CAHIER 9424

THE ADEQUACY OF ASYMPTOTIC APPROXIMATIONS IN THE NEAR-INTEGRATED AUTOREGRESSIVE MODEL WITH DEPENDENT ERRORS

Pierre PERRON[1]

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

December 1994

C.P. 6128, succursale Centre-ville   Télécopieur (FAX) : (514) 343-5831
Montréal (Québec) H3C 3J7   Courrier électronique (E-Mail) : econo@tornade.ERE.Umontreal.CA

RÉSUMÉ

On considère l'estimateur des moindres carrés normalisé du paramètre dans un modèle autorégressif quasi-intégré avec erreurs dépendantes. Dans un premier temps, on analyse sa distribution asymptotique ainsi que son expansion asymptotique jusqu'à un ordre de $O_p(T^{-1})$. La fonction génératrice de moment limite est dérivée, ce qui nous permet de calculer différentes quantités de la distribution par intégration numérique. Des simulations sont utilisées pour vérifier la fiabilité de la distribution asymptotique lorsque les erreurs sont correlées. On s'attarde à deux cas particuliers : erreurs MA(1) et erreurs AR(1). Les approximations asymptotiques s'avèrent être inadéquates lorsque la racine MA s'approche de moins un et lorsque la racine AR s'approche de moins un et de un. Cette analyse théorique permet d'expliquer et de comprendre les résultats de simulations de Schwert (1989) et de DeJong, Nankervis, Savin et Whiteman (1992) concernant le niveau et la puissance du test de Phillips et Perron (1988). Un article associé, Nabeya et Perron (1994), présente des structures asymptotiques alternatives dans les cas où la distribution asymptotique habituelle ne donne pas une approximation adéquate de la distribution en échantillon fini.

Mots-clés : modèle quasi-intégré, convergence fonctionnelle, expérience de simulations, processus de racine unitaire, modèles autorégressifs à moyenne mobile.

ABSTRACT

We consider the normalized least squares estimator of the parameter in a nearly integrated first-order autoregressive model with dependent errors. In a first step, we consider its asymptotic distribution as well as asymptotic expansion up to order $O_p(T^{-1})$. We derive a limiting moment generating function which enables us to calculate various distributional quantities by numerical integration. A simulation study is performed to assess the adequacy of the asymptotic distribution when the errors are correlated. We focus our attention on two leading cases: MA(1) errors and AR(1) errors. The asymptotic approximations are shown to be inadequate as the MA root gets close to minus one and as the AR root approaches either minus one or one. Our theoretical analysis helps to explain and understand the simulation results of Schwert (1989) and DeJong, Nankervis, Savin and Whiteman (1992) concerning the size and power of Phillips and Perron's (1988) unit root test. A companion paper, Nabeya and Perron (1994), presents alternative asymptotic frameworks in the cases where the usual asymptotic distribution fails to provide an adequate approximation to the finite sample distribution.

Key words: near-integrated model, functional weak convergence, simulation experiment, unit root process, autoregressive moving-average models.

1. INTRODUCTION

In an attempt to cover more general time series structures, it has become popular in econometrics to consider models which permit both the regressors and the errors to have substantial heterogeneity and dependence over time. On a theoretical level, this advance has become possible due to a new class of functional central limit theorems which provide the required asymptotic results. An integrated treatment can be found in White (1984) and Gallant and White (1988). This approach made possible the analysis of a wide class of models with substantial relaxation of the standard conditions. Examples include time series models with unit roots (e.g.,Phillips (1987a)), testing for structural change in a general nonlinear framework (e.g., Andrews and Fair (1988)), and cointegration (e.g., Phillips and Ouliaris (1990)). However, little is known about the adequacy of the limiting distributions as an approximation to the finite sample distribution in such a general framework. This paper is a step in a systematic examination of this issue. We consider the leading case of a dynamic first-order autoregressive model when the errors are allowed to be dependent and provide a detailed analysis of the behavior of the associated ordinary least squares estimator. To be more precise, we consider the first-order stochastic difference equation:

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

with y_0 fixed. The least-squares estimator of α based on the sequence $\{y_t\}_0^T$ is:

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

The distribution of $\hat{\alpha}$ has been extensively studied, especially in the case where the errors $\{u_t\}$ are uncorrelated. Mann and Wald (1943) and Rubin (1950) showed that $T^{1/2}(\hat{\alpha} - \alpha)(1 - \alpha^2)^{-1/2}$ has a limiting $N(0,1)$ distribution when $|\alpha| < 1$. White (1958) showed that when $|\alpha| > 1$, the limiting distribution of $|\alpha|^T(\alpha^2 - 1)^{-1} (\hat{\alpha} - \alpha)$ is Cauchy provided that $y_0 = 0$ and $\{u_t\}$ is normal. White also considered the cased $|\alpha| = 1$ and showed that the limiting distribution of $T(\hat{\alpha} - 1)$ can be expressed in terms of the ratio of two functionals of a Wiener process (see also Phillips (1987a)). The case of the unit root ($\alpha = 1$) has attracted a great deal of attention. The asymptotic distribution of $T(\hat{\alpha} - 1)$ has been tabulated by Dickey (1976) via simulation methods (see also Fuller (1976) and Dickey and Fuller (1979)) and by Evans and Savin (1981a) using numerical integration. Evans and Savin (1981b) showed how the standard limiting distributions fail to provide adequate approximations to the exact distribution when α is close to but not equal to one.

Recently, a class of models which specifically deal with the presence of a root close to, but not necessarily equal to one, has been studied. These specify a near–integrated process where the autoregressive parameter is defined by $\alpha = \exp(c/T)$. The constant c being a measure of the deviation from the unit root case, the model 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. An expression for the limiting distribution of $T(\hat\alpha - \alpha)$ has been derived by Bobkoski (1983), Cavanagh (1986), Chan and Wei (1987) and Phillips (1987b) [1]. In the near–integrated context, with errors that are weakly dependent, Phillips (1987b) showed that (under some conditions to be made precise later) :

$$T(\hat\alpha - \alpha) \Longrightarrow \left\{ \int_0^1 J_c(r)dW(r) + \delta \right\} \left\{ \int_0^1 J_c(r)^2 dr \right\}^{-1} , \tag{1.3}$$

with $\delta = (\sigma^2 - \sigma_u^2)/(2\sigma^2)$, $\sigma^2 = \lim_{T\to\infty} E(T^{-1}S_T^2)$, $S_t = \Sigma_{j=1}^t u_j$, $\sigma_u^2 = \lim_{T\to\infty} T^{-1}E(\Sigma_{t=1}^T u_t^2)$, $J_c(r) = \int_0^r \exp((r - s)c)dW(r)$ with $W(r)$ the unit Wiener process.

Tabulations of the limiting distribution (1.3) with $\delta = 0$ have been obtained by Cavanagh (1986), Chan (1988), Nabeya and Tanaka (1990) and Perron (1989) using different procedures. These studies also assess its adequacy as an approximation to the finite sample distribution when α is near 1. They show the approximation to be quite good in the case where $y_0 = 0$. Perron (1991a,b) also considers a continuous time approximation which performs well even in the case where $y_0 \neq 0$. These asymptotic distributions provide substantial improvements over the traditional asymptotic framework when α is in the vicinity of one because they are continuous with respect to the autoregressive parameter α.

However, most of the evidence about the adequacy of the approximation pertains to the case $\delta = 0$, i.e with no correlation in the residuals [2]. The purpose of this paper is to investigate this adequacy when the errors are dependent. Section 2 discusses the limiting distribution of $T(\hat\alpha - \alpha)$ and extends Phillips' (1987c) $O_p(T^{-1})$ expansion to the near–integrated setting. The results of Perron (1991a) are used to derive appropriate limiting moment–generating functions. These permit the calculation of the cumulative distribution functions and the moments of the asymptotic distributions. In Section 3, we present simulation experiments to compare the asymptotic results with their finite sample counterparts. We concentrate on two leading cases, namely:

MA(1) errors : $\qquad u_t = e_t + \theta e_{t-1}$, $\tag{1.4}$

$$\text{AR(1) errors:} \qquad u_t = \rho u_{t-1} + e_t, \tag{1.5}$$

with $e_t \sim$ i.i.d. $N(0, \sigma_e^2)$ random variables. The results show the asymptotic distribution to be a very poor guide to the finite sample distribution, even for large sample sizes, when θ (in the MA case) or ρ (in the AR case) are close to -1. When ρ is close to 1, the approximation is not as bad but the approach to the limiting distribution is slow. Section 4 presents results for the $O_p(1)$ asymptotic approximation in the cases where a constant or a constant and a trend are included in regression (1.1). Section 5 provides some concluding comments and an appendix contains mathematical derivations.

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

This section considers the limiting distribution of the normalized least–squares estimator $T(\hat{\alpha} - \alpha)$ in the near–integrated model with dependent errors. We also consider its asymptotic expansion up to order $O_p(T^{-1})$. As a matter of notation, we denote weak convergence in distribution by '$\Rightarrow$' and equality in distribution by '$\overset{d}{=}$'. The asymptotic analyses below can be obtained under various conditions upon the error structure. We consider those of 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. When considering the asymptotic expansion of order $O_p(T^{-1})$ the following additional restriction will be imposed on the sequence of errors $\{u_t\}$.

ASSUMPTION 2: $\{u_t\}_1^\infty$ *is a Gaussian weakly stationary sequence.*

Consider the construction of random elements defined by: $X_T(r) = T^{-1/2}\sigma^{-1}S_{[Tr]} = T^{-1/2}\sigma^{-1}S_{j-1}$ $[(j-1)/T \le r < j/T; (j = 1,...,T)]$. From Herrndorf (1984), the following functional central limit theorem holds under Assumption 1:

$$X_T(r) \Rightarrow W(r). \tag{2.1}$$

Phillips (1987b) proved (1.3) under the conditions of Assumption 1 using the result (2.1). We consider here an extension of his (1987c) result concerning the asymptotic expansion up to order $O_p(T^{-1})$ of $T(\hat{a} - a)$. The following Theorem is proved in the Appendix.

