ESTIMATION OF ERROR COMPONENTS MODEL WITH ARMA (p,q) TIME COMPONENT - AN EXACT GLS APPROACH

by

Lung-Fei Lee

Discussion Paper No. 78-104, November 1978

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

September 1978

Estimation of Error Components Model With ARMA (p,q) Time Component - An Exact GLS Approach

by

Lung-Fei Lee*

1. Introduction

The use of Error Components Models in pooling cross-sections and time series data has attracted the attention of many econometricians, see Balestra and Nerlove [1966], Wallace and Hussain [1969], Maddala [1971] among others. One of the widely used models is the classical error components regression model,

$$y_{it} = \alpha + x_{it}\beta + u_i + v_t + w_{it} \qquad \begin{matrix} i=1,\ldots,N \\ t=1,\ldots,T \end{matrix} \tag{1.1}$$

where x_{it} are $1 \times K$ vectors of exogenous variables. The regression errors in (1.1) are composed of three independent error components -- the component u_i is associated with the cross sectional units, v_t associated with time and a residual component w_{it} in both the time and cross sectional dimensions. The classical model that has been analysed assumes that all the error components have zero means, finite variances and, in addition,

*This research was supported in part by a Summer Research Grant from the Graduate School, University of Minnesota, Minneapolis, and by Grant SOC-78-07304 from the National Science Foundation.

(1) $E(u_i u_{i'}) = 0$ for $i \neq i'$,

(2) $E(v_t v_{t'}) = 0$ for $t \neq t'$ and

(3) $E(w_{it} w_{i't'}) = 0$ for $i \neq i'$ or $t \neq t'$.

When the cross sectional units are well separated, assumption (1) is well-justified. Assumption (3) is also well-justified if w_{it} consists mainly of white noises. However, assumption (2) can hardly be justified as v_t consists of unobservable time variables and it is well-known, Cochrane and Orcutt [1949], that time series variables are serially correlated.

In this paper, we will relax the restricted assumption (2) to allow general serially correlation for v_t. We consider model (1.1) with assumption (2) replaced by assumption v_t is ARMA (p, q). Related models can be found in Tiao and Ali [1971] and Revankar [1977]. Tiao and Ali [1971] analyzed from a Bayesian viewpoint a two error components model. Their model is much simplified than ours; only two components are allowed and v_t is ARMA (1,1). Revankar [1977] analyzed a three error components model with simple AR(1) time component v_t. Revankar's approach required Cochrane and Orcutt type transformations twice and hence efficiency is lost. This is so, especially for large number of cross sectional units but relatively short time period. In addition, it is not clear how his approach can be generalized for more complicated processes other than AR(1).

In this paper, we will derive the exact and operational GLS estimates and ML estimates for our model.

2. Construction of exact GLS estimator.

2.1. Model Decomposition:

Specifically, we consider the error component model (1.1) with ARMA (p, q) time component v_t,

$$\phi(B)v_t = \Theta(B)\varepsilon_t \tag{2.1}$$

where B is a back-shift operator. $\phi(B)$ and $\Theta(B)$ are polynomials in B with degrees p and q respectively, given by

$$\phi(B) = 1 - \phi_1 B,\ldots,\phi_p B^p$$

$$\Theta(B) = 1 - \Theta_1 B,\ldots,\Theta_q B^q .$$

ε_t's are uncorrelated random variables with mean zero and finite variance σ_ε^2. All the roots of $\phi(B)$ lie outside the unit circle to guarantee v_t to be stationary. Furthermore, to avoid model redundancy, $\phi(B)$ and $\Theta(B)$ do not have any common root.

When $p = 0$ and $q = 0$, model (1.1) with (2.1) is the classical error components model analyzed by Wallace and Hussain [1969], Nerlove [1971] and Maddala [1971]. The covariance matrix Σ for the disturbances in this classical model is $\sigma_w^2 I_N \otimes I_T - \sigma_u^2 I_N \otimes \ell_T \ell_T' - \sigma_v^2 \ell_N \ell_N' \otimes I_T$. Its inverse has been derived in Wallace and Hussain [1969], Henderson [1971] and Nerlove [1971]. Even in this simplest case, its deviation is never a trivial problem. In our general model, direct approach does not seem possible. In the following paragraphs, we adapt another approach which was motivated by Maddala's interpretation [1971] of GLS estimator in the classical error components model and our previous analysis on the issue between fixed effects and random effects model, Lee [1978], and Tiao and Ali [1971]. In this section, we assume that all the parameters σ_u^2, σ_w^2, ϕ, Θ and σ^2 are known.

Let $\bar{y}_{\cdot t} = \frac{1}{N}\Sigma_{i=1}^{N} y_{it}$, $\bar{y}_{i\cdot} = \frac{1}{T}\Sigma_{t=1}^{T} y_{it}$, $\bar{y} = \frac{1}{NT}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T} y_{it}$.; similarly $\bar{x}_{\cdot t}$, $\bar{x}_{i\cdot}$, $\bar{x}$, $\bar{w}_{\cdot t}$, $\bar{w}_{i\cdot}$, $\bar{w}$ etc. are defined. Without lost of generality, assume $\bar{x} = 0$. The samples y_{it} can be decomposed into three subsamples, $\bar{y}_{\cdot t}$, $\bar{y}_{i\cdot} - \bar{y}$ and $y_{it} - \bar{y}_{\cdot t} - \bar{y}_{i\cdot} + \bar{y}$. It follows from Lee [1978], equation (1.1) can be decomposed as

$$\bar{y}_{\cdot t} = \alpha + \bar{x}_{\cdot t}\beta + v_t + \bar{u} + \bar{w}_{\cdot t} \qquad t=1,\ldots,T \tag{2.2}$$

$$\bar{y}_{i\cdot} - \bar{y}_{1\cdot} = (\bar{x}_{i\cdot} - \bar{x}_{1\cdot})\beta + u_i - u_1 + \bar{w}_{i\cdot} - \bar{w}_1 \qquad i=2,\ldots,N \tag{2.3}$$

$$y_{it}^d = x_{it}^d\beta + w_{it} - \bar{w}_{\cdot t} - \bar{w}_{i\cdot} + \bar{w} \qquad i=1,\ldots,N;\ t=1,\ldots,T \tag{2.4}$$

