



**On the Exact Moments of Asymptotic Distributions in an Unstable Ar(1)
with Dependent Errors**

Jesus Gonzalo; Jean-Yves Pitarakis

International Economic Review, Volume 39, Issue 1 (Feb., 1998), 71-88.

Stable URL:

<http://links.jstor.org/sici?sici=0020-6598%28199802%2939%3A1%3C71%3AOTEMOA%3E2.0.CO%3B2-P>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

International Economic Review is published by Economics Department of the University of Pennsylvania. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ier_pub.html.

International Economic Review

©1998 Economics Department of the University of Pennsylvania

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

ON THE EXACT MOMENTS OF ASYMPTOTIC DISTRIBUTIONS
IN AN UNSTABLE AR(1) WITH DEPENDENT ERRORS*

BY JESÚS GONZALO AND JEAN-YVES PITARAKIS^{†1}

Universidad Carlos III de Madrid, Spain
University of Reading, United Kingdom

In this paper we derive the exact moments of asymptotic distributions of the OLS estimate and t -statistic in an unstable AR(1) with dependent errors. We also study the relationship between the number of lagged dependent variables required for matching the distribution moments in the 'approximately i.i.d. errors' model with those occurring in the 'purely i.i.d.' model.

1. INTRODUCTION

Many results regarding the impact of dependent errors on test statistics and coefficient estimates in the context of nonstationary time series have been established via Monte-Carlo simulations. In an extensive empirical study Schwert (1989), for instance, considered the distortions induced by the presence of moving average errors on various test statistics within a nonstationary AR(1) model. More recently, Agiakoglou and Newbold (1992) studied the size properties of the augmented Dickey-Fuller (ADF) test under a moving average error structure. Since the relevant asymptotic distributions are expressed in terms of stochastic integrals involving Wiener processes, they are not directly usable for computational purposes. This partly explains why most studies adopted the computationally demanding simulation approach.

However, since most quantities of interest are expressed in terms of ratios of stochastic integrals, which themselves are the limits of properly normalized quadratic forms, it is possible to obtain their exact moments using their joint moment generating function (MGF thereafter). The recent unit root literature contains many attempts to obtain exact distributional results via the relevant characteristic function and the use of Gurland's (1948) inversion theorem (Evans and Savin 1981, 1984, Perron 1989, Hisamatsu and Maekawa 1994, and Abadir 1993a, 1995a and 1995b, among others). Most of these studies, however, used the i.i.d. errors assumption and focused solely on the normalized ordinary least squares (OLS) estimate of the autoregressive coefficient. More recently Perron (1994) also considered MA and AR error structures in a near integrated framework. His objectives differed from ours in

* Manuscript received March 1995; final revision May 1996.

[†] Email: jgonzalo@eltrond.uc3m.es; j.pitarakis@reading.ac.uk

¹ We would like to thank two anonymous referees for their useful comments. The first author acknowledges support from the Spanish Secretary of Education DGICYT program No. PB950298, and the second author thanks the European Union Human Capital and Mobility Program No. ERBCHBICT941677 for financial support. All errors are solely ours.

that he focused mainly on obtaining exact distributional results for the normalized OLS coefficient and on studying the adequacy of the asymptotic approximations.

Typically, the two most important quantities arising in the nonstandard asymptotic distributions are $\int_0^1 W(r) dW(r)$ and $\int_0^1 W(r)^2 dr$, where $W(r)$ denotes a standard Brownian Motion. In this paper we will analyze the impact of the presence of dependent errors on autoregressive coefficient estimates and corresponding t -statistics by focusing on the influence of the error process on the exact moments of the asymptotic distributions and indirectly on inferences. Thus, we will establish theoretically and exactly various empirically observed facts with the aim of better understanding the behavior of these asymptotic distributions when the error structure is not i.i.d. Our framework allows us to distinguish explicitly between different types of error processes.

It is important to point out that this paper is not an alternative to finite sample methods such as Monte-Carlo or bootstrapping, but rather a complementary framework that provides an exact analysis of the asymptotic behavior of the relevant test statistics. This is particularly important given the difficulty of quantifying the distortions persisting at the asymptotic level with simulation methods.

As a byproduct, we extend a result by Le Breton and Pham (1989) regarding the exact asymptotic bias of the OLS estimate in a nonstationary AR(1), to the case where the error process is not i.i.d. In addition, we investigate the connection between the magnitude of the coefficients of the error process and the number of lags of the dependent variable necessary to dampen their effects on the original asymptotic distribution. Our results can also be viewed as a preliminary step for the design of Bartlett corrections to the test statistics in the context of nonstationary processes.

The plan of this paper is as follows. Section 2 presents the general framework and theoretical tools. Section 3 focuses on quantitative results, and Section 4 concludes.

2. FRAMEWORK AND METHODOLOGY

We consider the following first-order autoregressive process

$$(1) \quad \Delta X_t = \alpha X_{t-1} + u_t$$

$$(2) \quad \phi(L)u_t = \theta(L)\epsilon_t$$

where $\phi(L) = 1 - \rho L$, $\theta(L) = 1 + \theta L$ and $X_0 = 0$. We further assume that ϵ_t is i.i.d. Gaussian with mean zero and variance σ_ϵ^2 , and that the two polynomials in the lag operator L have no roots in common and satisfy the usual stability and invertibility assumptions. For notational simplicity, and with no loss of generality we also put $\sigma_\epsilon^2 = 1$. The quantities of interest are the OLS estimate of α in (1) and the corresponding t -statistic denoted t_α . In what follows, we distinguish between an AR(1) process for u_t (i.e., $\theta = 0$), an MA(1) and mixed ARMA(1,1) process, respectively. The various asymptotic distributions of the quantities of interest when $\alpha = 0$ are gathered in the following lemma where ' \Rightarrow ' denotes convergence in distribution.

