

Asymptotic approximations in the near-integrated model with a non-zero initial condition

PIERRE PERRON¹ AND COSME VODOUNOU²

¹*Department of Economics, Boston University, 270 Bay State Rd.,
Boston, MA 02215, USA
E-mail: perron@bu.edu*

²*Institut National de la Statistique et de l'Analyse Économique, Bénin
E-mail: vodounoc@caramail.com*

Received: October 2000

Summary This paper considers various asymptotic approximations in the near-integrated first-order autoregressive model with a non-zero initial condition. We first extend the work of Knight and Satchell (1993), who considered the random walk case with a zero initial condition, to derive the expansion of the relevant joint moment generating function in this more general framework. We also consider, as alternative approximations, the stochastic expansion of Phillips (1987c) and the continuous-time approximation of Perron (1991a). We assess, via a Monte Carlo simulation study, the extent to which these alternative methods provide adequate approximations to the finite sample distribution of the least-squares estimator in a first-order autoregressive model. The results show that, when the initial condition is non-zero, Perron's (1991a) continuous-time approximation performs very well while the others only offer improvements when the initial condition is zero.

Keywords: *Edgeworth expansion, Continuous-time asymptotics, Stochastic expansion, Distribution function, Autoregressive model.*

1. INTRODUCTION

It is often the case that a nuisance parameter has a substantial effect on the distribution of some statistic in finite samples but that this effect disappears from the leading term of the asymptotic distribution. For that reason, several refinements of the usual asymptotic framework have been suggested which retain the effect of such nuisance parameters. A standard example is the Edgeworth expansion which has been applied to several cases. In this paper, we consider the standard first-order autoregressive model where the initial condition is one such nuisance parameter that substantially affects the finite sample distribution of the least-squares estimator but whose effect is not present in the leading term of the asymptotic distribution (henceforth, simply referred to as the limiting distribution).

The autoregressive model of order one has attracted considerable attention in both the statistics and econometrics literature, especially concerning the study of various asymptotic refinements. It is simple enough to yield tractable solutions, yet complex enough to generate interesting problems. It is something of a base case by which we can learn useful features about the

properties of alternative asymptotic refinements that can be helpful in guiding our efforts in more complex and empirically relevant models.

Since the pioneering works of Mann and Wald (1943) and Rubin (1950) establishing the root- T asymptotic normality of the least-squares estimator in the stationary case much of the research has focused on the case where the autoregressive coefficient, say α , is equal to one; see White (1958), Fuller (1976), Dickey and Fuller (1979), Evans and Savin (1981), Phillips (1987a), and Phillips and Perron (1988) among others. A feature of interest in this case is that the rate of convergence is T (the sample size) and that the limit distribution can be expressed in terms of a functional of a Wiener process. Also, White (1958) has shown in the explosive case that the rate of convergence is even faster (and depends on α), and that the limit distribution is non-standard and depends on the value of the initial condition. When the latter is zero the limit is that of a Cauchy random variable.

Two features emerge from these asymptotic results which contrast with the finite sample distribution of the estimator: the discontinuity (of both the rate of convergence and the limit) at the boundary $\alpha = 1$ and the invariance of the distribution with respect to the initial condition (unless α is greater than one). Since the finite sample distribution is continuous with respect to α and is influenced by the initial condition, this suggests that these limiting distributions should be poor approximations to the finite sample distribution when α is not equal but near to one. This problem has effectively been demonstrated by Evans and Savin (1981b).

To alleviate the first problem concerning the discontinuity at one, the research has concentrated on the so-called near-integrated model where the autoregressive coefficient is specified as being local to unity, for example using the specification $\alpha = \exp(c/T)$ where T is the sample size and c is some non-centrality parameter. When $c = 0$, we recover the unit root case; when $c < 0$, the process is said to be locally stationary and when $c > 0$, it is locally explosive. The resulting limiting distribution of the least-squares estimator is continuous with respect to the non-centrality parameter c and can be expressed as a functional of an Ornstein–Uhlenbeck process. Also, it provides a very good approximation to the finite sample distribution when the initial condition is zero. See Bobkoski (1983), Cavanagh (1986), Chan and Wei (1987), Larson (1995), Phillips (1987b), Perron (1989), and Tanaka (1996). The limiting distribution remains, however, invariant with respect to the initial condition and, accordingly, the adequacy of the approximation deteriorates when this initial value is non-zero.

A few asymptotic approximations have been proposed to capture the effect of this initial condition while retaining the continuity at $\alpha = 1$. One is the stochastic expansion of Phillips (1987c) and the other is the continuous-time approximation of Perron (1991a,b).

This paper contributes to this literature by first considering the Edgeworth expansion in this near-integrated context. Phillips (1977) showed that in the stationary case the Edgeworth expansion is not satisfactory when the autoregressive coefficient is close to one. Satchell (1984) considered the Edgeworth expansion to order $O(T^{-1/2})$ for a fixed initial condition in the stationary and explosive cases and Vodounou (1997b) extended his results to order $O(T^{-1})$ for the stationary case (the expansion to order $O(T^{-1})$ is necessary for the initial condition to affect the distribution). Also, Knight and Satchell (1993) have shown that the Edgeworth expansion performs well in the unit root case when the initial condition is zero. Our aim is to see how it performs in the general near-integrated case with a non-zero initial condition and compare its relative merits, as an approximation to the finite sample distribution, with Phillip's stochastic expansion and Perron's continuous-time approximation.

The rest of this paper is structured as follows. Section 2 presents the model and the estimator as well as some finite sample distributional results. Section 3 presents the three asymptotic

approximations and the relevant moment generating functions used to numerically evaluate the corresponding distribution functions. Section 4 contains a study of the adequacy of the various approximations for a selected range of cases. Section 5 offers some concluding comments and a mathematical appendix contains some technical derivations.

2. MODEL AND ESTIMATOR

We consider the AR(1) model defined as follows:

$$\begin{aligned}
 y_t &= \alpha y_{t-1} + u_t, \\
 u_t &\sim N(0, \sigma^2), \\
 \alpha &= \exp(c/T), \\
 y_0/\sigma &= \kappa \quad \text{fixed}, \\
 t &= 1, \dots, T.
 \end{aligned}
 \tag{1}$$

Here the scaled initial value y_0/σ is considered fixed at some constant κ , the errors u_t are independent and normally distributed with mean 0 and variance σ^2 and c is some real fixed coefficient. The autoregressive coefficient α is said to be local to unity. When $c = 0$, we have a unit root; when $c < 0$, the process is locally stationary, and when $c > 0$ it is locally explosive. We consider the *OLS* estimator of α defined by

$$\hat{\alpha} = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2}.
 \tag{2}$$

This is also the maximum likelihood estimator conditional upon y_0 . The estimate $\hat{\alpha}$ being invariant to a scale transformation of the data, we can, without loss of generality, suppose that $\sigma^2 = 1$, hence $u_t \sim N(0, 1)$ and $y_0 = \kappa$. We are interested in the distribution of the statistic

$$T(\hat{\alpha} - \alpha) = \frac{T^{-1} \left(\sum_{t=1}^T y_t y_{t-1} - \alpha \sum_{t=1}^T y_{t-1}^2 \right)}{T^{-2} \sum_{t=1}^T y_{t-1}^2}.
 \tag{3}$$

Using matrix notation and the vector $y = (y_1, y_2, \dots, y_T)'$, we can write this statistic as

$$T(\hat{\alpha} - \alpha) = \frac{y' Ay / T + (y_1 \kappa - \kappa^2 \alpha) / T}{y' B y / T^2 + \kappa^2 / T^2},
 \tag{4}$$

where A and B are T by T symmetric matrices (and B is positive semi-definite) defined by (see, for example, White (1958), Evans and Savin (1981b))

$$A = \begin{bmatrix} -\alpha & 1/2 & 0 & \dots & 0 \\ 1/2 & -\alpha & \ddots & \ddots & \vdots \\ 0 & 1/2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -\alpha & 1/2 \\ 0 & \dots & 0 & 1/2 & 0 \end{bmatrix}$$

and

$$B = \begin{bmatrix} I_{T-1} & \tilde{O}' \\ \tilde{O} & 0 \end{bmatrix}$$

where I_{T-1} is the identity matrix of dimension $T - 1$ and \tilde{O} is a 1 by $(T - 1)$ vector of zeros. We use the notation

$$X = y' Ay/T + (y_1 \kappa - \kappa^2 \alpha)/T, \quad (5)$$

$$H = y' By/T^2 + \kappa^2/T^2, \quad (6)$$

and define the joint moment generating function of (X, H) by $m_\kappa(u, v) = E(\exp\{uX + vH\})$. From the result of White (1958), we have

$$m_\kappa(u, v) = \exp\left\{\kappa^2\left(\frac{v}{T^2} - \alpha\frac{u}{T} - \frac{\alpha^2}{2}\right) + \frac{\kappa^2}{2}\left(\frac{u}{T} + \alpha\right)^2 \frac{D(T-1)}{D(T)}\right\} D(T)^{-1/2}, \quad (7)$$

where $D(T)$ satisfies the second-order difference equation

$$D(T) = pD(T-1) - q^2D(T-2),$$

with initial conditions $D(1) = 1$ and $D(2) = p - q^2$ and where $p = 1 + \alpha^2 - 2v/T^2 + 2\alpha u/T$ and $q = -(\alpha + u/T)$. Solving this system, we have

$$D(T) = \frac{1-s}{r-s}r^T + \frac{1-r}{s-r}s^T, \quad (8)$$

where r and s are the roots of the equation $x^2 - px + q^2 = 0$. Expression (7) allows to compute the joint characteristic function of the pair (X, H) and, using Gurland's (1948) inversion theorem, to obtain the exact distribution of the statistic $T(\hat{\alpha} - \alpha)$.

3. ALTERNATIVE ASYMPTOTIC APPROXIMATIONS

We now discuss three different asymptotic approximations which apply to this near-integrated context that are specifically influenced by the value of the initial condition. These are: an extension of the Edgeworth approximation studied by Knight and Satchell (1993), the stochastic expansion considered by Phillips (1987c) and the continuous-time approximation analysed by Perron (1991a).

3.1. The Edgeworth approximation

Here, we extend the work of Knight and Satchell (1993) who considered the case $c = 0$ with a zero initial condition. We first consider the joint moment generating function of the pair (X, H) defined by (5) and (6). The result is given in the following theorem.

