Journal of the Korean Statistical Society (1998), **27**: *2, pp 165–187*

Testing for a unit root in an AR(p) signal observed with MA(q) noise when the MA parameters are unknown †

Dong-Bin Jeong[1] and Sahadeb Sarkar[2]

ABSTRACT

Shin and Sarkar (1993, 1994) studied the problem of testing for a unit root in an AR(p) signal observed with MA(q) noise when the MA parameters are known. In this paper we consider the case when the MA parameters are unknown and to be estimated. Test statistics are defined using unit root parameter estimates based on three different estimation methods of Hannan and Rissanen (1982), Kohn (1979) and Shin and Sarkar (1995). An AR(p) process contaminated by MA(q) noise is a restricted ARMA model, for which Shin and Sarkar (1995) derived an easy-to-compute Newton- Raphson estimator. The two-stage estimation procedure of Hannan and Rissanen (1982) is used to compute initial parameter estimates in implementing the iterative estimation methods of both Shin and Sarkar (1995) and Kohn (1979). In a simulation study we compare the relative performance of these unit root tests with respect to both size and power for p=q=1.

Keywords: Monte Carlo study; Newton-Raphson estimator; Restricted maximum likelihood estimator; Standard Brownian motion; Unit root test; Vector ARMA model

1. INTRODUCTION

Applications of the scalar or vector autoregressive moving average (ARMA) model can be found in various fields such as economics, survey sampling and engineering. The ARMA models with autoregressive unit roots are known to provide good stochastic approximations for many nonstationary time series. Sometimes

†The research of the first author was supported in part by a grant from Office of Research Support at Kangnung National University, Korea

[1]Department of Statistics, Kangnung National University, Kangnung, 210-702, Korea.

[2]IIM Calcutta, OM Group, Joka, D. H. Road, Post Box 16757, Calcutta 700 027, India and Department of Statistics, Oklahoma State University, Stillwater, OK, 74078-0595, USA.

the model parameters are restricted by some natural constraints. Shin and Sarkar (1995) discussed several examples of such restricted ARMA models in the areas of economics (DeJong and Husted 1993), signal measurement in the engineering sciences (Shin and Sarkar 1993), and rotational sampling as conducted, for instance, by the Current Population Survey by the US Bureau of Census (see Shin 1993 and Miazaki and Dorea 1993). In the Current Population Survey a 4-8-4 rotational sampling scheme in which sampling units entering the sample in month t are observed in months t, t+1, t+2, t+3 and t+12, t+13, t+14, t+15 before being removed from the sample forever. Under such a sampling scheme instead of observing the m×1 true values $\mathbf{z_t}$, one observes $\mathbf{y_t}$ which contains a sampling error $\mathbf{u_t}$. Shin and Sarkar (1995, Section 3) argue that one may model $\mathbf{z_t}$ and $\mathbf{u_t}$ as AR(p) and MA(r) processes respectively:

$$\mathbf{y_t} = \mathbf{z_t} + \mathbf{u_t}, \quad a(L)\mathbf{z_t} = \varepsilon_\mathbf{t}, \quad \mathbf{u_t} = c(L)\mathbf{w_t} \tag{1.1}$$

where ε_t and $\mathbf{w_t}$ are two independent sequences of independent, identically distributed (iid) random vectors with variance-covariance matrices $\Omega_{\varepsilon\varepsilon}$ and Ω_{ww} respectively,

$$a(L) = I + A_1 L + \cdots + A_p L_p, \quad c(L) = I + C_1 L + \cdots + C_r L^r,$$

I is the m× m identity matrix, $A_1, A_2, \ldots, A_p, C_1, C_2, \ldots, C_r$ are m×m matrices of unknown parameters, and L denotes the lag operator such that $L^k \mathbf{y_t} = \mathbf{y_{t-k}}$. Estimation problem of the univariate version of the model (1.1) has been studied by many authors including Pagano (1974), Sakai and Arase (1979), Dunsmuir (1979), Shin (1993), Miazaki and Dorea (1993), Binder and Hidiroglou (1987). Shin and Sarkar (1993, 1994) investigate testing for a unit root in an AR(p) signal contaminated by MA(q) noise when the MA parameters are known. They show that their unit root tests perform very well in small samples and compare favorably with the Said and Dickey (1985) tests with respect to both sizes and powers through a simulation study for the case p=q=1.

In this paper we consider model (1.1) for m=1, and study tests for an autoregressive unit root when the moving average parameters of the MA(q) noise are unknown. We examine three test statistics based on different methods of estimation of the unit root parameter using the univariate version of the works of Hannan and Rissanen (1982), Kohn (1979) and Shin and Sarkar (1995). Since an AR(p) process disturbed by MA(q) noise is a restricted ARMA model, parameters can be estimated using the Newton- Raphson estimator given by Shin and Sarkar (1995) as an approximation of the restricted maximum likelihood estimator. The two-stage estimation procedure of Hannan and Rissanen (1982) can be used to obtain starting parameter estimates. We also consider parameter estimation using the iterative estimation method of Kohn (1979), again using the Hannan-Rissanen procedure to generate initial parameter estimates.

The remainder of the paper is organized as follows. In Section 2 three different methods of estimation are introduced. Section 3 discusses the asymptotic distributions of parameter estimators and the unit root test statistics under the unit root hypothesis. Simulation results on the relative performance of our proposed tests for p=q=1 are presented in Section 4. Finally, some concluding remarks are given in Section 5.

2. PARAMETER ESTIMATION

In this section we briefly review three different estimation methods under the following general restricted m-vector ARMA(p,q) model, parameter estimation in which was studied by Kohn (1979) and Shin and Sarkar (1995):

$$a(L)\mathbf{y_t} = b(L)\mathbf{e_t}, \quad t = 1, 2, \dots, n, \tag{2.1}$$

where $\{\mathbf{e_t}\}$ is a sequence of iid random vectors with variance-covariance matrix Ω, and

$$a(L) = I + A_1 L + \cdots + A_p L_p, \quad b(L) = I + B_1 L + \cdots + B_q L^q. \tag{2.2}$$

It is assumed that (p, q) are known, nonnegative integers, and $A = (A_1|A_2|\cdots|A_p)$ $B = (B_1|B_2|\cdots|B_q)$ are respectively m×mp and m×mq matrices of unknown parameters. Let $\{\mathbf{y_t}\}$, t = 1,..., n be the set of m×1 observation vectors and $\{\mathbf{e_t}\}$ a sequence of iid m×1 random vectors with mean zero and a nonsingular variance-covariance matrix Ω. For stationarity, invertibility and identifiability, Shin and Sarkar also assumed that all the roots of det[a(L)] lie outside the unit circle and A_p is of full rank. The vector of restrictions on parameters A, B and Ω is defined through

$$\mathbf{f}(A, B, \Omega) = \mathbf{0} \tag{2.3}$$

where $\mathbf{f}$ is a k×1 vector of differentiable functions. Let $\alpha = (\alpha_1', \cdots, \alpha_p')'$, $\alpha_i = \mathrm{vec}(A_i')$, $\beta = (\beta_1', \cdots, \beta_q')'$, $\beta_i = \mathrm{vec}(B_i'), \theta = (\alpha', \beta')', \eta = \mathrm{vech}(\Omega)$, and $\xi = (\alpha', \beta', \eta')'$, where for a matrix R, vec(R) represents the vector obtained by piling up the columns of R and vech(R) denotes the vector obtained by stacking the columns of R with all elements above the diagonal removed. Letting $\mathbf{e_t} = 0$ for $t \leq 0$, (2.1) can be solved to get

