LECTURE 6

Panel Data and the Analysis of Covariance

The Models

In the previous chapter, we have considered a model in the form of

$$y_{tj} = \mu + \gamma_t + \delta_j + \varepsilon_{tj}, \tag{1}$$

wherein $t = 1, \ldots, T$ and $j = 1, \ldots, M$ are, respectively, indices of temporal and spatial location. The index j is related to individual persons or to individual units of production such as farms or factories, and the index t corresponds to the time of observation.

There are various ways in which the information of the MT sample points can be formatted. First, there is the matrix format. In terms of the index notation, the TM equations are represented by

$$(y_{tj}e_t^j) = \mu(e_t^j) + (\gamma_t e_t^j) + (\delta_j e_t^j) + (\varepsilon_{tj}e_t^j). \tag{2}$$

Here, e_t^j denotes a basis matrix of order $T \times M$ that has a single unit in the tth row and the jth column and zeros elsewhere. The parentheses denote summations in respect of the indices of the basis. Thus, for example, (e_t^j) denotes a matrix of order $T \times M$ that has a unit in every position, whereas $(\gamma_t e_t^j)$ denotes a matrix in which the tth row contains M repetitions of the element γ_t.

In ordinary matrix notation, the set of TM equations becomes

$$Y = \mu\iota_T\iota_M' + \gamma\iota_M' + \iota_T\delta' + \mathcal{E}, \tag{3}$$

where $Y = [y_{tj}]$ and $\mathcal{E} = [\varepsilon_{tj}]$ are matrices of order $T \times M$, $\gamma = [\gamma_1, \ldots, \gamma_T]'$ and $\delta = [\delta_1, \ldots, \delta_M]'$ are vectors of orders T and M respectively, and ι_T and ι_M are vectors of units whose orders are indicated by their subscripts.

The second format that it is appropriate to consider represents a vectorised version of the matrix equation. In the index notation, this is denoted by

$$(y_{tj}e_{jt}) = \mu(e_{jt}) + (e_{jt}^{t})(\gamma_t e_t) + (e_{jt}^{j})(\delta_j e_j) + (\varepsilon_{tj}e_{jt}). \tag{4}$$

Using the notation of the Kronecker product, this can be rendered as

$$Y^c = \mu(\iota_M \otimes \iota_T) + (\iota_M \otimes I_T)\gamma + (I_M \otimes \iota_T)\delta + \mathcal{E}^c. \tag{5}$$

The latter can also be obtained by applying the rule

$$Y^c = (AXB')^c = (B \otimes A)X^c \tag{6}$$

to the matrix equation (3).

The model can now be elaborated by introducing a function $x_{t.}\beta_{.tj} = \sum_k x_{tk}\beta_{ktj}$ comprising K explanatory variables or regressors. Of course, if the elements β_{ktj} were to vary across all of the indices, then there would be no chance of making any reasonable inference about their values, unless some further assumptions could be made regarding the nature of this variation.

The Unrestricted Model

Without further ado, we shall make the assumption that $\beta_{ktj} = \beta_{kj}$ for all t, which is to say that there is no temporal variation in these coefficients. If, in addition, it can be assumed that $\gamma_t = 0$ for all t, then the model can be written as

$$\begin{aligned}(y_{tj}e_{jt}) &= \mu(e_{jt}) + (\delta_j e_{jt}) + (\{x_{tk}\beta_{kj}\}e_{jt}) + (\varepsilon_{tj}e_{jt})\\ &= \mu(e_{jt}) + (\delta_j e_{jt}) + (x_{tk}e_{jt}^{jk})(\beta_{kj}e_{jk}) + (\varepsilon_{tj}e_{jt}).\end{aligned} \tag{7}$$

Here, the braces which surround the expression $\{x_{tk}\beta_{kj}\}$ are to indicate that a sum has been taken over the repeated index k.

