The Annals of Statistics
1989, Vol. 17, No. 1, 419–428

TESTING FOR A UNIT ROOT NONSTATIONARITY IN MULTIVARIATE AUTOREGRESSIVE TIME SERIES

BY NICOLAOS G. FOUNTIS AND DAVID A. DICKEY

North Carolina State University

The characteristic equation of a multiple autoregressive time series involves the eigenvalues of a matrix equation which determine if the series is stationary. Suppose one eigenvalue is 1 and the rest are less than 1 in magnitude. We show that ordinary least squares may be used to estimate the matrices involved and that the largest estimated eigenvalue has distributional properties that allow us to test this unit root hypothesis using critical values tabulated by Dickey (1976). See also Fountis (1983). If a single unit root is suspected, a model can be fit whose parameters are constrained to produce an exact unit root. This is the vector analog of differencing in the univariate case. In the fitting process, canonical series can be computed thus extending the work of Box and Tiao (1977) to the unit root case.

1. Introduction. Consider the multivariate first-order autoregressive [AR(1)] process defined by the rule

(1.1) $$Y_t = AY_{t-1} + e_t, \qquad t = 1, 2, \dots,$$

where

$$Y_t = [Y_{1,t}, Y_{2,t}, \dots, Y_{k,t}]^{\mathrm{T}},$$
$$e_t = [e_{1,t}, e_{2,t}, \dots, e_{k,t}]^{\mathrm{T}},$$
$$Y_0 = \emptyset,$$

and $\{e_t;\ t = 1, 2, \dots\}$ is a sequence of independent identically distributed multivariate normal variates with mean $\emptyset$ (a vector of 0's) and variance matrix Σ. Henceforth we assume

(1.2) $$\sum_{i=0}^{k-1} A^i \Sigma (A^i)^{\mathrm{T}}$$

is of full rank. In linear system theory, (1.2) is the definition of controllability of $[A, \Sigma^{1/2}]$.

We also assume

(1.3) A has an eigenvalue 1 and the rest less than 1 in magnitude.

There exists a real matrix R such that

$$V \equiv R^{-1}AR = \begin{bmatrix} 1 & \emptyset^{\mathrm{T}} \\ \emptyset & D \end{bmatrix},$$

where the eigenvalues of the $(k-1) \times (k-1)$ matrix D are those of A which

Received June 1983; revised April 1988.
AMS 1980 *subject classifications*. Primary 62M10; secondary 62M02.
Key words and phrases. Time series, nonstationary, unit root, multivariate.

are less than 1 in magnitude. The matrix R^{-1} provides a transformation

$$R^{-1}Y_t \equiv Z_t = \left[z_{1,t}, Z_{2,t}^{\mathrm{T}}\right]^{\mathrm{T}} \tag{1.4}$$

for which

$$Z_t = VZ_{t-1} + \varepsilon_t, \tag{1.5}$$

where

$$\varepsilon_t = R^{-1}e_t = \left[\varepsilon_{1,t}, \varepsilon_{2,t}^{\mathrm{T}}\right]^{\mathrm{T}}.$$

Equation (1.5) can be written as

$$\begin{aligned} z_{1,t} &= z_{1,t-1} + \varepsilon_{1,t}, \\ Z_{2,t} &= DZ_{2,t-1} + \varepsilon_{2,t}. \end{aligned} \tag{1.6}$$

Write the appropriately partitioned covariance matrix of ε_t as

$$R^{-1}\Sigma(R^{-1})^{\mathrm{T}} \equiv J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}$$

and note that

$$\begin{aligned} \sum_{i=0}^{k-1} V^iJ(V^i)^{\mathrm{T}} &= \begin{bmatrix} kJ_{11} & J_{12}\sum_{i=0}^{k-1}(D^i)^{\mathrm{T}} \\ \sum_{i=0}^{k-1} D^iJ_{21} & \sum_{i=0}^{k-1} D^iJ_{22}(D^i)^{\mathrm{T}} \end{bmatrix} \\ &= R^{-1}\left[\sum_{i=0}^{k-1} A^i\Sigma(A^i)^{\mathrm{T}}\right](R^{-1})^{\mathrm{T}}. \end{aligned}$$

Now Σ is positive definite and R is full rank so $J_{11} > 0$ and $\sum_{i=0}^{k-1} D^iJ_{22}(D^i)^{\mathrm{T}}$ is full rank.

