ON THE ISSUES OF FIXED EFFECTS VS. RANDOM EFFECTS ECONOMETRIC MODELS WITH PANEL DATA

by

Lung-Fei Lee

Discussion Paper No. 78-101, June 1978

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

ON THE ISSUES OF FIXED EFFECTS VS. RANDOM EFFECTS ECONOMETRIC MODELS WITH PANEL DATA

by Lung-Fei Lee

1. Introduction

The analysis of cross-section and time-series data has had a long history. The traditional model for pooling has been based on the equation

$$y_{it} = \alpha_i + \gamma_t + x_{it}\beta + w_{it} \qquad i=1,\ldots,N;\ t=1,\ldots,T. \tag{1.1}$$

Assuming α's and γ's are fixed parameters, it is the covariance model (CV) or least squares with dummy variables (LSDV) model. Assuming α's and γ's to be random, it is the variance components model (VC). The latter model has been discussed at length by Balestra and Nerlove (1966), Nerlove (1971), Wallace and Hussain (1969), Maddala (1971), etc. These two different points of view lead to two different estimators, the CV model has resulted in the "within" (or covariance) estimator whereas the VC model has lead to a generalized least squares estimator (GLS).

Mundlak (1978) objects to the distinction between fixed effects and random effects models on the ground that the rules for deciding whether an effect is fixed or random are arbitrary. He argued that the whole literature which has been based on an imaginary difference between the two estimators is based on an incorrect specification

---

This research has been supported by National Science Foundation under grant SOC-78-07304. I would like to thank C. A. Sims and G. S. Maddala for valuable comments. Responsibility for any errors is my own.

which ignores the correlation between the effects α's and γ's and the explanatory variables. He argued also that when the model is properly specified, the GLS is identical to the "within" estimator and therefore there is only one estimator.

While Mundlak is correct in arguing that the distinction between the two models are arbitrary, there are still many problems remaining. One common problem is that some of the explanatory variables are individual-specific and do not vary over time, or time specific and do not vary across individuals. This has been a common experience with the estimation of earnings functions from panel data (see, e.g. Weiss and Lillard (1977)). Another example is in the estimation of production functions with data from individual firms in Lee and Pitt [1978]. In such cases if one uses the covariance method all these variables have to be ignored since they are collinear with the dummy variables for the different units. This problem does not arise in the variance components framework so long as there is some between group variation among these variables.[1/] Another problem is that if there were no correlation between the effects and the explanatory variables, as contrary to the extreme assumptions made in Mundlak, the GLS and "within" estimators would not be identical. In this case, the question why there are two competing estimators raised by Mundlak [1978] has not been answered at all. Thus the issues of fixed effects vs. random effects models are far from resolved.

---

[1] I am indebted to Professor G. S. Maddala for raising this question.

In the paper, we are trying to investigate the above problems. The paper is organized as follows. In section 2, we begin with the estimation problems of the coefficients of individual specific factors which are invariant over time in both the VC model and CV model. In section 3, the models are generalized to include both time invariant individual specific factors and individual invariant time specific factors. In section 4, we will discuss the issue of fixed effects vs. random effects models. In the final section, we draw our conclusions.

2. Estimation and Analysis of Models with Specific Variables

2.1 GLS and Extend Covariance Estimation Procedures

Let us consider the linear equation

$$(2.1.1) \qquad y_{it} = \alpha + z_i\delta + x_{it}\beta + u_i + w_{it} \qquad \begin{matrix} i=1,\ldots,N \\ t=1,\ldots,T \end{matrix}$$

where x_{it} is a $1 \times k$ vector of explanatory variables which vary over time, z_i is a $1 \times p$ vector of individual specific variables which vary across individual units but do not vary over time. We assume that w_{it} are i.i.d.,

$$(2.1.2) \qquad E(w_{it}|z_j, x_{jt'} \text{ all } j, t') = 0 ,$$

$$E(w_{it}^2|z_j, x_{jt'} \text{ all } j, t') = \sigma_w^2$$

for all i and t. u_i in the above equation represents the specific effect associated with the i^{th} unit in the underlying population. Following Mundlak [1978], the effects u_i can be assumed to be random from the outset and one views the LSDV inference as a conditional inference conditional on the effects that are in the sample. We assume that u_i are i.i.d.,

$$(2.1.3) \qquad E(u_i|z_j, x_{jt'} \text{ all } j, t') = 0 ,$$

$$E(u_i^2|z_j, x_j \text{ all } j, t') = \sigma_u^2$$

and u_i and w_{it} are independent for all i and t. To simplify notations, we assume further that $\frac{1}{NT} \Sigma_{i=1}^N \Sigma_{t=1}^T x_{it} = 0$ and $\frac{1}{N} \Sigma_{i=1}^N z_i = 0$.

Unconditional on the effects u_i, the model is similar to the traditional variance components (TVC) model investigated in Balestra and Nerlove (1966), Nerlove (1971), Wallace and Hussain (1969) and Maddala (1971), etc., with the exception that time invariant variables z_i are introduced. In this VC framework, all the coefficients in the model can be estimated by GLS procedure. In matrix notations, equation (2.1.1) is

$$y = \ell_N \otimes \ell_T \alpha + Z\otimes\ell_T \delta + X\beta + u\otimes\ell_T + w \tag{2.1.4}$$

where $\otimes$ denotes the Kronecker product, ℓ_N, ℓ_T are $N \times 1$ and $T \times 1$ vectors with all elements unity, y is an $NT \times 1$ vector, X is an $NT \times k$ matrix of variables x_{it}, Z is an $N \times p$ matrix of z_i, etc. The covariance matrix of the disturbances in (2.1.4) is $\Omega = \sigma_w^2 I_N \otimes I_T + \sigma_u^2 I_N \otimes \ell_T \ell_T'$ and its inverse is (see Graybill [1961], Wallace and Hussain [1969], Nerlove [1971]) $\Omega^{-1} = \frac{1}{\sigma_w^2}(I_N \otimes I_T - \frac{\sigma_u^2}{\sigma_w^2 + T\sigma_u^2} I_N \otimes \ell_T \ell_T')$ where I_N and I_T are $N \times N$ and $T \times T$ identity matrices. The GLS estimate of the coefficients in (2.1.4) if σ_u^2 and σ_w^2 are known, is

$$\begin{pmatrix} \alpha \\ \delta \\ \beta \end{pmatrix}_G = \begin{pmatrix} NT\Theta & 0 & 0 \\ 0 & \Theta TZ'Z & \Theta TZ'\bar{X}^{in} \\ 0 & \Theta T\bar{X}^{in'}Z & W_{xx} + \Theta B_{xx} \end{pmatrix}^{-1} \begin{pmatrix} \Theta NT\bar{y} \\ \Theta TZ'\bar{y}^{in} \\ W_{xy} + \Theta B_{xy} \end{pmatrix} \tag{2.1.5}$$

where $\Theta = \frac{\sigma_w^2}{\sigma_w^2 + T\sigma_u^2}$, $\bar{X}^{in'} = (\bar{x}_{1.}' \dots \bar{x}_{N.}')$ an $k \times N$ matrix with

$\bar{x}_{i\cdot} = \frac{1}{T}\Sigma_{t=1}^{T} x_{it}$, $\bar{y}^{in'} = (\bar{y}_{1\cdot},\ldots,\bar{y}_{N\cdot})$ an $1 \times N$ vector with

$\bar{y}_{i\cdot} = \frac{1}{T}\Sigma_{t=1}^{T} y_{it}$, $\bar{y} = \frac{1}{NT}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T} y_{it}$, $B_{xx} = T\ \Sigma_{i=1}^{N} \bar{x}_{i\cdot}'\bar{x}_{i\cdot}$,

$W_{xx} = X'X - B_{xx}$, $B_{xy} = T\ \Sigma_{i=1}^{N} \bar{x}_i'\bar{y}_i$ and $W_{xx} = X'y - B_{xy}$. Thus the matrix B_{xx} contains the sum of squares and sums of products between groups and W_{xx} is the corresponding matrix within groups.

Conditional on the effects u in the sample, u is a vector of unknown parameters. It is more convenient in this case to write equation (2.1.4) as

$$y = \ell_N \otimes \ell_T \alpha + I_N \otimes \ell_T u + Z \otimes \ell_T \delta + X\beta + w \tag{2.1.6}$$

Equation (2.1.6) is the CV model and the coefficient vector β can be estimated by the covariance method. However when one uses the covariance method, the variables Z have to be ignored since they are collinear with $I_N \otimes \ell_T$, i.e., $z_i\delta$ and u_i can not be identified in (2.1.6). The equation that can be estimated is

$$y = \ell_N \otimes \ell_T v + X\beta + w \tag{2.1.6$'$}$$

where v is a $N \times 1$ vector of unknown parameters. The covariance estimates of β and v are

$$\hat{\beta}_c = W_{xx}^{-1} W_{xy} \quad \text{and} \tag{2.1.7}$$

$$\hat{v}_i = \bar{y}_{i\cdot} - \bar{x}_{i\cdot}\hat{\beta}_c \qquad i=1,\ldots,N \tag{2.1.8}$$

While it is true that δ can not be estimated from equation (2.1.6), it can be estimated from the given sample in this approach. What one needs to utilize is the relation,

$$v_i = \alpha + z_i\delta + u_i \qquad i=1,\ldots,N \tag{2.1.9}$$

from (2.1.6) and (2.1.6)'.

