The Annals of Statistics
1991, Vol. 19, No. 3, 1143–1154

ESTIMATION OF THE PARAMETERS OF LINEAR TIME SERIES MODELS SUBJECT TO NONLINEAR RESTRICTIONS[1]

BY NEERCHAL K. NAGARAJ AND WAYNE A. FULLER

University of Maryland at Baltimore County and Iowa State University

Least squares estimators of the parameters of a linear time series model, where the parameters are constrained by a set of nonlinear restrictions, are studied. The model may contain lags of the dependent variable as regressors and the sums of squares of the explanatory variables may grow at different rates as the sample size increases. The estimation procedures can be applied to a regression model with an error process that satisfies either a stationary or a nonstationary autoregression.

1. Introduction. We consider estimation of the parameter vector $\boldsymbol{\gamma}$ for the model

$$Y_t = \mathbf{Z}_t\boldsymbol{\gamma} + e_t, \tag{1.1}$$

$$\mathbf{f}(\boldsymbol{\gamma}) = \mathbf{0}, \tag{1.2}$$

where $\mathbf{Z}_t = (\psi_{t1}, \dots, \psi_{tq}, Y_{t-1}, \dots, Y_{t-p})$, $\boldsymbol{\gamma}$ is a k-dimensional column vector, ψ_{ti}, $i = 1, 2, \dots, q$, are explanatory variables and $\mathbf{f}(\boldsymbol{\gamma}) = [f_1(\boldsymbol{\gamma}), \dots, f_r(\boldsymbol{\gamma})]'$ is a vector of functional restrictions on $\boldsymbol{\gamma}$. Our interest centers on the case where the $f_i(\boldsymbol{\gamma})$ are nonlinear in $\boldsymbol{\gamma}$. An important feature of the model is that the vector of regressors $\mathbf{Z}_t$ contains lags of the dependent variable. The $\{\psi_{ti}\}$ can be fixed or random sequences. In Section 2, it is assumed that the e_t and ψ_{ti} satisfy conditions such that the properly normalized unrestricted least squares estimator of $\boldsymbol{\gamma}$ has a limiting distribution. For example, the $\{e_t\}$ might be a sequence of independently and identically distributed $(0, \sigma^2)$ random variables or the e_t might be martingale differences with $2 + \delta$, $\delta > 0$, moments.

The characteristic polynomial associated with model (1.1) is

$$m^p - \sum_{j=1}^{p} \gamma_{q+j} m^{p-j} = 0. \tag{1.3}$$

Let the roots of (1.3) be $m_1, m_2, \dots, m_p$ with $|m_1| \geq |m_2| \geq \cdots \geq |m_p|$. The properties of Y_t depend on the nature of the ψ_{ti}, on the distributions of the e_t and on the roots of the polynomial.

Fuller, Hasza and Goebel (1981) derived the asymptotic distribution of the least squares estimator of the unrestricted vector $\boldsymbol{\gamma}$ for the three cases: $|m_1| < 1$, $|m_1| = 1$ and $|m_1| > 1$. The stationary case, when all the roots are

Received March 1988; revised August 1990.
[1]Research partly supported by Joint Statistical Agreements JSA 86-2 and JSA 87-1 with the U.S. Bureau of the Census.
AMS 1980 *subject classifications*. Primary 62M10; secondary 62J02, 62F12.
Key words and phrases. Least squares, nonlinear estimation, nonstationary, time series.

less than 1 in absolute value, has been studied by Mann and Wald (1943), Anderson and Rubin (1950), Hannan and Nicholls (1972) and others. The case when the largest root of the polynomial (1.3) is greater than or equal to 1 in absolute value has been studied by Rubin (1950), White (1958), Fuller (1976), Hasza (1977), Dickey and Fuller (1979), Phillips and Durlauf (1986), Chan and Wei (1988) and others.

The linear model with nonlinear restrictions can also be written as a nonlinear model with the restrictions incorporated into the model equation. The limiting distribution of the nonlinear least squares estimator when the error process $\{e_t\}$ satisfies some mixing conditions are derived in White and Domowitz (1984) and in Wooldridge (1986). Wu (1981) considers nonlinear estimation when the sums of squares of the partial derivatives increase at rates that are not necessarily proportional to n. General discussions of nonlinear estimation are contained in Gallant (1987) and Gallant and White (1988).

We consider least squares estimation of the parameters of model (1.1) and (1.2). The results are an extension of those in Fuller, Hasza and Goebel (1981) in that there are nonlinear restrictions on the parameters of the model. The results represent an extension of existing nonlinear results in that the model may contain some regressors, such as time trend, with sum of squares increasing at a rate different from that of other variables in the model.

2. Asymptotic properties of the least squares estimator. In this section we obtain the limiting distribution of the restricted least squares estimator as a function of the limiting distribution of the unrestricted least squares estimator. The vector $\mathbf{Z}_t$ may contain polynomials in time, a random walk or other variables whose sums of squares increase at a rate other than n. Also if $|m_1| \geq 1$ the sums of squares of Y_{t-1} will increase at a rate other than n. In such situations, it is generally necessary to transform the $\mathbf{Z}$-vectors to obtain a nondegenerate limiting distribution for the unrestricted least squares estimator.

Let a sequence of nonsingular transformations of the vectors $\mathbf{Z}_t$ be defined by the sequence of, possibly random, matrices $\mathbf{A}_n$. Let

$$\mathbf{X}_{tn} = \mathbf{Z}_t \mathbf{A}'_n. \tag{2.1}$$

Using the transformation $\mathbf{A}_n$, the model given by (1.1) and (1.2) can be written as

$$Y_t = \mathbf{X}_{tn}\boldsymbol{\theta}_n + e_t, \tag{2.2}$$

$$f_i(\mathbf{A}'_n\boldsymbol{\theta}_n) = \mathbf{0}, \qquad i = 1, 2, \dots, r, \tag{2.3}$$

where $\boldsymbol{\theta}'_n = \boldsymbol{\gamma}'\mathbf{A}_n^{-1}$ is a $1 \times k$ vector, $\mathbf{X}_{tn}$ is a $1 \times k$ vector and $k = p + q$. The matrix $\mathbf{A}_n$ will depend on the nature of the explanatory variables and on the form of the problem. Generally, $\mathbf{A}_n$ is not unique.