LEMMA 1. *The controllability assumption* (1.2) *and the unit eigenvalue assumption* (1.3) *imply*

(i) $$n^{-1}\sum_{t=0}^{n} Z_{2,t}Z_{2,t}^{\mathrm{T}} \to \sum_{i=0}^{\infty} D^iJ_{22}(D^i)^{\mathrm{T}} \equiv \Phi \quad a.s.,$$

(ii) $$\sum_{t=0}^{n} z_{1,t}Z_{2,t}^{\mathrm{T}} = O_p(n),$$

(iii) $$\sum_{t=0}^{n} z_{1,t}\varepsilon_{2,t} = O_p(n).$$

PROOF. Result (i) follows from Theorem 2 of Lai and Wei (1985). Results (ii) and (iii) follow by arguments similar to those of Lemma 3.4.3 in Chan and Wei (1988). See also Dickey and Fuller (1979) and Tiao and Tsay (1983). □

2. Distribution of estimators. The matrix A is a $k \times k$ real matrix whose eigenvalues determine whether the sequence $\{Y_t\}$ is stationary. If all of the eigenvalues of A are less than 1 in magnitude, we will say that the series is stationary. Otherwise we will say that it is nonstationary. This differs slightly from the convention for stationary series because we have fixed the initial value Y_0. Given the observations $Y_1, Y_2, \dots, Y_n$, define the $k \times n$ data matrices

$$Y = [Y_t;\ t = 1, \dots, n] \quad \text{and} \quad Y_L = [Y_{t-1};\ t = 1, \dots, n].$$

The ordinary least-squares estimate of A is given by

$$\hat{A}_n = (YY_L^{\mathrm{T}})(Y_L Y_L^{\mathrm{T}})^{-1}.$$

Define $Z = R^{-1}Y$, $Z_L = R^{-1}Y_L$ and $\hat{V}_n \equiv (ZZ_L^{\mathrm{T}})(Z_L Z_L^{\mathrm{T}})^{-1} = R^{-1}\hat{A}_n R$.

LEMMA 2. *The controllability assumption* (1.2) *and unit eigenvalue assumption* (1.3) *imply* $\hat{V}_n \to V$ *a.s.*

PROOF. By Corollary 3 of Lai and Wei (1985), $\hat{A}_n \to A$ a.s. Hence $\hat{V}_n = R^{-1}\hat{A}_n R \to R^{-1}AR = V$ a.s. □

By similarity, the eigenvalues of $\hat{V}_n$ and $\hat{A}_n$ are the same. Now let $\hat{\lambda}_n$ be the eigenvalue of $\hat{A}_n$ (or $\hat{V}_n$) with largest magnitude. Lemma 2, and its assumptions (1.2) and (1.3), imply

$$\hat{\lambda}_n \to 1 \quad \text{a.s.} \tag{2.1}$$

Let

$$W_{1,n} = \sum_{t=2}^{n} z_{1,t-1}^2, \qquad W_{2,n} = \sum_{t=2}^{n} z_{1,t} Z_{2,t-1}, \qquad W_{3,n} = \sum_{t=2}^{n} Z_{2,t-1} Z_{2,t-1}^{\mathrm{T}},$$

$$W_{4,n} = \sum_{t=2}^{n} z_{1,t-1}\varepsilon_{1,t}, \qquad W_{5,n} = \sum_{t=2}^{n} Z_{2,t-1}^{\mathrm{T}}\varepsilon_{1,t}, \qquad W_{6,n} = \sum_{t=2}^{n} z_{1,t-1}\varepsilon_{2,t}$$

and

$$W_{7,n} = \sum_{t=2}^{n} Z_{2,t-1}\varepsilon_{2,t}^{\mathrm{T}}.$$

Then

$$\hat{V}_n - V = \begin{bmatrix} W_{4,n} & W_{5,n} \\ W_{6,n} & W_{7,n} \end{bmatrix} \begin{bmatrix} W_{1,n} & W_{2,n}^{\mathrm{T}} \\ W_{2,n} & W_{3,n} \end{bmatrix}^{-1}.$$

Following Chan and Wei (1988) we obtain the following lemma.

