LEAST SQUARES ESTIMATION OF THE LINEAR MODEL WITH AUTOREGRESSIVE ERRORS*

NEERCHAL K. NAGARAJ** AND WAYNE A. FULLER†

Abstract. A Monte Carlo study of the least squares estimator of the regression model with autocorrelated errors is presented. The model contains a stationary explanatory variable and a random walk explanatory variable. The error model is a first order autoregressive model and the unit root case is included in the simulations. The limiting distribution of the regression pivotals for the basic model are normal, while the statistics for the autoregressive coefficient have a distribution that depends on the true parameter. The agreement between the Monte Carlo results and the asymptotic theory depends upon the autoregressive coefficient and on the nature of the explanatory variable.

Key words. Least squares, nonlinear estimation, Monte Carlo, time series.

AMS(MOS) subject classifications. Primary 62M10; secondary 62J02, 62F12.

1. Introduction. The regression model with autocorrelated errors is a natural model to use in many situations where the regression variables are observed over time. The basic model can be written as

$$Y_t = X_t\beta + u_t, \qquad t = 1, 2, \ldots \tag{1.1}$$

$$u_t = \sum_{i=1}^{p} \alpha_i u_{t-i} + e_t, \quad t = 1, 2, \ldots \tag{1.2}$$

where X_t is a q-dimensional row vector of explanatory variables, u_t is the unobserved error and $\{e_t\}$ is a sequence of zero mean random variables that are independent, or that satisfy conditions, such as those of martingale differences, that lead to behavior similar to that of independent random variables. Let

$$m^p - \sum_{i=1}^{p} \alpha_i m^{p-1} = 0 \tag{1.3}$$

be the characteristic equation associated with the autoregressive process.

By substituting the definition of u_t from (1.1) into (1.2), we obtain

$$\begin{aligned} Y_t &= X_t\beta + \sum_{i=1}^{p} \alpha_i (Y_{t-i} - X_{t-i}\beta) + e_t \\ &= f(Z_t, \eta) + e_t, \end{aligned} \tag{1.4}$$

where $Z_t = (X_t, Y_{t-1}, X_{t-1}, \ldots, Y_{t-p}, X_{t-p})$ and $\eta' = (\beta', \alpha_1, \alpha_2, \ldots, \alpha_p)$. Given a sample of n observations (Y_t, Z_t), $t = 1, 2, \ldots, n$, the problem is to estimate η.

Model (1.4) can also be written as

$$Y_t = X_t\beta + \sum_{i=1}^{p} X_{t-i}\zeta_i + \sum_{i=1}^{p} \alpha_i Y_{t-i} + e_t, \tag{1.5}$$

where $\zeta_i = -\alpha_i\beta$. Model (1.5) is linear in the parameters β, ζ_i, and α_i. However, the matrix of sums of squares and products for n observations on the vector $(X_t, X_{t-1}, \ldots, X_{t-p})$ may be singular. For example, the matrix will be singular if the intercept is included in the model. Let θ be the portion of $(\beta', \zeta_1', \zeta_2', \ldots, \zeta_p', \alpha_1, \alpha_2, \ldots, \alpha_p)$ associated with the part of the vector $(X_t, \ldots, X_{t-p})$ that has a nonsingular sum of squares and products matrix. We call the ordinary least squares estimator of θ, denoted by $\hat{\theta}$, the unrestricted least squares estimator.

If the largest root of the characteristic equation (1.3) is less than one in absolute value, Theorem 1 of Fuller, Hasza and Goebel (1981) can be used to show that the limiting distribution of the regression pivotal for the ordinary least squares estimator of an element of θ is normal for a wide range of explanatory variables.

If the largest root is equal to one, the limiting distribution of the least squares coefficient associated with that root is not normal and the limiting distribution depends on the explanatory variables in the equation. See Dickey and Fuller (1979), Fuller (1984), Fuller, Hasza and Goebel (1981), Phillips and Durlauf (1986), and Chan and Wei (1988).