The set of T realisations of the jth equation can be written as

$$\begin{aligned}y_{.j} &= \mu\iota_T + \delta_j\iota_T + X\beta_{.j} + \varepsilon_{.j}\\ &= \alpha_j\iota_T + X\beta_{.j} + \varepsilon_{.j},\end{aligned} \tag{8}$$

where $\alpha_j = \mu + \delta_j$. This is a classical regression equation of the sort that can be estimated by ordinary least-squares regression. The full set of M such equations can be compiled to give the following system:

(9)

$$\begin{bmatrix} y_{.1}\\ y_{.2}\\ \vdots\\ y_{.M}\end{bmatrix} = \begin{bmatrix}\iota_T & 0 & \dots & 0\\ 0 & \iota_T & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \iota_T\end{bmatrix}\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_M\end{bmatrix} + \begin{bmatrix} X & 0 & \dots & 0\\ 0 & X & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & X\end{bmatrix}\begin{bmatrix}\beta_{.1}\\ \beta_{.2}\\ \vdots\\ \beta_{.M}\end{bmatrix} + \begin{bmatrix}\varepsilon_{.1}\\ \varepsilon_{.2}\\ \vdots\\ \varepsilon_{.M}\end{bmatrix}.$$

Using the Kronecker product, this can be rendered as

$$Y^c = (I_M \otimes \iota_T)\alpha + (I_M \otimes X)B^c + \mathcal{E}^c. \tag{10}$$

A useful elaboration of this model, which costs little in terms of added difficulty, is to allow the matrix X to vary between the M equations. Then, in place of the variables x_{tk}, there are elements x_{tkj} bearing the spatial subscript j. In that case, equation (9) is replaced by

$$
\begin{aligned}
\begin{bmatrix} y_{.1} \\ y_{.2} \\ \vdots \\ y_{.M} \end{bmatrix} &= \begin{bmatrix} \iota_T & 0 & \dots & 0 \\ 0 & \iota_T & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \iota_T \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_M \end{bmatrix} \\
&\quad + \begin{bmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & X_M \end{bmatrix} \begin{bmatrix} \beta_{.1} \\ \beta_{.2} \\ \vdots \\ \beta_{.M} \end{bmatrix} + \begin{bmatrix} \varepsilon_{.1} \\ \varepsilon_{.2} \\ \vdots \\ \varepsilon_{.M} \end{bmatrix}.
\end{aligned} \tag{11}
$$

It may be, for example, that the equations, which explain farm production in M regions, comprise explanatory variables whose measured values vary from region to region.

There is no obvious notation that will allow the structure of the matrix of explanatory variables to be expressed in a concise manner. However, the individual equations of (11) that are indexed by j and which pertain to a specific regions are separable. Thus, the generic the equation can be written as

$$
y_{.j} = \iota_T \alpha_j + X_j \beta.j + \varepsilon_{.j}. \tag{12}
$$

The Model with Individual Fixed Effects

Within the context of this model, there are some more restrictive hypotheses to be considered. The first of these, which is denoted by

$$
H_\beta : \beta_{.1} = \beta_{.2} = \cdots = \beta_{.M}, \tag{13}
$$

asserts that the slope parameters of all M of the regression equations are equal. This condition gives rise to a model in the form of

$$
\begin{bmatrix} y_{.1} \\ y_{.2} \\ \vdots \\ y_{.M} \end{bmatrix} = \begin{bmatrix} \iota_T & 0 & \dots & 0 \\ 0 & \iota_T & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \iota_T \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_M \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_M \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_{.1} \\ \varepsilon_{.2} \\ \vdots \\ \varepsilon_{.M} \end{bmatrix}, \tag{14}
$$

wherein each of the M equations is distinguished by having a particular value for the intercept. This equation can be rendered as

$$
Y^c = (I_M \otimes \iota_T)\alpha + X\beta + \mathcal{E}^c, \tag{15}
$$

where $X' = [X_1', X_2', \ldots, X_M']$. Equation (14) represents the starting point of many textbook accounts of panel data models, where it is common to write $(I_M \otimes \iota_T) = D$ for the so-called matrix of dummy variables associated with the intercept terms.