Fuller, Hasza and Goebel (1981) defined a transformation for the model with $\mathbf{Z}_t$ containing different types of explanatory variables and for various values of $|m_1|$ of characteristic equation (1.3). Related transformations have been utilized by Fuller (1976), Section 8.5, Hasza (1977), Sims, Stock and Watson (1990) and Chan and Wei (1988).

The Fuller, Hasza and Goebel transformation accomplishes several things. First, the sample correlation matrix of $\mathbf{X}_{tn}$ will have a nonsingular limit in a number of cases in which the correlation matrix of the $\mathbf{Z}_t$ does not. Second, the transformation isolates the largest roots of the characteristic equation as the last elements in the transformed parameter vector. The limiting distribution of the estimator of the largest root will sometimes differ from the limiting distribution of the remaining estimators when the largest root is greater than or equal to 1. In many cases the remaining elements of the vector have a normal distribution in the limit. The transformation is illustrated in the example at the end of this section.

We now introduce the assumptions used in our derivation. We require the functions defining the restrictions to satisfy the usual smoothness conditions.

ASSUMPTION 1. The functions $f_i(\boldsymbol{\gamma})$, $i = 1, 2, \ldots, r$, are continuous and twice differentiable in a region about $\boldsymbol{\gamma}^0$, where $\boldsymbol{\gamma}^0 = (\gamma_1^0, \gamma_2^0, \ldots, \gamma_k^0)'$ is the true value of the parameter vector.

The derivatives of the restrictions play an important role in the limiting distribution. Let $\mathbf{g}_n(\boldsymbol{\theta}) = [g_{1n}(\boldsymbol{\theta}), \ldots, g_{rn}(\boldsymbol{\theta})]'$, where $g_{in}(\boldsymbol{\theta}) = f_i(\mathbf{A}'_n \boldsymbol{\theta})$, $i = 1, 2, \ldots, r$, and $\boldsymbol{\theta}_n^{0\prime} = \boldsymbol{\gamma}^{0\prime}\mathbf{A}_n^{-1}$. Let

$$\mathbf{D}_n(\boldsymbol{\theta}) = \frac{\partial \mathbf{g}_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \tag{2.4}$$

be the $r \times k$ matrix of partial derivatives, where the ijth entry is the partial derivative of $g_{in}(\boldsymbol{\theta})$ with respect to θ_j. If there is a redundancy among the restrictions (1.2), then the rank of the matrix $\mathbf{D}_n(\boldsymbol{\theta}_n^0)$ will be less than r. Possible redundancies are removed by assumption.

ASSUMPTION 2. The matrix $\mathbf{D}_n(\boldsymbol{\theta}_n^0)$ is of rank r with probability 1.

Let the ordinary unrestricted least squares estimator of $\boldsymbol{\theta}_n$ be

$$\hat{\boldsymbol{\theta}}_n = \left(\sum_{t=1}^{n} \mathbf{X}'_{tn}\mathbf{X}_{tn} \right)^{-1} \sum_{t=1}^{n} \mathbf{X}'_{tn} Y_t. \tag{2.5}$$

The estimator defined in (2.5) minimizes the quantity

$$Q_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} (Y_t - \mathbf{X}_{tn}\boldsymbol{\theta})^2. \tag{2.6}$$

A number of papers containing results on the limiting distribution of the least squares estimator were cited in Section 1. Because different theorems correspond to different situations, the regularity conditions for the theorems are different. To cover a wide range of models, we simply assume that the properly normalized unrestricted least squares estimator has a limiting distribution.

ASSUMPTION 3. The matrix $\Sigma_{t=1}^{n}\mathbf{X}'_{tn}\mathbf{X}_{tn}$ is positive definite with probability 1 for $n > k$. There exists a sequence of diagonal matrices $\{\mathbf{H}_n = \operatorname{diag}(h_{11n}, \dots, h_{kkn})\}$ such that (i) $h_{iin} \to_P \infty$ as $n \to \infty$ for each i, (ii) $\mathbf{H}_n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0)$ converges in distribution to a nondegenerate random vector and (iii) the roots of $\mathbf{B}_n$ and $\mathbf{B}_n^{-1}$ are order 1 in probability, where

$$\mathbf{B}_n = \mathbf{H}_n^{-1/2} \sum_{t=1}^{n} \mathbf{X}'_{tn}\mathbf{X}_{tn}\mathbf{H}_n^{-1/2}.$$

In many applications $\mathbf{H}_n$ is chosen to be $\operatorname{diag}\{\Sigma_{t=1}^{n}\mathbf{X}'_{tn}\mathbf{X}_{tn}\}$. Then the matrix $\mathbf{B}_n$ of the assumption is in the form of a correlation matrix.

Let $\tilde{\boldsymbol{\theta}}_n$ denote the least squares estimator obtained by minimizing the sum of squared errors given in (2.6) subject to the restrictions (2.3). Then $\tilde{\boldsymbol{\theta}}_n$ is a value minimizing the Lagrangian

$$Q_n(\boldsymbol{\theta}) + 2 \sum_{j=1}^{r} \lambda_{jn} g_{jn}(\boldsymbol{\theta}). \tag{2.7}$$

We first prove the consistency of the nonlinear least squares estimator $\bar{\boldsymbol{\theta}}_n$ defined in (2.7). Consistency follows from the fact that the order of the error in $\tilde{\boldsymbol{\theta}}_n$ is no greater than the order of the error in $\hat{\boldsymbol{\theta}}_n$, where $\hat{\boldsymbol{\theta}}_n$ is the unrestricted least squares estimator.