THEOREM 1: Let $J_c(\tau) = \int_0^\tau exp((\tau-s)c)dW(\tau)$, $\delta = (\sigma^2 - \sigma_w^2)/(2\sigma^2)$ with $\sigma_u^2 = \lim_{T\to\infty} T^{-1}E(\Sigma_{t=1}^T u_t^2)$ and σ^2 as defined in Assumption 1; $\gamma = y_0/(\sigma T^{1/2})$, $\nu^2 = 2\pi f_{u^2}(0)$, with $f_{u^2}(0)$ the non-normalized spectral density function of $\{u_t^2 - E(u_t^2)\}$ at frequency 0; η is a $N(0,1)$ variable independent of the Wiener process $W(\tau)$. Under Assumptions 1 and 2:

$$T(\hat{a} - a) \stackrel{d}{=} H(c,\gamma)/K(c,\gamma) + O_p(T^{-1}),$$

where
$$H(c,\gamma) = \int_0^1 J_c(\tau)dW(\tau) + \delta + \gamma\int_0^1 exp(c\tau)dW(\tau) - (\nu/(2\sigma^2 T^{1/2}))\eta,$$

and
$$K(c,\gamma) = \int_0^1 J_c(\tau)^2 d\tau + 2\gamma\int_0^1 exp(c\tau)J_c(\tau)d\tau.$$

With the normalization $\sigma_e^2 = 1$, the following specifications of the various parameters apply. For the MA(1) case: $\sigma_u^2 = 1 + \theta^2$, $\sigma^2 = (1 + \theta)^2$, hence $\delta = \theta/(1 + \theta)^2$ and $\gamma = y_0$ $/[T^{1/2}(1 + \theta)]$. Also, $\nu^2 = 2(1 + 4\theta^2 + \theta^4)$. In the AR(1) case we have: $\sigma_u^2 = (1 - \rho^2)^{-1}$, $\sigma^2 = (1 - \rho)^{-2}$; hence $\delta = \rho/(1 + \rho)$, $\gamma = y_0(1 - \rho)/T^{1/2}$ and $\nu^2 = 2(1 + \rho^2)/(1 - \rho^2)^3$. We now characterize the joint moment-generating function of $\{H(c,\gamma), K(c,\gamma)\}$, which we denote by $MGF(v,u) = E[\exp\{vH(c,\gamma) + uK(c,\gamma)\}]$.

THEOREM 2: Let $g = \nu/(2\sigma^2 T^{1/2})$, $d = \gamma^2(exp(2c) - 1)/2c$ and $\lambda = (c^2 + 2cv - 2u)^{1/2}$ the joint moment-generating function of $\{H(c,\gamma), K(c,\gamma)\}$ is given by:

$$MGF(v,u) = exp(v\delta - ud + v^2g^2/2) M_{c,\gamma}(v,u),$$

where
$$M_{c,\gamma}(v,u) = \psi_c(v,u)exp\{- (\gamma^2/2)(v + c - \lambda)(1 - exp(v + c + \lambda)\psi_c^2(v,u))\},$$

and
$$\psi_c(v,u) = \left\{2\lambda exp(- (v + c))/[(\lambda + v + c)exp(-\lambda)+(\lambda - v - c)exp(\lambda)]\right\}^{1/2}.$$

Theorem 2 allows direct computation, by numerical integration, of the cumulative distribution function and the moments of the asymptotic distribution. Let the joint characteristic function of $\{H(c,\gamma), K(c,\gamma)\}$ be $CF(v,u) = MGF(iv,iu) = E[\exp\{ivH(c,\gamma) + iuK(c,\gamma)\}]$. From Gurland (1948), the limiting distribution function of $T(\hat{a} - a)$ is:

$$F(z) = (1/2) - (1/\pi) \int_0^\infty IM[CF(v,-vz)/v]dv, \tag{2.2}$$

with $IM(\cdot)$ the imaginary part of the complex number. The moments of the asymptotic distribution can be obtained using Mehta and Swamy's (1978) result, such that:

$$E[H(c,\gamma)/K(c,\gamma)]^r = \Gamma(r)^{-1} \int_0^\infty u^{r-1} \left\{ \partial^r MGF(v,-u)/\partial v^r \right\}_{v=0} du . \tag{2.3}$$

These results allow computation of distributional quantities for a variety of processes. The usual asymptotic distribution (1.3) can be obtained by letting $g = \gamma = 0$, then:

$$MGF(v,u) = \exp(v\delta)\psi_c(v,u) , \tag{2.4}$$

where $\psi_c(v,u)$ is defined in Theorem 2. The next section makes extensive use of Theorem 2, (2.2) and especially (2.4), to calculate the asymptotic distribution of $T(\hat{\alpha} - \alpha)$ when the errors are $MA(1)$ or $AR(1)$ processes. The numerical integrations were performed using the subroutine QDAG of the IMSL library. The bounds of integration are $(\epsilon, \bar{V})$ where $\bar{V}$ is set such that the square of the integrand evaluated at $\bar{V}$ is less than ϵ. The error criterion for the numerical integration was also set at ϵ. For most experiments we used $\epsilon = 1.0E-07$ [3].

3. APPROXIMATING THE EXACT DISTRIBUTION OF $T(\hat{\alpha} - \alpha)$

This section considers the adequacy of the asymptotic distribution as an approximation to the finite sample distribution. The asymptotic values are obtained using the results of Section 2 and the finite sample values are obtained via simulations. As discussed in the introduction, we consider two leading cases where the error sequence is either an $MA(1)$ or an $AR(1)$ process. As we will see, these cases are sufficient to provide a rich characterization of the relationship between the finite sample results and their asymptotic counterparts. Given that our aim is mainly oriented towards studying the effect of correlation in the errors we consider only the case where $y_0 = 0$.

The experiment is as follows. For each of the MA and AR cases we consider three values of c, namely $c = 0$ (unit root), $- 5$ (locally stationary) and 2 (locally explosive). For each we consider the following specifications for the errors [4]: a) $MA(1)$ case, $\theta = - .9, -.7, -.5, -.3$ and $.5$; for the $AR(1)$ case, $\rho = -.9, -.5, .5, .9$ and $.95$. We consider sample sizes of length $T = 25, 50, 100, 500, 1000$ and 5000 (though we do not report results for every sample sizes in each cases). The finite sample results are obtained using $10,000$ replications

(5,000 when T = 5,000). The asymptotic results for the cumulative distribution function are obtained using (2.4) for the $O_p(1)$ asymptotic framework [5].

3.A. The MA(1) Case.

Consider first the percentage points of the distribution of $T(\hat{a} - a)$ in the MA(1) case presented in Tables 1 for $c = 0$. With $\theta = .5$, the asymptotic approximation is excellent, even for small sample sizes. Further experiments (not reported) confirm this adequacy for any process with positively autocorrelated MA(1) errors.

When θ is negative the picture is rather different. Here the adequacy of the approximation depends very much upon the magnitude of θ and deteriorates quite rapidly as θ approaches -1. For different values of θ the asymptotic distribution adequately approximates the finite sample distribution for the following sample sizes: $\theta = -.3$, T $\geq$ 500; $\theta = -.5$, T $\geq$ 1000 ; $\theta = -.7$, T $\geq$ 5000; and when $\theta = -.9$ the asymptotic distribution is still quite far away from the exact distribution when T = 5000. The differences are substantial. For example, the 1% point of the exact distribution when $\theta = -.9$ and T = 50 corresponds approximately to the 95% point of the asymptotic distribution.

The same qualitative results hold when c is -5 or 2. Figures 1 and 2 present results for the case $\theta = -.9$. The main difference is that the approximation is marginally better with an explosive process ($c = 2$) and marginally worse when it is stationary ($c = -5$) [6].

3.B. The AR(1) Case.

Consider now the case where the errors are AR(1). The results for $c = 0$ are presented in Table 2. The picture is quite different from the MA(1) case. For $\rho < 0$, the approximation is again inadequate and worsens as ρ approaches $- 1$. However, for comparable values of ρ and θ the approach of the finite sample distribution to the asymptotic distribution is faster in the AR(1) case than it is in the MA(1) case. For $\rho = -.5$, the asymptotic approximation is adequate for T $\geq$ 500, and for $\rho = -.9$ when T $\geq$ 1000. For smaller sample sizes there are important discrepancies especially when $\rho = -.9$. However, the discrepancies are not as severe as in the MA(1) case. For example, when T = 50 and $\rho = -.9$, the 1% point of the exact distribution corresponds roughly to the 30% point of the asymptotic distribution (in the MA(1) case the first percentage point of the exact distribution with $\theta = -.9$ and T = 50 corresponds to the 90% point of the asymptotic distribution). Similar qualitative results hold in the stationary ($c = -5$) and explosive ($c =$

2) cases. When $c = -5$, the adequacy is marginally inferior and when $c = 2$, it is marginally superior. These results are illustrated in Figures 3 ($c = -5$) and 4 ($c = 2$).

The AR(1) case with positive autocorrelation offers yet a different picture. First, unlike all the cases considered so far the approach of the finite sample distribution to its asymptotic counterpart is from a density with a larger spread to one with a smaller spread. Secondly, the differences between the finite sample and asymptotic percentage points are not substantial. For example when $\rho = .9$ and $T = 25$, the tenth percentage point of the exact distribution corresponds roughly to the fifth percentage point of the asymptotic distribution. Nevertheless, what is interesting, and different from the MA(1) case with positive coefficient, is the fact that the approach to the asymptotic value is quite slow. The tail of the exact distribution is not well approximated until $T = 500$.

When $c = -5$ and $c = 2$, the same qualitative behavior remains as in the case where c $= 0$. The case $\rho = .9$ is illustrated in Figures 5 ($c = -5$) and 6 ($c = 2$). The approximation is again marginally better with $c = 2$ and marginally worse with $c = -5$. However, some differences emerge. When $c = -5$ we notice a difference from the cases where $\rho > 0$ and c $= 0$ or 2. The right tail is better approximated by the asymptotic distribution even for quite small sample sizes (e.g. $\rho = .9$ and $T = 25$). Secondly, when compared to the positive MA case, there is much more movement in the percentage points as c varies.

The results above suggest that the quality of the asymptotic approximation is heavily dependent on the nature and extent of the correlation in the residuals. With negative autocorrelation, the approximation becomes rapidly useless as the magnitude of this correlation increases, both in the MA(1) and AR(1) cases. In the AR(1) case with positive autocorrelation, the discrepancies are not as severe but the approach to the asymptotic distribution remains quite slow. Only in the MA(1) case with positive autocorrelation is the approximation adequate, indeed as good as in the case without correlation.

3.C. The Moments of the Distributions.

We also performed an extensive analysis of the behavior of the mean and variance of the exact and asymptotic distributions. For reason of space constraint, we report only the figures for values of c and θ or ρ presented in Tables 1 and 2. The asymptotic results were obtained using numerical integration of the function in (2.3) [7].

Consider first the mean of the distribution when $c = 0$. For the MA case and $\theta \geq 0$, the

the asymptotic value is basically equal to the exact value when $T \geq 100$. When $T < 100$, the discrepancies are minor. When $\theta < 0$, the picture is different. As θ approaches minus one, it takes an increasingly larger sample size for the mean of the exact distribution to correspond to its asymptotic value. When $\theta = -.1$, a sample of size 100 still appears enough but there are larger discrepancies with smaller sample sizes compared to the case where $\theta \geq 0$. When θ is between $-.2$ and $-.5$, a sample of size 500 is needed to ensure a satisfactory approximation. When $\theta = -.6$, a sample of size 1000 is needed and when $\theta = -.7$ or $-.8$, the corresponding figure is $T = 5000$. Finally, when $\theta = -.9$ or $-.95$ even a sample size as large as 5,000 is not sufficient to provide an adequate approximation. Of particular interest is the fact that when $\theta \leq -.5$, the mean of the finite sample distribution changes very rapidly as T increases. Hence, for this part of the parameter space, the asymptotic distribution provides a very bad approximation to the mean of the exact distribution when the sample size is not very large.

