Panel Data Models with Temporal Effects of Individual Characteristics

Young Hoon Lee*

In most of the panel data literature, the role of individual-variant intercepts is to control for unobservable individual specific effects. The unobservables which are represented by the individual effect should have influences on the dependent variable that are constant over time but varying over individuals. The primary focus of this study is on the construction of a regression model that allows time-varying effects of individual specific components on the dependent variable.

We discuss fixed effects and random effects and derive the estimators that are analogous to the within and GLS estimators of the standard panel data model. We derive the asymptotic properties of the generalized within and GLS estimators. Furthermore, we construct test statistics for the hypothesis that the individual effect has a constant coefficient over time (*JEL* Classification; C23).

I. Introduction

Panel data are data that have both a cross-sectional and a time-series dimmension. Panel data are potentially useful for several reasons. At the most basic level, observing each individual repeatedly is a way of increasing the total number of observations. Also, some parameters may be estimated more readily from cross-sectional information and others from time-series information. For example, in budget studies it is often argued that prices display little cross-sectional variation, so that precise estimation of price elasticities requires time-series information, while real incomes display little temporal variation, so

*Korea Economic Research Institute. 28-1, Yoido-dong, Youngdeungpo-ku, Seoul, 151-010, FAX. 785-0270, TEL. 786-6572. This is based on a part of my Ph.D. dissertation submitted to the Michigan State University, 1991. I am indebted to Peter Schmidt for helpful advice at various stages of the research. However, all errors are of my own making.
[**Seoul Journal of Economics** 1994, Vol. 7, No. 1]

that precise estimation of income elasticities requires cross-sectional information. Panel data contain both types of information and therefore may be very useful.

However, in this study we will be concerned specifically with techniques that are useful when N is large and T is small. Such cases are common in labor economics as well as consumption and investment studies, since many longitudinal data sets contain thousands of individuals but only a few time periods of data per individual. In such cases the usual motivation for the use of panel data is to control for possible biases due to unobservable individual characteristics. For example, Mundlak (1961) considered a Cobb-Douglas production function for farms, and was concerned about possible biases due to differences across farms in soil quality, an unobserved variable that affects output and may be correlated with the inputs. More recently, many labor economists have estimated wage equations and have been concerned with possible biases due to differences across individuals in unobserved ability.

The existing panel data literature has dealt extensively with the problem of avoiding biases due to unobservables like soil quality or ability, by assuming the unobservables to be time invariant. The standard model (henceforth, called 'the simple model') that is used is the regression model with individual effects;

$$Y_{it} = X_{it}'\beta + \alpha_i + \varepsilon_{it} \quad i = 1,\ldots,N, \quad t = 1,\ldots,T. \tag{1}$$

Here Y_{it} is the dependent variable; X_{it} is a $K \times 1$ vector of explanatory variables; β is a $K \times 1$ vector of parameters (regression coefficients); α_i is the unobserved individual effect, which is time invariant (does not depend on t); and ε_{it} is the random error. The errors ε_{it} are assumed to be independently and identically distributed (i.i.d.) with $E(\varepsilon_{it}) = 0$ and $Var(\varepsilon_{it}) = \sigma^2$.

In this model the unobserved individual characteristics represented by the individual effect α_i are assumed to have the same effect on the dependent variable Y_i in all time periods. The motivation for this study is that this assumption is unnecessarily strong, and we will relax it. Specifically, we will allow the effect of α_i on Y_i to vary over time, though we will require that the temporal pattern of the effect of α_i on Y_i must be the same for all individuals. On the other point of view, we may consider α_i as a response of each individual to time-varying changes (for example, macro shocks). Specially, we will consider the model (henceforth, called 'the general model').

$$\begin{aligned} Y_{it} &= X_{it}'\beta + \delta_{it} + \varepsilon_{it} \\ &= X_{it}'\beta + \theta_t\alpha_i + \varepsilon_{it} \quad i = 1,\ldots,N, \quad t = 1,\ldots,T, \end{aligned} \tag{2}$$

where $\delta_{it} = \theta_t\alpha_i$.