Define the least squares estimator of η, denoted by $\tilde{\eta}$ of model (1.4), to be the value of η that minimizes

$$Q_n(\eta) = \sum_{t=1}^{n} [Y_t - f(Z_t, \eta)]^2. \tag{1.6}$$

We sometimes call this estimator the restricted estimator.

The limiting properties of the least squares estimator of η depend on the properties of the sequence $\{X_t\}$, on the properties of the sequence $\{e_t\}$ and on the roots of the characteristic equation. Nagaraj and Fuller (1989) have given a theorem for the estimation of a linear model subject to nonlinear restrictions that is applicable to the model defined by (1.1) and (1.2). The theorem permits the sum of squares of the explanatory variables to grow at different rates. For example, the vector X_t could contain a stationary variable, a time trend and (or) a random walk.

An interesting special case of model (1.1, 1.2) is the model with a random walk explanatory variable. Let

(1.7) $$Y_t = \beta_0 + \beta_1 Z_t + u_t,$$

or as

$$
\begin{aligned}
Y_t &= \beta_0(1-\rho) + (\beta_1 - \rho\beta_1 + \rho\beta_1^0)Z_{t-1} + \beta_1(Z_t - Z_{t-1}) \\
&\qquad + \rho(Y_{t-1} - \beta_1^0 Z_{t-1}) + e_t, \\
&= X_t\gamma + e_t\ ,
\end{aligned} \tag{1.8}
$$

where β_1^0 is the true value of β_1, $\gamma = [\beta_0(1-\rho), \beta_1 - \rho\beta_1 + \rho\beta^0, \beta_1, \rho]'$, and $X_t = [1, Z_{t-1}, Z_t - Z_{t-1}, Y_t - \beta_1^0 Z_{t-1}]$. The transformed version of the model in (1.8) is only used to identify the limiting distribution. It cannot be used in the actual estimation because β_1^0 is unknown. The transformation is used to define the limiting distribution in some important cases. For example, if $\rho = 1$, then the correlation between Z_t and Z_{t-1} converges to one. Therefore, the limit of the normalized sums of squares and product matrix of the original variables is singular. The transformation to the parameter vector γ removes the singularity.

Let $\rho = 1$ and write the transformed model as

$$
Y_t = \gamma_0 + \gamma_1 Z_{t-1} + \gamma_2(Z_t - Z_{t-1}) + \gamma_3(Y_{t-1} - \beta_1^0 Z_{t-1}) + e_t,
$$

where the true value of γ_0 is zero. Then it can be shown that the regression pivotal for the unrestricted least squares estimator of γ_1 converges in distribution to

$$
\frac{T_{ab} - W_a T_b - 1/2(\Gamma_{bb} - W_b^2)^{-1}(\Gamma_{ba} - W_b W_a)(T_b^2 - 1 - 2T_b W_b)}{[\Gamma_{aa} - W_a^2 - (\Gamma_{bb} - W_b^2)^{-1}(\Gamma_{ba} - W_b W_a)]^{1/2}}, \tag{1.9}
$$

where

$$
\begin{aligned}
(\Gamma_{aa}, \Gamma_{bb}, \Gamma_{ab}) &= \sum_{i=1}^{\infty} \zeta_i^2 (a_i^2, b_i^2, a_i b_i), \\
T_{ab} &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} v_{ij} a_i b_j, \\
T_b &= \sum_{i=1}^{\infty} 2^{1/2}\zeta_i b_i, \\
(W_a, W_b) &= \sum_{i=1}^{\infty} 2^{1/2}\zeta_i^2 (a_i, b_i), \\
v_{ij} &= 2[\zeta_j + \zeta_i]^{-1}\zeta_i^2\zeta_j, \\
(a_i, b_i) \sim NI(0, I) \quad &\text{and} \quad \zeta_i = (-1)^{i+1}2[(2i-1)\pi]^{-1}.
\end{aligned}
$$

The regression pivotal for γ_3 in the unrestricted model with $\rho = 1$ converges in distribution to

Phillips (1986) and Phillips and Durlauf (1986) for a different representation of the limiting distribution.

