_p(T)$. Hence :

$$\text{(A.4)} \qquad T^{-1}\Sigma_1^T y_t^2 \Rightarrow \sigma_e^2 + \sigma_e^2 \, \delta^2 \int_0^1 J_c(r)^2 dr \, .$$

Consider now the sum $T^{-1}\Sigma_1^T y_{t-1}u_t$. Using (2.1) and (2.2) we can write :

$$T^{-1}\Sigma_1^T y_{t-1}u_t = a_T T^{-1}\Sigma_1^T e_t e_{t-1} - a_T(1 - \delta T^{-1/2})\, T^{-1}\Sigma_1^T e_{t-1}^2$$

$$+ b_T T^{-1}\Sigma_1^T X_{t-1}e_t - (1 - \delta T^{-1/2})b_T T^{-1}\Sigma_1^T X_{t-1}e_{t-1},$$

where a_T and b_T are defined in (A.2) and (A.3). We can show that $\Sigma_1^T e_t e_{t-1}$ is $O_p(T^{1/2})$, $\Sigma_1^T X_{t-1}e_t$ and $\Sigma_1^T X_{t-1}e_{t-1}$ are $O_p(T)$ and that $T^{-1}\Sigma_1^T e_t^2 \to \sigma_e^2$ (in probability) as $T \to \infty$. Using these results and the fact that $a_T \to 1$ and $b_T \to 0$ as $T \to \infty$, we obtain :

$$T^{-1}\Sigma_1^T y_{t-1}u_t \to -\sigma_e^2 . \tag{A.5}$$

This proves Theorem 1 using $\hat{\alpha} - \alpha = T^{-1}\Sigma_1^T y_{t-1}u_t / T^{-1}\Sigma_1^T y_{t-1}^2$ with (A.4), (A.5) and the fact that $\alpha = \exp(c/T) \to 1$ as $T \to \infty$. □

Proof of Theorem 2: From Theorem 1, $\lim_{T\to\infty} P[\hat{\alpha} \le x] = P[A(\delta,c) \le x]$. Note that $A(\delta,c) \le x$ is equivalent to $Z_x^1 \equiv x + \delta^2(x-1)Y \ge 0$ where $Y \equiv \int_0^1 J_c(r)^2 dr$. The moment–generating function of Y can be obtained as a special case of Phillips (1987, eq. (A1) corrected for a misprint) (see also Perron (1989, p. 244) and Nabeya and Tanaka (1990, Theorem 4)). It is given by $MG_Y(v) = \exp(-c/2)(\cosh(\mu) - c\sinh(\mu)/\mu)^{-1/2}$ where $\mu = (c^2 - 2v)^{1/2}$. The characteristic function of Z_x^1 is obtained upon substituting $v = (x-1)\delta^2 wi$ and multiplying $MG_Y(v)$ by $\exp(iwx)$.

Proof of Theorem 3 : We prove Theorem 3 under general "mixing" conditions on the sequence $\{e_t\}$. An explicit statement of these conditions can be found in, e.g., Phillips and Perron (1988). Suffice it to say that they are sufficient to ensure the application of a functional central limit theorem, i.e. they are such that $T^{-1/2}S_{[Tr]} \Rightarrow \sigma W(r)$ where $\sigma^2 = \lim_{T\to\infty} E[T^{-1}S_T^2]$ with $S_t = \Sigma_{j=1}^t e_t$. We start with the following Lemma concerning the sample moments of $\{y_t\}$ under the nearly twice integrated framework of Section 3.

LEMMA A.1 : *Suppose that $\{y_t\}$ is a sequence of random variables generated according to (3.1) and (3.2), let $\sigma^2 = \lim_{T\to\infty} E[T^{-1}S_T^2]$ and $Q_c(J_\phi(r))$ be as defined in (3.4). Then as $T \to \infty$:*

a) $T^{-3/2}y_{[Tr]} \Rightarrow \sigma\, Q_c(J_\phi(r))$;

b) $T^{-4}\Sigma_1^T y_t^2 \Rightarrow \sigma^2 \int_0^1 Q_c(J_\phi(r))^2 dr$;

c) $T^{-3}\Sigma_1^T y_{t-1}u_t \Rightarrow (\sigma^2/2)\left\{Q_c(J_\phi(1))^2 - 2c\int_0^1 Q_c(J_\phi(r))^2 dr\right\}$.

Proof : We first define the random process $X_T(r)$ as :

$$X_T(r) = \sigma^{-1}T^{-1/2}S_{[Tr]} = \sigma^{-1}T^{-1/2}S_{j-1} \qquad (j-1)/T \le r < j/T$$

$$(j = 1,\ldots,T)$$

$$X_T(1) = \sigma^{-1}T^{-1/2}S_T \;;$$

where $S_j = \Sigma_{t=1}^j e_t$. Given the assumed conditions on the sequence $\{e_t\}$ we have $X_T(r) \Rightarrow W(r)$, the unit Wiener process. To prove part (a), we assume, for simplicity, that $y_0 = e_0 = 0$, and using (3.1) and (3.2) we have $(1 - \exp(\phi/T)L)(1 - \exp(c/T)L)y_t = e_t$, where L is the lag operator. We can therefore write y_t as :

$$y_t = \Sigma_{k=0}^t \exp(c(t-k)/T)\Sigma_{j=0}^k \exp(\phi(k-j)/T)\, e_j \,.$$

Then :

$$\begin{aligned}
T^{-3/2}y_{[Tr]} &= T^{-3/2}\Sigma_{k=0}^{[Tr]}\exp(c([Tr]-k)/T)\Sigma_{j=0}^k \exp(\phi(k-j)/T)\, e_j \\
&= \Sigma_{k=0}^{[Tr]}\int_{(k-1)/T}^{k/T}\exp(c([Tr]-k)/T)\left\{\sigma\Sigma_{j=0}^k \int_{(j-1)/T}^{j/T}\exp(\phi(k-j)/T)dX_T(s)\right\}dv \\
&= \sigma\int_0^r \exp(c(r-v))\int_0^v \exp(\phi(v-s))dX_T(s)dv \\
&= \sigma\int_0^r \exp(c(r-v))\left\{X_T(v) + \phi\int_0^v \exp(\phi(v-s))X_T(s)ds\right\}dv \\
&\Rightarrow \sigma\int_0^r \exp(c(r-v))\left\{W(v) + \phi\int_0^v \exp(\phi(v-s))W(s)ds\right\}dv \\
&\equiv \sigma\int_0^r \exp(c(r-v))J_\phi(v)dv \equiv Q_c(J_\phi(r)).
\end{aligned}$$

This proves part (a). To prove part (b), we have :

$$T^{-4}\Sigma_1^T y_t^2 = T^{-4}\Sigma_{t=1}^T\left\{\Sigma_{k=0}^t \exp(c(t-k)/T)\Sigma_{j=0}^k \exp(\phi(k-j)/T)\, e_j\right\}^2$$
$$= T^{-1}\Sigma_{t=1}^T\left\{\sigma\int_0^{t/T}\exp(c(t/T-v))\int_0^v \exp(\phi(v-s))dX_T(s)dv\right\}^2$$
$$= \sigma^2\int_0^1\left\{\int_0^r \exp(c(r-v))\int_0^v \exp(\phi(v-s))dX_T(s)dv\right\}^2 dr$$
$$\Rightarrow \sigma^2\int_0^1 Q_c(J_\phi(r))^2 dr\ ;$$

using arguments similar to those of part (a). To prove part (c), note that squaring (3.1), summing over t and rearranging, we have :

(A.6) $$T^{-3}\Sigma_1^T y_{t-1}u_t =$$
$$(1/2)\exp(-2c/T)\left\{T^{-3}y_T^2 - T(\exp(2c/T)-1)T^{-4}\Sigma_1^T y_{t-1}^2 - T^{-3}\Sigma_1^T u_t^2\right\}.$$

Note that, from parts (a) and (b), $T^{-3}y_T^2 \Rightarrow \sigma^2 Q_c(J_\phi(1))^2$ and $T^{-4}\Sigma_1^T y_{t-1}^2 \Rightarrow \sigma^2\int_0^1 Q_c(J_\phi(r))^2 dr$. We also have $T^{-2}\Sigma_1^T u_t^2 \Rightarrow \sigma^2\int_0^1 J_\phi(r)^2 dr$ given that $\{u_t\}$ is a nearly integrated process with non-centrality parameter ϕ, hence $T^{-3}\Sigma_1^T u_t^2 \to 0$ (in probability). Taking the limit of (A.6) and noting that $T(\exp(2c/T)-1) \to 2c$, we have :

$$T^{-3}\Sigma_1^T y_{t-1}u_t \Rightarrow (\sigma^2/2)\left\{Q_c(J_\phi(1))^2 - 2c\int_0^1 Q_c(J_\phi(r))^2 dr\right\}.$$

To prove Theorem 3, simply note that $T(\hat{\alpha}-\alpha) = T^{-3}\Sigma_1^T y_{t-1}u_t / T^{-4}\Sigma_1^T y_{t-1}^2$, hence :

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

Since $T(\hat{\alpha}-1) = T(\hat{\alpha}-\alpha) - T(1-\alpha)$ and $T(1-\alpha) \to -c$, we have :

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

Proof of Theorem 4 : We first deal with the case

(A.7) $$c \neq 0\,,\ \phi \neq 0 \text{ and } c^2 \neq \phi^2\,.$$

Note that the inequality $T(\hat{\alpha} - 1) \leq x$ is equivalent to $xV_T - U_T \geq 0$, where (since $y_{-1} = y_0 = 0$) :

(A.8) $$U_T = 2T^{-3}(\Sigma_{t=2}^T y_t y_{t-1} - \Sigma_{t=2}^T y_{t-1}^2)\,,$$

(A.9) $$V_T = 2T^{-4}\Sigma_{t=2}^T y_{t-1}^2\,.$$

Defining $\alpha_1 = \exp(c/T)$ and $\alpha_2 = \exp(\phi/T)$ we have, using (3.3) and $y_{-1} = y_0 = 0$:

$$y_t = \Sigma_{j=1}^t (\alpha_1 - \alpha_2)^{-1}({\alpha_1}^{t-j+1} - {\alpha_2}^{t-j+1})e_j \qquad (t = 1, 2, \ldots).$$

Substituting in (A.8) and (A.9) we obtain :

