June 27, 1985

Estimation for Autoregressive Processes with Several Unit Roots

Sastry G. Pantula
North Carolina State University
Raleigh, NC 27695-8203

ABSTRACT

Let Y_t satisfy the stochastic difference equation $Y_t = \sum_{j=1}^{p} \alpha_j Y_{t-j} + e_t$, for $t = 1, 2, \ldots$, where the e_t are independent and identically distributed random variables and the initial conditions $(Y_{-p+1}, Y_{-p+2}, \ldots, Y_0)$ are fixed constants. It is assumed that the true, but unknown, roots $m_1, m_2, \ldots, m_p$ of $m^p - \sum_{j=1}^{p} \alpha_j m^{p-j} = 0$ satisfy $m_1 = m_2 = \ldots = m_d = 1$ and $|m_j| < 1$ for $j = d+1, \ldots, p$. We consider a reparameterization of the model for Y_t that is convenient for testing the hypothesis H_d: $m_1 = m_2 = \ldots = m_d = 1$, and $|m_j| < 1$ for $j = d+1, \ldots, p$. We consider the likelihood ratio type "F-statistics" to test the hypothesis H_d. We characterize the asymptotic distributions of the "F-statistics" under various alternative hypotheses. Using these asymptotic results, we obtain a sequential testing criterion that is asymptotically consistent. Finally, estimated percentiles for the distributions of the test statistics are obtained for $d \leq 5$, by Monte Carlo methods.

1. Introduction

Let the time series $\{Y_t\}$ satisfy

$$Y_t = \sum_{j=1}^{p} \alpha_j Y_{t-j} + e_t, \quad t = 1, 2, \ldots, \tag{1.1}$$

where $\{e_t\}_{t=1}^{\infty}$ is a sequence of independent random variables with mean zero and variance σ^2. It is assumed that the e_t are either identically distributed or that $E\{|e_t|^{4+\delta}\} < M$ for some $\delta > 0$ and all t. It is further assumed that the initial conditions $(Y_{-p+1}, \ldots, Y_0)$ are known constants. For the statistics considered in this paper, without loss of generality we assume that $\sigma^2 = 1$ and $Y_{-p+1} = \ldots = Y_0 = 0$. The time series is said to be an autoregressive process of order p. Let

$$m^p - \sum_{j=1}^{p} \alpha_j m^{p-j} = 0 \tag{1.2}$$

be the characteristic equation of the process. The roots of (1.2), denoted by $m_1, m_2, \ldots, m_p$ are the characteristic roots of the process. Assume $|m_1| \geq |m_2| \geq \ldots \geq |m_p|$.

Let the observations $Y_1, Y_2, \ldots, Y_n$ be available. It is assumed that $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_p)'$ and σ^2 are unknown. The least squares estimator $\hat{\boldsymbol{\alpha}}$ of $\boldsymbol{\alpha}$ is obtained by regressing Y_t on $Y_{t-1}, \ldots, Y_{t-p}$. Lai and Wei (1983) established the strong consistency of $\hat{\boldsymbol{\alpha}}$ under a very general set of conditions. If $|m_j| < 1$, $j = 1, 2, \ldots, p$, then Y_t converges to a weakly stationary process as $t \to \infty$. In the stationary case, the asymptotic properties of $\hat{\boldsymbol{\alpha}}$ and of related "F-type" likelihood ratio test statistics are well known. See Mann and Wald (1943), Anderson (1959) and Hannan and Heyde (1972).

Dickey (1976), Dickey and Fuller (1979), and Fuller (1979) have considered testing for a single unit root in a p-th order autoregressive process (i.e., $m_1 = 1$ and $|m_j| < 1$ for $j = 2, 3, \ldots, p$). Hasza (1977) and Hasza and Fuller (1979) considered a p-th order autoregressive process with $m_1 = m_2 = 1$ and $|m_j| < 1$ for $j = 3, 4, \ldots, p$. Hasza and Fuller (1979) characterized the distribution of "F-type" statistic for testing $m_1 = m_2 = 1$. Dickey, Hasza and Fuller (1984) obtained similar results for testing the unit roots in seasonal time series. Sen (1985) considered the distribution of symmetric estimators and that of the estimated characteristic roots under the assumption $m_1 = m_2 = 1$. With the exception of Sen (1985), no one has considered the distribution of the test statistics under the alternative hypotheses. For a second order process, Sen (1985) studied the asymptotic distribution of the regression type t-statistic for testing the hypothesis H_1: $m_1 = 1$, $|m_2| < 1$, under the assumption H_2: $m_1 = m_2 = 1$. He observed that the probability of rejecting the hypothesis H_1, in favor of the hypothesis H_0: $|m_1| < 1$, $|m_2| < 1$, is higher under H_2 than under H_1. That is, it is more likely we conclude, incorrectly, that the process is stationary when there are really two unit roots present.

We propose "F-type" statistics for testing the hypothesis H_d: $m_1 = m_2 = \ldots = m_d = 1$ and $|m_j| < 1$, $j = d+1, \ldots, p$. We characterize the asymptotic distribution of the "F-type" statistics under the hypotheses H_d, $d = 0, 1, \ldots, p$. We use these asymptotic distributions to obtain an asymptotically consistent sequential procedure to test H_d versus H_{d-1}. The procedure is very general and it successfully answers the question "how many times should one difference the series to attain stationarity?"

In Section 2 we present the notation and the main results. In Section 3 we present the estimated percentiles of "F-type" statistics and some possible extensions. We present an example in Section 4. The proofs of the results are presented in the Appendix.

2. Notation and Main Results

We consider the following reparameterization of the model (1.1),

$$Y_{p,t} = \sum_{i=1}^{p} \beta_i Y_{i-1,t-1} + e_t \,, \tag{2.1}$$

where $Y_{i,t} = (1-B)^i Y_t$ = ith difference of Y_t; and B is the backshift operator. After expanding $(1-B)^i$ in powers of B, if we compare (2.1) with (1.1), we get

$$\alpha_i = \left[\sum_{\ell=i}^{p} \beta_\ell \binom{\ell-1}{i-1} + \binom{p}{i}\right] (-1)^{i-1} \tag{2.2}$$

for $i = 1, 2, \ldots, p$. Therefore, we can write

$$\boldsymbol{\alpha} = \mathbf{T}\boldsymbol{\beta} + \mathbf{c} \tag{2.3}$$

where T_{ij} is an upper triangular matrix with $T_{ij} = (-1)^{i-1}\binom{j-1}{i-1}$ for $j \geq i$ and $\mathbf{c} = (c_1, c_2, \ldots, c_p)'$ with $c_i = (-1)^{i-1}\binom{p}{i}$. Since the diagonal elements of $\mathbf{T}$ are ± 1, $\mathbf{T}$ is nonsingular and we can also obtain $\boldsymbol{\beta}$ from $\boldsymbol{\alpha}$.

It is easy to see that $\alpha_{r+1} = \alpha_{r+2} = \ldots = \alpha_p = 0$ if and only if $\beta_{r+1} = \beta_{r+2} = \ldots = \beta_p = -1$. So if one would like to test whether the order of the process is r instead of p one could use the F-test for testing $\beta_{r+1} = \beta_{r+2} = \ldots = \beta_p = -1$ in the regression (2.1). Our main purpose of the paper, however, is to test the hypothesis that there are exactly d unit roots. So, we assume that we "know" the order p of the process. (For all our results, all we need to assume is that the order of the process is less than or equal to p.)

Now, suppose $m_1 = m_2 = \ldots = m_d = 1$ and $|m_j| < 1$ for $j = d+1,\ldots,p$. Then we can write (1.1) as,

$$(1-B)^d \prod_{i=d+1}^{p} (1 - m_i B) Y_t = e_t \quad . \tag{2.4}$$

Define,

$$Z_t = (1-B)^d Y_t = Y_{d,t}$$

and

$$W_t = \prod_{i=d+1}^{p} (1 - m_i B) Y_t \quad . \tag{2.5}$$

Then, W_t is a dth order autoregressive process with d unit roots and Z_t is a (p-d)th order process with stationary characteristic roots. Note that,

$$\prod_{i=d+1}^{p} (1 - m_i B) Z_t = e_t$$

and we can write Z_t as

$$Z_t = \sum_{j=d+1}^{p} \theta_j Z_{t-j-d} + e_t \tag{2.6}$$

or as

$$(1-B)^{p-d} Z_t = \sum_{j=d+1}^{p} \eta_j (1-B)^{j-d-1} Z_{t-1} + e_t \quad . \tag{2.7}$$

Since Z_t is $(1-B)^d Y_t$ we get,

$$Y_{p,t} = \sum_{j=d+1}^{p} \eta_j Y_{j-1,t-1} + e_t \quad . \tag{2.8}$$

Comparing (2.8) with (2.1), we get $\beta_1 = \beta_2 = \dots = \beta_d = 0$ and $\beta_j = \eta_j$ for $j \geq d+1$. Therefore, $m_1 = \dots = m_d = 1$ if and only if $\beta_1 = \beta_2 = \dots = \beta_d = 0$. The assumption that $|m_j| < 1$ for $j = d + 1, \dots, p$ imposes extra conditions on η_j. In particular, $\eta_{d+1} = \beta_{d+1} < 0$.

Let $\hat{\boldsymbol{\beta}}$ denote the ordinary least squares estimator of $\boldsymbol{\beta}$ obtained by regressing $Y_{p,t}$ on $Y_{0,t-1}, Y_{1,t-1}, \dots, Y_{p-1,t-1}$. That is,

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}'\mathbf{Y}_{(p)} \tag{2.9}$$

where

$$\boldsymbol{\Phi}_t = (Y_{0,t-1}, \dots, Y_{p-1,t-1})' ,$$

$$\boldsymbol{\Phi} = (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \dots, \boldsymbol{\Phi}_n)' ,$$

and

$$\mathbf{Y}_{(p)} = (Y_{p,1}, Y_{p,2}, \dots, Y_{p,n})' .$$

Define the "F-statistic" for testing $\beta_1 = \beta_2 = \dots = \beta_i = 0$ by

$$F_{i,n}(p) = \frac{\hat{\boldsymbol{\beta}}'_{(i)} \mathbf{C}^{-1}_{(i)} \hat{\boldsymbol{\beta}}_{(i)}}{i \; (\mathrm{MSE})_n} \tag{2.10}$$

where

$\mathbf{C}_{(i)}$ = (ixi) submatrix consisting of the first i rows and i columns of $(\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1}$,

$\hat{\boldsymbol{\beta}}_{(i)} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_i)'$

and

$$\begin{aligned} \mathrm{MSE}_n &= \text{mean square error of the regression} \\ &= (n-p)^{-1} \sum_{t=1}^{n} [Y_{p,t} - \hat{\boldsymbol{\beta}}'\boldsymbol{\Phi}_t'\mathbf{Y}_{(p)}]^2 . \end{aligned}$$

We use H_d to denote the hypothesis

$$H_d\colon m_1 = m_2 = \dots = m_d = 1 \text{ and } |m_j| < 1, \text{ for } j = d+1, d+2, \dots, p.$$

We now present the asymptotic properties of $F_{i,n}(p)$ under H_d.