where $y_{it}^d = y_{it} - \bar{y}_{\cdot t} - \bar{y}_{i\cdot} + \bar{y}$ and $x_{it}^d = x_{it} - \bar{x}_{\cdot t} - \bar{x}_{i\cdot}$; or equivalently, (2.3) can be rewritten as

$$\bar{y}_{i\cdot} - \bar{y} = (\bar{x}_{i\cdot} - \bar{x})\beta + u_i - \bar{u} + \bar{w}_{i\cdot} - \bar{w} \qquad i=1,\ldots,N \tag{2.3$'$}$$

since $\bar{y}_{i\cdot} - \bar{y} = \bar{y}_{i\cdot} - \bar{y}_{1\cdot} - \frac{1}{N}\Sigma_{i=1}^{N}(\bar{y}_{i\cdot} - \bar{y}_{1\cdot})$. Since the error components in (1.1) are independent, the errors in (2.2) - (2.4) are uncorrelated across equations. To estimate our model (1.1), it is convenient to analyze equations (2.2)-(2.4).

2.2 Within Estimator and Between Groups Estimator

Let $Q = I_N \otimes I_T - \frac{1}{T} I_N \otimes \ell_T\ell_T' - \frac{1}{N}\ell_N\ell_N' \otimes I_T + \frac{1}{NT}\ell_N\ell_N' \otimes \ell_T\ell_T'$ be the covariance transformation where ℓ_T, ℓ_N are $T \times 1$ and $N \times 1$ column vectors of unity. Coefficient β in (2.4) can be estimated as

$$\hat{\beta}_w = (X'QX)^{-1}X'Qy \tag{2.5}$$

where $X' = (x'_{11},\ldots,x'_{1T},\ldots,x'_{N1},\ldots,x'_{NT})$ and $y' = (y_{11},\ldots,y_{1T},\ldots y_{N1},\ldots,y_{NT})$. $\hat{\beta}_w$ is known as the covariance or within estimate. The variance matrix of $\hat{\beta}_w$ is

$$E(\hat{\beta}_w-\beta)(\hat{\beta}_w-\beta)' = \sigma_w^2(X'QX)^{-1} \tag{2.6}$$

Let ξ_2 be a $(N-1) \times 1$ column vector with i^{th} component $\xi_{2i} = u_{i+1} - u_1 + \bar{w}_{i+1.} - \bar{w}_{1.}$. It follows

$$V_2 = E(\xi_2\xi_2') = (\sigma_u^2 + \frac{\sigma_w^2}{T})(I_{N-1} + \ell_{N-1}\ell'_{N-1})$$

and its inverse $V_2^{-1} = (\sigma_u^2 + \frac{\sigma_w^2}{T})^{-1}\ (I_{N-1} - \frac{1}{N}\ \ell_{N-1}\ell'_{N-1})$. Equation (2.3) can be estimated by GLS procedure which gives, after simplification,

$$\hat{\beta}_I = (\Sigma_{i=1}^N \bar{x}'_{i.}\bar{x}_{i.})^{-1}(\Sigma_{i=1}^N \bar{x}'_{i.}\bar{y}_{i.}) \tag{2.7}$$

with covariance matrix,

$$E(\hat{\beta}_I - \beta)(\hat{\beta}_I - \beta)' = (\sigma_u^2 + \frac{\sigma_w^2}{T})(\Sigma_{i=1}^N \bar{x}'_{i.}\bar{x}_{i.})^{-1} \tag{2.8}$$

The estimator $\hat{\beta}_I$ is the between group estimator.

2.3 Analysis of Between Time Equation

Let ξ_3 be a $T \times 1$ column vector with t^{th} component $\xi_{3t} = v_t + \bar{u} + \bar{w}_{.t}$. It follows

$$E(\xi_3\xi_3') = \frac{\sigma_w^2}{N}\ I_T + \frac{\sigma_u^2}{N}\ \ell_T\ell_T' + \Omega_v$$

where Ω_v is the $T \times T$ covariance matrix of time component v_t. Deviation of the inverse matrix of $E(\xi_3\xi_3')$ is the main problem for this equation. Numerical solution is computationally inefficient when T is large.

It has been pointed out in T. W. Anderson [1971], Tiao and Ali [1971], O. D. Anderson [1975, 1977] and most recently Ansley et al [1977] that the sum of two independent moving average process of order q_1 and q_2 is itself a moving average process of order $q \leqq \max (q_1, q_2)$.

It is straightforward to generalize this lemma to the summation of two independent ARMA process,

<u>Lemma</u>: Let $\{v_t\}$ and $\{u_t\}$ be two uncorrelated ARMA (p_1, q_1) and ARMA (p_2, q_2) stochastic process;

$$\phi_{p_1}(B)v_t = \theta_{q_1}(B)\varepsilon_t, \quad \psi_{p_2}(B)u_t = \lambda_{q_2}(B)\zeta_t$$

where all the roots of $\phi_{p_1}(B)$ and $\psi_{p_2}(B)$ lie outside the unit circle. Then the summation $z_t = u_t + v_t$ has a ARMA $(p_1 + p_2, K)$ representation, $\phi^*_{p_1+p_2}(B)z_t = \Theta^*_K(B)\xi_t$ where $K \leq \max (p_1+q_2, p_2+q_1)$ and $\phi^*_{p_1+p_2}(B) = \phi_{p_1}(B)\psi_{p_2}(B)$.

Furthermore, if $\{u_t\}$ and $\{v_t\}$ are independent normal processes, then $\{z_t\}$ is normal and its associated random shocks have independent normal distributions.