LEMMA 3. *Assuming controllability* (1.2) *and a unit root* (1.3),

$$\begin{bmatrix} n^{-2} & W_{1,n} \\ n^{-1} & W_{4,n} \\ n^{-1/2} & W_{5,n} \\ n^{-1/2} & W_{7,n} \end{bmatrix} \xrightarrow{\mathscr{L}} \begin{bmatrix} \Gamma \\ \xi \\ C \end{bmatrix},$$

where $\Gamma = \sum_{j=1}^{\infty}\gamma_j H_j^2$, $\xi = 0.5(T_1^2 - J_{11})$, $T_1 = \sqrt{2}\sum_{j=1}^{\infty}\gamma_j^{1/2}(-1)^{j+1}H_j$, $\gamma_j = [2/(2j-1)\pi]^2$, H_j *iid* $N(0, J_{11})$ *and* $C \sim N(0, J \otimes \Phi)$ *independent of* Γ *and* ξ. *Chan and Wei* (1988) *and White* (1958) *give expressions for* Γ *and* ξ *as functionals of a Weiner process.*

Now partition $\hat{V}_n$ as

$$\hat{V}_n = \begin{bmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{bmatrix}.$$

LEMMA 4. *Assuming* (1.2) *and* (1.3),

(a) $$\|\hat{V}_{21}\| = O_p(n^{-1}),$$

(b) $$\|\hat{V}_{12}\| = O_p(n^{-1/2})$$

and

(c) $$n(\hat{V}_{11} - 1) \xrightarrow{\mathscr{L}} \xi/\Gamma,$$

where ξ/Γ *has the limit distribution given in Fuller* (1976), *page* 371.

PROOF.

$$\begin{bmatrix} n(\hat{V}_{11} - 1) & n^{1/2}\hat{V}_{12} \\ n\hat{V}_{21} & n^{1/2}(\hat{V}_{22} - D) \end{bmatrix}$$

$$= (\hat{V}_n - V)\begin{bmatrix} n & \phi^{\mathrm{T}} \\ \phi & n^{1/2}I \end{bmatrix},$$

$$= \begin{bmatrix} n^{-1}W_{4,n} & n^{-1/2}W_{5,n} \\ n^{-1}W_{6,n} & n^{-1/2}W_{7,n} \end{bmatrix}\begin{bmatrix} n^{-2}W_{1,n} & n^{-3/2}W_{2,n}^{\mathrm{T}} \\ n^{-3/2}W_{2,n} & n^{-1}W_{3,n} \end{bmatrix}^{-1}.$$

By Lemmas 1 and 3,

$$\left[\begin{bmatrix} n^{-2}W_{1,n} & n^{-3/2}W_{2,n}^{\mathrm{T}} \\ n^{-3/2}W_{2,n} & n^{-1}W_{3,n} \end{bmatrix}, n^{-1}W_{4,n}\right] \xrightarrow{\mathscr{L}} \left[\begin{bmatrix} \Gamma & \phi^{\mathrm{T}} \\ \phi & \Phi \end{bmatrix}, \xi\right].$$