(A.10) $$\begin{aligned} U_T = T^{-1}\Sigma_{j=1}^T\Sigma_{k=1}^T T^{-2}(\alpha_1 - \alpha_2)^{-2}\Big\{ &- 2(\alpha_1^{|j-k|} - \alpha_1^{2T-j-k+2})/(1 + \alpha_1) \\ &- 2(\alpha_2^{|j-k|} - \alpha_2^{2T-j-k+2})/(1 + \alpha_2) \\ &+ (2 - \alpha_1 - \alpha_2)(\alpha_1^{|j-k|} + \alpha_2^{|j-k|} - \alpha_1^{T-j+1}\alpha_2^{T-k+1} \\ &- \alpha_1^{T-k+1}\alpha_2^{T-j+1})/(1 - \alpha_1\alpha_2)\Big\} e_j e_k \end{aligned}$$

$$= T^{-1}\Sigma_{j=1}^T\Sigma_{k=1}^T B_N(j,k)e_j e_k \qquad \text{(say)},$$

(A.11) $$\begin{aligned} V_T = T^{-1}\Sigma_{j=1}^T\Sigma_{k=1}^T 2T^{-3}(\alpha_1 - \alpha_2)^{-2}\Big\{ &(\alpha_1^{|j-k|} - \alpha_1^{2T-j-k+2})/(1 - \alpha_1^2) \\ &+ (\alpha_2^{|j-k|} - \alpha_2^{2T-j-k+2})/(1 - \alpha_2^2) \\ - (\alpha_1^{|j-k|} + \alpha_2^{|j-k|} - \alpha_1^{T-j+1}\alpha_2^{T-k+1} &- \alpha_1^{T-k+1}\alpha_2^{T-j+1})/(1 - \alpha_1\alpha_2)\Big\} e_j e_k \end{aligned}$$

$$= T^{-1}\Sigma_{j=1}^T\Sigma_{k=1}^T B_D(j,k)e_j e_k \qquad \text{(say)}.$$

The coefficients $B_N(j,k)$ and $B_D(j,k)$ ($j, k = 1, ..., T$) in the summations (A.10) and (A.11) can be approximated uniformly by the continuous functions $K_N(s,t)$ and $K_D(s,t)$, respectively, in the sense that $\lim_{T\to\infty} \max_{j,k} |B_N(j,k) - K_N(j/T,k/T)| = 0$, and similarly for $K_D(s,t)$, where

$$K_N(s,t) = [\exp(c(1-s)) - \exp(\phi(1-s))][\exp(c(1-t)) - \exp(\phi(1-t))]/(c-\phi)^2$$

$$\begin{aligned} K_D(s,t) = \Big\{ & [\exp(c(2-s-t)) - \exp(c|s-t|)]/c + [\exp(\phi(2-s-t)) - \exp(\phi|s-t|)]/\phi \\ & + 2[\exp(c|s-t|) + \exp(\phi|s-t|) - \exp(c(1-s)+\phi(1-t)) \\ & - \exp(c(1-t)+\phi(1-s))]/(c+\phi)\Big\}/(c-\phi)^2. \end{aligned}$$

We note that $K_N(s,t)$ and $K_D(s,t)$ are real-valued symmetric functions defined on the interval $[0,1]\mathrm{x}[0,1]$, and that both are positive definite in the sense that $\int_0^1\int_0^1 K_N(s,t)f(s)f(t)dsdt \geq 0$ for any real-valued continuous function $f(t)$ on the interval $[0,1]$, and similarly for $K_D(s,t)$.

Applying results in Nabeya and Tanaka (1988, 1990), the characteristic function of the limiting distribution of $xV_T - U_T$ as $T \to \infty$ is given by $\{D_x^2(2iw)\}^{-1/2}$ where $D_x^2(\lambda)$ is the Fredholm determinant for the kernel $xK_D(s,t) - K_N(s,t)$. To find $D_x^2(\lambda)$ consider the integral equation

(A.12) $$f(t) = \lambda\int_0^1 \{xK_D(s,t) - K_N(s,t)\}f(s)ds .$$

It is readily seen that $f(t)$ satisfies the following fourth order differential equation :

(A.13) $$f^{(4)}(t) - (c^2+\phi^2)f''(t) + (c^2\phi^2 - 2\lambda x)f(t) = 0 .$$

The characteristic equation associated with (A.13) is :

$$\mu^4 - (c^2+\phi^2)\mu^2 + (c^2\phi^2 - 2\lambda x) = 0,$$

which has the solutions $\pm\ \mu_1$ and $\pm\ \mu_2$ as defined in Theorem 4. If $\pm\ \mu_1$ and $\pm\ \mu_2$ are four different numbers, the general solution to (A.13) is given by :

$$\text{(A.14)} \qquad f(t) = b_1\exp(\mu_1 t) + b_2\exp(-\mu_1 t) + b_3\exp(\mu_2 t) + b_4\exp(-\mu_2 t)\ .$$

Using the right hand side of (A.14) for (A.12), the terms involving $\exp(\mu_1 t)$, $\exp(-\mu_1 t)$, $\exp(\mu_2 t)$ and $\exp(-\mu_2 t)$ cancel out. Equating the coefficients of $\exp(ct)$, $\exp(c(1-t))$, $\exp(\phi t)$ and $\exp(\phi(1-t))$ on both sides of (A.12), we obtain a system of linear homogeneous equations in b_1 to b_4. From the determinant of the coefficient matrix of the equations, we obtain $D_x^2(\lambda)$ as specified in Theorem 4. Note that the result holds if any of the conditions in (A.7) is violated since $D_x^2(\lambda)$ is continuous with respect to c and ϕ.

Remarks : 1) Note that $xK_D(s,t) - K_N(s,t)$ is negative definite if $x \le 0$. This implies that the support of the limiting distribution of $T(\hat{\alpha} - 1)$ is limited to the interval $[0, \infty)$. Hence, the limiting distribution of $T(\hat{\alpha} - \alpha)$ is restricted to the interval $[-c, \infty)$. 2) The result of Theorem 4 was first given in Nabeya (1987) as a conjecture. At that time the author could not verify the condition that, for every eigenvalue λ, the multiplicity for the integral equation (A.12) and the order of zero to $D_x(\lambda)$ should be the same. This condition has now been verified by assessing the validity of the equations

$$\int_0^1 \{xK_D(s,s) - K_N(s,s)\}ds = -D_x'(0),$$

$$\int_0^1\int_0^1 \{xK_D(s,t) - K_N(s,t)\}^2 ds dt = \{D_x'(0)\}^2 - D_x''(0),$$

using the computerized algebra REDUCE. The details of the proof of other conditions are described in Nabeya (1987).

Proof of Theorem 5 : To prove Theorem 5, we proceed with a series of Lemmas concerning various sample moments of the data. For ease of notation, assume without loss of generality, that the sample size T is an even number and let $m = T/2$. Also let $\alpha = \exp(c/T)$ in (4.1) and $\rho = \exp(\phi/T)$ in (4.2). The first Lemma is concerned with the asymptotic distribution of sample moments involving different subsets of the data, i.e.

separating the sequence $\{y_t\}$ and $\{u_t\}$ into two subsets corresponding to whether the time index t is even or odd.

LEMMA A.2 : *Let the functions $A(r)$ and $B(r)$ be as defined in Theorem 5 and consider a sequence of random variables $\{y_t\}$ defined by (4.1) and (4.2). Then as $T \to \infty$:*

a) For $[Tr]$ an even number : $T^{-1/2} y_{[Tr]} \Rightarrow 2^{-3/2} A(r)$;

b) $T^{-2} \Sigma_{k=1}^{m} u_{2k-1}^2 \Rightarrow (\sigma_e^2/4) \int_0^1 B(r)^2 dr$;

c) $T^{-2} \Sigma_{t=1}^{T} u_t^2 \Rightarrow (\sigma_e^2/2) \int_0^1 B(r)^2 dr$;

d) $T^{-2} \Sigma_{k=1}^{m} y_{2k-2}^2 \Rightarrow (\sigma_e^2/16) \int_0^1 A(r)^2 dr$;

e) $T^{-2} \Sigma_{k=1}^{m} y_{2k-2} u_{2k-1} \Rightarrow -(\sigma_e^2/8) \int_0^1 A(r) B(r) dr$.

Proof : To prove part (a), note that from (4.1) : $y_t = \Sigma_{j=1}^{t} \alpha^{t-j} u_j$. Hence, for t an even number, we have $y_{2k} = \Sigma_{j=1}^{2k} \alpha^{2k-j} u_j$ $(k = 1,\dots,m)$. Separating the sequence $\{u_j\}$ according to whether j is even or odd we have :

$$y_{2k} = \Sigma_{j=1}^{k} \alpha^{2k-2j} u_{2j} + \Sigma_{j=1}^{k} \alpha^{2k-2j+1} u_{2j-1} \, . \tag{A.15}$$

Now define the following variables :

$$X_{1,k} = \Sigma_{j=1}^{k} (\rho^2)^{k-j} e_{2j} \, , \tag{A.16}$$

$$X_{2,k} = \Sigma_{j=1}^{k} (\rho^2)^{k-j} e_{2j-1} . \tag{A.17}$$

Note that $X_{1,k}$ and $X_{2,k}$ are independent nearly integrated random processes with noncentrality parameter ϕ given that $\rho^2 = \exp(2\phi/T) = \exp(\phi/m)$ and that the random sequences $\{e_{2j}\}_{j=1}^{m}$ and $\{e_{2j-1}\}_{j=1}^{m}$ are independent by assumption (since the innovation sequence $\{e_t\}_{t=1}^{T}$ is i.i.d.). It is straightforward to show that :

(A.18) $$u_{2k} = X_{1,k} - \rho X_{2,k} ,$$

(A.19) $$u_{2k-1} = X_{2,k} - (1/\rho)X_{1,k} + (1/\rho)e_{2k} .$$

Using (A.15) through (A.19) we deduce that :

$$T^{-1/2}y_{2k} = T(1-\alpha/\rho)T^{-3/2}\Sigma_{j=1}^{k}(\alpha^2)^{k-j}X_{1,j}$$

(A.20) $$+ T(\alpha-\rho)T^{-3/2}\Sigma_{j=1}^{k}(\alpha^2)^{k-j}X_{2,j} + (\alpha/\rho)T^{-1/2}\Sigma_{j=1}^{k}(\alpha^2)^{k-j}e_{2j}.$$

Noting that $\alpha^2 = \exp(2c/T) = \exp(c/m)$ and $\rho^2 = \exp(2\phi/T) = \exp(\phi/m)$, using standard limiting arguments, we have (see the proof of Lemma A.1 (a)) :

$$m^{-3/2}\Sigma_{j=1}^{k}(\alpha^2)^{k-j}X_{1,j} \Rightarrow \sigma_e\int_0^r \exp(c(r-s))\int_0^s \exp(\phi(s-v))dW_1(v)ds \equiv Q_c(J_{\phi,1}(r))$$

and, similarly,

$$m^{-3/2}\Sigma_{j=1}^{k}(\alpha^2)^{k-j}X_{2,j} \Rightarrow \sigma_e\int_0^r \exp(c(r-s))\int_0^s \exp(\phi(s-v))dW_2(v)ds \equiv Q_c(J_{\phi,2}(r)) ,$$

with $W_1(v)$ and $W_2(v)$ independent Wiener processes. Given that $m^{-1/2}\ \Sigma_{j=1}^{k}(\alpha^2)^{k-j}e_{2j} \Rightarrow \int_0^r \exp(\phi(r-v))dW_1(v) \equiv J_{c,1}(r)$, $T(1-\alpha/\rho) \to (\phi - c)$, $T(\alpha-\rho) \to (c-\phi)$, $\alpha \to 1$ and $\rho \to 1$ as $T \to \infty$, we obtain from (A.20) :

$$T^{-1/2}y_{2k} \Rightarrow 2^{-3/2}(\phi-c)Q_c(J_{\phi,1}(r)) + 2^{-3/2}(c-\phi)Q_c(J_{\phi,2}(r)) + 2^{-1/2}J_{c,1}(r)$$

$$\equiv 2^{-3/2}A(r) .$$

To prove part (b), first note that from (4.2) $u_{2k-1} = \Sigma_{j=1}^{2k-1}(-\rho)^{2k-j-1}e_j$. Separating this sum into ones that involve even and odd values of j we have :

$$u_{2k-1} = \Sigma_{j=1}^{k}(\rho^2)^{k-j}e_{2j-1} - (1/\rho)\Sigma_{j=1}^{k-1}(\rho^2)^{k-j}e_{2j}$$

(A.21)
$$= X_{2,k} - (1/\rho)X_{1,k} + (1/\rho)e_{2k}$$
$$= X_{2,k} - X_{1,k} + o_p(T^{1/2}) ,$$

given that $\rho \to 1$ as $T \to \infty$. Hence, we have :

$$T^{-2}\Sigma_{k=1}^{m}u_{2k-1}^2 = (1/4)m^{-2}\Sigma_{k=1}^{m}u_{2k-1}^2 = (1/4)m^{-2}\Sigma_{k=1}^{m}(X_{2,k} - X_{1,k} + o_p(T^{1/2}))^2$$

$$\Rightarrow (\sigma_e^2/4) \int_0^1[J_{\phi,2}(r) - J_{\phi,1}(r)]^2dr \equiv (\sigma_e^2/4) \int_0^1 B(r)^2dr .$$