LEMMA 2.1. Case $\rho = 0$ (MA(1) errors):

1. $T(\hat{\alpha} - \alpha) \Rightarrow \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W(r)^2 dr} + \frac{\theta}{(1 + \theta)^2} \frac{1}{\int_0^1 W(r)^2 dr}.$
2. $t_{\hat{\alpha}} \Rightarrow \frac{(1 + \theta)}{(1 + \theta^2)^{1/2}} \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W(r)^2 dr)^{1/2}} + \frac{\theta}{(1 + \theta)(1 + \theta^2)^{1/2}} \frac{1}{(\int_0^1 W(r)^2 dr)^{1/2}}.$

Case $\theta = 0$ (AR(1) errors):

3. $T(\hat{\alpha} - \alpha) \Rightarrow \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W(r)^2 dr} + \frac{\rho}{1 + \rho} \frac{1}{\int_0^1 W(r)^2 dr}.$
4. $t_{\hat{\alpha}} \Rightarrow \frac{(1 + \rho)^{1/2}}{(1 - \rho)^{1/2}} \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W(r)^2 dr)^{1/2}} + \frac{\rho}{(1 - \rho)^{1/2}(1 + \rho)^{1/2}} \frac{1}{(\int_0^1 W(r)^2 dr)^{1/2}}.$

Case $\theta \neq 0$ and $\rho \neq 0$ (ARMA(1, 1) errors):

5. $T(\hat{\alpha} - \alpha) = \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W(r)^2 dr} + f_1(\rho, \theta) \frac{1}{\int_0^1 W(r)^2 dr}.$
6. $t_{\hat{\alpha}} \Rightarrow f_2(\rho, \theta) \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W(r)^2 dr)^{1/2}} + f_3(\rho, \theta) \frac{1}{(\int_0^1 W(r)^2 dr)^{1/2}},$

where the functions f_i ($i = 1, 2, 3$) above are defined as follows

$$f_1(\rho, \theta) = \frac{(\theta + \rho)(1 + \rho\theta)}{(1 + \rho)(1 + \theta)^2},$$

$$f_2(\rho, \theta) = \frac{(1 + \rho)^{1/2}}{(1 - \rho)^{1/2}} \frac{(1 + \theta)}{(1 + \theta^2 + 2\rho\theta)^{1/2}},$$

$$f_3(\rho, \theta) = \frac{(1 + \rho\theta)(\rho + \theta)}{(1 + \rho)^{1/2}(1 - \rho)^{1/2}(1 + \theta)(1 + \theta^2 + 2\rho\theta)^{1/2}}.$$

PROOF. Follows from Phillips (1987, Theorem 3.1). □

Noting that when the errors are i.i.d. the asymptotic distributions are

$$\frac{\int_0^1 W(r) dW(r)}{\int_0^1 W^2(r) dr} \quad \text{and} \quad \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W^2(r) dr)^{1/2}}$$

for $T(\hat{\alpha} - 1)$ and $t_{\hat{\alpha}}$, respectively. The extra components appearing in the above lemma illustrate the impact of dependent errors on inferences, when we mistakenly assume them to be i.i.d. Clearly, the signs and magnitudes of the parameters of the error processes will play a crucial role.

All the above distributions are expressed in terms of ratios of stochastic integrals. Indeed, letting $Q_1 = \int_0^1 W(r) dW(r)$ and $Q_2 = \int_0^1 W(r)^2 dr$, the ratios of interest are Q_1/Q_2 , $Q_1/\sqrt{Q_2}$, $1/Q_2$ and $1/\sqrt{Q_2}$, respectively.

LEMMA 2.2. *The joint moment generating function $\phi(u, v)$ of Q_1 and Q_2 has the following representation*

$$\phi(u, v) = e^{-u/2} \left[\cosh(-2v)^{1/2} - \frac{u}{(-2v)^{1/2}} \sinh(-2v)^{1/2} \right]^{-1/2}$$

PROOF. Follows from White (1958, p. 1193). \square

We can now apply a result found by Sawa (1978). Letting $\phi(u, v)$ denote the joint MGF of Q_1 and Q_2 , the m th order moment of Q_1/Q_2 is given by

$$E \left[\left(\frac{Q_1}{Q_2} \right)^m \right] = \Gamma^{-1}(m) \int_0^\infty v^{m-1} \left[\frac{\partial^m \phi(u, -v)}{\partial u^m} \right]_{u=0} dv$$

provided the moments exist, and where $\Gamma(\cdot)$ is the gamma function, and m is a positive integer. Given the form of the ratios of interest to us, the above expression is not directly applicable for obtaining the exact moments of quantities such as $Q_1/\sqrt{Q_2}$ or Q_2^{-m} . However, using a result found in Davies et al. (1985), we can extend Sawa's result to the above cases as well.

LEMMA 2.3.

$$E \left[\left(\frac{Q_1}{\sqrt{Q_2}} \right)^m \right] = \Gamma^{-1} \left(\frac{m}{2} \right) \int_0^\infty v^{m/2-1} \left[\frac{\partial^m \phi(u, -v)}{\partial u^m} \right]_{u=0} dv,$$

$$E \left[\left(\frac{1}{\sqrt{Q_2}} \right)^m \right] = \Gamma^{-1} \left(\frac{m}{2} \right) \int_0^\infty v^{m/2-1} [\phi(u, -v)]_{u=0} dv.$$

PROOF. Writing $Q_1^m = [\partial^m e^{uQ_1} / \partial u^m]_{u=0}$ and $Q_2^{-m/2} = \Gamma^{-1}(m/2) \int_0^\infty v^{m/2-1} e^{-vQ_2} dv$, the result follows from Fubini's theorem noting that $E[Q_1^m e^{-vQ_2}] = E[\partial^m e^{-uQ_1 - vQ_2} / \partial u^m]_{u=0} = [\partial^m \phi(u, -v) / \partial u^m]_{u=0}$. \square

We can therefore directly use the joint moment generating function of Q_1 and Q_2 in Lemma 2.2 to solve the above integrals. This is the objective of the next section, where we also investigate the behavior of the moments with θ and ρ . It is important to point out that the existence of any one side of the expressions appearing in Lemma 2.3 implies the existence of the other (see Evans and Savin 1981, p. 768).

3. NUMERICAL RESULTS

3.1. *Original (Nonaugmented) Model.* The following lemma presents the exact numerical value of the expectations (moments) of the ratios of interest. We use the following notational convention, letting $\mu^n(A/B)$ denote $E((A/B)^n)$. In what follows, we will focus solely on the first moment and variance. Higher order moments can also be easily obtained using Lemma 2.3.