From (2.1.8) $E(\hat{v}|X, Z, u) = v$ and since $E(u|X,Z) = 0$, it follows from (2.1.9) that

$$\begin{aligned} E(\hat{v}|X, Z) &= E_u E(\hat{v}|X,Z,u) \\ &= E_u(v|X,Z) \\ &= \ell_N \cdot \alpha + Z\delta \end{aligned} \tag{2.1.10}$$

where E_u is the expectation operator of u conditional on X and Z. Equation (2.1.10) can be written as a regesssion equation

$$\hat{v} = \ell_N \cdot \alpha + Z\delta + \zeta \tag{2.1.11}$$

where $E(\zeta|Z,X) = 0$. Ordinary least squares can then be applied to (2.1.11) which gives

$$\hat{\alpha}_c = \frac{1}{N}\Sigma_{i=1}^N \hat{v}_i \tag{2.1.12}$$

$$\hat{\delta}_c = (Z'Z)^{-1}Z'\hat{v} \quad \text{or equivalently} \tag{2.1.13}$$

$$\hat{\alpha}_c = \bar{y} \tag{2.1.12)'}$$

$$\hat{\delta}_c = (Z'Z)^{-1}Z'(\bar{y}^{in} - \bar{X}^{in}\hat{\beta}_c) \tag{2.1.13)'}$$

Let us call this estimation procedure an extended covariance procedure and the estimates $\hat{\alpha}_c$ and $\hat{\delta}_c$ of α and δ extended covariance estimators. In other econometric models, this kind of two steps procedures has been proposed in Wachter [1970] and analyzed in Amemiya [1976].

Thus all the parameters in equation (2.1.1) can be estimated in both the random and fixed effects models. It is interesting to compare these two estimation procedures. The VC model in (2.1.4) differs from the TVC model of Balestra and Nerlove (1966) in that invariant variables Z are included. In the TVC model, GLS estimate is a weighted average of "within" and "between" estimates as argued in Maddala [1971], and for fixed N, as T goes to infinite, GLS estimate will tend to the "within" estimate. However in our model, the "within" estimator for δ does not exist. One may wonder what will happen to our GLS estimates in (2.1.5) when T goes to infinite. From (2.1.5), it is obvious that

$$\text{(2.1.14)} \qquad \hat{\alpha}_G = \bar{y}$$

and

$$\text{(2.1.15)} \qquad \begin{pmatrix} \delta \\ \beta \end{pmatrix}_G = \begin{pmatrix} \Theta\frac{1}{N}Z'Z & \Theta\frac{1}{N}Z'\bar{X}^{in} \\ \Theta\frac{1}{N}\bar{X}^{in'}Z & \frac{1}{NT}W_{xx} + \Theta\frac{1}{NT}B_{xx} \end{pmatrix}^{-1} \begin{pmatrix} \Theta\frac{1}{N}Z'\bar{y}^{in} \\ \frac{1}{NT}W_{xy} + \Theta\frac{1}{NT}B_{xy} \end{pmatrix}$$

Assuming that the inverse matrix in (2.1.15) does exist, (2.1.15) implies with the well-known formulae for inversion of a partitioned matrix

$$\text{(2.1.16)} \qquad \hat{\beta}_G = \left(\frac{1}{NT}W_{xx} + \Theta\frac{1}{N}\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{X}^{in}\right)^{-1}\left(\frac{1}{NT}W_{xy} + \Theta\frac{1}{N}\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{y}^{in}\right)$$

$$\text{(2.1.17)} \qquad \hat{\delta}_G = (Z'Z)^{-1}Z'(\bar{y}^{in} - \bar{X}^{in}\hat{\beta}_G)$$

The GLS estimate $\hat{\beta}_G$ in (2.1.16) is a weighted average of the "within" estimate and a "between" estimate of β. As $\Theta = 0$, the GLS $\hat{\beta}_G$ and $\hat{\delta}_G$ correspond to the extended covariance estimates $\hat{\beta}_c$ and $\hat{\delta}_c$ in (2.1.7) and (2.1.13)'. The ordinary least squares (OLS) estimates of (2.1.4) are

$$\hat{\beta}_L = (\frac{1}{NT} W_{xx} + \frac{1}{N}\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{X}^{in})^{-1}(\frac{1}{NT}W_{xy}+\frac{1}{N}\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{y}^{in})$$

and $\hat{\delta}_L = (Z'Z)^{-1}Z'(\bar{y}^{in} - \bar{X}^{in}\hat{\beta}_L)$. Thus $\hat{\beta}_G$ and $\hat{\delta}_G$ would be the OLS estimates corresponding to $\Theta = 1$. As $T \to \infty$ (and hence $\Theta \to 0$), if $\frac{1}{NT}W_{xx}$ and $\frac{1}{N}\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{X}^{in}$ have well-defined probability limits, $\operatorname*{plim}_{T\to\infty} \sqrt{T}(\hat{\beta}_G - \hat{\beta}_c) = 0$ and $\operatorname*{plim}_{T\to\infty} (\hat{\delta}_G - \hat{\delta}_c) = 0$ for fixed N, or $\operatorname*{plim}_{\substack{N\to\infty\\T\to\infty}}\sqrt{NT}\,(\hat{\beta}_G - \hat{\beta}_c) = 0$ and $\operatorname*{plim}_{\substack{N\to\infty\\T\to\infty}}\sqrt{N}\,(\hat{\delta}_G - \hat{\delta}_c) = 0$, i.e., the GLS procedure and the extended covariance procedure are asymptotically equivalent.

The covariance matrices of $\hat{\beta}_G$ and $\hat{\delta}_G$ are

$$(2.1.18) \quad E((\hat{\beta}_G - \beta)^2|Z,X) = \sigma_w^2(W_{xx} + \Theta T\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{X}^{in})^{-1} \quad \text{and}$$

$$(2.1.19) \quad E((\hat{\delta}_G - \delta)^2|Z,X) = \sigma_w^2(\Theta TZ'Z - \Theta TZ'\bar{X}^{in}(W_{xx} + \Theta B_{xx})^{-1}\Theta T\bar{X}^{in'}Z)^{-1}\ .$$

As u and w are independent, $E((\hat{\beta}_c - \beta)^2|X,Z,u) = E((\hat{\beta}_c - \beta)^2|X,Z)$; the covariance matrices of $\hat{\beta}_c$ and $\hat{\delta}_c$ are

$$(2.1.20) \quad E((\hat{\beta}_c - \beta)^2|Z,X) = \sigma_w^2 W_{xx}^{-1} \quad \text{and}$$

$$(2.1.21) \quad E((\hat{\delta}_c-\delta)^2|Z,X) = \frac{\sigma_w^2}{\Theta T}(Z'Z)^{-1} + \sigma_w^2(Z'Z)^{-1}Z'\bar{X}^{in}W_{xx}^{-1}\bar{X}^{in'}Z(Z'Z)^{-1}$$

It follows that $E((\hat{\beta}_c-\beta)^2|Z,X) - E((\hat{\beta}_G-\beta)^2|Z,X)$ is positive semi-definite, i.e., $\hat{\beta}_G$ is more efficient than $\hat{\beta}_c$, whenever $(I-Z(Z'Z)^{-1}Z')\bar{X}^{in} \neq 0$. To compare (2.1.19) with (2.1.21), one notes that

$$(2.1.22) \qquad (\Theta TZ'Z - \Theta TZ'\bar{X}^{in}(W_{xx} + \Theta B_{xx})^{-1}\Theta T\bar{X}^{in'}Z)^{-1}$$
$$= (\Theta TZ'Z)^{-1} + (Z'Z)^{-1}Z'\bar{X}^{in}(W_{xx}+\Theta T\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{X}^{in})^{-1}$$
$$\bar{X}^{in'}Z(Z'Z)^{-1}$$

It follows from (2.1.19), (2.1.21) and (2.1.22) that whenever $(I-Z(Z'Z)^{-1}Z')\bar{X}^{in} = 0$, $\hat{\delta}_G$ is more efficient than $\hat{\delta}_c$.