<u>Theorem 1</u>: Assume that Y_t is defined by (1.1). Then under the hypothesis H_d,

$$F_{i,n}(p) \xrightarrow{\mathcal{D}} \begin{cases} F_i(d) & \text{if } i \leq d \\ \infty & \text{if } i > d \end{cases} \tag{2.11}$$

where

$$F_i(d) = \frac{1}{i} \mathbf{g}^{*\prime}_{(i),d} [\mathbf{G}_d^{(ii)}]^{-1} \mathbf{g}^*_{(i),d}$$

and the elements of the vector $\mathbf{g}^*_{(i),d}$ and the matrix $\mathbf{G}_d^{(ii)}$ are functions of $\mathbf{g}_d$ and $\mathbf{G}_d$ defined in (A.9). The elements of $\mathbf{g}_d$ and $\mathbf{G}_d$ are functions of weighted chi-squares and normal variables. (See Lemma 4 of the Appendix.)

Notice that the limiting distribution of $F_{i,n}(p)$ is independent of the order p. Let us denote $F_d(d,\alpha)$ to be the 100(1-α)% percentile of the distribution of $F_d(d)$. Since, under H_d, $F_{i,n}(d)$ convergence in probability to infinity for $i > d$, we suggest the following criterion:

> Reject H_d in favor of H_{d-1} , if $F_{i,n}(p) > F_i(i,\alpha)$ for $i = d, d+1,\ldots,p$.

Note that,

$$\begin{aligned}\lim_{n\to\infty} P_{H_d}[\text{Rejecting } H_d] &= \lim_{n\to\infty} P_{H_d}[F_{d,n}(p) > F_d(d,\alpha),\ F_{d+1,n}(p) > F_{d+1}(d+1,\alpha), \\ &\qquad \ldots,\ F_{p,n}(p) > F_p(p,\alpha)] \\ &= \lim_{n\to\infty} P_{H_d}[F_{d,n}(p) > F_d(d,\alpha)] \\ &= \alpha\ .\end{aligned}$$

Also, for $j > d$

$$\lim_{n\to\infty} P_{H_j}[\text{Rejecting } H_d] \leq \alpha$$

and for $j < d$

$$\lim_{n\to\infty} P_{H_j}[\text{Rejecting } H_d] = 1.$$

That is, if we use our sequential procedure, asymptotically the chance that we reject the hypothesis that there are exactly d unit roots in favor of the hypothesis that there are exactly d-1 unit roots is (i) α, if there are exactly d unit roots, (ii) less than or equal to α, if there are more than d unit roots present and (iii) one, if there are less than d unit roots. In this sense, our procedure is asymptotically consistent.

Consider, for example, a second order process. Suppose we wish to test the hypothesis that there is one unit root. A common practice is to use the t-statistic ("equivalently" $F_{1,n}(2)$) for testing $\beta_1 = 0$. Dickey and Fuller (1979) presented the empirical percentiles for the distribution of the t-statistic, under the hypothesis H_1: $m_1 = 1$ and $|m_2| < 1$. Note that rejecting H_1 may mean that $|m_1| < 1$, $|m_2| < 1$ or $m_1 = m_2 = 1$. (We are excluding the possibility of $|m_i| > 1$ or $m_i = -1$.) However, in practice, the rejection of H_1 is misconstrued as the former. This is because "intuitively" we hope that if there are two unit roots then the test for one unit root will indicate that there is a unit root. Sen (1985), using simulation methods, has shown that, a level α test based on the t-statistic rejected H_1 more than $100\alpha\%$ of the time when there are in fact two unit roots. Our criterion takes into account that exactly d of the roots are equal to one and the remaining are less than one in modulus and tests the hypotheses sequentially. However, if one is "sure" that there are at most s unit roots (e.g., say s = 5) then one may modify the criterion for testing $H_d(d \leq s)$ to "reject H_d in favor of H_{d-1} if $F_{i,n}(p) > F_i(i,\alpha)$ for i = d, d+1,...,s."

In practical applications it is common to include an intercept term in the model (2.1), i.e., consider

$$Y_{p,t} = \beta_0 + \sum_{i=1}^{p} \beta_i Y_{i-1,t-1} + e_t . \tag{2.12}$$

Let $\hat{\beta}^{(+)}$ denote the ordinary least squares estimator of $\beta^{(+)} = (\beta_0, \beta_1, \ldots, \beta_p)'$

obtained by regressing $Y_{p,t}$ on 1, $Y_{0,t-1}$, $Y_{1,t-1}$, ..., $Y_{p-1,t-1}$. Denote the regression "F-statistic" for testing $\beta_1 = \beta_2 = \ldots = \beta_i = 0$ in the regression (2.12) by $F^{(1)}_{i,n}(p)$. Similarly, let $F^{(2)}_{i,n}(p)$ denote the "F-statistic" for testing $\beta_0 = \beta_1 = \ldots = \beta_i = 0$. We now present the asymptotic properties of $F^{(1)}_{i,n}(p)$ and $F^{(2)}_{i,n}(p)$ under the hypothesis H_d.

Theorem 2: Assume that Y_t is defined by (1.1). Then, under the hypothesis H_d,

$$F^{(j)}_{i,n}(p) \longrightarrow \begin{cases} F^{(j)}_i(d) & \text{if } i \leqq d \\ \infty & \text{if } i > d \end{cases}$$

for j = 1, 2, where

$$F^{(1)}_i(d) = \frac{1}{i}\, \mathbf{h}^{*'}_{(i),d}\, [\mathbf{H}^{(ii)}_{d,1}]^{-1}\, \mathbf{h}^{*}_{(i),d}$$

$$F^{(2)}_i(d) = \frac{1}{(i+1)}\, \mathbf{h}^{**'}_{(i),d}\, [\mathbf{H}^{(ii)}_{d,2}]^{-1}\, \mathbf{h}^{**}_{(i),d}$$

and the elements of the vectors $\mathbf{h}^{*}_{(i),d}$ and $\mathbf{h}^{**}_{(i),d}$ and those of the matrices $\mathbf{H}^{(ii)}_{d,1}$ and $\mathbf{H}^{(ii)}_{d,2}$ are functions of the vector $\mathbf{h}_d$ and the matrix $\mathbf{H}^{*}_d$ defined in (A.11). The elements of $\mathbf{h}_d$ and $\mathbf{H}^{*}_d$ are functions of weighted chi-square and normal variables. (See the proof of Theorem 2 in the Appendix.)

Therefore, an alternative criterion is to reject H_d in favor of H_{d-1} if $F^{(1)}_{i,n}(p) > F^{(1)}_i(i,\alpha)$ for i = d, d+1, ..., p. One may use $F^{(2)}_{i,n}(p)$ in place of $F^{(1)}_{i,n}(p)$ in the above criterion. However, based on the power studies of Hasza and Fuller (1979) and Dickey and Fuller (1981) we recommend the use of $F^{(1)}_{i,n}(p)$ over $F^{(2)}_{i,n}(p)$. From the results of Hasza (1977), it follows that if Y_t satisfy (2.12) with $\beta_0 \neq 0$, then under the hypothesis H_d, $F^{(1)}_{d,n}(p)$ converges in distribution to $d^{-1}\chi^2_d$, where χ^2_d denotes a chi-square random variable with d degrees of freedom.

3. Simulation for d=5

We first present the elements of $\mathbf{G}_5$ and $\mathbf{g}_5$ which will be used in obtaining the empirical percentiles of the random variables $F_i(d)$, $F^{(1)}_i(d)$ and $F^{(2)}_i(d)$.

From (A.9) and the results of Lemma 1, we get

$$G_5 = \begin{bmatrix} Q_5 & & & & \\ \tfrac{1}{2}N_5^2 & Q_4 & & & \text{Symmetric} \\ N_4N_5-Q_4 & \tfrac{1}{2}N_4^2 & Q_3 & & \\ N_3N_5-\tfrac{1}{2}N_4^2 & N_3N_4-Q_3 & \tfrac{1}{2}N_3^2 & Q_2 & \\ N_2N_5-N_3N_4+Q_3 & N_2N_4-\tfrac{1}{2}N_3^2 & N_2N_3-Q_2 & \tfrac{1}{2}N_2^2 & Q_1 \end{bmatrix},$$

and

$$g_5 = \begin{bmatrix} N_1N_5-N_2N_4+\tfrac{1}{2}N_3^2 \\ N_1N_4-N_2N_3+Q_2 \\ N_1N_3-\tfrac{1}{2}N_2^2 \\ N_1N_2-Q_1 \\ \tfrac{1}{2}(N_1^2-1) \end{bmatrix},$$

where

$$Q_1 = Q_1^* ,$$

$$Q_2 = Q_2^* - N_2^2 + 2N_2N_3,$$

$$Q_3 = Q_3^* - 1/3\ N_2^2 + 2N_2N_1^* ,$$

$$Q_4 = Q_4^* - 2/15\ N_2^2 - N_4^2 + 2/3\ N_2N_4 - N_2(N_5 - N_2^*) + 2N_4N_5,$$

$$Q_5 = Q_5^* - 17/315\ N_2^2 - 1/3\ N_4^2 + 4/15\ N_2\ N_4 + \tfrac{1}{2}(2N_4 - N_2)(N_5 - N_2^*) + 1/3\ N_2N_3^* ,$$

$$N_1 = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} \gamma_i V_i ; \quad N_2 = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} \gamma_i^2\ V_i ;$$

$$N_3 = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} [\gamma_i^2 - \gamma_i^3] V_i ; \quad N_4 = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} [\tfrac{1}{2}\gamma_i^2 - \gamma_i^4] V_i ;$$

$$N_5 = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} [1/6\ \gamma_i^2 - \gamma_i^4 + \gamma_i^5] V_i ; N_1^* = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} [1/3\ \gamma_i^2 - \gamma_i^5] V_i ;$$

$$N_2^* = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} [1/10\ \gamma_i^2 - 1/3\ \gamma_i^4 + \gamma_i^5 - 2\gamma_i^7] V_i ;$$

$$N_3^* = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} [1/42\ \gamma_i^2 - 1/5\ \gamma_i^4 + 3\gamma_i^7 - 6\gamma_i^9] V_i ;$$

$$Q_j^* = \sum_{i=1}^{\infty} \gamma_i^{2j}\ V_i^2 , \quad j \geq 1$$

and

$$\gamma_i = 2[(2i-1)\pi]^{-1}(-1)^{i+1} .$$

Also,

$$\mathbf{H}_5^* = \begin{bmatrix} 1 & \boldsymbol{\rho}_5' \\ \boldsymbol{\rho}_5 & \mathbf{G}_5 \end{bmatrix}$$

and

$$\mathbf{h}_5 = \begin{bmatrix} N_1 \\ \mathbf{g}_5 \end{bmatrix}$$

where

$$\boldsymbol{\rho}_5' = (N_6, N_5, N_4, N_3, N_2) ,$$

and

$$N_6 = 2^{\frac{1}{2}} \sum_{i=1}^{\infty} [1/24 \gamma_i^2 - 0.5 \gamma_i^4 + \gamma_i^6] V_i .$$

All of the F and the $F^{(j)}$ statistics can be computed using the elements of $\mathbf{H}_5^*$ and $\mathbf{h}_5$.