LEMMA 1. *Let Assumptions* 1, 2 *and* 3 *hold. Then*

$$\mathbf{H}_n^{1/2}\left(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right) = O_p(1).$$

PROOF. Note that

$$\left|\mathbf{H}_n^{1/2}\left(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n\right)\right|^2 \le \phi_{kn}^{-1}\left(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n\right)'\mathbf{H}_n^{1/2}\mathbf{B}_n\mathbf{H}_n^{1/2}\left(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n\right),$$

where ϕ_{kn} is the smallest root of $\mathbf{B}_n$. The inequality also holds if $\tilde{\boldsymbol{\theta}}_n$ is replaced with $\boldsymbol{\theta}_n^0$. Because

$$Q_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} \left(Y_t - \mathbf{X}_{tn}\hat{\boldsymbol{\theta}}_n\right)^2 + \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)'\mathbf{H}_n^{1/2}\mathbf{B}_n\mathbf{H}_n^{1/2}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right),$$

because $\boldsymbol{\theta}_n^0$ satisfies the restriction and because $\tilde{\boldsymbol{\theta}}_n$ minimizes $Q_n(\boldsymbol{\theta})$ subject to the restrictions (2.3), we have

$$\left(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n\right)'\mathbf{H}_n^{1/2}\mathbf{B}_n\mathbf{H}_n^{1/2}\left(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n\right) \le \left(\boldsymbol{\theta}_n^0 - \hat{\boldsymbol{\theta}}_n\right)'\mathbf{H}_n^{1/2}\mathbf{B}_n\mathbf{H}_n^{1/2}\left(\boldsymbol{\theta}_n^0 - \hat{\boldsymbol{\theta}}_n\right).$$

Hence,

$$\left|\mathbf{H}_n^{1/2}\left(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right)\right|^2 \le 2\phi_{kn}^{-1}\phi_{1n}\left|\mathbf{H}_n^{1/2}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right)\right|^2.$$

By Assumption 3, $\phi_{1n} = O_p(1)$, $\phi_{kn}^{-1} = O_p(1)$ and $\mathbf{H}_n^{1/2}(\boldsymbol{\theta}_n^0 - \hat{\boldsymbol{\theta}}_n) = O_p(1)$. Therefore, $\mathbf{H}_n^{1/2}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0) = O_p(1)$. □

The system of equations associated with the Lagrangian (2.7) can be written as

$$\left(\sum_{t=1}^{n} \mathbf{X}'_{tn}\mathbf{X}_{tn}\right)\hat{\boldsymbol{\theta}} + \mathbf{D}'_n(\boldsymbol{\theta})\boldsymbol{\lambda}_n = \sum_{t=1}^{n} \mathbf{X}'_{tn}Y_t, \qquad \mathbf{g}_n(\boldsymbol{\theta}) = \mathbf{0}, \tag{2.8}$$

where $\boldsymbol{\lambda}_n = (\lambda_{1n}, \dots, \lambda_{rn})'$ is the vector of multipliers. Expanding each component of $\mathbf{g}_n(\boldsymbol{\theta})$ in a first-order Taylor expansion around the true value $\boldsymbol{\theta}_n^0$ and using the facts that $\mathbf{g}_n(\boldsymbol{\theta}_n^0) = \mathbf{0}$ and that $\tilde{\boldsymbol{\theta}}_n$ is a solution for (2.8), we have

$$\left(\sum_{t=1}^{n} \mathbf{X}'_{tn}\mathbf{X}_{tn}\right)\left(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right) + \mathbf{D}'_n\left(\tilde{\boldsymbol{\theta}}_n\right)\boldsymbol{\lambda}_n = \sum_{t=1}^{n} \mathbf{X}'_{tn}e_t, \qquad \mathbf{D}_n\left(\overset{*}{\boldsymbol{\theta}}_n\right)\left[\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right] = 0, \tag{2.9}$$

where $\mathbf{D}_n(\overset{*}{\boldsymbol{\theta}}_n) = \overset{*}{\mathbf{D}}_n = (\overset{*}{\mathbf{d}}_{1n}, \dots, \overset{*}{\mathbf{d}}_{rn})$, $\overset{*}{d}_{ijn}$ is the partial derivative of $g_{in}(\boldsymbol{\theta})$ with respect to θ_j evaluated at $\boldsymbol{\theta} = \overset{*}{\boldsymbol{\theta}}_{in}$ and $\overset{*}{\boldsymbol{\theta}}_{in}$, $i = 1, 2, \dots, r$, are points on the line segment joining $\boldsymbol{\theta}_n^0$ and $\tilde{\boldsymbol{\theta}}_n$.