The Pooled Model

The second hypothesis, which is denoted by

$$(16) \qquad H_\alpha : \alpha_1 = \alpha_2 = \cdots = \alpha_M,$$

asserts that all of the intercepts have the same value. It is unlikely that one would ever wish to maintain this hypothesis without asserting H_β at the same time. The combined hypothesis $H_\gamma = H_\alpha \cap H_\beta$ gives rise to a pooled model in the form of

$$(17) \qquad \begin{bmatrix} y_{.1} \\ y_{.2} \\ \vdots \\ y_{.M} \end{bmatrix} = \begin{bmatrix} \iota_T \\ \iota_T \\ \vdots \\ \iota_T \end{bmatrix} \alpha + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_M \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_{.1} \\ \varepsilon_{.2} \\ \vdots \\ \varepsilon_{.M} \end{bmatrix}.$$

This equation can be rendered as

$$(18) \qquad Y^c = \iota_{MT}\alpha + X\beta + \mathcal{E}^c,$$

where, as before, $X' = [X_1', X_2', \ldots, X_M']$ and where ι_{MT} is a long vector consisting of MT units. This has the structure of the equation of an classical regression model, for which ordinary least-squares estimation is efficient.

The Least-Squares Estimates of the Models

Estimation of the Unrestricted Model

It will be assumed that the disturbances ε_{tj} are distributed independently and identically with $E(\varepsilon_{tj}) = 0$ and $V(\varepsilon_{tj}) = \sigma^2$ for all t, j. Under these assumptions, the equations of the unrestricted model represented by (11) are wholly separable, and the parameters of the jth equation may be estimated efficiently by ordinary least-squares regression.

The regression procedure can be applied directly to the generic equation of (11). Alternatively, the intercept term can be eliminated from the equation by taking the deviations of the data about their respective sample means. Traditionally, this approach has been preferred on the grounds that it results in more accurate computations.

In that case, the estimates of the parameters of the jth equation are

$$(19) \qquad \begin{aligned} &\hat{\beta}_{.j} = \left\{X_j'(I - P_T)X_j\right\}^{-1} X_j'(I - P_T)y_{.j} \quad \text{and} \\ &\hat{\alpha}_j = \bar{y}_j - \bar{x}_{j.}\hat{\beta}_{.j}, \end{aligned}$$

where

$$I - P_T = I - \iota_T(\iota_T'\iota_T)^{-1}\iota_T' \tag{20}$$

is the operator that transforms a vector of T observations into the vector of their deviations about the mean.

The residual sum of squares from the jth regression is given by

$$S_j = y_{.j}'(I - P_T)y_{.j} - y_{.j}'(I - P_T)X_j\{X_j'(I - P_T)X_j\}^{-1}X_j'(I - P_T)y_{.j}; \tag{21}$$

and, therefore, from the separability of the M regressions, it follows that the residual sum of squares, obtained from fitting the multi-equation model of (11) to the data, is just

$$S = \sum_j S_j. \tag{22}$$

The formulae of (19) are familiar from the treatment of the linear regression model $(y; i\alpha + X\beta, \sigma^2 I)$, which can be regarded as a particular instance of the partitioned model $(y; X_1\beta_1 + X_2\beta_2, \sigma^2 I)$.

It may be recalled that one way of developing the ordinary least-squares estimator of β_2 in the partitioned model depends on transforming the equation $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$ by the matrix $I - P_1$, where $P_1 = X_1(X_1'X_1)^{-1}X_1'$ is the orthogonal projector on the manifold of X_1. The effect of this transformation is to annihilate the term $X_1\beta_1$, which leads to the equation $(I - P_1)y = (I - P_1)X_2\beta_2 + (I - P_1)\varepsilon$. When ordinary least-squares regression is applied to the transformed equation, there is

$$\hat{\beta}_2 = \left\{X_2'(I - P_1)X_2\right\}^{-1}X_2'(I - P_1)y. \tag{23}$$

The same estimator may be derived by applying ordinary least-squares regression to the equation $Q'y = Q'X_2\beta_2 + Q'\varepsilon$, where Q is a matrix of orthonormal vectors such that $QQ' = (I - P_T)$. Since $D(Q'\varepsilon) = \sigma^2 I_{T-k_1}$, it follows that the equation fulfils the assumptions of the classical linear model; and a standard form of the Gauss–Markov theorem will serve to demonstrate the efficiency of the estimator $\hat{\beta}_2$.