Theorem 1. Let $\{y_t\}$ be generated by (1) and denote by $MGF(u, v)$ the joint moment generating function of the pair (X, H) defined by (5) and (6). Then

$$MGF(u, v) = \exp(-r_1/2) \left(\cos \delta - r_1 \frac{\sin \delta}{\delta} \right)^{-1/2} \times \left\{ 1 + \frac{F_1(u, v, c)}{4T} \frac{\sin \delta}{\delta} + \kappa^2 \frac{F_2(u, v, c)}{2T} \frac{\sin \delta}{\delta} - c^2 \frac{F_3(u, v, c)}{4T} \cos \delta \right\} + O(T^{-2})$$

where

$$F_1(u, v, c) = \frac{(u^2 + 2v)(2c^2 + c^3 - 2uv + 2cu^2 + 2cu - 2v - 2cv + 3c^2u)}{\delta^2(\cos \delta - r_1 \frac{\sin \delta}{\delta})},$$

$$F_2(u, v, c) = \frac{u^2 + 2v}{\cos \delta - r_1 \frac{\sin \delta}{\delta}},$$

$$F_3(u, v, c) = \frac{u^2 + 2v}{\delta^2(\cos \delta - r_1 \frac{\sin \delta}{\delta})},$$

and $r_1 = c + u$, $\delta^2 = 2v - 2cu - c^2$.

The expansion presented above and its proof were contained in most parts in some draft notes, Knight and Satchell (1990). We, nevertheless, present the full proof in the appendix since that draft contained some technical errors (hence, a different expression) and was never published.

In what follows we shall denote by $MGF_{KS}(u, v)$ the approximation to the joint moment generating function of the pair (X, H) obtained by retaining terms that are of higher order than $O(T^{-2})$. That is, we have

$$MGF_{KS}(u, v) = \exp(-r_1/2) \left(\cos \delta - r_1 \frac{\sin \delta}{\delta} \right)^{-1/2} \times \left\{ 1 + \frac{F_1(u, v, c)}{4T} \frac{\sin \delta}{\delta} + \kappa^2 \frac{F_2(u, v, c)}{2T} \frac{\sin \delta}{\delta} - c^2 \frac{F_3(u, v, c)}{4T} \cos \delta \right\}. \tag{9}$$

The use of Gurland's (1948) inversion theorem leads to the following approximation for the distribution function $\Pr(T(\hat{\alpha} - \alpha) < z)$:

$$F_{KS}(z) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty IM \left(\frac{cf_{KS}(u, -uz)}{u} \right) du, \tag{10}$$

where $cf_{KS}(u, v) = MGF_{KS}(iu, iv)$ is the characteristic function with $i = \sqrt{-1}$ and where $IM(\cdot)$ denotes the imaginary part of the argument.¹ We close by noting that the specification

¹As in Knight and Satchell (1993), we can write the distribution function as $F_{KS}(z) = 1/2 - (1/\pi)[FUNC D0(z) + (1/4T)FUNC D1(z) + (c^2/4T)FUNC D2(z) - (\kappa^2/2T)FUNC D3(z)]$. In practice, the four functions, $FUNC Di(z)$ are numerically integrated separately and the final value of the distribution function obtained using this relation. The values of these functions for a grid of z are presented in Vodounou (1997a). They are not reproduced here to conserve space.

$\alpha = \exp(c/T)$ instead of $\alpha = 1 + c/T$ is not innocuous. Use of the latter would have given a slightly different expression for the joint moment generating function.

3.2. The stochastic expansion

We consider now the stochastic expansion derived by Phillips (1987c) and also studied by Perron (1996). It is given by

$$T(\hat{\alpha} - \alpha) \stackrel{d}{=} Z(c, \gamma)/K(c, \gamma) + O_p(T^{-1}) \tag{11}$$

where

$$\begin{aligned} Z(c, \gamma) &= \int_0^1 J_c(r)dW(r) + \gamma \int_0^1 \exp(cr)dW(r) - \eta/\sqrt{2T}; \\ K(c, \gamma) &= \int_0^1 J_c(r)^2 dr + 2\gamma \int_0^1 \exp(cr)J_c(r)dr. \end{aligned}$$

Here, η is a $N(0, 1)$ random variable independent of the Wiener process $W(r)$, $J_c(r)$ is the Ornstein–Uhlenbeck process defined by $J_c(r) = \int_0^r \exp(c(r - s))dW(s)$ and $\gamma = \kappa/T^{1/2}$. We approximate the distribution of $T(\hat{\alpha} - \alpha)$ by the distribution of the random variable $Z(c, \gamma)/K(c, \gamma)$. The joint moment generating function of the pair $(Z(c, \gamma), K(c, \gamma))$ was derived by Perron (1996) in a more general context and it is shown in the appendix that it can be represented as

$$\begin{aligned} MGF_{PCB}(u, v) &= \exp(-r_1/2) \left(\cos \delta - r_1 \frac{\sin \delta}{\delta} \right)^{-1/2} \exp \left\{ \frac{u^2}{4T} - \frac{\exp(2c) - 1}{2c} v\gamma^2 \right\} \\ &\times \exp \left\{ -\frac{\gamma^2 v}{2} (r_1 - \lambda) \left[1 - \frac{\exp(-\lambda)}{\cos \delta - \frac{r_1}{\delta} \sin \delta} \right] \right\} \end{aligned} \tag{12}$$

where δ and r_1 are defined in Theorem 1 and $\lambda^2 = -\delta^2$. The corresponding approximation to the distribution function of $T(\hat{\alpha} - \alpha)$ is obtained using the inversion formula (10) with $cf_{PCB}(u, v) = MGF_{PCB}(iu, iv)$ instead of $cf_{KS}(u, v)$ in which case the distribution function is denoted $F_{PCB}(z)$.

The expansion (11) is not, strictly, a stochastic expansion to order $O_p(T^{-1})$ since it retains some terms that are $O_p(T^{-1})$. The expansion that keeps only terms that are of order $O_p(T^{-1/2})$ is

$$\frac{Z(c, \gamma)}{K(c, \gamma)} = X_0 + X_1 T^{-1/2} + O_p(T^{-1}). \tag{13}$$

where

$$X_0 = H_0/K_0, \tag{14}$$

$$X_1 = \frac{1}{K_0} \left(H_1 - K_1 \frac{H_0}{K_0} \right), \tag{15}$$

with $H_0 = \int_0^1 J_c(r)dW(r)$, $H_1 = \kappa \int_0^1 \exp(cr)dW(r) - \eta/\sqrt{2}$, $K_0 = \int_0^1 J_c(r)^2 dr$, and $K_1 = 2\kappa \int_0^1 \exp(cr)J_c(r)dr$. However, obtaining the moment generating function of the variable $X_0 + X_1 T^{-1/2}$ is more complex and we continue to work with the random variable $Z(c, \gamma)/K(c, \gamma)$ as the approximation. We will see later that excluding all terms that are $O_p(T^{-1})$ would deteriorate the approximation.

3.3. Continuous-time approximation

Consider a continuous-time process y_t defined by the stochastic differential equation

$$dy_t = \theta y_t dt + \sigma_1 dW(t), \quad y_0 = \kappa \sigma_1, \quad t \in [0, N], \tag{16}$$

where θ and σ are unknown parameters, κ is a fixed constant and N is the span of the data. The stochastic process $\{y_t\}$ has the following discrete time representation:

$$y_{th} = \alpha_h y_{(t-1)h} + u_{th}, \quad y_0 = \kappa \sigma_1, \quad t = 1, \dots, T, \tag{17}$$

where $u_{th} \sim N(0, \sigma_1^2(\exp(2\theta h) - 1)/2\theta)$ and h is the sampling interval such that $Th = N$ (see, for example, Bergstrom (1984), Phillips (1987a), Perron (1991a,b)) The autoregressive coefficient is $\alpha_h = \exp(\theta h)$. Hence, (17) is equivalent (up to a scale transformation) to model (1) with $c = \theta N$.

Perron (1991a) considered the distribution of the least-squares estimators $\hat{\theta}$ and $\hat{\alpha}_h$ defined, respectively, by $\hat{\theta} = \int_0^N y_t dy_t / \int_0^N y_t^2 dt$ and $\hat{\alpha}_h = \sum_{t=1}^T y_{th} y_{(t-1)h} / \sum_{t=1}^T y_{(t-1)h}^2$ when y_t and y_{th} are generated by (16) and (17). He obtained the results:

$$\begin{aligned} N(\hat{\theta} - \theta) &\stackrel{d}{=} A(\gamma, c)/B(\gamma, c), \\ T(\hat{\alpha}_h - \alpha_h) &\Rightarrow A(\gamma, c)/B(\gamma, c), \end{aligned} \tag{18}$$

the convergence result being obtained with N fixed and h converging to 0. The variables are defined by

$$\begin{aligned} A(\gamma, c) &= \int_0^1 J_c(r) dW(r) + \gamma \int_0^1 \exp(cr) dW(r), \\ B(\gamma, c) &= \int_0^1 J_c(r)^2 dr + 2\gamma \int_0^1 \exp(cr) J_c(r) dr + \gamma^2 (\exp(2c) - 1)/2c, \end{aligned}$$

with $\gamma = \kappa/T^{1/2}$ and $c = \theta N$. Here, we consider approximating the distribution of $T(\hat{\alpha} - \alpha)$ by the distribution of the random variable $A(\gamma, c)/B(\gamma, c)$ which we denote by $F_{PP}(z)$. From Perron (1991a), we have

$$F_{PP}(z) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty IM\left(\frac{cf_{PP}(u, -uz)}{u}\right) du,$$

where $cf_{PP}(u, v) = MGF_{PP}(iu, iv)$, with $MGF_{PP}(u, v)$ the joint moment generating function of the pair $(A(\gamma, c), B(\gamma, c))$ given by

$$MGF_{PP}(u, v) = \psi_c(u, v) \exp\left\{-\frac{\gamma^2}{2}(u + c - \delta)(1 - \exp(u + c + \delta))\psi_c^2(u, v)\right\}$$

with

$$\psi_c^2(u, v) = \frac{2\delta \exp(-u - c)}{(\delta + u + c) \exp(-\delta) + (\delta - u - c) \exp(\delta)},$$

where $\delta^2 = 2v - 2cu - c^2$. A little algebra shows that an equivalent expression is given by

$$MGF_{PP}(u, v) = \exp(-r_1/2) \left(\cos \delta - r_1 \frac{\sin \delta}{\delta} \right)^{-1/2} \exp \left\{ \frac{\gamma^2(u^2 + v)}{2(\cos \delta - r_1 \sin \delta/\delta)} \frac{\sin \delta}{\delta} \right\} \quad (19)$$

with $r_1 = c + u$.

Remark 1. The continuous-time asymptotic distribution described above is identical to the leading term of the asymptotic distribution of the least-squares estimate in an autoregression of order one when the data generating process is specified by (1) with $y_0/\sigma = \kappa T^{1/2}$ (see Tanaka, 1990). The continuous-time asymptotic framework can be seen as providing a theoretical justification for why the ad hoc specification with y_0 modelled as increasing with the sample size delivers adequate results. Note that it also provides a theoretical justification for the specification $\alpha = \exp(c/T)$ when modelling nearly integrated processes.