LEMMA 3.1. *Letting $x = (2v)^{1/2}$ and $\omega(n) = (-1)^n \Gamma(\frac{1}{2} + n)/n!$, we have*

$$\begin{aligned}
 \mu_{Q_1/Q_2}^1 &= -\frac{1}{2} \int_0^\infty \frac{x}{(\cosh x)^{1/2}} dx + \frac{1}{2} \int_0^\infty \frac{\sinh x}{(\cosh x)^{3/2}} dx \\
 &= -\frac{2\sqrt{2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)}{(1+4n)^2} + 1 = -1.78143 \\
 \mu_{Q_1/Q_2}^2 &= \int_0^\infty \left(\frac{x^3}{8(\cosh x)^{1/2}} - \frac{x^2 \sinh x}{4(\cosh x)^{3/2}} + \frac{3x(\sinh x)^2}{8(\cosh x)^{5/2}} \right) dx \\
 &= \frac{12\sqrt{2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)}{(1+4n)^4} \\
 &\quad - \frac{2^{5/2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)(1+2n)((5+4n)^3 - (1+4n)^3)}{(1+4n)^3(5+4n)^3} \\
 &\quad + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)(3+2n)(1+2n)}{(1+4n)^2} \\
 &\quad + \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)(3+2n)(1+2n)}{(9+4n)^2} \\
 &\quad - \frac{\sqrt{2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)(3+2n)(1+2n)}{(5+4n)^2} = 13.2857 \\
 \mu_{1/Q_2}^1 &= \int_0^\infty \frac{x}{(\cosh x)^{1/2}} dx \\
 &= \frac{4\sqrt{2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)}{(1+4n)^2} = 5.56286
 \end{aligned}$$

$$\begin{aligned}
\mu_{1/Q_1}^2 &= \int_0^\infty \frac{x^3}{2(\cosh x)^{1/2}} dx \\
&= \frac{96}{\sqrt{2}\pi} \sum_{n=0}^\infty \frac{\omega(n)}{(1+4n)^4} = 67.8312 \\
\mu_{Q_1/\sqrt{Q_1}}^1 &= \frac{1}{\sqrt{2}\pi} \left(\int_0^\infty \frac{\sinh x}{x(\cosh x)^{3/2}} dx - \int_0^\infty \frac{1}{(\cosh x)^{1/2}} dx \right) \\
&= -\frac{2}{\pi} \sum_{n=0}^\infty \frac{\omega(n)}{(1+4n)} + \frac{1}{\pi} \sum_{n=0}^\infty \omega(n) \left(\log\left(\frac{1+4n}{8}\right) + 2 \log \frac{\Gamma\left(\frac{1+4n}{8}\right)}{\Gamma\left(\frac{5+4n}{8}\right)} \right) \\
&= -0.4231 \\
\mu_{1/\sqrt{Q_2}}^1 &= \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{2}}{(\cosh x)^{1/2}} dx \\
&= \frac{4}{\pi} \sum_{n=0}^\infty \frac{\omega(n)}{(1+4n)} = 2.09211 \\
\mu_{Q_1/\sqrt{Q_2}}^2 &= \int_0^\infty \left(\frac{x}{4(\cosh x)^{1/2}} - \frac{\sinh x}{2(\cosh x)^{3/2}} + \frac{3(\sinh x)^2}{4x(\cosh x)^{5/2}} \right) dx \\
&= \frac{\sqrt{2}}{\sqrt{\pi}} \sum_{n=0}^\infty \frac{\omega(n)}{(1+4n)^2} - 1 \\
&\quad + \frac{3\sqrt{2}}{4\sqrt{\pi}} \sum_{n=0}^\infty \omega(n)(1+2n) \left(\log\left(\frac{1+4n}{8}\right) + 2 \log \frac{\Gamma\left(\frac{1+4n}{8}\right)}{\Gamma\left(\frac{5+4n}{8}\right)} \right) \\
&\quad - \frac{3\sqrt{2}}{4\sqrt{\pi}} \sum_{n=0}^\infty \omega(n)(1+2n) \left(\log\left(\frac{5+4n}{8}\right) + 2 \log \frac{\Gamma\left(\frac{5+4n}{8}\right)}{\Gamma\left(\frac{9+4n}{8}\right)} \right) = 1.1417
\end{aligned}$$

$$\begin{aligned}
\mu_{Q_1/Q_2}^1 &= \int_0^\infty \left(\frac{-x^3}{4(\cosh x)^{1/2}} + \frac{x^2 \sinh x}{2(\cosh x)^{3/2}} \right) dx \\
&= -\frac{24\sqrt{2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)}{(1+4n)^4} + \frac{2^{7/2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{\omega(n)(1+2n)((5+4n)^3 - (1+4n)^3)}{(1+4n)^3(5+4n)^3} \\
&= -22.7899.
\end{aligned}$$

PROOF. Follows by writing the hyperbolic functions as $1/2(e^x \pm e^{-x})$, expanding the terms appearing as $(1 + e^{-2x})^{-a/2}$ and integrating termwise. For the integrands involving x in the denominator an intermediate result from Gradshteyn and Ryzhik (1980, p. 361) was used. \square

Using the above lemma we can assess the exact impact of θ and ρ , on the first and second moments, as well as the variance of the asymptotic distributions in Lemma 2.1. Note that the value of $\mu_{Q_1/\sqrt{Q_2}}^2$ in Lemma 3.1 represents the mean of the likelihood ratio statistic and corresponds to the figure obtained by Larsson (1994) and Nielsen (1995). Our value is more accurate, however, since they used numerical integration techniques instead of proceeding analytically. Letting $g_l(\rho, \theta)$ denote the first moment of the distributions in cases $l = 1, 2, 3, 4, 5, 6$ of Lemma 2.1, we have

$$\begin{aligned}
g_1(\theta) &= -1.78143 + \frac{\theta}{(1+\theta)^2} 5.56286, \\
g_2(\theta) &= -0.42310 \frac{(1+\theta)}{(1+\theta^2)^{1/2}} + \frac{\theta}{(1+\theta)(1+\theta^2)^{1/2}} 2.09211, \\
g_3(\rho) &= -1.78143 + \frac{\rho}{(1+\rho)} 5.56286, \\
g_4(\rho) &= -0.42310 \frac{(1+\rho)^{1/2}}{(1-\rho)^{1/2}} + \frac{\rho}{(1-\rho)^{1/2}(1+\rho)^{1/2}} 2.09211, \\
g_5(\rho, \theta) &= -1.78143 + \frac{(\theta+\rho)(1+\rho\theta)}{(1+\rho)(1+\theta)^2} 5.56286, \\
g_6(\rho, \theta) &= -0.42310 f_2(\rho, \theta) + f_3(\rho, \theta) 2.09211
\end{aligned}$$

for the pure moving average, autoregressive and ARMA processes, respectively. The following tables illustrate in more detail the exact quantitative impact of ρ and θ on the location and variance of the asymptotic distributions. It is worth emphasizing the fact that our figures are *exact*. The use of a purely numerical Monte-Carlo approach would have required days of computing time in order to lead to numbers close to ours. It is indeed difficult to make conjectures about the distortions that might continue to hold even asymptotically unless we are willing to bear an extremely