We normalize the system (2.9) so that we obtain the error in the unrestricted least squares estimator $\hat{\boldsymbol{\theta}}_n$ on the right-hand side and so that the system can be solved for the error in the estimator $\tilde{\boldsymbol{\theta}}_n$. To this end we rewrite (2.9) as

$$\begin{bmatrix} \mathbf{B}_n & \tilde{\mathbf{R}}'_n \\ \overset{*}{\mathbf{R}}_n & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{H}_n^{1/2}\left(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right) \\ \mathbf{G}_n^{-1/2}\boldsymbol{\lambda}_n \end{bmatrix} = \begin{bmatrix} \mathbf{H}_n^{-1/2}\sum_{t=1}^{n}\mathbf{X}'_{tn}e_t \\ \mathbf{0} \end{bmatrix}, \tag{2.10}$$

where $\mathbf{G}_n = [\mathbf{D}_{0n}(\Sigma\mathbf{X}'_{tn}\mathbf{X}_{tn})^{-1}\mathbf{D}'_{0n}]^{-1}$, $\tilde{\mathbf{R}}_n = \mathbf{G}_n^{1/2}\tilde{\mathbf{D}}_n\mathbf{H}_n^{-1/2}$, $\overset{*}{\mathbf{R}}_n = \mathbf{G}_n^{1/2}\overset{*}{\mathbf{D}}_n\mathbf{H}_n^{-1/2}$, $\mathbf{D}_{0n} = \mathbf{D}_n(\boldsymbol{\theta}^0)$, $\tilde{\mathbf{D}}_n = \mathbf{D}_n(\tilde{\boldsymbol{\theta}}_n)$, $\overset{*}{\mathbf{D}}_n$ is defined in (2.9), $\mathbf{D}_n(\boldsymbol{\theta})$ is defined in (2.4), $\mathbf{B}_n$ is defined in Assumption 3 and $\mathbf{G}_n^{1/2}$ is the symmetric positive-definite square root of $\mathbf{G}_n$.

Because the elements of $\mathbf{H}_n$ are not necessarily of the same order in probability, some additional restrictions on the matrix $\mathbf{D}_n(\boldsymbol{\theta})$ are required. Let

$$\boldsymbol{\Omega}_n = \Big[\mathbf{G}_n^{1/2}\left(\hat{\mathbf{D}}_n - \mathbf{D}_{0n}\right)\mathbf{H}_n^{-1/2}, \mathbf{G}_n^{1/2}\left(\tilde{\mathbf{D}}_n - \mathbf{D}_{0n}\right)\mathbf{H}_n^{-1/2}, \mathbf{G}_n^{1/2}\left(\overset{*}{\mathbf{D}}_n - \mathbf{D}_{0n}\right)\mathbf{H}_n^{-1/2}\Big], \tag{2.11}$$

where $\hat{\mathbf{D}}_n = \mathbf{D}_n(\hat{\boldsymbol{\theta}}_n)$ and $\mathbf{D}_{0n}$, $\tilde{\mathbf{D}}_n$ and $\overset{*}{\mathbf{D}}_n$ are the matrices defined in (2.10).

ASSUMPTION 4. The sequence $\{\boldsymbol{\Omega}_n\}$ defined in (2.11) converges to $\mathbf{0}$ in probability.

If the regressors $\{\psi_{ti}: i = 1, 2, \dots, q\}$ and the Y-process are such that the transformation $\mathbf{A}_n = \mathbf{I}$ produces a nondegenerate limiting distribution for $\hat{\boldsymbol{\theta}}_n$, Assumption 4 will hold. Assumption 4 is verified for the regression model with first-order autoregressive errors in the example at the end of this section.

To derive the limiting distribution of the estimator $\tilde{\boldsymbol{\theta}}_n$ in terms of the limiting distribution of $\hat{\boldsymbol{\theta}}_n$, the following lemma is required.

LEMMA 2. *Let*

$$[\hat{\mathbf{T}}_n, \mathbf{T}_n] = \left[\begin{pmatrix} \mathbf{B}_n & \tilde{\mathbf{R}}'_n \\ \overset{*}{\mathbf{R}}_n & \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{B}_n & \mathbf{R}'_n \\ \mathbf{R}_n & \mathbf{0} \end{pmatrix}\right], \tag{2.12}$$

where $\mathbf{R}_n = \mathbf{G}_n^{1/2}\mathbf{D}_{0n}\mathbf{H}_n^{-1/2}$ *and* $\mathbf{B}_n$, $\tilde{\mathbf{R}}_n$ *and* $\overset{*}{\mathbf{R}}_n$ *are defined in* (2.10). *Under Assumptions* 1, 2, 3 *and* 4,

(i) $\mathbf{T}_n^{-1} = O_p(1)$ *and* $\hat{\mathbf{T}}_n^{-1} = O_p(1)$,
(ii) $\mathbf{T}_n^{-1} - \hat{\mathbf{T}}_n^{-1} \to_P \mathbf{0}$.

PROOF. We have

$$\mathbf{T}_n^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{B}_n^{-1}\mathbf{R}'_n \\ \mathbf{0} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{B}_n^{-1} & \mathbf{0} \\ \mathbf{0} & -(\mathbf{R}_n\mathbf{B}_n^{-1}\mathbf{R}'_n) \end{pmatrix} \times \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{R}_n\mathbf{B}_n^{-1} & \mathbf{I} \end{pmatrix}. \tag{2.13}$$

By the definitions, $\mathbf{R}_n\mathbf{B}_n^{-1}\mathbf{R}'_n = \mathbf{I}$. Also by Assumption 3, $\mathbf{B}_n^{-1} = O_p(1)$.

Since $\mathbf{R}_n\mathbf{B}_n^{-1/2}\mathbf{B}_n^{-1/2}\mathbf{R}'_n = \mathbf{I}$, the sum of squares of the elements of $\mathbf{R}_n\mathbf{B}_n^{-1/2}$ is bounded above by r, where r is the number of restrictions. Therefore, $\mathbf{R}_n\mathbf{B}_n^{-1/2} = O_p(1)$ and $\mathbf{R}_n\mathbf{B}_n^{-1} = \mathbf{R}_n\mathbf{B}_n^{-1/2}\mathbf{B}_n^{-1/2} = O_p(1)$. Hence, from (2.13), it follows that $\mathbf{T}_n^{-1} = O_p(1)$. To prove that $\hat{\mathbf{T}}_n^{-1} = O_p(1)$, we observe that

$$\operatorname{Prob}\left\{\left(\|\mathbf{T}_n^{-1}\|\right)^{-1} > \|\mathbf{T}_n - \hat{\mathbf{T}}_n\|\right\} \to 1 \tag{2.14}$$