$$\begin{aligned}\mathbf{e_t}(\theta) &= \mathbf{y_t} + A_1\mathbf{y_{t-1}} + \cdots + A_p\mathbf{y_{t-p}} - B_1\mathbf{e_{t-1}}(\theta) - \cdots - B_q\mathbf{e_{t-p}}(\theta) \\ &= \mathbf{y_t} + \sum_{i=1}^{p}(I \otimes \mathbf{y}_{\mathbf{t-1}}')\alpha_i - \sum_{i=1}^{p}(I \otimes \mathbf{e}_{\mathbf{t-1}}'(\theta))\beta_i\end{aligned}$$

where $\otimes$ denotes the kronecker product. In order to develop the conditional Gaussian likelihood of $\{\mathbf{y_t}\}$, t = 1,. . . ,n, the initial observations $\mathbf{y_0}$, $\mathbf{y_{-1}}, \ldots,$ $\mathbf{y_{1-p}}$ are assumed to be available and fixed and also the initial disturbances $\mathbf{e_0}$, $\mathbf{e_{-1}}, \ldots,$ e_{1-q} are assumed to be zero. Then the negative logarithm of the conditional Gaussian likelihood function of $\{\mathbf{y_t},\ t = 1, \ldots, n\}$ is approximated (Shin, 1993) by

$$L_n = L_n(\xi) = \frac{1}{2}\sum_{t=1}^{n}\mathbf{e_t}'(\theta)\Omega^{-1}\mathbf{e_t}(\theta) + \frac{1}{2}\mathrm{nlogdet}(\Omega)$$

As shown in Shin and Sarkar (1995, Sec 3), $\mathbf{y_t}$ of (1.1) can be written as a restricted ARMA(p, p+q) model

$$a(L)\mathbf{y_t} = b(L)\mathbf{e_t}$$

where $\{\mathbf{e_t}\}$ is a sequence of iid random vectors with a variance-covariance matrix Ω and the parameter restrictions are defined by

$$b(L)\Omega[b(L^{-1})]' - \Omega_{\varepsilon\varepsilon} - a(L)c(L)[c(L^{-1})]'[a(L^{-1})]' = \mathbf{0}.$$

For more details see Section 3 of Shin and Sarkar (1995).

2.1. Hannan and Rissanen's method

We may apply the first and second stages of the multivariate version of Hannan and Rissanen (1982) procedure ignoring parameter restrictions to compute an initial estimate $\tilde{\xi}$ of ξ. In the first stage, regressing $\mathbf{y_t}$ on $\mathbf{y_{t-1}}, \cdots, \mathbf{y_{t-s}}$, where s is chosen suitably so that the lag size is large enough for approximating $\mathbf{y_t}$ by previous observations. The residuals e_t are computed as estimates of the $\tilde{\mathbf{e}}_\mathbf{t}$. In the second stage, estimates $\tilde{A}$ and $\tilde{B}$ are computed by regressing $\mathbf{y_t}$ on $\{\mathbf{y_{t-1}}, \cdots, \mathbf{y_{t-p}}\}$ and $\{\tilde{\mathbf{e}}_{\mathbf{t-1}}, \cdots, \tilde{\mathbf{e}}_{\mathbf{t-q}}\}$. Then $\tilde{\theta} = \{(\text{vec}(\tilde{A}_1'))', \ldots, (\text{vec}(\tilde{A}_p'))', \ (\text{vec}(\tilde{B}_1'))', \ldots, (\text{vec}(\tilde{B}_q'))'\}$, $\tilde{\Omega} = n^{-1} \sum \mathbf{e_t}(\tilde{\theta})\mathbf{e_t}'(\tilde{\theta})$ and $\tilde{\eta} = \text{vech}(\tilde{\Omega})$.

2.2. Kohn's method

For a general vector linear time series model Kohn (1979) proved the strong consistency and asymptotic normality of parameter estimates obtained by maximizing an approximation to a Gaussian likelihood, where the observations are not necessarily assumed to be normally distributed. To estimate the parameters based on model (2.1)- (2.2) he used a Newton-Raphson procedure which takes restriction (2.3) into account. In particular, an efficient estimator is obtained if the iteration is initialized with a $\sqrt{n}$-consistent estimator. Let the first partial derivatives $\partial L_n/\partial\xi$ of L_n be denoted by L_ξ. Letting $W_{\theta,t} = \partial\mathbf{e_t}(\theta)/\partial\theta'$, $W_{\theta,t} = (W_{\alpha,t}|W_{\beta,t})$ is a m$\times\{m^2(p+q)\}$ matrix, and $L_\theta = \partial L_n/\partial\theta = \sum W_{\theta,t}'\Omega^{-1}\mathbf{e_t}(\theta)/n$. Analysis of the derivatives $L_\eta = \partial L_n/\partial\eta$ and $L_{\eta\eta} = \partial^2 L_n/\partial\eta\partial\eta'$ is based on Theorem 4.B.2 of Fuller (1987, p. 398). By (4.B.20) of Fuller (1987, p. 400),

$$L_\eta = -\Gamma^{-1}\text{vech}\left[\sum \mathbf{e_t}(\theta)\mathbf{e_t}'(\theta) - n\Omega\right]$$

where $\Gamma = 2\Psi(\Omega \otimes \Omega)\Psi'$ and Ψ is defined by $\text{vech}(\Omega) = \Psi\text{vec}(\Omega)$. Furthermore, it can be shown that

$$L_\xi = \left[\sum \mathbf{e_t}'(\theta)\Omega^{-1}W_{\theta,t}| - \text{vech}(\sum \mathbf{e_t}(\theta)e_t'(\theta) - n\Omega)'\Gamma^{-1}\right]' = H_\xi, \text{ say.}$$

A Newton-Raphson procedure for estimating the model (2.1)-(2.2), without restriction (2.3), is

$$\xi^+ = \tilde{\xi} - \left(\tilde{H}_{\xi\xi}\right)^{-1}\tilde{H}_\xi \tag{2.4}$$

where $\tilde{\xi}$ is an initial estimator of ξ. In order to get estimator $\tilde{\xi}$, the first and second stages of multivariate version of Hannan-Rissanen (1982) procedure may be applied to (2.4). Now the Newton-Raphson estimator ξ^+ in (2.4) is modified to accommodate restrictions (2.3) using the Largrangian multiplier method. The maximum likelihood estimator in this case minimize L_n subject to the restriction (2.3) and can be obtained by minimizing $L_n + \mathbf{f}'\lambda$, $\lambda \in R^k$ with respect to ξ and λ. Letting $G = \partial\mathbf{f}/\partial\xi'$, the following equations are obtained:

$$\tilde{L}_{\xi\xi}(\xi - \tilde{\xi}) + \tilde{G}'\lambda = -\tilde{L}_{\xi},\ \tilde{G}(\xi - \tilde{\xi}) = -\tilde{\mathbf{f}}$$

where $\tilde{L}_{\xi}$, $\tilde{L}_{\xi\xi}$, $\tilde{G}$ and $\tilde{\mathbf{f}}$ are the values of L_{ξ}, $L_{\xi\xi}$, G and $\mathbf{f}$ evaluated at $\xi = \tilde{\xi}$, the initial estimate. Therefore, the Newton-Raphson estimator is obtained as

$$\begin{bmatrix} \hat{\xi} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \tilde{\xi} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \tilde{L}_{\xi\xi} & \tilde{G}' \\ \tilde{G} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{L}_{\xi} \\ \tilde{\mathbf{f}} \end{bmatrix}. \tag{2.5}$$

For the square matrix in (2.5) to be nonsingular, it is assumed that G has full rank and this is equivalent to having no redundancy in the set of restrictions $\mathbf{f} = \mathbf{0}$.