By Lemma 3 and direct moment computations on $W_{6,n}$,

$$\|n^{-1/2}W_{5,n}\| = O_p(1), \qquad \|n^{-1}W_{6,n}\| = O_p(1)$$

and

$$\|n^{-1/2}W_{7,n}\| = O_p(1),$$

which completes the proof. □

THEOREM 1. *Under assumptions* (1.2) *and* (1.3),

$$n(\hat{\lambda}_n - 1) \xrightarrow{\mathscr{L}} \xi/\Gamma.$$

PROOF. Let $f(\lambda) = |\hat{V}_n - \lambda I|$. Then $f(\hat{\lambda}_n) = 0$. Since $\hat{V}_{22} - \hat{\lambda}_n I \to_{n\to\infty} D - I$ a.s. by Lemma 2, eventually $\hat{V}_{22} - \hat{\lambda}_n I$ is nonsingular and

$$f(\hat{\lambda}_n) = |(\hat{V}_{11} - \hat{\lambda}_n) + \hat{V}_{12}(\hat{V}_{22} - \hat{\lambda}_n I)^{-1}\hat{V}_{21}|\,|\hat{V}_{22} - \hat{\lambda}_n I|.$$

However $|\hat{V}_{22} - \hat{\lambda}_n I| \to_{n\to\infty} |D - I| \neq 0$ a.s. Hence eventually

$$\hat{V}_{11} - \hat{\lambda}_n = -\hat{V}_{12}(\hat{V}_{22} - \hat{\lambda}_n I)^{-1}\hat{V}_{21}.$$

Consequently,

$$|\hat{V}_{11} - \hat{\lambda}_n| = O_p(\|\hat{V}_{12}\|\,\|\hat{V}_{21}\|) = O_p(n^{-3/2}),$$

by Lemma 4. Therefore

$$\begin{aligned} n(\hat{\lambda}_n - 1) &= n(\hat{\lambda}_n - \hat{V}_{11}) + n(\hat{V}_{11} - 1) \\ &= O_p(n^{-1/2}) + n(\hat{V}_{11} - 1) \xrightarrow{\mathscr{L}} \xi/\Gamma \end{aligned}$$

by Lemma 4. □

COROLLARY 1. *Let*

$$Y_t = B_1 Y_{t-1} + \cdots + B_p Y_{t-p} + \varepsilon_t, \tag{2.2}$$

where ε_t *are iid* $N(0, \Sigma)$ *with* Σ *positive definite. Assume that the characteristic equation*

$$\left|\lambda^p I - \sum_{j=1}^{p} \lambda^{p-j} B_j\right| = 0$$

has one root $\lambda_1 = 1$ *and the other* $pk - 1$ *roots less than* 1 *in modulus. Let* $\hat{B}_{jn}$ *be the least-squares estimate of* B_j *based on observations* $(Y_0, \dots, Y_n)$ *and define* $\hat{\lambda}_n$ *to be the root of*

$$\left|\lambda^p I - \sum_{j=1}^{p} \lambda^{p-j} \hat{B}_{jn}\right| = 0$$

with the largest modulus. Then

$$n(\hat{\lambda}_n - 1) \xrightarrow{\mathscr{L}} \xi/\Gamma.$$

PROOF. Let

$$X_t = \begin{bmatrix} Y_t \\ \vdots \\ Y_{t-p+1} \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 & \cdots & & B_p \\ I & \cdots & & \phi \\ \vdots & \ddots & & \vdots \\ \phi & \cdots & I & \phi \end{bmatrix}, \qquad e_t = \begin{bmatrix} \varepsilon_t \\ \phi \\ \vdots \\ \phi \end{bmatrix}.$$

Then B satisfies (1.3) and the covariance matrix of e_t is $\begin{bmatrix} \Sigma & \phi \\ \phi & \phi \end{bmatrix} = \Sigma^*$. It is known that $(B, \Sigma^{*1/2})$ satisfies (1.2) [cf. Lai and Wei (1985)]. Now observe that