Estimates of the percentiles of the F-statistics are presented in Tables 3.1, 3.2 and 3.3. For the finite sample sizes the F-statistics were computed for samples generated using the model (1.1), with $e_t \sim NID(0,1)$ and $Y_0 = Y_{-1} = \ldots = Y_{-4} = 0$. To estimate the percentiles of the limiting distributions of the test statistics we use the method of simulation employed by Dickey (1976) and Hasza (1977). Briefly, the method consists of approximating the sequence $\{\gamma_i\}_{i=1}^{\infty}$ by a finite sequence. Using the Monte Carlo methods sample distribution functions of the statistics are generated by appropriately truncating the infinite series in the definitions of N_j's, N_j^*'s and Q_j^*'s. The percentiles were "smoothed" by fitting a regression function to the original set of estimates.

The estimated standard errors of the estimated percentiles are generally less than 0.90 percent of the table entry for the limiting distributions and are generally less than 1.50 percent of the table entry for the finite sample cases. Fifty thousand indpendent sample statistics were used in constructing the percentiles for the asymptotic distribution.

In our discussion, we assumed that the initial conditions are fixed. The limiting distribution does not depend on the initial condition $(Y_{-p+1},\ldots,Y_0)$, but these values will influence the small sample distribution. Also, the criterion we proposed in Section 2 for testing H_d involved $F_d(d,\alpha)$ rather than $F_{d,n}(d,\alpha)$ where $F_{d,n}(d,\alpha)$ denotes the $100(1-\alpha)\%$ empirical percentile of the statistic $F_{d,n}(d)$. For small samples, one should use $F_{d,n}(d,\alpha)$ instead of $F_d(d,\alpha)$. Therefore, our criterion is to reject H_d in favor of H_{d-1} if $F_{i,n}(p) > F_{i,n}(i,\alpha)$ for $i = d, d+1,\ldots,p$. The asymptotic results of Section 2 are still valid even though we replace $F_i(i,\alpha)$ by $F_{i,n}(i,\alpha)$.

Consider, for example, a second order autoregressive process $\{Y_t\}$. Suppose $n = 50$. We regress $Y_{2,t} = (1-B)^2 Y_t$ on $Y_{0,t-1} = Y_{t-1}$ and $Y_{1,t-1} = (1-B)Y_{t-1}$, to obtain $\hat{\beta}_1$ and $\hat{\beta}_2$. Compute the regression F-statistics $F_{1,n}(2)$ and $F_{2,n}(2)$ for testing the hypotheses, (i) $\beta_1 = 0$ and (ii) $\beta_1 = \beta_2 = 0$, respectively. Let $\alpha = 0.10$. Hasza and Fuller (1979) suggest that one should reject H_2: two unit roots (in favor of H_1: exactly one unit root) if $F_{2,n}(2) > 2.82$. Also, a common practice is to reject H_1 in favor of stationarity if $F_{1,n}(2) > 3.01$. Our criterion differs from above in the sense we reject H_1 in favor of H_0: no unit roots, only if $F_{1,n}(2) > 3.01$ <u>and</u> $F_{2,n}(2) > 2.82$. If we do not add the condition $F_{2,n}(2) > 2.82$, then with probability greather than 0.10 we will be concluding that the process is stationary when in fact it has two unit roots. However, if we are absolutely sure that there is at most one unit root then we may use the criterion of rejecting H_1 in favor of H_0 if $F_{1,n}(2) > 3.01$.

TABLE 3.1

Empirical Percentiles for the F-Statistics

	n	0.50	0.80	0.90	0.95	0.975	0.99
$F_{1,n}(1)$	25	0.58	1.89	3.04	4.34	5.74	7.80
	50	0.59	1.89	3.01	4.23	5.54	7.38
	100	0.60	1.89	2.99	4.18	5.42	7.16
	250	0.60	1.89	2.98	4.15	5.35	7.02
	500	0.60	1.89	2.97	4.14	5.32	6.97
	∞	0.61	1.88	2.96	4.13	5.28	6.91
$F_{2,n}(2)$	25	0.95	2.04	2.88	3.76	4.71	5.98
	50	0.97	2.02	2.82	3.62	4.45	5.59
	100	0.98	2.02	2.79	3.55	4.32	5.38
	250	0.98	2.01	2.77	3.50	4.24	5.23
	500	0.98	2.01	2.76	3.49	4.22	5.17
	∞	0.99	2.01	2.75	3.47	4.19	5.10
$F_{3,n}(3)$	25	1.15	2.22	2.97	3.73	4.49	5.57
	50	1.18	2.20	2.88	3.55	4.21	5.11
	100	1.19	2.19	2.83	3.46	4.07	4.88
	250	1.20	2.18	2.81	3.41	3.99	4.75
	500	1.20	2.18	2.80	3.39	3.96	4.70
	∞	1.20	2.17	2.80	3.39	3.94	4.66
$F_{4,n}(4)$	25	1.29	2.35	3.07	3.80	4.56	5.60
	50	1.32	2.31	2.95	3.56	4.17	4.97
	100	1.34	2.29	2.89	3.45	3.99	4.67
	250	1.35	2.28	2.86	3.39	3.88	4.51
	500	1.35	2.28	2.85	3.37	3.85	4.46
	∞	1.35	2.28	2.84	3.35	3.84	4.46
$F_{5,n}(5)$	25	1.37	2.44	3.17	3.90	4.64	5.68
	50	1.41	2.38	3.02	3.60	4.16	4.93
	100	1.43	2.36	2.94	3.46	3.95	4.58
	250	1.44	2.34	2.90	3.38	3.84	4.40
	500	1.45	2.34	2.88	3.36	3.81	4.36
	∞	1.45	2.34	2.87	3.36	3.83	4.38

Table 3.2 Empirical Percentiles of the $F^{(1)}$-Statistics

Statistic	n	0.50	0.80	0.90	0.95	0.975	0.99
$F_{1,n}^{(1)}(1)$	25	2.36	4.99	6.95	8.96	10.98	13.84
	50	2.41	4.94	6.74	8.54	10.36	12.76
	100	2.43	4.91	6.65	8.35	10.04	12.24
	250	2.44	4.91	6.60	8.24	9.84	11.93
	500	2.45	4.91	6.58	8.24	9.78	11.83
	∞	2.45	4.91	6.58	8.21	9.69	11.76
$F_{2,n}^{(1)}(2)$	25	2.55	4.43	5.79	7.15	8.56	10.51
	50	2.56	4.30	5.49	6.60	7.69	9.14
	100	2.57	4.24	5.34	6.35	7.33	8.59
	250	2.58	4.21	5.25	6.22	7.14	8.33
	500	2.58	4.20	5.23	6.18	7.09	8.27
	∞	2.59	4.19	5.20	6.15	7.06	8.23
$F_{3,n}^{(1)}(3)$	25	2.68	4.39	5.56	6.78	8.00	9.68
	50	2.67	4.19	5.16	6.11	7.03	8.23
	100	2.67	4.08	4.96	5.78	6.56	7.54
	250	2.67	4.02	4.85	5.60	6.30	7.17
	500	2.67	4.01	4.81	5.55	6.22	7.07
	∞	2.67	3.99	4.79	5.52	6.19	7.06
$F_{4,n}^{(1)}(4)$	25	2.80	4.51	5.67	6.83	8.05	9.76
	50	2.76	4.20	5.11	5.96	6.74	7.83
	100	2.74	4.05	4.84	5.55	6.20	7.06
	250	2.73	3.97	4.69	5.33	5.95	7.70
	500	2.73	3.94	4.65	5.27	5.88	6.61
	∞	2.72	3.93	4.63	5.26	5.84	6.55
$F_{5,n}^{(1)}(5)$	25	2.40	3.87	4.87	5.90	6.91	8.39
	50	2.34	3.51	4.23	4.92	5.60	6.46
	100	2.32	3.37	3.98	4.54	5.08	5.73
	250	2.31	3.29	3.86	4.36	4.83	5.41
	500	2.30	3.28	3.84	4.32	4.77	5.34
	∞	2.30	3.26	3.82	4.29	4.73	5.29