When $(I-Z(Z'Z)^{-1}Z')\bar{X}^{in} = 0$, $\hat{\beta}_G = \hat{\beta}_c$ and $\hat{\delta}_G = \hat{\delta}_c$ i.e., GLS estimates are the same as the extended covariance estimates. This result generalizes the result obtained in Mundlak [1978]. However Mundlak's motivation is quite different from ours. Mundlak argued, in the TCV model, the GLS approach has completely neglected the consequences of the correlation which might exist between the effects and the included explanatory variables. He introduced an auxiliary regression

$$(2.1.23) \qquad \alpha_i = \bar{x}_{i.}\delta + u_i$$

in addition to equation (1.1). This, in turn, implies equation (2.1.1) when time effects are ignored with $z_i = \bar{x}_{i.}$. In this approach, from practical point of view, the auxiliary regression introduced in (2.1.23) is not necessarily a proper procedure to correct specification errors since it ignores completely the degree of correlation between α_i and x_{it}. If the correlation between α_i and x_{it}

occurs, it is basically an omitted variables problem. This can be correlated if appropriately omitted variables z_i are included in (2.1.1). However the omitted variables z_i need not be $\bar{x}_{i.}$ at all. Such an example can be found in Lee and Pitt [1978]. Also if α_i is only correlated with some but not all the included explanatory variables, (2.1.23) is again not appropriate. Whenever Z in (2.1.1) does not contain a submatrix which spans the column space of $\bar{X}^{in}$ in R^N, GLS estimates of β will not be the same as the covariance estimate.

The GLS estimate is prefered to the covariance estimate when the model is properly specified since it utilizes certain "between" group information in $\bar{X}^{in}$. However the extended covariance approach may be prefered if the correlation between effects and explanatory variables persists.[2] In this case, between group estimates and hence GLS estimates are biased but not the covariance estimates.

2.2 More Complicated Covariance Estimator

Instead of estimating equation (2.1.11) by simple OLS procedure, it can be estimated by the more complicated GLS procedure which takes account the non-diagonal covariance matrix of ζ.

From (2.1.6)', (2.1.7) and (2.1.8),

$$\zeta = u + (\tfrac{1}{T} I_N \otimes \ell_T' - \bar{X}^{in} W_{xx}^{-1} X' Q) w \tag{2.2.1}$$

where $Q = I_{NT} - \frac{1}{T} I_N \otimes \ell_T \ell_T'$ is the covariance transformation matrix. It follows that the covariance matrix of ζ is

---

[2] E.g. the omitted variables are unobservable. Recently Hausman [1977] has proposed an interesting specification error test in a wide class of econometric models. His test can be used to test such possible correlation in the pooled model (see Hausman [1977], section 3).

(2.2.2)
$$V \equiv E(\zeta\zeta' \,|\, X,Z)$$
$$= \frac{\sigma_w^2}{T\Theta} I_N + \sigma_w^2 \bar{X}^{in} W_{xx}^{-1} \bar{X}^{in'}$$

Equation (2.1.11) can be estimated by GLS procedure which gives, if σ_w^2 and σ_u^2 are known,

(2.2.3)
$$\begin{pmatrix}\hat{\alpha}\\ \delta\end{pmatrix}_{cG} = \{\begin{pmatrix}\ell_N'\\ Z'\end{pmatrix} V^{-1}(\ell_N \; Z)\}^{-1}\begin{pmatrix}\ell_N'\\ Z'\end{pmatrix} V^{-1}\hat{v}$$

It is easy to check that

(2.2.4)
$$V^{-1} = \frac{T\Theta}{\sigma_w^2} I_N - \frac{T^2\Theta^2}{\sigma_w^2} \bar{X}^{in}(W_{xx} + \Theta B_{xx})^{-1}\bar{X}^{in}$$

With (2.2.4) we have

$$\ell_N' V^{-1} Z = 0,$$

$$Z'V^{-1}Z = \frac{T\Theta}{\sigma_w^2} Z'Z - \frac{T^2\Theta^2}{\sigma_w^2} Z'\bar{X}^{in}(W_{xx}+\Theta B_{xx})^{-1}\bar{X}^{in'}Z$$

and
$$Z'V^{-1}\hat{v} = \frac{T\Theta}{\sigma_w^2} Z'[I_N - T\Theta\bar{X}^{in}(W_{xx}+\Theta B_{xx})^{-1}\bar{X}^{in'}](\bar{y}^{in}-\bar{X}^{in}W_{xx}^{-1}W_{xy})$$

$$= \frac{T\Theta}{\sigma_w^2} Z'[\bar{y}^{in}-\bar{X}^{in}(W_{xx}+\Theta B_{xx})^{-1}(W_{xy}+\Theta B_{xy})]$$

Hence

(2.2.5)
$$\hat{\delta}_{cG} = [\Theta Z'Z - T\Theta^2 Z'X^{in}(W_{xx}+\Theta B_{xx})^{-1}\bar{X}^{in}Z]^{-1}\Theta Z'[\bar{y}^{in}-\bar{X}^{in}(W_{xx} + \Theta B_{xx})^{-1}(W_{xy} + \Theta B_{xy})]$$

which is exactly $\hat{\delta}_G$ in (2.1.17) written in an equivalent form from (2.1.15). Similarly one can check that $\hat{\alpha}_{cG}$ in (2.2.3) is equal to $\hat{\alpha}_G$. Hence we conclude that the BLUE estimate of δ can be derived from the CV approach. From this, it implies also that even if X and u were correlated, $\hat{\delta}_G$ would be unbiased.

3. Analysis of models including time effects and time specific variables

3.1 GLS and Extended Covariance Estimation Procedures

In this section, we would like to generalize our model to include time effects and time specific explanatory variables which vary over time but do not vary across individual units. We consider the following model

$$(3.1.1) \qquad y_{it} = \alpha + z_i\delta + r_t\gamma + x_{it}\beta + u_i + v_t + w_{it}$$

We assume that u_i, v_t and w_{it} are mutually independent, u_i is i.i.d. for different i, $E(u_i|z_j, r_t, x_{jt},$ all $j, t) = 0$, $E(u_i^2|z_j, r_t, x_{jt},$ all $j, t) = \sigma_u^2$ for all i, and similar conditions hold for v_t and w_{it}. Furthermore, to simplify notations, we assume $\Sigma_{i=1}^N \Sigma_{t=1}^T x_{it} = 0$, $\Sigma_{i=1}^N z_i = 0$ and $\Sigma_{t=1}^T r_t = 0$.

Unconditional on the effects u_i and v_t, it is a VC model and it can be estimated by GLS procedure.

The covariance matrix of the errors in (3.1.1) is $\Sigma = \sigma_w^2 I_{NT} + \sigma_u^2 I_N \otimes \ell_T\ell_T' + \sigma_v^2 \ell_N\ell_N' \otimes I_T$ and as derived in Wallace and Hussain (1969) and Nerlove (1971) its inverse is

$$(3.1.2) \qquad \Sigma^{-1} = \frac{1}{\sigma_w^2}(I_{NT} - \gamma_1 I_N \otimes \ell_T\ell_T' - \gamma_2 \ell_N\ell_N' \otimes I_T + \gamma_3 \ell_N\ell_N' \otimes \ell_T\ell_T')$$

where

$$\gamma_1 = \frac{\sigma_u^2}{\sigma_w^2 + T\sigma_u^2}, \quad \gamma_2 = \frac{\sigma_v^2}{\sigma_w^2 + N\sigma_v^2} \quad \text{and}$$

$$\gamma_3 = \frac{\sigma_u^2\sigma_v^2}{(\sigma_w^2 + T\sigma_u^2)(\sigma_N^2 + N\sigma_v^2)} \left(\frac{2\sigma_w^2 + N\sigma_v^2 + T\sigma_u^2}{\sigma_w^2 + N\sigma_v^2 + T\sigma_u^2}\right) .$$

The GLS estimates, if σ_u^2, σ_v^2 and σ_w^2 are known, are

(3.1.3) $$\hat{\alpha}_G = \frac{1}{NT} \Sigma_{i=1}^{N}\Sigma_{t=1}^{T} y_{it}$$

and

(3.1.4) $$\begin{pmatrix} \hat{\delta} \\ \gamma \\ \beta_G \end{pmatrix} = \begin{pmatrix} \Theta_1 TZ'Z & 0 & \Theta_1 TZ'\bar{X}^{in} \\ 0 & \Theta_2 NR'R & \Theta_2 NR'\bar{X}^{bt} \\ \Theta_1 T\bar{X}^{in'}Z & \Theta_2 N\bar{X}^{bt'}R & W_{xx} + \Theta_1 T\bar{X}^{in'}\bar{X}^{in} + \Theta_2 N\bar{X}^{bt'}\bar{X}^{bt} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \Theta_1 TZ'\bar{y}^{in} \\ \Theta_2 NR'\bar{y}^{bt} \\ X'y - \gamma_1 T^2\bar{X}^{in'}\bar{y}^{in} - \gamma_2 N^2\bar{X}^{bt'}\bar{y}^{bt} \end{pmatrix}$$

where $\Theta_1 = 1 - T\gamma_1$, $\Theta_2 = 1 - N\gamma_2$, $\Theta_3 = NT\gamma_3 - 1$,

$R' = (r_1', \ldots, r_T')$, $\bar{X}^{bt'} = (\bar{x}_{\cdot 1}', \ldots, \bar{x}_{\cdot T}')$, $\bar{x}_{\cdot t}' = \frac{1}{N} \Sigma_{i=1}^{N} x_{it}$,

$\bar{y}^{bt'} = (\bar{y}_{.1}, \ldots, \bar{y}_{.T})$, $y_{\cdot t} = \frac{1}{N} \Sigma_{i=1}^{N} y_{it}$, and

$W_{xx} = X'X - T\bar{X}^{in'}\bar{X}^{in} - N\bar{X}^{bt'}\bar{X}^{bt}$.