as $n \to \infty$ because $\|\mathbf{T}_n - \hat{\mathbf{T}}_n\| \to 0$ in probability and $\|\mathbf{T}_n^{-1}\| = O_p(1)$, where $\|\mathbf{B}\|$ denotes a matrix norm of $\mathbf{B}$. By (4.24) of Kato (1966), page 31, $\|\mathbf{T}_n^{-1}\|^{-1} > \|\mathbf{T}_n - \hat{\mathbf{T}}_n\|$ implies

$$\|\hat{\mathbf{T}}_n^{-1}\| \le \left(1 - \|\mathbf{T}_n - \hat{\mathbf{T}}_n\|\,\|\mathbf{T}_n^{-1}\|\right)^{-1}\left(\|\mathbf{T}_n^{-1}\|\right),$$

and we conclude that $\hat{\mathbf{T}}_n^{-1} = O_p(1)$. See Lemma 3 of Section 3.2 of Nagaraj (1986). Finally, (ii) follows from (i) by the identity,

$$\hat{\mathbf{T}}_n^{-1} - \mathbf{T}_n^{-1} = \hat{\mathbf{T}}_n^{-1}(\mathbf{T}_n - \hat{\mathbf{T}}_n)\mathbf{T}_n^{-1}. \qquad \square$$

We now give the limiting distribution for $\tilde{\boldsymbol{\theta}}_n$.

THEOREM 1. *Under Assumptions* 1, 2, 3 *and* 4,

(i) $\mathbf{H}_n^{1/2}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0)$ *has the same limiting distribution as*

$$\mathbf{M}_{11n}\mathbf{H}_n^{1/2}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right). \tag{2.15}$$

(ii) $\mathbf{G}_n^{-1/2}\boldsymbol{\lambda}_n$ *has the same limiting distribution as*

$$\mathbf{G}_n^{-1/2}\left(\mathbf{D}_{0n}\mathbf{H}_n^{-1/2}\mathbf{B}_n^{-1}\mathbf{H}_n^{-1/2}\mathbf{D}'_{0n}\right)^{-1}\mathbf{D}_{0n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right), \tag{2.16}$$

where

$$\mathbf{M}_{11n} = \left[\mathbf{I} - \mathbf{B}_n^{-1}\mathbf{H}_n^{-1/2}\mathbf{D}'_{0n}\left(\mathbf{D}_{0n}\mathbf{H}_n^{-1/2}\mathbf{B}_n^{-1}\mathbf{H}_n^{-1/2}\mathbf{D}'_{0n}\right)^{-1}\mathbf{D}_{0n}\mathbf{H}_n^{-1/2}\right], \tag{2.17}$$

where $\mathbf{B}_n$ *and* $\mathbf{H}_n$ *are defined in Assumption* 3.

Proof. By Lemma 1, $(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0) \to_P \mathbf{0}$. Using (2.10), Assumption 3 and that $\boldsymbol{\gamma}^0$ is in the interior of the parameter space, we obtain

$$\begin{bmatrix} \mathbf{H}_n^{1/2}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0) \\ \mathbf{G}_n^{-1/2}\boldsymbol{\lambda}_n \end{bmatrix} = \begin{bmatrix} \mathbf{B}_n & \mathbf{R}'_n \\ \mathbf{R}_n & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{H}_n^{-1/2}\Sigma\mathbf{X}'_{tn}e_t \\ 0 \end{bmatrix} + o_p(1) \tag{2.18}$$

by Lemma 2. Using the expression for the inverse of a partitioned matrix, we have result (i).

Also from (2.18),

$$\mathbf{G}_n^{-1/2}\boldsymbol{\lambda}_n = \left(\mathbf{R}_n\mathbf{B}_n^{-1}\mathbf{R}'_n\right)^{-1}\mathbf{R}_n\mathbf{H}_n^{1/2}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right) + o_p(1) \tag{2.19}$$

and result (ii) follows. □

Because $\mathbf{H}_n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0) = O_P(1)$, the residual mean square obtained from the nonlinear regression is a consistent estimator of the variance of the process $\{e_t\}$. We state the result as a corollary.

Corollary 1. *Let* $s^2 = (n - k + r)^{-1}\sum_{t=1}^{n}(Y_t - \mathbf{X}_{tn}\tilde{\boldsymbol{\theta}}_n)^2$, *where* $\tilde{\boldsymbol{\theta}}_n$ *is the estimator defined by* (2.10). *Let* $\{e_t\}$ *be a sequence of independently and identically distributed* $(0, \sigma^2)$ *random variables. Then under Assumptions* 1 *through* 5, s^2 *converges to* σ^2 *in probability.*

A common test for the validity of the restrictions imposed on the regression coefficients is the ratio of the difference between the residual sum of squares for the full model and the residual sum of squares for the restricted model to the residual mean square for the full model. The limiting distribution of this test is given in Corollary 2.