Table 3.3 Empirical Percentiles of the $F^{(2)}$-Statistics

Statistic	n	0.50	0.80	0.90	0.95	0.975	0.99
$F^{(2)}_{1,n}(1)$	25	1.71	3.09	4.12	5.16	6.29	7.77
	50	1.72	3.00	3.94	4.87	5.81	7.02
	100	1.72	2.96	3.85	4.72	5.57	6.66
	250	1.72	2.94	3.80	4.64	5.44	6.46
	500	1.72	2.94	3.79	4.61	5.39	6.40
	∞	1.72	2.94	3.78	4.58	5.36	6.37
$F^{(2)}_{2,n}(2)$	25	2.05	3.34	4.26	5.20	6.18	7.56
	50	2.04	3.20	3.99	4.75	5.50	6.48
	100	2.03	3.13	3.86	4.54	5.20	6.06
	250	2.03	3.10	3.79	4.43	5.06	5.86
	500	2.03	3.09	3.76	4.40	5.02	5.82
	∞	2.03	3.08	3.75	4.38	4.99	5.78
$F^{(2)}_{3,n}(3)$	25	2.30	3.61	4.52	5.46	6.43	7.73
	50	2.26	3.40	4.14	4.86	5.57	6.50
	100	2.25	3.30	3.96	4.58	5.17	5.92
	250	2.24	3.24	3.86	4.42	4.94	5.61
	500	2.23	3.22	3.82	4.37	4.88	5.52
	∞	2.23	3.21	3.80	4.36	4.86	5.51
$F^{(2)}_{4,n}(4)$	25	2.49	3.89	4.86	5.80	6.81	8.26
	50	2.43	3.59	4.32	5.02	5.65	6.55
	100	2.40	3.44	4.07	4.65	5.18	5.87
	250	2.38	3.36	3.93	4.44	4.95	5.55
	500	2.37	3.33	3.89	4.39	4.89	5.47
	∞	2.37	3.31	3.87	4.38	4.85	5.41
$F^{(2)}_{5,n}(5)$	25	2.63	4.14	5.18	6.23	7.34	8.87
	50	2.53	3.72	4.44	5.14	5.83	6.71
	100	2.49	3.54	4.15	4.71	4.25	5.91
	250	2.47	3.45	4.02	4.52	4.99	5.57
	500	2.47	3.43	3.99	4.48	4.93	5.49
	∞	2.46	3.41	3.97	4.44	4.89	5.44

In addition, we also computed $F_{i,n}(d,\alpha)$ for $1 \leq i < d \leq 5$. (Tables for $F_{i,n}(d,\alpha)$ are not included.) It is interesting to note that the empirical percentiles $F_{i,n}(d,\alpha)$ increase in d for a fixed i and $\alpha \leq 0.5$. For example, $F_{1,100}(1, 0.05) = 4.18$, $F_{1,100}(2, 0.05) = 5.04$, $F_{1,100}(3, 0.05) = 6.20$, $F_{1,100}(4, 0.05) = 6.77$ and $F_{1,100}(5, 0.05) = 7.14$. This indicates that if one uses the criterion, "reject the hypothesis that there is exactly one unit root in favor of stationarity if $F_{1,n}(p) > F_{1,n}(1,\alpha)$," then, with probability greater than α, one would conclude that the process is stationary when really there are several unit roots.

Except for some tedious bookkeeping, the results can be extended to regressions involving time trend in the model.

4. Example

Pankratz (1983) lists 70 consecutive monthly volume of commercial bank real estate loans, in billions of dollars, starting from January 1973. By fitting a sixth order autoregressive process we concluded that a third order autoregressive process provides an adequate fit for the data. Regressing $Y_{3,t} = (1-B)^3 Y_t = Y_t - 3Y_{t-1} + 3Y_{t-2} - Y_{t-3}$ on Y_{t-1}, $Y_{1,t-1} = (1-B)Y_{t-1} = Y_{t-1} - Y_{t-2}$ and $Y_{2,t-1} = (1-B)^2 Y_{t-1} = Y_{t-1} - 2Y_{t-2} + Y_{t-3}$, we get

$$\underset{}{Y_{3,t}} = \underset{(0.00094)}{0.00139Y_{t-1}} - \underset{(0.0795)}{0.1045Y_{1,t-1}} - \underset{(0.1233)}{1.3061Y_{2,t-1}} + e_t \qquad (4.1)$$

with $\hat{\sigma}^2 = 0.0839$. The numbers in the parentheses are estimated standard errors calculated by the usual regression formulas used in most computer regression routines.

From (4.1) we calculated the F-statistics for testing the relevant hypotheses. The calculated values are $F_{1,70}(3) = 2.19$, $F_{2,70}(3) = 1.19$ and $F_{3,70}(3) = 47.29$, where $F_{i,70}(3)$ is the regression F-statistic for testing the hypothesis that the first i coefficients of (4.1) are zero. Since 47.29 is greater than any of the table values for $F_{3,50}(3)$ and $F_{3,100}(3)$, we reject the null hypothesis that there are three unit roots.

Note that, under the hypothesis H_2: there are two unit roots and the third root is less than one in absolute value, the statistics $F_{2,n}(3)$ and $F_{2,n}(2)$ are asymptotically equivalent. Comparing the calculated value of $F_{2,70}(3) = 1.19$ with $F_{2,50}(2, 0.20) = 2.02$ and $F_{2,100}(2, 0.20) = 2.02$, we fail to reject the hypothesis H_2 at 20% level of significance. Finally, we fitted a first order autoregressive model for $Y_{2,t} = (1-B)^2 Y_t$ to obtain

$$(1 + \underset{(0.12)}{0.36}B)(1-B)^2 Y_t = e_t ,$$

with $\hat{\sigma}^2 = 0.0842$. Same conclusions were obtained when an intercept was included in the model (4.1).

ACKNOWLEDGEMENTS

I wish to express my thanks to Professor D. Boos, D. Dickey and W. Fuller for their helpful comments. I am very grateful to Dr. Dickey and Dr. Sen for letting me use their Monte Carlo programs.

APPENDIX

We prove Theorem 1 in several steps. First we consider a dth order autoregressive process with d unit roots. For this process we obtain the limiting distributions of $F_{i,n}(d)$, i = 1, 2, ..., d. Then extend the results to a pth order process with d unit roots. Define,

$$S_{i,t} = S_{i,t-1} + S_{i-1,t} \qquad (A.1)$$

$$= \sum_{j=1}^{t} S_{i-1,j} ,$$

$$S_{0,t} = e_t ,$$

for i = 1, 2, ..., d and t = 1, 2, ..., where e_t is as defined in Section 1. Note that,

$$(1-B)^j S_{i,t} = S_{i-j,t}$$

for $j \leq i$ and $S_{i,t}$ is an ith order autoregressive process with i unit roots.

Consider $W_t = S_{d,t}$, a dth order autoregressive process with d unit roots. In the parameterization of (2.1), we can write W_t as

$$W_{d,t} = \sum_{i=1}^{d} \beta_i W_{i-1,t-1} + e_t \qquad (A.2)$$

where

$$W_{i,t} = (1-B)^i W_t = S_{d-i,t} ,$$

and $\beta_1 = \beta_2 = \ldots = \beta_d = 0$. Let $\tilde{\boldsymbol{\beta}}$ denote the ordinary least squares estimator of $\boldsymbol{\beta}$ in the regression of W_{dt} on $W_{0,t-1}$, $W_{1,t-1}$, ..., $W_{d-1,t-1}$. That is,

$$\tilde{\boldsymbol{\beta}} = \mathbf{B}_{d,n}^{-1} \mathbf{b}_{d,n} \qquad (A.3)$$

where the (ij)th element of $\mathbf{B}_{d,n}$ is

$$\sum_{t=1}^{n-1} W_{i-1,t} W_{j-1,t} = \sum_{t=1}^{n-1} S_{d-i+1,t} S_{d-j+1,t}$$

and the ith coordinate of $\mathbf{b}_{d,n}$ is

$$\sum_{t=1}^{n-1} W_{i-1,t}\, W_{d,t+1} = \sum_{t=1}^{n-1} S_{d-i+1,t}\, S_{0,t+1} .$$

We now obtain the orders in probability of the elements of $\mathbf{B}_{d,n}$ and $\mathbf{b}_{d,n}$. We also obtain recursive relationships between the elements.

Lemma 1: Let $S_{i,t}$ be as defined in (A.1). Then,

(i) $S_{i,t} = [(i-1)!]^{-1} \sum_{r=1}^{t} (t-r+1)(t-r+2)\ldots(t-r+i-1)e_r$,

(ii) $n^{-(i-\frac{1}{2})} S_{i,n} \xrightarrow{\mathcal{D}} N_i \sim N(0,\sigma_i^2)$, $\sigma_i^2 = [(2i-1)\{(i-1)!\}^2]^{-1}$,

(iii) $\sum_{t=1}^{n-1} S_{i,t}\, S_{j,t} = 0_p(n^{i+j})$ for i, j $\geq$ 1,

(iv) $\sum_{t=1}^{n-1} S_{i,t}\, S_{0,t+1} = 0_p(n^{i})$, $i \geq 1$,

(v) $\sum_{t=1}^{n-1} S_{i,t}\, S_{i-1,t} = \frac{1}{2} S_{i,n-1}^2 + 0_p(n^{2i-2})$, $i \geq 2$,

(vi) $\sum_{t=1}^{n-1} S_{i,t}\, S_{i-k,t} = S_{i,n-1}\, S_{i-k+1,n-1} - \sum_{t=1}^{n-1} S_{i-1,t}\, S_{i-k+1,t} + 0_p(n^{2i-k-1})$,
for $2 \leq k \leq i$.

Proof: The results (i) and (ii) are straightforward. Now,

$$n^{-2i} E[\sum_{t=1}^{n} S_{i,t}^2] = [(i-1)!]^{-2} \sum_{t=1}^{n} \sum_{k=1}^{t} \frac{1}{n} \cdot \frac{1}{n} \cdot [\frac{k}{n}]^{2(i-1)}$$

$$\to [(i-1)!]^{-2} \cdot [2i(2i-1)]^{-1} ,$$

as $n \to \infty$. Therefore,

$$\sum_{t=1}^{n} S_{i,t}^2 = 0_p(n^{2i}),$$

and hence $\sum_{t=1}^{n} S_{i,t}\, S_{j,t} = 0_p(n^{i+j})$.