4. ADEQUACY OF THE APPROXIMATIONS

We now discuss to what extent the different approximations described are good approximations to the exact distribution of $T(\hat{\alpha} - \alpha)$ especially when the initial condition is non-zero. To that extent, we need to numerically evaluate the distribution functions $F_{KS}(z)$, $F_{PCB}(z)$ and $F_{PP}(z)$.² We start with some technical descriptions of how these evaluations are performed.

4.1. Numerical evaluation of the distribution functions

The evaluation of the approximate distribution functions at some point z requires the numerical integration of the relevant functions. This involves two special problems. The first is related to the presence of the square root of the complex number $(\cos \delta - r_1 \frac{\sin \delta}{\delta})^{-1}$. Using the principal root does not ensure the continuity of the integrand. One must therefore carefully assess which root to select along the integration path to ensure continuity. The solution to this problem is explained in detail in Perron (1989).

The second difficulty is how to ensure a given desired precision when the integrand may show wide fluctuations and multiple crossings at zero. Here, we follow the suggestion of Knight and Satchell (1993). If the integrand shows large variations near zero but reduces rapidly when evaluated at a higher number, the integral is computed over a range $[\epsilon, \bar{V}]$, where ϵ is some small number that we set at $\epsilon = 1.0E - 7$ and the upper bound is such that the absolute value of the integrand evaluated at this upper bound is less than ϵ (see also, Perron (1989, 1991a), Hisamatsu and Maekawa (1994)).

However, when the integrand shows oscillatory behaviour that dissipates slowly, Knight and Satchell (1993) suggest integrating between the points at which the integrand crosses the zero axis. Here, the upper bound of integration depends on the values of the successive integrations.

²In the case of the stochastic expansion, obtaining the distribution function $F_{PCB}(z)$ by numerical integration of the characteristic function $cf_{PCB}(u, v)$ given by (12) yielded unstable or non-convergent results in some parts of the parameter space considered. For this case, to ensure adequate results, we resorted to simulations to obtain the relevant distribution functions. We used 30 000 replications of the random variable $Z(c, \gamma)/K(c, \gamma)$ where the Wiener process was approximated by 500 steps of a sequence of *i.i.d.* $N(0, 1)$ and η was drawn as an independent $N(0, 1)$.

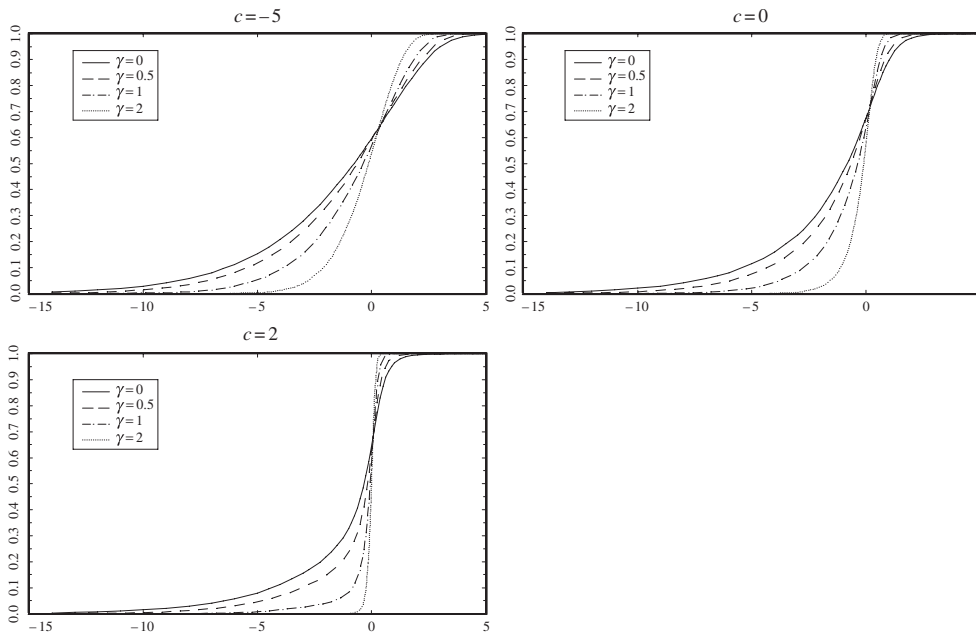


Figure 1. Effect of the initial condition for different values of c . Exact distribution for $T = 25$.

The method adopted is to stop whenever an integral between two points at which the integrand crosses zero is less than $1.0E - 5$.

The strategies discussed above are used to compute the distribution functions for various values of c , κ and T . In all cases, the distribution functions are evaluated at 50 equidistant points z over the range $[-14, 5]$. We exclude the point $z = -c$ since the integral is ill-defined in this case; we consider instead the evaluation at some close neighbour.

4.2. The effects of the initial condition on the distributions

In this section, we consider the effect of variations in the initial condition on the finite sample distribution and the three approximations considered. To obtain the finite sample distribution, we used 30 000 replications of the statistic (3) when the model is generated by (1). We consider three values of the non-centrality parameter c , namely $c = -5, 0$ and 2 (covering the stationary, unit root and explosive cases). For the initial condition, we consider the values $\gamma (= y_0/T^{1/2}) = 0, 0.5, 1.0$ and 2.0 .

The results are presented in Figures 1–4 for the case $T = 25$. Consider first Figure 1 which presents the results for the exact distribution. As the initial condition increases, the exact distribution gets more concentrated around zero and the tails of the distributions, accordingly, get shorter. This concentration is more pronounced as c increases.

Consider now the effects on the asymptotic approximations. An increase in the initial condition has a similar effect on the continuous-time and Edgeworth approximations but the effect is clearly different on the approximation given by the stochastic expansion. The variations in

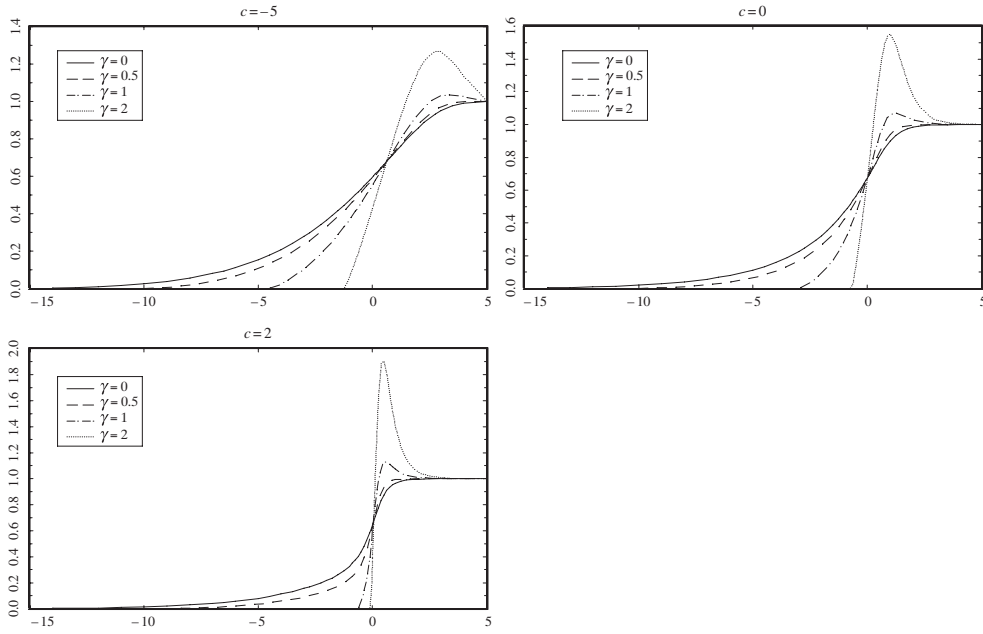


Figure 2. Effect of the initial condition for different values of c . Edgeworth expansion for $T = 25$.

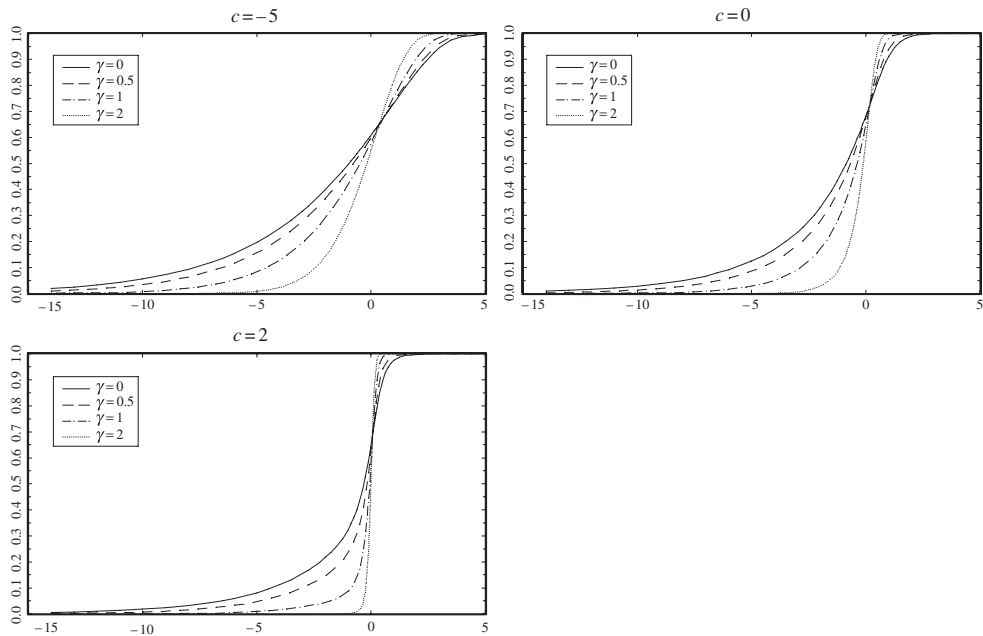


Figure 3. Effect of the initial condition for different values of c . Continuous-time approximation for $T = 25$.