TABLE 1
 $T(\hat{\alpha} - \alpha)$

θ or ρ	$g_1(\theta)$	$g_3(\rho)$	$v_1(\theta)$	$v_3(\rho)$
-0.9	-502.4388	-51.8472	302104.71	3329.83
-0.7	-45.0481	-14.7614	2528.36	297.00
-0.1	-2.4682	-2.3995	15.23	14.67
0.0	-1.7814	-1.7814	10.11	10.11
0.1	-1.3217	-1.2757	7.32	7.06
0.7	-0.4340	0.5092	3.34	1.18
0.9	-0.3946	0.8536	3.21	0.92

TABLE 2
 $t_{\hat{\alpha}}$

θ or ρ	$g_2(\theta)$	$g_4(\rho)$	$v_2(\theta)$	$v_4(\rho)$
-0.9	-14.0270	-4.4167	53.9669	5.9559
-0.7	-4.1031	-2.2284	5.2340	1.7764
-0.1	-0.6102	-0.5930	0.9640	0.9625
0.0	-0.4231	-0.4231	0.9626	0.9626
0.1	-0.2739	-0.2575	0.9854	0.9893
0.7	0.1165	1.0435	1.1598	2.4115
0.9	0.1391	2.4754	1.1756	7.2124

heavy computational burden. We also believe that our approach leads to a better understanding of the distinct impact of the structure of the error process.

In Tables 1 and 2 above, $v_i(\theta)$ and $v_i(\rho)$ denote the variances in the MA(1) and AR(1) cases, respectively. The number -1.78143 appearing in the row corresponding to $\rho = \theta = 0$ (Table 1) is the asymptotic bias of the OLS estimate in the unit root model with i.i.d. errors (Le Breton and Pham 1989) and -0.4231 (Table 2) represents the mean of the asymptotic distribution of the t -statistic in the same context. Clearly, both cases illustrate the more frequent occurrence of negative values due to the presence of the unit root.

Particularly interesting are the magnitudes of the directions of shift and the differences between the autoregressive and moving average error processes. In Table 1, we can clearly observe the drastic shift to the left of the asymptotic distribution of $T(\hat{\alpha} - 1)$ as θ and ρ tend towards -1, the impact being much stronger in the case of moving average errors as judged by *both* the mean and variance. Indeed, we can observe an important influence of the error structure on the distributional shifts and variance changes—an autoregressive structure being much less distortionary than a moving average one when the parameters are negative and large in absolute value.

These shifts explain the severe size distortions occurring in the presence of dependent errors when inferences are based on the asymptotic distributions derived under the i.i.d. errors assumption. Both types of error structures will produce severe size distortions, and except when $\theta \rightarrow 1$, the MA process causes greater distortion.

Positive parameter values imply a shift rightward of the asymptotic distributions and therefore an easier wrong acceptance of the unit root hypothesis when the i.i.d.-based distributions are used for inferences. Our results suggest that for the

$T(\hat{\alpha} - \alpha)$ statistic, the 'undersizedness' will be more severe under an autoregressive error structure, where *both* the mean and variance are more severely modified relative to the MA case (see Table 1). When both ρ and θ are close to 0, however, their respective effect on the asymptotic distribution is quite comparable in magnitude.

The behavior of the asymptotic distribution of the t -statistic (Table 2) displays less pronounced displacements with θ or ρ . Although the directions are similar to the ones occurring in the $T(\hat{\alpha} - \alpha)$ case, they are also weaker. Clearly, this supports the view that the t -statistic is more reliable than the normalized OLS estimate for making inferences in such frameworks. Although Dickey and Fuller (1979) suggested that the $T(\hat{\alpha} - \alpha)$ statistic is more powerful than the t -statistic, our previous finding supports Schwert's (1989) Monte-Carlo-based claim that the latter is more robust to model misspecifications. It will display better-size properties when the model contains dependent errors.

The fact that the t -statistic is better behaved under dependent errors will also be reinforced in the next section where we analyze the impact of augmenting the model on the behavior of the two test statistics. Regarding the differences in the impact of the two types of error processes, it is also worth observing that, for the t -statistic, when the parameters are positive, their increase leads to a faster shift of the mean (rightward) as well as a faster increase in the variance under an autoregressive error structure. On the other hand, when negative values are considered, the mean decreases and the variance increases faster under the MA process.

Finally, results pertaining to the mixed ARMA case are summarized in Figures 1a-1d. It is difficult to argue that such a mixed process will necessarily lead to more pronounced changes in the asymptotic distributions than in the pure AR or MA