In our equation (2.2), v_t and $\bar{w}_{\cdot t}$ are independent ARMA (p,q) and ARMA (0,0) process. Hence $z_t = v_t + \bar{w}_{\cdot t}$ is a ARMA (p,K) with $K \leq \max(p,q)$ for any fixed N. Explicitly, we have

$$\phi_p(B) z_t = \theta_q(B)\varepsilon_t + \phi_p(B)\bar{w}_{\cdot t}$$

$$= \theta_K^*(B)\eta_t \tag{2.9}$$

where $\theta_K^*(B) = \eta_t - \theta_1^*\eta_{t-1} - \cdots - \theta_K^*\eta_{t-K}$. The parameters θ_i^*, $i=1,\ldots,K$ and σ_η^2 satisfy the following nonlinear equations,

$$\Sigma_{j=0}^{K}\theta_j^{*2}\sigma_\eta^2 = \Sigma_{j=0}^{q}\theta_j^2\sigma_\varepsilon^2 + \Sigma_{j=0}^{p}\phi_j^2\ \sigma_w^2/N, \quad \theta_0^* = \theta_0 = \phi_0 = 1 \tag{2.10}$$

$$\Sigma_{j=0}^{K-i}\ \theta_j^*\theta_{j+i}^*/\Sigma_{j=0}^{K}\theta_j^{*2} = \lambda\Sigma_{j=0}^{q-i}\theta_j\theta_{j+i}/\Sigma_{j=0}^{q}\theta_j^2 +$$

$$(1-\lambda)\ \Sigma_{j=0}^{p-i}\phi_j\phi_{j+i}/\Sigma_{j=0}^{p}\phi_j^2 \qquad i=1,\ldots,K \tag{2.11}$$

where $\lambda = \Sigma_{j=0}^{q}\theta_j^2\sigma_\varepsilon^2/(\Sigma_{j=0}^{q}\theta_j^2\sigma_\varepsilon^2 + \Sigma_{j=0}^{p}\phi_j^2\ \frac{\sigma_w^2}{N})$. The left hand expressions in (2.10) and (2.11) are respectively the variance and autocorrelations of the MA(K) process in (2.9). The lemma guarantees that solutions $\sigma_\eta^2, \theta_1^*, \ldots, \theta_K^*$ in (2.10) and (2.11) exist. It is known, Ansley et al [1977] that solutions may not be unique. However, as far as the covariance matrix of z_t in (2.9) is concerned, any solution will be immaterial since all the solutions determine the same covariance matrix. Equation (2.2) becomes

$$\bar{y}_{\cdot t} = \alpha + \bar{x}_{\cdot t}\beta + \bar{u} + z_t \qquad t=1,\ldots,T \tag{2.2)'$$

Let Ω_z be the $T \times T$ covariance matrix of z. The covariance matrix of $\xi_{3t} = \bar{u} + z_t$ is $V_3 = E(\xi_3\xi_3') = \frac{\sigma_u^2}{N} \ell_T\ell_T' + \Omega_z$. The inverse matrix Ω_z^{-1} has been derived explicitly in Shaman [1973] for the stationary general ARMA process with invertible MA polynomial. The expression Ω_z^{-1} derived from Shaman's technique seems quite complicated and no expression is given for the determinant of the matrix. Shaman's approach may not be applicable to our case since it is not known whether the ARMA representation of z_t has invertible MA polynomial even though $\theta(B)$ in (2.1) is invertible. The approach in Newbold [1974] who generalized Box and Jenkins [1970] backward forecasting technique which requires only stationarity is useful for our purpose. The deviation for Ω_z^{-1} is outlined in the Appendix. The only complication involved in this approach is to decompose a $p+q$ symmetric matrix. However, this can be easily solved by numerical technique as p and q are usually small in practice. It follows, the inverse of V_3 is

$$V_3^{-1} = \Omega_z^{-1} - \Omega_z^{-1}\ell_T\ell_T'\Omega_z^{-1}/(\frac{N}{\sigma_u^2} + \ell_T'\Omega_z^{-1}\ell_T)$$

Equation (2.2)' can be estimated by GLS procedure which gives

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}_b = \begin{bmatrix} \ell_T'V_3^{-1}\ell_T & \ell_T'V_3^{-1}\bar{X}^{bt} \\ \bar{X}^{bt'}V_3^{-1}\ell_T & \bar{X}^{bt'}V_3^{-1}\bar{X}^{bt} \end{bmatrix}^{-1} \begin{bmatrix} \ell_T'V_3^{-1}\bar{y}^{bt} \\ \bar{X}^{bt'}V_3^{-1}\bar{y}^{bt} \end{bmatrix} \tag{2.12}$$

where

$$\bar{X}^{bt'} = (\bar{x}'_{.1},\ldots,\bar{x}'_{.T}), \qquad \bar{y}^{bt'} = (\bar{y}_{.1},\ldots,\bar{y}_{.T})$$

Equation (2.12) implies

$$\hat{\beta}_b = \left(\bar{X}^{bt'}V_3^{-1}\bar{X}^{bt} - \frac{\bar{X}^{bt'}V_3^{-1}\ell_T\ell_T'V_3^{-1}\bar{X}^{bt}}{\ell_T V_3^{-1}\ell_T}\right)^{-1}\left(\bar{X}^{bt'}V_3^{-1}\bar{y}^{bt} - \frac{\bar{X}^{bt'}V_3^{-1}\ell_T\ell_T'V_3^{-1}\bar{y}^{bt}}{\ell_T'V_3^{-1}\ell_T}\right) \tag{2.13}$$

with covariance matrix,

$$E(\hat{\beta}_b-\beta)(\hat{\beta}_b-\beta)' = \left[\bar{X}^{bt'}\left(V_3^{-1} - \frac{V_3^{-1}\ell_T\ell_T'V_3^{-1}}{\ell_T'V_3^{-1}\ell_T}\right)\bar{X}^{bt}\right]^{-1} \tag{2.14}$$

The estimate in (2.13) is similar to the between time estimate in the classical error components model. This estimate utilizes all the sample information between time in our model and it is the proper between time estimate of (1.1) in the presence of ARMA time component.