$$X_t = BX_{t-1} + e_t. \tag{2.3}$$

The least-squares estimate of B is

$$\begin{aligned} \hat{B}_n &= \sum_{t=1}^{n} X_t X_{t-1}^{\mathrm{T}} \left(\sum_{t=1}^{n} X_{t-1} X_{t-1}^{\mathrm{T}} \right)^{-1} \\ &= B + \sum_{t=1}^{n} \begin{bmatrix} \varepsilon_t \\ \phi \\ \vdots \\ \phi \end{bmatrix} X_{t-1}^{\mathrm{T}} \left(\sum_{t=1}^{n} X_{t-1} X_{t-1}^{\mathrm{T}} \right)^{-1} \\ &= B + \begin{bmatrix} \sum_{t=1}^{n} \varepsilon_t X_{t-1}^{\mathrm{T}} \left(\sum_{t=1}^{n} X_{t-1} X_{t-1}^{\mathrm{T}} \right)^{-1} \\ \phi \end{bmatrix} \\ &= \begin{bmatrix} \hat{B}_{1n} & \cdots & & \hat{B}_{pn} \\ I & \cdots & & \phi \\ \vdots & \ddots & & \vdots \\ \phi & \cdots & I & \phi \end{bmatrix}. \end{aligned}$$

This implies that $\hat{\Lambda}_n$ is the eigenvalue of $\hat{B}_n$ with the largest modulus. Clearly Theorem 1 is applicable. □

3. Stationarity transformation. Reparameterize (2.2) as

$$\nabla Y_t = -CY_{t-1} - \sum_{j=1}^{p-1} W_j \nabla Y_{t-j},$$

where $C = I - B_1 - \cdots - B_p$ and $W_j = B_{j+1} + B_{j+2} + \cdots + B_p$. Note that C has exactly one zero eigenvalue. Let g_1 be the right eigenvector $Cg_1 = \phi$. This implies $g_1 = (\Sigma B_i)g_1$ so that $BG_1 = G_1$, where B is given in (2.3) and $G_1^{\mathrm{T}} = (g_1^{\mathrm{T}}, g_1^{\mathrm{T}}, g_1^{\mathrm{T}}, \dots, g_1^{\mathrm{T}})^{\mathrm{T}}$. Thus we can find a real matrix R such that $R^{-1}BR$ is block diagonal with upper-left block 1 and with the first column of R being G_1. Let l^{T} be the left eigenvector of C corresponding to the zero eigenvalue. Then there is a matrix T with $T^{-1}CT$ having 0 in all first row and all first column entries and with the first row of T^{-1} being l^{T}. Let the first row of R^{-1} be partitioned as $(\gamma_1^{\mathrm{T}}, \gamma_2^{\mathrm{T}}, \gamma_3^{\mathrm{T}}, \ldots, \gamma_p^{\mathrm{T}})$. Since this is a left eigenvector of B corresponding to the unique unit eigenvalue, we see that $\gamma_i^{\mathrm{T}} = \gamma_1^{\mathrm{T}} B_i + \gamma_{i+1}^{\mathrm{T}}$, $i = 1, 2, \ldots, p$, where $\gamma_{p+1} = \emptyset$. Summing over i we get $\gamma_1^{\mathrm{T}}(I - B_1 - B_2 - \cdots - B_p) = \emptyset$ so that γ_1 can be taken to equal l and all other γ's can be obtained from l^{T} and the B_i's.

The transformation matrix T transforms Y_t into $T^{-1}Y_t$ which has a unit root process (not necessarily a pure random walk) as first entry and a stationary vector process forming the other $k - 1$ entries. To see this, note that the matrix R^{-1} transforms X_t into $Z_t = R^{-1}X_t$ which has a random walk z_{1t} as its first entry and remaining entries forming a stationary process Z_{2t}. Define the $k \times kp$ matrix $H = [I, \emptyset, \emptyset, \ldots, \emptyset]$ and note from (2.3) that $HX_t = Y_t$. We have $T^{-1}Y_t = T^{-1}HRZ_t$ and we note that the first column of $T^{-1}HR$ is $(1, 0, 0, \ldots, 0)^{\mathrm{T}}$ which shows that only the first element of $T^{-1}Y_t$ involves z_{1t}. Granger and Weiss (1983) refer to the situation where linear combinations of nonstationary series are stationary as "cointegration." The last $k - 1$ rows of T^{-1} provide cointegrating vectors in our case.