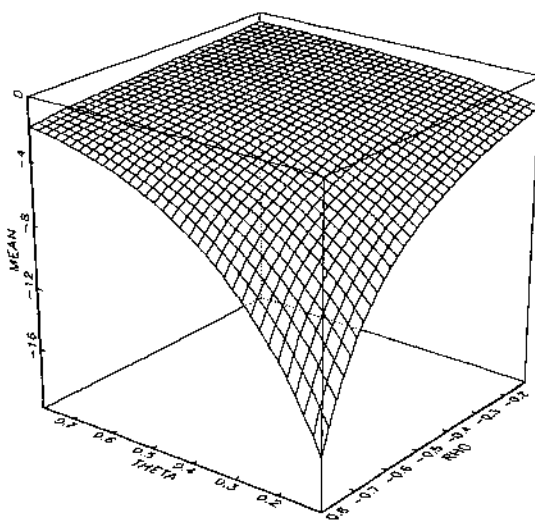


FIGURE 1a

MEAN OF $T(\hat{\alpha} - \alpha)$, ARMA(1, 1) ERRORS, $\theta > 0$, $\rho < 0$

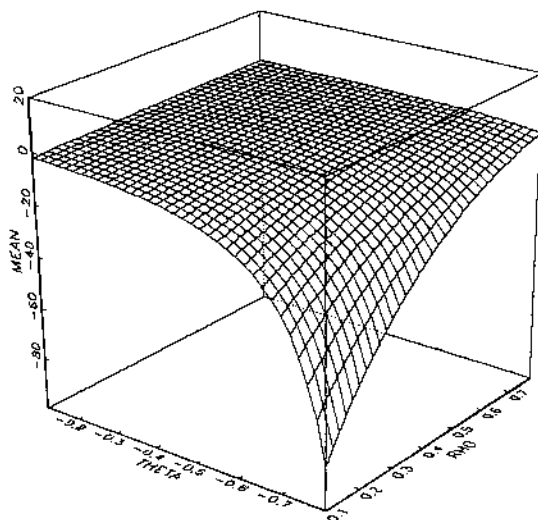


FIGURE 1b

MEAN OF $T(\hat{\alpha} - \alpha)$, ARMA(1, 1) ERRORS, $\theta < 0$, $\rho > 0$

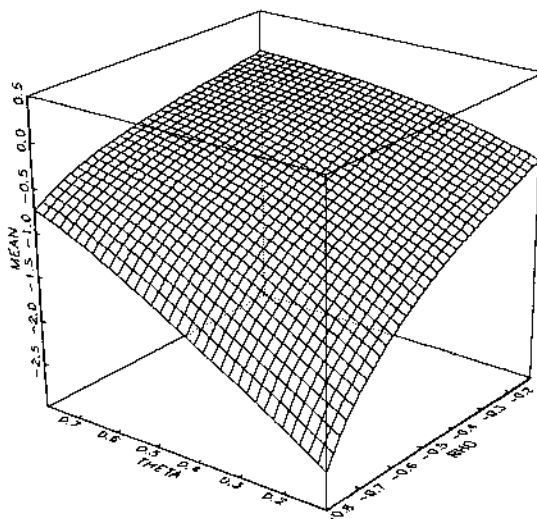


FIGURE 1c

MEAN OF $t_{\hat{\alpha}}$, ARMA(1, 1) ERRORS, $\theta > 0$, $\rho < 0$

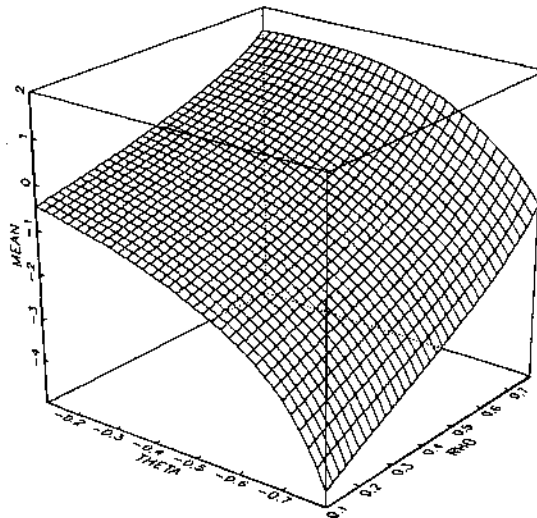


FIGURE 1d

MEAN OF $t_{\hat{\alpha}}$, ARMA(1,1) ERRORS, $\theta < 0$, $\rho > 0$

cases. Indeed, this will depend on the mix of values taken by θ and ρ together, and especially on their respective signs. It might happen, for instance, that a large positive ρ considerably dampens the influence of a large negative θ , leading to distributions that remain closer to the i.i.d. case than say when both θ and ρ are moderate but negative. Overall, the t -statistic displays less pronounced deviations from the i.i.d. error-based asymptotic distribution.

3.2. Normal Approximations. Given that we obtained the exact mean and variance of the asymptotic distributions of $t_{\hat{\alpha}}$ and $T(\hat{\alpha} - \alpha)$, it is natural to inquire about the quality of a normal approximation to these nonstandard distributions. The unit root literature has often raised this question by comparing the left tails of the Dickey-Fuller distributions to the ones of the standard normal (Evans and Savin 1981, Abadir 1995b, etc.). In Abadir (1995b), for instance, the author shows that shifting the standard normal probability density function (*pdf*) by -0.3 provides a very accurate approximation to the unit root density of the t -statistic. Here we will consider the comparison with $N(\mu_{\infty}, \sigma_{\infty}^2)$, where μ_{∞} and σ_{∞}^2 are the mean and variance of the correct nonstandard asymptotic distributions. For the $t_{\hat{\alpha}}$ statistic, we have $\mu_{1,\infty} = -0.4231$ and $\sigma_{1,\infty}^2 = 0.9626$, and for $T(\hat{\alpha} - \alpha)$, $\mu_{2,\infty} = -1.78143$ and $\sigma_{2,\infty}^2 = 10.11$. Another motivation behind these calculations is to obtain approximate numerical values for the magnitudes of the size distortions implied by the shifts in the moments when the i.i.d.-based asymptotic distributions are used for inferences. In Abadir (1995b) the author derived the exact density of the asymptotic distribution

TABLE 3
NORMAL APPROXIMATIONS*

	$N(-0.4231, 0.9626)$	DF_1	$N(-1.78143, 10.11)$	DF_2
2.5%	-2.35	-2.227	-8.01	-7.382
5%	-2.04	-1.941	-7.03	-5.685
10%	-1.68	-1.617	-5.85	-4.040

* DF_1 and DF_2 denote the left tail 5% critical values of the nonstandard (Dickey and Fuller 1979) asymptotic distribution of $t_{\hat{\alpha}}$ and $T(\hat{\alpha} - \alpha)$, the values of which have been obtained from Abadir (1993a) and Abadir (1995b), respectively.

of the t -statistic under the assumption of i.i.d. errors, but to our knowledge exact results under dependent errors are not available in the literature. Using the *pdf* transformation theorem, together with the joint density of Q_1 and Q_2 from Abadir (1995a), one could also obtain the exact densities of the variates in Lemma 2.1. This however, is beyond the scope of this paper and will be investigated in further research.

Table 3 displays the relevant normal and 'exact' DF critical values. Clearly, a suitable normalization leads to a very accurate normal approximation for the t -statistic at all relevant percentage points. The closeness of these distributions can partly justify the use of the normal approximation in order to obtain approximate estimates of the size distortions under different error structures. However, since the approximation for the normalized OLS coefficient is less accurate, we will concentrate solely on the t -statistic. The fact that the approximation based on the t -statistic is more accurate confirms an important finding by Abadir (1992), who showed that the lower tail of the t -statistic decays much faster than that of the normalized OLS coefficient.