Also,

$$V[\sum_{t=1}^{n} S_{it}\, S_{0,t+1}] = \sum_{t=1}^{n} E(S_{it}^2) = 0(n^{2i})$$

and we get (iv). Now, for $i \geq 2$,

$$\sum_{t=1}^{n-1} S_{i,t}^2 = \sum_{t=1}^{n-1} S_{i,t-1}^2 + \sum_{t=1}^{n-1} S_{i-1,t}^2 + 2 \sum_{t=1}^{n-1} S_{i,t-1} S_{i-1,t}$$

and hence

$$\sum_{t=1}^{n-1} S_{i,t-1}S_{i-1,t} = \tfrac{1}{2} S_{i,n-1}^2 - \tfrac{1}{2} \sum_{t=1}^{n-1} S_{i-1,t}^2 .$$

Therefore,

$$\sum_{t=1}^{n-1} S_{i,t} S_{i-1,t} = \sum_{t=1}^{n-1} S_{i,t-1}S_{i-1,t} + \sum_{t=1}^{n-1} S_{i-1,t}^2$$

$$= \tfrac{1}{2} S_{i,n-1}^2 + 0_p(n^{2i-2}) .$$

Note that, for $2 \leq k \leq i$,

$$\sum_{t=1}^{n-1} S_{i,t}S_{i-k,t} = \sum_{t=1}^{n-1} S_{i,t-1} S_{i-k,t} + 0_p(n^{2i-k-1}) .$$

Now,

$$\sum_{t=1}^{n-1} S_{i,t-1}S_{i-k+1,t} = \sum_{t=1}^{n-1} S_{i,t-1}S_{i-k+1,t-1} + \sum_{t=1}^{n-1} S_{i,t-1}S_{i-k,t}$$

and hence

$$\sum_{t=1}^{n-1} S_{i,t-1} S_{i-k+1,t-1} = \sum_{t=1}^{n-1} S_{i,t-1}S_{i-k+1,t} - \sum_{t=1}^{n-1} S_{i,t-1} S_{i-k,t} .$$

Also,

$$\sum_{t=1}^{n-1} S_{i,t} S_{i-k+1,t} = \sum_{t=1}^{n-1} S_{i,t-1} S_{i-k+1,t} + \sum_{t=1}^{n-1} S_{i-1,t}S_{i-k+1,t}$$

and hence

$$S_{i,n-1} S_{i-k+1,n-1} = \sum_{t=1}^{n-1} S_{i-1,t} S_{i-k+1,t} + \sum_{t=1}^{n-1} S_{i,t-1} S_{i-k,t} .$$

Therefore,

$$\sum_{t=1}^{n-1} S_{i,t-1} S_{i-k,t} = S_{i,n-1} S_{i-k+1,n-1} - \sum_{t=1}^{n-1} S_{i-1,t} S_{i-k+1,t} . \qquad \square$$

From Lemma 1, it follows that, for $0 \leq j \leq i$ and $i \geq 2$,

$$\sum_{t=1}^{n-1} S_{i,t-1}S_{j,t-1} = \sum_{r=0}^{k-1}(-1)^r S_{i-r,n-1}S_{j+r+1,n-1} + (-1)^k \sum_{t=1}^{n-1} S_{i-k,t-1}^2$$

$$+ 0_p(n^{i+j-1}) , \qquad \text{if } i-j \text{ is even}$$

$$= \sum_{r=0}^{k-1}(-1)^r S_{i-r,n-1} S_{j+r+1,n-1} + (-1)^k \tfrac{1}{2} S^2_{i-k,n-1}$$

$$+ 0_p(n^{i+j-1}), \text{if } i-j \text{ is odd},$$

where k = the integer part of $\frac{1}{2}(i-j)$.

We now obtain the order in probability results for $\mathbf{B}^{-1}_{d,n}$.

Lemma 2: Let $\mathbf{D}_{d,n} = \text{diag}\{n^d, n^{d-1}, \ldots, n\}$. Define,

$$\mathbf{G}_{d,n} = \mathbf{D}^{-1}_{d,n} \mathbf{B}_{d,n} \mathbf{D}^{-1}_{d,n}$$

and

$$\mathbf{g}_{d,n} = \mathbf{D}^{-1}_{d,n} \mathbf{b}_{d,n} .$$

Then,

$$\mathbf{G}_{dn} = 0_p(1)$$

$$\mathbf{g}_{dn} = 0_p(1)$$

and the determinant of $\mathbf{G}_{dn}$ is bounded away from 0 in probability.

Proof: From Lemma 1, it is clear that $\mathbf{G}_{dn} = 0_p(1)$ and $\mathbf{g}_{dn} = 0_p(1)$. We prove the result on determinants through induction on d. For d = 1,

$$\mathbf{G}_{1,n} = n^{-2}\sum_{t=1}^{n} S^2_{1,t}$$

$$\geq n^{-3} S^2_{2,n}$$

and since $n^{-3}S^2_{2,n}$ convergence in distribution to a constant times a chi-square random variable with one degree of freedom, we get $G_{1,n}$ is bounded away from zero.

Assume now that the det $[\mathbf{G}^{-1}_{d-1,n}]$ is bounded in probability. Note that $\lambda_{max}(\mathbf{G}_{d-1,n}) \leq \text{trace}\,[\mathbf{G}_{d-1,n}] = 0_p(1)$. Also, since we assumed that det $[\mathbf{G}^{-1}_{d-1,n}]$ is bounded in probability we get $\lambda^{-1}_{min}(\mathbf{G}_{d-1,n})$ is bounded in probability. We now show that det $[\mathbf{G}^{-1}_{d,n}]$ is $0_p(1)$. Note that

$$\det[\mathbf{G}_{d,n}] = \det[\mathbf{G}_{d-1,n}]\, n^{-2d}\Big[\sum_{t=1}^{n-1} S^2_{dt} - \mathbf{q}'_n \mathbf{G}^{-1}_{d-1,n}\mathbf{q}_n\Big]$$

where $\mathbf{q}_n = (n^{-(d-1)} \sum_{t=1}^{n-1} S_{d,t}S_{d-1,t}, \ldots, n^{-1} \sum_{t=1}^{n-1} S_{d,t}\, S_{1,t})'$. Let $\theta(n)$ denote the vector of regression coefficients in the regression of S_{dt} on $\mathbf{c}_t' = (n^{-(d-1)} S_{d-1,t}, \ldots, n^{-1}S_{1,t})$. Then,

$$\theta(n) = \mathbf{G}_{d-1,n}^{-1}\mathbf{q}_n$$

and

$$\det[\mathbf{G}_{d,n}] = \det[\mathbf{G}_{d-1,n}]n^{-2d} \sum_{t=1}^{n-1} [S_{d,t} - \mathbf{c}_t'\, \theta(n)]^2 .$$

Note that

$$\|\mathbf{q}_n\|^2 = \mathbf{q}_n'\mathbf{q}_n \geq [n^{-(d-1)} \sum_{t=1}^{n-1} S_{d,t}S_{d-1,t}]^2$$

$$= n^{-(2d-2)}\, \tfrac{1}{4}S_{d,n-1}^4 + 0_p(n^{2d-1}) .$$

Since $n^{-(2d-1)}S_{d,n-1}^2$ converges in distribution to $\sigma_d^2\, \chi_1^2$, we get $n^{-2d}\, \|\mathbf{q}_n\|^2$ is bounded away from zero. Also

$$\|\mathbf{q}_n\|^2 = \sum_{i=1}^{d-1} [n^{-(d-i)} \sum_{t=1}^{n-1} S_{d,t}S_{d-i,t}]^2$$

$$= 0_p(n^{2d}).$$

Also,

$$\lambda_{max}^{-2}[\mathbf{G}_{d-1,n}]\|\mathbf{q}_n\|^2 \leq \|\theta(n)\|^2 \leq \lambda_{min}^{-2}[\mathbf{G}_{d-1,n}]\|\mathbf{q}_n\|^2 .$$

Therefore,

$$\|\theta(n)\| = 0_p(n^d)$$

and

$$\|\theta(n)\|^{-1} = 0_p(n^{-d}) .$$

Now, for $n^* < n$,

$$n^{-2d} \sum_{t=d}^{n-1} [S_{d,t} - \mathbf{c}_t'\theta(n)]^2 \geq n^{-2d}[\theta(n^*) - \theta(n)]'\mathbf{G}_{d-1,n^*} [\theta(n^*) - \theta(n)]$$

$$\geq \lambda_{min}[\mathbf{G}_{d-1,n^*}]n^{-2d}[\|\theta(n^*) - \theta(n)\|^2]$$

$$\geq \lambda_{min}[\mathbf{G}_{d-1,n^*}]n^{-2d}[\|\theta(n^*)\| - \|\theta(n)\|]^2 .$$

We now show that $n^{-2d}[\|\theta(n^*)\| - \|\theta(n)\|]^2$ is bounded away from zero. Given $\epsilon > 0$, find $M_1 > 0$, $M_2 > 0$ and N_1 large such that

$$P[M_1 n^d \leq \|\theta(n)\| \leq M_2 n^d] > 1 - \epsilon,$$

for $n > N_1$. Let,

$$n^* = \text{integer part of } \left[n \left(\frac{M_1}{2M_2} \right)^{\frac{1}{d}} \right]$$

and

$$N = \text{integer part of } \left[3N_1 \left(\frac{2M_2}{M_1} \right)^{\frac{1}{d}} \right].$$

Then, for $n > N$, $n^* > N_1$ and

$$P[\|\theta(n^*)\| < \frac{M_1}{2} n^d] > 1 - \epsilon.$$

Therefore, for $n > N$,

$$P[\{\|\theta(n)\| - \|\theta(n^*)\|\}^2 > \frac{M_1}{4} n^{2d}] > 1 - 2\epsilon.$$

Since $\lambda_{min}[\mathbf{G}_{d-1,n^*}]$ is bounded away from zero, we get $\det[\mathbf{G}_{d,n}]$ is bounded away from zero in probability. ||

We now obtain the asymptotic distributions of $\mathbf{G}_{d,n}$ and $\mathbf{g}_{d,n}$. Define,

$$\mathbf{A}_n = \mathbf{L}_n'\mathbf{L}_n \tag{A.4}$$

where $\mathbf{L}_n$ is an (n-1)x(n-1) lower triangular matrix with every element below and on the diagonal is equal to 1. Then,

$$\mathbf{L}_n' = \mathbf{1}\mathbf{1}' + \mathbf{I}_{n-1} - \mathbf{L}_n \tag{A.5}$$

where $\mathbf{1} = (1,1,\ldots,1)'$: (n-1)x1, $\mathbf{I}_{n-1}$ is an identity matrix of size (n-1). Define $\mathbf{e}_n = (e_1, e_2, \ldots, e_{n-1})'$. Let $\lambda_{1,n} > \lambda_{2,n} > \ldots > \lambda_{n-1,n}$ denote the eigenvalues of $\mathbf{A}_n$ and let $\mathbf{x}_{in} = (x_{in1}, x_{in2}, \ldots, x_{in,n-1})'$ denote the eigenvector associated with λ_{in}. Define the orthogonal transformation of $\mathbf{e}_n$ into $\mathbf{u}_n = (u_{1n}, u_{2n}, \ldots, u_{n-1,n})'$ by

$$u_{in} = \sum_{t=1}^{n-1} x_{int}\, e_t.$$

Hasza and Fuller (1979) established that, for a fixed k,

$$(u_{1n}, u_{2n}, \ldots, u_{kn})' \xrightarrow{\mathcal{D}} (V_1, V_2, \ldots, V_k)' \sim N(\mathbf{0}, \mathbf{I}_k).$$

Following the arguments of Hasza and Fuller (1979) it can be shown that

$$n^{-2k} \mathbf{e}_n' \mathbf{A}_n^k \mathbf{e}_n \xrightarrow{\mathcal{D}} \sum_{i=1}^{\infty} \gamma_i^{2k} V_i^2 = Q_k^*$$

where $\gamma_i = 2[(2i - 1)\pi]^{-1} (-1)^{i+1}$ and $\{V_i\}$ is a sequence of normal independent (0,1) random variables. Also, the random variables N_k defined in Lemma 1 may be written as $N_k = \sum_{i=1}^{\infty} \delta_{ki} V_i$ for some constants $\{\delta_{ki}\}$. Now we obtain the limiting distribution of the diagonal elements of $\mathbf{G}_{d,n}$.