In the case of system under (11), the intercept terms α_j are eliminated from the individual equations by premultiplying them by $(I - P_T)$. The intercept terms may be eliminated from the full system of equations by premultiplying it by

$$\begin{aligned} W &= I_M \otimes (I_T - P_T) = I_{MT} - (I_M \otimes P_T) \\ &= I_{MT} - D(D'D)^{-1}D'. \end{aligned} \tag{24}$$

Estimation of the Model with Individual Fixed Effects

Now, consider fitting the model under (14), which may be regarded as a variant of the model under (11) that has been subjected to the restrictions of H_β of (13), which asserts that the slope parameters of the regressions are the same in every region.

In this case, the efficient estimates are obtained by treating the system of equations as a whole; and it continues to be appropriate to take the data in deviation form. To eliminate the intercept terms, the individual equations of (14), which are of the form $y_{.j} = \alpha_j \iota_T + X_j\beta + \varepsilon_j$ are multiplied by the operator $I - P_T$, which creates deviations about sample means. The result is

$$(I - P_T)y_{.j} = (I - P_T)X_j\beta + (I - P_T)\varepsilon_j, \tag{25}$$

within which the tth equation is

$$y_{tj} - \bar{y}_j = (x_{.tj} - \bar{x}_{.j})\beta + (\varepsilon_{tj} - \bar{\varepsilon}_j). \tag{26}$$

To obtain an efficient estimate of β, the full set of TM mean-adjusted equations must be taken together. Once the estimate of β available, the individual intercept terms can be obtained. Thus, the efficient estimates of the parameters are given by the

$$\begin{aligned}\hat{\beta}_W &= \Big[\sum_j X_j'(I - P_T)X_j\Big]^{-1}\Big[\sum_j X_j'(I - P_T)y_j\Big]\\ &= \Big[X'\{I_M \otimes (I_T - P_T)\}X\Big]^{-1} X'\{I_M \otimes (I_T - P_T)\}y \quad \text{and}\\ \hat{\alpha}_j &= \bar{y}_{j.} - \bar{x}_j\hat{\beta}_W, \quad j = 1, \ldots, M,\end{aligned} \tag{27}$$

where $X' = [X_1', X_2', \ldots, X_M']$ and $y' = [y_{.1}' y_{.2}', \ldots, y_{.M}']$. The estimator $\hat{\beta}_W$ is the result of applying ordinary least-squares regression to an equation derived by premultiplying (14) by the projection matrix W of (24), which serves to annihilate the intercept terms.

The residual sum of squares from fitting the model of (14) is given by

$$\begin{aligned}S_\beta &= y'Wy - y'WX(X'WX)^{-1}X'Wy, \quad \text{where}\\ W &= I_M \otimes (I_T - P_T).\end{aligned} \tag{28}$$

The estimator $\hat{\beta}_W$ of (27) makes use only of the information conveyed by the deviations of the data points from their means within the M groups. For this reason, it is often called the *within-groups estimator.*

There exists another estimator, complementary to the *within-groups estimator,* that collapses the information within the groups by replacing it by the group means. The operator that averages over the sample of T observations is $P_T = \iota_T(\iota_T'\iota_T)^{-1}\iota_T'$. Applying it to the generic equation of (14) gives

$$\begin{aligned} P_T y_{.j} &= P_T\iota_T\alpha_{.j} + P_TX_j\beta + P_T\varepsilon_j \quad \text{or} \\ \iota_T\bar{y} &= \iota_T\alpha_j + \iota_T\bar{x}_{.j}\beta + \bar{\iota}_T\varepsilon_j. \end{aligned} \tag{29}$$

This equation comprises T redundant repetitions of the equation

$$\bar{y}_j = \alpha + \bar{x}_{.j}\beta + (\alpha_j - \alpha + \bar{\varepsilon}_j), \tag{30}$$

where $\alpha = \sum_j \alpha_j/M$ is a global or averaged intercept term, and where the deviation of the jth intercept α_j from global value has been joined with the averaged disturbance term. It may be observed that, by adding together equations (26) and (30), we recover the generic equation of the jth group.