Obviously, we also need to check the validity of the approximation when the original DF distributions are shifted due to the presence of MA or AR errors. For this reason, the following tables also include direct DF-based size estimates obtained via numerical simulations. The previously obtained moments of the various asymptotic distributions were very informative about the directions of the shifts of these distributions and gave an overall intuition about the seriousness and magnitude of the distortions. However, in order to obtain a more precise numerical value of, for example, the probability of rejecting the null when true, we need to compute the relevant tail areas.

Our main objective here is to illustrate the connection between the shifts in the moments and the implied 'new tail area.' Going back to our previous point, we saw for instance that when $\theta = -0.7$, we have $\mu_{\infty} = -4.1031$ and $\sigma_{\infty}^2 = 5.2340$ for the standard t -statistic. We can therefore compute the implied size distortion via the following probability

$$P[X \leq -1.941 | X \approx N(-4.1031, 5.2340)]$$

where -1.941 corresponds to the 5 per cent DF critical value (see Table 3). Similar probabilities can be obtained for a whole range of values for θ . Table 4 summarizes

TABLE 4
APPROXIMATE AND EMPIRICAL SIZE ESTIMATES

θ or ρ	SIZE (MA(1))	SIZE (AR(1))	DF(MA(1))	DF(AR(1))
-0.9	94.95%	84.38%	100 %	86.48%
-0.7	82.64%	58.32%	84.04%	52.86%
-0.1	8.53%	8.69%	8.52%	7.96%
0.0	5.94%	5.94%	5.01%	5.01%
0.1	4.55%	4.46%	2.90%	2.89%
0.7	2.74%	2.68%	0.56%	0.00%
0.9	2.68%	4.95%	0.33%	0.00%

some of the implied approximate size estimates for the t -statistic under a 5 per cent nominal size. The last two columns display the counterparts of the size estimates obtained by Monte-Carlo simulations with $N = 10000$ replications and a sample size of 5000 observations.

The above numbers confirm that the undersizedness (due to both AR or MA error processes with $(\rho, \theta) > 0$) is more serious when the errors are characterized by an autoregressive structure. On the other hand, an MA process leads to greater oversizedness when the parameters are negative. These results unanimously confirm our moments-based analysis of Section 3.1. In addition, our results based on the normal approximation were able to give an accurate description of the size distortions implied by the shifts in the moments, for the vast majority of cases.

3.3. Augmented Model. In practice, in order to be able to continue using the distributions corresponding to $\theta = 0$ and/or $\rho = 0$ even when the error process is not i.i.d., one adds lagged changes of the dependent variable to the right hand side of (1). This has the effect of whitening the error process, which can then be assumed to be approximately i.i.d. In applied work, an important issue is then the selection of the truncation lag. Indeed, in order for inferences to be based on Q_1/Q_2 , even when say $\theta \neq 0$ or $\rho \neq 0$, the lag length needs to satisfy certain speed conditions (Said and Dickey 1984, Ng and Perron 1995), and will therefore play a crucial role on the quality of inferences even asymptotically. In order to shed some light on this issue, we computed the asymptotic distributions of $\hat{\alpha}$ and $t_{\hat{\alpha}}$ in (1) when $k = 1$ and 2, and where k denotes the truncation lag. We can therefore analyze explicitly the relationship between the lag length and the magnitudes of θ or ρ . The estimated model is given by

$$\Delta X_t = \alpha_1 X_{t-1} + \sum_{j=1}^k \gamma_j \Delta X_{t-j} + u_t.$$

The following lemma presents the different distributions under the hypothesis that $\alpha_1 = 0$ and for a given lag length.

LEMMA 3.2. *Case $k = 1$ and $MA(1)$:*

$$\begin{aligned}
 1. \quad T(\hat{\alpha}_1 - \alpha_1) &\Rightarrow \left(1 - \frac{\theta}{(1 + \theta^2)}\right) \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W(r)^2 dr} - \frac{\theta^2}{(1 + \theta)^2(1 + \theta^2)} \frac{1}{\int_0^1 W(r)^2 dr}, \\
 2. \quad t_{\hat{\alpha}_1} &\Rightarrow \left(\frac{1 - \theta^2}{1 + \theta^2}\right)^{1/2} \frac{(1 + \theta)(1 + \theta^2 - \theta)}{(1 - \theta^6)^{1/2}} \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W(r)^2 dr)^{1/2}} \\
 &\quad - \left(\frac{1 - \theta^2}{1 + \theta^2}\right)^{1/2} \frac{\theta^2}{(1 + \theta)(1 - \theta^6)^{1/2}} \frac{1}{(\int_0^1 W(r)^2 dr)^{1/2}}.
 \end{aligned}$$

Case $k = 1$ and $AR(1)$:

$$\begin{aligned}
 3. \quad T(\hat{\alpha}_1 - \alpha_1) &\Rightarrow (1 - \rho) \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W(r)^2 dr}, \\
 4. \quad t_{\hat{\alpha}_1} &\Rightarrow \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W(r)^2 dr)^{1/2}}.
 \end{aligned}$$

Case $k = 2$ and $MA(1)$:

$$\begin{aligned}
 5. \quad T(\hat{\alpha}_1 - \alpha_1) &\Rightarrow \frac{1 + \theta^2}{1 + \theta^2 + \theta} \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W(r)^2 dr} \\
 &\quad + \frac{\theta^3}{(1 + \theta)^2(1 + \theta^2 - \theta)(1 + \theta^2 + \theta)} \frac{1}{\int_0^1 W(r)^2 dr}, \\
 6. \quad t_{\hat{\alpha}_1} &\Rightarrow \frac{(1 + \theta)(1 + \theta^2)(1 - \theta^6)^{1/2}}{(1 + \theta^2 + \theta)(1 - \theta^8)^{1/2}} \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W(r)^2 dr)^{1/2}} \\
 &\quad + \frac{(\theta^3)(1 - \theta^6)^{1/2}}{(1 + \theta^2 + \theta)(1 + \theta^2 - \theta)(1 + \theta)(1 - \theta^8)^{1/2}} \frac{1}{(\int_0^1 W(r)^2 dr)^{1/2}}.
 \end{aligned}$$

Case $k = 2$ and $AR(1)$:

$$\begin{aligned}
 7. \quad T(\hat{\alpha}_1 - \alpha_1) &\Rightarrow (1 - \rho) \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W(r)^2 dr}, \\
 8. \quad t_{\hat{\alpha}_1} &\Rightarrow \frac{\int_0^1 W(r) dW(r)}{(\int_0^1 W(r)^2 dr)^{1/2}}.
 \end{aligned}$$

PROOF. Follows from Phillips (1987, Theorem 3.1) and the continuous mapping theorem. \square

We can now compare the moments of the above distributions in order to quantify the strength of the lag length for a given magnitude of θ and ρ . Table 5 displays the first moments of the distributions of $T(\hat{\alpha}_1 - \alpha_1)$ under no lags added, one lag, and two lags, respectively, in both the MA and AR cases.