2.3. Shin and Sarkar's method

Based on Kohn's method, Shin and Sarkar (1995) considered an approximation $H_{\xi\xi}$ of $L_{\xi\xi}$ in (2.5) to compute an easier-to-compute Newton-Raphson estimator that approximates the restricted ML estimator. Shin and Sarkar defined

$$H_{\xi\xi} = \text{diag}[\sum W'_{\theta,t}\Omega^{-1}W_{\theta,t}, -n\Gamma^{-1}]. \tag{2.6}$$

Approximation $\sum W'_{\theta,t}\Omega^{-1}W_{\theta,t}$ of $\partial^2 L_n/\partial\theta\partial\theta'$ is well explained in Reinsel et al. (1992), and by (4.B.22) of Fuller (1987) $L_{\eta\eta} = -n\Gamma^{-1} + o_p(n)$. Hence, $L_{\xi\xi} = H_{\xi\xi} + o_p(n)$.

2.4. Asymptotic distribution

Let $I(\theta)$ be the information matrix of θ in the unrestricted model (2.1)-(2.2) and $V = \text{diag}[I(\theta), \Gamma^{-1}]$. Then the limiting distribution of the estimator $\hat{\xi}$ obtained by Kohn's and Shin-Sarkar's methods follows from Shin and Sarkar (1995) under the regularity conditions assumed therein:

Theorem 2.1. *(due to Shin and Sarkar (1995, Theorem 1)). Assume that the model (2.1)-(2.3) holds, and assume that the* $\mathbf{e_t}$*'s are normal,* $\mathbf{f}$ *is continuously differentiable and* G *is nonsingular. Suppose that the initial estimator* $\tilde{\xi}$ *of* ξ *satisfies* $\sqrt{n}(\tilde{\xi}-\xi)=O_p(1)$. *Let* $(\hat{\xi},\hat{\lambda})$ *be as defined in (2.5). Then,*

$$\begin{bmatrix} \sqrt{n}(\hat{\xi}-\xi) \\ \frac{1}{\sqrt{n}}\hat{\lambda} \end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} V & G' \\ G & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} V & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} V & G' \\ G & \mathbf{0} \end{bmatrix}^{-1} \right).$$

Observe that

$$\begin{bmatrix} V & G' \\ G & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} V & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} V & G' \\ G & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} V^{-1}-V_gV_{gg}V_g & \mathbf{0} \\ \mathbf{0} & V_{gg} \end{bmatrix} \tag{2.7}$$

where $V_g = GV^{-1}$ *and* $V_{gg} = (GV^{-1}G')^{-1}$. *The upper left block of the variance covariance matrix in the right hand side of (2.7) corresponding to* ξ *is the same as the upper-left block of the matrix* $\begin{bmatrix} V & G' \\ G & \mathbf{0} \end{bmatrix}^{-1}$. *Since* $I(\theta)$ *is consistently estimated by* $n^{-1}\sum \hat{W}'_{\theta,t}\hat{\Omega}^{-1}\hat{W}_{\theta,t}$, *the variance-covariance matrix of* $\hat{\xi}$ *is estimated by the upper-left block of the matrix* $\begin{bmatrix} \hat{H}_{\xi\xi} & \hat{G}' \\ \hat{G} & \mathbf{0} \end{bmatrix}^{-1}$.

3. APPLICATION TO THE CASE p=q=1

For ease of presentation, we illustrate application of the estimation methods in constructing unit root test statistics under the univariate model (1.1) for the case p=1 and q=1, i.e.,

$$y_t = z_t + u_t, \quad a(L)z_t = \varepsilon_t, \quad u_t = c(L)w_t \tag{3.1}$$

where

$$c(L) = 1 + C_1L, \quad a(L) = 1 + A_1L \tag{3.2}$$

The null hypothesis of interest is $\mathrm{H}_0 : A_1 = -1$. The model (3.1) can be expressed as follows:

$$(1 + A_1 L)\mathbf{y_t} = (1 + B_1 L + B_2 L^2)e_t, \tag{3.3}$$

subject to the restrictions $\mathbf{f} = (f_1,\ f_2)' = [\sigma_e^2(B_1 + B_1B_2) - (A_1 + C_1 + A_1^2C_1 + A_1C_1^2), \sigma_e^2B_2 - A_1C_1]' = (0,0)'$, where $\{e_t\}$ is a sequence of iid $(0, \sigma_e^2)$ random variables. Now (3.1) and (3.2) can be expressed as follows:

$$e_t^* = e_t^*(\theta) = y_t + A_1 y_{t-1} - B_1 e_{t-1}(\theta) - B_2 e_{t-2}(\theta) \tag{3.4}$$

and

$$L_n^* = L_n^*(\xi) = \sum \mathbf{e}_t^*(\theta)^2/2\sigma_e^2 + n/2(\log_e \sigma_e^2).$$

Applying the formulae given for Kohn's method (Sec 2.2) to the model (3.1)-(3.3), we obtain

$$L_\xi = \partial L_n/\partial \xi = \begin{bmatrix} \sum W_{A_1,t}e_t^*/\sigma_e^2 \\ \sum W_{B_1,t}e_t^*/\sigma_e^2 \\ \sum W_{B_2,t}e_t^*/\sigma_e^2 \\ \frac{\sum e_t{}^{*2}}{2(\sigma_e^2)^2} + \frac{n}{2\sigma_e^2} \end{bmatrix},$$

and by differentiating (3.4) we compute the elements of $W_{A_1,t}, W_{B_1,t}$ and $W_{B_2,t}$ recursively as

$$\begin{aligned} W_{A_1,t} &= y_{t-1} - B_1 W_{A_1,t-1} - B_2 W_{A_1,t-2} \\ W_{B_1,t} &= -e_{t-1}^* - B_1 W_{B_1,t-1} - B_2 W_{B_1,t-2} \end{aligned}$$

and

$$W_{B_2,t} = -e_{t-2}^* - B_1 W_{B_2,t-1} - B_2 W_{B_2,t-2}.$$

Note that $L_{\xi\xi} = \partial^2 L_n/\partial\xi\partial\xi'$ can be expressed as follows:

$$L_{\xi\xi} = \begin{bmatrix} \mathbf{P} & \mathbf{q}' \\ \mathbf{q} & s \end{bmatrix},$$

where

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

$$\begin{aligned}
p_{11} &= \sum (W_{A_1,t}^2 + e_t^* WW_{A_1,t})/\sigma_e^2 \\
p_{12} &= \sum (W_{A_1,t} W_{B_1,t} + e_t^* WW_{A_1,B_1})/\sigma_e^2 \\
p_{13} &= \sum (W_{A_1,t} W_{B_2,t} + e_t^* WW_{A_1,B_2,t})/\sigma_e^2 \\
p_{21} &= \sum (W_{A_1,t} W_{B_1,t} + e_t^* WW_{A_1,B_1,t})/\sigma_e^2 \\
p_{22} &= \sum (W_{B_1,t}^2 + e_t^* WW_{B_1,t})/\sigma_e^2 \\
p_{23} &= \sum (W_{B_1,t} W_{B_2,t} + e_t^* WW_{B_1,B_2,t})/\sigma_e^2 \\
p_{31} &= \sum (W_{A_1,t} W_{B_2,t} + e_t^* WW_{A_1,B_2,t})/\sigma_e^2 \\
p_{32} &= \sum (W_{B_1,t} W_{B_2,t} + e_t^* WW_{B_1,B_2,t})/\sigma_e^2 \\
p_{33} &= \sum (W_{B_2,t}^2 + e_t^* WW_{B_2,t})/\sigma_e^2
\end{aligned}$$