To prove part (c), note that $\Sigma_{t=1}^{T}u_t^2 = \Sigma_{k=1}^{m}u_{2k}^2 + \Sigma_{k=1}^{m}u_{2k-1}^2$. In a manner similar to part (b), it is easy to show that $u_{2k} = X_{1,k} - \rho X_{2,k}$. Hence $\Sigma_{k=1}^{m}u_{2k}^2 = \Sigma_{k=1}^{m}[X_{1,k} - \rho X_{2,k}]^2$. Using (A.21) we have :

$$\text{(A.22)}\ \Sigma_{t=1}^{T}u_t^2 = (1 + 1/\rho^2)\Sigma_{k=1}^{m}X_{1,k}^2 + (\rho^2 + 1)\Sigma_{k=1}^{m}X_{2,k}^2 - 2(\rho + 1/\rho)\Sigma_{k=1}^{m}X_{1,k}X_{2,k}$$
$$+ (1/\rho^2)\ \Sigma_{k=1}^{m}e_{2k}^2 + (2/\rho)\ \Sigma_{k=1}^{m}X_{2,k}e_{2k} - (2/\rho^2)\Sigma_{k=1}^{m}X_{1,k}e_{2k} .$$

It is easy to verify that the last three terms in (A.22) are $O_p(T)$. Hence using standard convergence arguments and the fact that $\rho \to 1$ as $T \to \infty$:

$$T^{-2}\ \Sigma_{t=1}^{T}u_t^2 = (1/2)\ m^{-2}\Sigma_{k=1}^{m}[X_{1,k} - X_{2,k}]^2 + o_p(1)$$

$$\Rightarrow (\sigma_e^2/2) \int_0^1[J_{\phi,1}(r) - J_{\phi,2}(r)]^2dr \equiv (\sigma_e^2/2) \int_0^1 B(r)^2dr .$$

The proof of part (d) follows straightforwardly from part (a) using the fact that $T^{-2}\Sigma_{k=1}^{m}y_{2k-2}^2 = (1/4)m^{-2}\Sigma_{k=1}^{m}y_{2k-2}^2$ and that $m^{-1/2}y_{2k} \Rightarrow (1/2)A(r)$. To prove part (e) note that (using $y_0 = 0$) :

$$T^{-2}\Sigma_{k=1}^{m}y_{2k-2}u_{2k-1} = T^{-2}\Sigma_{k=1}^{m-1}y_{2k}u_{2k+1} = T^{-2}\Sigma_{k=1}^{m}y_{2k}u_{2k+1} - T^{-2}y_Tu_{T+1}$$

$$= T^{-2}\Sigma_{k=1}^{m} y_{2k}(-\rho u_{2k} + e_{2k+1}) - T^{-2} y_T u_{T+1}$$

(A.23) $$= -T^{-2}\Sigma_{k=1}^{m} y_{2k} u_{2k} + o_p(1) ,$$

given that $\rho \to 1$ as $T \to \infty$ and that both $T^{-2}\Sigma_{k=1}^{m} y_{2k} e_{2k+1}$ and $T^{-2} y_T u_{T+1}$ are $o_p(1)$. Now using (A.20) and the fact that $u_{2k} = X_{1,k} - \rho X_{2,k}$ we have :

$$T^{-2}\Sigma_{k=1}^{m} y_{2k} u_{2k} = (1/4)\, m^{-2}\Sigma_{k=1}^{m} [X_{1,k} - \rho X_{2,k}][(1-\alpha/\rho)\, \Sigma_{j=1}^{k} (\alpha^2)^{k-j} X_{1,j} + (\alpha - \rho)\Sigma_{j=1}^{k} (\alpha^2)^{k-j} X_{2,j} + (\alpha/\rho)\Sigma_{j=1}^{k} (\alpha^2)^{k-j} e_{2j}]$$

$$\Rightarrow (\sigma_e^2/8) \int_0^1 \left\{(\phi - c)[Q_c(J_{\phi,1}(r)) - Q_c(J_{\phi,2}(r))] + 2J_{c,1}(r)\right\} \left\{J_{\phi,1}(r) - J_{\phi,2}(r)\right\} dr \equiv (\sigma^2/8) \int_0^1 A(r)B(r)dr.$$

This proves part (e) using (A.23). The next Lemma characterizes the limiting distribution of the numerator and denominator of $(\hat{\alpha} - \alpha)$, namely $T^{-2}\Sigma_{t=1}^{T} y_{t-1} u_t$ and $T^{-2}\Sigma_{t=1}^{T} y_{t-1}^2$.

LEMMA A.3 : *Let the functions $A(r)$ and $B(r)$ be as defined in Theorem 5 and consider a sequence of random variables $\{y_t\}$ defined by (4.1) and (4.2). Then as $T \to \infty$:*

a) $T^{-2}\Sigma_{t=1}^{T} y_{t-1}^2 \Rightarrow (\sigma_e^2/8) \int_0^1 \left\{[A(r) - B(r)]^2 + B(r)^2\right\} dr$;

b) $T^{-2}\Sigma_{t=1}^{T} y_{t-1} u_t \Rightarrow -(\sigma_e^2/4) \int_0^1 B(r)^2 dr$.

To prove part (a), first note that $\Sigma_1^T y_{t-1}^2 = \Sigma_{k=1}^{m} y_{2k-1}^2 + \Sigma_{k=1}^{m} y_{2k-2}^2$. Using the fact that $y_{2k-1} = \alpha y_{2k-2} + u_{2k-1}$, we deduce that :

$$\Sigma_1^T y_{t-1}^2 = (\alpha^2 + 1)\Sigma_{k=1}^{m} y_{2k-2}^2 + 2\alpha\Sigma_{k=1}^{m} y_{2k-2} u_{2k-1} + \Sigma_{k=1}^{m} u_{2k-1}^2.$$

Using Lemma A.2 (b,d and e), we deduce that :

$$T^{-2}\Sigma_1^T y_{t-1}^2 \Rightarrow (\sigma_e^2/8)\int_0^1 A(r)^2 dr - (\sigma_e^2/4)\int_0^1 A(r)B(r)dr + (\sigma_e^2/4)\int_0^1 B(r)^2 dr$$

$$= (\sigma_e^2/8)\int_0^1\left\{[A(r) - B(r)]^2 + B(r)^2\right\} dr ,$$

as required. To prove part (b), note that using derivations similar to those used to obtain (A.6), we have :

$$T^{-2}\Sigma_1^T y_{t-1}u_t = (1/2\alpha)[T^{-2}y_T^2 - T^{-2}(\alpha^2 - 1)\Sigma_1^T y_{t-1}^2 - T^{-2}\Sigma_1^T u_t^2] .$$

Note that $\alpha \to 1$ and $T(\alpha^2 - 1) \to 2c$ as $T \to \infty$, $y_T^2 = O_p(T)$ using Lemma A.2 (a) and $\Sigma_1^T y_{t-1}^2 = O_p(T^2)$ using part (a). Hence :

$$T^{-2}\Sigma_1^T y_{t-1}u_t = -(1/2)T^{-2}\Sigma_1^T u_t^2 + o_p(1) \Rightarrow -(\sigma_e^2/4)\int_0^1 B(r)^2 dr ,$$

using Lemma A.2 (c). □

The proof of Theorem 5 follows using the fact that $\hat{\alpha} = \alpha + T^{-2}\Sigma_1^T y_{t-1}u_t/T^{-2}\Sigma_1^T y_{t-1}^2$ with Lemma A.3 and noting that $\alpha \to 1$ as $T \to \infty$. □

Proof of Theorem 6 : Using the notation $U_T^* = 2T^{-2}\Sigma_{t=2}^T y_t y_{t-1}$, $V_T^* = 2T^{-2}\Sigma_{t=2}^T y_{t-1}^2$, $\alpha_1 = \exp(c/T)$ and $\alpha_2 = -\exp(\phi/T)$, we have from (A.8) through (A.11) :

$$V_T^* + U_T^* = TU_T + 2T^2V_T$$

(A.24) $$= T^{-1}\Sigma_{j=1}^T\Sigma_{k=1}^T\left\{2T^{-1}(\alpha_1 - \alpha_2)^{-2}(\alpha_1^{|j-k|} - \alpha_1^{2T-j-k+2})/(1 - \alpha_1)\right\}e_je_k + O_p(T^{-1}),$$

$$V_T^* - U_T^* = -TU_T$$

(A.25) $$= T^{-1}\Sigma_{j=1}^T\Sigma_{k=1}^T\left\{2T^{-1}(\alpha_1 - \alpha_2)^{-2}(\alpha_2^{|j-k|} - \alpha_2^{2T-j-k+2})/(1 + \alpha_2)\right\}e_je_k + O_p(T^{-1}).$$

Let $(W_1(\cdot), W_2(\cdot))$ be a pair of independent Brownian motions defined by $(T^{-1/2}\Sigma_{j=1}^{[Tr]}e_j, T^{-1/2}\Sigma_{j=1}^{[Tr]}(-1)^j e_j) \Rightarrow (W_1(r), W_2(r))$ $(0 \leq r \leq 1)$. Let $B_1(j,k)$ (j, k = 1, ..., T) be the coefficients associated with $e_j e_k$ in the sum defined by (A.24) and let $B_2(j,k)$ (j, k, = 1, ..., T) be the coefficients associated with $(-1)^j e_j (-1)^k e_k$ in (A.25). We have :

$$B_1(j,k) = 2T^{-1}(\alpha_1 - \alpha_2)^{-2}(\alpha_1^{|j-k|} - \alpha_1^{2T-j-k+2})/(1 - \alpha_1) ,$$

$$B_2(j,k) = 2T^{-1}(\alpha_1 - \alpha_2)^{-2}(|\alpha_2|^{|j-k|} - |\alpha_2|^{2T-j-k+2})/(1 - |\alpha_2|).$$

Let $a_1 = c$ and $a_2 = \phi$ and define the functions

$$K_i(s,t) = (-1/2a_i)(\exp(a_i|s-t|) - \exp(a_i(2-s-t))) . \qquad (i = 1, 2)$$

We have :

$$\lim_{T\to\infty} \max_{j,k} |B_i(j,k) - K_i(j/T,k/T)| = 0 . \qquad (i = 1, 2)$$

Therefore, $(V_T^* + U_T^*, V_T^* - U_T^*)$ converge jointly in distribution to (Z_1, Z_2) as $T \to \infty$, where

$$Z_i = \int_0^1\int_0^1(-1/2a_i)[\exp(a_i|s-t|) - \exp(a_i(2-s-t))]dW_i(s)dW_i(t) . \qquad (i = 1,2)$$

The characteristic functions of Z_1 and Z_2 are given, for i = 1, 2, by (see Nabeya and Tanaka (1990)) :

$$E[\exp(iwZ_i)] = \exp(-a_i/2)\left\{\cosh\{(a_i^2 - iw)^{1/2}\} - a_i \sinh\{(a_i^2 - iw)^{1/2}\}/(a_i^2 - iw)^{1/2}\right\}^{-1/2}.$$

Since Z_1 and Z_2 are independent, the joint limiting characteristic function of (U_T^*, V_T^*) becomes :

$$E[\exp(i(u+v)Z_1/2)]E[\exp(i(v-u)Z_2/2)] =$$

$$\exp(-(c+\phi)/2)\Big\{[\cosh\{(c^2 - i(u+v)/2)^{1/2}\} - c \sinh\{(c^2 - i(u+v)/2)^{1/2}\}/(c^2 - i(u+v)/2)^{1/2}] [\cosh\{(\phi^2 - i(v-u)/2)^{1/2}\} - \phi \sinh\{(\phi^2 - i(v-u)/2)^{1/2}\}/(\phi^2 - i(v-u)/2)^{1/2}]\Big\}^{-1/2}.$$

Since $\hat{\alpha} \leq x$ is equivalent to $xV_T^* - U_T^* \geq 0$, the theorem is proved substituting $u = -w$ and $v = wx$.