Let ξ_t denote the first element of $T^{-1}Y_t$. Note that ξ_t can be expressed as a linear combination $\beta_0 z_{1t} + \alpha_0^{\mathrm{T}} Z_{2t}$ with $\beta_0 \neq 0$ using the arguments of the preceding paragraph. Similarly note that $Y_t = HRZ_t$ implies that the ith element of Y_t can be expressed as $Y_{it} = \beta_i z_{1t} + \alpha_i^{\mathrm{T}} Z_{2t}$, where $g_1 = (\beta_1, \beta_2, \ldots, \beta_k)^{\mathrm{T}}$ is the previously mentioned eigenvector of C. Using the orders of convergence of sums of squares and cross products for stationary and unit root processes, a simple regression of x_{it} on ξ_t produces a regression coefficient converging to β_i/β_0. Now $x_{it} - (\beta_i/\beta_0)\xi_t = (\alpha_i^{\mathrm{T}} - (\beta_i/\beta_0)\alpha_0^{\mathrm{T}})Z_{2t}$ so the residuals from this regression are (approximately) stationary. Plots of these residual series show interrelationships of component series adjusted for the system nonstationarity. This regression approach was originally suggested to us by David Findley.

Note that $n^{-2}\Sigma Y_t Y_t^{\mathrm{T}}$ converges to a random $(\lim n^{-2}\Sigma z_{1t}^2)$ multiple of $g_1 g_1^{\mathrm{T}}$ so g_1 is approximated by the eigenvector of $n^{-2}\Sigma Y_t Y_t^{\mathrm{T}}$ associated with the largest in magnitude eigenvalue of $\Sigma Y_t Y_t^{\mathrm{T}}$. This is an extension of Box and Tiao's (1977) approach for the stationary case.

4. Example. Consider the U.S. birthrate as measured in births per thousand married women from 1948 through 1980. A logarithmic transformation seems to be appropriate. Let $y_{1,t}$ be the log of birthrate for mothers age 20–24 and let $y_{2,t}$ be the log of birthrate for mothers age 25–29. The data were obtained from the U.S. Bureau of the Census. Define the column vector $Y_t = (y_{1,t}, y_{2,t})^{\mathrm{T}}$. A plot of the data is given by the solid lines in Figure 1.

Our estimated equation is

$$
\begin{aligned}
Y_t &= \begin{bmatrix} -0.1874 \\ 0.2925 \end{bmatrix} + \begin{bmatrix} 1.4085 & 0.1829 \\ 0.2907 & 1.3060 \end{bmatrix} Y_{t-1} \\
&\quad - \begin{bmatrix} 0.5727 & -0.0226 \\ 0.1797 & 0.4797 \end{bmatrix} Y_{t-2} + \varepsilon_t.
\end{aligned}
\tag{4.1}
$$

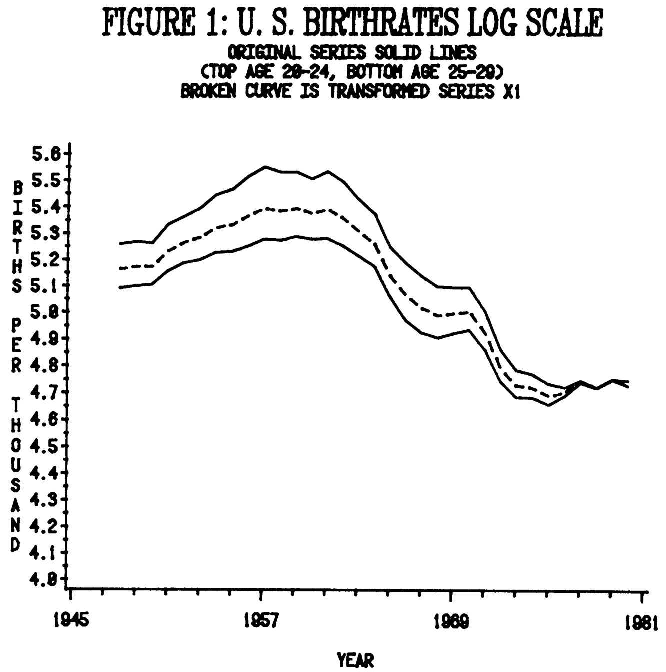

FIG. 1.