Consider now the AR(1) case. When $\rho \leq 0$, the picture is similar to that of the MA case, except that the discrepancies between the mean of the asymptotic and exact distributions are not as severe. When ρ is between $-.1$ and $-.5$, the exact mean attains its asymptotic value when T reaches somewhere between $T = 100$ and $T = 500$. When ρ is between $-.6$ and $-.95$, the correspondence is achieved with a sample size somewhere between $T = 500$ and $T = 1000$. When $\rho > 0$, the picture is different. With ρ between .1 and .5, a sample size as small as 50 is enough to provide an adequate approximation. When $\rho = .7$, a sample of size 100 is needed, and with $\rho = .9$ or .95, the corresponding figure is $T = 500$. Also of interest is the fact that when ρ is between .5 and .95, the mean is positive (unlike all the other cases considered). Note that, in all cases considered, the exact mean approaches its asymptotic counterpart in a monotonically decreasing way [8].

With $c = -5$, a locally stationary process, the asymptotic approximation is, in general, less good than with $c = 0$ for both the MA and AR cases. That is, for a given θ or ρ and sample size T, the discrepancy between the exact mean and its asymptotic counterpart is greater. A feature that is different, however, is the fact that when θ or ρ are greater than .3, the approach to the asymptotic value is achieved in a monotonically increasing way. When $c = 2$, the locally explosive case, the general features are similar but the discrepancies between the exact and asymptotic results are not as severe, compared to both cases where $c = 0$ or -5. As was the case with $c = 0$, the approach to the asymptotic value is monotonically increasing with T. A small difference from earlier cases is that, unless T is small and $\rho = .9$ or .95, the mean of the distribution is negative.

Tables 1 and 2 also present results for the variance of the distribution. For the case of a unit root process, $c = 0$ with MA(1) errors (Table 1), the general features are similar to the ones for the mean of the distribution. An interesting difference is that when $\theta < 0$ the approach to the asymptotic value appears to be slower. For example, when $\theta = -.7$ and T $= 1000$, the mean of the exact distribution is quite close to the mean of the asymptotic distribution, but the variance is still quite far away. Of particular interest is the fact that, for a given θ, the exact variance approaches its asymptotic value in a monotonically increasing way. However, for a given sample size, the variance does not increase monotonically as θ approaches -1. For example, with T $= 50$ the variance of the distribution of $T(\hat{\alpha} - \alpha)$ is 66.38 ($\theta = -.5$), 142.37 ($\theta = -.7$) and 119.64 ($\theta = -.9$) [9].

Much of the same features apply in the AR case (Table 2); in particular the discrepancies between the exact and asymptotic values are not as severe as for the MA(1) case. Some interesting features are, however, different from the MA(1) case. First, for a given sample size, the variance is monotonically increasing as ρ approaches -1. More importantly, when $\rho \geq .3$ the exact variance approaches its asymptotic value in a monotonically decreasing way unlike the case with negatively correlated MA(1) or AR(1) errors. This feature will prove of interest when considering the asymptotic expansions. The behavior of the variance when $c = -5$ or 2 is similar to that when $c = 0$. As was the case for the mean, the discrepancies are more severe with $c = -5$ and less so with $c = 2$. Apart from this, a notable difference is that in the MA(1) case with $\theta \geq .1$ and $c = 2$, the exact variance now decreases monotonically towards its asymptotic value.

3. D. The $O_p(T^{-1})$ Expansion.

Note first that the use of the $O_p(T^{-1})$ asymptotic expansion does not provide any improvement over the usual $O_p(1)$ asymptotic distribution when considering the mean of the distribution. Indeed both of them yield the same mean since the $O_p(T^{-1})$ expansion does not provide any location adjustment given that the extra term η has mean zero and is independent of the Wiener process $W(r)$ present in the other components.

The $O_p(T^{-1})$ asymptotic expansion provides, in contrast, an adjustment to the variance of the asymptotic distribution. However, given the independence of η and the Wiener process $W(r)$, the $O_p(T^{-1})$ expansion yields a *higher* variance than the $O_p(1)$ asymptotic distribution. As we saw, for most cases, the exact variance is *smaller* than the asymptotic variance. In these cases, the asymptotic expansion should provide a *less*

accurate approximation to the exact variance than does the $O_p(1)$ asymptotic distribution. Since the asymptotic expansion does not provide, in most cases, an adjustment to the mean of the asymptotic distribution, it follows that it should provide a *poorer* approximation to the exact distribution as a whole. An instance where it could yield a better approximation to the exact variance is when the variance approaches its asymptotic value in a decreasing way. This occurs with AR(1) errors and $\rho \geq 0.3$. Even in those cases, the expansion would provide a better approximation overall only if the mean is fairly stable as T increases. In our experimental setting, this is the case only when ρ is near 0.5 and $c = 0$ or 2.

We performed many calculations related to the distribution of this asymptotic expansion. As an illustration of the qualitative results, we present in Tables 1 and 2 numerical values for $T = 50$ (other results are available upon request). Overall, the $O_p(T^{-1})$ expansion provides a worse approximation than the standard $O_p(1)$ asymptotic distribution. There are, however, several interesting features that emerge.

Consider the case of MA(1) errors. The first striking feature is that the asymptotic distribution provides an improvement in the right tail of the distribution but a worse approximation in the left tail. This asymmetric feature is interesting because it is indeed in the left tail that the asymptotic distribution is so inaccurate for negative moving–average coefficients. The expansion can therefore be viewed, in a sense, as providing a useful correction where it is easy to do so, i.e. where the discrepancies between the exact and standard asymptotic distributions are least severe. Other noticeable features include the fact that the improvement is better as we move further in the right tail and worsens as the moving–average coefficient varies from positive to negative.

Consider now the case of AR(1) errors. When $\rho = -.5$, we observe the same pattern as with negative MA(1) errors; the expansion provides a worse approximation in the left tail but is surprisingly accurate in the right tail. When $\rho = -.9$, the approximation is worse in the left tail and better for the median. However, we now observe a substantial overcorrection in the right tail such that the expansion is again a worse approximation than the standard asymptotic distribution.

As discussed above, the expansion is likely to provide a better approximation to the overall distribution in the case of AR(1) errors with positive coefficient. We now discuss this case in more details. With $\rho = .5$, the asymptotic expansion provides a substantial improvement in the left tail of the distribution. For example, the first percentage point

with $T = 50$ has an exact value of -4.47; the $O_p(1)$ asymptotic value is -4.23 while the $O_p(T^{-1})$ value is -4.53. When ρ is closer to 1, the asymptotic expansion still provides an improvement in the left tail of the distribution but not as significant. This is due to the fact that the $O_p(T^{-1})$ expansion provides no adjustment to the mean of the distribution and as ρ approaches 1 the mean of the distribution decreases substantially as T increases. These observations are consistent with the fact that the $O_p(T^{-1})$ expansion also provides an adequate adjustment to the variance of the distribution when $\rho = 0.5$, but that it deteriorates rapidly as ρ approaches 1. The results concerning the right tail of the distribution are quite different. Here, the $O_p(T^{-1})$ expansion provides no significant improvement over the usual $O_p(1)$ asymptotic distribution for any value of ρ considered.

The results are similar when $c = -5$, namely an improvement in the left tail of the distribution especially when $\rho = .5$. The approximation for the variance is better when $\rho = .5$ but the improvement diminishes as ρ gets closer to one. With $c = 2$ and $\rho = .5$, the expansion provides a slight improvement for the 1% and 2.5 % points, but a worse approximation for the 5% and 10% points. The improvements in the right tail of the distribution and in the variance are marginal. When $\rho = .9$ or $.95$, there is very little improvement; indeed there is little change in the $O_p(T^{-1})$ distribution as T changes.

3. E. A Possible Explanation Related to Further Research.

Some insights about the results described above can be gained by looking at the behavior of the parameter δ in the asymptotic distribution given in Theorem 1. In the MA case $\delta = \theta/(1 + \theta)^2$ which becomes unbounded when θ approaches -1, and decreases to 0 as θ increases. In the AR(1) case $\delta = \rho/(1 + \rho)$ which diverges to $-\infty$ as ρ approaches -1 but at a smaller rate than in the MA(1) case. When ρ approaches 1, δ approaches $1/2$.

These considerations lead to the following conjectures. First, the asymptotic distribution is a bad approximation when θ approaches -1 because the asymptotic distribution of $T(\hat{\alpha} - \alpha)$ with θ approaching -1 is unbounded. However, when θ approaches 1 the asymptotic distribution is still valid. In the AR(1) case, the asymptotic distribution of $T(\hat{\alpha} - \alpha)$ is again unbounded if $\rho \to -1$ as $T \to \infty$. However, the rate at which ρ may approach -1 to obtain a non–degenerate local asymptotic distribution is higher in the AR(1) case than it is in the MA(1) case. This would explain the relatively smaller discrepancies in the AR(1) case for a given equal value of θ and ρ. On the other hand, when

ρ approaches $+ 1$ as $T \to \infty$, $T(\hat{\alpha} - \alpha)$ still has a non–degenerate asymptotic distribution but different from that given by (1.3). These conjectures are verified in a companion paper, Nabeya and Perron (1994), where the relevant local asymptotic distributions are derived and tabulated. It is shown that these local asymptotic distributions provide very good approximations to the finite sample distributions. Their characterization also helps in further understanding the reasons for the failures of the usual asymptotic distribution.

4. EXTENSIONS TO CASES WITH FITTED INTERCEPT AND TREND.

In this section we consider extensions to cases where an intercept or an intercept and a trend are included in the first–order autoregression. Given the nature of the results discussed for the leading case in the previous sections, we concentrate on the quality of the standard $O_p(1)$ approximation (the asymptotic expansions and the moments of the distribution are not covered). To be more specific consider the following regressions:

$$y_t = \mu + \alpha y_{t-1} + u_t, \tag{4.1}$$

$$y_t = \mu + \beta t + \alpha y_{t-1} + u_t. \tag{4.2}$$

The least–squares estimator of α in (4.1) is denoted $\hat{\alpha}_2$ and the one from (4.2) by $\hat{\alpha}_3$. We study the behavior of the statistics $T(\hat{\alpha}_i - \alpha)$ $(i = 2, 3)$ when the series $\{y_t\}$ is generated by (1.1) with $\alpha = \exp(c/T)$ and $y_0 = 0$. The cumulative distribution functions of the limiting distributions of $T(\hat{\alpha}_i - \alpha)$ can be obtained from results in Nabeya and Tanaka (1990) (see also Nabeya and Sørensen (1993)) and are stated in the following Lemma.