TABLE I : Percentage Points of the Distribution of $T(\hat{\alpha}-\alpha)$; $\alpha = \exp(c/T)$, $c = 0.0$; Nearly White Noise Model.

Asymptotic values based on $\delta = T^{1/2}(1+\theta)$; finite sample values in parentheses.

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%
$\theta = -0.90$									
T=25	–24.81	–24.73	–24.65	–24.53	–23.31	–19.25	–17.68	–16.30	–14.73
	(–34.69)	(–32.80)	(–31.05)	(–28.95)	(–21.23)	(–13.05)	(–10.90)	(–9.14)	(–7.42)
T=50	–49.17	–48.93	–48.62	–48.16	–43.66	–31.29	–27.35	–24.19	–20.89
	(–61.57)	(–58.22)	(–55.76)	(–52.61)	(–39.95)	(–23.90)	(–19.87)	(–16.52)	(–13.64)
T=100	–96.75	–95.75	–94.63	–92.89	–77.49	–45.54	–37.65	–31.91	–26.40
	(–110.28)	(–106.09)	(–101.79)	(–96.30)	(–71.38)	(–37.56)	(–29.85)	(–24.68)	(–19.33)
T=500	–427.00	–408.93	–389.68	–361.59	–203.88	–71.65	–53.89	–42.84	–33.48
	(–427.96)	(–407.79)	(–385.31)	(–355.24)	(–192.30)	(–63.81)	(–45.28)	(–33.94)	(–24.94)
T=1000	–744.25	–691.85	–638.78	–566.38	–256.12	–77.17	–56.94	–44.74	–34.58
	(–733.45)	(–675.98)	(–625.27)	(–553.72)	(–234.38)	(–67.17)	(–49.53)	(–38.65)	(–29.70)
T=5000	–1837.34	–1549.38	–1305.06	–1036.31	–322.20	–82.25	–59.69	–46.40	–35.63
	(–1867.76)	(–1570.88)	(–1275.06)	(–999.41)	(–315.94)	(–92.19)	(–71.18)	(–57.97)	(–43.94)
$\theta = -0.70$									
T=25	–23.21	–22.72	–22.18	–21.33	–15.12	–6.77	–5.29	–4.31	–3.44
	(–30.34)	(–27.72)	(–25.47)	(–22.73)	(–11.72)	(–3.43)	(–2.24)	(–1.47)	(–0.78)
T=50	–43.36	–41.66	–39.84	–37.19	–21.67	–7.83	–5.92	–4.71	–3.69
	(–49.25)	(–44.76)	(–40.64)	(–35.63)	(–16.52)	(–4.41)	(–3.05)	(–2.10)	(–1.34)
T=100	–76.42	–71.38	–66.25	–59.21	–27.67	–8.50	–6.29	–4.95	–3.83
	(–77.81)	(–70.13)	(–62.69)	(–53.22)	(–20.46)	(–5.23)	(–3.45)	(–2.47)	(–1.59)
T=500	–196.22	–166.42	–140.88	–112.57	–35.55	–9.12	–6.63	–5.15	–3.96
	(–162.58)	(–135.44)	(–115.70)	(–92.08)	(–27.12)	(–5.86)	(–3.80)	(–2.45)	(–1.46)
T=1000	–245.11	–199.78	–164.27	–126.77	–36.84	–9.21	–6.67	–5.18	–3.97
	(–197.33)	(–160.71)	(–133.73)	(–102.39)	(–26.75)	(–5.72)	(–3.92)	(–2.82)	(–1.91)
$\theta = -0.50$									
T=25	–20.58	–19.56	–18.48	–16.92	–8.88	–2.95	–2.20	–1.74	–1.36
	(–24.90)	(–21.54)	(–18.88)	(–15.64)	(–5.47)	(–0.63)	(–0.07)	(0.36)	(0.77)
T=50	–35.08	–32.13	–29.31	–25.57	–10.80	–3.14	–2.30	–1.81	–1.40
	(–34.69)	(–30.10)	(–25.55)	(–20.45)	(–6.36)	(–0.83)	(–0.21)	(0.17)	(0.56)
T=100	–53.99	–47.33	–41.45	–34.32	–12.10	–3.24	–2.36	–1.84	–1.41
	(–47.17)	(–39.19)	(–32.54)	(–25.19)	(–6.86)	(–0.95)	(–0.32)	(0.07)	(0.43)
T=500	–94.55	–76.21	–62.02	–47.31	–13.40	–3.32	–2.40	–1.87	–1.43
	(–62.77)	(–49.76)	(–41.22)	(–31.25)	(–7.88)	(–1.00)	(–0.35)	(0.07)	(0.45)

TABLE II : Distribution of $T(\hat{\alpha}-\alpha)$; $\alpha = \exp(c/T)$; Nearly White Noise Model

Asymptotic values based on $\delta = T^{1/2}(1+\theta)$; finite sample values in parentheses.

A : Mean

c	θ	T=25	T=50	T=100	T=500	T=1000	T=5000
0.0	–0.90	–22.51	–41.50	–73.02	–210.72	–291.68	–456.60
		(–21.13)	(–39.07)	(–69.00)	(–201.45)	(–275.74)	(–447.14)
	–0.70	–14.54	–22.12	–31.01	–49.83	–54.88	
		(–12.43)	(–18.49)	(–25.32)	(–39.19)	(–42.33)	
	–0.50	–9.45	–12.73	–15.80	–20.37	–21.24	
		(–6.92)	(–8.77)	(–10.39)	(–12.65)	(–13.77)	
–5.0	–0.90	–24.45	–47.88	–91.96	–354.94	–561.33	–1103.46
		(–22.07)	(–45.08)	(–88.46)	(–341.04)	(–539.09)	(–1018.31)
	–0.70	–20.98	–36.49	–58.54	–119.19	–138.92	
		(–16.25)	(–30.63)	(–48.70)	(–91.69)	(–106.82)	
	–0.50	–16.62	–25.48	–35.25	–52.53	–56.27	
		(–11.13)	(–16.76)	(–22.25)	(–29.71)	(–32.20)	
2.0	–0.90	–17.17	–28.60	–45.37	–107.63	–141.02	–206.18
		(–13.82)	(–23.96)	(–36.64)	(–101.14)	(–133.57)	(–183.44)
	–0.70	–8.10	–11.41	–15.10	–22.56	–24.51	
		(–6.89)	(–9.41)	(–12.45)	(–18.06)	(–19.61)	
	–0.50	–4.73	–6.07	–7.29	–9.06	–9.39	
		(–3.80)	(–4.46)	(–5.23)	(–6.09)	(–6.31)	

B : Variance

c	θ	T=25	T=50	T=100	T=500	T=1000	T=5000
0.0	–0.90	5.20	46.22	326.00	11380.04	34048.96	167052.12
		(37.34)	(119.64)	(482.88)	(11573.26)	(33716.06)	(158516.00)
	–0.70	28.73	116.25	361.89	1928.47	2796.55	
		(52.61)	(142.37)	(352.78)	(1347.55)	(1882.63)	
	–0.50	26.66	73.38	157.03	405.98	475.96	
		(36.21)	(66.38)	(111.29)	(192.09)	(258.00)	
–5.0	–0.90	0.10	1.44	18.27	3071.33	16813.96	213117.59
		(22.11)	(48.76)	(119.51)	(3965.71)	(19248.78)	(235949.10)
	–0.70	3.44	28.38	164.09	2358.53	4302.75	
		(28.73)	(82.83)	(257.62)	(1882.48)	(3257.43)	
	–0.50	8.73	43.03	144.62	677.08	901.64	
		(32.35)	(74.40)	(150.83)	(349.85)	(450.90)	
2.0	–0.90	35.76	185.27	796.71	11931.33	28353.66	103245.87
		(82.30)	(263.61)	(904.97)	(11469.52)	(27448.45)	(92427.60)
	–0.70	43.31	127.07	310.54	1205.10	1634.98	
		(56.21)	(126.45)	(279.97)	(843.09)	(1129.94)	
	–0.50	25.74	57.56	106.55	235.49	260.28	
		(29.59)	(46.39)	(75.04)	(116.79)	(133.04)	

TABLE III.A : The Distribution of $T(\hat{\alpha}-\alpha)$; $\alpha = \exp(c/T)$, $c = 0.0$; AR Errors , $u_t = \rho u_{t-1} + e_t$.

Nearly Doubly Integrated Model; asymptotic values based on $\phi = T(\rho-1)$; finite sample values in parentheses.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%	Mean	Var
$\rho = 0.50$									
T=25	0.0009	0.0208	0.0793	0.9573	2.5095	3.1668	4.7209	1.17	1.08
	(−5.28)	(−2.67)	(−1.73)	(0.27)	(1.76)	(2.31)	(3.46)	(0.10)	(2.55)
T=50	0.0007	0.0169	0.0650	0.8705	2.3757	3.0131	4.5331	1.09	0.99
	(−4.47)	(−2.47)	(−1.61)	(0.24)	(1.65)	(2.11)	(3.26)	(0.09)	(2.08)
T=100	0.0006	0.0152	0.0586	0.8260	2.2979	2.9098	4.4079	1.04	0.92
	(−4.67)	(−2.44)	(−1.63)	(0.20)	(1.62)	(2.06)	(3.26)	(0.07)	(2.08)
$\rho = 0.90$									
T=25	0.0025	0.0586	0.2067	1.4141	2.9947	3.6672	5.2185	1.56	1.28
	(−1.66)	(−0.57)	(−0.28)	(1.25)	(2.83)	(3.44)	(4.86)	(1.27)	(1.67)
T=50	0.0014	0.0350	0.1292	1.1806	2.7781	3.4522	5.0213	1.37	1.21
	(−1.04)	(−0.38)	(−0.17)	(1.04)	(2.61)	(3.25)	(4.63)	(1.14)	(1.36)
T=100	0.0009	0.0230	0.0870	0.9989	2.5666	3.2299	4.7922	1.21	1.11
	(−0.73)	(−0.28)	(−0.12)	(0.86)	(2.33)	(3.00)	(4.47)	(1.01)	(1.14)
T=500	0.0006	0.0152	0.0586	0.8260	2.2979	2.9098	4.4079	1.04	0.92
	(−0.58)	(−0.26)	(−0.12)	(0.69)	(2.15)	(2.73)	(4.07)	(0.87)	(0.98)
$\rho = 0.95$									
T=25	0.0039	0.0905	0.3034	1.6238	3.1601	3.8266	5.3546	1.73	1.31
	(−1.35)	(−0.40)	(−0.11)	(1.52)	(3.07)	(3.73)	(5.06)	(1.53)	(1.68)
T=50	0.0025	0.0586	0.2067	1.4141	2.9947	3.6672	5.2185	1.56	1.28
	(−0.69)	(−0.21)	(−0.05)	(1.33)	(2.89)	(3.56)	(5.02)	(1.42)	(1.42)
T=100	0.0014	0.0350	0.1292	1.1806	2.7781	3.4522	5.0213	1.37	1.21
	(−0.43)	(−0.13)	(−0.02)	(1.11)	(2.63)	(3.29)	(4.76)	(1.24)	(1.22)
T=500	0.0007	0.0169	0.0650	0.8705	2.3757	3.0131	4.5331	1.09	0.99
	(−0.28)	(−0.10)	(−0.03)	(0.79)	(2.29)	(2.89)	(4.43)	(0.99)	(1.01)
T=1000	0.0006	0.0152	0.0586	0.8260	2.2979	2.9098	4.4079	1.04	0.92
	(−0.28)	(−0.11)	(−0.04)	(0.71)	(2.24)	(2.78)	(4.28)	(0.94)	(0.97)

TABLE III.B : The Distribution of $T(\hat{\alpha}-\alpha)$; $\alpha = \exp(c/T)$, $c = -5.0$; AR Errors , $u_t = \rho u_{t-1} + e_t$.