<u>Lemma 3</u>: Let $N_{i,(n)} = n^{-(i-\frac{1}{2})} S_{i,n-1}$, where $S_{i,n-1}$ is defined in (A.1). Then, for $k \geq 1$,

$$n^{-2k} \mathbf{A}_n^k \mathbf{e}_n = \sum_{i=1}^{k} \mathbf{a}_{n,i}^{(k)} N_{2i,(n-1)} + (-1)^k n^{-2k} \mathbf{S}_n^{(2k)} + 0_p(n^{-1}) \qquad \text{(A.6)}$$

where $\mathbf{S}_n^{(2k)} = (S_{2k,1}, S_{2k,2}, \ldots, S_{2k,n-1})'$, the <u>jth</u> element of $\mathbf{a}_{n,i}^{(k)}$ is of the form $n^{-\frac{1}{2}} \sum_{\ell=0}^{2(k-i)} v_{k,i,\ell} \left(\frac{j}{n}\right)^{2(k-i)-\ell} + 0(n^{-1})$ and $v_{k,i,\ell}$ is a finite positive constant.

<u>Proof</u>: We prove this using induction on k. For k =1, from Dickey (1977),

$$n^{-2} \mathbf{A}_n \mathbf{e}_n \doteq n^{-\frac{1}{2}} \mathbf{1} N_{2,(n-1)} - \mathbf{S}_n^{(2)} + 0_p(n^{-1})$$

and hence (3.6) holds. Now, we assume that (A.6) holds for k and show that it also holds for k + 1. Note that,

$$n^{-(2k+1)} \mathbf{L}_n \mathbf{A}_n^k \mathbf{e}_n = n^{-1} \sum_{i=1}^{k} \mathbf{L}_n \mathbf{a}_{n,i}^{(k)} N_{2i,(n-1)} + (-1)^k n^{-(2k+1)} \mathbf{L}_n \mathbf{S}_n^{(2k)} + 0_p(n^{-1})$$

$$= \sum_{i=1}^{k} \mathbf{f}_{n,i}^{(k)} N_{2i,(n-1)} + (-1)^k n^{-(2k+1)} \mathbf{S}_n^{(2k+1)} + 0_p(n^{-1}) \qquad \text{(A.7)}$$

where

$$\mathbf{f}_{n,i}^{(k)} = n^{-1} \mathbf{L}_n \mathbf{a}_{n,i}^{(k)} .$$

Note that the <u>jth</u> element of $\mathbf{f}_{n,i}^{(k)}$ is

$$f_{n,i,j}^{(k)} = n^{-1} \sum_{r=1}^{j} a_{n,i,r}^{(k)}$$

$$= n^{-\frac{1}{2}} \sum_{\ell=0}^{2(k-i)+1} v_{k,i,\ell}^{*} \left(\frac{j}{n}\right)^{2(k-i)+1-\ell} + 0(n^{-1})$$

where $v_{k,i,\ell}^{*}$ is a constant. Now,

$$n^{-(2k+2)} \mathbf{A}_n^{k+1} \mathbf{e}_n = n^{-1}\mathbf{L}_n' \, n^{-(2k+1)} \mathbf{L}_n \mathbf{A}_n^k \mathbf{e}_n$$

$$= n^{-1}[\mathbf{1}\mathbf{1}' + \mathbf{I}_{n-1} - \mathbf{L}_n] \, \mathbf{L}_n \mathbf{A}_n^k \mathbf{e}_n$$

$$= n^{-1} \sum_{i=1}^{k} [\mathbf{1}\mathbf{1}' + \mathbf{I}_{n-1} - \mathbf{L}_n] \mathbf{f}_{n,i}^{(k)} N_{2i,(n-1)}$$

$$+ (-1)^k n^{-(2k+2)} \mathbf{1}\mathbf{1}'\mathbf{S}_n^{(2k+1)} + (-1)^k n^{-(2k+2)} [\mathbf{I}_{n-1} - \mathbf{L}_n] \mathbf{S}_n^{(2k+1)}$$

$$+ 0_p(n^{-1})$$

$$= \sum_{i=1}^{k} \mathbf{a}_{n,i}^{(k+1)} N_{2i,(n-1)} + (-1)^k n^{-(2k+2)} S_{2k+2,n-1}$$

$$+ (-1)^k n^{-(2k+2)} [\mathbf{S}_n^{(2k+1)} - \mathbf{S}_n^{(2k+2)}] + 0_p(n^{-1})$$

where

$$\mathbf{a}_{n,i}^{(k+1)} = n^{-1}[\mathbf{1}(\mathbf{1}'\mathbf{f}_{n,i}^{(k)}) + \mathbf{f}_{n,i}^{(k)} - \mathbf{L}_n \, f_{n,i}^{(k)}] \;, \; i = 1, 2, \ldots, k.$$

Note that the <u>jth</u> element of $\mathbf{a}_{n,i}^{(k+1)}$ has the form,

$$a_{n,i,j}^{(k+1)} = n^{-\frac{1}{2}} \sum_{\ell=0}^{2(k-i)+2} v_{k+1,i,\ell} \left(\frac{j}{n}\right)^{2(k-i)+2-\ell} + 0(n^{-1}).$$

Therefore,

$$n^{-(2k+2)} \mathbf{A}_n \mathbf{e}_n = \sum_{i=1}^{k+1} \mathbf{a}_{n,i}^{(k+1)} N_{2i,(n-1)} + (-1)^{k+1} \mathbf{S}_n^{(2k+2)} + 0_p(n^{-1}) \; . \quad \Box$$

Note, from (A.6) and (A.7) we get

$$n^{-4k} \sum_{t=1}^{n-1} S_{2k,t}^2 = n^{-4k} \mathbf{e}_n'\mathbf{A}_n^{2k}\mathbf{e}_n - \sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{a}_{n,i}^{(k)'} \mathbf{a}_{n,j}^{(k)} N_{2i,(n-1)} N_{2j,(n-1)}$$

$$- 2(-1)^k \sum_{i=1}^{k} N_{2i,(n-1)} n^{-2k} \mathbf{a}_{n,i}^{(k)'} \mathbf{S}_n^{(2k)}$$

$$+ 0_p(n^{-1}),$$

and

$$n^{-(4k+2)} \sum_{t=1}^{n-1} S_{2k+1,t}^2 = n^{-(4k+2)} \mathbf{e}_n' \mathbf{A}_n^{2k+1} \mathbf{e}_n - \sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{f}_{n,i}^{(k)'} \mathbf{f}_{n,j}^{(k)} N_{2i,(n-1)} N_{2j,(n-1)}$$

$$- 2(-1)^k \sum_{i=1}^{k} N_{2i,(n-1)} \, n^{-(2k+1)} \mathbf{f}_{n,i}^{(k)'} \mathbf{s}_n^{(2k+1)}$$

$$+ 0_p(n^{-1}).$$

It can be shown that,

$$n^{-r-\frac{1}{2}} \sum_{t=1}^{n-1} \left(\frac{t}{n}\right)^m S_{r,t} \xrightarrow{\mathcal{D}} N_{m,r}^* \sim N(0, \sigma_{m,r}^{*2})$$

for any integers m, r $\geq$ 0. In fact, one can express

$$N_{m,r}^* = \sum_{i=1}^{\infty} \delta_{m,r,i}^* V_i$$

where $V_i \sim NID(0,1)$. Therefore,

$$n^{-2k} \mathbf{a}_{ni}^{(k)'} \mathbf{s}_n^{(2k)} = n^{-2k} \sum_{t=1}^{n-1} a_{n,i,t}^{(k)} S_{2k,t}$$

$$= \sum_{\ell=0}^{2(k-i)} v_{k,i,\ell} \, n^{-2k-\frac{1}{2}} \sum_{t=1}^{n-1} \left(\frac{t}{n}\right)^{2(k-i)-\ell} S_{2k,t} + 0_p(n^{-1})$$

$$\xrightarrow{\mathcal{D}} \sum_{\ell=0}^{2(k-i)} v_{k,i,\ell} \, N_{2(k-i)-\ell,2k}^* = N_{i,k}^{(a)}, \text{ say.}$$

Similarly,

$$n^{-(2k+1)} \mathbf{f}_{n,i}^{(k)'} \mathbf{s}_n^{(2k+1)} \xrightarrow{\mathcal{D}} \sum_{\ell=0}^{2(k-i)+1} v_{k,i,\ell}^* \, N_{2(k-i)-\ell+1,2k}^* = N_{i,k}^{(f)}.$$

Also, there exist finite positive constants $a_{i,j,k}^*$ and $f_{i,j,k}^*$ such that

$$\lim_{n\to\infty} \mathbf{a}_{n,i}^{(k)'} \mathbf{a}_{n,i}^{(k)} = a_{i,j,k}^*$$

and

$$\lim_{n\to\infty} \mathbf{f}_{n,i}^{(k)'} \mathbf{f}_{n,j}^{(k)} = f_{i,j,k}^*.$$

Therefore,

$$n^{-4k} \sum_{t=1}^{n-1} S_{2k,t}^2 \xrightarrow{\mathcal{D}} Q_{2k} = Q_{2k}^* - \sum_{i=1}^{k} \sum_{j=1}^{k} a_{i,j,k}^* N_{2i} N_{2j}$$

$$+ 2(-1)^{k+1} \sum_{i=1}^{k} N_{2i} N_{i,k}^{(a)}$$

and

$$n^{-(4k+2)} \sum_{t=1}^{n-1} S_{2k+1,t}^2 \xrightarrow{\mathcal{D}} Q_{2k} = Q_{2k+1}^* - \sum_{i=1}^{k} \sum_{j=1}^{k} f_{i,j,k}^* N_{2i} N_{2j}$$

$$+ 2(-1)^{k+1} \sum_{i=1}^{k} N_{2i} N_{i,k}^{(f)} .$$

Now, using the results of Lemma 1, we get for $0 \leq j \leq i$ and $i \geq 2$,

$$n^{-(i+j)} \sum_{t=1}^{n-1} S_{i,t-1} S_{j,t-1} \xrightarrow{\mathcal{D}} P_{d,i,j}$$

where

$$P_{d,i,j} = \begin{cases} \sum_{r=0}^{k-1} (-1)^r N_{i-r} N_{j+r+1} + (-1)^k Q_k , & \text{if } i-j \text{ is even} \\ \sum_{r=0}^{k-1} (-1)^r N_{i-r} N_{j+r+1} + (-1)^k \tfrac{1}{2} N_{i-k}^2 , & \text{if } i-j \text{ is odd,} \end{cases}$$

and k = integer part of $\tfrac{1}{2}(i-j)$. Note also that

$$n^{-1} \sum_{t=1}^{n-1} S_{1,t} S_{0,t+1} \xrightarrow{\mathcal{D}} \tfrac{1}{2}(N_1^2 - 1) .$$

Therefore,

$$\mathbf{G}_{d,n} \xrightarrow{\mathcal{D}} \mathbf{G}_d$$

and

$$\mathbf{g}_{d,n} \xrightarrow{\mathcal{D}} \mathbf{g}_d \qquad (A.9)$$

where $\mathbf{G}_d$ and $\mathbf{g}_d$ consist of the corresponding limiting variables. We now present the limiting distributions of the "F-type" statistics.