The first three columns above display the moments of the asymptotic distribution of $T(\hat{\alpha}_1 - \alpha_1)$ under MA(1) errors when $k=0, 1$, and 2 , respectively. When θ is negative and as we increase k , we get closer and closer to the moment of the i.i.d. error-based asymptotic distribution. The picture is very different when we focus on positive values for θ . Indeed, in this latter case there is an initial improvement as we move from $k=0$ to $k=1$, but as we go further to $k=2$, the situation deteriorates and hence adding lags will not always be beneficial when inferences are based on $T(\hat{\alpha}_1 - \alpha_1)$. For instance, when $\theta > 0$, the optimal number of lags seems to be $k=1$. It is important to observe that this does not mean that using $k=3$ or further will not lead to any improvement. The point is that an increase in k does not yield to a *strict* improvement. This phenomenon has also occurred in Monte Carlo results, where the interest lay in determining the empirical size of the tests under MA(1) errors as k was being increased. (See Agiakoglou and Newbold 1992, pp. 474–475. In their study an increase in the number of lags from, say, $k=1$ to $k=8$ was shown to improve the size, but no systematic improvement occurred when the lag length increased from $k=1$ to $k=2$ or 3). The last three columns above focus on the AR(1) errors case. Clearly, fewer lags are required for whitening the error process for similar magnitudes of ρ . Table 6 displays the equivalent numbers for the t -statistic.

TABLE 5
 $T(\hat{\alpha}_1 - \alpha_1)$

θ or ρ	$g_1(\theta)_{k=0}$	$g_1(\theta)_{k=1}$	$g_1(\theta)_{k=2}$	$g_3(\rho)_{k=0}$	$g_3(\rho)_{k=1}$	$g_3(\rho)_{k=2}$
-0.9	-502.45	-251.61	-167.99	-51.85	-3.38	-3.38
-0.7	-45.04	-22.94	-15.61	-14.76	-3.03	-3.03
-0.1	-2.47	-2.03	-1.98	-2.40	-1.96	-1.96
0.0	-1.78	-1.78	-1.78	-1.78	-1.78	-1.78
0.1	-1.32	-1.65	-1.62	-1.28	-1.60	-1.60
0.7	-0.43	-1.58	-0.83	0.51	-0.53	-0.53
0.9	-0.39	-1.59	-0.73	0.85	-0.18	-0.18

TABLE 6
 $t_{\hat{\alpha}_1}$

θ or ρ	$g_2(\theta)_{k=0}$	$g_2(\theta)_{k=1}$	$g_2(\theta)_{k=2}$	$g_4(\rho)_{k=0}$	$g_4(\rho)_{k=1}$	$g_4(\rho)_{k=2}$
-0.9	-14.0270	-8.075	-5.6858	-4.4167	-0.4231	-0.4231
-0.7	-4.1031	-2.3010	-1.5695	-2.2284	-0.4231	-0.4231
-0.1	-0.6102	-0.4410	-0.4249	-0.5930	-0.4231	-0.4231
0.0	-0.4231	-0.4231	-0.4231	-0.4231	-0.4231	-0.4231
0.1	-0.2739	-0.4380	-0.4216	-0.2575	-0.4231	-0.4231
0.7	0.1165	-0.729	-0.2374	1.0435	-0.4231	-0.4231
0.9	0.1391	-0.768	-0.1918	2.4754	-0.4231	-0.4231

Again, the first three columns in Table 6 focus on the MA(1) case, and the remaining columns on AR(1). Noting that if the true process is a simple random walk with i.i.d. errors and we instead fit the random walk by also adding one lagged change of the dependent variable to the right-hand side, we implicitly have a random walk DGP with AR(1) errors. This is why the column corresponding to AR(1) errors with $k = 1$ shows a perfect match of the first moment for any value of ρ . The same also happens with $k = 2$, since asymptotically the inclusion of extra lags (beyond the true number) does not affect the asymptotic distribution of the t -statistic. An interesting point also arises by looking at the asymptotic distribution of $T(\hat{\alpha}_1 - \alpha_1)$ under AR(1) errors when $k = 1$ or 2 (cases 3 and 7 in Lemma 3.2). Indeed, we notice that although one lag is enough for the t -statistic to be brought to the i.i.d. distribution case (cases 4 and 8), the $T(\hat{\alpha}_1 - \alpha_1)$ statistic will always remain displaced with respect to the i.i.d. distribution, no matter how many lags we use. This can be intuitively explained by the fact that the latter statistic does not take the variance into account and therefore its use will always lead to more severely distorted inferences. Table 7 displays the variances of the asymptotic distributions of the t -statistic in the MA(1) case for the augmented model.

We can now compute the implied size distortions using the normal approximation. We focus solely on MA errors since we previously showed that the distortions induced by the presence of AR errors are neutralized when the number of lags is equal to or greater than one. The first three columns of Table 8 display the size estimates obtained via the normal approximation, using the relevant mean and variance of the asymptotic distributions for various values of the lag length. The last three columns are again the 'exact' counterparts obtained via Monte Carlo simulations using the correct Dickey and Fuller distribution.