Nearly Doubly Integrated Model; asymptotic values based on $\phi = T(\rho-1)$; finite sample values in parentheses.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%	Mean	Var
$\rho = 0.50$									
T=25	5.0001	5.0036	5.0144	5.3900	6.9715	7.6807	9.3238	5.75	0.89
	(−3.73)	(−1.07)	(0.10)	(2.83)	(4.45)	(4.98)	(6.20)	(2.50)	(3.74)
T=50	5.0001	5.0031	5.0125	5.3463	6.8375	7.5253	9.1371	5.69	0.79
	(−2.85)	(−0.44)	(0.56)	(3.05)	(4.57)	(5.08)	(6.21)	(2.78)	(3.02)
T=100	5.0001	5.0029	5.0116	5.3237	6.7542	7.4235	9.40026	5.65	0.71
	(−2.77)	(−0.33)	(0.69)	(3.17)	(4.63)	(5.13)	(6.43)	(2.89)	(3.03)
$\rho = 0.90$									
T=25	5.0004	5.0086	5.0340	5.6642	7.4108	8.1289	9.7649	5.99	1.13
	(1.16)	(2.82)	(3.37)	(4.59)	(6.30)	(6.92)	(8.33)	(4.71)	(1.69)
T=50	5.0002	5.0053	5.0211	5.5091	7.2184	7.9429	9.5966	5.87	1.04
	(2.77)	(3.63)	(3.96)	(4.79)	(6.39)	(7.10)	(8.47)	(5.01)	(1.18)
T=100	5.0002	5.0039	5.0155	5.4111	7.0259	7.7411	9.3909	5.77	0.92
	(3.33)	(4.02)	(4.25)	(4.87)	(6.37)	(7.01)	(8.42)	(5.12)	(0.95)
T=500	5.0001	5.0029	5.0116	5.3237	6.7542	7.4235	9.40026	5.65	0.71
	(3.92)	(4.31)	(4.48)	(4.94)	(6.32)	(6.92)	(8.34)	(5.21)	(0.75)
$\rho = 0.95$									
T=25	5.0006	5.0146	5.0564	5.8420	7.5616	8.2660	9.8784	6.12	1.19
	(1.69)	(3.25)	(3.72)	(4.95)	(6.64)	(7.27)	(8.64)	(5.05)	(1.68)
T=50	5.0004	5.0086	5.0340	5.6642	7.4108	8.1289	9.7649	5.99	1.13
	(3.30)	(4.02)	(4.29)	(5.10)	(6.79)	(7.52)	(8.91)	(5.35)	(1.24)
T=100	5.0002	5.0053	5.0211	5.5091	7.2184	7.9429	9.5966	5.87	1.04
	(3.96)	(4.41)	(4.57)	(5.13)	(6.75)	(7.43)	(9.02)	(5.43)	(1.01)
T=500	5.0001	5.0031	5.0125	5.3463	6.8375	7.5253	9.1371	5.69	0.79
	(4.44)	(4.66)	(4.76)	(5.12)	(6.59)	(7.28)	(8.72)	(5.44)	(0.79)
T=1000	5.0001	5.0029	5.0116	5.3237	6.7542	7.4235	9.40026	5.65	0.71
	(4.48)	(4.68)	(4.76)	(5.12)	(6.49)	(7.15)	(8.55)	(5.42)	(0.71)

TABLE III.C : The Distribution of $T(\hat{\alpha}-\alpha)$; $\alpha = \exp(c/T)$, $c = 2.0$; AR Errors , $u_t = \rho u_{t-1} + e_t$.

Nearly Doubly Integrated Model; asymptotic values based on $\phi = T(\rho-1)$; finite sample values in parentheses.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%	Mean	Var
$\rho = 0.50$									
T=25	–1.9887	–1.7541	–1.2933	0.1518	1.1684	1.6809	3.0489	0.10	1.00
	(–5.85)	(–3.36)	(–2.43)	(–0.03)	(0.93)	(1.32)	(2.42)	(–0.40)	(2.39)
T=50	–1.9900	–1.7783	–1.3453	0.1005	1.0775	1.5633	2.8803	0.04	0.95
	(–5.07)	(–3.34)	(–2.37)	(–0.06)	(0.83)	(1.19)	(2.21)	(–0.40)	(1.93)
T=100	–1.9906	–1.7900	–1.3712	0.0751	1.0277	1.4961	2.7740	0.00	0.92
	(–5.39)	(–3.24)	(–2.39)	(–0.07)	(0.77)	(1.09)	(2.13)	(–0.45)	(1.98)
$\rho = 0.90$									
T=25	–1.9796	–1.5934	–0.9817	0.4545	1.5606	2.1345	3.5565	0.42	1.12
	(–3.11)	(–2.19)	(–1.35)	(0.41)	(1.55)	(2.13)	(3.49)	(0.29)	(1.55)
T=50	–1.9848	–1.6829	–1.1487	0.2949	1.3741	1.9266	3.3444	0.25	1.08
	(–2.63)	(–2.07)	(–1.42)	(0.25)	(1.31)	(1.88)	(3.24)	(0.15)	(1.27)
T=100	–1.9880	–1.7418	–1.2676	0.1773	1.2094	1.7320	2.1160	0.13	1.02
	(–2.74)	(–2.05)	(–1.48)	(0.14)	(1.12)	(1.58)	(2.81)	(0.03)	(1.12)
T=500	–1.9906	–1.7900	–1.3712	0.0751	1.0277	1.4961	2.7740	0.00	0.92
	(–2.40)	(–2.03)	(–1.56)	(0.03)	(0.99)	(1.39)	(2.67)	(–0.08)	(1.03)
$\rho = 0.95$									
T=25	–1.9737	–1.4987	–0.8196	0.6005	1.7109	2.2978	3.7112	0.56	1.15
	(–2.83)	(–2.04)	(–1.11)	(0.59)	(1.75)	(2.33)	(3.68)	(0.48)	(1.53)
T=50	–1.9796	–1.5934	–0.9817	0.4545	1.5606	2.1345	3.5565	0.42	1.12
	(–2.41)	(–1.88)	(–1.18)	(0.44)	(1.56)	(2.16)	(3.56)	(0.36)	(1.29)
T=100	–1.9848	–1.6829	–1.1487	0.2949	1.3741	1.9266	3.3444	0.25	1.08
	(–2.25)	(–1.89)	(–1.30)	(0.27)	(1.31)	(1.83)	(3.12)	(0.19)	(1.14)
T=500	–1.9900	–1.7783	–1.3453	0.1005	1.0775	1.5633	2.8803	0.04	0.95
	(–2.16)	(–1.92)	(–1.47)	(0.07)	(1.05)	(1.48)	(2.78)	(–0.02)	(1.00)
T=1000	–1.9906	–1.7900	–1.3712	0.0751	1.0277	1.4961	2.7740	0.00	0.92
	(–2.15)	(–1.87)	(–1.41)	(0.04)	(1.00)	(1.45)	(2.57)	(–0.02)	(0.93)

TABLE IV : The Distribution of $T(\hat{\alpha}-\alpha)$; $\alpha = \exp(c/T)$; AR Errors , $u_t = \rho u_{t-1} + e_t$, $\rho = -0.90$.

Nearly Seasonally Integrated Model; Asymptotic values based on $\phi = -T(1+\rho)$; finite sample values in parentheses.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%	Mean	Var
c = 0.0									
T=25	–43.66	–38.27	–34.05	–14.82	–3.72	–2.49	–1.22	–16.99	128.03
	(–44.24)	(–38.31)	(–33.76)	(–13.89)	(–2.70)	(–1.39)	(–0.04)	(–16.24)	(134.04)
T=50	–77.38	–63.67	–54.42	–20.77	–5.19	–3.55	–1.83	–25.67	361.66
	(–76.33)	(–62.43)	(–53.52)	(–19.36)	(–4.16)	(–2.42)	(–0.71)	(–24.35)	(359.69)
T=100	–123.81	–94.23	–77.07	–25.91	–6.43	–4.49	–2.43	–34.70	826.52
	(–122.62)	(–93.53)	(–75.20)	(–23.93)	(–5.36)	(–3.34)	(–1.34)	(–33.10)	(826.35)
T=500	–230.21	–150.92	–114.79	–32.30	–7.89	–5.66	–3.28	–49.31	2432.05
	(–214.48)	(–146.79)	(–112.24)	(–31.24)	(–6.58)	(–4.09)	(–1.71)	(–47.75)	(2284.21)
c = –5.0									
T=25	–41.46	–38.95	–36.98	–25.80	–11.76	–8.45	–3.62	–25.48	88.08
	(–42.21)	(–39.69)	(–37.58)	(–25.90)	(–11.74)	(–8.71)	(–3.60)	(–25.29)	(91.39)
T=50	–81.10	–73.67	–68.53	–45.00	–21.47	–16.33	–8.90	–45.24	306.80
	(–81.36)	(–73.48)	(–68.17)	(–44.22)	(–20.83)	(–15.84)	(–8.38)	(–44.42)	(308.97)
T=100	–142.23	–123.07	–111.39	–67.47	–32.22	–25.11	–15.01	–70.05	895.27
	(–142.18)	(–123.01)	(–110.87)	(–66.12)	(–31.37)	(–24.38)	(–15.03)	(–68.69)	(899.28)
T=500	–317.55	–244.17	–208.34	–107.98	–50.04	–39.79	–25.69	–120.52	4120.31
	(–302.56)	(–233.83)	(–199.97)	(–104.99)	(–46.22)	(–35.38)	(–21.05)	(–115.80)	(3805.87)
c = 2.0									
T=25	–42.90	–33.89	–26.87	–6.07	–2.60	–2.38	–2.18	–10.69	102.84
	(–44.03)	(–33.32)	(–25.56)	(–4.15)	(–0.34)	(0.00)	(0.47)	(–8.86)	(115.69)
T=50	–72.00	–51.89	–38.72	–7.13	–2.81	–2.53	–2.26	–14.46	262.04
	(–70.20)	(–50.24)	(–36.36)	(–5.08)	(–0.55)	(–0.22)	(0.20)	(–12.20)	(264.31)
T=100	–109.10	–71.09	–50.09	–7.87	–2.97	–2.66	–2.34	–18.20	543.28
	(–109.43)	(–71.88)	(–49.03)	(–5.47)	(–0.73)	(–0.39)	(0.04)	(–16.09)	(573.18)
T=500	–184.11	–101.44	–65.90	–8.62	–3.16	–2.81	–2.46	–24.00	1385.69
	(–176.34)	(–99.11)	(–63.53)	(–6.47)	(–0.89)	(–0.50)	(–0.09)	(–21.86)	(1360.10)