Other notations in (3.1.4) are the same as defined in previous sections.

Conditional on the effects u_i and v_t in the sample, u and v are constant vectors of unknown parameters. It is convenient to rewrite (3.1.1) as

$$y_{it} = \alpha + u_i + v_t + z_i\delta + r_t\gamma + x_{it}\beta + w_{it} \qquad (3.1.5)$$

However in this fixed effects framework, the invariant factors $z_i\delta$ and $r_t\gamma$ cannot be identified in (3.1.5) with u_i and v_t. Furthermore, the dummy variables for the effects are multicollinear. The function that can be estimated is

$$y_{it} = u_i^* + v_t^* - \frac{1}{T}\Sigma_{t=1}^T v_t^* + x_{it}\beta + w_{it} \qquad (3.1.6)$$

with

$$u_i^* = \alpha + z_i\delta + \frac{1}{T}\Sigma_{t=1}^T v_t + u_i \quad \text{and} \qquad (3.1.7)$$

$$v_t^* = (r_t - r_1)\gamma + v_t - v_1 \; .$$

The covariance estimates of the coefficients in (3.1.6) are

$$\hat{\beta}_c = W_{xx}^{-1}W_{xy} \qquad (3.1.8)$$

where $W_{xy} = X'y - T\bar{X}^{in'}y - N\bar{X}^{bt'}y;$

$$\hat{u}_i^* = \bar{y}_{i.} - \bar{x}_{i.}\hat{\beta}_c \qquad i=1,\ldots,N \qquad (3.1.9)$$

and

$$\hat{v}_t^* = y_{.t} - \bar{y}_{.1} + (\bar{x}_{.t} - \bar{x}_{.1})\hat{\beta}_c \qquad t=2,\ldots,T \qquad (3.1.10)$$

The coefficients of the specific variables can be estimated as follows. From equations (3.1.7), (3.1.9) and (3.1.10),

$$\hat{u}_i^* = \alpha + z_i\delta + \frac{1}{T}\Sigma_{t=1}^T v_t + u_i + \xi_{1i} \qquad i = 1,\ldots,N \qquad (3.1.11)$$

and

$$\hat{v}_t^* = (r_t - r_1)\gamma + v_t - v_1 + \xi_{2t} \qquad t=2,\ldots,T \tag{3.1.12}$$

where $E(\bar{v} + u_i + \xi_{1i}|Z,X) = 0$ and $E(v_t - v_1 + \xi_{2t}|Z,X) = 0$ with $\bar{v}=\frac{1}{T}\Sigma_{t=1}^T v_t$. Equation (3.1.11) can be estimated by OLS procedure which neglects the complicated covariance structure introduced by ξ_{1i}. The OLS estimates of α and δ in (3.1.11) are

$$\hat{\alpha}_c = \frac{1}{N}\Sigma_{i=1}^N(\bar{y}_i - \bar{x}_i\hat{\beta}_c) \tag{3.1.13}$$

$$\hat{\delta}_c = (Z'Z)^{-1}Z'(\bar{y}^{in} - \bar{X}^{in}\hat{\beta}_c) \tag{3.1.14}$$

While equation (3.1.12) can be estimated by OLS, it is more interesting to estimate (3.1.12) by GLS procedure with matrix $I_{T-1} + \ell_{T-1}\ell_{T-1}'$, which is proportional to the covariance matrix of $v_t - v_1$, $t = 2,\ldots,T$, as the covariance matrix for the disturbances in (3.1.12). The inverse matrix of $I_{T-1} + \ell_{T-1}\ell_{T-1}'$ is

$$(I_{T-1} + \ell_{T-1}\ell_{T-1}')^{-1} = I_{T-1} - \frac{1}{T}\ell_{T-1}\ell_{T-1}' \tag{3.1.15}$$

With this procedure, γ will be estimated by

$$\hat{\gamma}_c = (R'R)^{-1}R'(\bar{y}^{bt} - \bar{X}^{bt}\hat{\beta}_c) \tag{3.1.16}$$

These two estimation procedures can be compared. The GLS estimates in (3.1.4) can be written explicitly as

$$\begin{aligned} \hat{\beta}_G = \{&\frac{1}{NT}W_{xx} + \Theta_1 \frac{1}{N}\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{X}^{in} + \Theta_2 \frac{1}{T}\bar{X}^{bt'} \\ &(I-R(R'R)^{-1}R')\bar{X}^{bt}\}^{-1} \\ \{&\frac{1}{NT} W_{xy} + \Theta_1 \frac{1}{N}\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{y}^{in} + \Theta_2 \frac{1}{T}\bar{X}^{bt'} \\ &(I-R(R'R)^{-1}R')\bar{y}^{bt}\} \end{aligned} \tag{3.1.17}$$

(3.1.18) $$\hat{\delta}_G = (Z'Z)^{-1}Z'(\bar{y}^{in} - \bar{X}^{in}\hat{\beta}_G)$$

(3.1.19) $$\hat{\gamma}_G = (R'R)^{-1}R'(\bar{y}^{bt} - \bar{X}^{bt}\hat{\beta}_G)$$

As $N \to \infty$, $T \to \infty$, $\underset{\substack{N\to\infty \\ T\to\infty}}{\text{plim}}\sqrt{NT}\,(\hat{\beta}_G - \hat{\beta}_c) = 0$, $\underset{\substack{N\to\infty \\ T\to\infty}}{\text{plim}}\sqrt{N}\,(\hat{\delta}_G - \hat{\delta}_c) = 0$ and $\underset{\substack{N\to\infty \\ T\to\infty}}{\text{plim}}\sqrt{T}\,(\hat{\gamma}_G - \hat{\gamma}_c) = 0$, so the two estimation procedures are asymptotically equivalent. For a given finite sample, GLS estimates which utilize the "within" sample variation, a modified "between" groups variation and a modified "between" time sample variation, are more efficient than the extended covariance estimates $\hat{\beta}_c$, $\hat{\delta}_c$ and $\hat{\gamma}_c$ with an exception. When the column space of $\bar{X}^{in}$ is a subspace of the column space of Z and the column space of $\bar{X}^{bt}$ is a subspace of the column space of R, the two procedures are identical. This characterizes Mundlak's result [1978] with time effects.

3.2 More Complicated Covariance Estimation Procedures

Instead of estimating (3.1.11) by simple OLS procedure, it can be estimated by more complicated GLS procedure.

Let us denote the vector of disturbances in (3.1.11) by

(3.2.1) $$\begin{aligned} e_1 &= \bar{v}\ell_N + u + \xi_1 \\ &= \bar{v}\ell_N + u + \bar{w}^{in} - \bar{X}^{in}(\hat{\beta}_c - \beta) . \end{aligned}$$

The covariance matrix of e_1 is

(3.2.2) $$\begin{aligned} V_1 &= E(e_1e_1'|X,Z,R) \\ &= \frac{\sigma_v^2}{T}\ell_N\ell_N' + \frac{\sigma_w^2}{T\Theta_1} I_N + \sigma_w^2\bar{X}^{in}W_{xx}^{-1}\bar{X}^{in'} \end{aligned}$$

Let us denote $A = \frac{\sigma_v^2}{T} \ell_N \ell_N' + \frac{\sigma_w^2}{T\Theta_1} I_N$. The inverses of A and V_1 are

$$(3.2.3) \qquad A^{-1} = \frac{T\Theta_1}{\sigma_w^2} (I_N - \frac{\sigma_v^2}{\sigma_w^2 + \sigma_u^2 T + N\sigma_v^2} \ell_N \ell_N')$$

and

$$(3.2.4) \qquad V_1^{-1} = A^{-1} - A^{-1}\bar{X}^{in}(\frac{1}{\sigma_w^2} W_{xx} + \bar{X}^{in'} A^{-1} \bar{X}^{in})^{-1} \bar{X}^{in'} A^{-1}$$