The test consists of finding the solutions m to

$$\begin{aligned}|m^2 I - \hat{B}_1 \quad m - \hat{B}_2| &= (m - 0.943589)(m - 0.716984) \\ &\quad \times (m^2 - 1.0539695m + 0.4121333) \\ &= 0.\end{aligned}$$

Then compare $33(0.943589 - 1) = -1.86$ to the left tail critical values of the top display of Table 8.5.1 of Fuller (1976). The unit root hypothesis cannot be rejected at any reasonable level of significance.

From (4.1) we get an estimate $\hat{C} = I - \hat{B}_1 - \hat{B}_2$, where $\hat{B}_1$ and $\hat{B}_2$ are the estimated coefficient matrices. Computing the eigenvalues and vectors gives

$$\begin{aligned}\hat{T}^{-1}\hat{C}\hat{T} &= \begin{bmatrix} 0.63303 & 0.83493 \\ 0.60788 & -0.85342 \end{bmatrix} \begin{bmatrix} 0.1642 & -0.2055 \\ -0.1110 & 0.1737 \end{bmatrix} \begin{bmatrix} 0.814504 & 0.796860 \\ 0.58016 & -0.60417 \end{bmatrix} \\ &= \begin{bmatrix} 0.01784 & 0 \\ 0 & 0.32006 \end{bmatrix}.\end{aligned}$$

Computing $T^{-1}Y_t \equiv U_t$ and the first difference ∇U_t we fit (2.3), using least squares to get the coefficients for each row.

$$\begin{aligned}\nabla U_t &= \begin{bmatrix} 0.1256 \\ -0.3636 \end{bmatrix} + \begin{bmatrix} -0.0179 & 0 \\ 0 & -0.3201 \end{bmatrix} U_{t-1} \\ &\quad + \begin{bmatrix} 0.6416 & 0.1751 \\ -0.0868 & 0.4109 \end{bmatrix} \nabla U_{t-1} + \eta_t.\end{aligned} \tag{4.2}$$

In the univariate case the coefficient on U_{t-1} would be a scalar near 0 and an exact unit root for U_t would be imposed by setting this coefficient to 0 in (4.2), that is, omitting U_{t-1} from the regression. The vector analog is to omit U_{t-1} from the first row regression in (4.2). If we also leave out the intercept in the first row (0.1256 has standard error 0.2367) we get

$$\nabla U_t = \begin{bmatrix} 0 \\ -0.3636 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -0.3201 \end{bmatrix} U_{t-1} + \begin{bmatrix} 0.6431 & 0.0904 \\ -0.0868 & 0.4109 \end{bmatrix} \nabla U_{t-1} + \eta_t. \tag{4.3}$$

Now the entire matrix system can easily be recast into the original Y scale using $\hat{T}$ on the restricted estimates (4.3),

$$Y_t = \begin{bmatrix} -0.2897 \\ 0.2197 \end{bmatrix} + \begin{bmatrix} 1.3765 & 0.2549 \\ 0.2679 & 1.3573 \end{bmatrix} Y_{t-1} - \begin{bmatrix} 0.5315 & 0.0373 \\ 0.1504 & 0.5224 \end{bmatrix} Y_{t-2} + \varepsilon_t.$$

This model contains an exact unit root and its forecasts are, then, the analog of forecasts from a differenced univariate series.