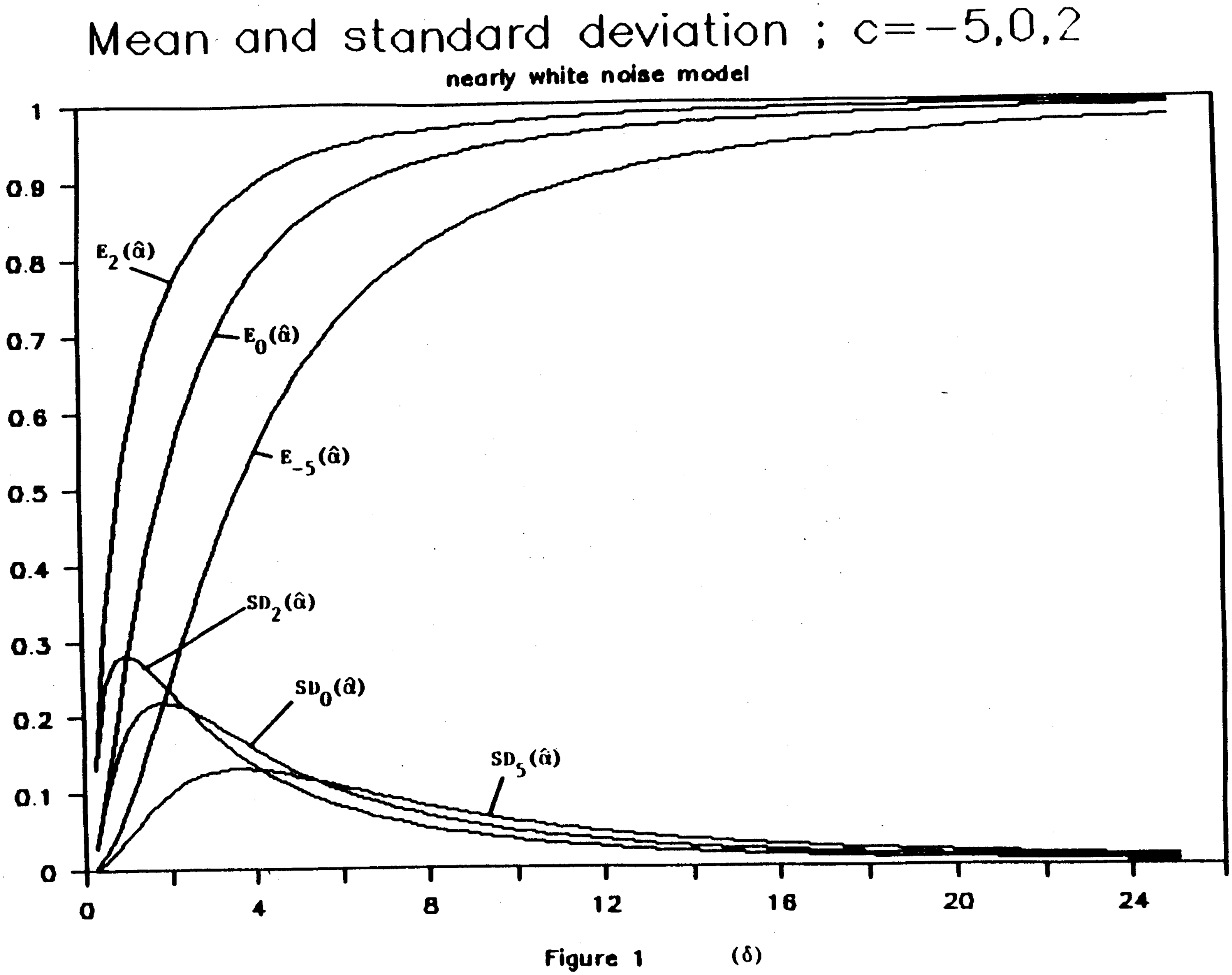


Figure 1

Notes: $E_x(\hat{\alpha})$ denotes the asymptotic mean of $\hat{\alpha}$ when c=x. Similarly, $SD_x(\hat{\alpha})$ denotes the asymptotic standard deviation of $\hat{\alpha}$ when c=x. Each curve is drawn from calculations at 100 equidistant points.

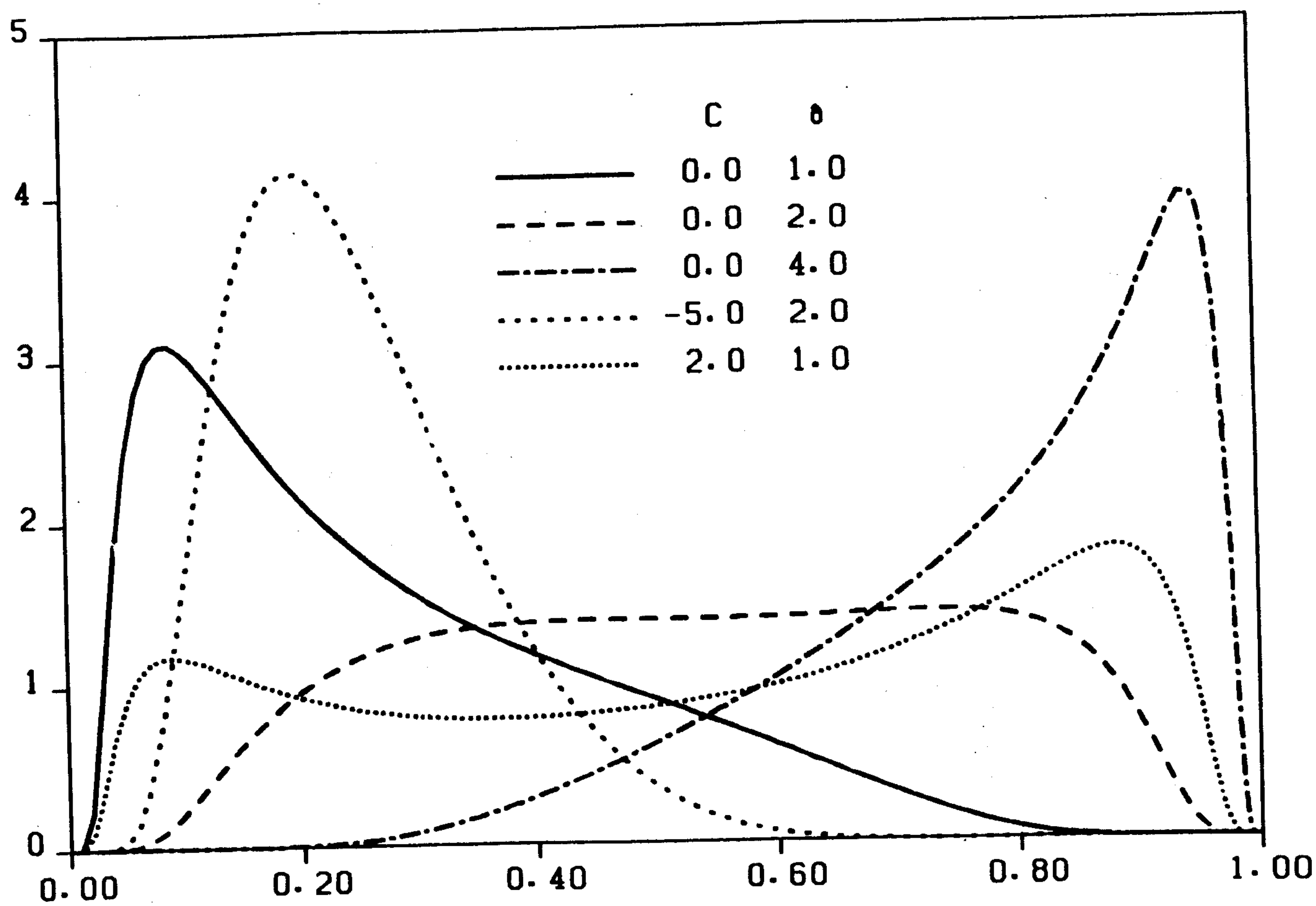


Figure 2: Limiting density functions of $\hat{\alpha}$; nearly white noise model.

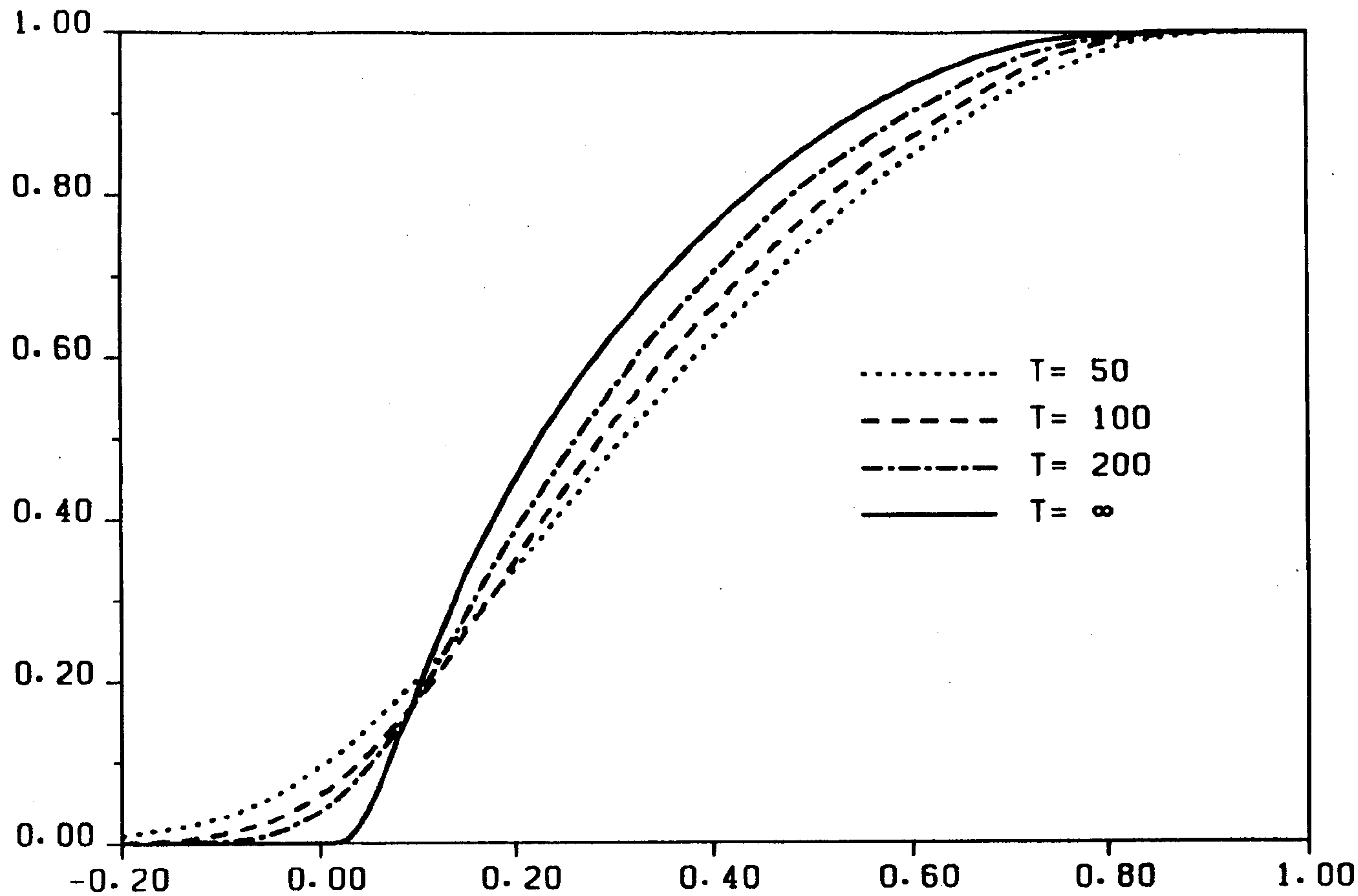
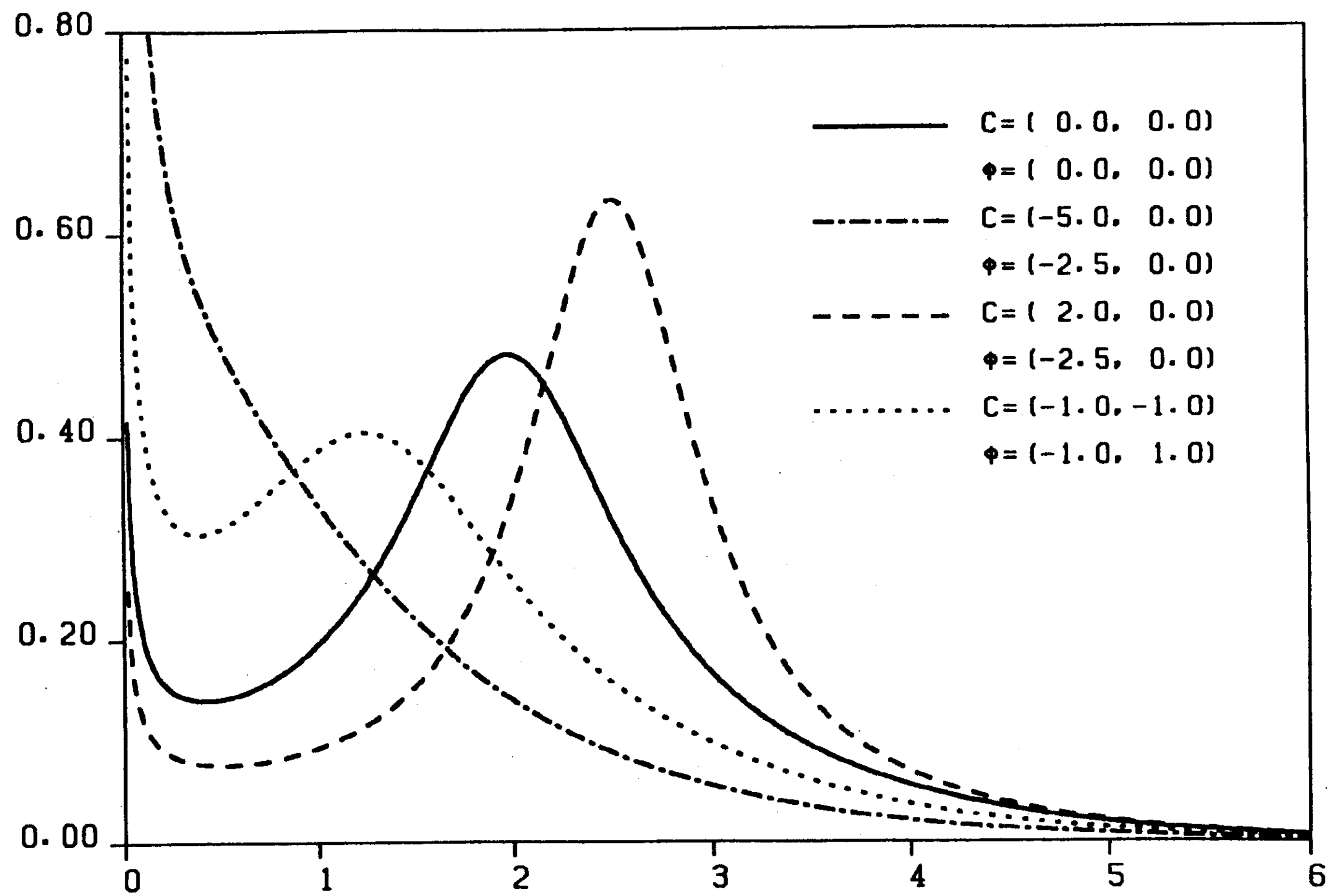