2.4 Exact GLS Estimator

The three estimates in (2.5), (2.7) and (2.13) can be weighted by their precision matrices, i.e., inverses of the covariance matrices in (2.6), (2.8), and (2.14). Equivalently, this pooled estimate can be derived from the mixed estimation procedure, Theil and Goldberger [1961], applied to the three equations (2.2) - (2.4). The pooled estimate of β in (1.1) is

$$\hat{\beta}_p = [\frac{1}{\sigma_w^2} X'QX + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'}\bar{X}^{in} + \bar{X}^{bt'}(V_3^{-1} - \frac{V_3^{-1}\ell_T\ell_T'V_3^{-1}}{\ell_T'V_3^{-1}\ell_T})\bar{X}^{bt}]^{-1} \cdot$$

$$[\frac{1}{\sigma_w^2} X'QX\hat{\beta}_w + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'}\bar{X}^{in}\ \hat{\beta}_I + \bar{X}^{bt'}(V_3 - \frac{V_3^{-1}\ell_T\ell_T'V_3^{-1}}{\ell_T'V_3^{-1}\ell_T})\bar{X}^{bt}\hat{\beta}_b]$$

$$= [\frac{1}{\sigma_w^2} X'QX + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'}\bar{X}^{in} + \bar{X}^{bt'}(V_3^{-1} - \frac{V_3^{-1}\ell_T\ell_T'V_3^{-1}}{\ell_T'V_3^{-1}\ell_T})\bar{X}^{bt}]^{-1} \cdot$$

$$[\frac{1}{\sigma_w^2} X'Qy + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'}\bar{y}^{in} + \bar{X}^{bt'}(V_3^{-1} - \frac{V_3^{-1}\ell_T\ell_T'V_3^{-1}}{\ell_T'V_3^{-1}\ell_T})\bar{y}^{bt}] \tag{2.15}$$

with covariance matrix,

$$E(\hat{\beta}_p - \beta)(\hat{\beta}_p - \beta)' = [\frac{1}{\sigma_w^2}X'QX + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'}\bar{X}^{in} + \bar{X}^{bt'}(V_3^{-1} - \frac{V_3^{-1}\ell_T\ell_T'V_3^{-1}}{\ell_T'V_3^{-1}\ell_T}) \cdot$$

$$\cdot\ \bar{X}^{bt}]^{-1} . \tag{2.16}$$

The pooled estimate $\hat{\beta}_p$ is a weighted average of within, between group and between time estimates. Hence $\hat{\beta}_p$ utilizes all the within, between group and between time information in the sample and the structure in (2.1). It remains to check that $\hat{\beta}_p$ is the exact GLS estimate derived from (1.1) and (2.1).

Instead of pooling the three estimates in (2.5), (2.7) and (2.13), one can pool the estimates in (2.5), (2.7) and (2.12). The pooled estimate for β as well as α is

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}_p = \begin{bmatrix} \ell_T' V_3^{-1} \ell_T & \ell_T V_3^{-1} \bar{X}^{bt} \\ \bar{X}^{bt'} V_3^{-1} \ell_T & \frac{1}{\sigma_w^2} X'QX + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'} \bar{X}^{in} + \bar{X}^{bt'} V_3^{-1} \bar{X}^{bt} \end{bmatrix}^{-1} \cdot$$

$$\begin{bmatrix} \ell_T' V_3^{-1} \bar{y}^{bt} \\ \frac{1}{\sigma_w^2} X'Qy + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'} \bar{y}^{in} + \bar{X}^{bt'} V_3^{-1} \bar{y}^{bt} \end{bmatrix} \tag{2.17}$$

with covariance matrix

$$V\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}_p = \begin{bmatrix} \ell_T' V_3^{-1} \ell_T & \ell_T V_3^{-1} \bar{X}^{bt} \\ \bar{X}^{bt'} V_3^{-1} \ell_T & \frac{1}{\sigma_w^2} X'QX + \frac{1}{\sigma_u^2 + \frac{\sigma_w^2}{T}} \bar{X}^{in'} \bar{X}^{in} + \bar{X}^{bt'} V_3^{-1} \bar{X}^{bt} \end{bmatrix}^{-1} \tag{2.18}$$

It is obvious that the estimate of β in (2.15) is the same estimate in (2.17). Consider the matrix,

$$A = \frac{1}{\sigma_w^2} Q + \frac{1}{T\sigma_u^2 + \sigma_w^2} \frac{1}{T} I_N \otimes \ell_T \ell_T' - \frac{1}{T\sigma_u^2 + \sigma_w^2} \frac{1}{NT} \ell_N \ell_N' \otimes \ell_T \ell_T' +$$

$$\frac{1}{N^2} (\ell_N \otimes I_T) V_3^{-1} (\ell_N' \otimes I_T)$$

$$= \frac{1}{\sigma_w^2} I_N \otimes I_T - \frac{\sigma_u^2}{\sigma_w^2(\sigma_w^2 + T\sigma_u^2)} I_N \otimes \ell_T \ell_T' + \frac{1}{N} \frac{\sigma_u^2}{\sigma_w^2(\sigma_w^2 + T\sigma_u^2)} \ell_N \ell_N' \otimes \ell_T \ell_T'$$

$$+ \frac{1}{N^2} (\ell_N \otimes I_T) \Omega_z^{-1} (\ell_N' \otimes I_T) - \frac{1}{N^2} (\ell_N \otimes I_T) \Omega_z^{-1} \ell_T \ell_T' \Omega_z^{-1} (\ell_N' \otimes I_T) / (\frac{N}{\sigma_u^2} + \ell_T' \Omega_z^{-1} \ell_T)$$

$$- \frac{1}{N\sigma_w^2} \ell_N \ell_N' \otimes I_T \tag{2.19}$$

As $\bar{X}^{in} = \frac{1}{T}(I_N \otimes \ell_T')X$, $\bar{y}^{in} = \frac{1}{T}(I_N \otimes \ell_T')y$, $\bar{X}^{bt} = \frac{1}{N}(\ell_N' \otimes I_T)X$,

$\bar{y}^{bt} = \frac{1}{N}(\ell_N' \otimes I_T)y$ and $(\ell_N \ell_N' \otimes \ell_T \ell_T')X = 0$, one can easily check

$$\ell_{NT}' A \ell_{NT} = \ell_T' V_3^{-1} \ell_T, \quad \ell_{NT}' AX = \ell_T' V_3^{-1} \bar{X}^{bt}, \quad \ell_{NT}' Ay = \ell_T' V_3^{-1} \bar{y}^{bt}$$

$$X'AX = \frac{1}{\sigma_w^2} X'QX + \frac{T}{\sigma_w^2 + T\sigma_u^2} \bar{X}^{in\prime} \bar{X}^{in} + \bar{X}^{bt\prime} V_3^{-1} \bar{X}^{bt}$$

and $$X'Ay = \frac{1}{\sigma_w^2} X'Qy + \frac{T}{\sigma_w^2 + T\sigma_u^2} \bar{X}^{in\prime} \bar{y}^{in} + \bar{X}^{bt\prime} V_3^{-1} \bar{y}^{bt} .$$