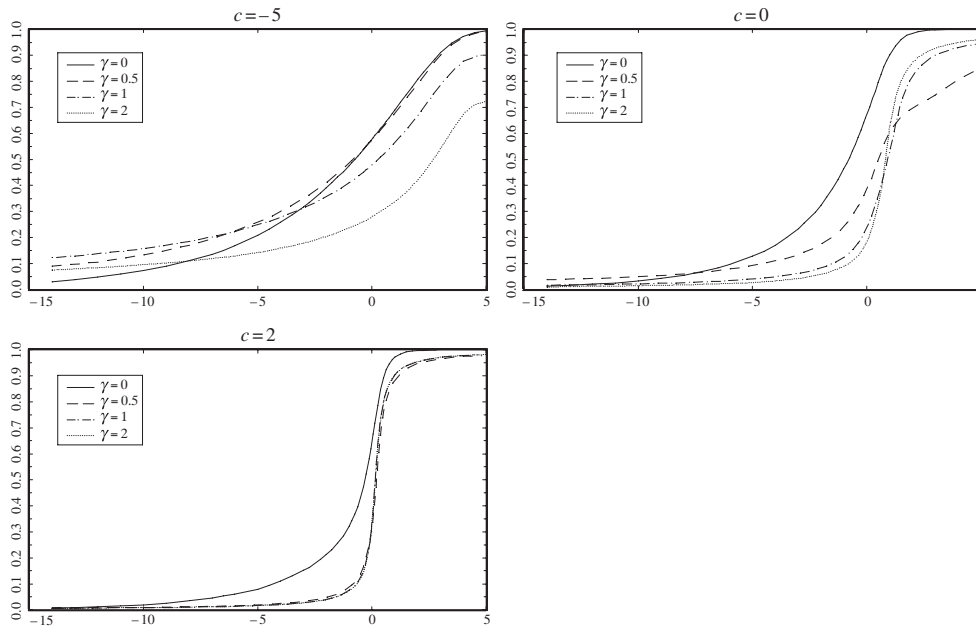


Figure 4. Effect of the initial condition for different values of c . Stochastic expansion for $T = 25$.

the continuous-time approximation mirror closely those of the exact distribution. On the other hand, when $\gamma \geq 1$, the approximation provided by the Edgeworth expansion deteriorates in the right tail and takes values greater than one so that the approximation no longer provides a proper distribution function. The variations, induced by changes in γ , on the approximation provided by the stochastic expansion are very different from those of the exact distribution. For a fixed c less than or equal to zero, we note a substantial increase in the dispersion of the distribution as γ increases (though some non-monotonicities are observed). When $c \geq 0$, we also observe a rightward shift as γ increases and, especially when $c = 2$, the distributions become invariant to γ as γ increases.

Note that, when evaluated at 0, all distribution functions are equal. At this point, the distribution functions are equivalent to the probability that the numerator X defined by (5) be less than or equal to 0. The continuous-time approximation provides no finite sample adjustment to this probability over the limiting distribution. It can be shown that the stochastic and Edgeworth expansions provide an adjustment to the numerator, over the first-order term, that is symmetric around 0 (for example, the term $\eta/\sqrt{2}$ with $\eta \sim N(0, 1)$ for the stochastic expansion). Hence, $Pr[X < 0]$ is the same in all three cases.

The effect of a change in the non-centrality parameter c on the exact and approximate distributions are similar to those induced by changes in the initial condition. For a given γ , these distributions (except for the stochastic expansion) become more concentrated around zero as c increases and this concentration becomes more important as γ increases.

4.3. Adequacy of the approximations

The quality of the approximations depends on the values of the parameters c and γ . To analyse this issue, we compare the approximations to the exact distribution for $T = 10, 25, 50$ and 100 and for the same values of c and γ considered above. The results are in Tables 1–4 and Figures 5–7. We consider in turn the cases where γ is 0 and non-zero.

Case with $\gamma = 0$. In the case where the initial condition is zero the continuous-time approximation is equivalent to the standard asymptotic distribution without any correction factor. This is not the case for the approximations based on the Edgeworth and the stochastic expansions which provide finite sample correction factors. The adequacy of the usual asymptotic distribution in the case of a zero initial condition was studied in detail in Perron (1989). He showed that it is not particularly adequate when $T = 10$ or 25 and $c = -5$ but that the adequacy increases when c increases. In particular, when $T = 25$ and $c = -1$, it is satisfactory in the right tail and excellent overall when $c = 5$ and $T = 10$. This study also showed that the usual limiting distribution is satisfactory even with $c < 0$ when T reaches 50.

As seen in Figures 5–7, the Edgeworth expansion, unlike the stochastic expansion, provides an improvement over the standard asymptotic (or continuous-time) distribution when $T = 10$ or 25 , especially in the locally stationary ($c < 0$) or integrated ($c = 0$) cases. When $T = 10$ and $c = -5$, the stochastic expansion is a bad approximation (Table 4) while the continuous-time and Edgeworth approximations are satisfactory in the right tail. The Edgeworth expansion is the best approximation in the left tail. When $T = 25$, the Edgeworth approximation is excellent, the stochastic expansion provides a slight improvement in the right tail but is not satisfactory in the left tail. As c increases while remaining negative (for example, $c = -2$ and -1), the continuous-time and stochastic expansion approximations remain satisfactory in the right tail but the Edgeworth expansion is best overall, especially if $T = 25$ (Figures 5–7). Similar conclusions apply with a unit root ($c = 0$). When the model is locally explosive ($c > 0$), the three approximations share the common feature of being excellent for $T = 25$ and $c = 2$ (Figure 7). However, for $T = 10$, the continuous-time and stochastic expansion approximations are not so good in the extreme tails. But these shortcomings disappear when c increases.

Case with $\gamma \neq 0$. When the initial condition is non-zero, the continuous-time approximation is, in general, the better one. Indeed, when $T \geq 25$, γ fixed and $c \leq 0$, the continuous-time approximation is satisfactory, especially in the right tail (Figures 5, 6). Furthermore, the quality of this approximation increases when c increases. For example, when $c = 2$ and $\gamma = 2$ (see Figure 7), it is excellent for $T \geq 10$. Similar conclusions apply when considering a fixed c and varying γ : i.e. a good approximation in the right tail and significant improvements in the left tail as γ increases. In particular, when $T = 50$, the continuous-time approximation is excellent for $c = 2$ and any γ .

The Edgeworth and stochastic expansions offer a contrasting picture, especially when $\gamma \geq 1$ (for all T considered). Figures 5–7 show that as the initial condition increases, these approximations seriously deteriorate. For moderate values (e.g. $\gamma = 0.5$), the stochastic expansion is still not a good approximation. On the other hand, the Edgeworth expansion is satisfactory in the right tail and seems to better capture the median of the exact distribution when $T = 10$ or 25 . When T increases (e.g. $T = 50$) with $c < 0$, the continuous-time and Edgeworth approximations are good in the right tail but somewhat lacking in the left. The quality of the approximation increases as c

Table 1. Percentage points of the exact distribution of $T(\hat{\alpha} - \alpha)$.

	1%	2.5%	5%	10%	90%	95%	97.5%	99%
$c = -5$								
$\gamma = 0.0$								
$T = 10$	-9.13	-7.66	-6.38	-4.96	2.33	2.91	3.48	4.25
$T = 25$	-12.83	-10.23	-8.36	-6.40	2.40	3.00	3.44	4.03
$T = 50$	-14.67	-11.68	-9.44	-7.04	2.43	3.00	3.46	3.97
$T = 100$	-15.76	-12.41	-9.91	-7.34	2.44	3.03	3.48	3.99
$\gamma = 0.5$								
$T = 10$	-7.75	-6.56	-5.44	-4.23	2.07	2.56	2.96	3.44
$T = 25$	-10.54	-8.71	-7.12	-5.42	2.17	2.72	3.14	3.58
$T = 50$	-12.06	-9.53	-7.84	-5.94	2.21	2.77	3.21	3.65
$T = 100$	-13.31	-10.49	-8.31	-6.27	2.21	2.78	3.19	3.61
$\gamma = 1.0$								
$T = 10$	-5.64	-4.76	-3.94	-3.05	1.69	2.08	2.40	2.71
$T = 25$	-7.41	-6.05	-5.00	-3.88	1.82	2.26	2.61	2.93
$T = 50$	-8.49	-6.93	-5.62	-4.23	1.86	2.34	2.68	3.04
$T = 100$	-8.97	-7.19	-5.96	-4.46	1.87	2.34	2.71	3.06
$\gamma = 2.0$								
$T = 10$	-3.28	-2.71	-2.25	-1.73	1.17	1.46	1.69	1.94
$T = 25$	-4.05	-3.31	-2.71	-2.10	1.28	1.59	1.87	2.15
$T = 50$	-4.46	-3.68	-3.03	-2.32	1.32	1.66	1.94	2.23
$T = 100$	-4.70	-3.83	-3.12	-2.39	1.33	1.68	1.96	2.25
$c = 0$								
$\gamma = 0.0$								
$T = 10$	-9.92	-8.07	-6.59	-4.91	1.17	1.63	2.11	2.78
$T = 25$	-11.82	-9.41	-7.40	-5.35	1.01	1.41	1.79	2.32
$T = 50$	-12.58	-9.93	-7.65	-5.52	0.97	1.34	1.68	2.11
$T = 100$	-13.36	-10.26	-7.93	-5.68	0.96	1.31	1.65	2.08
$\gamma = 0.5$								
$T = 10$	-7.89	-6.39	-5.23	-3.90	0.86	1.19	1.51	1.92
$T = 25$	-9.59	-7.58	-5.90	-4.35	0.79	1.10	1.37	1.77
$T = 50$	-9.95	-7.97	-6.13	-4.45	0.80	1.09	1.35	1.69
$T = 100$	-10.46	-8.16	-6.42	-4.59	0.77	1.05	1.31	1.61
$\gamma = 1.0$								
$T = 10$	-5.25	-4.21	-3.36	-2.47	0.63	0.84	1.00	1.22
$T = 25$	-5.89	-4.64	-3.68	-2.65	0.60	0.79	0.96	1.17
$T = 50$	-6.56	-5.13	-4.00	-2.79	0.60	0.80	0.96	1.16
$T = 100$	-6.63	-5.18	-4.06	-2.89	0.59	0.78	0.94	1.13

Table 1—continued

	1%	2.5%	5%	10%	90%	95%	97.5%	99%
$\gamma = 2.0$								
$T = 10$	-2.43	-1.87	-1.46	-1.06	0.43	0.55	0.65	0.76
$T = 25$	-2.59	-2.00	-1.57	-1.10	0.42	0.53	0.63	0.74
$T = 50$	-2.72	-2.06	-1.58	-1.13	0.42	0.53	0.63	0.74
$T = 100$	-2.72	-2.11	-1.61	-1.15	0.41	0.53	0.62	0.72
$c = 2$								
$\gamma = 0.0$								
$T = 10$	-10.06	-7.94	-6.17	-4.28	0.74	1.10	1.48	2.11
$T = 25$	-11.22	-8.42	-6.31	-4.30	0.59	0.88	1.18	1.60
$T = 50$	-11.77	-8.74	-6.53	-4.37	0.55	0.81	1.08	1.43
$T = 100$	-11.74	-8.77	-6.55	-4.39	0.55	0.80	1.07	1.42
$\gamma = 0.5$								
$T = 10$	-7.60	-5.80	-4.46	-2.83	0.46	0.66	0.87	1.16
$T = 25$	-8.42	-6.30	-4.69	-2.96	0.41	0.58	0.77	1.05
$T = 50$	-8.95	-6.49	-4.75	-3.02	0.40	0.57	0.75	1.01
$T = 100$	-9.21	-6.60	-4.80	-3.00	0.38	0.55	0.72	0.96
$\gamma = 1.0$								
$T = 10$	-4.36	-2.56	-1.42	-0.79	0.29	0.39	0.48	0.61
$T = 25$	-4.31	-2.87	-1.62	-0.83	0.26	0.36	0.45	0.56
$T = 50$	-4.70	-3.04	-1.68	-0.86	0.26	0.35	0.43	0.54
$T = 100$	-5.04	-3.16	-1.74	-0.89	0.25	0.34	0.42	0.53
$\gamma = 2.0$								
$T = 10$	-0.56	-0.44	-0.34	-0.25	0.17	0.21	0.26	0.31
$T = 25$	-0.59	-0.43	-0.33	-0.24	0.15	0.20	0.24	0.29
$T = 50$	-0.60	-0.45	-0.34	-0.24	0.15	0.20	0.24	0.28
$T = 100$	-0.60	-0.42	-0.33	-0.24	0.15	0.19	0.23	0.27

increases, most notably for the continuous-time approximation which is excellent when $c = -1$ or $c \geq 0$.