With the covariance matrix in (3.2.2), (3.1.11) can be estimated by GLS procedure. The GLS estimate is, if σ_w^2, σ_u^2 and σ_v^2 are known,

$$(3.2.5) \qquad \begin{pmatrix} \hat{\alpha} \\ \hat{\delta} \end{pmatrix}_{CG} = \{\begin{pmatrix} \ell_N' \\ Z' \end{pmatrix} V_1^{-1} (\ell_N \ Z)\}^{-1} \begin{matrix} \ell_N' \\ Z' \end{matrix} \ V^{-1} u^*$$

It is easy to check that $\ell_N' A^{-1} \bar{X}^{in} = 0$ and $\ell_N' A^{-1} Z = 0$. After simplification,

$$(3.2.6) \qquad \hat{\delta}_{CG} = \{\frac{T\Theta_1}{\sigma_w^2} Z'Z - \frac{T^2\Theta_1^2}{\sigma_w^2} Z'\bar{X}^{in}(W_{xx} + \Theta_1 B_{xx})^{-1} \bar{X}^{in'} Z\}^{-1} \cdot$$

$$\{\frac{T\Theta_1}{\sigma_w^2} Z'\hat{u}^* - \frac{T^2\Theta_1^2}{\sigma_w^2} Z'\bar{X}^{in}(W_{xx} + \Theta_1 B_{xx})^{-1} \bar{X}^{in'} \hat{u}^*\}$$

The covariance matrix of $\hat{\delta}_{CG}$ is

$$(3.2.7) \qquad Var(\hat{\delta}_{CG}) = \{ \frac{T\Theta_1}{\sigma_w^2} Z'Z - \frac{T^2\Theta_1^2}{\sigma_w^2} Z'\bar{X}^{in}(W_{xx} + T\Theta_1 \bar{X}^{in'} \bar{X}^{in})^{-1} \bar{X}^{in'} Z\}^{-1}$$

$$= \frac{\sigma_w^2}{T\Theta_1}(Z'Z)^{-1} + \sigma_w^2 (Z'Z)^{-1} Z'\bar{X}^{in}(W_{xx} + T\Theta_1 \bar{X}^{in'}$$

$$(I - Z(Z'Z)^{-1}Z')\bar{X}^{in})^{-1} \bar{X}^{in'} Z(Z'Z)^{-1} .$$

Equation (3.1.12) can also be estimated with more complicated GLS procedure. Let us denote the vector of disturbances in (3.1.12) by

$$(3.2.8)\qquad e_2 = \begin{pmatrix} v_2 - v_1 \\ \vdots \\ v_T - v_1 \end{pmatrix} + \xi_2$$

$$= \begin{pmatrix} v_2 - v_1 \\ \vdots \\ v_T - v_1 \end{pmatrix} + \begin{pmatrix} \hat{v}_2{}^* - v_2{}^* \\ \vdots \\ \hat{v}_T{}^* - v_T{}^* \end{pmatrix}$$

$$= \begin{pmatrix} v_2 - v_1 \\ \vdots \\ v_T - v_1 \end{pmatrix} + \begin{pmatrix} \bar{w}_{\cdot 2} - \bar{w}_{\cdot 1} \\ \vdots \\ \bar{w}_{\cdot T} - \bar{w}_{\cdot 1} \end{pmatrix} + \begin{matrix} \bar{x}_{\cdot 2} - \bar{x}_{\cdot 1} \\ \vdots \\ \bar{x}_{\cdot T} - \bar{x}_{\cdot 1} \end{matrix} \quad (X'QX)^{-1}X'Qw$$

where $Q = I_{NT} - \frac{1}{N} I_N \otimes \ell_T \ell_T{}' - \frac{1}{N} \ell_N \ell_N{}' \otimes I_T + \frac{1}{NT} \ell_N \ell_N{}' \otimes \ell_T \ell_T{}'$ is the covariance transformation. The covariance matrix of e_2 is

$$(3.2.9)\qquad V_2 = E(e_2 e_2' | X,Z,R)$$

$$= \frac{\sigma_w^2}{N\Theta_2}(I_{T-1} + \ell_{T-1}\ell_{T-1}') + \sigma_w^2 \begin{pmatrix} \bar{x}_{\cdot 2} - \bar{x}_{\cdot 1} \\ \vdots \\ \bar{x}_{\cdot T} - \bar{x}_{\cdot 1} \end{pmatrix} (X'QX)^{-1}$$

$$(\bar{x}'_{\cdot 2} - \bar{x}'_{\cdot 1}, \ldots, \bar{x}_{\cdot T} - \bar{x}_{\cdot 1})$$

Since $(I_{T-1} + \ell_{T-1}\ell_{T-1}')^{-1} = I_{T-1} - \frac{1}{T} \ell_{T-1}\ell_{T-1}'$, the inverse matrix of V_2 is, after simplification,

$$(3.2.10)\qquad V_2^{-1} = \frac{N\Theta_2}{\sigma_w^2} (I_{T-1} - \frac{1}{T} \ell_{T-1}\ell_{T-1}') - \frac{N\Theta_2}{\sigma_w^2} \begin{pmatrix} \bar{x}_{\cdot 2} \\ \vdots \\ \bar{x}_{\cdot T} \end{pmatrix} \cdot$$

$$[\frac{1}{\sigma_w^2} X'QX + \frac{N\Theta_2}{\sigma_w^2} \bar{X}^{bt'}\bar{X}^{bt}]^{-1} \frac{N\Theta_2}{\sigma_w^2} (\bar{x}'_{\cdot 2}, \ldots, \bar{x}'_{\cdot T})$$

With the covariance matrix V_2, (3.1.12) can be estimated by GLS procedure,

$$\hat{\gamma}_{CG} = \{(r_2'-r_1',\ldots,r_T'-r_1')V_2^{-1}\begin{pmatrix} r_2-r_1 \\ \vdots \\ r_T-r_1 \end{pmatrix}\}^{-1}(r_2'-r_1',\ldots,r_T'-r_1')V^{-1}\hat{v}^* . \tag{3.2.11}$$

The covariance matrix of $\hat{\gamma}_{CG}$ is

$$\begin{aligned} \operatorname{var}(\hat{\gamma}_{CG}) &= \{\frac{N\Theta_2}{\sigma_w^2} R'R - \frac{N^2\Theta_2^2}{\sigma_w^2} R'\bar{X}^{bt}(X'Q\,X + N\Theta_2\bar{X}^{bt'}\bar{X}^{bt})^{-1}\bar{X}^{bt'}R\}^{-1} \\ &= \frac{\sigma_w^2}{N\Theta_2}(R'R)^{-1} + \sigma_w^2(R'R)^{-1}R'\bar{X}^{bt}(W_{xx}+N\Theta_2\bar{X}^{bt'} \\ &\qquad (I-R(R'R)^{-1}R')\bar{X}^{bt})^{-1}\bar{X}^{bt'}R(R'R)^{-1} . \end{aligned} \tag{3.2.12}$$

It is interesting to compare the estimates in (3.2.6) and (3.2.11) with the GLS estimates in (3.1.4). The covariance matrix of $(\hat{\delta}_G', \hat{\gamma}_G')$ in (3.1.4) is

$$\begin{aligned} \operatorname{var}\begin{pmatrix} \hat{\delta}_G \\ \hat{\gamma}_G \end{pmatrix} &= \sigma_w^2 \begin{pmatrix} \frac{1}{\Theta_1 T}(Z'Z)^{-1} & 0 \\ 0 & \frac{1}{\Theta_2 N}(R'R)^{-1} \end{pmatrix} \\ &+ \sigma_w^2 \begin{pmatrix} (Z'Z)^{-1}Z'\bar{X}^{in} \\ (R'R)^{-1}R'\bar{X}^{bt} \end{pmatrix} [W_{xx} + \Theta_1 T\bar{X}^{in'}(I-Z(Z'Z)^{-1}Z')\bar{X}^{in} \\ &+ \Theta_2 N\bar{X}^{bt'}(I-R(R'R)^{-1}R')\bar{X}^{bt}]^{-1}(\bar{X}^{in}Z(Z'Z)^{-1},\ \bar{X}^{bt'}R(R'R)^{-1}) \end{aligned} \tag{3.2.13}$$

Comparing (3.2.7), (3.2.12) with (3.2.13), it follows that $\mathrm{var}(\hat{\delta}_{CG})$ - $\mathrm{var}(\hat{\delta}_G)$ is positive semi-definite unless $(I-R(R'R)^{-1}R')\bar{X}^{bt} = 0$, and var $(\hat{\gamma}_{CG})$ - $\mathrm{var}(\hat{\gamma}_G)$ is positive semi-definite unless $(I-Z(Z'Z)^{-1}Z')\bar{X}^{in}=0$. When $(I-R(R'R)^{-1}R')\bar{X}^{bt} = 0$ and $(I-Z(Z'Z)^{-1}Z')\bar{X}^{in} = 0$, the estimates derived from the two approaches are the same. In general, the GLS estimates of δ and γ derived from CV model will be more efficient.