LEMMA 1: *Let* $F_i(z) = \lim_{T \to \infty} P[T(\hat{\alpha}_i - \alpha) \leq z]$ $(i = 2, 3)$ *where* $\hat{\alpha}_2$ *and* $\hat{\alpha}_3$ *are the OLS estimates of* α *in* (4.1) *and* (4.2), *respectively. Assume that* $\{y_t\}_0^T$ *is generated by* (1.1) *with* $\alpha = exp(c/T)$ *and let* δ *be defined as in Theorem 1. We have:*

$$F_i(z) = (1/2) + (1/\pi) \int_0^\infty (1/v) IM(cf_i(v, z)) dv \qquad (i = 2, 3)$$

where $cf_i(v, z) = exp(i(1 - 2\delta)v - c/2)[D_i(2iv, z)]^{-1/2}$ *with*

$$D_2(\lambda, z) = c^2 \cosh(\mu)/\mu^2 + ((1 - \cosh(\mu))/\mu^4)(2c^2\lambda - 4c\lambda(z + c) - 2\lambda^2)$$
$$+ (\sinh(\mu)/\mu^3)(-c^3 + c^2\lambda - 2\lambda(z + c) - \lambda^2)$$

$$D_g(\lambda, z) = (cosh(\mu)/\mu^6)\{c^6 - 12c^4\lambda + 2\lambda(z+c)(-c^4 + 12c^3 - 12c^2) + 48\lambda^2(c^2 - 3c + 3)\}$$

$$+ (1 - cosh(\mu)/\mu^8)\{4\lambda(-c^6 + 6c^4) + 16\lambda^2(c^4 - 3c^3 - 3c^2 + 18c - 18)$$

$$+ 8\lambda(z + c)(c^5 - 3c^4 + c^4\lambda - 4\lambda^2(c^2 - 3c + 3))$$

$$+ 16\lambda^2(z + c)^2(-c^3 + 3c^2 - 6c + 6)\}$$

$$+ (sinh(\mu)/\mu^7)\{-c^7 + \lambda(c^6 + 24c^4) + 4\lambda^2(-c^4 + 3c^3 - 27c^2 + 72c - 72)$$

$$+ 2\lambda(z + c)(c^5 - 4c^4 - 12c^3 + 12c^2 - c^4\lambda + 4\lambda^2(c^2 - 3c + 3))$$

$$+ 16\lambda^2(z + c)^2(c^2 - 3c + 3)\},$$

where $\mu = (c^2 - 2\lambda(z + c))^{1/2}$.

Lemma 1 allows us to obtain percentage points of the limiting distributions using numerical integrations [10]. Results concerning the distribution of $T(\hat{\alpha}_2 - \alpha)$ where an intercept is included are presented for the leading case $c = 0$ in Table 3.a for MA(1) errors and in Table 3.b for AR(1) errors. We consider only a subset of the cases discussed earlier, namely $\theta = -.9, -.7$ and $-.5$ in the MA(1) case; and $\rho = -.9, -.5$ and $.9$ in the AR(1) case.

Consider first results pertaining to the MA(1) case in Table 3.a. The qualitative picture is similar to the case where no intercept is included in the regression. The quality of the asymptotic approximation deteriorates as the moving-average parameter θ approaches -1 and becomes inadequate even for very large sample sizes when θ is $-.9$. An interesting feature to note is that for small values of T (e.g. $T = 50, 100$) the differences in the exact critical values are small with and without an intercept included. A shift of the distribution to the left occurs as T increases and this shift is substantial for the asymptotic distribution. The fact that the leftward movements of the distributions are greater than in the no-intercept case as T increases implies larger discrepancies between the exact and asymptotic distributions. For example, when $T = 100$ and $\theta = -0.9$, the 1% point of the exact distribution corresponds roughly to the 99% point of the asymptotic distribution; when $T = 100$ and $\theta = -0.7$ the 5% point of the exact distribution corresponds roughly to the 50% point of the asymptotic distribution. When $\theta = -.5$, the leftward shifts in the distribution are relatively important for both the finite sample and asymptotic distributions. Nevertheless, the adequacy of the asymptotic approximation is worse compared to the case where no intercept is included.

The same qualitative results hold when $c = -5$ or 2. Figures 7 and 8 present results for the case $\theta = -.9$. Compared to the case where $c = 0$, the main difference is again that the approximation is marginally better with an explosive process ($c = 2$) and marginally worse when it is stationary ($c = -5$). Compared to the case with no intercept (see Figures 1 and 2), it is evident that for each T the adequacy of the asymptotic approximation is worse when an intercept is included for both the explosive and stationary cases.

Consider now results pertaining to the case where the errors are AR(1) presented in Table 3.b. For negative values of ρ, the qualitative results are similar to those in the MA(1) case. The adequacy of the asymptotic approximation is worse when an intercept is included. For example, with $\rho = -.9$ and $T = 100$, the 1% point of the exact distribution corresponds roughly to the 10% point of the asymptotic distribution when no intercept is included whereas if one is included the 1% point of the exact distribution corresponds roughly to the 30% point of the asymptotic distribution. For $\rho = -.5$, the asymptotic distribution could be considered an adequate approximation to the exact distribution for T greater than 100 in the no–intercept case. Now it is adequate only for T greater than 500. When $\rho = .9$ one first notes the leftward shift of both the exact and asymptotic distributions compared to the case where no intercept is included. Again the asymptotic distribution is a good approximation to the exact distribution when $T = 500$. It is, however, less accurate than in the no–intercept case for smaller values of T (e.g. $T = 50$).

Figures 9 and 10 present results for the case $c = -5$ with $\theta = -.9$ and .9. Compared to the case where $c = 0$, the main difference is again that the approximation is marginally worse. Compared to the case with no intercept (see Figures 3 and 5), it is evident that the adequacy of the asymptotic approximation is worse when an intercept is included.

The results concerning the distribution of $T(\hat{\alpha}_3 - \alpha)$ for the case where an intercept and a trend are included are first presented for the leading case $c = 0$ in Table 4.a for MA(1) errors and in Table 4.b for AR(1) errors. We consider the same subset of specifications as in the intercept only case.

The qualitative picture is again similar except that now the discrepancies between the exact and asymptotic distributions are even bigger for any sample size. Consider the following examples. First, when $\theta = -.9$ and $T = 1000$, the 1% point of the exact distribution corresponds roughly to the 90% point of the asymptotic distribution. Second, even when T is as large as 5000 the differences are substantial, viz. the 1% point of the

exact distribution corresponds roughly to the 20% point of the asymptotic distribution. The inaccuracies are not as severe when $\theta = -.7$ or $-.5$ but are nonetheles more important than in the case where no intercept or only an intercept is included in the regression.

These dramatic differences between the exact and asymptotic distributions are well captured in Figures 11 and 12 which present results for the case $\theta = -.9$ with $c = -5$ and 2, respectively. Comparing with Figures 1 and 2 (no intercept case) and Figures 7 and 8 (intercept included), one clearly sees a gradual deterioration in the quality of the approximation as the number of included deterministic regressors increases.

Consider now results pertaining to the case of AR(1) errors presented in Table 4.b for the leading case $c = 0$. For negative values of ρ, the qualitative results are similar to the those in the MA(1) case. The adequacy of the asymptotic approximation is worse when an intercept and a trend are included. For example, when $\rho = -.9$ and $T = 50$, the 1% point of the exact distribution corresponds roughly to the 50% point of the asymptotic distribution in the no trend case whereas if a trend is included it corresponds roughly to the 90% point. These discrepancies are well illustrated for the case $c = -5$ in Figure 13 (compare to Figures 3 and 9) which shows again the deterioration in the quality of the approximation as the number of deterministic components included in the autoregression increases. For positive values of the autoregressive coefficient (with results presented for $\rho = .9$), we observe, as in earlier cases, that the asymptotic distribution is a good approximation to the exact distribution for $T = 500$. However, we see that for smaller sample sizes, the quality of the approximation deteriorates as the number of included deterministic components increases, more so as T gets smaller. This feature is well illustrated in Figure 14 which considers the case $c = -5$ and $\rho = .9$ to be compared to Figures 5 and 10.

5. CONCLUDING COMMENTS

We characterized and tabulated the asymptotic distribution of the normalized least – squares estimator in a nearly integrated autoregressive process allowing dependence in the errors with emphasis on MA(1) and AR(1) processes. These were sufficient to provide a rich array of cases. Special attention was given to the adequacy of the standard asymptotic distribution as an approximation in finite samples. We showed that, in a substantial part of the parameter space, the approximation is seriously inadequate. An implication of our results is the inherent danger associated with an asymptotic framework that allows very general conditions with respect to the type of dependence permitted. There appears to be a

tradeoff between generality in the conditions and the sample size needed for the asymptotic distribution to provide a reasonable guide to the finite sample distribution. While this conclusion is drawn from a simple model, it should extend to more general models.

Our analysis can also be viewed as a step in a more complete analysis of the adequacy of asymptotic distributions of tests allowing general dependence and heterogeneity. Examples include the class of unit root tests proposed by Phillips (1987a) and Phillips and Perron (1988) as well as those involving the Newey–West (1987) correction in more general structural models. In practice, the statistics are used with correction factors that eliminate asymptotically the dependence of the asymptotic distributions upon nuisance parameters. Our analysis has direct implications for the behavior of such transformed statistics. Suppose that the correction factor adequately approximates the asymptotic correction necessary to eliminate the dependence upon the nuisance parameters. Our results would still imply an inadequate corrected statistic as the asymptotic distribution of the uncorrected part is far from the finite sample distribution in an important range of the parameter space. In a sense, the finite sample distribution would bear an adequate correspondence to the asymptotic distribution by fortuitous cancellation of approximation errors for the distribution of the original statistic and the correction factor. Such a situation is unlikely to occur as demonstrated in the case of unit root tests in the simulations of Schwert (1989) and Dejong, Nankervis, Savin and Whiteman (1992) and the analytical results of Pantula (1991). Their results about size distortions in finite samples can be explained by the poor approximation provided by the asymptotic distribution of

$T(\hat{\alpha} - 1)$ which forms the basis of the transformed statistics. These issues are analyzed analytically for the Phillips–Perron (1988) test and related ones in Ng and Perron (1993) using the local asymptotic framework laid out in Nabeya and Perron (1994).

FOOTNOTES

* This paper is a revised version of Sections 2, 3 and 4 of Perron (1990b). The other parts of that earlier draft are now the object of a companion paper, Nabeya and Perron (1994). Some material in the present paper were initially investigated as part of a joint project with Peter C.B. Phillips in the spring of 1986; I which to thank him for his help during this initial stage. I also wish to thank Seiji Nabeya for useful comments. 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.

[1] This framework has been quite useful in studying various problems such as the power of tests of a unit root under local alternatives (Phillips (1987b), Phillips and Perron (1988) and Perron (1990a)), the derivation of confidence intervals when a is near unity (Cavanagh (1986) and Stock (1991)) and the power of tests of a unit root with a continuum of observations (Perron (1989)).

[2] An exception is Nabeya and Tanaka (1990) which contains some results when the errors are correlated.

[3] In principle, the computation of the exact finite sample distribution is possible using Imhof's (1961) routine or a variant of it. However, given the sample sizes analyzed here (up to 5,000) such an approach becomes practically infeasible.

[4] A much wider range of experiments were performed. For the sake of brevity we report only a subset of the results. Some comments made in the text pertain to the full set of results, however. These are available upon request.