with

$$WW_{A_1,B_1} = \frac{\partial W_{A_1,t}}{\partial B_1} = -W_{A_1,t-1} - B_1 WW_{A_1,B_1,t-1} - B_2 WW_{A_1,B_2,t-2},$$

$$WW_{A_1,B_2,t} = \frac{\partial W_{A_1,t}}{\partial B_2} = -W_{A_1,t-2} - B_1 WW_{A_1,B_1,t-1} - B_2 WW_{A_1,B_2,t-2},$$

$$WW_{B_1,t} = \frac{\partial W_{B_1,t}}{\partial B_1} = -2W_{B_1,t-1} - B_1 WW_{B_1,t-1} - B_2 WW_{B_1,t-2},$$

$$WW_{B_1,B_2,t} = \frac{\partial W_{B_1,t}}{\partial B_2} = -W_{B_2,t-1} - B_1 WW_{B_1,B_2,t-1} - W_{B_1,t-2} - B_2 WW_{B_1,B_2,t-2},$$

$$WW_{B_2,t} = \frac{\partial W_{B_2,t}}{\partial B_2} = -2W_{B_2,t-1} - B_1 WW_{B_2,t-1} - B_2 WW_{B_2,t-2},$$

$$q = \left[-\sum e_t^* W_{A_1,t}/(\sigma_e^2)^2 - \sum e_t^* W_{B_1,t}/(\sigma_e^2)^2 - \sum e_t^* W_{B_2,t}/(\sigma_e^2)^2 \right]$$

and

$$s = \sum (e_t^*)^2/(\sigma_e^2)^3 - n/(\sigma_e^2)^2.$$

Furthermore,

$$G = \partial \mathrm{f}/\partial \xi' = \begin{bmatrix} -1 - 2A_1C_1 & \sigma_e^2 + B_2\sigma_e^2 & B_1\sigma_e^2 & B_1 + B_1B_2 \\ -C_1 & 0 & \sigma_e^2 & B_2 \end{bmatrix}.$$

Now we can compute Kohn's estimate for model (3.3), by applying (2.5) with G, f and $L_{\xi\xi}$ described above. Similarly we can compute Shin-Sarkar's estimate for model (3.3) by using (2.6) with $H_{\xi\xi} = \text{diag}\ [\sum W^2_{A_1,t}/\sigma^2_e,\ \sum W^2_{B_1,t}/\sigma^2_e,\ \sum W^2_{B_2,t}/\sigma^2_e,\ n/2\sigma^4_e]$.
As a consequence of Theorem 2.1, the limiting distributions of the test statistics $n(\hat{A}_1 - 1)$ and $\hat{\tau} = \hat{A}_1/\text{s.e}(\hat{A}_1)$ under the unit root null hypothesis are given by the following.

Theorem 3.1. *Under the model (3.3) with restrictions* $f_1 = \sigma^2_e(B_1 + B_1B_2) - (A_1 + C_1 + A_1^2C_1 + A_1C_1^2) = 0$, $f_2 = \sigma^2_eB_2 - A_1C_1 = 0$. *Let* $\hat{A}_1$ *denote the estimate of* A_1 *using either Kohn's method or Shin-Sarkar's method. The limiting distribution of* $n(\hat{A}_1 - 1)$ *and* $\hat{\tau}$ *under the null hypothesis* $H_0 : A_1 = -1$ *are given by* $\frac{\frac{1}{2}\{W(1)^2-1\}}{\int_0^1 W(r)^2dr}$ *and* $\frac{\frac{1}{2}\{W(1)^2-1\}}{\sqrt{\int_0^1 W(r)^2dr}}$ *respectively, where* $W(\cdot)$ *denotes the standard Brownian motion on [0, 1].*

4. SIMULATION RESULTS

We now consider a Monte Carlo study on the power functions of our unit root tests under model (3.1)-(3.3). Ten thousand replications were simulated for different combinations of $(n,\ A_1,\ C_1,\ \sigma^2_e)$ and for nominal levels 1%, 5% and 10%. We used sample size n = 25, 50, 100, 250, A_1 = 1.00, 0.99, 0.95, 0.90 and 0.70, C_1 = -0.5, 0, 0.5, and the value of σ^2_e was set to 0.2, 1.0 and 5.0. The normal random numbers $\{\varepsilon_t\}$ and $\{w_t\}$ were generated by the subroutine DRNNOA of the IMSL package. The value of σ^2_w was set to one. The value of C_1 was set to -0.5, 0 and 0.5, and the value of σ^2_e was set to 0.2, 1.0 and 5.0. For different $(n,\ C_1,\ \sigma^2_e)$ combinations independent samples were used. The computed values of the tests $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ were compared to the theoretical 1%, 5% and 10% left tail critical values tabulated by Dickey and Fuller (see Fuller 1976, p. 371 and p. 373). Simulation results on the empirical levels and powers of the test statistics $n(\hat{A}_1 - 1)$ and $\hat{\tau}$ are presented in Tables 4.1-4.6. We now discuss our findings based on the tables.

First we consider the common points among three methods: (1) When we fix n and increase σ^2_e or fix σ^2_e and increase n, the empirical powers get bigger and empirical sizes get closer to the nominal level for both $n(\hat{A}_1 - 1)$ and $\hat{\tau}$. In particular, for large samples (n = 100, 250), the empirical sizes of each of three methods are very close to the desired nominal levels except that $\sigma^2_e = 0.2$. (2) For $\sigma^2_e = 0.2$, i.e., when the signal is weak, especially when $C_1 = 0.5$, the size tends to be larger than the nominal level. For fixed $(n,\ C_1,\ \sigma_e^2)$ the powers are monotone functions of A_1. (3) The comparative performance of the tests for nominal levels 1%, 5%, and 10% are similar.

Next we discuss the differences among three methods: (1) For the "normalized" unit root test statistic $n(\hat{A}_1 - 1)$, the empirical sizes of the test using Kohn's estimate are very close to those using Shin and Sarkar's estimate. For the Kohn's method the sizes are usually slightly smaller. As expected, the empirical powers based on Shin and Sarkar's method generally are larger than those based on Kohn's due to the larger empirical sizes. There are no big differences between normalized tests based on Kohn's and Shin and Sarkar's estimates. (2) For the t-test statistic $\hat{\tau}$, the sizes using Kohn's estimates are in general much smaller than those using Shin and Sarkar's. The powers for Shin and Sarkar's method become closer as the sample size gets larger and C_1 gets smaller.

5. CONCLUDING REMARKS

For small sample sizes 25 and 50, all three different methods (H-R, Kohn, S-S) lead to test statistics for which the tabulated critical values given by Dickey and Fuller (1979) are not appropriate. Numerical results of Schwert (1989) show that the distribution of unit root test statistics in unrestricted ARMA models containing a nonzero MA component can be very different from those tabulated by Dickey and Fuller (1979). That this is also true in case of restricted ARMA models can be seen in Tables 4.7-4.8 for p=q=1 case.