The use of this general model could improve empirical studies in several area. In the stochastic frontier estimation, Schmidt and Sickles (1984) applied the simple panel data model (without θ_t in (2) so that equal to (1)) in which interfirm differences in time-invariant individual effects were interpreted as measures of technical inefficiency. However, they needed the assumption that technical inefficiency is time-invariant. A main focus in this area after Schmidt and Sickles has been on how to relax this strong assumption. Kumbhakar (1990) and Cornwell, Schmidt and Sickles (1990) have proposed panel data models that allow technical inefficiency to change over time, but rather structured ways. Kumbhakar specified a model in which $\delta_{it} = r(t)\ \alpha_i$ where $r(t) = [1 + \exp\ (bt + ct^2)]^{-1}$ (a specific function of time). Cornwell, Schmidt, and Sickles also suggested the specific case that α_{it} is quadratic in time, so that $\delta_{it} = \alpha_{i0} + \alpha_{i1}\ t + \alpha_{i2}\ t^2$. The use of this general panel data model (eq (2)) in a stochastic frontier estimation allows more flexibility in the way that technical inefficiency changes over time and this is a fairly direct relaxation of Kumbhakar's model and the model of Cornwell, Schmidt and Sickles as well as the model of Schmidt and Sickles. The general model includes Kumbhakar's model as a special case when $\theta_t = r(t)$. Models with multiple components ($\delta_{it} = \sum_{j=1}^{G}\theta_{tj}\ \alpha_{ij}$) are also identified and could be estimated(see Appendix B). With G =3, the model of Cornwell, Schmidt, and Sickles corresponds to $\theta_{t1} = 1$, $\theta_{t2} = $ t, and $\theta_{t3} = t^2$ for all t and the simple model obviously corresponds to $\theta_t = 1$ for all t. All three models are special cases of this general model and are testable.

Panel data models are widely used in studying liquidity constraints of consumption and corporate investment. Zeldes (1989) and Whited (1992) tested liquidity constraints of consumption and of corporate investment, respectively, both by using Euler's equation. Whited argued that macro shocks, which are identical for all individuals but time-variant, could influence on the behavior of investment. He derived the Euler's equations of investment with macro shock effects (macro shock effects = θ_t without α_i in eq(2)).

An assumption are imposed on his model that each individual's response to a macro shock is equal to one another. This assumption may not be true in the real world, but could be relaxed simply by using the general panel data model since θ_t represents a macro shock at t and α_i represents the individuals i's response to a macro shock. We may also test the hypothesis that responses to a macro shock are identical for all individual. The model of Zeldes can also be improved by the same way as Whited in studying liquidity constraint of consumption.

This model requires a normalization, and we set $\theta_1 = 1$. Compared to the model (1), the new model introduces the $(T-1)$ new parameters θ_2, $\theta_3,\ldots,\theta_T$ to represent the effect of α_i on Y_{it} for $t=2, 3,\ldots,T$ relative to the effect of α_i on Y_{i1}.

As a matter of notation, let $Y_i = (Y_{i1}, Y_{i2},\ldots,Y_{iT})'$, $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2},\ldots,\varepsilon_{iT})'$ and $X_i = (X'_{i1}, X'_{i2},\ldots, X'_{iT})'$, each representing the T observations for person i. Then we can write equation (2) as

$$Y_i = X_i\beta + \xi\alpha_i + \varepsilon_i \quad i = 1,\ldots,N \tag{3}$$

where $\xi = (1,\theta')'$, $\theta = (\theta_2, \theta_3,\ldots,\theta_T)'$.

The simple model (1) thus corresponds to the case that $\theta_2 = \theta_3\ldots = \theta_T = 1$, or equivalently that θ(or ξ) is a vector of ones. As we shall see, this is a testable proposition in our model.

The model we consider can also be compared to the two-way analysis of covariance model that includes both individual and time effects. That model can be written as

$$Y_{it} = X'_{it}\beta + \alpha_i + \theta_t + \varepsilon_{it} \quad i = 1,\ldots,N, \quad t = 1,\ldots,T. \tag{4}$$

The number of parameters in (4) is exactly the same as in our model (3), but the models are different. Our interpretation of (4) is that it is suitable in cases in which there are relevant unobservable variables that vary over time but not over individuals; it does not handle the case that our model is designed for, in which the effects of unobservable individual characteristics vary over time. Compared to the two-way analysis of covariance model (4), our model (3) is more difficult to estimate, because it is nonlinear. However, unlike the analysis of covariance model, our model allows for inclusion of observables that are time invariant or invariant over individuals, a considerable advantage in some applications.