[5] Special care must be taken with the numerical integration since it involves the square root of a complex valued quantity. The use of the principal value may not ensure the continuity of the integrand. The numerical integration must be performed over Reimann surfaces consisting here of two planes. The method is described in more detail in Perron (1989).

[6] The full set of results is available upon request. Tabulated critical values for an extended range of values for θ are available in the working paper version.

[7] The specifications are basically the same as in the numerical integration of the distribution function, except that here the integrand does not involve complex valued quantities, so only straightforward numerical integration routines are needed.

[8] Some slight exceptions to this rule occur for large values of T because of sampling variability induced by the simulations.

[9] These features are further discussed in Nabeya and Perron (1994). In particular, the alternative asymptotic framework derived in that paper can explain them.

[10] These were performed in double precision FORTRAN using the subroutine DCADRE of the IMSL library. Again, the bounds are $(\epsilon, \bar{V})$ where $\bar{V}$ is such that the square of the integrand evaluated at $\bar{V}$ is less than ϵ. The error for the numerical integration was also set at ϵ. For most experiments we used $\epsilon = 1.0E{-}07$.

MATHEMATICAL APPENDIX

Proof of Theorem 1: The proof relies on Theorem 3.1 of Phillips (1987c) which shows that under the conditions of Assumptions 1 and 2 we have:

$$X_T(r) \stackrel{d}{=} W(r) + O_p(T^{-1}) . \tag{A.1}$$

where $\stackrel{d}{=}$ signifies equality in distribution. Using (A.1) we can prove the following Lemma related to the sample moments of $\{y_t\}$.

LEMMA A.1: *Let $\{y_t\}$ be generated by (1.1) with $\alpha = exp(c/T)$ and let the innovation sequence $\{u_t\}$ satisfy the conditions of Assumptions 1 and 2, then:*

a) $T^{-1/2}y_{[Tr]} \stackrel{d}{=} \sigma J_c(r) + T^{-1/2}exp(cr)y_0 + O_p(T^{-1})$;

b) $T^{-3/2}\Sigma_{t=1}^T y_t \stackrel{d}{=} \sigma\int_0^1 J_c(r)dr + T^{-1/2}y_0(exp(c) - 1)/c + O_p(T^{-1})$;

c) $T^{-2}\Sigma_{t=1}^T y_t^2 \stackrel{d}{=} \sigma^2\int_0^1 J_c(r)^2 dr + 2T^{-1/2}\sigma y_0\int_0^1 exp(cr)J_c(r)dr + O_p(T^{-1})$;

d) $T^{-1}\Sigma_{t=1}^T y_{t-1}u_t \stackrel{d}{=} \sigma^2\int_0^1 J_c(r)dW(r) + (\sigma^2 - \sigma_u^2) + T^{-1/2}\sigma y_0\int_0^1 exp(cr)dW(r)$
$$- T^{-1/2}(\nu/2\sigma^2)\eta + O_p(T^{-1}) .$$

Proof: The proof follows closely that of Phillips (1987c, Lemma 4.2). Using (1.1):

$$T^{-1/2}y_{[Tr]} = T^{-1/2}\Sigma_{j=1}^{[Tr]} exp(([Tr] - j)c)/T) u_j + T^{-1/2}exp([Tr]c/T)y_0$$
$$= \sigma\left\{X_T(r) + c\int_0^r exp((r-s)c)X_T(s)ds\right\} + T^{-1/2}exp([Tr]c/T)y_0.$$

Using (A.1), we deduce that :

$$T^{-1/2}y_{[Tr]} \stackrel{d}{=} \sigma\left\{W(r) + c\int_0^r exp((r - s)c)W(s)ds\right\} + T^{-1/2}exp(cr)y_0 + O_p(T^{-1})$$
$$= \sigma J_c(r) + T^{-1/2}exp(cr)y_0 + O_p(T^{-1}) ;$$

using the fact that $J_c(r) \equiv W(r) + c\int_0^r exp((r - s)c)W(s)ds$. The proof of parts (b) and (c) are analogous and omitted. To prove part (d) note that squaring (1.1), summing over t and rearranging we obtain:

$$T^{-1} \Sigma_{t=1}^{T} y_{t-1} u_t =$$

$$(1/2) \exp(-2c/T) \left\{ T^{-1} y_T^2 - T^{-1} y_0^2 - T(\exp(2c/T) - 1) T^{-2} \Sigma_{t=1}^{T} y_{t-1}^2 - T^{-1} \Sigma_{t=1}^{T} u_t^2 \right\}.$$

Note that :

$$T^{-1} \Sigma_{t=1}^{T} u_t^2 = T^{-1/2} [T^{-1/2} \Sigma_{t=1}^{T} (u_t^2 - \sigma_u^2)] + \sigma_u^2 \overset{d}{=} T^{-1/2} \zeta + \sigma_u^2 + O_p(T^{-1}) ;$$

where $\zeta \sim N(0, \nu^2)$ (see Phillips (1987c), Lemma 4.2). Hence using parts (a), (b) and (c):

$$T^{-1} \Sigma_{t=1}^{T} y_{t-1} u_t \overset{d}{=} (1/2) \left\{ [\sigma J_c(1) + T^{-1/2} \exp(c) y_0]^2 \right.$$

$$- 2c \left[\sigma^2 \int_0^1 J_c(r)^2 dr + 2T^{-1/2} \sigma y_0 \int_0^1 \exp(cr) J_c(r) dr \right]$$

$$\left. - \sigma_u^2 - T^{-1/2} \zeta \right\} + O_p(T^{-1})$$

$$\overset{d}{=} (\sigma^2/2) [J_c(1)^2 - 2c \int_0^1 J_c(r)^2 dr - 1] + (\sigma^2 - \sigma_u^2)/2$$

$$+ T^{-1/2} [\sigma \exp(c) y_0 J_c(1) - 2\sigma c y_0 \int_0^1 \exp(cr) J_c(r) dr - \zeta/2] + O_p(T^{-1}).$$

The result follows by noting that $(1/2)[J_c(1)^2 - 2c \int_0^1 J_c(r)^2 dr - 1] \overset{d}{=} \int_0^1 J_c(r) dW(r)$ and that $\exp(c) J_c(1) - 2c \int_0^1 \exp(cr) J_c(r) dr \overset{d}{=} \int_0^1 \exp(cr) dW(r)$ (see Perron (1991a)); and using the fact that $\nu \eta \overset{d}{=} \zeta$. Theorem 1 follows using (1.2) and (c–d) of Lemma A.1. □

Proof of Theorem 2: We can write $H(c, \gamma) = A(c, \gamma) + \delta - g\eta$, and $K(c, \gamma) = B(c, \gamma) - d$,

where $\qquad A(c, \gamma) = \int_0^1 J_c(r) dW(r) + \gamma \int_0^1 \exp(cr) dW(r)$,

and $\qquad B(\gamma, c) = \int_0^1 J_c(r)^2 dr + 2\gamma \int_0^1 \exp(cr) J_c(r) dr + \gamma^2 (\exp(2c) - 1)/2c$.

We have:

$$MGF(v, u) = E[\exp\{v(A(c, \gamma) + \delta - g\eta) + u(B(c, \gamma) - d)\}]$$

$$= \exp(v\delta - ud) E[\exp\{-vg\eta\}] E[\exp\{vA(c, \gamma) + uB(c, \gamma)\}],$$

since η is independent of the Wiener process $W(r)$, δ and d are fixed constants. With η a $N(0,1)$ random variable, $E[\exp(-vg\eta)] = \exp(v^2 g^2/2)$ and the result follows using Theorem 2 of Perron (1991a) who showed that $E[\exp\{vA(c, \gamma) + uB(c, \gamma)\}] = M_{c, \gamma}(v, u)$. □

REFERENCES

Andrews, D.W. and R. Fair, 1988, Inference in nonlinear econometric model with structural change, Review of Economic Studies 55, 615–640.

Bobkoski, M.J., 1983, Hypothesis testing in nonstationary time series, Unpublished PhD dissertation (Univeristy of Wisconsin, Department of Statistics, Madison, WI).

Cavanagh, C., 1986, Roots local to unity, manuscript, (Harvard University, Department of Economics, Cambridge, MA).

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.

DeJong, D.N., J.C. Nankervis, N.E. Savin and C.H. Whiteman, 1992, The power problems of unit root tests in time series with autoregressive errors, Journal of Econometrics 53, 323–343.

Dickey, D.A., 1976, Estimation and hypothesis testing for nonstationary time series, Unpublished Ph.D. dissertation (Iowa State University, Department of Economics, Ames).

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.

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

Evans, G.B.A. and N.E. Savin, 1981b, Testing for unit roots: 1, Econometrica 49, 753–779.

Fuller, W.A., 1976, Introduction to statistical time series (John Wiley, New York).

Gallant, A.R. and H. White, 1988, A unified theory of estimation and inference for nonlinear dynamic models (Basil Blackwell, Oxford).

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

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.

Mann, H.B. and A. Wald, 1943, On the statistical treatment of linear stochastic difference equations, Econometrica 11, 173–220.

Mehta, 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. and P. Perron, 1994, Local asymptotic distributions related to the AR(1) model with dependent errors, forthcoming in Journal of Econometrics.

Nabeya, S. and B.E. Sørensen, 1993, Asymptotic distributions of the least squares estimators and test statistics in the near unit root model with non-zero initial value and local drift and trend, manuscript (Brown University, Department of Economics, Providence, RI).

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.

Ng, S. and P. Perron, 1993, Useful modifications to some unit root tests with dependent errors and their local asymptotic properties, manuscript (Université de Montréal, Département de Sciences Economiques, Montréal, Canada).

Newey, W. and K. West, 1987, A simple positive definite heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.

Pantula, S. G., 1991, Asymptotic distributions of the unit root tests when the process is nearly stationary, Journal of Business and Economic Statistics 9, 63–71.

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, Tests of joint hypotheses for time series regression with a unit root," in T.B. Fomby and G.F. Rhodes (eds), Advances in econometrics, 8: Cointegration, spurious regression, and unit roots (JAI Press: Greenwich, CT), 135–159.

Perron, P., 1990b, The adequacy of limiting distributions in the AR(1) model with dependent errors, Econometric research program memo No. 349 (Princeton University, Department of Economics, Princeton, NJ).

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.

Phillips, P.C.B., 1987a, Time series regression with unit roots, Econometrica 55, 277–302.

Phillips, P.C.B., 1987b, Towards a unified asymptotic theory for autoregression, Biometrika 74, 535–47.

Phillips, P.C.B., 1987c, Asymptotic expansions in nonstationary vector autoregressions, Econometric Theory 3, 45–68.

Phillips, P.C.B. and S. Ouliaris, 1990, Asymptotic properties of residual based tests for cointegration, Econometrica 58, 165–193.

Phillips, P.C.B. and P. Perron, 1988, Testing for a unit root in time series regression, Biometrika 75, 335–346.

Rubin, H., 1950, Consistency of maximum likelihood estimates in the explosive case, in T.C. Koopmans (ed.), Statistical inference in dynamic economic models (Wiley, New York).