The estimates in (3.2.5) and (3.2.11) are single equation estimates. As e_1 in (3.2.1) and e_2 in (3.2.8) have nonzero correlation, it is possible to derive more efficient estimates by seemingly unrelated regression approach (Zellner [1962]), applied to (3.1.11) and (3.1.12). The covariance matrix of e_1 and e_2 is

$$E(e_1e_2') = -\sigma_w^2\bar{X}^{in}W_{xx}^{-1}(\bar{x}_{\cdot 2}' - \bar{x}_{\cdot 1}',\ldots,\bar{x}_{\cdot T}' - \bar{x}_{\cdot 1}') \tag{3.2.14}$$

Hence the covariance matrix of (e_1', e_2') is

$$V = \begin{pmatrix} \frac{\sigma_v^2}{T}\ell_N\ell_N' + \frac{\sigma_w^2}{T\Theta_1}I_N & 0 \\ 0 & \frac{\sigma_w^2}{N\Theta_2}(I_{T-1}+\ell_{T-1}\ell_{T-1}') \end{pmatrix} \tag{3.2.15}$$

$$+ \sigma_w^2\begin{pmatrix} \bar{X}^{in} \\ \bar{x}_{\cdot 2} - \bar{x}_{\cdot 1} \\ \vdots \\ \bar{x}_{\cdot T} - \bar{x}_{\cdot 1} \end{pmatrix} W_{xx}^{-1}(\bar{X}^{in'}\ \bar{x}_{\cdot 2}'-\bar{x}_{\cdot 1}',\ldots,\bar{x}_{\cdot T}'-\bar{x}_{\cdot 1}')$$

The inverse matrix of V in (3.2.15) is

(3.2.16) $$V^{-1} = \frac{1}{\sigma_w^2}\begin{pmatrix} T\Theta_1\left(I_N - \frac{\sigma_v^2}{\sigma_w^2+\sigma_u^2 T+N\sigma_v^2}\ell_N\ell_N'\right) & 0 \\ 0 & N\Theta_2\left(I_{T-1}-\frac{1}{T}\ell_{T-1}\ell_{T-1}'\right)\end{pmatrix}$$

$$-\frac{1}{\sigma_w^2}\begin{pmatrix} T\Theta_1\bar{X}^{in} \\ N\Theta_2\bar{\bar{x}}_{\cdot 2} \\ \vdots \\ N\Theta_2\bar{\bar{x}}_{\cdot T}\end{pmatrix}(W_{xx}+T\Theta_1\bar{X}^{in'}\bar{X}^{in}+N\Theta_2\bar{X}^{bt'}\bar{X}^{bt})^{-1}$$

$$(T\Theta_1\bar{X}^{in'}\ N\Theta_2\bar{\bar{x}}_{\cdot 2}',\ \ldots, N\Theta_2\bar{\bar{x}}_{\cdot T}')$$

The seeming unrelated regression estimate derived from (3.1.11) and (3.1.12) is

(3.2.17) $$\begin{pmatrix}\hat{\delta} \\ \hat{\gamma}\end{pmatrix}_{CS} = \left\{\begin{pmatrix} Z' & 0,\ldots\ldots,0 \\ 0 & r_2'-r_1',\ldots,r_T'-r_1'\end{pmatrix} V^{-1}\begin{pmatrix} Z' & 0 \\ 0 & r_2-r_1 \\ \vdots & \vdots \\ 0 & r_T-r_1\end{pmatrix}\right\}^{-1}$$

$$\begin{pmatrix} Z' & 0 & 0 \\ 0 & r_2'-r_1',\ldots,r_T'-r_1'\end{pmatrix} V^{-1}\begin{pmatrix}\hat{u}^* \\ \hat{v}^*\end{pmatrix}$$

After simplification with V^{-1} in (3.2.16), it is

$$(3.2.18) \qquad \begin{pmatrix}\hat{\delta}\\ \gamma\end{pmatrix}_{CS} = \left\{\begin{pmatrix}T\Theta_1 Z'Z & 0\\ 0 & N\Theta_2 R'R\end{pmatrix} - \begin{pmatrix}T\Theta_1 Z'\bar{X}^{in}\\ N\Theta_2 R'\bar{X}^{bt}\end{pmatrix}\right.$$

$$\left.(W_{xx}+T\Theta_1\bar{X}^{in'}\bar{X}^{in}+N\Theta_2\bar{X}^{bt'}\bar{X}^{bt})^{-1}(T\Theta_1\bar{X}^{in'}Z \quad N\Theta_2\bar{X}^{bt'}R)\right\}^{-1}$$

$$\left\{\begin{pmatrix}T\Theta_1 Z'(\bar{y}^{in}-\bar{X}^{in}\hat{\beta}_c)\\ N\Theta_2 R'(\bar{y}^{bt}-\bar{X}^{bt}\hat{\beta}_c)\end{pmatrix} - \begin{pmatrix}T\Theta_1 Z'\bar{X}^{in}\\ N\Theta_2 R'\bar{X}^{bt}\end{pmatrix}\right.$$

$$(W_{xx}+T\Theta_1\bar{X}^{in'}\bar{X}^{in}+N\Theta_2\bar{X}^{bt'}\bar{X}^{bt})^{-1}(T\Theta_1\bar{X}^{in'}(\bar{y}^{in}-\bar{X}^{in}\hat{\beta}_c)$$

$$\left.+ N\Theta_2\bar{X}^{bt'}(\bar{y}^{bt}-\bar{X}^{bt}\hat{\beta}_c))\right\}$$

It is of interest to compare (3.2.18) with the GLS estimates from VC model in (3.1.4). The estimates $\hat{\delta}_G$ and $\hat{\gamma}_G$ in (3.1.18) and (3.1.19) can be written in an alternative form from (3.1.4)

$$(3.2.19) \qquad \begin{pmatrix}\hat{\delta}\\ \gamma\end{pmatrix}_{G} = \left\{\begin{pmatrix}T\Theta_1 Z'Z & 0\\ 0 & N\Theta_2 R'R\end{pmatrix} - \begin{pmatrix}T\Theta_1 Z'\bar{X}^{in}\\ N\Theta_2 R'\bar{X}^{bt}\end{pmatrix}\right.$$

$$\left.(W_{xx}+T\Theta_1\bar{X}^{in'}\bar{X}^{in}+N\Theta_2\bar{X}^{bt'}\bar{X}^{bt})^{-1}(T\Theta_1\bar{X}^{in'}Z \quad N\Theta_2\bar{X}^{bt'}R)\right\}^{-1}$$

$$\left\{\begin{pmatrix}\Theta_1 TZ'\bar{y}^{in}\\ \Theta_2 NR'\bar{y}^{bt}\end{pmatrix} - \begin{pmatrix}\Theta_1 TZ'\bar{X}^{in}\\ \Theta_2 NR'\bar{X}^{bt}\end{pmatrix}(W_{xx}+\Theta_1 T\bar{X}^{in'}\bar{X}^{in} +\right.$$

$$\left.\Theta_2 N\bar{X}^{bt'}\bar{X}^{bt})^{-1}(W_{xy}+\Theta_1 T\bar{X}^{in'}\bar{y}^{in}+\Theta_2 N\bar{X}^{bt'}\bar{y}^{bt}\right\}.$$

It is not difficult to check that expressions in (3.2.18) and (3.2.19) are identical. Hence the seemingly unrelated estimates of δ and γ in the CV approach give the same estimates from GLS procedure in the VC approach. This implies also that even if the effects u and v were correlated with X variables, the GLS estimate $\hat{\delta}_G$ and $\hat{\gamma}_G$ would be unbiased.

4. The fixed effects model vs. the random effects model.

In the previous sections, we have demonstrated that it is feasible to estimate the coefficients of specific invariant explanatory variables not only in the VC models but also in the CV models. While accepting the view that fixed effects inference as a conditional inference conditional on the effects in the sample and hence unified the random and fixed effects approaches in a well defined framework, there still lead to two different estimators. Mundlak [1978] argued in the traditional models that the difference was due to specification error and derived his unique estimator results by introducing equation (2.1.23). As argued in section 2, this is rather an extreme point of view. When the specification errors are properly corrected, the two estimators may still be different.