Using the results of Nagaraj and Fuller (1989), it can be shown that the limiting distribution for the estimator of γ_2 under the restricted model is the same as under the unrestricted model. The limiting distribution of the regression pivotal for the estimator of ρ under the restricted model is the distribution of

$$\hat{\tau}_\mu = \frac{1/2(T_b^2 - 1 - 2T_b W_b)}{(\Gamma_{bb} - W_b^2)^{1/2}},$$

where the distribution of $\hat{\tau}_\mu$ is tabulated in Fuller (1976).

Because the model (1.4) is a restricted version of the model (1.5), it is natural to consider a test of the restrictions. One test is constructed by analogy to the F-test in linear regression. This test is

$$F = (d_f - d_r)^{-1}(n - d_f)[Q_n(\hat{\theta})]^{-1}[Q_n(\tilde{\eta}) - Q_n(\hat{\theta})], \tag{1.11}$$

where d_f is the number of parameters in the full model associated with θ, $d_r = p+q$ is the number of parameters in the restricted model and $Q_n(\hat{\theta})$ is the residual sum of squares from the unrestricted model. Provided the unrestricted estimator has a limiting distribution, Corollary 2 of Nagaraj and Fuller (1989) can be used to obtain the limiting distribution of the test statistic. In model (1.7) with $\rho = 1$, the limiting distribution of the test statistic is the square of (1.9).

We use the Monte Carlo method to study the small sample behavior of the least squares estimator for a regression model with two regressors and an intercept. The error in the regression is a first order autoregression. One of the regressors is the normal random walk and the other is a sequence of independent $N(0,1)$ random variables.

There are some Monte Carlo experiments in the literature related to model (1.1). For example, Rao and Griliches (1969) study the model with a stationary regressor and stationary first order autoregressive error process. A regression model with a regressor which follows a random walk and errors which follow a stationary first order autoregression was considered by Krämer (1986). The Monte Carlo experiment in Krämer (1986) compared the ordinary least squares estimator to the corresponding generalized least squares estimators constructed with the true parameter of the error process. The generalized least squares estimator is the best linear unbiased estimator for the regression parameters. However, the generalized least squares estimator is generally unattainable in practice because it requires knowledge of the variance-covariance matrix of the errors.

2. Monte Carlo Study. The model of our study is an extension of the model considered by Krämer (1986). Our model contains two regressors and we include nonstationary autoregressive errors in our study. The model is where $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 1$, and $e_t \sim NI(0,1)$. The sequence $\{X_{1t}\}$ in the model (2.1) is a random walk generated by the following stochastic difference equation,

$$\begin{aligned} X_{1t} &= X_{1,t-1} + w_t, \quad && t = 1, 2, \ldots, n \\ &= 0, && t = 0 \end{aligned}$$

where $w_t \sim NI(0,1)$. The X_{2t} are $NI(0,1)$ random variables independent of w_t. The sequence $\{u_t\}$ is independent of the sequence $\{X_{1t}, X_{2t}\}$. The sum of squares of X_{1t} is order in probability n^2 and the sum of squares of X_{2t} is order in probability n. We rewrite the model as

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + u_1, \\ Y_t &= \beta_0(1-\rho) + \beta_1 X_{1t} - \beta_1 \rho X_{1,t-1} + \beta_2 X_{2t} - \beta_2 \rho X_{2,t-1} + \rho Y_{t-1} + e_t, \\ & \qquad t = 2, 3, \ldots, n. \end{aligned} \tag{2.3}$$

Because of the degeneracy of model (2.3) associated with $\rho = 1$, we defined a new parameter $\beta_* = \beta_0(1-\rho)$. Although the parameter β_* is zero under model (2.1) with $\rho = 1$, we estimated β_* in all cases.