Schwert, G.W., 1989, Tests for unit roots: a monte carlo investigation, Journal of Business and Economic Statistics 7, 147–160.

Stock, J.H., 1991, Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series, Journal of Monetary Economics 28, 435–459.

White, H., 1984, Asymptotic theory for econometrician (Academic Press, New York).

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

TABLE 1: The Distribution of $T(\hat{a} - a)$; $a = \exp(c/T)$, $c = 0.0$, MA Errors, $u_t = e_t + \theta e_{t-1}$.

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%	Mean	Var
$\theta = -0.90$											
T=50	-61.57	-58.22	-55.76	-52.61	-39.95	-23.90	-19.87	-16.52	-13.64	-39.07	119.64
T=100	-110.28	-106.09	-101.79	-96.30	-71.38	-37.56	-29.85	-24.68	-19.33	-69.00	482.88
T=500	-427.96	-407.79	-385.31	-355.24	-192.30	-63.81	-45.28	-33.94	-24.94	-201.45	11573.26
T=1000	-733.45	-675.98	-625.27	-553.72	-234.38	-67.17	-49.53	-38.65	-29.70	-275.74	33716.06
T=5000	-1867.76	-1570.88	-1275.06	-999.41	-315.94	-92.19	-71.18	-57.97	-43.94	-447.14	158516.00
T=∞	-2625.73	-2036.37	-1601.94	-1181.47	-310.56	-74.68	-53.65	-41.39	-31.46	-502.44	302105.93
O_p(T=50)	-3311.05	-2217.16	-1673.02	-1199.96	-300.32	-69.54	-48.83	-36.66	-26.66	-502.44	335316.86
$\theta = -0.70$											
T=50	-49.25	-44.76	-40.64	-35.63	-16.52	-4.41	-3.05	-2.10	-1.34	-18.49	142.37
T=100	-77.81	-70.13	-62.69	-53.22	-20.46	-5.23	-3.45	-2.47	-1.59	-25.32	352.78
T=500	-162.58	-135.44	-115.70	-92.08	-27.12	-5.86	-3.80	-2.45	-1.46	-39.19	1347.55
T=1000	-197.33	-160.71	-133.73	-102.39	-26.75	-5.72	-3.92	-2.82	-1.91	-42.33	1882.63
T=∞	-239.25	-185.36	-145.63	-107.17	-27.50	-5.90	-3.96	-2.83	-1.90	-45.05	2528.21
O_p(T=50)	-256.08	-193.33	-149.26	-107.97	-26.59	-5.46	-3.55	-2.42	-1.49	-45.05	2796.20
$\theta = -0.50$											
T=50	-34.69	-30.10	-25.55	-20.45	-6.36	-0.83	-0.21	0.17	0.56	-8.77	66.38
T=100	-47.17	-39.19	-32.54	-25.19	-6.86	-0.95	-0.32	0.07	0.43	-10.39	111.29
T=500	-62.77	-49.76	-41.22	-31.25	-7.88	-1.00	-0.35	0.07	0.45	-12.65	192.09
T=1000	-76.00	-58.10	-45.59	-34.38	-7.97	-0.94	-0.36	0.08	0.45	-13.77	258.00
T=∞	-71.60	-55.32	-43.32	-31.71	-7.63	-1.02	-0.39	0.00	0.36	-12.91	231.35
O_p(T=50)	-75.77	-57.39	-44.27	-31.90	-7.38	-0.90	-0.27	0.13	0.50	-12.91	253.74
$\theta = -0.30$											
T=50	-22.93	-18.54	-15.17	-11.34	-2.67	0.28	0.69	0.99	1.36	-4.34	27.03
T=100	-27.44	-21.43	-17.22	-12.61	-2.76	0.26	0.65	0.93	1.26	-4.77	35.08
T=500	-29.43	-22.96	-18.81	-13.98	-3.08	0.23	0.63	0.93	1.19	-5.28	43.82
T=∞	-31.36	-24.13	-18.79	-13.62	-2.89	0.23	0.61	0.89	1.21	-5.19	46.49
O_p(T=50)	-32.95	-24.92	-19.15	-13.69	-2.79	0.28	0.66	0.96	1.29	-5.19	50.35
$\theta = 0.50$											
T=50	-6.83	-5.35	-4.06	-2.84	-0.13	1.35	1.77	2.18	2.77	-0.49	3.44
T=100	-7.58	-5.67	-4.23	-2.91	-0.16	1.33	1.73	2.16	2.71	-0.54	3.79
T=500	-7.19	-5.47	-4.33	-3.07	-0.22	1.30	1.71	2.13	2.63	-0.60	3.75
T=∞	-7.35	-5.54	-4.21	-2.91	-0.16	1.31	1.71	2.11	2.64	-0.55	3.73
O_p(T=50)	-7.77	-5.75	-4.30	-2.93	-0.14	1.32	1.74	2.15	2.69	-0.55	4.01

TABLE 2: The Distribution of $T(\hat{a} - a)$; $a = \exp(c/T)$, $c = 0.0$; AR Errors, $u_t = \rho u_{t-1} + e_t$.

	1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95.0%	97.5%	99.0%	Mean	Var
$\rho = -0.90$											
T=50	-76.33	-69.41	-62.43	-53.52	-19.36	-4.16	-2.42	-1.46	-0.71	-24.35	359.69
T=100	-122.62	-108.94	-93.53	-75.20	-23.93	-5.36	-3.34	-2.21	-1.34	-33.10	826.35
T=500	-214.48	-178.94	-146.79	-112.24	-31.24	-6.58	-4.09	-2.83	-1.71	-47.75	2284.21
T=1000	-273.27	-216.22	-175.23	-130.44	-32.63	-6.81	-4.65	-3.42	-2.35	-53.65	3359.83
T=∞	-274.69	-212.87	-167.27	-123.14	-31.71	-6.93	-4.71	-3.41	-2.35	-51.85	3329.55
O$_\rho$(T=50)	-364.82	-262.57	-194.42	-133.57	-26.28	-2.21	1.35	7.27	21.63	-51.85	5662.27
$\rho = -0.50$											
T=25	-24.84	-20.50	-17.28	-13.48	-3.54	0.08	0.52	0.85	1.24	-5.31	33.31
T=50	-30.92	-25.19	-20.46	-15.46	-3.86	-0.05	0.42	0.74	1.09	-6.11	47.49
T=100	-36.92	-29.36	-23.45	-17.22	-3.99	-0.11	0.36	0.66	0.98	-6.74	64.12
T=500	-40.14	-31.21	-25.53	-19.16	-4.42	-0.13	0.34	0.64	0.94	-7.46	79.20
T=∞	-42.61	-32.84	-25.64	-18.67	-4.21	-0.14	0.30	0.60	0.91	-7.34	83.94
O$_\rho$(T=50)	-45.63	-34.37	-26.36	-18.83	-4.02	-0.04	0.40	0.71	1.05	-7.34	94.12
$\rho = 0.50$											
T=25	-5.28	-3.60	-2.67	-1.73	0.27	1.76	2.31	2.82	3.46	0.10	2.55
T=25	-4.47	-3.38	-2.47	-1.61	0.24	1.65	2.11	2.62	3.26	0.09	2.08
T=50	-4.67	-3.42	-2.44	-1.63	0.20	1.62	2.06	2.53	3.26	0.07	2.08
T=100	-4.32	-3.15	-2.47	-1.67	0.14	1.57	2.05	2.50	3.14	0.03	1.96
T=500	-4.23	-3.15	-2.35	-1.57	0.19	1.57	2.03	2.49	3.11	0.07	1.92
O$_\rho$(T=50)	-4.53	-3.29	-2.41	-1.58	0.20	1.59	2.05	2.52	3.15	0.07	2.05
$\rho = 0.90$											
T=25	-1.66	-0.95	-0.57	-0.28	1.25	2.83	3.44	4.06	4.86	1.27	1.67
T=25	-1.04	-0.64	-0.38	-0.17	1.04	2.61	3.25	3.91	4.63	1.14	1.36
T=50	-0.73	-0.45	-0.28	-0.12	0.86	2.33	3.00	3.54	4.47	1.01	1.14
T=100	-0.58	-0.40	-0.26	-0.12	0.69	2.15	2.73	3.34	4.07	0.87	0.98
T=500	-0.53	-0.36	-0.23	-0.12	0.67	2.08	2.65	3.23	4.03	0.85	0.93
O$_\rho$(T=50)	-0.68	-0.43	-0.26	-0.11	0.68	2.08	2.66	3.24	4.04	0.85	0.94
$\rho = 0.95$											
T=25	-1.35	-0.70	-0.40	-0.11	1.52	3.07	3.73	4.33	5.06	1.53	1.68
T=50	-0.69	-0.37	-0.21	-0.05	1.33	2.89	3.56	4.24	5.02	1.42	1.42
T=100	-0.43	-0.25	-0.13	-0.02	1.11	2.63	3.29	3.87	4.76	1.24	1.22
T=500	-0.28	-0.17	-0.10	-0.03	0.79	2.29	2.89	3.51	4.43	0.99	1.01
T=1000	-0.28	-0.19	-0.11	-0.04	0.71	2.24	2.78	3.42	4.28	0.94	0.97
T=∞	-0.22	-0.14	-0.08	-0.01	0.73	2.15	2.73	3.33	4.14	0.93	0.90
O$_\rho$(T=50)	-0.34	-0.19	-0.10	-0.02	0.73	2.15	2.74	3.34	4.14	0.93	0.91

TABLE 3.a: The Distribution of $T(\hat{a} - a)$; Demeaned Case ;
$a = 1$; MA Errors , $u_t = e_t + \theta e_{t-1}$.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%
$\theta = -0.90$							
T=50	-65.20	-59.72	-57.18	-47.45	-37.11	-33.90	-28.27
T=100	-116.51	-108.90	-104.50	-88.19	-68.54	-62.53	-50.62
T=500	-457.57	-425.06	-405.42	-306.70	-171.65	-135.29	-84.14
T=1000	-820.46	-741.38	-690.78	-444.78	-204.56	-160.61	-108.17
T=5000	-2279.55	-1784.56	-1498.46	-673.53	-288.54	-236.37	-159.01
T=∞	-3978.96	-2506.52	-1978.06	-761.62	-260.23	-195.62	-119.40
$\theta = -0.70$							
T=50	-55.08	-48.22	-44.25	-29.68	-15.22	-12.26	-7.65
T=100	-89.32	-76.82	-69.41	-41.74	-18.74	-14.60	-9.49
T=500	-205.89	-159.47	-135.26	-63.86	-22.83	-16.59	-9.08
T=1000	-280.21	-203.31	-168.13	-68.48	-22.86	-17.33	-11.11
T=∞	-337.30	-226.71	-180.03	-69.65	-23.79	-17.87	-10.99
$\theta = -0.50$							
T=50	-42.04	-34.49	-29.95	-15.39	-5.98	-4.41	-2.28
T=100	-58.50	-45.64	-38.48	-17.89	-6.48	-4.76	-2.73
T=500	-85.99	-61.82	-50.14	-21.19	-7.04	-4.85	-2.33
T=1000	-104.93	-71.41	-57.51	-21.70	-6.88	-5.14	-2.95
T=∞	-101.45	-68.44	-54.41	-21.12	-7.06	-5.20	-2.92