It follows that (2.17) and (2.18) can be rewritten as

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}_P = \begin{bmatrix} \ell_{NT}' A \ell_{NT} & \ell_{NT}' AX \\ X' A \ell_{NT} & X'AX \end{bmatrix}^{-1} \begin{bmatrix} \ell_{NT}' Ay \\ X'Ay \end{bmatrix} \tag{2.17$'$}$$

$$V\begin{pmatrix}\hat{\alpha}\\ \beta\end{pmatrix}_p = \begin{bmatrix}\ell'_{NT}A\ell_{NT} & \ell'_{NT}AX\\ X'A\ell_{NT} & X'AX\end{bmatrix}^{-1} \tag{2.18)'}$$

From (2.17)' and (2.18)' $A = \Sigma^{-1}$, this is so, since

$$\begin{pmatrix}\hat{\alpha}\\ \beta\end{pmatrix}_p - \begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{bmatrix}\ell'_{NT}A\ell_{NT} & \ell'_{NT}AX\\ X'A\ell_{NT} & X'AX\end{bmatrix}^{-1}\begin{bmatrix}\ell'_{NT}A & \zeta\\ X'A & \zeta\end{bmatrix}$$

where ζ is the $NT \times 1$ vector of disturbances $\zeta_{it} = u_i + v_t + w_{it}$. It implies

$$V\begin{pmatrix}\hat{\alpha}\\ \beta\end{pmatrix}_p = \left(\begin{bmatrix}\ell'_{NT}\\ X'\end{bmatrix}A[\ell_{NT} X]\right)^{-1}\begin{bmatrix}\ell'_{NT}\\ X'\end{bmatrix} A\Sigma A\ [\ell_{NT}\ X]\left(\begin{bmatrix}\ell'_{NT}\\ X'\end{bmatrix}A[\ell_N\ X]\right)^{-1} \tag{2.20}$$

It follows from (2.18)' and (2.20) that $A\Sigma A = A$ and $A = \Sigma^{-1}$. Hence we conclude that the estimates in (2.17) is the exact GLS estimate for our model (1.1) and (2.1). The inverse of the $NT \times NT$ covariance matrix of the disturbance ζ is derived in (2.19).

From (2.15) and (2.16), it is obvious that $\underset{\substack{N\to\infty\\ T\to\infty}}{\text{plim}}\ \sqrt{NT}\ (\hat{\beta}_p - \beta)$ $= \underset{\substack{N\to\infty\\ T\to\infty}}{\text{plim}}\ \sqrt{NT}\ (\hat{\beta}_w - \beta)$ i.e., asymptotically, the GLS estimate and the within estimate are equivalent. This conclusion is similar to the result in the classical error components model, Wallace and Hussain [1969].

3. Estimation with Unknown Covariance Parameters.

The parameters of disturbances in (1.1) and (2.1) are assumed to be known in the last section to simplify our arguments. That is rarely the case in practice. To estimate all parameters in the model, one of the asymptotic efficient procedures is the maximum likelihood approach. To derive the likelihood function and the determinant of Σ, all the error components u_i, ε_t and w_{it} are assumed to be normal.

As the samples can be decomposed into independent subsamples $\bar{y}^{bt}$, $\bar{y}^{in} - \bar{y}$ and Qy, the likelihood function of y can be written as the product,

$$L(\alpha, \beta, \phi, \theta, \sigma_u^2, \sigma_\varepsilon^2, \sigma_w^2|y, X) \tag{3.1}$$

$$= L_1(\beta, \sigma_w^2|Qy, QX)L_2(\beta,\sigma_u^2,\sigma_w^2|\bar{y}^{in}-\bar{y}, \bar{X}^{in} - \bar{X})L_3(\alpha,\beta,\phi,\theta,\sigma_u^2,\sigma_\varepsilon^2,\sigma_w^2|\bar{y}^{bt},\bar{X}^{bt})$$

where L_1, L_2 and L_3 are the likelihood functions of equations (2.4), (2.3)' and (2.2) respectively, or equivalently (2.4), (2.3) and (2.2). The determinants of the covariance matrices in (2.3) and (2.2)' can easily be derived;

$$|V_2| = N(\sigma_u^2 + \frac{\sigma_w^2}{T})^{N-1}, \qquad |V_3| = |\Omega_z|(1 + \frac{\sigma_u^2}{N} \ell_T'\Omega_z^{-1}\ell_T)$$

The determinant $|\Omega_z|$ can be derived from Newbold's technique and is available in the appendix. It follows

$$L_2(\beta,\sigma_u^2,\sigma_w^2|\bar{y}^{in}-\bar{y},\ \bar{X}^{in}-\bar{X}) = \frac{1}{(2\Pi)^{\frac{N-1}{2}} N^{1/2}(\sigma_u^2 + \frac{\sigma_w^2}{T})^{\frac{N-1}{2}}} \exp$$

$$\{-\frac{T}{2(\sigma_u^2 T + \sigma_w^2)} (\tilde{y}-\tilde{X}\beta)'(I_{N-1} - \frac{1}{N}\ell_{N-1}\ell_{N-1}')(\tilde{y}-\tilde{X}\beta)\}$$

$$= \frac{1}{(2\Pi)^{\frac{N-1}{2}} N^{1/2}(\sigma_u^2 + \frac{\sigma_w^2}{T})^{\frac{N-1}{2}}} \exp \ \{-\frac{T}{2(\sigma_u^2 T + \sigma_w^2)} \cdot$$

$$(\bar{y}^{in} - \bar{y}\ell_N - \bar{X}^{in}\beta)'(\bar{y}^{in} - \bar{y}\ell_N - \bar{X}^{in}\beta)\} \tag{3.2}$$

where $\tilde{y} = \begin{bmatrix} \bar{y}_{2\cdot} - \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{N\cdot} - \bar{y}_{1\cdot} \end{bmatrix}$, $\tilde{X} = \begin{bmatrix} \bar{x}_{2\cdot} - \bar{x}_{1\cdot} \\ \vdots \\ \bar{x}_{N\cdot} - \bar{x}_{1\cdot} \end{bmatrix}$

$$L_3(\alpha,\beta,\phi,\theta,\sigma_u^2,\ \sigma_\varepsilon^2,\ \sigma_w^2|\bar{y}^{bt},\ \bar{X}^{bt})$$

$$= \frac{1}{(2\Pi)^{T/2}|\Omega_z|^{1/2}(1 + \frac{\sigma_u^2}{N} \ell_T'\Omega_z^{-1}\ell_T)^{1/2}} \exp\ \{-\frac{1}{2(\frac{N}{\sigma_u^2} + \ell_T'\Omega_z^{-1}\ell_T)} \cdot$$

$$(\bar{y}^{bt}-\alpha\ell_T-\bar{X}^{bt}\beta)'((\frac{N}{\sigma_u^2}+\ell_T'\Omega_z^{-1}\ell_T)\Omega_z^{-1}-\Omega_z^{-1}\ell_T\ell_T'\Omega_z^{-1})(\bar{y}^{bt} - \alpha\ell_T - \bar{X}^{bt}\beta)\} \tag{3.3}$$