Let $\theta = (\beta_*, \beta_1, -\beta_1\rho, \beta_2, -\beta_2\rho, \rho)$ and let $\hat{\theta}$ be the unrestricted least squares estimator of θ. The correlation between X_{1t} and $X_{1,t-1}$ approaches one as the sample size increases. Therefore, it is necessary to transform the independent variables with a transformation such as that described in Fuller, Hasza and Goebel (1981) to define the limiting distribution of the transformed unrestricted least squares estimator. The limiting distribution of the transformed parameter vector depends upon the parameters. In particular, the limiting distribution of the estimator of ρ (the coefficient of Y_{t-1}) depends on the value of ρ. The limiting distribution of the regression pivotal for ρ is given in (1.10). The addition of the stationary process to the model (1.7) does not alter the limiting distribution of the estimator of ρ.

It can be shown that assumptions of Theorem 1 of Nagaraj and Fuller (1989) are satisfied for the model (2.1) and (2.2) for all values of ρ in the interval $[-1, 1]$. Hence, the restricted least squares estimator, properly normalized, has a limiting distribution.

Samples were generated for several values of ρ in the range -1 to 1. Because the values of ρ close to 1 or -1 are more interesting than the values of ρ close to zero, more values of ρ close to the boundary were used in the study.

The nonlinear least squares estimators were obtained using the Gauss-Newton method. Estimation of the model (2.3) by the Gauss-Newton method consists of repeated regressions using derivatives evaluated at the estimates from the previous step as independent variables, and using the residuals $\hat{e}_t$ from the previous step as

Using the last $n-1$ rows of the Table 1, four Gauss-Newton iterations were performed. The first iteration used the coefficients of $(X_{1t}, X_{2t}, Y_{t-1})$ in the regression of Y_t on $(X_{1t}, X_{1,t-1}, X_{2t}, X_{2,t-1}, Y_{t-1})$ as start values for β and ρ, respectively. The estimate of ρ was restricted to the interval $[-1, 1]$ at the end of each iteration. If the estimate of ρ was less than one in absolute value at the end of four iterations then the fifth iteration was performed using all n observations. During the last iteration, the first observation was weighted by the factor $(1-\rho^2)^{1/2}$ as illustrated in Table 1. If the estimate of ρ was either -1 or 1 at the end of the four iterations, the fifth iteration was based on only the last $n-1$ observations of Table 1. The nonlinear least squares estimator obtained by this procedure is denoted by $(\tilde{\beta}_*, \tilde{\beta}_1, \tilde{\beta}_2, \tilde{\rho})$.

Table 1. Matrix used in estimation.

t	$\beta_0(1-\rho) = \beta_*$	β_1	β_2	ρ
1	$(1-\rho^2)^{1/2}(1-\rho)^{-1}$	$(1-\rho^2)^{1/2}X_{11}$	$(1-\rho^2)^{1/2}X_{21}$	0
2	1	$X_{12} - \rho X_{11}$	$X_{22} - \rho X_{21}$	u_1
3	1	$X_{13} - \rho X_{12}$	$X_{23} - \rho X_{22}$	u_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	1	$X_{1n} - \rho X_{1n-1}$	$X_{2n} - \rho X_{2n-1}$	u_{n-1}

The usual "t-statistics" are constructed using the estimated standard errors obtained at the last iteration. For example, the t-statistic, $t_{\tilde{\beta}_1}$, for β_1 is

$$t_{\tilde{\beta}_1} = [\tilde{V}\{\tilde{\beta}_1\}]^{-1/2}(\tilde{\beta}_1 - \beta_1^0),$$

where $\tilde{V}\{\tilde{\beta}_1\}$ is the second diagonal element of the inverse of the sum of squares and cross products matrix for the derivatives multiplied by s^2, and s^2 is the residual mean square obtained at the last iteration.

For model (2.1, 2.2), the "t-statistics" corresponding to the regression parameters $(\beta_0, \beta_1, \beta_2)$ are distributed as standard normal random variables, asymptotically, for $|\rho| < 1$. The "t-statistic" corresponding to the parameter ρ also has a limiting normal distribution when the true value of ρ is strictly less than one in absolute value.