TABLE 3.b: The Distribution of $T(\hat{a} - a)$; Demeaned Case ;
$a = 1$; AR Errors , $u_t = \rho u_{t-1} + e_t$.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%
$\rho = -0.90$							
T=50	-83.87	-73.53	-67.38	-38.29	-14.82	-11.11	-5.89
T=100	-140.38	-117.86	-103.19	-52.11	-19.45	-14.60	-8.51
T=500	-279.43	-209.68	-172.48	-74.80	-25.20	-18.29	-9.90
T=1000	-357.98	-258.37	-204.48	-80.05	-25.75	-19.40	-12.86
T=∞	-385.74	-260.19	-206.58	-79.93	-27.32	-20.51	-12.54
$\rho = -0.50$							
T=50	-39.27	-30.13	-25.00	-10.94	-3.70	-2.54	-0.94
T=100	-48.86	-35.20	-28.64	-11.93	-3.92	-2.64	-1.22
T=500	-56.08	-39.64	-31.68	-13.05	-4.07	-2.72	-0.97
T=1000	-66.18	-44.94	-35.63	-13.20	-4.10	-2.83	-1.37
T=∞	-61.08	-41.12	-32.70	-12.68	-4.10	-2.89	-1.30
$\rho = 0.90$							
T=25	-4.50	-2.94	-2.26	0.01	2.43	3.18	4.76
T=50	-4.03	-2.71	-2.16	-0.18	2.24	3.10	4.64
T=100	-4.18	-2.85	-2.18	-0.26	2.02	2.78	4.53
T=500	-4.52	-3.00	-2.26	-0.32	1.83	2.52	4.18
T=1000	-4.77	-3.10	-2.31	-0.31	1.71	2.46	4.09
T=∞	-4.78	-3.03	-2.25	-0.29	1.71	2.44	4.00

TABLE 4.a: The Distribution of $T(\hat{\alpha} - \alpha)$; Detrended Case ;
$\alpha = 1$; MA Errors , $u_t = e_t + \theta e_{t-1}$.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%
$\theta = -0.90$							
T=50	-66.99	-62.31	-59.65	-50.55	-41.35	-38.81	-33.44
T=100	-120.76	-113.66	-110.15	-96.14	-81.92	-77.88	-70.61
T=500	-481.24	-457.07	-442.54	-382.56	-298.77	-271.18	-218.21
T=1000	-872.31	-812.90	-780.05	-631.57	-417.01	-355.29	-263.09
T=5000	-2676.75	-2324.64	-2076.49	-1196.41	-601.78	-493.03	-348.43
T=∞	-5377.20	-3883.05	-3252.54	-1629.49	-752.34	-600.56	-401.37
$\theta = -0.70$							
T=50	-60.49	-54.52	-51.25	-39.76	-27.64	-24.12	-18.24
T=100	-99.63	-89.58	-83.76	-62.40	-40.42	-34.77	-26.12
T=500	-252.88	-210.87	-188.78	-115.11	-60.87	-49.59	-33.62
T=1000	-347.75	-271.94	-238.93	-135.45	-61.02	-48.07	-31.97
T=∞	-492.19	-355.00	-296.95	-149.05	-69.37	-55.83	-35.77
$\theta = -0.50$							
T=50	-49.92	-42.80	-39.34	-26.10	-14.52	-12.10	-7.85
T=100	-72.44	-60.40	-53.90	-33.21	-17.54	-14.12	-9.84
T=500	-114.36	-89.39	-76.93	-41.49	-20.00	-16.04	-10.41
T=1000	-137.48	-100.82	-86.42	-45.22	-19.17	-15.00	-9.63
T=∞	-145.05	-106.77	-89.74	-44.99	-20.88	-16.74	-10.83

TABLE 4.b: The Distribution of $T(\hat{\alpha} - \alpha)$; Detrended Case ;
$\alpha = 1$; AR Errors , $u_t = \rho u_{t-1} + e_t$.