The derivation of $L_1(\beta,\ \sigma_w^2|Qy,\ QX)$ is as follows. Conditional on $u_1,\ldots,u_N$ and $v_1,\ldots,v_T$, the conditional likelihood function for model (1.1) is

$$L(\alpha,\beta,\sigma_w^2|y,\ X,\ u,\ v) = \frac{1}{(2\Pi)^{\frac{NT}{2}}\sigma_w^{NT}} \exp\{-\frac{1}{2\sigma_w^2}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}\cdot$$

$$(y_{it} - \alpha - x_{it}\beta - u_i - v_t)^2\}\ .$$

It is easy to check that

$$\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}(y_{it} - \alpha - x_{it}\beta - u_i - v_t)^2$$

$$= \Sigma_{i=1}^{N}\Sigma_{t=1}^{T}(y_{it}^d - x_{it}^d\beta)^2 + N\ \Sigma_{t=1}^{T}(\bar{y}_{\cdot t} - \alpha - \bar{x}_{\cdot t}\beta - v_t - \bar{u})^2$$

$$+ T\ \Sigma_{i=1}^{N}(\bar{y}_{i.} - \bar{y} - \bar{x}_{i.}\beta - u_i + \bar{u})^2$$

and

$$\Sigma_{i=1}^{N}(\bar{y}_{i.} - \bar{y} - \bar{x}_{i.}\beta - u_i + \bar{u})^2$$

$$= (\tilde{y} - \bar{y}_{1.}\ell_{N-1} - \tilde{x}\beta + \bar{x}_{1.}\ell_{N-1}\beta - \tilde{u} + u_1\ell_{N-1})'(I_{N-1} - \frac{1}{N}\ell_{N-1}\ell_{N-1}')$$

$$(\tilde{y} - \bar{y}_{1.}\ell_{N-1} - \tilde{x}\beta + \bar{x}_{1.}\ell_{N-1}\beta - \tilde{u} + u_1\ell_{N-1})$$

where $\tilde{y}' = (\bar{y}_{2.},\ldots,\bar{y}_{N.})$, $\tilde{x}' = (\bar{x}_{2.}',\ldots,\bar{x}_{N.}')$ and $\tilde{u}' = (u_2,\ldots,u_N)$. As $|(\sigma_w^2/T)(I_{N-1} + \ell_{N-1}\ell_{N-1}')| = |(T/\sigma_w^2)(I_{N-1} - \frac{1}{N}\ell_{N-1}\ell_{N-1}')|^{-1} = N(\sigma_w^2/T)^{N-1}$, it follows

$$L(\alpha,\beta,\sigma_w^2|y,X,u,v)$$

$$= \frac{1}{(2\Pi\sigma_w^2)^{\frac{NT-N-T+1}{2}} N^{\frac{T-1}{2}} T^{\frac{N-1}{2}}} \exp\{-\frac{1}{2\sigma_w^2}(y-X\beta)'Q(y-X\beta)\}\cdot$$

$$\frac{1}{(2\Pi\sigma_w^2)^{T/2} N^{-T/2}} \exp\{-\frac{N}{2\sigma_w^2}\Sigma_{t=1}^T(\bar{y}_{\cdot t}-\alpha-\bar{x}_{\cdot t}\beta - v_t - \bar{u})^2\}\cdot$$

$$\frac{1}{(2\Pi\sigma_w^2)^{\frac{N-1}{2}} T^{-(N-1)/2} N^{1/2}} \exp\{-\frac{T}{2\sigma_w^2}\Sigma_{i=1}^N(\bar{y}_{i\cdot}-\bar{y}-\bar{x}_{i\cdot}\beta-u_i+\bar{u})^2\}$$

Hence

$$L_1(\alpha,\beta,\sigma_w^2|Qy,\ QX) = \frac{1}{(2\Pi\sigma_w^2)^{\frac{NT-N-T+1}{2}} N^{\frac{T-1}{2}} T^{\frac{N-1}{2}}} \exp$$

$$\{-\frac{1}{2\sigma_w^2}(y-X\beta)'Q(y-X\beta)\} \tag{3.4}$$

is the likelihood function for equation (2.4).

The likelihood function of y is

$$L(\alpha,\beta,\phi,\theta,\sigma_u^2,\sigma_\varepsilon^2,\sigma_w^2|y,X)$$

$$= \frac{1}{(2\Pi)^{\frac{NT}{2}} N^{\frac{T-1}{2}} \sigma_w^{NT-N-T+1} (\sigma_w^2+T\sigma_u^2)^{\frac{N-1}{2}} |\Omega_z|^{1/2} (N+\sigma_u^2\ell_T'\Omega_z^{-1}\ell_T)^{1/2}}\cdot$$

$$\exp\{-\frac{1}{2\sigma_w^2}[(y-X\beta)'Q(y-X\beta) + \frac{\sigma_w^2}{\frac{N}{(\sigma_u^2 + \ell_T'\Omega_z^{-1}\ell_T)}} (\bar{y}^{bt}-\alpha\ell_T - \bar{X}^{bt}\beta)'\cdot$$

$$((\frac{N}{\sigma_u^2} + \ell_T'\Omega_z^{-1}\ell_T)\Omega_z^{-1}-\Omega_z^{-1}\ell_T\ell_T'\Omega_z^{-1})(\bar{y}^{bt}-\alpha\ell_T-\bar{X}^{bt}\beta) +$$

$$\frac{T\sigma_w^2}{(\sigma_w^2+T\sigma_u^2)} (\bar{y}^{in} - \bar{y}\ell_N - \bar{X}^{in}\beta)'(\bar{y}^{in}-\bar{y}\ell_N-\bar{X}^{in}\beta)]\} \quad (3.1)'$$

As a joint product, the determinant of Σ is,

$$|\Sigma| = N^{T-1} \sigma_w^{2(NT-N-T+1)}(\sigma_w^2+T\sigma_u^2)^{N-1} (N+\sigma_u^2\ell_T'\Omega_z^{-1}\ell_T)|\Omega_z| \quad (3.5)$$

It is obvious from (3.1)' that GLS estimate of β will be the same as ML estimate if the parameters in Σ were known. When they are unknown, they can be jointly estimated by maximum likelihood method. The ML estimates can be derived from iterative optimization techniques applied to (3.1)'.