If $\rho = 1$, the pivotal statistic corresponding to ρ has the same limiting distribution as the τ_μ statistic characterized in Fuller (1976, Chapter 8). If $\rho = -1$, the limiting distribution of the pivotal is the limiting distribution of $-\tau$. When $|\rho| = 1$, the limiting distributions of the "t-statistics" for $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are independent dard errors of the estimates in Table 4 are about 0.18.

Table 2. Empirical percentiles of $t_{\tilde{\beta}_1}$ for 1000 samples.

	$n = 25$		$n = 100$	
ρ	5%	95%	5%	95%
1.00	−2.29	2.23	−1.77	1.75
0.99	−2.25	2.20	−1.86	1.78
0.95	−2.42	2.35	−1.85	2.24
0.90	−2.29	2.77	−1.96	1.99
0.70	−2.27	2.56	−1.90	2.00
0.50	−2.30	2.14	−1.87	1.89
0.25	−1.97	1.86	−1.82	1.75
0.00	−1.94	1.91	−1.69	1.67
−0.25	−1.80	1.86	−1.69	1.71
−0.50	−1.76	1.71	−1.69	1.65
−0.70	−1.62	1.77	−1.61	1.58
−0.90	−1.59	1.60	−1.54	1.56
−0.95	−1.62	1.65	−1.48	1.66
−0.99	−1.56	1.74	−1.76	1.60
−1.00	−1.68	1.65	−1.75	1.65
$N(0,1)$	−1.65	1.65	−1.65	1.65

Table 3. Empirical percentiles of $t_{\tilde{\beta}_2}$ for 1000 samples.

	$n = 25$		$n = 100$	
ρ	5%	95%	5%	95%
1.00	−1.63	1.68	−1.64	1.66
0.99	−1.66	1.55	−1.53	1.67
0.95	−1.83	1.77	−1.53	1.63
0.90	−1.55	1.63	−1.63	1.68
0.70	−1.65	1.78	−1.58	1.58
0.50	−1.67	1.74	−1.60	1.64
0.25	−1.63	1.91	−1.75	1.55
0.00	−1.73	1.78	−1.80	1.73
−0.25	−1.73	1.76	−1.70	1.71
−0.50	−1.81	1.72	−1.58	1.64
−0.70	−1.60	1.66	−1.75	1.71
−0.90	−1.64	1.70	−1.65	1.71

Table 4. Empirical percentiles of $t_{\tilde{\rho}}$ for 1000 samples.

	$n = 25$		$n = 100$	
ρ	5%	95%	5%	95%
1.00	−3.27	0.00	−3.69	0.00
0.99	−9.37	0.11	−5.83	0.13
0.95	−4.06	0.39	−2.87	0.75
0.90	−3.15	0.52	−2.39	0.92
0.70	−2.52	0.96	−2.10	1.15
0.50	−2.38	1.31	−2.01	1.35
0.25	−2.16	1.21	−2.02	1.51
0.00	−2.23	1.33	−1.86	1.32
−0.25	−1.94	1.43	−1.75	1.43
−0.50	−1.80	1.29	−1.81	1.59
−0.70	−1.69	1.30	−1.64	1.65
−0.90	−1.44	1.60	−1.42	1.71
−0.95	−1.10	1.49	−1.28	1.62
−0.99	−0.55	1.54	−1.06	1.85
−1.00	0.00	1.68	0.00	1.87
$N(0,1)$	−1.65	1.65	−1.65	1.65
τ_μ	−3.00	0.00	−2.89	−0.05
$-\tau$	−1.33	1.95	−1.29	1.95

When $n = 25$, the percentiles for the pivotal associated with the random walk explanatory variable are close to those of a $N(0,1)$ variable if ρ is less than -0.70. When $n = 100$, the percentiles differ from those of the normal distribution by a considerable amount for ρ in the range 0.7 to 0.95, but the agreement is quite good for ρ less than zero. The deviation from normality is partly explained by the serious downward bias in $\tilde{\rho}$ for ρ close to 1.00.