Having gathering together the M equations of this form, an ordinary least-squares estimate of β, denoted $\hat{\beta}_B$ can be obtained which is described as the *between-groups estimator.* This estimator uses only the information conveyed by the variation amongst the M group means.

This would not constitute an efficient estimator of the slope parameters. However, the estimator has a purpose within context of the random effects model, to be described later.

The Pooled Estimator

Finally, consider fitting the model of (17) which, on the basis of the hypothesis H_γ, makes no distinction between the structures of the M equations. Let P_{MT} denote the projector $P_{MT} = \iota_{MT}(\iota_{MT}'\iota_{MT})^{-1}\iota_{MT}'$, where ι_{MT} is the summation vector of order MT. Then, the estimators of the parameters of the model can be written as

$$\begin{aligned} \hat{\beta}_G &= \{X'(I - P_{MT})X\}^{-1}X'(I - P_{MT})y \\ \hat{\alpha} &= \bar{y} - \bar{x}\hat{\beta}. \end{aligned} \tag{31}$$

The residual sum of squares from fitting the restricted model of (17) is given by

$$S_\gamma = (y - \alpha\iota_{MT} - X\hat{\beta})'(y - \alpha\iota_{MT} - X\hat{\beta}). \tag{32}$$

The Tests of the Restrictions

In order to test the various hypotheses, the following results are needed, which concern the distribution of the residual sum of squares from each of the regressions that have been considered:

$$\begin{aligned} &1. \quad \frac{1}{\sigma^2}S \sim \chi^2\{MT - M(K+1)\},\\ &2. \quad \frac{1}{\sigma^2}S_\beta \sim \chi^2\{MT - (K+M)\},\\ &3. \quad \frac{1}{\sigma^2}S_\gamma \sim \chi^2\{MT - (K+1)\}. \end{aligned} \tag{33}$$

The number of degrees of freedom in each of these cases is easily explained. It is simply the number of observations available in the vector $y' = [y'_{.1}, y'_{.2}, \ldots, y'_{.M}]$ less the number of parameters that are estimated in the particular model.

The hypothesis H_β can be tested by assessing the loss of fit that results from imposing the restrictions $\beta_1 = \beta_2 = \cdots = \beta_M$. The loss is given by $S_\beta - S$. The residual sum of squares S from the unrestricted model is the standard against which this loss is measured. The appropriate test statistic is therefore

$$F = \left\{ \frac{S_\beta - S}{(M-1)K} \bigg/ \frac{S}{MT - M(K+1)} \right\}, \tag{34}$$

which has a F distribution of $(M-1)K$ and $MT-M(K+1)$ degrees of freedom.

If the hypothesis H_β is accepted, then one might proceed to test the more stringent hypothesis $H_\gamma = H_\beta \cap H_\alpha$ which entails the additional restrictions of $H_\alpha : \alpha_1 = \alpha_2 = \cdots = \alpha_M$. The relevant test statistic in this case is given by

$$F = \left\{ \frac{S_\gamma - S_\beta}{M-1} \bigg/ \frac{S_\beta}{MT - (K+M)} \right\}, \tag{35}$$

which has a F distribution of $M-1$ and $MT-(K+M)$ degrees of freedom. The numerator of this statistic embodies a measure of the loss of fit that comes from imposing the additional restrictions of H_α.

The statistic of (35) tests the hypothesis H_γ within the context of an assumption that H_β is true. One might decide to test additionally, or even alternatively, the joint hypothesis $H_\gamma = H_\alpha \cap H_\beta$ within the context of the unrestricted model. The relevant statistic in that case would be given by

$$F = \left\{ \frac{S_\gamma - S}{(M-1)(K+1)} \bigg/ \frac{S}{MT - M(K+1)} \right\}. \tag{36}$$

The possibility has to be considered that, having accepted the hypotheses H_β and H_α on the strength of the values the F statistics under (34) and (35), we shall then discover that value of the statistic of (36) casts doubt on the joint hypothesis $H_\gamma = H_\beta \cap H_\alpha$. The possibility arises from the fact that critical region of the test of H_γ can never coincide with the critical region of the joint test implicit in the sequential procedure. However, if the critical value of the test H_γ has been appropriately chosen, then such a conflict in the results of the tests is an unlikely eventuality.