Figure 3: Nearly white noise asymptotic distribution of $\hat{\alpha}$ and the corresponding finite sample distributions; $c = 0$, $\delta = 1$.



<u>Figure 4</u>: Limiting density functions of $T(\hat{\alpha}-1)$; nearly doubly integrated model.

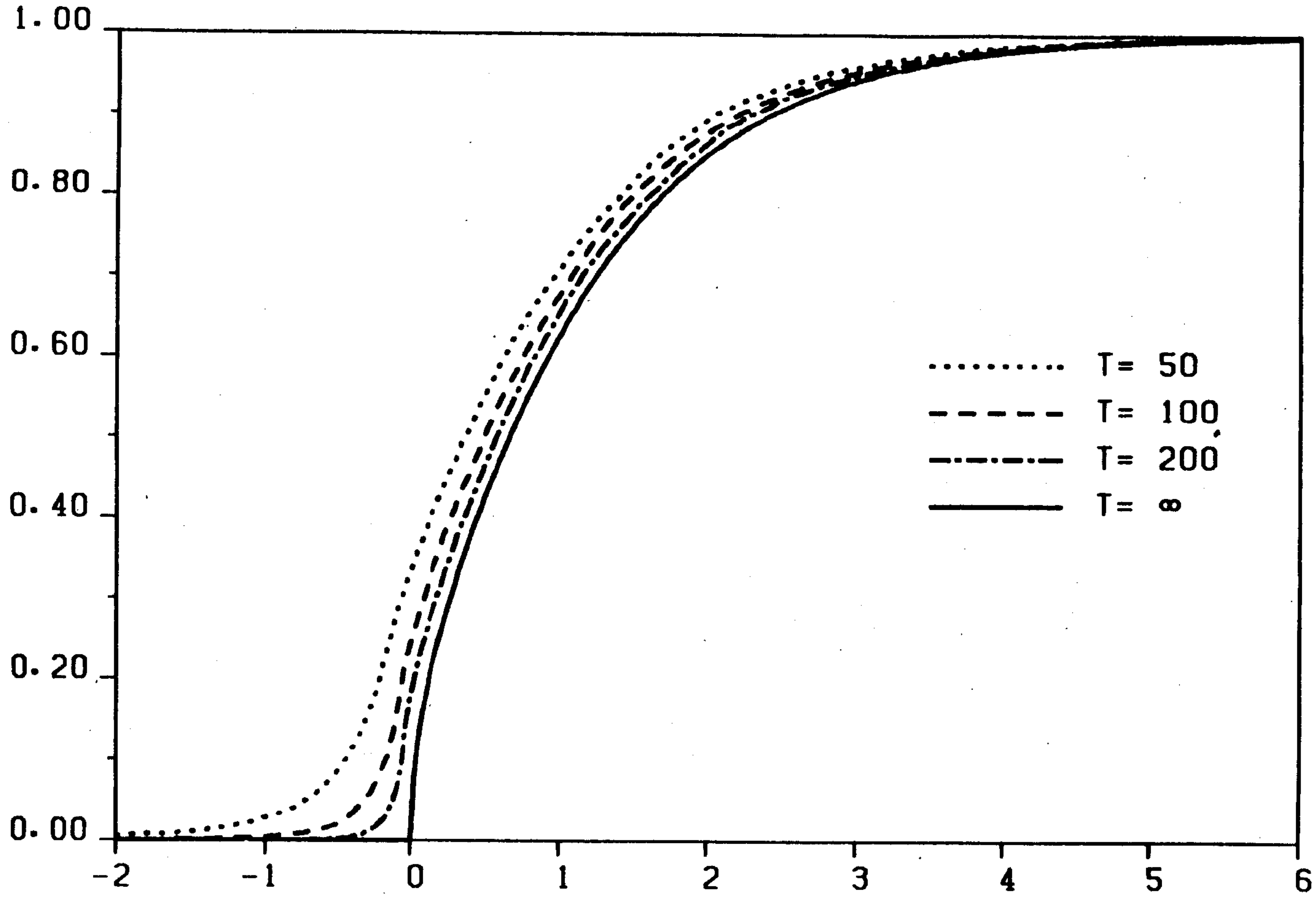


Figure 5: Nearly doubly integrated asymptotic distribution of $T(\hat{\alpha}-1)$ and the corresponding finite sample distributions; $c = -5$, $\phi = -2.5$.

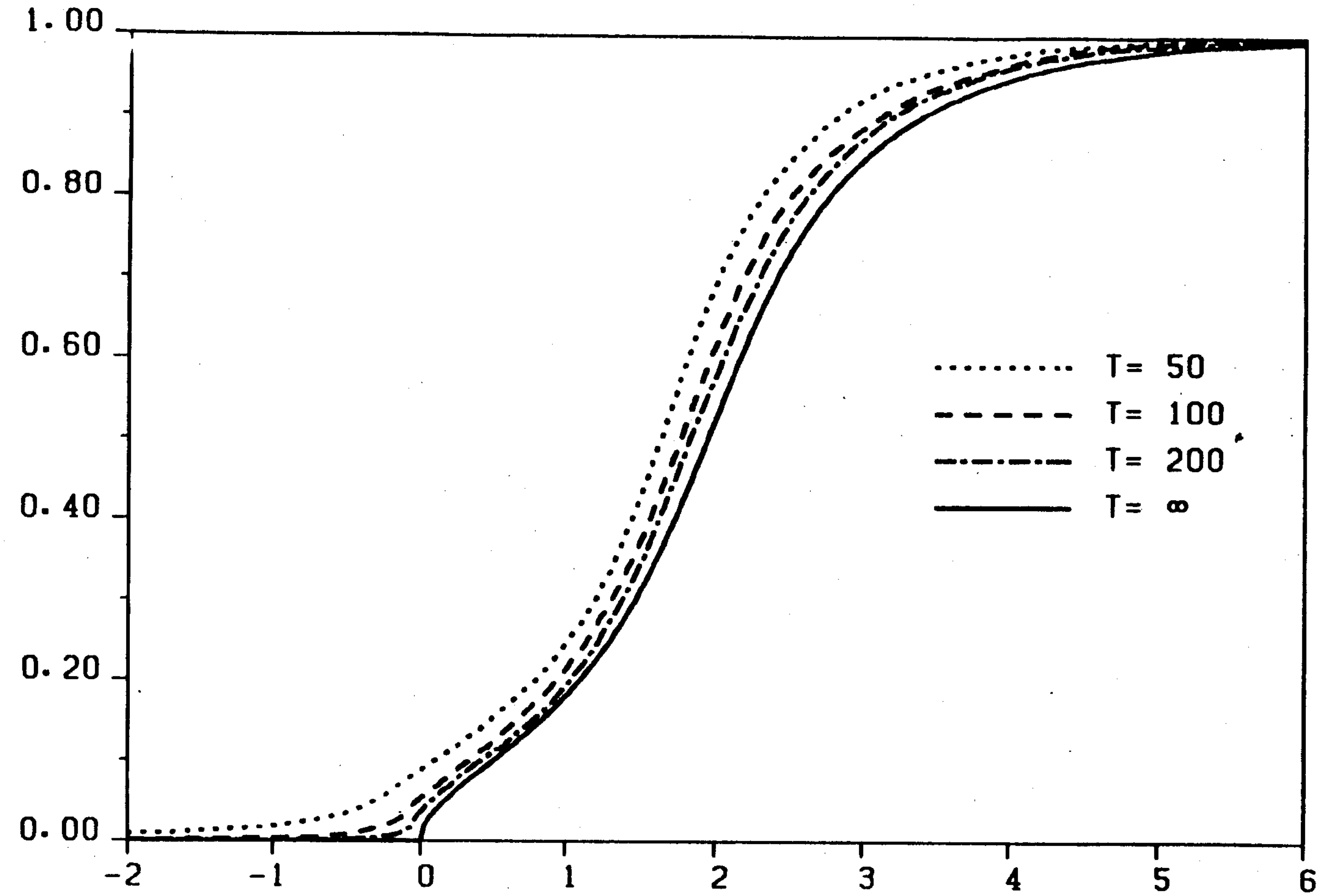


Figure 6: Nearly doubly integrated asymptotic distribution of $T(\hat{\alpha}-1)$ and the corresponding finite sample distributions; $c = \phi = 0$ with MA(1) errors, $e_t = \eta_t - (3/4)\eta_{t-1}$ ($\eta_t \sim$ i.i.d.N(0,1)).