The percentiles of $t_{\tilde{\beta}_2}$ for the coefficient of the normal $(0,1)$ explanatory variable are closer to the limiting distribution than are those of the coefficient of the random walk explanatory variable for all parameter configurations, except $n = 25$ and $\rho = -1$.

The regression pivotals for the estimated autocorrelation coefficient are given in Table 4. The small sample bias in the estimator is clear from this table. Because the random walk has a very high positive sample autocorrelation, the estimator of

The Monte Carlo distribution becomes closer to the normal approximation as the sample size is increased.

The generalized least squares estimator of $(\beta_0, \beta_1, \beta_2)$ was computed for each sample using the true value of ρ. The generalized least squares estimators of the regression parameters are conveniently obtained by a regression where the variables are defined in Table 1. The last $n-1$ observations are $(Y_t - \rho Y_{t-1}, 1, X_{1t} - \rho X_{1t-1}, X_{2t} - \rho X_{2t-1})$. For values of ρ other than -1 and 1 the first observation in the regression is that given in the first row of Table 1. When ρ is either 1 or -1 the first observation is deleted from the sample. We denote the generalized least squares estimator of (β_1, β_2) by $(\hat{\beta}_{1GLS}, \hat{\beta}_{2GLS})$.

For our model with estimated (β_1, β_2, ρ), it can be shown that the limiting distribution of the nonlinear least squares estimator of (β_1, β_2) is the same as the limiting distribution of the generalized least squares estimator constructed with known ρ. Table 5 compares the Monte Carlo variances of the nonlinear least squares estimator of (β_1, β_2) with the infeasible generalized least squares estimator of (β_1, β_2). The ratios of the empirical variance of the nonlinear least squares estimator to the empirical variance of the generalized least squares estimator for various values of ρ are given in the table. The convergence of the nonlinear least squares estimator to the limiting form is faster for the $N(0,1)$ variable than for the random walk explanatory variable. Convergence is slowest for ρ near to, but not too near to, one.

Table 5. Ratios of the variance of the nonlinear least squares estimator to the variance of the generalized least squares estimator.

	$n = 25$		$n = 100$	
ρ	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_1$	$\tilde{\beta}_2$
1.00	1.27	1.04	1.06	1.00
0.99	1.28	1.05	1.09	1.00
0.95	1.35	1.06	1.17	1.00
0.90	1.28	1.09	1.19	1.01
0.70	1.27	1.07	1.12	1.01
0.50	1.40	1.10	1.08	1.03
0.25	1.16	1.15	1.06	1.03
0.00	1.21	1.14	1.03	1.03
-0.25	1.12	1.09	1.00	1.04
-0.50	1.04	1.09	1.00	1.03
-0.70	1.04	1.04	1.00	1.01
-0.90	1.03	1.02	1.00	1.00

It can be shown that the limiting distribution of the test statistic (1.7) is that of a chi-square random variable with two degrees of freedom divided by its degrees of freedom when $|\rho| < 1$. This follows from the fact that the unrestricted estimators standardized by the square root of the sums of squares have limiting normal distributions.

If $\rho = 1$, the limiting distribution of the test statistic is the average of two random variables. The first variable in the test statistic is associated with the stationary variable and is a one-degree-of-freedom chi-square random variable. The second variable in the test statistic is independent of the first and the limiting distribution of the second variable is the distribution of the square of the pivotal given in (1.9). If $\rho = -1$, the limiting distribution of the test statistic is the average of two random variables, where the first is a chi-square random variable and the second is the square of variable like (1.9), but without the mean adjustment parts. The 5% and 95% points of the distribution of the test statistics are given in the last part of Table 6. The statistic denoted by $1/2(\chi^2 + \xi_\mu^2)$ is the statistic for $\rho = 1$. The percentage points are based upon a Monte Carlo run of size 10,000.

Table 6. Empirical percentiles of test statistic.