The likelihood function $L(\alpha,\beta,\phi,\theta,\sigma_u^2,\sigma_w^2,\sigma_\eta^2|y,X)$ in (3.1)' is maximized subject to the nonlinear constraints in (2.10) and (2.11). Equivalently, one can maximize the concentrated log likelihood function

$$L(\phi,\theta,\frac{\sigma_u^2}{\sigma_w^2},\frac{\sigma_\varepsilon^2}{\sigma_w^2}|y,X) = \text{const.} - \frac{1}{2}\ \ell n|\Sigma^*| - \frac{NT}{2}\ \ell n\ y'(\Sigma^{*-1}-\Sigma^{*-1}X \cdot$$
$$(X'\Sigma^{*-1}X)^{-1}X'\Sigma^{*-1})y \quad (3.6)$$

where $\Sigma^* = \frac{1}{\sigma_w^2}\Sigma$, subject to the nonlinear constraints:

$$\Sigma_{j=0}^K\ \theta_j^{*2}\ \sigma_\eta^2/\sigma_w^2 = \Sigma_{j=0}^q\theta_j^2\ \sigma_\varepsilon^2/\sigma_w^2 + \Sigma_{j=0}^p\ \phi_j^2/N, \quad \theta_0^* = \theta_0 = \phi_0 = 1 \quad (2.10)'$$

$$\Sigma_{j=0}^{K-i}\theta_j^*\ \theta_{j+i}^*/\Sigma_{j=0}^K\theta_j^{*2} = \mu\Sigma_{j=0}^{q-i}\theta_j\theta_{j+i}/\Sigma_{j=0}^q\theta_j^2 + (1-\mu)\Sigma_{j=0}^{p-i}\phi_j\phi_{j+i}/\Sigma_{j=0}^p\phi_j^2 \quad (2.11)'$$

$$i=1,\ldots,K$$

where $\mu = \Sigma_{j=0}^q\theta_j^2(\sigma_\varepsilon^2/\sigma_w^2)/(\Sigma_{j=0}^q\theta_j^2(\sigma_\varepsilon^2/\sigma_w^2) + \Sigma_{j=0}^p\phi_j^2/N)$, to obtain the maximum likelihood estimates $\tilde{\phi}, \tilde{\theta}$, $\widetilde{(\sigma_u^2/\sigma_w^2)}$ and $\widetilde{(\sigma_\varepsilon^2/\sigma_w^2)}$. The maximum likelihood estimates of α, β and σ_w^2 can then be computed as

$$\begin{bmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix} = \left(\begin{bmatrix} \ell'_{NT} \\ X' \end{bmatrix} \tilde{\Sigma}^{*-1} [\ell_{NT} \; X] \right)^{-1} \begin{bmatrix} \ell'_{NT} \\ X' \end{bmatrix} \tilde{\Sigma}^{*-1} y \tag{3.7}$$

$$\tilde{\sigma}_w^2 = \frac{1}{NT}(y - \tilde{\alpha}\ell_{NT} - X\tilde{\beta})'\tilde{\Sigma}^{*-1}(y - \tilde{\alpha}\ell_{NT} - X\tilde{\beta}) \tag{3.8}$$

For ARMA process associated with low degrees p and q, solutions $\theta_1^*, \ldots, \theta_n^*$ and σ_η^2 can be explicitly derived as functions of ϕ, θ, σ_ε^2 and σ_u^2 from (2.10) and (2.11). The likelihood function in (3.1)' or (3.6) can be maximized with unconstrained optimization techniques. Good initial consistent estimates may also be derived easily.

For example, if v_t in (2.1) is ARMA (1,0), K=1 and one set of solutions of σ_η^2 and θ_1^* is

$$\theta_1^* = \frac{1}{2\phi_1}\left(1 + \frac{N\sigma_\varepsilon^2}{\sigma_w^2} + \phi_1^2 - \left[\left(1 + \frac{N\sigma_\varepsilon^2}{\sigma_w^2} + \phi_1^2\right)^2 - 4\phi_1^2\right]^{1/2}\right) \tag{3.9}$$

$$\sigma_\eta^2 = \frac{\phi_1}{\theta_1^*}\frac{\sigma_w^2}{N} \tag{3.10}$$

For the ARMA (1, 0) process in (2.1), simple initial consistent parameters for Σ can be derived as follows. The estimates of variance components σ_w^2 and σ_u^2 can be derived as in Graybill [1961] or Wallace and Hussain [1969]. Let $(\hat{\alpha}_{LS}, \hat{\beta}_{LS})$ be the OLS estimates of (α, β) in (1.1) and let $\hat{\zeta}_{it} = y_{it} - \hat{\alpha}_{LS} - x_{it}\hat{\beta}_{LS}$ be the estimated residuals. σ_w^2 and σ_u^2 can be estimated as

$$\hat{\sigma}_w^2 = \frac{1}{(N-1)(T-1)} \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} (\hat{\zeta}_{it} - \hat{\bar{\zeta}}_{i.} - \hat{\bar{\zeta}}_{.t})^2 \tag{3.11}$$

$$\hat{\sigma}_u^2 = \frac{1}{N-1} \Sigma_{i=1}^{N} \hat{\bar{\zeta}}_{i.}^2 - \frac{1}{T} \hat{\sigma}_w^2 \tag{3.12}$$

where $\bar{\zeta}_{i.} = \frac{1}{T} \Sigma_{t=1}^{T} \zeta_{it}$ and $\bar{\zeta}_{.t} = \frac{1}{N} \Sigma_{i=1}^{N} \zeta_{it}$. Since ϕ_1 and σ_ε^2 satisfy the relations,