Some analytical explanations. It is instructive to consider the behaviour of the various approximations in the case of a large initial condition, i.e. a large value of γ (which is here equivalent to a small- σ approximation). From Perron (1991a), we have that, when γ is large, the continuous-time approximation implies the approximation

$$T(\hat{\alpha} - \alpha) \stackrel{a}{\sim} N\left(0, \frac{2c}{\gamma^2(\exp(2c) - 1)}\right). \quad (20)$$

This approximation is consistent with the main features of the finite sample distribution. In particular, it implies that as γ increases the distribution of $T(\hat{\alpha} - \alpha)$ becomes more symmetric (due

Table 2. Percentage points from the Edgeworth approximation.

	1%	2.5%	5%	10%	90%	95%	97.5%	99%
$c = -5$								
$\gamma = 0.0$								
$T = 10$	-7.34	-6.71	-5.90	-4.75	2.28	2.82	3.26	3.91
$T = 25$	-11.84	-9.98	-8.26	-6.31	2.38	2.97	3.43	4.00
$T = 50$	-14.24	-11.53	-9.32	-6.98	2.42	3.02	3.48	4.02
$T = 100$	-15.52	-12.35	-9.88	-7.34	2.44	3.04	3.50	4.03
$\gamma = 0.5$								
$T = 10$	-6.42	-5.93	-5.28	-4.30	2.10	2.56	2.89	3.19
$T = 25$	-9.07	-7.99	-6.82	-5.34	2.19	2.69	3.06	3.41
$T = 50$	-10.55	-9.01	-7.52	-5.78	2.22	2.74	3.11	3.48
$T = 100$	-11.48	-9.60	-7.91	-6.02	2.23	2.76	3.14	3.51
$\gamma = 1.0$								
$T = 10$	-3.35	-3.19	-2.95	-2.51	1.74	2.04	2.21	2.33
$T = 25$	-4.18	-3.95	-3.61	-3.04	1.78	2.10	2.30	2.43
$T = 50$	-4.56	-4.29	-3.90	-3.26	1.79	2.13	2.33	2.47
$T = 100$	-4.82	-4.52	-4.09	-3.40	1.80	2.14	2.35	2.49
$\gamma = 2.0$								
$T = 10$	-1.10	-1.06	-0.98	-0.84	1.12	1.25	1.31	1.35
$T = 25$	-1.19	-1.14	-1.06	-0.91	1.15	1.28	1.35	1.39
$T = 50$	-1.22	-1.17	-1.09	-0.94	1.16	1.29	1.36	1.41
$T = 100$	-1.24	-1.19	-1.11	-0.95	1.16	1.30	1.37	1.42
$c = 0.0$								
$\gamma = 0.0$								
$T = 10$	-9.42	-7.92	-6.50	-4.87	1.14	1.57	1.98	2.48
$T = 25$	-11.77	-9.32	-7.36	-5.35	1.01	1.40	1.77	2.25
$T = 50$	-12.71	-9.86	-7.69	-5.52	0.97	1.34	1.69	2.15
$T = 100$	-13.21	-10.14	-7.86	-5.61	0.95	1.31	1.65	2.09
$\gamma = 0.5$								
$T = 10$	-6.60	-5.80	-4.90	-3.77	0.89	1.25	1.61	2.12
$T = 25$	-8.04	-6.74	-5.50	-4.10	0.79	1.08	1.34	1.68
$T = 50$	-8.73	-7.15	-5.74	-4.22	0.76	1.03	1.26	1.53
$T = 100$	-9.13	-7.37	-5.87	-4.29	0.74	1.00	1.22	1.47
$c = 2$								
$\gamma = 0.0$								
$T = 10$	-9.03	-7.46	-5.82	-4.05	0.70	1.01	1.34	1.80
$T = 25$	-10.83	-8.25	-6.23	-4.22	0.59	0.87	1.15	1.56

Table 2—continued

	1%	2.5%	5%	10%	90%	95%	97.5%	99%
$T = 50$	-11.42	-8.55	-6.39	-4.28	0.56	0.81	1.08	1.46
$T = 100$	-11.73	-8.71	-6.47	-4.31	0.54	0.79	1.04	1.41
$\gamma = 0.5$								
$T = 10$	-6.03	-5.00	-3.92	-2.62	0.47	0.71	0.99	1.43
$T = 25$	-6.88	-5.46	-4.15	-2.68	0.39	0.55	0.72	0.97
$T = 50$	-7.27	-5.65	-4.24	-2.70	0.36	0.51	0.64	0.82
$T = 100$	-7.48	-5.75	-4.28	-2.71	0.35	0.49	0.61	0.76
$\gamma = 1.0$								
$T = 10$	-0.63	-0.58	-0.52	-0.42	0.21	0.24	0.27	0.28
$T = 25$	-0.62	-0.58	-0.52	-0.42	0.19	0.22	0.24	0.26
$T = 50$	-0.61	-0.57	-0.51	-0.42	0.18	0.22	0.24	0.25
$T = 100$	-0.60	-0.56	-0.50	-0.41	0.18	0.21	0.23	0.24

Table 3. Percentage points from the continuous-time approximation.

	1%	2.5%	5%	10%	90%	95%	97.5%	99%
$c = -5$								
$\gamma = 0.0$	-16.74	-13.17	-10.46	-7.72	2.46	3.06	3.52	4.03
0.5	-13.90	-10.98	-8.76	-6.49	2.27	2.83	3.24	3.68
1.0	-9.54	-7.61	-6.11	-4.57	1.93	2.40	2.76	3.12
2.0	-4.87	-3.95	-3.22	-2.44	1.38	1.74	2.01	2.30
$c = 0.0$								
$\gamma = 0.0$	-13.69	-10.44	-8.04	-5.71	0.93	1.29	1.61	2.03
0.5	-10.96	-8.35	-6.43	-4.57	0.76	1.05	1.30	1.63
1.0	-6.85	-5.22	-4.02	-2.86	0.59	0.78	0.94	1.13
2.0	-2.74	-2.09	-1.61	-1.15	0.41	0.53	0.62	0.73
$c = 2$								
$\gamma = 0.0$	-12.06	-8.88	-6.56	-4.34	0.52	0.76	1.01	1.36
0.5	-9.26	-6.71	-4.84	-3.02	0.38	0.54	0.71	0.97
1.0	-4.88	-3.13	-1.74	-0.88	0.25	0.34	0.42	0.53
2.0	-0.59	-0.43	-0.33	-0.24	0.15	0.19	0.23	0.27

to the normality) and more concentrated around zero (since the variance decreases). Also, the approximation (20) is consistent with the finite sample feature that this concentration is more rapid (as γ increases) the larger c is.

On the other hand, the stochastic expansion (11) yields the following approximation for a large γ :

$$T(\hat{\alpha} - \alpha) \stackrel{a}{\sim} \frac{\int_0^1 \exp(cr) dW(r)}{2 \int_0^1 \exp(cr) J_c(r) dr}.$$

Table 4. Percentage points from the stochastic expansion.

	1%	2.5%	5%	10%	90%	95%	97.5%	99%
$c = -5.0$								
$\gamma = 0.0$								
$T = 10$	-22.16	-17.17	-13.30	-9.39	3.57	4.39	5.12	6.25
$T = 25$	-19.14	-14.92	-11.77	-8.50	2.89	3.53	4.05	4.64
$T = 50$	-17.82	-14.10	-11.20	-8.18	2.65	3.28	3.75	4.28
$T = 100$	-17.01	-13.59	-10.89	-7.98	2.54	3.17	3.61	4.12
$\gamma = 0.5$								
$T = 10$	-58.66	-33.86	-22.55	-13.48	3.63	4.43	5.19	6.56
$T = 25$	-54.92	-31.98	-21.26	-12.64	2.99	3.62	4.14	4.74
$T = 50$	-52.81	-31.49	-20.77	-12.39	2.78	3.39	3.85	4.39
$T = 100$	-52.91	-30.96	-20.35	-12.27	2.69	3.28	3.72	4.24
$\gamma = 1.0$								
$T = 10$	-188.37	-75.35	-38.23	-18.02	6.16	35.91	76.95	188.23
$T = 25$	-184.23	-75.75	-37.73	-17.69	4.74	36.85	77.01	188.18
$T = 50$	-181.85	-77.18	-37.24	-17.68	4.44	37.42	76.04	191.98
$T = 100$	-179.39	-77.62	-37.47	-17.67	4.32	37.62	76.14	195.74
$\gamma = 2.0$								
$T = 10$	-123.89	-47.98	-22.87	-9.39	17.66	31.82	59.99	143.58
$T = 25$	-122.39	-47.04	-22.76	-9.33	17.75	31.48	60.07	143.88
$T = 50$	-124.83	-47.22	-22.85	-9.43	17.74	31.76	59.64	143.51
$T = 100$	-124.66	-47.23	-23.07	-9.40	17.80	31.80	60.38	144.05
$c = 0.0$								
$\gamma = 0.0$								
$T = 10$	-16.01	-11.98	-8.86	-6.03	1.07	1.48	1.90	2.40
$T = 25$	-14.73	-11.15	-8.40	-5.87	0.98	1.37	1.71	2.18
$T = 50$	-14.21	-10.94	-8.23	-5.84	0.96	1.34	1.66	2.13
$T = 100$	-14.06	-10.79	-8.13	-5.83	0.95	1.33	1.64	2.09
$\gamma = 0.5$								
$T = 10$	-58.53	-21.45	-10.19	-4.56	7.07	12.75	23.38	57.00
$T = 25$	-58.83	-21.28	-10.04	-4.51	6.87	12.38	22.83	55.02
$T = 50$	-60.09	-21.42	-10.00	-4.47	6.79	12.36	22.66	53.72
$T = 100$	-58.02	-21.45	-10.04	-4.46	6.80	12.31	22.87	54.12
$\gamma = 1.0$								
$T = 10$	-24.41	-9.14	-4.03	-1.49	3.06	5.30	9.92	24.20
$T = 25$	-24.34	-8.91	-3.99	-1.48	3.01	5.22	9.83	23.68
$T = 50$	-24.39	-8.86	-3.96	-1.47	3.00	5.21	9.71	23.33
$T = 100$	-24.78	-8.88	-3.90	-1.47	2.98	5.23	9.69	23.43