The advantage of using Shin and Sarkar's method lies in the fact that it is much easier than Kohn's for computation in terms of inverting the Hessian matrix. If p, q are bigger than the case considered in our simulation study gain in computational ease will be even more noticeable. In terms of statistic $n(\hat{A}_1 - 1)$, the Shin and Sarkar method can be a good alternative to the Kohn's method when we test for a unit root in an AR(1) signal observed with MA(1) noise. On the other hand, the Kohn's method is preferable while using the t-test statistic $\hat{\tau}$.

ACKNOWLEDGMENTS

The authors would like to thank Dr. D. W. Shin for kindly suggesting the problem and wish to thank the SERI supercomputer center for providing the CRAY C90 access for our research.

Table 4.1 : Empirical Size(%) of $n(\hat{A}_1 - 1)$ for nomial level ($A_1 = 1$)

(a) nominal level 0.01

Sample Size	Method \ c_1	$\sigma_e^2 = 0.2$ -0.5	0.0	0.5	$\sigma_e^2 = 1.0$ -0.5	0.0	0.5	$\sigma_e^2 = 5.0$ -0.5	0.0	0.5
	HR	17.1	19.5	23.6	6.9	7.5	9.5	3.6	4.2	4.4
25	Kohn	16.5	18.3	20.8	6.6	7.4	8.7	3.6	4.1	4.3
	SS	17.0	18.5	23.6	6.8	7.4	9.5	3.6	4.1	4.3
	HR	11.5	10.4	12.8	4.2	4.1	4.7	2.5	2.6	3.0
50	Kohn	11.2	10.2	12.5	3.9	3.8	4.4	2.4	2.5	2.5
	SS	11.5	10.3	12.0	4.1	3.8	4.7	2.5	2.5	2.9
	HR	6.4	5.5	7.7	2.6	2.4	2.8	1.9	1.6	1.8
100	Kohn	6.0	5.2	7.4	2.5	2.3	2.4	1.8	1.6	1.7
	SS	6.3	5.3	6.1	2.6	2.4	2.7	1.9	1.6	1.8
	HR	2.9	2.3	3.2	1.4	1.5	1.3	1.3	1.2	1.1
250	Kohn	2.4	2.0	2.5	1.3	1.4	1.2	1.3	1.0	1.0
	SS	2.8	2.1	2.9	1.4	1.4	1.2	1.3	1.1	1.1

(b) nominal level 0.05

	HR	26.1	27.8	31.2	13.6	14.0	15.9	8.9	9.4	10.2
25	Kohn	25.4	26.6	29.0	13.2	13.9	15.2	8.8	9.4	10.0
	SS	25.7	27.7	31.1	13.5	13.9	15.8	8.8	9.3	10.1
	HR	18.5	18.8	20.3	9.6	9.5	10.6	7.5	7.5	7.6
50	Kohn	18.0	17.8	20.0	9.2	9.1	9.8	7.4	7.2	7.4
	SS	18.4	16.3	18.2	9.5	9.3	10.1	7.3	7.4	7.5
	HR	13.0	11.7	14.9	7.7	6.8	8.1	6.6	5.6	6.3
100	Kohn	12.3	10.8	13.7	7.5	6.6	7.3	6.5	5.5	6.2
	SS	12.9	11.2	12.8	7.6	6.8	7.8	6.5	5.6	6.3
	HR	8.0	7.1	8.6	5.9	5.7	6.0	5.5	5.4	5.3
250	Kohn	7.1	6.5	7.3	5.8	5.6	5.9	5.4	5.3	5.2
	SS	7.9	6.6	8.1	5.8	5.7	5.8	5.5	5.4	5.2

Table 4.2 : Empirical Size(%) of $\hat{\tau}$ for nomial level ($A_1 = 1$)

(a) nominal level 0.01

Sample Size	Method (c_1)	$\sigma_e^2 = 0.2$ -0.5	0.0	0.5	$\sigma_e^2 = 1.0$ -0.5	0.0	0.5	$\sigma_e^2 = 5.0$ -0.5	0.0	0.5
25	HR	21.1	26.3	31.0	8.1	9.5	14.1	4.0	4.6	5.5
	Kohn	11.8	6.5	11.5	5.2	3.1	3.3	3.3	3.1	2.4
	SS	20.8	26.1	30.2	8.0	9.4	14.0	3.9	4.6	5.5
50	HR	13.9	14.8	14.3	4.7	4.7	8.0	2.8	2.7	3.4
	Kohn	6.6	2.2	3.4	2.5	1.3	1.5	2.1	2.0	1.7
	SS	13.8	14.6	13.0	4.6	4.6	8.0	2.7	2.7	3.3
100	HR	8.1	9.0	6.8	3.0	2.9	5.4	1.0	1.8	2.1
	Kohn	2.6	1.1	2.5	1.7	1.1	1.3	1.6	1.4	1.4
	SS	8.0	8.7	5.7	2.9	2.9	5.3	1.9	1.7	2.0
250	HR	3.5	3.3	4.1	1.5	1.6	2.5	1.3	1.2	1.1
	Kohn	1.0	0.8	2.0	1.0	1.0	0.9	1.1	1.0	1.0
	SS	3.3	3.2	3.7	1.5	1.5	2.4	1.2	1.2	1.1

(b) nominal level 0.05

25	HR	29.8	33.9	35.9	15.0	16.6	20.6	9.6	10.1	11.4
	Kohn	17.6	11.4	16.2	10.0	7.8	7.8	8.2	8.0	7.5
	SS	29.2	33.5	34.4	14.9	16.4	20.3	9.5	10.0	11.4
50	HR	21.0	20.6	20.6	10.4	10.6	13.8	7.7	7.7	8.3
	Kohn	10.7	6.9	9.6	6.6	5.6	6.2	6.8	6.4	6.2
	SS	20.7	20.2	18.3	10.3	10.5	13.7	7.6	7.6	8.2
100	HR	15.1	14.7	16.1	8.1	7.8	10.7	6.7	5.9	6.6
	Kohn	6.6	5.0	9.4	5.8	4.9	5.7	6.0	5.2	5.6
	SS	15.0	14.4	14.2	8.0	7.6	10.6	6.6	5.8	6.6
250	HR	8.2	7.3	10.8	6.1	6.0	7.3	5.5	5.4	5.4
	Kohn	4.3	4.7	7.0	5.0	5.0	5.3	5.2	5.1	4.9
	SS	8.2	7.1	9.9	6.0	5.9	7.3	5.4	5.4	5.3