$$E(\zeta_{it}^2) = \sigma_u^2 + \sigma_v^2 + \sigma_w^2 \tag{3.13}$$

$$E(\zeta_{it}\zeta_{jt'}) = \delta_{ij}\sigma_u^2 + \phi_1^{|t-t'|}\sigma_v^2 \tag{3.14}$$

where $\delta_{ij} = 1$ if $i = j$; $\delta_{ij} = 0$, otherwise, σ_v^2 can be estimated as

$$\hat{\sigma}_v^2 = \frac{1}{NT} \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \hat{\zeta}_{it}^2 - \hat{\sigma}_u^2 - \hat{\sigma}_w^2 \tag{3.15}$$

and ϕ_1 can be estimated as

$$\hat{\phi}_1 = \frac{1}{\hat{\sigma}_v^2 N} \left(\frac{1}{N(T-1)} \Sigma_{t=1}^{T-1} \Sigma_{i=1}^{N} \Sigma_{j=1}^{N} \zeta_{it}\, \zeta_{jt-1} - \hat{\sigma}_u^2\right) \tag{3.16}$$

These initial estimates can be used to start the iteration, or if only the coefficient β is concerned, they can be used to construct two step GLS estimate. To test statistical hypothesis, Wald test or Maximum likelihood ratio test can be used.

Appendix: Inverse and Determinant of the Covariance Matrix of ARMA (p,q).

Newbold [1974] has derived explicitly the exact likelihood function for general ARMA (p,q) process. Implicitly, the inverse and determinant are derived.

Consider the ARMA (p, q) process

$$v_t - \phi_1 v_{t-1},\dots,\phi_p v_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1},\dots,\theta_q \varepsilon_{t-q} \qquad t=1,\dots,T. \tag{A.1}$$

Define a $p + q$ column vector e,

$$e_{1-p-j} = \varepsilon_{1-j} \qquad j=1,\dots,q$$

$$e_{1-j} = v_{1-j} \qquad j=1,\dots,p$$

Then (A.1) can be rewritten as

$$\binom{e}{\varepsilon} = \binom{0}{L} v + \binom{I_{p+q}}{K} e \tag{A.2}$$

where the matrices L and K, defined from the above equation and (A.1), are $T \times T$ and $T \times (p+q)$ matrices involving ϕ and θ only. Let

$$\sigma_\varepsilon^2 \Omega = E(ee')$$

where for any particular model in (A.1), Ω is readily calculated. Let T be nonsingular $(p+q)$ matrix such that $T\Omega T' = I$. It follows

$$\begin{pmatrix} u_* \\ \varepsilon \end{pmatrix} = \begin{pmatrix} 0 \\ L \end{pmatrix} v + \begin{pmatrix} I_{p+q} \\ KT^{-1} \end{pmatrix} u_* \tag{A.3}$$

where $u_* = Te$. Denote (A.3) as

$$\begin{pmatrix} u_* \\ \varepsilon \end{pmatrix} = Hv + Zu_* \tag{A.4}$$

The inverse of the covariance matrix $V = E(vv')$ is

$$V^{-1} = \frac{1}{\sigma_\varepsilon^2} H'(I_{T*} - Z(Z'Z)^{-1}Z)H \tag{A.5}$$

where $T* = T + p + q$; the determinant of V is

$$|V| = \sigma_\varepsilon^{2T} |Z'Z| \tag{A.6}$$

As $Z'Z$ is a $(p+q) \times (p+q)$ matrices and, in practice, p and q are small integers, the computation of the inverse and determinant of V does not pose any difficulty.

References

Anderson, O. D., (1975) "On a Lemma Associated with Box, Jenkins and Granger", Journal of Econometrics, Vol. 3, pp. 151-156.

_______, (1977) "An Inequality and a Lemma Revisited", Journal of Econometrics, vol. 6, pp. 135-140.

Anderson, T. W., (1971) The Statistical Analysis of Time Series, Wiley, New York.

Ansley, C. F., W. A. Spivey and W. J. Wrobleski, (1977) "On the Structure of Moving Average Processes", Journal of Econometrics, Vol. 6, pp. 121-134.

Balestra, P. and M. Nerlove, (1966) "Pooling Cross-Section and Time-Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas", Econometrica, pp. 585-612.

Box, G.E.P. and G. M. Jenkins, (1970) Time Series Analysis, Forecasting and Control, San Francisco: Holden Day.

Cochrane, D., and G. H. Orcutt, (1949) "Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms", Journal of the American Statistical Association, Vol. 44, pp. 32-61.

Graybill, F. A. (1961) An Introduction to Linear Statistical Models, McGraw-Hill Book Company, Inc., New York.

Henderson, C. R., Jr., (1971) "Comment on 'The Use of Error Components Models in Combining Cross Section with Time Series Data'", Econometrica, Vol. 39, p. 397.

Lee, L. F., (1978) "On the Issues of Fixed Effects vs. Random Effects Econometric Models with Panel Data", Discussion Paper no. 78-101, June 1978, Center for Economic Research, Department of Economics, University of Minnesota.

Maddala, G. S., (1971) "The Use of Variance Components Models in Pooling Cross-Section and Time-Series Data", Econometrica, pp. 341-358.

Nerlove, M., (1971) "A Note on Error Components Models", Econometrica, Vol 39, pp. 383-396.

Newbold, P., (1974) "The Exact Likelihood Function for a Mixed Autoregressive-Moving Average Process", Biometrika 61, pp. 423-426.

Revankar, N. S., (1977) "Error Components Models with Serially Correlated Time Effects", Discussion paper, State University of New York, Buffalo, October 1977.

Shaman, P., (1973) "On the inverse of the Covariance Matrix for an Autoregressive-moving Average Processes", Biometrika 60, pp. 193-196.

Theil, H., and A. S. Goldberger, (1961) "On Pure and Mixed Statistical Estimation in Economics", International Economic Review, 2, pp. 65-78.

Tiao, G. C. and M. M. Ali, (1971) "Analysis of Correlated Random Effects: Linear Model with Two Random Components", Biometrika 58, pp. 37-51.

Wallace, T. D., and A. Hussain, (1969) "The Use of Error Component Models in Combing Cross-Section and Time Series Data", Econometrica, pp. 55-72.