The answer to this question is in fact quite simple as demonstrated below. It lies in the specification and estimation of the fixed effects in the CV models, i.e., the estimates in (2.1.8) or (3.1.9) and (3.1.10). To simplify expressions, consider the traditional CV model without time effects, under similar assumptions as in (2.1.1)

$$y_{it} = \alpha + u_i + x_{it}\beta + w_{it} \tag{4.1}$$

Let

$$x^e_{it} = x_{it} - \bar{x}_{i\cdot} \tag{4.2}$$

be the explanatory variables deviated from their group means. In terms of variables in (4.2), (4.1) becomes

$$(4.1)' \qquad y_{it} = \alpha + u_i + \bar{x}_{i.}\beta + x^e_{it}\beta + w_{it}$$

$$= u^*_i + x^e_{it}\beta + w_{it}$$

where

$$(4.3) \qquad u^*_i = \alpha + \bar{x}_{i.}\beta + u_i$$

With this transformation, $\bar{x}_i$ is introduced explicitly as a vector of specific variables in our model. Instead of estimating $\alpha + u_i$ in (4.1) as in the traditional procedure, we estimate u^*_i in (4.1)' by the covariance method which gives

$$(4.4) \qquad \hat{u}^*_i = \bar{y}_{i.} \qquad i=1,\ldots,N$$

and

$$(4.5) \qquad \hat{\beta}_c = W^{-1}_{xx}W_{xy}$$

Comparing (4.4) and (4.5) with (2.1.7) and (2.1.8), one notes that the covariance estimates of β are the same, but the estimates of specific effects are different. Let $\hat{u}^{*\prime} = (\hat{u}^*_1,\ldots,\hat{u}^*_N)$. It can be easily shown that

$$(4.6) \qquad E((\hat{u}^* - u^*)(\hat{u}^* - u^*)'|X, u) = \frac{\sigma^2_w}{T} I_N$$

$$(4.7) \qquad E((\hat{\beta}_c-\beta)(\hat{\beta}_c-\beta)'|X, u) = E((\hat{\beta}_c-\beta)(\hat{\beta}_c-\beta)'|X) = \sigma^2_w W^{-1}_{xx}$$

and

$$(4.8) \qquad E((\hat{u}^*-u^*)(\hat{\beta}_c-\beta)'|X, u) = 0.$$

While β can be estimated from (4.1)' as in (4.5), it can also be estimated from (4.3) with estimates $\hat{u}_i^*$ in (4.4). Thus there is information in (4.3) which has not been taken into account in the estimation of equation (4.1)'. To properly utilize this information the mixed estimation procedure introduced in Theil and Goldberger [1960] and Theil [1963] can be used. It follows from equations (4.3), (4.4) and (4.5),

$$\hat{\beta}_c = \beta + \xi_1 \tag{4.9}$$

$$\hat{u}^* = \ell_N\alpha + \bar{X}^{in}\beta + u + \xi_2 \tag{4.10}$$

where $\xi_1 = \hat{\beta}_c - \beta$ and $\xi_2 = \hat{u}^* - u^*$. Under the assumptions, i.e., u and w are independent, and the covariance structures in (4.6) - (4.8),

$$E(\xi_1|X) = 0, \quad E(u + \xi_2|X) = 0 \tag{4.11}$$

$$E(\xi_1^2|X) = \sigma_w^2 W_{xx}^{-1} \tag{4.12}$$

$$E((u + \xi_2)^2|X) = (\sigma_u^2 + \frac{\sigma_w^2}{T})I_N \tag{4.13}$$

and

$$E((u+\xi_2)\xi_1|X) = 0 \ . \tag{4.14}$$

By the mixed estimation procedure, coefficients α and β can be estimated as follows,

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}_M = \left\{ \begin{bmatrix} \ell_N' & 0 \\ \bar{X}^{in'} & I_k \end{bmatrix} \begin{pmatrix} (\sigma_u^2 + \frac{\sigma_w^2}{T})I_N & 0 \\ 0 & \sigma_w^2 W_{xx}^{-1} \end{pmatrix}^{-1} \begin{bmatrix} \ell_N & \bar{X}^{in} \\ 0 & I_k \end{bmatrix} \right\}^{-1} \begin{bmatrix} \ell'_N & 0 \\ \bar{X}^{in'} & I_k \end{bmatrix} \begin{pmatrix} (\sigma_u^2 + \frac{\sigma_w^2}{T})I_N & 0 \\ 0 & \sigma_w^2 W_{xx}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \hat{u}^* \\ \hat{\beta}_c \end{pmatrix} \tag{4.15}$$

After simplification,

$$\hat{\alpha}_M = \bar{y} \tag{4.16}$$

and

$$\hat{\beta}_M = (W_{xx} + \frac{\sigma_w^2}{\sigma_w^2 + T\sigma_u^2} T\bar{X}^{in'}\bar{X}^{in})^{-1}(W_{xy} + \frac{\sigma_w^2}{\sigma_w^2 + T\sigma_u^2} T\bar{X}^{in'}\bar{y}^{in}) \tag{4.17}$$

which are exactly the GLS estimates as derived from the VC model.

Let us now consider the traditional covariance model with both individual and time effects, under similar assumptions in section 3,

$$y_{it} = \alpha + u_i + v_i + x_{it}\beta + w_{it} \qquad i=1,\ldots,N;\ t=1,\ldots,T. \tag{4.18}$$

Let

$$x_{it}^d = x_{it} - \bar{x}_{i\cdot} - \bar{x}_{\cdot t} \tag{4.19}$$

be the explanatory variables deviated from their group and time means. In terms of (4.19), (4.18) can be rewritten as

$$y_{it} = u_i^* + v_t^* - \frac{1}{T}\Sigma_{t=1}^T v_t^* + x_{it}^d\beta + w_{it} \tag{4.20}$$

with

$$u_i^* = \alpha + \bar{x}_{i\cdot}\beta + u_i + \frac{1}{T}\Sigma_{t=1}^T v_t \tag{4.21}$$

$$v_t^* = (\bar{x}_{\cdot t} - \bar{x}_{\cdot 1})\beta + v_t - v_1 \quad . \tag{4.22}$$

The covariance estimates of (4.20) are

$$\hat{\beta}_c = W_{xx}^{-1}W_{xy} \tag{4.23}$$

$$\hat{u}_i^* = \bar{y}_{i\cdot} \qquad i=1,\ldots,N \tag{4.24}$$

$$\hat{v}_t^* = \bar{y}_{\cdot t} - \bar{y}_{\cdot 1} \qquad t=2,\ldots,T \tag{4.25}$$

and the covariance matrices of these estimates are

$$(4.26) \qquad E((\hat{u}^*-u^*)(\hat{u}^*-u^*)'|X,u,v) = \frac{\sigma_w^2}{T} I_N$$

$$(4.27) \qquad E((\hat{v}^*-v^*)(\hat{v}^*-v^*)'|X,u,v) = \sigma_w^2 (I_{T-1} - \frac{1}{T}\ell_{T-1}\ell'_{T-1})^{-1}$$

$$(4.28) \qquad E((\hat{\beta}_c-\beta)(\hat{\beta}_c-\beta)'|X,u,v) = E((\hat{\beta}_c-\beta)(\hat{\beta}_c-\beta)'|X) = \sigma_w^2 W_{xx}^{-1}$$

and

$$(4.29) \qquad E((\hat{u}^*-u^*)(\hat{v}^*-v^*)'|X,u,v) = 0 \qquad E((\hat{u}^*-u^*)(\hat{\beta}_c-\beta)'|X,u,v) = 0$$

$$E((\hat{v}^*-v^*)(\hat{\beta}_c-\beta)'|X,u,v) = 0$$

It follows from the above relations (4.21) - (4.29),

$$(4.30) \qquad \hat{\beta}_c = \beta + \xi_1$$

$$(4.31) \qquad \hat{u}_i^* = \alpha + \bar{x}_{i.}\beta + \bar{v} + u_i + \xi_{2i} \qquad i=1,\ldots,N$$

$$(4.32) \qquad \hat{v}_t^* = (\bar{x}_{.t} - \bar{x}_{.1})\beta + v_t - v_1 + \xi_{3t} \qquad t=2,\ldots,T$$

with

$$(4.33) \qquad E(\xi_1^2|X) = \sigma_w^2 W_{xx}^{-1}$$

$$(4.34) \qquad E((\bar{v}\ell_N + u + \xi_2)^2|X) = (\sigma_u^2 + \frac{\sigma_w^2}{T})I_N + \frac{1}{T}\sigma_v^2\ell_N\ell'_N$$

$$(4.35) \qquad E((\tilde{v}-v_1\ell_{T-1} + \xi_3)^2|X) = (\sigma_v^2 + \frac{\sigma_w^2}{N})(I_{T-1} + \ell_{T-1}\ell'_{T-1})$$

and

$$(4.36) \qquad E(\xi_1(\bar{v}\ell_N + u + \xi_2)'|X) = 0, \quad E(\xi_1(\tilde{v}-v_1\ell_{T-1} + \xi_3)'|X) = 0,$$

$$E((\bar{v}\ell_N + u + \xi_2)(\tilde{v}-v_1\ell_{T-1} + \xi_3)'|X) = 0 \ .$$

where $\tilde{v}' = (v_2,\ldots,v_T)$. The mixed estimation of (4.30) - (4.32) with covariance matrices (4.33) - (4.36) gives