Table 4.3 : Empirical Power(%) of $n(\hat{A}_1 - 1)$ for nomial level 0.01

Sample Size	A_1	Method \ c_1	$\sigma_e^2 = 0.2$ -0.5	0.0	0.5	$\sigma_e^2 = 1.0$ -0.5	0.0	0.5	$\sigma_e^2 = 5.0$ -0.5	0.0	0.5
25	0.99	HR	18.8	19.6	24.4	7.4	7.7	9.6	3.8	4.7	5.3
		Kohn	18.1	21.8	28.8	7.3	8.2	10.3	3.3	4.7	5.3
		SS	18.9	20.1	26.7	7.5	7.7	10.2	3.8	4.7	5.3
	0.90	HR	34.7	37.6	45.7	17.7	18.0	20.2	10.6	10.5	11.8
		Kohn	32.3	39.3	49.2	17.2	18.0	20.8	10.5	10.6	11.7
		SS	34.9	38.9	48.8	17.7	18.0	21.3	10.7	10.5	11.8
50	0.99	HR	14.3	12.3	14.2	5.4	5.0	5.4	3.2	3.0	3.1
		Kohn	14.2	14.2	18.4	5.5	5.1	5.6	3.4	3.2	3.5
		SS	14.5	13.1	14.3	5.4	5.2	5.5	3.2	3.3	3.2
	0.90	HR	42.7	40.4	46.3	22.0	21.2	23.6	15.6	15.8	16.4
		Kohn	41.7	42.7	51.0	22.0	21.3	23.2	16.1	16.0	16.1
		SS	42.8	41.3	46.6	22.1	21.6	25.8	15.6	15.9	16.4
100	0.99	HR	9.6	8.3	9.7	3.7	3.5	3.7	2.4	2.6	2.5
		Kohn	9.3	8.1	11.1	3.6	3.4	3.5	2.5	2.7	2.5
		SS	9.8	8.4	9.5	3.7	3.6	3.7	2.5	2.7	2.6
	0.90	HR	58.1	57.1	60.5	40.9	40.8	44.5	36.7	35.9	36.6
		Kohn	56.7	55.8	62.6	40.5	39.6	41.2	36.7	35.9	36.5
		SS	58.1	57.2	59.7	41.3	40.8	44.9	36.8	35.9	36.7
250	0.99	HR	7.3	6.2	9.4	4.3	4.2	5.0	3.8	3.8	4.0
		Kohn	6.5	5.9	7.5	4.4	4.3	4.7	3.8	3.8	4.0
		SS	7.4	6.2	9.1	4.4	4.3	5.0	3.8	3.8	4.0
	0.90	HR	83.1	84.7	88.8	93.8	94.5	95.2	96.3	96.4	96.0
		Kohn	83.5	87.2	90.0	93.9	94.6	95.2	96.3	96.4	96.1
		SS	83.3	84.4	88.3	93.9	94.5	95.2	96.3	96.4	96.0

Table 4.4 : Empirical Power(%) of $n(\hat{A}_1 - 1)$ for nomial level 0.05

			$\sigma_e^2 = 0.2$			$\sigma_e^2 = 1.0$			$\sigma_e^2 = 5.0$		
Sample		c_1	-0.5	0.0	0.5	-0.5	0.0	0.5	-0.5	0.0	0.5
Size	A_1	Method									
25	0.99	HR	29.0	28.5	33.1	14.7	14.5	16.4	10.0	10.3	11.1
		Kohn	28.2	30.6	37.3	14.6	14.9	16.7	10.1	10.5	11.0
		SS	28.9	29.2	34.4	14.7	14.5	16.9	10.1	10.3	11.1
	0.90	HR	49.5	50.9	58.3	32.4	30.9	34.7	24.4	23.0	24.5
		Kohn	48.3	52.9	61.9	32.3	30.6	34.9	24.5	23.3	24.5
		SS	49.6	52.0	59.1	32.4	31.0	35.5	24.5	23.0	24.5
50	0.99	HR	22.7	20.1	21.5	12.2	11.5	12.3	9.5	8.8	9.3
		Kohn	22.1	22.1	27.4	12.2	11.3	12.2	9.6	9.1	9.2
		SS	22.7	20.3	21.2	12.2	11.7	12.3	9.5	9.0	9.3
	0.90	HR	57.7	55.3	58.5	42.2	41.6	44.3	37.3	37.3	37.8
		Kohn	57.2	58.1	64.5	42.0	41.7	43.1	37.5	37.7	37.7
		SS	57.6	56.1	58.6	42.5	41.8	44.3	37.8	37.3	37.8
100	0.99	HR	18.8	17.0	20.9	10.7	10.9	11.9	9.0	9.5	9.4
		Kohn	18.0	16.8	20.7	10.6	10.7	11.1	9.0	9.5	9.3
		SS	18.8	17.0	19.9	10.7	11.0	11.9	9.0	9.5	9.4
	0.90	HR	74.4	74.3	78.7	71.8	72.2	75.3	72.2	72.0	72.3
		Kohn	74.2	74.5	79.3	71.8	71.9	73.2	72.2	72.0	72.2
		SS	74.2	74.3	77.5	72.0	72.2	75.3	72.2	72.1	72.3
250	0.99	HR	19.6	17.7	23.0	15.8	15.9	17.0	15.3	15.2	15.7
		Kohn	17.7	17.7	19.7	16.0	15.8	16.6	15.3	15.2	15.8
		SS	19.8	17.8	22.0	15.9	15.9	17.0	15.3	15.2	15.8
	0.90	HR	90.0	93.3	95.2	99.5	99.6	99.7	99.9	99.9	99.9
		Kohn	91.7	94.6	95.9	99.6	99.6	99.8	99.9	99.9	99.9
		SS	90.8	93.7	95.9	99.5	99.6	99.7	99.9	99.9	99.9

Table 4.5 : Empirical Power(%) of $\hat{\tau}$ for nomial level 0.01

Sample Size	A_1	Method / c_1	$\sigma_e^2 = 0.2$ -0.5	0.0	0.5	$\sigma_e^2 = 1.0$ -0.5	0.0	0.5	$\sigma_e^2 = 5.0$ -0.5	0.0	0.5
25	0.99	HR	13.7	6.8	13.6	5.9	3.2	3.9	3.5	3.4	2.9
		Kohn	12.9	6.9	14.3	5.5	3.2	4.0	3.4	3.4	2.9
		SS	22.8	27.1	33.8	8.8	9.7	15.0	4.2	5.1	6.3
	0.90	HR	23.0	13.1	25.3	13.1	7.3	8.2	9.4	8.0	7.0
		Kohn	20.7	12.1	24.5	12.2	6.8	7.9	9.2	7.8	6.8
		SS	42.4	49.5	57.5	20.6	22.1	31.8	11.7	11.7	14.2
50	0.99	HR	8.1	3.0	4.2	3.4	2.3	2.1	2.9	2.4	2.3
		Kohn	8.0	3.4	5.1	3.7	2.1	2.2	3.1	2.5	2.4
		SS	17.6	18.4	15.9	6.1	6.4	9.3	3.4	3.5	3.8
	0.90	HR	24.3	11.1	16.4	14.4	9.8	9.9	13.9	13.1	12.3
		Kohn	23.1	11.2	15.4	13.4	10.0	9.2	13.7	12.8	11.8
		SS	50.8	52.2	46.7	25.1	26.5	36.4	17.1	17.4	18.8
100	0.99	HR	3.9	1.6	4.1	2.3	1.6	2.2	2.1	2.0	1.9
		Kohn	3.8	1.6	4.1	2.2	1.6	2.0	2.2	2.0	1.9
		SS	12.1	13.2	9.4	3.9	4.5	7.6	2.6	2.7	2.6
	0.90	HR	26.1	16.1	22.9	26.2	25.0	29.0	33.1	31.8	31.4
		Kohn	24.9	15.5	21.5	25.7	24.0	25.9	33.2	31.9	31.1
		SS	66.6	61.6	69.8	44.9	47.5	60.9	37.8	37.5	39.6
250	0.99	HR	3.2	3.2	8.8	3.2	3.0	3.9	3.8	3.4	3.5
		Kohn	2.9	2.8	6.0	3.2	2.9	3.5	3.8	3.4	3.6
		SS	9.7	8.5	11.1	4.7	4.8	8.3	4.1	3.9	4.3
	0.90	HR	50.3	63.0	80.3	87.9	90.1	91.3	95.5	95.5	95.1
		Kohn	48.2	60.6	75.2	87.8	90.1	90.9	95.5	95.5	95.2
		SS	80.6	71.2	82.5	94.7	96.1	91.9	96.4	96.7	96.6