Models with Two-way Fixed Effects

The model of H_β can be elaborated by including the parameters γ_t which represent the temporal variation that is experienced by all J individuals. The equation of the model now assumes the form of

$$(37)\quad (y_{tj}e_{jt}) = \mu(e_{jt}) + (x_{tjk}e_{jt}^{k})(\beta_k e_k) + (e_{jt}^{t})(e_t\gamma_t) + (e_{jt}^{j})(e_j\delta_j) + (\varepsilon_{tj}e_{jt}).$$

Comparison with equation (7) reveals the present assumption that $\beta_{kj} = \beta_k$ for all k, which is to say that all j individuals share the same slope coefficients. Therefore, the system of equations as a whole can be represented by

$$(38)\qquad Y^c = \mu\iota_{MT} + X\beta + (\iota_M \otimes I_T)\gamma + (I_M \otimes \iota_T)\delta + \mathcal{E}^c.$$

The matrix $[\iota_{MT}, X, \iota_M \otimes I_T, I_M \otimes \iota_T]$, which contains the regressors of the model is, in fact, singular, by virtue of the linear dependence that exists between the columns of its submatrix $[\iota_{MT}, \iota_M \otimes I_T, I_M \otimes \iota_T]$. This dependence is made clear by writing the equation

$$(39)\qquad (\iota_M \otimes I_T)(1 \otimes \iota_T) = (I_M \otimes \iota_T)(\iota_M \otimes 1) = \iota_{MT}.$$

Therefore, the parameters μ, β, γ, δ will not be estimable as a whole unless some restrictions are introduced. It is natural to impose the conditions that $\iota_T'\gamma = \sum_t \gamma_t = 0$ and that $\iota_{MT}'\delta = \sum_j \delta_j = 0$.

To derive the ordinary least-squares estimate of β, we can begin by transforming equation (38) in such a way as to eliminate the parameters γ, δ, μ. This can be accomplished by premultiplying the equation by the matrix

$$(40)\qquad \begin{aligned} W &= [I_{MT} - (I_M \otimes P_T)][I_{MT} - (P_M \otimes I_M)] \\ &= I_{MT} - (I_M \otimes P_T) - (P_M \otimes I_M) + (I_M \otimes P_T)(P_M \otimes I_M). \end{aligned}$$

The two factors commute. The first factor $I_{MT} - (I_M \otimes P_T)$ has the effect of annihilating the term $(I_M \otimes \iota_T)\delta$ whilst the second factor $I_{MT} - (P_M \otimes I_M)$ has the effect of annihilating $(\iota_M \otimes I_T)\gamma$. However, since W is a symmetric idempotent matrix, it can be written in the form of $W = QQ'$, where Q is a matrix of order $MT \times (MT - M - T)$ consisting or orthonormal vectors. Therefore, the equation may, be transformed, with equal effect, by premultiplying it by Q' to obtain the system

$$Q'Y^c = Q'X\beta + \mathcal{E}^c. \tag{41}$$

The latter fulfils the assumptions of the classical linear model. It follows that the efficient estimator of β is given by

$$\begin{aligned}\hat{\beta} &= (X'QQ'X)^{-1}X'QQ'y \\ &= (X'WX)^{-1}X'Wy.\end{aligned} \tag{42}$$

Models with Random Effects

An alternative way of accommodating temporal and individual effects is to regard them as random variables rather than as fixed constants. To signify the difference in approach, the relevant equation will be denoted by

$$(y_{tj}e_{jt}) = \mu e_{jt} + (x_{jtk}e_{jt}^k)(\beta_k e_k) + (e_{jt}^t)(e_t\zeta_t) + (e_{jt}^j)(e_j\eta_j) + (\varepsilon_{jt}e_{jt}). \tag{43}$$

Here, ζ, which replaces γ, is a vector of random effects that vary through time and η, which replaces δ, is vector of effects that vary between individuals.

These effects are now part of the disturbance structure of the model and, as such, they must be uncorrelated with the systematic part of its structure if the ordinary methods of regression analysis are to be valid.