Corollary 2. *Let* $Q(\hat{\boldsymbol{\theta}}_n)$ *denote the residual sum of squares from the unrestricted model and let* $Q_n(\tilde{\boldsymbol{\theta}}_n)$ *denote the residual sum of squares from the restricted model. Let* $Q(\boldsymbol{\theta}_n^{\ddagger})$ *denote the residual sum of squares from the model satisfying* d *of the* r *restrictions, where* $0 \le d < r$. *Let* $\mathbf{D}'_{0n} = (\mathbf{D}'_{0n1}, \mathbf{D}'_{0n2})$, *where* $\mathbf{D}_{0n1}$ *contains the restrictions associated with* $\boldsymbol{\theta}_n^{\ddagger}$. *Let*

$$\boldsymbol{\Phi}_{nnij} = \mathbf{D}_{0ni}\mathbf{H}_n^{-1/2}\mathbf{B}_n^{-1}\mathbf{H}_n^{-1/2}\mathbf{D}'_{0nj}$$

and let $\boldsymbol{\Phi}_{nn}$ *be the* $r \times r$ *matrix composed of the four submatrices* $\boldsymbol{\Phi}_{nn11}$, $\boldsymbol{\Phi}_{nn12}$, $\boldsymbol{\phi}_{nn21}$ *and* $\boldsymbol{\Phi}_{nn22}$. *Then under the assumptions of Theorem* 1,

$$\hat{\sigma}^{-2}\left[Q(\tilde{\boldsymbol{\theta}}_n) - Q(\hat{\boldsymbol{\theta}}_n)\right] - \sigma^{-2}\boldsymbol{\lambda}'_n\mathbf{G}_n^{-1}\boldsymbol{\lambda}_n = o_p(1), \tag{2.20}$$

$$\begin{aligned} &\hat{\sigma}^{-2}\left[Q(\tilde{\boldsymbol{\theta}}_n) - Q(\boldsymbol{\theta}_n^{\ddagger})\right] \\ &\quad = \sigma^{-2}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right)'\mathbf{A}'_{n2}\mathbf{V}_{nn22}^{-1}\mathbf{A}_{n2}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0\right) + o_p(1), \end{aligned} \tag{2.21}$$

where $\mathbf{A}'_{n2} = \mathbf{D}'_{0n2} - \mathbf{D}'_{0n1}\mathbf{\Phi}^{-1}_{nn11}\mathbf{\Phi}_{nn12}$,

$$\mathbf{V}_{nn22} = \mathbf{\Phi}_{nn22} - \mathbf{\Phi}_{nn21}\mathbf{\Phi}^{-1}_{nn11}\mathbf{\Phi}_{nn12},$$

and $\hat{\sigma}^2 = (n-k)^{-1}Q(\hat{\boldsymbol{\theta}}_n)$ *is the residual mean square for the full model. It is understood that* $\mathbf{V}_{nn22} = \mathbf{\Phi}_{nn22} = \mathbf{\Phi}_{nn}$ *if* $d = 0$.

PROOF. We have

$$\sum_{t=1}^{n}\left(Y_t - \mathbf{X}_{tn}\tilde{\boldsymbol{\theta}}_n\right)^2 = \sum_{t=1}^{n}\left(Y_t - \mathbf{X}_{tn}\hat{\boldsymbol{\theta}}_n\right)^2 + \left(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n\right)\sum_{t=1}^{n}\mathbf{X}'_{tn}\mathbf{X}_{tn}\left(\tilde{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n\right),$$

where we have used the fact that the residuals from the unrestricted regression are orthogonal to the columns of the model matrix. Then, using (2.18),

$$(2.22)\qquad Q\left(\tilde{\boldsymbol{\theta}}_n\right) - Q\left(\hat{\boldsymbol{\theta}}_n\right) = \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^0_n\right)'\mathbf{D}'_{0n}\mathbf{\Phi}^{-1}_{nn}\mathbf{D}_{0n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^0_n\right) + o_p(1).$$

Also from (2.19),

$$\mathbf{G}_n^{-1/2}\boldsymbol{\lambda}_n = \left(\mathbf{D}_{0n}\mathbf{H}_n^{-1/2}\mathbf{B}_n^{-1}\mathbf{H}_n^{-1/2}\mathbf{D}'_{0n}\right)^{-1/2}\mathbf{D}_{0n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^0_n\right) + o_p(1).$$

Because the mean square error from the unrestricted model is a consistent estimator of σ^2, (2.20) and (2.21) for $d = 0$ follow.

The sum of squares on the right of (2.22) can be partitioned into that associated with the first d restrictions and that due to the remaining $(r-d)$ restrictions after adjusting for the first d restrictions. This partition gives result (2.21) for $d > 0$. □

For many applications, $\mathbf{B}_n$ is converging to a constant matrix and $\mathbf{H}_n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^0_n)$ is converging in distribution to a normal vector. In these cases, the statistic of (2.21) has a chi-square distribution with $r - d$ degrees of freedom in the limit.

EXAMPLE 1. To illustrate some of the ideas associated with our results, we consider the regression model with autocorrelated errors. To treat a nonstandard case, we let the explanatory variable be a random walk. The model is

$$(2.23)\qquad Y_t = \beta Z_t + U_t \quad\text{and}\quad U_t = \rho U_{t-1} + e_t,$$

where $Z_t = \sum_{i=1}^{t} d_i$, $\{d_t\}$ is a sequence of independent identically distributed $(0, \sigma_{dd})$ random variables, $\{e_t\}$ is a sequence of independent identically distributed $(0, \sigma_{ee})$ random variables and $\{d_t\}$ is independent of $\{e_t\}$. The model can be written as

$$(2.24)\qquad Y_t = \gamma_1 Z_{t-1} + \gamma_2 Z_t + \gamma_3 Y_{t-1} + e_t,$$

where $\gamma_2 = \beta$, $\gamma_3 = \rho$, $\gamma_1 = -\rho\beta$ and the restriction is $\gamma_1 + \gamma_2\gamma_3 = 0$. Because the correlation of Z_t and Z_{t-1} is tending to 1, it is necessary to transform the variables to obtain a nondegenerate limiting distribution for the ordinary least squares estimator. For purposes of defining the limiting distribution, let

$$\mathbf{X}_{tn} = \left(x_{t1}, x_{t2}, \hat{U}_{t-1}\right) = \left(Z_{t-1}, Z_t - Z_{t-1}, Y_{t-1} - \delta_n Z_{t-1}\right),$$

where $\delta_n = (\Sigma_{t=1}^n Z_{t-1}^2)^{-1}\Sigma_{t=1}^n Z_{t-1}Y_{t-1}$. The transformed model becomes

$$Y_t = \mathbf{X}_{tn}\boldsymbol{\theta}_n + e_t, \tag{2.25}$$

where $\theta_{n1} = \beta - \rho\beta + \delta_n\rho$, $\theta_{n2} = \beta$, $\theta_{n3} = \rho$ and the restriction is $\theta_{n1} - \theta_{n2} - \theta_{n3}\delta_n + \theta_{n2}\theta_{n3} = 0$.