Lemma 4: Let $F_{i,n}(d)$ be as defined in (2.10) with $Y_t = W_t$. Then, for $i \leq d$,

$$F_{i,n}(d) \xrightarrow{\mathcal{D}} i^{-1} \mathbf{g}_{(i),d}^{*\prime} \left[\mathbf{G}_d^{(ii)}\right]^{-1} \mathbf{g}_{(i),d}^* = F_i(d)$$

where $\mathbf{G}_d^{(ii)}$ is the i x i upper left hand submatrix of $\mathbf{G}_d^{-1}$ and $\mathbf{g}_{(i),d}^*$ is an i x 1 vector consisting of the first i elements of $\mathbf{G}_d^{-1}\mathbf{g}_d$. (Note that $F_d(d) = d^{-1}\mathbf{g}_d'\mathbf{G}_d^{-1}\mathbf{g}_d$.)

<u>Proof</u>: Recall,

$$F_{i,n}(d) = [i(MSE)_n]^{-1} \tilde{\beta}'_{(i)} C^{-1}_{(i)} \tilde{\beta}_{(i)}$$

$$= [i(MSE)_n]^{-1} \tilde{\beta}' U_i^* [U_i^* B^{-1}_{d,n} U_i^*]^{-1} U_i^* \tilde{\beta}$$

where

$$U^*_{(i)} = (i \times i) \text{ left hand corner submatrix of } I_d$$

and

$$\tilde{\beta} = B^{-1}_{d,n} b_{d,n}$$

as defined in (A.3). Note that,

$$(MSE)_n = (n-d)^{-1} \sum_{t=1}^{n} W^2_{d,t} - \tilde{\beta}' b_{d,n}$$

$$= (n-d)^{-1} \sum_{t=1}^{n} e_t^2 + 0_p(n^{-1})$$

$$\xrightarrow{P} 1.$$

Also,

$$D_n \tilde{\beta} \xrightarrow{\mathcal{D}} G_d^{-1} g_{(d)}$$

and

$$D_n B^{-1}_{d,n} D_n \xrightarrow{\mathcal{D}} G_d^{-1}$$

where D_n = diagonal $\{n^d, n^{d-1}, \ldots, n\}$. Therefore,

$$F_{i,n}(d) \xrightarrow{\mathcal{D}} \frac{1}{i} g_d' G_d^{-1} U_i^* [G_d^{(ii)}]^{-1} U_i^* G_d^{-1} g_d \; . \qquad \square$$

Now we present the proof of the theorem.

<u>Proof of Theorem 1</u>: Recall that Y_t is a p<u>th</u> order process with d unit roots. Let W_t and Z_t be as defined in (2.5). Then, W_t is a d<u>th</u> order autoregressive process with d unit roots and Z_t is a (p - d)<u>th</u> order autoregressive process with stationary roots. We can express Z_t as,

$$Z_{p-d,t} = \sum_{j=d+1}^{p} \eta_j Z_{j-d-1,t-1} + e_t ,$$

where $Z_{i,t} = (1 - B)^i Z_t$. Note that $\eta_{d+1} < 0$. From Fuller (1976), it follows that

$$n^{\frac{1}{2}}(\tilde{\eta} - \eta) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{T}^{-1}\Gamma_Z^{-1} \mathbf{T}^{-1\prime}) \qquad (A.10)$$

where $\tilde{\eta}$ is the least squares estimator of η obtained by regressing $Z_{p-d,t}$ on $Z_{0,t-1}, \ldots, Z_{p-d-1,t-1}$; $\mathbf{T}$ is given in (2.3); the (i,j)th element of the matrix Γ_Z is $\gamma_Z(|i-j|)$ and $\gamma_Z(h) = \lim_{t\to\infty} \mathrm{Cov}(Z_t, Z_{t-h})$. Let $\hat{\beta}$ be obtained by regressing $Y_{p,t}$ on $Y_{0,t-1}, Y_{1,t-1}, \ldots, Y_{p-1,t-1}$, where $Y_{i,t} = (1-B)^i Y_t$. Note that, for $i \geq d$, $Y_{i,t} = (1-B)^{i-d} Z_t = Z_{i-d,t}$. Therefore, $\hat{\beta}$ is obtained by regressing $Z_{p-d,t}$ on $Y_{0,t-1}, Y_{1,t-1}, \ldots, Y_{d-1,t-1}, Z_{0,t-1}, Z_{1,t-1}, \ldots, Z_{p-d-1,t-1}$. Let us partition $\hat{\beta}$ as $(\hat{\beta}^{(1)\prime}, \hat{\eta}')'$. We can obtain $\hat{\eta}$ using several regressions. First, regress $Z_{p-d,t}$ on $\psi_t' = (Y_{0,t-1}, Y_{1,t-1}, \ldots, Y_{d-1,t-1})$. Let $R_{p-d,t} = Z_{p-d,t} - \psi_t' \hat{\eta}^{(p-s)}$, be the tth residual from this regression where $\hat{\eta}^{(p-s)}$ is the corresponding least squares estimator of η. Similarly, regress $Z_{i,t-1}$ on ψ_t' and obtain $R_{i,t}$ and $\hat{\eta}^{(i)}$, for $i = 0, 1, \ldots, p-d-1$. Finally, regress $R_{p-d,t}$ on $R_{0,t}, R_{1,t}, \ldots, R_{p-d-1,t}$, to obtain $\hat{\eta}$.

For $i < d$, let $S^*_{d-i,t} = Y_{i,t}$. Then,

$$Y_{d-i,t} = S^*_{i,t} = [(i-1)!]^{-1} \sum_{r=1}^{t} (t-r+1)(t-r+2)\ldots(t-r+i-1)Z_r .$$

Using Kronecker's lemma and the arguments of Lemma 1, it can be shown that,

$$\sum_{t=1}^{n} Y^2_{i,t} = O_p(n^{2d-2i}) ,$$

and

$$\sum_{t=1}^{n} Y_{i,t-1} Y_{j,t-1} = O_p(n^{2d-i-j})$$

for $0 \leq i,j \leq d-1$. Similarly, one can show that,

$$\sum_{t=1}^{n} Y_{i,t-1} Z_{t-k} = O_p(n^{d-i})$$

and hence

$$\sum_{t=1}^{n} Y_{i,t-1} Z_{j,t-1} = O_p(n^{d-i})$$

for $i = 0, 1, 2, \ldots, d-1$; $k = 0, 1, \ldots, p$; $j = 0, 1, \ldots, p-d$.

Now, consider, for $0 \leq i \leq j \leq d-1$,

$$\begin{aligned} n^{-(2d-i-j)} \sum_{t=1}^{n} W_{i,t-1} W_{j,t-1} &= n^{-(2d-i-j)} \sum_{t=1}^{n} [Y_{p-d+i,t} - \eta_{d+1} Y_{i,t-1} - \cdots \\ &\quad - \eta_p Y_{p-d+i-1,t}] \\ &\quad [Y_{p-d+j,t} - \eta_{d+1} Y_{j,t-1} - \cdots - \eta_p Y_{p-d+j-1,t}] \\ &= \eta_{d+1}^2 n^{-(2d-i-j)} \sum_{t=1}^{n} Y_{i,t-1} Y_{j,t-1} + O_p(n^{-1}). \end{aligned}$$

Also, for $0 \leq i \leq d-1$,

$$\begin{aligned} n^{-(d-i)} \sum_{t=1}^{n} W_{it} Z_{p-d,t+1} &= n^{-(d-i)} \sum_{t=1}^{n} [Y_{p-d+i,t} - \eta_{d+1} Y_{i,t-1} - \cdots - \eta_p Y_{p-d+i-1,t}] Z_{p-d,t+1} \\ &= -\eta_{d+1} n^{-(d-i)} \sum_{t=1}^{n} Y_{i,t-1} Z_{p-d,t+1} + O_p(n^{-1}) . \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbf{D}_n^{-1} \boldsymbol{\psi}'\boldsymbol{\psi} \mathbf{D}_n^{-1} &= \sum_{t=1}^{n} \mathbf{D}_n^{-1} \boldsymbol{\psi}_t \boldsymbol{\psi}_t' \mathbf{D}_n^{-1} \\ &= \eta_{d+1}^{-2} \mathbf{G}_{d,n} + O_p(n^{-1}) , \end{aligned}$$

$$\mathbf{D}_n^{-1} \sum_{t=1}^{n} \boldsymbol{\psi}_t Z_{p-d,t} = -\eta_{d+1}^{-1} \mathbf{g}_{d,n} + O_p(n^{-1}) ,$$

and, we get

$$D_n \hat{\eta}^{(i)} = [D_n^{-1}(\psi'\psi)D_n^{-1}]^{-1} D_n^{-1} \sum_{t=1}^{n} \psi_t Z_{i,t}$$

$$= o_p(1) \quad , i = 0, 1, \ldots, p-d-1$$

$$D_n \hat{\eta}^{(p-d)} = -\eta_{d+1} G_{d,n}^{-1} g_{d,n} + o_p(n^{-1})$$

$$= o_p(1)$$

where D_n = diagonal $\{n^d, n^{d-1}, \ldots, n\}$. Now, note that, for $0 \leq i, j \leq p-d-1$,

$$\sum_{t=1}^{n} R_{i,t}R_{j,t} = \sum_{t=1}^{n} [Z_{i,t-1} - \psi_t' \hat{\eta}^{(i)}][Z_{j,t-1} - \psi_t' \hat{\eta}^{(j)}]$$

$$= \sum_{t=1}^{n} Z_{i,t-1}Z_{j,t-1} + o_p(1).$$

Similarly,

$$\sum_{t=1}^{n} R_{i,t} R_{p-d,t} = \sum_{t=1}^{n} Z_{i,t-1} Z_{p-d,t} + o_p(1).$$