A plot of $y_{1,t}$, $y_{2,t}$ and the canonical unit root process ξ_t is given in Figure 1. The canonical series displays the overall trend in the data. We could plot the other canonical series but we feel it is more informative to regress each element $y_{i,t}$ on ξ_t and plotting the (singular) system of residuals as suggested earlier. A

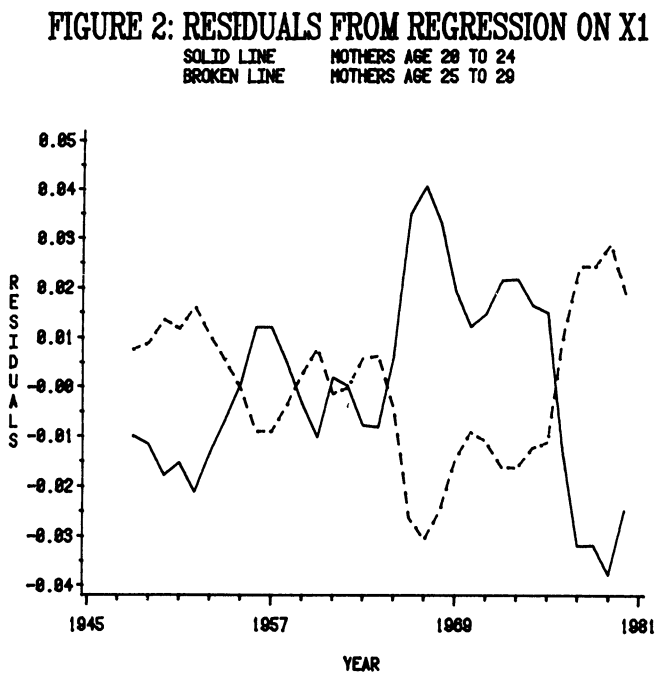

FIG. 2.

plot of the regression residuals for the birth series is given in Figure 2. A study of this graph helps to reveal changing preference over time for motherhood earlier in life versus later in life.

We do not claim that the adjusted series are white noise, only that they are stationary. Since the adjusted series taken together form a singular bivariate system, the autocorrelation structure of the two component series is the same. The first six autocorrelations are 0.83, 0.52, 0.23, 0.03, -0.05 and -0.08. An autoregressive model of order 2 with coefficients 1.3 and -0.5 fits the data well. The fitted autoregression has two complex roots with magnitude 0.75 which is consistent with an assumption of stationarity for these adjusted series. Whether 0.75 is significantly less than 1 would ideally be decided by a statistical hypothesis test, however it is unclear what the effect of our adjustment might be on available unit root tests.

Acknowledgment. The authors wish to thank an anonymous referee for suggesting the method in Sections 1 and 2. This led to greater generality than we originally achieved.

REFERENCES

BOX, G. E. P. and TIAO, G. C. (1977). A canonical analysis of multiple time series. *Biometrika* **64** 355–365.

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

DICKEY, D. A. (1976). Estimation and hypothesis testing in nonstationary time series. Ph.D. dissertation, Iowa State Univ.

DICKEY, D. A. and FULLER, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* **74** 423–431.

FOUNTIS, N. G. (1983). Testing for unit roots in multivariate autoregressions. Ph.D. dissertation, North Carolina State Univ.

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

GRANGER, C. W. J. and WEISS, A. A. (1983). Time series analysis of error-correction models. In *Studies in Econometrics, Time Series, and Multivariate Statistics* (S. Karlin, T. Amemiya and L. A. Goodman, eds.) 255–278. Academic, New York.

LAI, T. L. and WEI, C. Z. (1985). Asymptotic properties of multivariate weighted sums with applications to stochastic regression in linear dynamic systems. In *Multivariate Analysis VI* (P. R. Krishnaiah, ed.) 375–393. North-Holland, Amsterdam.

TIAO, G. C. and TSAY, R. S. (1983). Consistency properties of least squares estimates of autoregressive parameters in ARMA models. *Ann. Statist.* **11** 856–871.

WHITE, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.* **29** 1188–1197.

DEPARTMENT OF STATISTICS
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, NORTH CAROLINA 27695-8203