TABLE 7
 $V_{\infty}(t_{\hat{\alpha}_1})$

θ or ρ	$v(\theta)_{k=0}$	$v(\theta)_{k=1}$	$v(\theta)_{k=2}$
-0.9	53.967	18.332	9.422
-0.7	5.234	2.136	1.397
-0.1	0.964	0.961	0.963
0.0	0.963	0.963	0.963
0.1	0.985	0.962	0.963
0.7	1.160	0.982	0.995
0.9	1.176	0.990	1.008

TABLE 8
SIZE ESTIMATES (NORMAL VERSUS MONTE CARLO WHEN MA(1))

θ	$N_{k=0}$	$N_{k=1}$	$N_{k=2}$	$DF_{k=0}$	$DF_{k=1}$	$DF_{k=2}$
-0.9	94.95%	92.36%	88.88%	100%	99.54%	94.64%
-0.7	82.64%	59.48%	37.45%	84.04%	54.32%	35.46%
-0.1	8.53%	6.18%	6.06%	8.52%	5.42%	5.24%
0.0	5.94%	5.94%	5.94%	5.01%	5.00%	5.00%
0.1	4.55%	6.18%	5.94%	2.90%	5.10%	4.80%
0.7	2.74%	10.93%	4.27%	0.56%	10.70%	2.57%
0.9	2.68%	11.70%	4.01%	0.33%	11.56%	2.26%

It is interesting to observe the evolution of the size estimates when $\theta > 0$, where an increase from $k = 0$ to $k = 1$ seriously deteriorates the size properties of the t -statistic. This clearly highlights the importance of the lag length selection for carrying on ADF-based unit root tests and reinforces our argument that an increase of lag length does not always lead to a *strict* improvement of size probabilities.

4. CONCLUSION

In this paper our objective was to offer a complementary analysis to the usual Monte Carlo simulations for evaluating the properties of distributions arising in nonstationary autoregressions. More specifically, we investigated the impact of the presence of dependent errors on the asymptotic distributions of the two most important quantities used for testing. We then focused on the properties of the standard method used for whitening the error process by analyzing the connection between the magnitude and sign of the parameters and the number of necessary lags in order to legitimately use the i.i.d. error-based distributions. Our framework allows us to distinguish specifically between different types of error processes. In addition, we showed that a proper normal approximation to nonstandard asymptotic distributions could give valuable hints on the magnitude of size distortions. Our results can easily be generalized to the multivariate framework using a multivariate analog of $\phi(u, v)$ recently obtained by Abadir and Larsson (1996). This can also open the way to multivariate Edgeworth-type asymptotic expansions, as in Knight and Satchell (1993). Finally, our results can also be used as a starting point for constructing Bartlett corrections for unit root tests.

REFERENCES

- ABADIR, K., "A Distribution Generating Equation for Unit Root Statistics," *Oxford Bulletin of Economics and Statistics* 54 (1992), 305-323.
- , "The Limiting Distribution of the Autocorrelation Coefficient Under a Unit Root," *Annals of Statistics* 21 (1993a), 1058-1070.
- , "On the Asymptotic Power of Unit Root Tests," *Econometric Theory* 9 (1993b), 189-221.
- , "The Joint Density of Two Functionals of a Brownian Motion," *Mathematical Methods of Statistics*, 4 (1995a), 449-462.
- , "The Limiting Distribution of the t -ratio Under a Unit Root," *Econometric Theory* 11 (1995b), 775-793.
- AND R. LARSSON, "The Joint Moment Generating Function of Quadratic Forms in Multivariate Autoregressive Series," *Econometric Theory* 12 (1996), 682-704.
- AGIAKOGLU, C. AND P. NEWBOLD, "Empirical Evidence on Dickey-Fuller Type Tests," *Journal of Time Series Analysis* 13 (1992), 471-483.
- DAVIES, N., M.B. PATE, AND J.D. PETRUCELLI, "Exact Moments of the Sample Cross Correlations of Multivariate Autoregressive Moving Average Time Series," *Sankhya: The Indian Journal of Statistics* 47 (1985), Series B, 325-337.
- DICKEY, D.A. AND W.A. FULLER, "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association* 74 (1979), 427-431.
- EVANS, G.B.A. AND N.E. SAVIN, "Testing for Unit Roots: 1," *Econometrica* 49 (1981), 753-779.
- AND ———, "Testing for Unit Roots: 2," *Econometrica* 52 (1984), 1241-1269.
- GRADSHTEYN, I.S. AND I.M. RYZHIK, *Tables of Integrals, Series and Products* (New York: Academic Press, 1980).

- GURLAND, J., "Inversion Formulae for the Distribution of Ratios," *Annals of Mathematical Statistics* 19 (1948), 228-237.
- HISAMATSU, H. AND K. MAEKAWA, "The Distribution of the Durbin Watson Statistic in Integrated and Near Integrated Models," *Journal of Econometrics* 61 (1994), 367-382.
- KNIGHT, J.L. AND S.E. SACHELL, "Asymptotic Expansions for Random Walks with Normal Errors," *Econometric Theory* 9 (1993), 363-376.
- LARSSON, R., "Bartlett Corrections for Unit Root Test Statistics," Preprint 2, Institute of Mathematical Statistics, University of Copenhagen, 1994.
- LE BRETON, A., AND D.T. PHAM, "On the Bias of the Least Squares Estimator for the First Order Autoregressive Process," *Annals of the Institute of Statistical Mathematics* 41 (1989), 555-563.
- MAGNUS, J.R., "The Exact Moments of a Ratio of Quadratic Forms in Normal Variables," *Annales d'Economie et de Statistique* 4 (1986), 95-109.
- NG, S. AND P. PERRON, "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association* 90 (1995), 268-281.
- NIELSEN, B., "Bartlett Correction of the Unit Root Test in Autoregressive Models," Discussion Paper No. 98, Nuffield College, Oxford University, 1995.
- PERRON, P., "The Calculation of the Limiting Distribution of the Least Squares Estimator in a Near Integrated Model," *Econometric Theory* 5 (1989), 241-255.
- , "The Adequacy of Limiting Distributions in the AR(1) Model with Dependent Errors," CRDE Working Paper, Université de Montréal, December 1994.
- PHILLIPS, P.C.B., "Time Series Regression with a Unit Root," *Econometrica* 55 (1987), 277-301.
- SAID, S. AND DICKEY, D.A., "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order," *Biometrika* 71 (1984), 599-607.
- SAWA, T., "The Exact Moments of the Least Squares Estimator for the Autoregressive Model," *Journal of Econometrics* 8 (1978), 159-172.
- SCHWERT, G.W., "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics* 7 (1989), 147-159.
- WHITE, J.S., "The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case," *Annals of Mathematical Statistics* 29 (1958), 1188-1197.