$$\begin{pmatrix}\hat{\alpha}\\ \beta\end{pmatrix}_M = \left\{ \begin{bmatrix} \ell_N' & 0 & \cdots & 0 & 0 \\ \bar{X}^{in'} & \bar{x}'_{\cdot 2}-\bar{x}'_{\cdot 1}, & \ldots, & \bar{x}'_{\cdot T}-\bar{x}'_{\cdot 1} & I_k \end{bmatrix} \right.$$

$$\begin{pmatrix} (\sigma_u^2 + \frac{\sigma_w^2}{T})I_N + \frac{1}{T}\sigma_v^2 \ell_N \ell_N' & 0 & 0 \\ 0 & (\sigma_v^2 + \frac{\sigma_w^2}{N})(I_{T-1}+\ell_{T-1}\ell_{T-1}') & 0 \\ 0 & 0 & \sigma_w^2 W_{xx}^{-1} \end{pmatrix}^{-1}$$

$$\left. \begin{bmatrix} \ell_N & \bar{X}^{in} \\ 0 & \bar{x}_{\cdot 2}-\bar{x}_{\cdot 1} \\ 0 & \bar{x}_{\cdot T}-\bar{x}_{\cdot 1} \\ 0 & I_k \end{bmatrix} \right\}^{-1} \begin{bmatrix} \ell_N' & 0 & \cdots & 0 & 0 \\ \bar{X}^{in'} & \bar{x}'_{\cdot 2}-\bar{x}'_{\cdot 1}, & \ldots, & \bar{x}'_{\cdot T}-\bar{x}'_{\cdot 1} & I_k \end{bmatrix} \tag{4.37}$$

$$\begin{pmatrix} (\sigma_u^2 + \frac{\sigma_w^2}{T})I_N + \frac{1}{T}\sigma_v^2 \ell_N \ell_N' & 0 & 0 \\ 0 & (\sigma_v^2 + \frac{\sigma_w^2}{N})(I_{T-1}+\ell_{T-1}\ell_{T-1}') & 0 \\ 0 & 0 & \sigma_w^2 W_{xx}^{-1} \end{pmatrix}^{-1}$$

$$\begin{pmatrix} \hat{u}^* \\ \hat{v}_2^* \\ \vdots \\ \hat{v}_T^* \\ \hat{\beta} \end{pmatrix}$$

(4.37) can be simplified with the relations $[(\sigma_u^2 + \frac{\sigma_w^2}{T})I_N + \frac{1}{T}\sigma_v^2 \ell_N \ell_N']^{-1} =$

$\frac{T}{\sigma_w^2 + T\sigma_u^2}(I_N - \frac{\sigma_v^2}{\sigma_w^2 + T\sigma_u^2 + N\sigma_v^2}\ell_N \ell_N')$, and $(I_{T-1} + \ell_{T-1}\ell_{T-1}')^{-1} =$

$I_{T-1} - \frac{1}{T}\ell_{T-1}\ell_{T-1}'$. It follows

$$\hat{\alpha}_M = \frac{1}{N}\Sigma_{i=1}^N \hat{u}_i^* = \bar{y} \tag{4.38}$$

$$\hat{\beta}_M = (\frac{1}{\sigma_w^2} W_{xx} + \frac{T}{\sigma_w^2 + T\sigma_u^2}\bar{X}^{in'}\bar{X}^{in} + \frac{N}{\sigma_w^2 + N\sigma_v^2}\bar{X}^{bt'}\bar{X}^{bt})^{-1} \cdot \tag{4.39}$$

$$(\frac{1}{\sigma_w^2}W_{xy} + \frac{T}{\sigma_w^2 + T\sigma_w^2}\bar{X}^{in'}\bar{y}^{in} + \frac{N}{\sigma_w^2 + N\sigma_v^2}\bar{X}^{bt'}\bar{y}^{bt}) \quad .$$

which are the same as the GLS estimates derived from the VC model.

The analysis above can be generalized to include specific invariant explanatory variables as in (2.1.4) or (3.1.5). With the similar arguments, the mixed estimation procedure for the CV model gives the GLS estimates as derived from the VC models in (2.1.16) - (2.1.17) or (3.1.17) - (3.1.19).

Thus we can conclude that the fixed effects and the random effects approaches are no different.

5. Conclusions

In this paper, we have discussed some problems related to the fixed effects and random effects linear models with panel data. We have suggested some estimation methods to estimate models with specific explanatory variables which do not vary over time or cross individual units in both fixed effects and random effects models. The extended covariance estimation procedure introduced in the fixed effects model is well justified under the view that effects are random from the outset and the covariance inference as a conditional inference conditional on the effects that are in the sample as introduced in Mundlak [1978]. Once the effects in the sample have been estimated by the covariance procedure, unconditional on the effects the coefficients of the invariant variables can then be estimated. Asymptotically, these two approaches are the same. For finite samples, if the variances of the disturbances are known, the GLS estimates have smaller variances than the extended covariance estimates with one exception. This exception characterizes and generalizes Mundlak's specification error approach [1978] in which the GLS estimates and covariance estimates are identical. However, the extended covariance approach may be preferred when the omitted variables problem which gives correlation between the effects and the included explanatory variables persists. In this case, GLS estimates are biased but not the extended covariance estimates. More efficient estimation procedures to estimate the coefficients of the specific variables in the CV models are also discussed and compare with the GLS estimates in the VC models. It turns out that some of these estimation procedures give the same estimates as the GLS for those variables.

On the presumptions that there were specification errors and a particular pattern of correlation between the effects and the explanatory variables, namely the within group means of all the explanatory variables in the model, Mundlak pointed out there would be no difference between the fixed effects and the random effects approaches. While he is correct, his approach is rather an extreme case. No matter whether there are specification errors or not, in this paper, we have shown that the two approaches are indeed no different. The covariance and GLS estimates are different because the traditional covariance estimation technique does not properly utilize available sample information and there is one step beyond which has been neglected. With the extended covariance estimation procedure which utilizes the within group means as specific invariant variables, and the mixed estimation techniques, the two approaches are shown to be no different at all.

In this paper, we have only considered the linear models. Whether the two approaches are the same for nonlinear models is still an open question. In some nonlinear models, such as the probit model, the variance components specification may impose very difficult estimation problems but the covariance specification may not be so (see Hall [1977]). For those models, the several extended covariance estimation techniques proposed can be extended and more efficient estimates can be derived. The detail investigation is beyond the scope of this paper and will be reported elsewhere.

References

Amemiya, T., (1976): "A Note on Random Coefficients Model," Technical Report no. 225, Institute for Mathematical Studies in the Social Sciences, Stanford University.

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, Vol. 34, pp. 585-612.

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

Hall, B. H., (1977): "A General Framework for Time-Series-Cross-Section Estimation", Paper presented at the INSEE Conference on "Econometrics of Panel Data" held in Paris, France, August, 1977.

Hausman, J. A., (1977): "Specification Tests in Econometrics", manuscript, Department of Economics, M.I.T., forthcoming in Econometrica.

Hoch, I., (1962): "Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data," Econometrica, Vol. 30, pp. 34-53.

Kuh, E., (1959): "The Validity of Cross Sectionally Estimated Behavior Equations in Time Series Applications," Econometrica, Vol. 27, pp. 197-214.

Lee, L. F. and M. M. Pitt, (1978): "Pooling Cross-Section and Time Series Data in the Estimation of Stochastic Frontier Production Function Models", Discussion Paper No. 98, Center for Economic Research, University of Minnesota.

Maddala, G. S., (1971): "The Use of Variance Components Models in Pooling Cross Section and Time Series Data," Econometrica, Vol. 39, pp 341-358.

Mundlak, Y., (1961): "Empirical Production Functions Free of Management Bias," Journal of Farm Economics, Vol 43, No. 1, pp. 44-56.

Mundlak, Y., (1978): "On the Polling of Time Series and Cross Section Data", Econometrica, Vol. 46, pp. 69-85.

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

Nerlove, M., (1971a): "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Section Data", Econometrica, Vol. 39, pp. 359-383.

Theil, H., and A. S. Goldberger, (1960): "On Pure and Mixed Statistical Estimation in Economics," International Economics Review, Vol. 2, pp. 65-78.

Theil, H., (1963): "On the Use of Incomplete Prior Information in Regression Analysis," Journal of American Statistical Association, Vol. 58, pp. 401-414.

Wachter, M. L., (1970): "Relative Wage Equations for U.S. Manufacturing Industries 1947-1967", Review of Economics and Statistics, Vol. LII, pp. 405-510.

Wallace, T. D., and Ashiq Hussain, (1969): "The Use of Error Components Models in Combining Cross Section with Time Series Data," Econometrica, Vol 37, pp. 55-68.

Weiss, Y., and L. A. Lillard, (1977): "Experience, Vintage and Time Effects in the Growth of Earnings: American Scientists, 1960-1970," NBER Working Paper #138, revised, Feb. 1977.

Zellner, A., (1962): "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," Journal of the American Statistical Association, 57, pp. 348-368.