If $|\rho| < 1$, it can be shown that Theorem 1 of Fuller, Hasza and Goebel (1981) is applicable. It follows that the limiting distribution of $\mathbf{H}_n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0)$, where $\mathbf{H}_n = \operatorname{diag}(\Sigma_{t=1}^n x_{t1}^2, \Sigma_{t=1}^n x_{t2}^2, \Sigma_{t=1}^n \hat{U}_{t-1}^2)$, is that of a normal $(\mathbf{0}, \mathbf{I}\sigma_{ee})$ vector, where $\boldsymbol{\theta}_n = \mathbf{A}_n^{-1\prime}\boldsymbol{\gamma}$. The use of the δ_n part of the transformation is not required to obtain the limiting distribution when $|\rho| < 1$.

If $\rho = 1$, the first two elements of $\mathbf{H}_n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_n^0)$ are converging to $N(0, \sigma_{ee})$ random variables. Using the transformation given in Dickey and Fuller (1979), the third element, divided by the square root of the residual mean square,

$$\frac{\left(\Sigma \hat{U}_{t-1}^2\right)^{1/2}\left(\hat{\theta}_{n3} - 1\right)}{s} \to_L \frac{\frac{1}{2}\left(T_U^2 - 1\right) - \Gamma_{ZZ}^{-1}\Gamma_{ZU}\Upsilon_{ZU}}{\left(\Gamma_{UU} - \Gamma_{ZZ}^{-1}\Gamma_{ZU}^2\right)^{1/2}}, \tag{2.26}$$

where

$$(\Gamma_{ZZ}, \Gamma_{UU}, \Gamma_{ZU}, T_U) = \sum_{i=1}^{\infty}\left[\zeta_i^2\left(a_i^2, b_i^2, a_i b_i\right), 2^{1/2}\zeta_i b_i\right],$$

$$\Upsilon_{ZU} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} 2\left[\zeta_j + \zeta_i\right]^{-1}\zeta_i^2\zeta_j a_i b_j,$$

$(a_i, b_i) \sim NI(\mathbf{0}, \mathbf{I})$, $\zeta_i = (-1)^{i+1}2[(2i - 1)\pi]^{-1}$ and s^2 is the regression residual mean square. See Phillips (1986) for integral representations of Γ_{ZZ}, Γ_{UU}, Γ_{ZU} and Υ_{ZU}.

The vector of partial derivatives of the restriction with respect to $\boldsymbol{\theta}_n$ is $\mathbf{D}_n(\boldsymbol{\theta}) = (1, \theta_3 - 1, \theta_2 - \delta_n)$. If $\rho = \theta_{n3} = 1$, then $\mathbf{D}_n(\boldsymbol{\theta}_n^0) = (1, 0, \theta_{n2}^0 - \delta_n)$ and

$$G_n \doteq \left(\sum_{t=1}^n Z_{t-1}^2 \sum_{t=1}^n \hat{U}_{t-1}^2\right)\left(\sum_{t=1}^n U_{t-1}^2\right)^{-1} = O_p(n^2).$$

The error in $\hat{\theta}_{n3}$ is $O_p(n^{-1})$ and the error in $\hat{\theta}_{n2}$ is $O_p(n^{-1/2})$. The three diagonal elements of $\mathbf{H}_n^{-1}$ are $O_p(n^{-2})$, $O_p(n^{-1})$ and $O_p(n^{-2})$ in that order. Hence, $G_n^{1/2}(\hat{\mathbf{D}}_n - \mathbf{D}_{0n})\mathbf{H}_n^{-1/2} = O_p(n^{-1/2})$ and Assumption 4 is satisfied.

Because $\mathbf{B}_n \to \mathbf{I}$ in probability, the error in the normalized restricted least squares estimator of $\boldsymbol{\theta}_n^0$ is approximated by

$$\left[\mathbf{H}_n^{1/2} - \mathbf{H}_n^{-1/2}\mathbf{D}'_{0n}\left(\mathbf{D}_{0n}\mathbf{H}_n^{-1}\mathbf{D}'_{0n}\right)^{-1}\mathbf{D}_{0n}\right]\mathbf{L}_n,$$

where $\mathbf{L}_n = \mathbf{H}_n^{-1}\Sigma_{t=1}^n \mathbf{X}'_{tn} e_t$. The last two elements of $\tilde{\boldsymbol{\theta}}_n$ are the restricted least squares estimators of β and ρ. The large sample distribution of the restricted estimator of β is the same as that of the unrestricted least squares coefficient for θ_{2n} in model (2.25). If the coefficient estimating ρ is normalized with $(\Sigma U_{t-1}^2)^{1/2}$, the limiting distribution is that of $\hat{\tau}$ described by Dickey and Fuller (1979). Because the sum of squares of the derivatives of model (2.25) with respect to ρ divided by n^2 is converging to the limit of $n^{-2}\Sigma_{t=1}^n U_{t-1}^2$, the limiting distribution of the regression "t-statistic" for ρ is that of $\hat{\tau}$.

From (2.16), the limiting distribution of the test statistic for the restriction computed by analogy to the usual regression F-statistic has the limiting distribution

$$\frac{\left[\Upsilon_{ZU} - \frac{1}{2}\Gamma_{UU}^{-1}\Gamma_{ZU}\left(T_U^2 - 1\right)\right]^2}{\Gamma_{ZZ} - \Gamma_{UU}^{-1}\Gamma_{ZU}^2}, \tag{2.27}$$

where the variables are defined in (2.26).