Table 4.6 : Empirical Power(%) of $\hat{\tau}$ for nomial level 0.05

Sample Size	A_1	Method / c_1	$\sigma_e^2 = 0.2$ -0.5	0.0	0.5	$\sigma_e^2 = 1.0$ -0.5	0.0	0.5	$\sigma_e^2 = 5.0$ -0.5	0.0	0.5
25	0.99	HR	20.1	11.9	18.8	11.0	7.9	8.7	9.3	8.8	8.6
		Kohn	19.0	12.2	19.7	10.7	7.9	8.6	9.4	8.9	8.5
		SS	32.5	35.1	38.6	16.0	17.2	22.5	10.5	11.1	12.5
	0.90	HR	32.9	22.5	34.2	24.2	17.9	18.8	22.6	20.2	19.5
		Kohn	30.4	21.7	33.2	23.3	17.4	18.4	22.7	20.2	19.2
		SS	55.8	59.5	61.7	35.8	36.0	44.6	26.0	25.0	27.6
50	0.99	HR	13.5	7.7	10.9	8.8	7.4	7.9	9.3	7.7	7.8
		Kohn	13.3	8.4	12.0	9.2	7.4	7.7	8.6	7.9	7.7
		SS	25.6	25.4	22.0	13.2	13.3	16.7	9.9	9.4	10.0
	0.90	HR	35.8	24.4	31.0	31.3	29.0	30.7	35.0	34.7	33.5
		Kohn	34.4	24.3	30.4	30.2	28.5	29.3	35.4	34.8	33.1
		SS	63.7	61.5	54.6	45.5	47.2	53.5	39.7	39.2	40.7
100	0.99	HR	10.1	8.8	18.5	8.5	8.6	9.8	8.5	8.7	8.7
		Kohn	9.5	8.3	16.0	8.3	8.3	8.8	8.5	8.7	8.5
		SS	22.0	20.6	21.4	11.6	12.5	16.1	9.2	9.9	10.0
	0.90	HR	45.8	44.6	58.8	61.0	62.8	68.0	70.0	69.8	69.2
		Kohn	43.6	42.5	54.7	60.3	61.5	63.9	69.9	69.7	69.7
		SS	76.5	70.4	69.8	74.7	77.2	79.4	73.2	73.2	74.3
250	0.99	HR	12.8	14.5	24.8	14.1	14.2	16.0	15.0	14.7	15.5
		Kohn	11.1	13.0	18.5	14.1	14.1	15.2	15.1	14.7	15.4
		SS	22.0	18.0	26.4	16.5	16.9	20.8	15.6	15.3	16.3
	0.90	HR	78.4	87.4	92.8	98.7	99.2	97.2	99.8	99.8	99.8
		Kohn	77.7	87.5	91.7	98.8	99.2	99.5	99.8	99.8	99.8
		SS	86.4	87.1	91.0	99.4	99.5	97.2	99.9	99.9	99.8

Table 4.7 : Empirical Cumulative Distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$

(a) n=25

	Probability of a Smaller Value							
(σ_e^2, C_1)	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-46.78	-33.53	-25.30	-17.58	2.93	5.65	10.10	20.77
(0.2, 0.0)	-61.67	-42.89	-31.44	-21.93	2.84	5.63	10.47	20.42
(0.2, 0.5)	-57.06	-43.22	-35.27	-26.87	3.24	6.60	11.62	20.78
(1.0,-0.5)	-30.25	-20.47	-14.21	-9.32	1.63	2.51	3.57	6.13
(1.0, 0.0)	-32.31	-22.39	-15.72	-9.75	1.58	2.36	3.61	6.40
(1.0, 0.5)	-36.23	-25.48	-18.24	-11.42	1.58	2.42	3.65	6.26
(5.0,-0.5)	-21.78	-13.93	-10.15	-6.77	1.40	1.93	2.60	3.66
(5.0, 0.0)	-22.27	-15.42	-10.85	-7.02	1.39	1.95	2.57	3.57
(5.0, 0.5)	-22.89	-15.78	-11.13	-7.37	1.41	1.99	2.60	3.71

(b) n=50

	Probability of a Smaller Value							
(σ_e^2, C_1)	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-58.75	-38.39	-25.78	-14.72	1.89	3.15	5.73	12.97
(0.2, 0.0)	-54.25	-35.53	-23.34	-13.30	1.84	3.05	6.03	11.93
(0.2, 0.5)	-61.03	-43.09	-29.56	-15.89	2.32	4.87	9.58	16.65
(1.0,-0.5)	-25.83	-16.61	-11.50	-7.45	1.19	1.67	2.21	2.90
(1.0, 0.0)	-24.29	-15.84	-11.03	-7.31	1.19	1.65	2.10	2.80
(1.0, 0.5)	-27.41	-17.84	-12.42	-7.75	1.22	1.70	2.25	3.08
(5.0,-0.5)	-18.78	-12.96	-9.44	-6.41	1.10	1.51	1.94	2.46
(5.0, 0.0)	-17.86	-13.02	-9.42	-6.32	1.12	1.54	1.91	2.46
(5.0, 0.5)	-18.63	-13.72	-9.76	-6.53	1.14	1.62	2.03	2.59

Table 4.7(continue) Empirical Cumulative Distribution of $n(\hat{A}_1 - 1)$ for $A_1 = 1$

(c) n=100

	Probability of a Smaller Value							
(σ_e^2, C_1)	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-38.05	-24.53	-15.73	-9.77	1.16	1.67	2.22	3.02
(0.2, 0.0)	-32.98	-20.88	-14.00	-8.76	1.12	1.62	2.13	2.79
(0.2, 0.5)	-33.37	-20.89	-15.05	-9.57	1.06	1.54	2.09	3.04
(1.0,-0.5)	-19.60	-13.61	-9.95	-6.67	1.00	1.42	1.80	2.32
(1.0, 0.0)	-18.52	-13.07	-9.28	-6.24	1.01	1.43	1.89	2.34
(1.0, 0.5)	-19.72	-13.76	-10.10	-6.76	1.00	1.39	1.74	2.24
(5.0,-0.5)	-16.90	-11.72	-9.02	-6.19	1.00	1.39	1.76	2.23
(5.0, 0.0)	-15.84	-11.44	-8.39	-5.88	1.01	1.40	1.84	2.35
(5.0, 0.5)	-15.89	-11.92	-8.87	-6.09	1.0.	1.39	1.74	2.17

(d) n=250

	Probability of a Smaller Value							
(σ_e^2, C_1)	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-20.53	-14.41	-10.22	-6.76	1.04	1.44	1.87	2.39
(0.2, 0.0)	-17.86	-12.92	-9.46	-6.33	1.03	1.43	1.83	2.35
(0.2, 0.5)	-20.21	-14.24	-10.46	-7.02	0.97	1.39	1.78	2.31
(1.0,-0.5)	-15.25	-11.41	-8.59	-5.97	0.98	1.35	1.67	2.10
(1.0, 0.0)	-14.78	-11.17	-8.54	-5.91	0.95	1.33	1.72	2.14
(1.0, 0.5)	-14.62	-11.02	-8.54	-5.95	0.94	1.32	1.68	2.12
(5.0,-0.5)	-14.44	-11.05	-8.28	-5.73	0.97	1.34	1.65	2.06
(5.0, 0.0)	-14.28	-10.80	-8.26	-5.78	0.96	1.33	1.71	2.16
(5.0, 0.5)	-13.86	-10.45	-8.23	-5.80	0.96	1.32	1.67	2.09