Table 4—continued

	1%	2.5%	5%	10%	90%	95%	97.5%	99%
$\gamma = 2.0$								
$T = 10$	-15.09	-5.55	-2.50	-0.80	2.37	4.06	8.03	19.16
$T = 25$	-15.19	-5.56	-2.48	-0.80	2.36	3.99	7.83	19.31
$T = 50$	-15.26	-5.57	-2.44	-0.80	2.35	3.99	7.75	19.26
$T = 100$	-15.28	-5.56	-2.43	-0.80	2.34	4.01	7.78	19.22
$c = 2.0$								
$\gamma = 0.0$								
$T = 10$	-13.44	-9.48	-6.81	-4.24	0.55	0.80	1.11	1.52
$T = 25$	-12.74	-9.13	-6.71	-4.31	0.53	0.78	1.04	1.43
$T = 50$	-12.29	-9.15	-6.70	-4.35	0.53	0.77	1.02	1.39
$T = 100$	-12.10	-9.06	-6.69	-4.36	0.53	0.76	1.02	1.38
$\gamma = 0.5$								
$T = 10$	-10.18	-3.92	-1.89	-0.83	1.14	2.14	4.24	10.30
$T = 25$	-9.90	-3.90	-1.87	-0.82	1.13	2.14	4.19	9.99
$T = 50$	-9.75	-3.90	-1.86	-0.82	1.13	2.13	4.20	9.94
$T = 100$	-9.78	-3.88	-1.85	-0.82	1.13	2.13	4.18	9.88
$\gamma = 1.0$								
$T = 10$	-8.66	-3.35	-1.58	-0.69	0.96	1.85	3.52	9.17
$T = 25$	-8.50	-3.34	-1.59	-0.69	0.95	1.84	3.51	9.16
$T = 50$	-8.60	-3.33	-1.59	-0.69	0.95	1.85	3.50	9.14
$T = 100$	-8.57	-3.32	-1.59	-0.69	0.96	1.85	3.47	9.15
$\gamma = 2.0$								
$T = 10$	-8.61	-3.29	-1.54	-0.68	0.95	1.82	3.50	9.24
$T = 25$	-8.58	-3.28	-1.53	-0.67	0.95	1.82	3.50	9.27
$T = 50$	-8.56	-3.28	-1.54	-0.67	0.94	1.82	3.49	9.31
$T = 100$	-8.52	-3.28	-1.54	-0.67	0.94	1.81	3.50	9.33

This implies that as γ gets large the distribution of $T(\hat{\alpha} - \alpha)$ becomes less influenced by γ and, in the limit, invariant to it. As seen from the experimental results, this is not a feature of the distribution of $T(\hat{\alpha} - \alpha)$ in finite samples. If we had worked with the stochastic expansion that excluded all terms of order $O_p(T^{-1})$, i.e. considered the approximation

$$T(\hat{\alpha} - \alpha) \stackrel{a}{\sim} X_0 + X_1/T^{1/2}$$

with X_0 and X_1 defined by (14) and (15), things would have been even worse. Indeed, for large γ or large $\kappa = y_0/\sigma$ the approximation then becomes

$$\kappa \left[\frac{\int_0^1 \exp(cr) dW(r)}{\int_0^1 J_c(r)^2 dr} - 2 \frac{\int_0^1 \exp(cr) J_c(r) dr \int_0^1 J_c(r) dW(r)}{(\int_0^1 J_c(r)^2 dr)^2} \right]$$

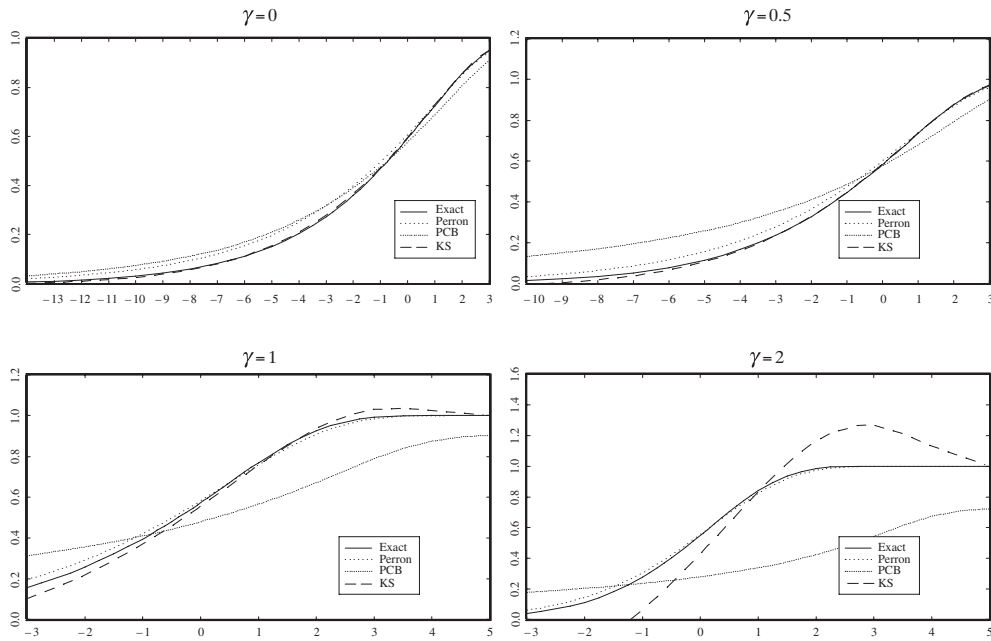


Figure 5. Approximations to the distribution of the normalized OLS coefficient. $T = 25$ and $c = -5$.

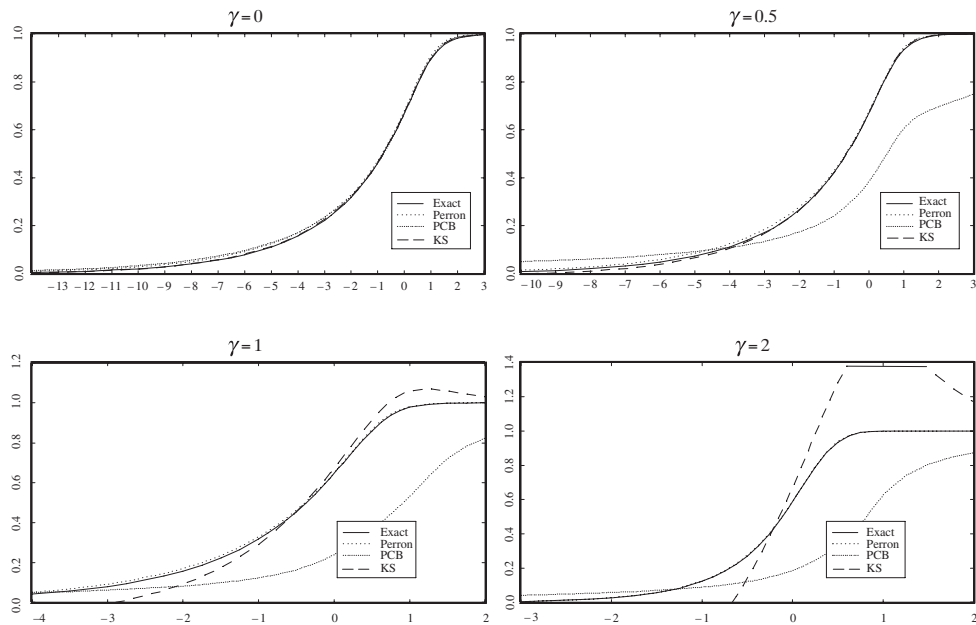


Figure 6. Approximations to the distribution of the normalized OLS coefficient. $T = 25$ and $c = 0$.

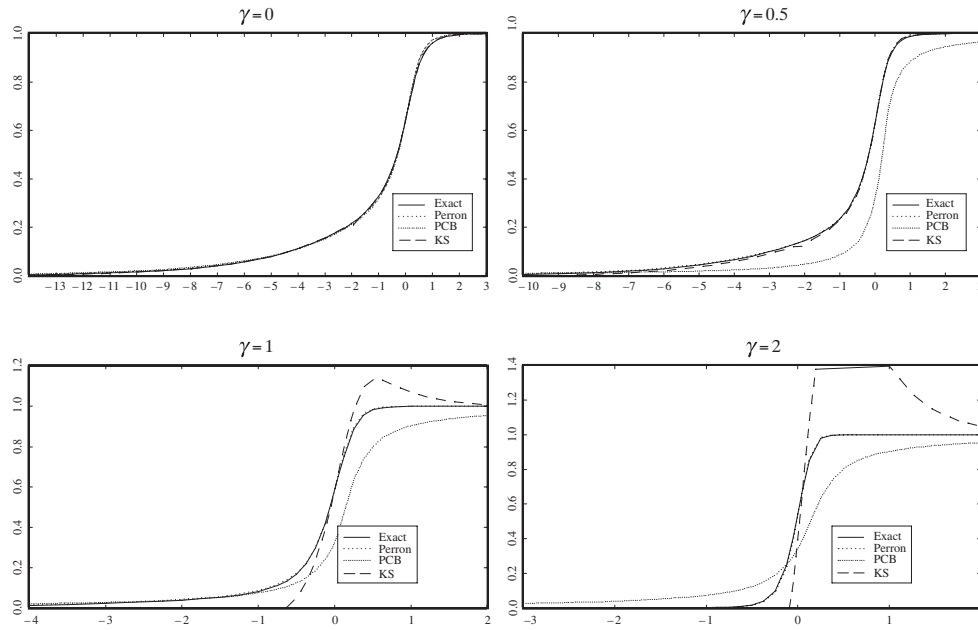


Figure 7. Approximations to the distribution of the normalized OLS coefficient. $T = 25$ and $c = 2$.

whose variance increases with an increase in the initial condition. This feature is the precise reverse of that present in the finite sample distribution of $T(\hat{\alpha} - \alpha)$.

5. CONCLUSION

To summarize our findings, the quality of the various approximations depends on the values of c and γ . When the initial condition is zero, the continuous-time approximation which corresponds to the standard asymptotic distribution is excellent for $c > 0$ as soon as $T = 25$, and even with as low a sample size as $T = 10$ when c increases. In this case, the additional corrections provided by the Edgeworth and stochastic expansions are negligible. However, when $c \leq 0$, the Edgeworth expansion provides a significant improvement over the standard asymptotic distribution, especially when $T = 25$. When the initial condition is non-zero, the continuous-time approximation is clearly the better approximation and more so as γ increases.