In summary, the estimator of ρ in the unrestricted problem has a distribution that depends on the nature of the regressor variable when $\rho = 1$. In the restricted problem, the distribution of the estimator of ρ is free of that dependence. In the restricted problem the standardized estimator of β has a limiting normal distribution and is uncorrelated with the estimator of ρ in the limit. The test of the restriction has a rather complicated distribution when $\rho = 1$. Nagaraj and Fuller (1989) give some Monte Carlo results for the model.

REMARK. We have discussed the estimation problem for a single equation. The results extend immediately to the parameters of a system of equations. Let the multivariate problem be written as

$$\mathbf{Y}_t = \mathbf{X}_{tn}\boldsymbol{\theta}_n + \mathbf{e}_t, \tag{2.28}$$

where $\mathbf{Y}_t$ is a c-dimensional row vector, $\mathbf{X}_{tn}$ is a k-dimensional row vector and $\boldsymbol{\theta}_n$ is a $k \times c$ matrix of parameters. Let

$$Q_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} (\mathbf{Y}_t - \mathbf{X}_{tn}\boldsymbol{\theta}_n)\mathbf{S}_n^{-1}(\mathbf{Y}_t - \mathbf{X}_{tn}\boldsymbol{\theta}_n)'$$

be the quadratic form to be minimized, where $\mathbf{S}_n$ is a fixed matrix or a consistent estimator of a fixed matrix. Then if the ordinary least squares estimator of $\boldsymbol{\theta}_n$ has a limiting distribution, Theorem 1 applies for nonlinear restrictions on vec $\boldsymbol{\theta}$.

The vector autoregression

$$\mathbf{Y}_t = \sum_{i=1}^{p} \mathbf{Y}_{t-i}\mathbf{A}_i + \mathbf{e}_t, \tag{2.29}$$

where the $\mathbf{A}_i$ are $c \times c$ parameter matrices is a specific example of a multivariate model. Let $\hat{\boldsymbol{\theta}}$ be the vector of least squares estimators of the elements of $\mathbf{A}_i$, $i = 1, 2, \dots, c$. Under certain conditions, the properly normalized elements of $\hat{\boldsymbol{\theta}}$ have a limiting distribution; see, for example, Hannan (1970), Phillips and Durlauf (1986), Sims, Stock and Watson (1990) and Chan and Wei (1988). The limiting distributions and our results can be used to obtain the limiting distributions of restricted estimators and of tests of the restriction. The restriction that there are m, where $m \leq c$, unit roots of the characteristic equation is an example. Johansen (1988), Phillips (1988), Fountis and Dickey (1989) and Reinsel and Ahn (1989) have given results for the vector autoregressive model with unit roots.

Acknowledgment. We thank the referees for useful comments and suggestions.

REFERENCES

ANDERSON, T. W. and RUBIN, H. (1950). The asymptotic properties of estimates of the parameters in a single equation in a complete system of stochastic equations. *Ann. Math. Statist.* **21** 570–572.

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

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** 427–431.

FOUNTIS, N. G. and DICKEY, D. A. (1989). Testing for a unit root nonstationarity in multivariate autoregressive time series. *Ann. Statist.* **17** 419–428.

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

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

GALLANT, A. R. (1987). *Nonlinear Statistical Models*. Wiley, New York.

GALLANT, A. R. and WHITE, H. (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Blackwell, New York.

HANNAN, E. J. (1970). *Multiple Time Series*. Wiley, New York.

HANNAN, E. J. and NICHOLLS, D. F. (1972). The estimation of mixed regression, autoregression, moving average and distributed lag models. *Econometrica* **40** 529–548.

HASZA, D. P. (1977). Estimation in nonstationary time series. Ph.D. dissertation, Iowa State Univ.

JOHANSEN, S. (1988). Statistical analysis of cointegration vectors. *J. Econom. Dynamics Control* **12** 231–254.

KATO, T. (1966). *Perturbation Theory for Linear Operators*. Springer, New York.

MANN, H. B. and WALD, A. (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica* **11** 173–220.

NAGARAJ, N. K. (1986). Estimation of stochastic difference equations with nonlinear restrictions. Ph.D. dissertation, Iowa State Univ.

NAGARAJ, N. K. and FULLER, W. A. (1989). Least squares estimation of the linear model with autoregressive errors. Unpublished manuscript.

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

PHILLIPS, P. C. B. (1988). Optimal inference in cointegrated systems. Cowles Foundation Discussion Paper 866, Yale Univ.

PHILLIPS, P. C. B. and DURLAUF, S. N. (1986). Multivariate time series with integrated variables. *Rev. Econom. Stud.* **53** 473–496.

REINSEL, G. C. and AHN, S. K. (1989). Asymptotic distribution of the likelihood ratio test for cointegration in the nonstationary vector AR model. Unpublished manuscript.

RUBIN, H. (1950). Consistency of maximum likelihood estimates in the explosive case. In *Statistical Inference in Dynamic Economic Models* (T. C. Koopmans, ed.) 356–364. Wiley, New York.

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

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

WHITE, H. and DOMOWITZ, I. (1984). Nonlinear regression with dependent observations. *Econometrica* **52** 143–161.

WOOLDRIDGE, J. M. (1986). Asymptotic properties of econometric estimators. Ph.D. dissertation, Univ. California, San Diego.

WU, C. F. (1981). Asymptotic theory of nonlinear least squares estimation. *Ann. Statist.* **9** 501–513.

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF MARYLAND
BALTIMORE COUNTY CAMPUS
BALTIMORE, MARYLAND 21228

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011