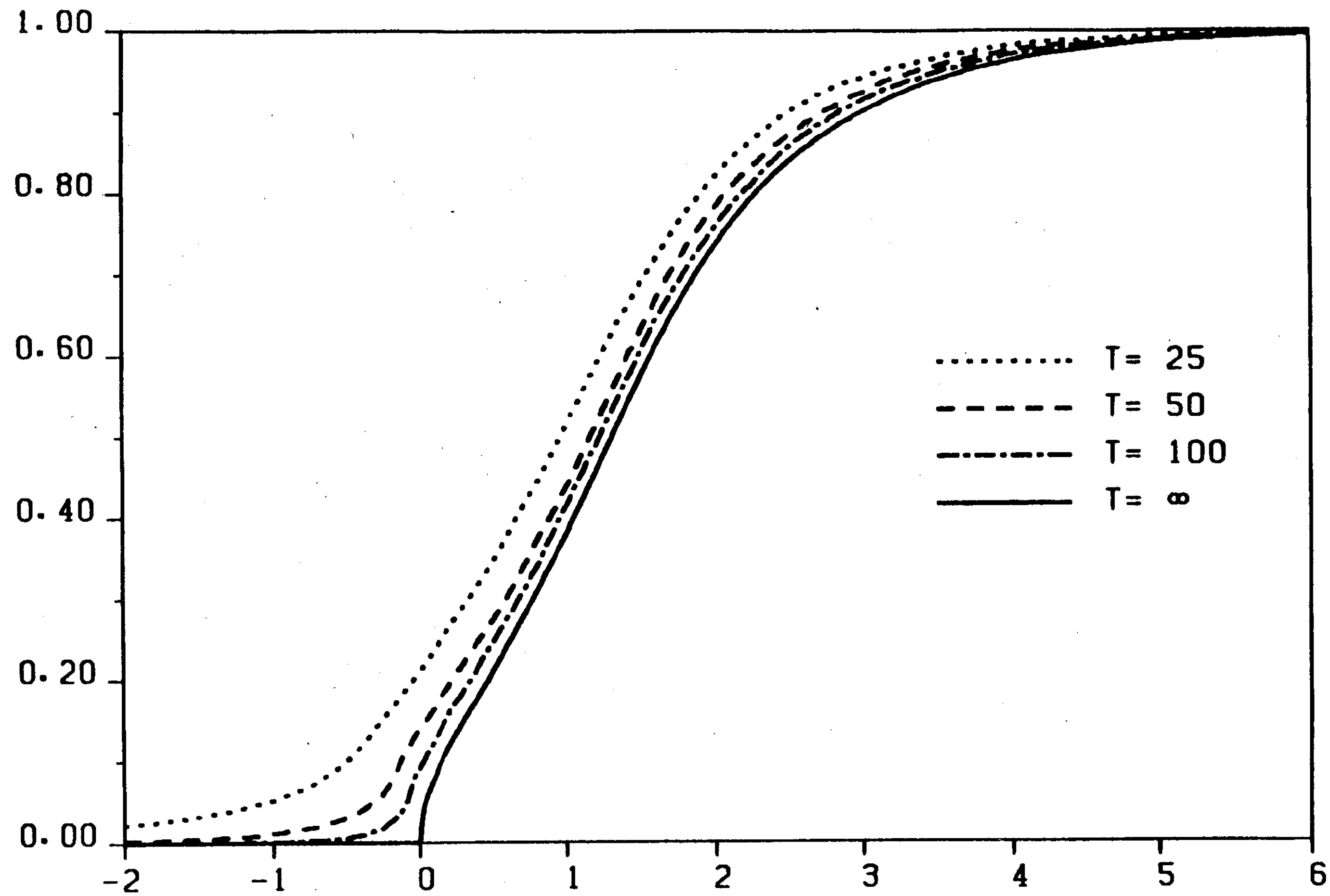


Figure 7: Nearly doubly integrated asymptotic distribution of $T(\hat{\alpha}-1)$ and the corresponding finite sample distributions; $c = -1-i$, $\phi = -1+i$ with AR(1) errors, $e_t = -(2/3)e_{t-1} + \eta_t$ (η_t~i.i.d.N(0,1)).

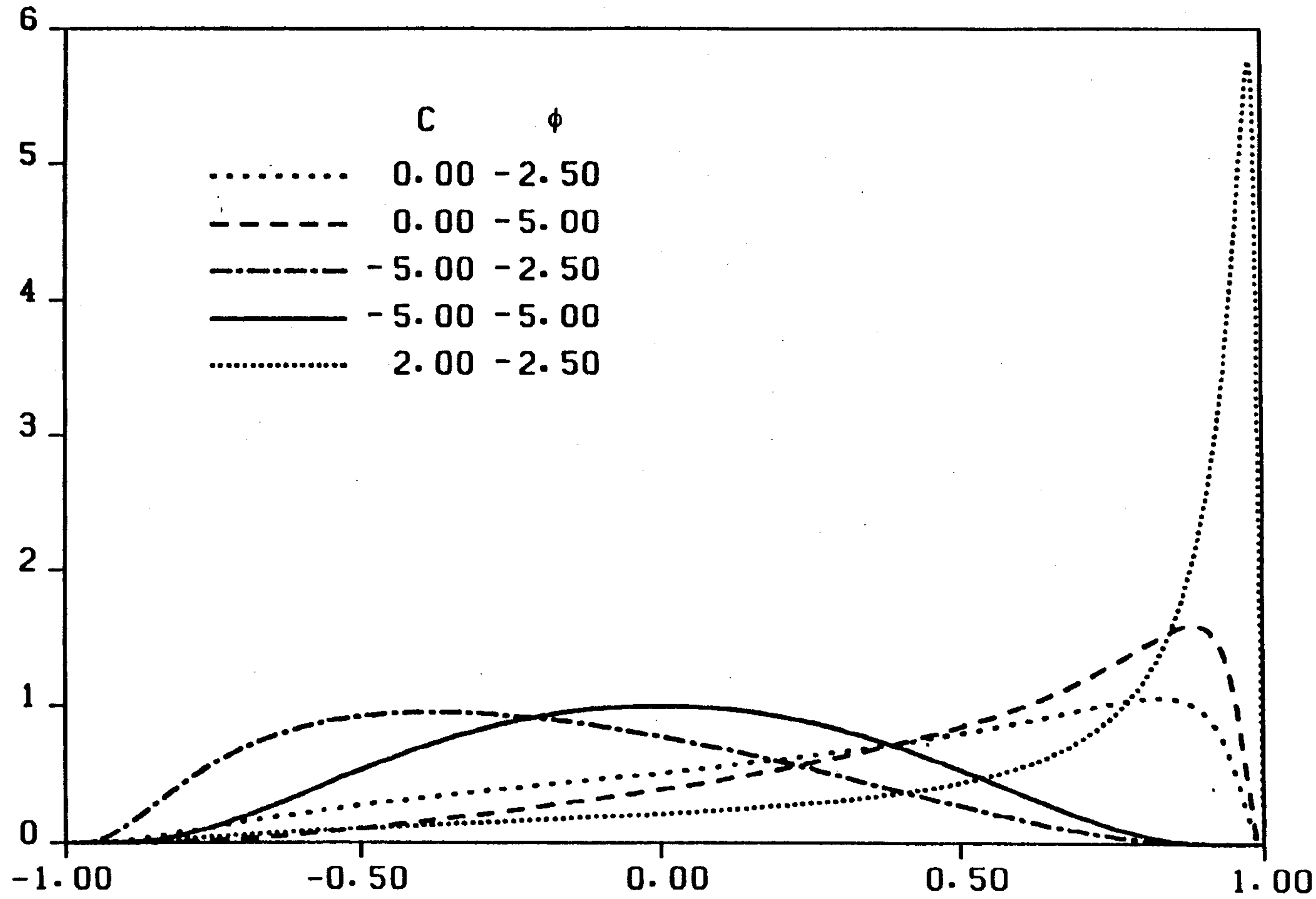


Figure 8: Limiting density of $\hat{\alpha}$; nearly seasonally integrated model.

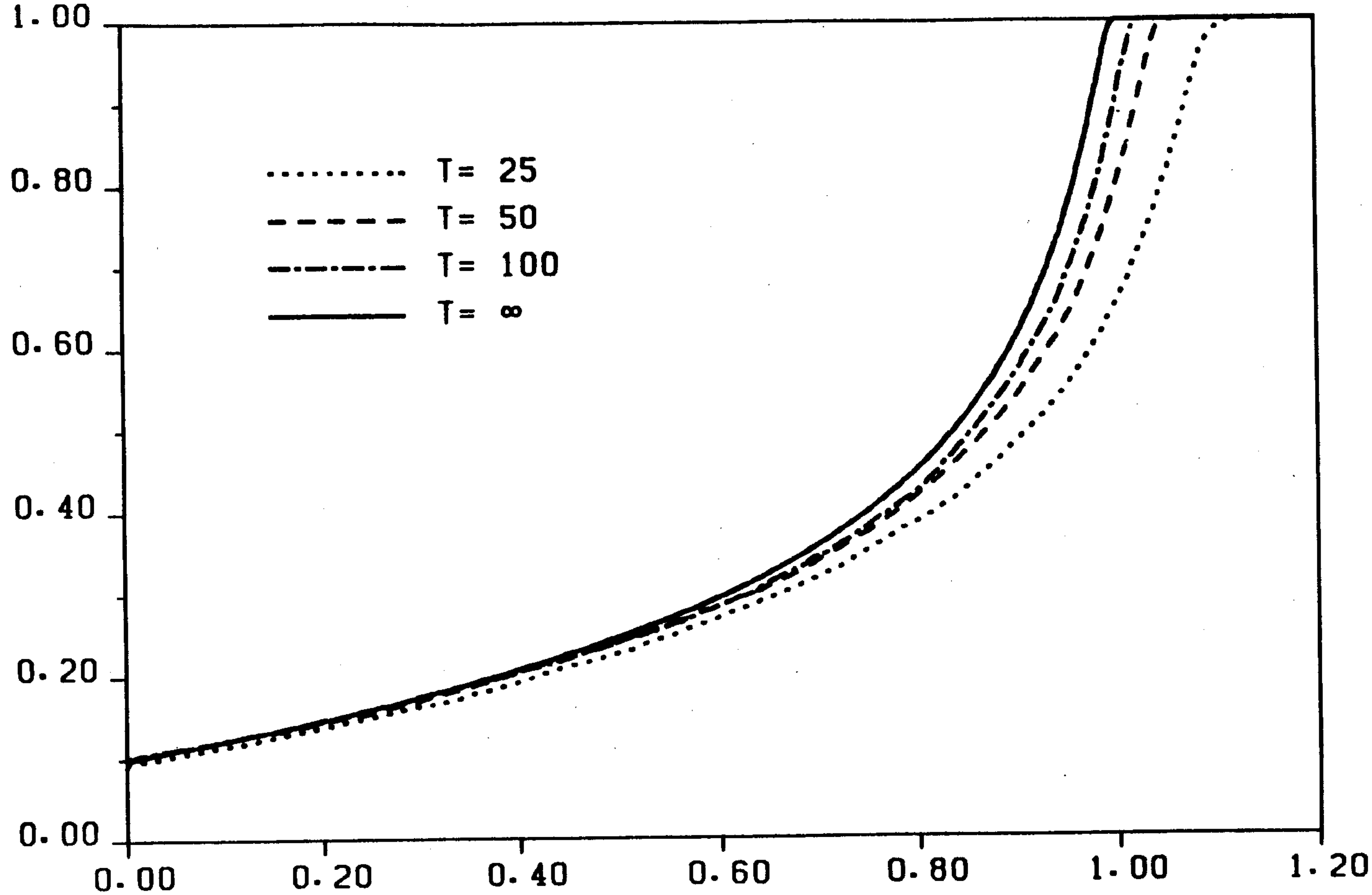


Figure 9: Asymptotic distribution of $\hat{\alpha}$ in the nearly seasonally integrated model and the corresponding finite sample distributions; $c = 2.0$, $\phi = -2.5$.