As mentioned in the introduction, the autoregressive model of order one, given that it is simple enough to yield tractable solutions, is something of a base case by which we can learn useful features about the properties of alternative asymptotic refinements that can be helpful in guiding our efforts in more complex and empirically relevant models. The results in this paper point to the fact that the continuous-time asymptotic framework appears to be an interesting avenue to explore when considering approximations to statistics in more complex models that deal with data that are substantially correlated. There are numerous possible applications where the finite sample distribution is not tractable yet the continuous-time asymptotic distribution is relatively easy to derive. This approach was successfully used in Perron and Vodounou (2000a,b) where

we analyse the power functions of several tests of market efficiency for stock returns when the relevant alternative is one where a series such as the dividend–price ratio (which is highly correlated and modelled as a nearly integrated process) explains returns. The asymptotic power functions derived yield useful insights about the effect of variations in some parameters (for example, the horizon over which returns are computed) on the power of the tests. Another example pertains to the analysis of the estimates of cointegrating vectors with mixed stock and flow data. Chambers (2001) uses a continuous-time framework which delivers interesting practical results. Continuous-time asymptotic approximations have also been applied to financial problems; see the collected work of Nelson (Rossi, 1996). No doubt many such applications of interest await further investigation along these lines.

ACKNOWLEDGEMENTS

This paper is drawn from chapter 1 of Vodounou's Ph.D. Dissertation at the Université de Montréal. Perron acknowledges financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC), the Natural Sciences and Engineering Council of Canada (NSERC), and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR). We also wish to thank John Knight for useful conversations and for providing us with the draft notes referred to as Knight and Satchell (1990).

REFERENCES

- Bergstrom, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In Z. Griliches and M. D. Intriligator (eds), *The Handbook of Econometrics*, vol. 2, pp. 1146–210. Amsterdam: North-Holland.
- Bobkoski, M. J. (1983). *Hypothesis Testing in Nonstationary Time Series*. Unpublished Ph.D. Dissertation, Department of Statistics, University of Wisconsin.
- Cavanagh, C. (1986). Roots local to unity. Manuscript, Harvard University.
- Chambers, M. J. (2001). Cointegration and sampling frequency. Manuscript, University of Essex.
- Chan, N. H. and C. Z. Wei (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *The Annals of Statistics* 15, 1050–63.
- 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, 427–31.
- Evans, G. B. and N. E. Savin (1981a). The calculation of the least squares estimator of the parameter in a random walk model. *The Annals of Statistics* 9, 1114–18.
- Evans, G. B. and N. E. Savin (1981b). Testing for unit roots: 1. *Econometrica* 49, 753–78.
- Fuller, W. A. (1976). *Introduction to Statistical Time Series*. New York: Wiley.
- Gurland, J. (1948). Inversion formulae for the distribution of ratios. *Annals of Mathematical Statistics* 19, 228–37.
- Hisamatsu, H. and K. Maekawa (1994). The distribution of the Durbin–Watson statistic in integrated and near-integrated models. *Journal of Econometrics* 61, 367–82.
- Knight, J. L. and S. E. Satchell (1990). Edgeworth expansion of the least squares estimator in the near integrated model. Unpublished draft notes.
- Knight, J. L. and S. E. Satchell (1993). Asymptotic expansions for random walks with normal errors. *Econometric Theory* 9, 363–76.

- Larson, R. (1995). The asymptotic distributions of some test statistics in near-integrated AR processes. *Econometric Theory* 11, 306–30.
- Mann, H. B. and A. Wald (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica* 11, 173–220.
- Perron, P. (1989). The calculation of the limiting distribution of the least squares estimator in a near-integrated model. *Econometric Theory* 5, 241–55.
- Perron, P. (1991a). A continuous time approximation to the unstable first-order autoregressive process: the case without an intercept. *Econometrica* 59, 211–36.
- Perron, P. (1991b). A continuous time approximation to the stationary first-order autoregressive model. *Econometric Theory* 7, 236–52.
- Perron, P. (1996). The adequacy of asymptotic approximations in the near-integrated autoregressive model with dependent errors. *Journal of Econometrics* 70, 317–50.
- Perron, P. and C. Vodounou (2000a). Tests of market efficiency: an analysis of their properties based on a continuous time asymptotic framework. Unpublished manuscript, Department of Economics, Boston University.
- Perron, P. and C. Vodounou (2000b). The variance ratio test: an analysis of size and power based on a continuous time asymptotic framework. Unpublished manuscript, Department of Economics, Boston University.
- Phillips, P. C. B. (1977). Approximations to some finite sample distributions associated with a first order stochastic difference equation. *Econometrica* 45, 463–85.
- 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 P. Perron (1988). Testing for a unit root in time series regression. *Biometrika* 75, 335–46.
- Rossi, P. E., ed. (1996). *Modelling Stock Market Volatility: Bridging the Gap to Continuous Time*. New York: Academic.
- Rubin, H. (1950). Consistency of maximum likelihood estimates in the explosive case. In T. C. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models*. New York: Wiley.
- Satchell, S. E. (1984). Approximation to the finite sample distribution for nonstable first order stochastic difference equations. *Econometrica* 52, 1271–89.
- Tanaka, K. (1990). The Fredholm approach to asymptotic inference on nonstationary and noninvertible time series models. *Econometric Theory* 6, 411–32.
- Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. New York: Wiley.
- Vodounou, C. (1997a). *Essais en économétrie théorique et en économétrie de la finance appliqués aux modèles dynamiques*. Unpublished Ph.D. Dissertation, Département de Sciences Économiques, Université de Montréal.
- Vodounou, C. (1997b). Adéquation des approximations d'Edgeworth dans le modèle AR(1) stable avec valeur initiale fixe. Unpublished manuscript, Université de Montréal.
- White, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Annals of Mathematical Statistics* 29, 1188–97.

MATHEMATICAL APPENDIX

Proof of Theorem 1. As a matter of notation, we let $\stackrel{a}{\approx}$ denote equality up to terms that are of order $O(T^{-a})$. We start with the expansion of the term $D(T)$ given by equation (8). Since r and s are the solutions to the equation $x^2 - px + q^2 = 0$, we have

$$\begin{aligned}
 r, s = & \left(1 + \alpha^2 + 2\alpha \frac{u}{T} - 2\frac{v}{T^2} \right) / 2 \\
 & \pm \left((1 - \alpha^2)^2 - 4\alpha(1 - \alpha^2)\frac{u}{T} - 4(1 - \alpha^2)\frac{u^2}{T^2} - 4(1 + \alpha^2)\frac{v}{T^2} \right. \\
 & \left. - 8\alpha \frac{uv}{T^3} + 4\frac{v^2}{T^4} \right)^{1/2} / 2
 \end{aligned} \tag{A.1}$$

Since $\alpha = e^{c/T}$,

$$\alpha \stackrel{3}{\approx} 1 + c/T + c^2/(2T^2), \tag{A.2}$$

and

$$\alpha^2 \stackrel{3}{\approx} 1 + 2c/T + 2c^2/T^2. \tag{A.3}$$

Substituting these expansions, we have

$$\begin{aligned}
 r \stackrel{2}{\approx} & 1 + \frac{c + u}{T} + \frac{c^2 + cu - v}{T^2} \\
 & + \frac{i\sqrt{2}}{T} [(v - uc - c^2/2) + (uv + vc - u^2c - 2uc^2 - c^3)/T]^{1/2},
 \end{aligned}$$

and s the complex conjugate of r . Using the notation

$$\begin{aligned}
 r_1 &= c + u, \\
 r_2 &= c(c + u) - v = cr_1 - v, \\
 \delta^2 &= 2v - 2cu - c^2 = c^2 - 2r_2,
 \end{aligned}$$

and the definition

$$r = a + ib = |r| \cos \theta + i|r| \sin \theta,$$

we obtain

$$|r| \cos \theta \stackrel{3}{\approx} 1 + \frac{r_1}{T} + \frac{r_2}{T^2},$$

and

$$\begin{aligned}
 |r| \sin \theta & \stackrel{2}{\approx} \frac{\sqrt{2}}{T} [\delta^2/2 + (uv + vc - u^2c - 2uc^2 - c^3)/T]^{1/2} \\
 & \stackrel{2}{\approx} \frac{\delta}{T} \left[1 - \frac{2r_1r_2}{\delta^2T} \right]^{1/2}.
 \end{aligned}$$

Using de Moivre's formula, we deduce that

$$r^T = |r|^T \cos(T\theta) + i|r|^T \sin(T\theta).$$

Note, from (8), that we can write $D(T)$ as

$$D(T) = 2\operatorname{Re}\left(\frac{1 - \bar{r}}{r - \bar{r}}r^T\right),$$

where \bar{r} is the complex conjugate of r . After some developments we can write

$$D(T) = |r|^T (\cos(T\theta) + X \sin(T\theta)),$$

where

$$X = (1 - |r| \cos \theta) / (|r| \sin \theta).$$

We now consider the expansion of X and θ . We have

$$\begin{aligned} X &\stackrel{\cong}{=} -\frac{T}{\delta} \left(\frac{r_1}{T} + \frac{r_2}{T^2} \right) \left(1 - \frac{2r_1r_2}{\delta^2 T} \right)^{-1/2} \\ &\stackrel{\cong}{=} -\frac{r_1}{\delta} \left(1 + \frac{r_2}{r_1 T} \right) \left(1 + \frac{r_1r_2}{\delta^2 T} \right) \\ &\stackrel{\cong}{=} -\frac{1}{\delta} \left[r_1 + \frac{r_2}{\delta^2 T} (r_1^2 + \delta^2) \right], \end{aligned}$$

and

$$\begin{aligned} \theta &= \tan^{-1} \left[\left(\frac{\delta}{T} \left(1 - \frac{2r_1r_2}{\delta^2 T} \right)^{1/2} \right) / \left(1 + \frac{r_1}{T} + \frac{r_2}{T^2} \right) \right] \\ &= \frac{\delta}{T} \left(1 - \frac{r_1r_2}{\delta^2 T} \right) \left(1 - \frac{r_1}{T} - \frac{r_2}{T^2} \right) \\ &\stackrel{\cong}{=} \frac{\delta}{T} \left(1 - \frac{r_1}{\delta^2 T} (\delta^2 + r_2) \right). \end{aligned}$$

Hence,

$$\cos(T\theta) \stackrel{\cong}{=} \cos \left(\delta - \frac{r_1}{\delta T} (\delta^2 + r_2) \right) = \cos \delta + \frac{r_1}{\delta T} (\delta^2 + r_2) \sin \delta,$$

and

$$\sin(T\theta) \stackrel{\cong}{=} \sin \left(\delta - \frac{r_1}{\delta T} (\delta^2 + r_2) \right) = \sin \delta - \frac{r_1}{\delta T} (\delta^2 + r_2) \cos \delta.$$