	1.0%	5.0%	10.0%	50.0%	90.0%	95.0%	99.0%
$\rho = -0.90$							
T=50	-89.14	-83.14	-78.70	-57.28	-32.35	-26.03	-17.65
T=100	-156.14	-139.86	-130.13	-86.66	-46.90	-37.68	-24.95
T=500	-362.53	-289.82	-251.28	-141.50	-68.93	-55.71	-36.29
T=1000	-461.75	-353.80	-303.10	-161.82	-70.29	-55.09	-35.96
T=∞	-551.84	-406.08	-340.61	-171.05	-79.57	-64.12	-43.16
$\rho = -0.50$							
T=50	-50.04	-40.90	-36.30	-21.18	-10.28	-8.22	-4.98
T=100	-64.15	-50.79	-43.89	-23.98	-11.47	-9.08	-5.85
T=500	-77.87	-59.27	-50.21	-25.95	-12.12	-9.55	-5.88
T=1000	-89.14	-63.74	-54.42	-27.85	-11.51	-8.90	-5.46
T=∞	-87.13	-64.20	-53.92	-27.05	-12.42	-9.88	-6.36
$\rho = 0.90$							
T=25	-8.92	-6.03	-4.94	-1.54	2.24	3.23	5.06
T=50	-6.60	-4.50	-3.64	-0.95	2.38	3.48	5.39
T=100	-5.72	-3.85	-3.01	-0.77	2.21	3.12	5.18
T=500	-5.22	-3.50	-2.63	-0.56	1.79	2.63	4.61
T=1000	-5.43	-3.55	-2.69	-0.58	1.72	2.60	4.56
T=∞	-5.54	-3.50	-2.62	-0.54	1.56	2.48	4.41
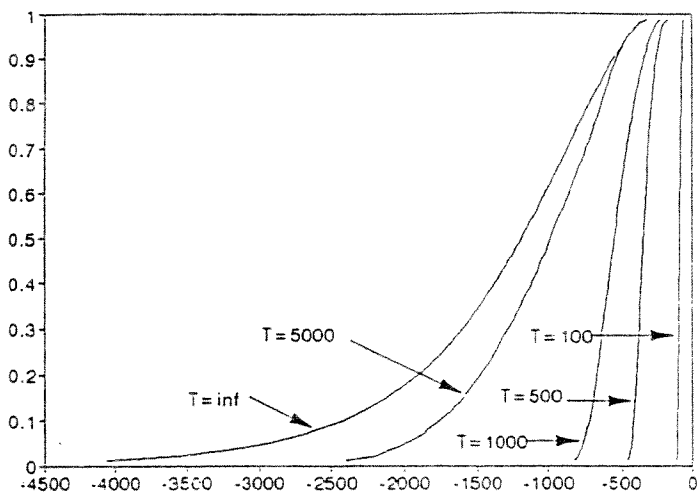
Figure 1: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; MA errors, $\theta = -.9$.
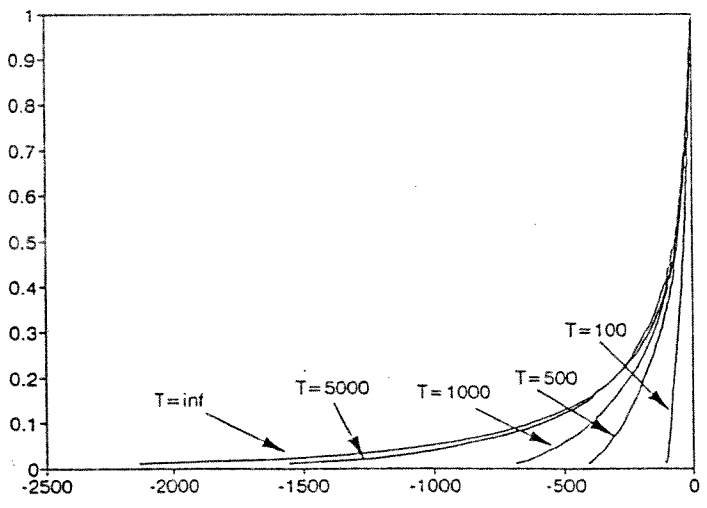
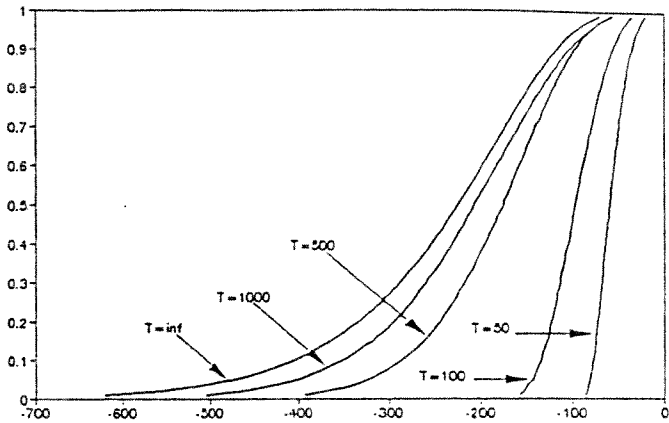**Figure 2:** The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = 2$; MA errors, $\theta = -.9$.
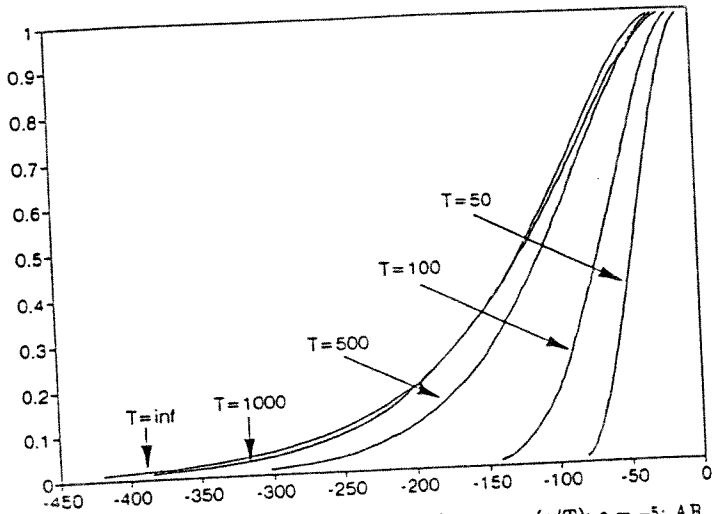
Figure 3: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; AR errors, $\rho = -.9$.
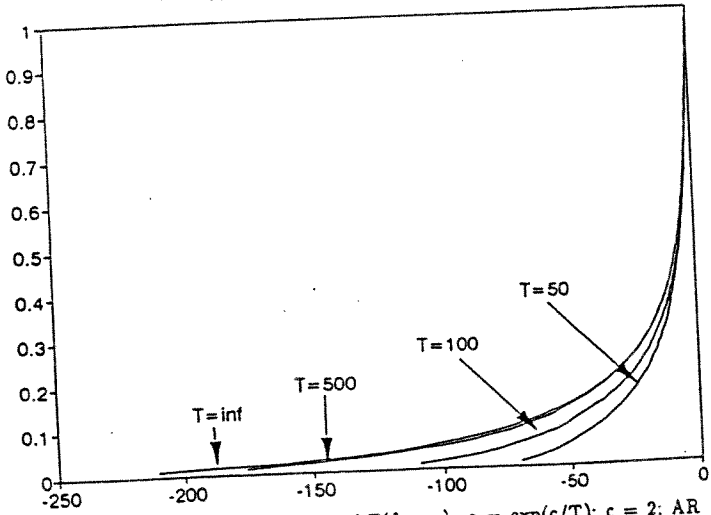
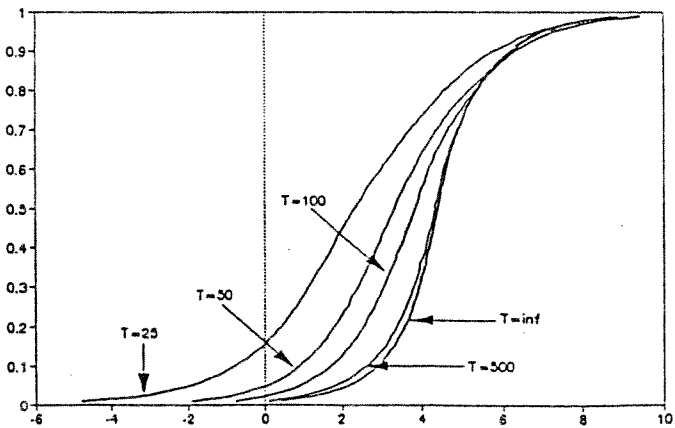**Figure 4:** The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = 2$; AR errors, $\rho = -.9$.
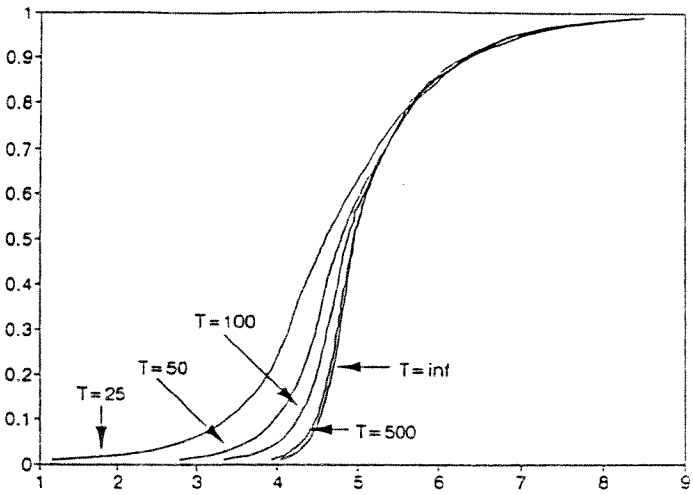
Figure 5: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; AR errors, $\rho = .9$.
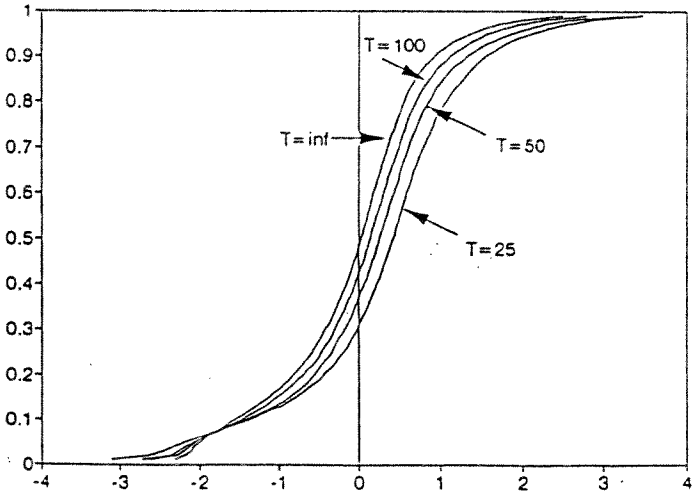
Figure 6: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = 2$; AR errors, $\rho = .9$.
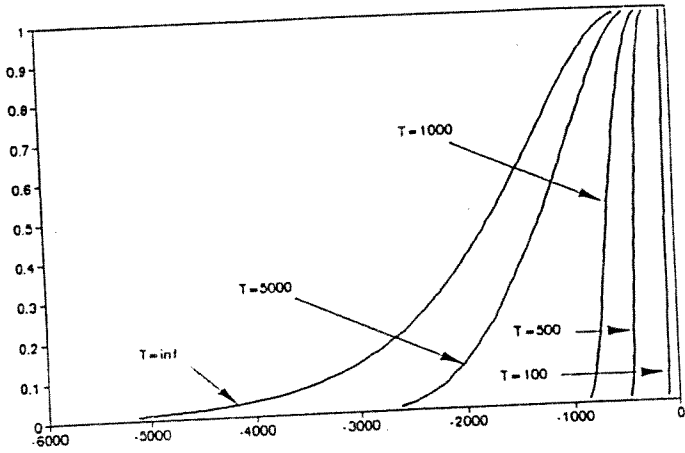
Figure 7: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; MA errors, $\theta = -.9$; Demeaned Case.
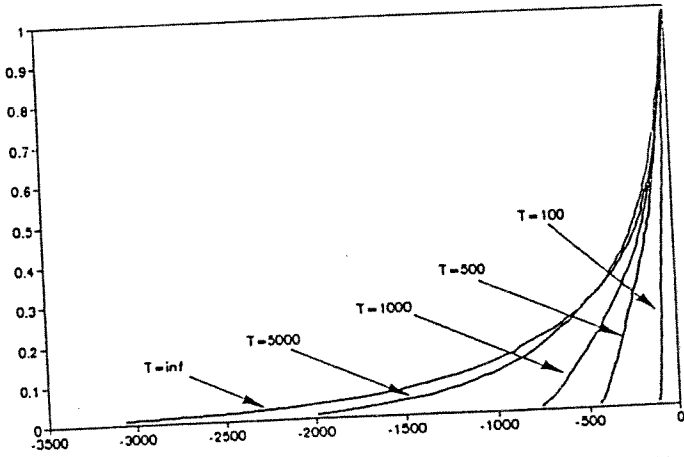
Figure 8: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = 2$; MA errors, $\theta = -.9$; Demeaned Case.
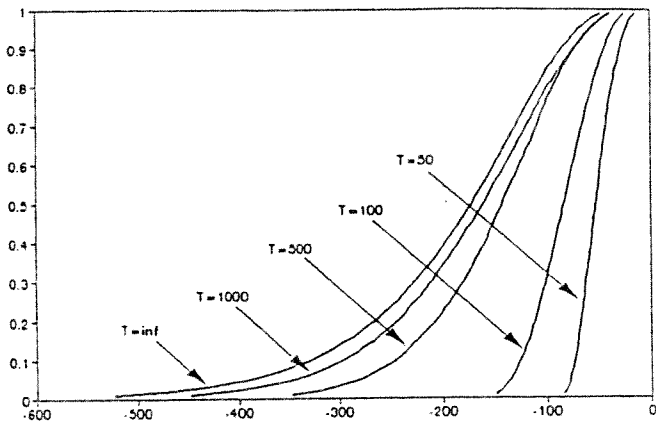
Figure 9: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; AR errors, $\rho = -.9$; Demeaned Case.
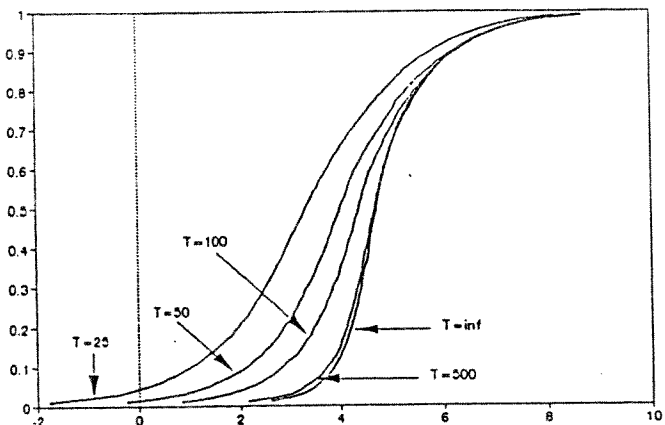
Figure 10: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; AR errors, $\rho = .9$; Demeaned Case.
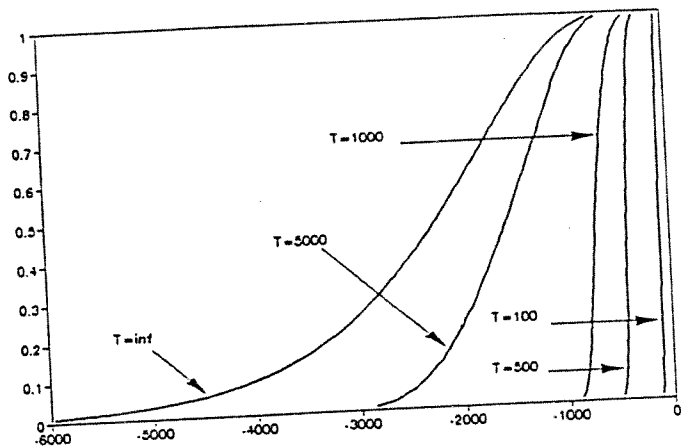
Figure 11: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; MA errors, $\theta = -.9$; Detrended Case.
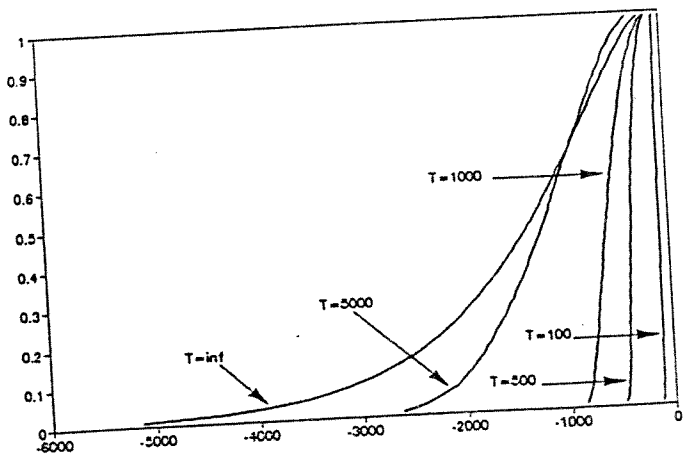
Figure 12: The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = 2$; MA errors, $\theta = -.9$; Detrended Case.

Figure 13:   The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; AR errors, $\rho = -.9$; Detrended Case.

Figure 14:   The Distribution of $T(\hat{\alpha} - \alpha)$; $\alpha = \exp(c/T)$; $c = -5$; AR errors, $\rho = .9$; Detrended Case.

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.

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

9336 : Kollmann, Robert, "The Duration of Unemployment as a Signal : Implications 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., Fabienne Rosenwald et Lise 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, 64 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.