Therefore, we get,

$$\hat{\eta} - \tilde{\eta} = o_p(n^{-1})$$

and

$$D_n[\hat{\beta}^{(1)} - (-\eta_{d+1})\tilde{\beta}] = o_p(n^{-1})$$

where $\tilde{\beta}$ is defined in (A.3), and $\tilde{\eta}$ is defined in (A.10). Also, for $i \leq d$,

$$F_{i,n}(p) = F_{i,n}^{*}(d) + o_p(n^{-1})$$

where $F_{i,n}^{*}(d)$ is the "F-statistic" for testing $\beta_1 = \beta_2 = \ldots = \beta_i = 0$ in the regression of $W_{d,t}$ on $W_{0,t-1}, W_{1,t-1}, \ldots, W_{d-1,t-1}$. Therefore,

$$F_{i,n}(p) \xrightarrow{\mathcal{D}} F_i(d).$$

Now, consider,

$$F_{d+1,n}(p) = \frac{1}{(d+1)(MSE)_n} (\hat{\boldsymbol\beta}'^{(1)}, \hat{\eta}'_{d+1})\, \mathbf{C}^{-1}_{(d+1)} \begin{pmatrix} \hat{\boldsymbol\beta}^{(1)} \\ \hat{\eta}_{d+1} \end{pmatrix}$$

where $\mathbf{C}_{(d+1)}$ is the (d+1) x (d+1) submatrix consisting of the first (d+1) rows and (d+1) columns of the matrix,

$$\boldsymbol\Lambda_n = \begin{bmatrix} \sum_{t=1}^{n} \boldsymbol\psi_t \boldsymbol\psi'_t & \sum_{t=1}^{n} \boldsymbol\psi_t \xi'_t \\ \sum_{t=1}^{n} \xi_t \boldsymbol\psi'_t & \sum_{t=1}^{n} \xi_t \xi'_t \end{bmatrix}^{-1}$$

and $\xi'_t = (Z_{0,t-1}, Z_{1,t-1}, \ldots, Z_{p-d-1,t-1})$. Let, $\mathbf{D}^*_n$ = diagonal $\{n^d, n^{d-1}, \ldots, n, n^{\frac{1}{2}}, \ldots, n^{\frac{1}{2}}\}$. Then,

$$\mathbf{D}^*_n \boldsymbol\Lambda_n \mathbf{D}^*_n \xrightarrow{\mathcal{D}} \begin{bmatrix} \eta^2_{d+1} G^{-1}_d & \mathbf{0}' \\ \mathbf{0} & \mathbf{T}^{-1} \Gamma^{-1}_Z \mathbf{T}^{-1\prime} \end{bmatrix} .$$

Therefore, under the assumption of H_d: exactly d unit roots

$$\frac{1}{n} F_{d+1,n}(p) \xrightarrow{P} \eta^2_{d+1} \lfloor (d+1)\, \omega^* \rfloor^{-1}$$

where ω^* is the (1,1)th element of $T^{-1}\Gamma_Z^{-1}T^{-1\prime}$. Therefore,

$$F_{d+1,n}(p) \xrightarrow{\mathcal{D}} \infty$$

under H_d. Now, since $F_{d+i,n}(p) \geq \frac{d+1}{d+i} F_{d+1,n}(p)$, for $i \geq 2$, we have that $F_{d+i,n} \xrightarrow{\mathcal{D}} \infty$, under H_d. □

We now outline the proof of Theorem 2.

Proof of Theorem 2: Let $\hat{\beta}^{(+)}$ be obtained by regressing $Y_{p,t}$ on $(1, \psi_t', \xi_t')$ where ψ_t and ξ_t be as defined in the proof of Theorem 1. Define,

$$\Lambda_n^{(1)} = \begin{bmatrix} n & \sum_{t=1}^{n} \psi_t' & \sum_{t=1}^{n} \xi_t' \\ \sum_{t=1}^{n} \psi_t & \sum_{t=1}^{n} \psi_t \psi_t' & \sum_{t=1}^{n} \psi_t \xi_t' \\ \sum_{t=1}^{n} \xi_t & \sum_{t=1}^{n} \xi_t \psi_t' & \sum_{t=1}^{n} \xi_t \xi_t' \end{bmatrix}^{-1}$$

$$D_n^{(1)} = \text{diagonal } \{n^{\frac{1}{2}}, n^d, n^{d-1}, \ldots, n, n^{\frac{1}{2}}, \ldots, n^{\frac{1}{2}}\} ,$$

and

$$\tau_n^{(1)} = \begin{bmatrix} \sum_{t=1}^{n} e_t \\ \sum_{t=1}^{n} \psi_t' e_t \\ \sum_{t=1}^{n} \xi_t' e_t \end{bmatrix} .$$

Then, it follows from the proof of Theorem 1 that

$$D_n^{(1)} \Lambda_n^{(1)} D_n^{(1)} \xrightarrow{\mathcal{D}} \begin{bmatrix} \Delta\, H_d^{*-1} \Delta & 0' \\ 0 & T^{-1}\, \Gamma_Z^{-1}\, T^{-1\prime} \end{bmatrix} ,$$

and

$$D_n^{(1)^{-1}}\, \tau_n^{(1)} \xrightarrow{\mathcal{D}} \Delta^{-1} h_d , \tag{A.11}$$

where

$$H_d^* = \begin{bmatrix} 1 & \rho_d' \\ \rho_d & G_d \end{bmatrix} ,$$

$$\rho_d = (N_{d+1}, N_d, \ldots, N_2)' ,$$

$$\Delta = \text{diagonal } (1, -\eta_{d+1}, -\eta_{d+1}, \ldots, -\eta_{d+1})$$

$$h_d = \begin{bmatrix} N_1 \\ g_d \end{bmatrix} ,$$

and the matrix $\mathbf{G}_d$ and the vector $\mathbf{g}_d$ are as defined in Theorem 1. Now, following the arguments used in proving Theorem 1, we can show that, under H_d,

$$F^{(1)}_{i,n}(p) \xrightarrow{\mathcal{D}} \begin{cases} F^{(1)}_i(d) & , i \leq d \\ \infty & , i > d \end{cases}$$

where

$F^{(1)}_{i,n}(p)$ = the F-statistic for testing $\beta_1 = \ldots = \beta_i = 0$ in the regression of (2.12),

$$F^{(1)}_i(d) = \frac{1}{i}\, \mathbf{h}^{*'}_{(i),d}\, [\mathbf{H}^{(ii)}_{d,1}]^{-1}\, \mathbf{h}^{*}_{(i),d} ,$$

$\mathbf{H}^{(ii)}_{d,1}$ = (i x i) submatrix consisting of the rows 2 to i+1 and columns 2 to i+1 of the matrix $\mathbf{H}^{*-1}_d$,

and

$\mathbf{h}^{*}_{(i),d}$ = i x 1 vector consisting of the elements 2 to i+1 of the vector $\mathbf{H}^{*-1}_d\, \mathbf{h}_d$.

Similarly,

$$F^{(2)}_{i,n}(p) \xrightarrow{\mathcal{D}} \begin{cases} F^{(2)}_i(d) & , i \leq d \\ \infty & , i > d \end{cases}$$

where

$F^{(2)}_{i,n}(p)$ = the F-statistic for testing $\beta_0 = \beta_1 = \ldots = \beta_i = 0$ in the regression of (2.12)

$$F^{(2)}_i(d) = \frac{1}{i+1}\, \mathbf{h}^{**'}_{(i),d}\, [\mathbf{H}^{(ii)}_{d,2}]^{-1}\, \mathbf{h}^{**}_{(i),d} ,$$

$\mathbf{H}^{(i,i)}_{d,2}$ = (i+1) x (i+1) submatrix consisting of the first (i+1) rows and (i+1) columns of the matrix $\mathbf{H}^{*-1}_d$,

and

$\mathbf{h}^{**}_{(i),d}$ = (i+1) x 1 vector consisting of the first (i+1) elements of $\mathbf{H}^{*-1}_d\, \mathbf{h}_d$. ||

REFERENCES

Anderson, T. W. (1959). On asymptotic distributions of parameters of stochastic difference equations. Annals of Mathematical Statistics, **30**, 676-687.

Dickey, D. A. (1976). Estimation and hypothesis testing in nonstationary time series. Ph.D. Thesis, Department of Statistics, Iowa State University, Ames, Iowa.

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

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, **49**, 1057-1072.

Dickey, D. A., Hasza, D. P., and Fuller, W. A. (1984). Testing for unit roots in seasonal time series. Journal of American Statistical Association, **79**, 355-367

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

Fuller, W. A. (1979). Testing the autoregressive process for a unit root. Proceedings of the 42nd Session of the International Statistical Institute, Manilla, Spain.

Hannan, E. J. and Heyde, C. C. (1972). On limit theorems for quadratic functions of discrete time series. Annals of Mathematical Statistics, **43**, 2058-2066.

Hasza, D. P. (1977). Estimation in nonstationary time series. Ph.D. Thesis, Iowa State University, Ames, Iowa.

Hasza, D. P. and Fuller, W. A. (1979). Estimation for autoregressive processes with unit roots. Annals of Statistics, **7**, 1106-1120.

Lai, T. L. and Wei, C. Z. (1983). Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters. Journal of Multivariate Analysis, **13**, 1-23.

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

Pankratz, A. (1983). Forecasting with Univariate Box-Jenkins Models: Concepts and Cases. Wiley, New York.

Rao, M. M. (1978). Asymptotic distribution of an estimator of the boundary parameter of an unstable process. Annals of Statistics, **6**, 185-190.

Sen, D. L. (1985). Robustness of single unit root test statistics in the presence of multiple unit roots. Ph.D. Thesis, North Carolina State University, Raleigh, North Carolina.