The advantage of the random-effects formulation is that it leads to a parsimonious parametrisation in which the effects are summarised by a pair of variance parameters in place of the TM coefficients of γ and δ. This economy, if it can be justified, should lead to more efficient estimates of the parameters of β.

A set of T realisations of all M equations is now written as

$$y = \mu\iota_{MT} + X\beta + (\iota_M \otimes I_T)\zeta + (I_M \otimes \iota_T)\eta + \varepsilon. \tag{44}$$

It is assumed that the random variables ζ_t, η_j and ε_{tj} are independently distributed with expectations of zero and with $V(\zeta_t) = \sigma_\zeta^2, V(\eta_j) = \sigma_\eta^2$ and $V(\varepsilon_{tj}) = \sigma_\varepsilon^2$. It follows that the dispersion matrix of the vector of disturbances in this model is given by

$$\Omega = \sigma_\zeta^2(\iota_M\iota_M' \otimes I_T) + \sigma_\eta^2(I_M \otimes \iota_T\iota_T') + \sigma_\varepsilon^2 I_{MT}. \tag{45}$$

A special case that is often considered arises when $\sigma_\zeta^2 = 0$, which is to say that there is no intertemporal variation in the structure of the stochastic disturbances. In that case, the dispersion matrix is of the form

$$\begin{aligned} \Omega &= \sigma_\eta^2(I_M \otimes \iota_T\iota_T') + \sigma_\varepsilon^2 I_{MT} \\ &= I_M \otimes (\sigma_\varepsilon^2 I_T + \sigma_\eta^2 \iota_T\iota_T) = I_M \otimes V. \end{aligned} \tag{46}$$

It can be confirmed by direct multiplication that the inverse of the matrix $V = \sigma^2\varepsilon I_T + \sigma_\eta^2\iota_T\iota_T$ is

$$V^{-1} = \frac{1}{\sigma_\varepsilon^2}\left(I_T - \frac{\sigma_\eta^2}{\sigma_\varepsilon^2 + T\sigma_\eta^2}\iota_T\iota_T'\right). \tag{47}$$

The inverse of the matrix Ω of (35) has a somewhat complicated structure. It takes the form of the form of

(48)

$$\Omega^{-1} = \frac{1}{\sigma_\varepsilon^2}\{I_{MT} - \lambda_1(\iota_M\iota_{MT}' \otimes I_T) + \lambda_2(I_M \otimes \iota_T\iota_T') + \lambda_3(\iota_M\iota_{MT}' \otimes \iota_T\iota_T)\}$$

$$\begin{aligned} \text{where}\quad \lambda_1 &= \sigma_\zeta^2(\sigma_\varepsilon^2 - M\sigma_\zeta^2)^{-1}, \\ \lambda_2 &= \sigma_\eta^2(\sigma_\varepsilon^2 - T\sigma_\eta^2)^{-1}, \\ \lambda_3 &= \lambda_1\lambda_2(2\sigma_\varepsilon^2 + M\sigma_\zeta^2 + T\sigma_\eta^2)(\sigma_\varepsilon^2 + M\sigma_\zeta^2 + T\sigma_\eta^2)^{-1}. \end{aligned}$$

Feasible Least-Squares Estimator of the Random Effects Model

In order to realise the generalised least squares estimators of the random effect models, it is necessary to derive estimators of the variances σ_ζ^2, σ_η^2 and σ_ε^2 of the error components. For simplicity, we shall continue to assume that $\sigma_\zeta^2 = 0$. Then, the appropriate estimator can be derived from the the *within-groups* and between groups estimators associated with the fixed effects model of equation (9). The estimators are

$$\begin{aligned} \hat\sigma_\varepsilon^2 &= \frac{(y - X\hat\beta_W)'(y - X\hat\beta_W)}{MT - (K+M)}, \\ \hat\sigma_B^2 &= \frac{(y - X\hat\beta_B)'(y - X\hat\beta_B)}{M - K} \quad \text{and} \\ \hat\sigma_\eta^2 &= \hat\sigma_B^2 - \frac{\hat\sigma_\varepsilon^2}{T}. \end{aligned} \tag{49}$$