	$n = 25$		$n = 100$	
ρ	5%	95%	5%	95%
1.00	0.07	4.83	0.08	4.33
0.99	0.09	5.01	0.08	5.48
0.95	0.09	5.33	0.07	4.26
0.90	0.07	4.81	0.08	3.63
0.70	0.06	3.76	0.06	3.05
0.50	0.05	3.61	0.06	3.05
0.25	0.05	3.43	0.05	3.21
0.00	0.05	3.42	0.05	3.31
−0.25	0.05	3.55	0.05	3.02
−0.50	0.05	3.39	0.05	3.00
−0.70	0.05	3.31	0.05	2.90
−0.90	0.06	3.31	0.04	2.85
−0.95	0.04	3.10	0.05	2.87
−0.99	0.06	3.09	0.06	3.46
−1.00	0.05	3.36	0.06	2.95
$F(2, n-6)$	0.05	3.55	0.05	3.09
$1/2(\chi^2 + \xi_\mu^2)$	0.04	3.79	0.04	3.79

Table 6 contains the 5% and 95% points of the Monte Carlo distribution of the test statistics based upon 1000 samples. For $\rho = 0$, the percentiles of the test statistic are close to those of the F-distribution. For ρ close to one, the 95-th percentage points for both sample sizes are beyond that of the F-distribution and beyond the limiting distribution of the test statistic for $\rho = 1$. For ρ close to negative one, the percentage points are close to those of the F-distribution. For $\rho = +1$, and $n = 100$ the observed percentage points are in reasonable agreement with those of the limiting distributions.

REFERENCES

AHTOLA, J.A., AND TIAO, G.C., *Parameter inference for a nearly nonstationary first-order autoregressive model*, Biometrika 74 (1984), pp. 263–272.

CHAN, N.H., *The parameter inference for nearly nonstationary time series*, J. Amer. Statist. Assoc. 83 (1988), pp. 857–862.

CHAN, N.H. AND WEI, C.Z., *Asymptotic inference for nearly nonstationary $AR(1)$ processes*, Ann. Statist. 15 (1987), pp. 1050–1063.

CHAN, N.H. AND WEI, C.Z., *Limiting distributions of least squares estimates of unstable autoregressive processes*, Ann. Statist. 16 (1988), pp. 367–401.

DICKEY, D.A., AND FULLER, W.A., *Distribution of the estimators for autoregressive time series with a unit root*, J. Amer. Statist. Assoc. 74 (1979), pp. 427–431.

DICKEY, D.A., AND FULLER, W.A., *Likelihood ratio statistics for autoregressive time series with a unit root*, Econometrica 49 (1981), pp. 1056–1072.

FULLER, W.A., *Introduction to Statistical Time Series*, Wiley, New York (1976).

FULLER, W.A., *Nonstationary autoregressive time series*, In Handbook of Statistics 5, Time Series in the Time Domain. (E.J. Hannan, P.R. Krishnaiah and M.M. Rao, Eds.) North Holland, Amsterdam (1984).

FULLER, W.A., HASZA, D.P., AND GOEBEL, J.J., *Estimation of the parameters of stochastic difference equations*, Ann. Statist. 9 (1981), pp. 531–543.

KRÄMER, W., *Least squares regression when the independent variable follows an ARIMA process*, J. Amer. Statist. Assoc. 81 (1986), pp. 150–154.

NAGARAJ, N.K. AND FULLER, W.A., *Estimation of the parameters of linear time series models subject to nonlinear restrictions*, Ann. Statist. 19 (1991), pp. 1143–1154.

PHILLIPS, P.C.B., *Understanding spurious regressions in econometric*, J. Econometrics 33 (1986), pp. 311–340.

PHILLIPS, P.C.B., AND DURLAUF, S.N., *Multivariate time series with integrated variables*, Review of Economic Studies 53 (1986), pp. 473–496.

RAO, P., AND GRILICHES, Z., *Small sample properties of several two stage regression methods in the context of auto-correlated errors*, J. Amer. Statist. Assoc. 64 (1969), pp. 253–272.

SIMS, C.A., STOCK, J.H. AND WATSON, M.W., *Inference in linear time series models with some unit roots*, Econometrica 58 (1990), pp. 113–144.