Since $|r| = ((|r| \cos \theta)^2 + (|r| \sin \theta)^2)^{1/2}$, we also have

$$\begin{aligned} |r| &\stackrel{\cong}{=} \left[\left(1 + \frac{r_1}{T} + \frac{r_2}{T^2} \right)^2 + \frac{\delta^2}{T^2} \left(1 - \frac{2r_1r_2}{\delta^2 T} \right) \right]^{1/2} \\ &\stackrel{\cong}{=} \left[1 + \frac{2r_1}{T} + \frac{r_1^2 + 2r_2 + \delta^2}{T^2} \right]^{1/2} \\ &\stackrel{\cong}{=} 1 + \frac{r_1}{T} + \frac{2r_2 + \delta^2}{2T^2}. \end{aligned}$$

It follows that

$$\begin{aligned} |r|^T &\stackrel{2}{=} \exp\left[T \ln\left(1 + \frac{r_1}{T} + \frac{2r_2 + \delta^2}{2T^2}\right)\right] \\ &\stackrel{2}{=} \exp\left[T\left(\frac{r_1}{T} + \frac{2r_2 + \delta^2}{2T^2} - \frac{r_1^2}{2T^2}\right)\right] \\ &\stackrel{2}{=} \exp(r_1) \exp\left(\frac{2r_2 + \delta^2 - r_1^2}{2T}\right) \\ &\stackrel{2}{=} \exp(r_1) \left[1 + \frac{2r_2 + \delta^2 - r_1^2}{2T}\right]. \end{aligned}$$

It remains to consider the expression $\cos(T\theta) + X \sin(T\theta)$. Using the results above,

$$\begin{aligned} \cos(T\theta) + X \sin(T\theta) &\stackrel{3}{=} \cos \delta + \frac{r_1}{\delta T} (\delta^2 + r_2) \sin \delta \\ &\quad - \frac{1}{\delta} \left[r_1 + \frac{r_2}{\delta^2 T} (r_1^2 + \delta^2) \right] \left[\sin \delta - \frac{r_1}{\delta T} (\delta^2 + r_2) \cos \delta \right]. \end{aligned}$$

Expanding this last expression leads to

$$\cos(T\theta) + X \sin(T\theta) \stackrel{2}{=} A_1 - \frac{A_2}{\delta^2 T},$$

with

$$\begin{aligned} A_1 &= \cos \delta - r_1 \frac{\sin \delta}{\delta}, \\ A_2 &= (r_2(r_1^2 + \delta^2) - \delta^2 r_1(\delta^2 + r_2)) \frac{\sin \delta}{\delta} - r_1^2(\delta^2 + r_2) \cos \delta. \end{aligned}$$

We finally get

$$\begin{aligned} D(T) &\stackrel{3}{=} \exp(r_1) \left[1 + \frac{2r_2 + \delta^2 - r_1^2}{2T}\right] \left[A_1 - \frac{A_2}{\delta^2 T}\right] \\ &\stackrel{2}{=} \exp(r_1) \left\{ A_1 - \frac{1}{2T\delta^2} (2A_2 - \delta^2 A_1(2r_2 + \delta^2 - r_1^2)) \right\}, \end{aligned}$$

and we deduce that

$$D(T)^{-1/2} \stackrel{2}{=} A_1^{-1/2} \exp(-r_1/2) \left\{ 1 + \frac{1}{4T\delta^2 A_1} (2A_2 - \delta^2 A_1(2r_2 + \delta^2 - r_1^2)) \right\}.$$

Expanding this expression, we have

$$\begin{aligned} D(T)^{-1/2} &\stackrel{2}{=} \exp(-r_1/2) \left(\cos \delta - r_1 \frac{\sin \delta}{\delta} \right)^{-1/2} \\ &\quad \times \left\{ 1 + \frac{(r_1^2 + \delta^2)(2r_2 - r_1\delta^2)}{4T\delta^2 (\cos \delta - r_1 \frac{\sin \delta}{\delta})} \frac{\sin \delta}{\delta} - \frac{(r_1^2 + \delta^2)(2r_2 + \delta^2)}{4T\delta^2 (\cos \delta - r_1 \frac{\sin \delta}{\delta})} \cos \delta \right\}. \end{aligned} \tag{A.4}$$

We now consider the component of the moment generating function involving $y_0 = \kappa$ which is

$$\Psi = \exp\left(\kappa^2 \frac{v}{T^2} - \kappa^2 \frac{u}{T} \alpha - \frac{\kappa^2 \alpha^2}{2}\right) \exp\left(\left(\frac{u}{T} + \alpha\right)^2 \frac{\kappa^2}{2} \frac{D(T-1)}{D(T)}\right).$$

Substituting α and α^2 by the approximations (A.2) and (A.3), we have

$$\exp\left(\kappa^2 \frac{v}{T^2} - \kappa^2 \frac{u}{T} \alpha - \frac{\kappa^2 \alpha^2}{2}\right) \stackrel{\approx}{=} \exp\left(-\frac{\kappa^2}{2}\right) \left\{1 - \frac{\kappa^2 r_1}{T}\right\},$$

and

$$\exp\left(\left(\frac{u}{T} + \alpha\right)^2 \frac{\kappa^2}{2} \frac{D(T-1)}{D(T)}\right) \stackrel{\approx}{=} \exp\left(\frac{\kappa^2}{2}\right) \left\{1 + \frac{\kappa^2(h + 2r_1)}{2T}\right\},$$

supposing that

$$D(T-1)/D(T) \stackrel{\approx}{=} 1 + h/T,$$

which is to be shown below. Hence, we have

$$\Psi \stackrel{\approx}{=} 1 + \frac{\kappa^2 h}{2T}.$$

To determine h , we first expand $D(T-1)$ as follows:

$$\begin{aligned} D(T-1) &= |r|^{T-1} (\cos((T-1)\theta) + X \sin((T-1)\theta)) \\ &= |r|^{T-1} [\cos(T\theta) + \theta \sin(T\theta) + X(\sin(T\theta) - \theta \cos(T\theta))] \\ &= |r|^{T-1} \{(1 - \theta X) \cos(T\theta) + (\theta + X) \sin(T\theta)\}. \end{aligned}$$

The expansions of $(1 - \theta X)$ and $(\theta + X)$ are, respectively,

$$1 - \theta X \stackrel{\approx}{=} 1 + \frac{r_1}{T},$$

and

$$\theta + X \stackrel{\approx}{=} -\frac{r_1}{\delta} + \frac{1}{\delta^3 T} (\delta^4 - r_2(r_1^2 + \delta^2)).$$

Hence, we have

$$\begin{aligned} D(T-1) &\stackrel{\approx}{=} |r|^{T-1} \left\{ \left(\cos \delta - r_1 \frac{\sin \delta}{\delta} \right) \right. \\ &\quad \left. + \frac{1}{\delta^2 T} (r_1 \delta^2 (r_2 + \delta^2) - r_2 (r_1^2 + \delta^2) + \delta^4) \frac{\sin \delta}{\delta} + \frac{1}{\delta^2 T} A_3 \cos \delta \right\}, \end{aligned}$$

with

$$A_3 = (r_1 \delta^2 + r_1^2 (r_2 + \delta^2)).$$

Using (A.4) and the expansion above for $D(T-1)$, we deduce that the expansion for $D(T-1)/D(T)$ equals the expression inside the brackets above divided by $|r|(A_1 - A_2/(\delta^2 T))$. After some manipulations, we obtain

$$\frac{D(T-1)}{D(T)} \stackrel{\approx}{=} 1 + \frac{r_1^2 + \delta^2}{T(\cos \delta - r_1 \frac{\sin \delta}{\delta})} \frac{\sin \delta}{\delta} \stackrel{\approx}{=} 1 + \frac{h}{T},$$

with

$$h = \frac{r_1^2 + \delta^2}{(\cos \delta - r_1 \frac{\sin \delta}{\delta})} \frac{\sin \delta}{\delta}.$$

Finally,

$$MGF(u, v) = \left(1 + \frac{\kappa^2(r_1^2 + \delta^2)}{2T(\cos \delta - r_1 \frac{\sin \delta}{\delta})} \frac{\sin \delta}{\delta} \right) D(T)^{-1/2}.$$

Combining this result with (A.4) and rearranging yields the result stated in the Theorem.

Derivation of (12). From Perron (1996), the joint moment generating function of $(Z(c, \gamma), K(c, \gamma))$ can be expressed as

$$MGF_{PCB}(u, v) = \exp\left(\frac{u^2}{4T} - \frac{v\kappa^2(\exp(2c) - 1)}{T \cdot 2c}\right) M_{c,\gamma}(u, v), \tag{A.5}$$

where

$$M_{c,\gamma}(u, v) = \Psi_c(u, v) \exp\left\{-\frac{\kappa^2}{2T}(u + c - \lambda)[1 - \exp(u + c - \lambda)\Psi_c^2(u, v)]\right\},$$

with

$$\Psi_c(u, v) = \{2\lambda \exp(-(c + u))/[(\lambda + u + c) \exp(-\lambda) + (\lambda - u - c) \exp(\lambda)]\}^{1/2},$$

and

$$\lambda^2 = -(2v - 2cu - c^2).$$

To obtain the alternative representation, we start with the following identities: $\exp(-\lambda) = ch(\lambda) - sh(\lambda)$ and $\exp(\lambda) = ch(\lambda) + sh(\lambda)$ which define the hyperbolic functions $ch(\lambda)$ and $sh(\lambda)$. We can write

$$(\lambda + u + c) \exp(-\lambda) + (\lambda - u - c) \exp(\lambda) = 2\lambda ch(\lambda) - 2r_1 sh(\lambda),$$

with $r_1 = c + u$ and we deduce that

$$\frac{(\lambda + u + c) \exp(-\lambda) + (\lambda - u - c) \exp(\lambda)}{2\lambda} = ch(\lambda) - r_1 \frac{sh(\lambda)}{\lambda}.$$

Using $\delta = i\lambda$, we have $\delta^2 = -\lambda^2 = 2v - 2cu - c^2$ and

$$ch(\lambda) - r_1 \frac{sh(\lambda)}{\lambda} = \cos \delta - r_1 \frac{\sin \delta}{\delta}.$$

It is then easy to show that $\Psi_c(u, v)$ can be written as

$$\Psi_c(u, v) = \exp(-(c + u)/2) \left(\cos \delta - \frac{r_1}{\delta} \sin \delta \right)^{-1/2},$$

and we deduce that

$$M_{c,\gamma}(u, v) = \exp(-r_1/2) \left(\cos \delta - \frac{r_1}{\delta} \sin \delta \right)^{-1/2} \times \exp\left\{-\frac{\kappa^2 v}{2T}(r_1 - \lambda) \left[1 - \frac{\exp(-\lambda)}{\cos \delta - \frac{r_1}{\delta} \sin \delta} \right]\right\}.$$

The expression (12) is obtained combining this last result with (A.5).