Table 4.8 : Empirical Cumulative Distribution of $\hat{\tau}$ for $A_1 = 1$

(a) n=25

	Probability of a Smaller Value							
$(\sigma_e^2,\ C_1)$	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-20.21	-9.47	-6.08	-4.17	4.34	20.11	74.92	4804
(0.2, 0.0)	-318.0	-15.72	-8.78	-5.63	4.06	17.90	70.03	4775
(0.2, 0.5)	-186.4	-28.78	-16.07	-10.22	4.80	23.64	109.2	12684
(1.0,-0.5)	-6.98	-4.41	-3.24	-2.40	1.92	3.59	7.90	25.87
(1.0, 0.0)	-7.79	-4.94	-3.67	-2.58	1.88	3.49	6.87	26.10
(1.0, 0.5)	-12.66	-7.29	-5.11	-3.39	1.81	3.23	6.28	22.32
(5.0,-0.5)	-4.33	-3.05	-2.45	-1.91	1.51	2.41	3.51	6.10
(5.0, 0.0)	-4.49	-3.35	-2.59	-1.96	1.57	2.57	3.64	6.30
(5.0, 0.5)	-5.00	-3.59	-2.73	-2.08	1.51	2.42	3.60	6.41

(b) n=50

	Probability of a Smaller Value							
$(\sigma_e^2,\ C_1)$	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-9.85	-6.37	-4.53	-3.16	2.21	4.90	20.79	217.5
(0.2, 0.0)	-10.04	-6.78	-4.97	-3.49	2.12	4.28	20.67	193.5
(0.2, 0.5)	-17.93	-11.97	-8.53	-4.18	2.80	15.86	96.14	9403
(1.0,-0.5)	-4.27	-3.20	-2.55	-1.97	1.29	1.99	2.78	3.95
(1.0, 0.0)	-4.24	-3.21	-2.57	-2.00	1.25	1.88	2.62	4.09
(1.0, 0.5)	-5.86	-4.17	-3.24	-2.31	1.30	1.96	2.70	3.98
(5.0,-0.5)	-3.33	-2.67	-2.21	-1.78	1.12	1.70	2.28	2.97
(5.0, 0.0)	-3.23	-2.68	-2.22	-1.76	1.15	1.63	2.18	2.94
(5.0, 0.5)	-3.46	-2.86	-2.34	-1.81	1.17	1.71	2.23	3.01

Table 4.8(continue) Empirical Cumulative Distribution of $\hat{\tau}$ for $A_1 = 1$

(c) n=100

	Probability of a Smaller Value							
$(\sigma_e^2,\ C_1)$	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-5.35	-4.00	-3.15	-2.40	1.14	1.76	2.47	3.94
(0.2, 0.0)	-6.09	-4.72	-3.47	-2.44	1.11	1.73	2.31	3.33
(0.2, 0.5)	-7.92	-3.32	-2.71	-2.21	1.06	1.70	2.37	3.51
(1.0,-0.5)	-3.32	-2.71	-2.26	-1.81	1.00	1.52	1.96	2.54
(1.0, 0.0)	-3.32	-2.70	-2.21	-1.77	1.02	1.50	1.98	2.55
(1.0, 0.5)	-4.18	-3.29	-2.68	-2.00	1.00	1.46	1.93	2.52
(5.0,-0.5)	-2.95	-2.45	-2.10	-1.71	0.98	1.44	1.85	2.33
(5.0, 0.0)	-2.90	-2.41	-2.03	-1.67	1.01	1.46	1.86	2.41
(5.0, 0.5)	-2.92	-2.48	-2.12	-1.71	1.01	1.41	1.82	2.34

(d) n=250

	Probability of a Smaller Value							
$(\sigma_e^2,\ C_1)$	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
(0.2,-0.5)	-3.55	-2.85	-2.33	-1.84	0.98	1.44	1.89	2.42
(0.2, 0.0)	-4.00	-2.90	-2.20	-1.73	1.01	1.43	1.86	2.38
(0.2, 0.5)	-3.32	-2.82	-2.41	-1.95	0.96	1.44	1.89	2.39
(1.0,-0.5)	-2.77	-2.38	-2.04	-1.67	0.94	1.37	1.79	2.21
(1.0, 0.0)	-2.77	-2.37	-2.05	-1.68	0.94	1.33	1.74	2.15
(1.0, 0.5)	-3.13	-2.58	-2.17	-1.75	0.94	1.37	1.74	2.16
(5.0,-0.5)	-2.68	-2.32	-1.99	-1.63	0.93	1.35	1.76	2.17
(5.0, 0.0)	-2.65	-2.29	-1.99	-1.63	0.93	1.32	1.72	2.09
(5.0, 0.5)	-2.64	-2.26	-1.98	-1.65	0.93	1.35	1.71	2.11

REFERENCES

Binder, D. A. and Hidiroglou M. A. (1987). Sampling in Time, *Handbook of Statistics*, 6, 187-211. New York: North Holland.

Dejong, D. N. and Husted, S. (1993). "Towards a reconciliation of the empirical evidence on the monetary approach to exchange rate determination," *Review of Economics and Statistics*, 123-129.

Dickey, D. A. and Fuller, W. A. (1979). "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association,* **74**, 427-431.

Dunsmuir, W. (1979). "A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise," *Annals of Statistics,* **7**, 490-506.

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

Fuller, W. A. (1987). *Measurement error models.* New York: John Wiley.

Hannan, E. J. and Rissanen. J. (1982). "Recursive estimation of mixed auto regressive moving average order," *Biometrika*, **69**, 81-94.

Kohn, R. (1979). "Asymptotic estimation and hypothesis testing results for vector linear time series models," *Econometrica*, **47**, 1005-1030.

Miazaki, E. S. and Dorea, C. C. Y. (1993). "Estimation for the parameters of a time series subject to the errors of rotational sampling," *Communications in Statistics: Theory and Methods*, **22**, 805-825.

Pagano, N. (1974). "Estimation of models of autoregressive signal plus white noise," *Annals of Statistics*, **2**, 99-108.

Reinsel. G. C., Basu, S. B. and Yap, S. F. (1992). "Maximum likelihood estimators in the multivariate autoregressive moving average model from a generalized least squares viewpoint," *Journal of Time Series Analysis*, **13**, 133-145.

Sakai, H. and Arase, M. (1979). "Recursive parameter estimation of an autoregressive process distributed by white noise," *International Journal of Control*, *30*, 949-966.

Said, S. E. and Dickey, D. A. (1985). "Hypothesis testing in ARIMA(p,1,q) models," *Journal of the American Statistical Association*, **80**, 369-374.

Schwert, G. W. (1989). "Tests for unit roots: A Monte Carlo investigation," *Journal of the American Statistical Association*, **83**, 147-159.

Shin, D. W. (1993). "Maximum likelihood estimation for autoregressive processes distributed by a moving average," *Journal of Time Series Analysis*, **14**, 629-643.

Shin, D. W. and Sarkar, S. (1993). "A note on testing for a unit root in an ARIMA(p,1,0) signal observed with MA(q) noise," *Statistics and Probability Letter*, **18**, 195-203.

Shin, D. W. and Sarkar, S. (1994). "Testing for a unit root in an ARIMA(p,1,0) signal observed with MA(q) noise," *Communications in Statistics: Theory and Methods*, **29**, 2643-2670.

Shin, D. W. and Sarkar, S. (1995). "Estimation of the multivariate autoregressive moving average having parameter restrictions and an application to rotational sampling," *Journal of Time Series Analysis*, **16**, 431-444.