We can also model the case of several possible interactions between time-invariant and individually invariant parameters, as in the model.

$$Y_i = X_i\beta + \sum_{g=1}^{G} \xi_g\alpha_{ig} + \varepsilon_i \quad i = 1,\ldots, N. \tag{5}$$

(4) is nested in (5) with G=2. In the area of liquidity constraints of

firm, (4) are well used as an Euler's equation. α_i and θ_t represents firm's characteristics and macro shocks, respectively. However, this model assumes that each firm responses identically to a macro shock. (5) makes (4) be more realistic since $\alpha_{2i}\theta_{2t}$ means the firms i's response (α_{2i}) to a macro shock (θ_{2t}).

The plan of this study is as follows. Section II discusses the fixed effect model in which the parameters α_i are treated as fixed. We derive a generalized within estimator, and we show its consistency and asymptotic distribution. Section III discusses the random effects model in which the individual effects α_i are treated as random. We derive the appropriate GLS estimator and prove that it is more efficient than the within estimator. Section IV considers tests of the hypothesis that ξ is a vector of ones, so that our model reduces to the simple panel data model. We present Lagrange Multiplier(*LM*), Likelihood ratio(*LR*) and Wald statistics for this hypothesis. Finally, section V gives our concluding remarks.

II. Fixed Effects

We may rewrite (3) with all *NT* observations as

$$Y = X\beta + (I_N \otimes \xi)\alpha + \varepsilon^1. \tag{6}$$

If (6) is the true relationship and $\xi \neq e_T$ where e_T is a $T \times 1$ vector of ones, the estimates of β from the simple model are not unbiased, since

$$E(\hat{\beta}_w) = \beta + [X'(I_N \otimes M_{e_T})X]^{-1} X'(I_N \otimes M_{e_T})(I_N \otimes \xi)\alpha \neq \beta.$$

Thus we expect the simple within estimates to be biased for the coefficients of those variables whose temporal variation is correlated with the temporal variation in the effect of α on Y.

The generalization of the within transformation is to premultiply (6) by the idempotent matrix $(I_N \otimes M_\xi)$ that is defined as $M_\xi = I_T - P_\xi$ and $P_\xi = \xi(\xi'\xi)^{-1}\xi'$.

That is, the transformed regression model is expressed by

$$(I_N \otimes M_\xi)Y = (I_N \otimes M_\xi)X\beta + (I_N \otimes M_\xi)\varepsilon, \tag{7}$$

since $M_\xi \xi = 0$. The individual effects are deleted by taking deviations

[1]A distributional assumption for $\varepsilon_{it}(\varepsilon_{it}$: i.i.d. $N(0, \sigma^2))$ is necessary since the fourth moment of ε_{it} appears in the calculation of the covariance matrix of the estimator.

from individual weighted means ($P_\xi Y$ and $P_\xi X$) instead of taking differences from individual means in the simple model.

We may not apply OLS to (7) since $M_\xi Y_i$ and $M_\xi X_i$ are not observables; $M_\xi Y_i$ and $M_\xi X_i$ include the parameter vector ξ. Instead, we construct an objective function which will be minimized with respect to β and θ. This objective function is simply the error sum of squares of (7):

$$CSSE = \sum_{i=1}^{N} (Y_i - X_i\beta)' M_\xi (Y_i - X_i\beta). \tag{8}$$

The reason that we denote this objective function *CSSE* is that it is the same as the (concentrated) error sum of squares of (6). By taking derivatives of (8) with respect to β and θ the first order conditions are obtained as

$$\frac{\partial CSSE}{\partial \beta} = -2\sum_{i=1}^{N} X_i' M_\xi (Y_i - X_i\beta) = 0 \tag{9}$$

$$\begin{aligned}\frac{\partial CSSE}{\partial \theta} = -\frac{2}{\xi'\xi}\Bigg[&\sum_{i=1}^{N} (Y_i - X_i\beta)'\xi(Y_{i.} - X_{i.}\beta) \\ &- \sum_{i=1}^{N} (Y_i - X_i\beta)' P_\xi (Y_i - X_i\beta)\theta\Bigg] = 0\end{aligned} \tag{10}$$

where $Y_{i.} = (Y_{i2}, Y_{i3}, \ldots, Y_{iT})'$, $X_{i.} = (X_{i2}', X_{i3}', \ldots, X_{iT}')'$.

The solutions of the first order conditions are the following;

$$\tilde{\beta}_w = (X'(I_N \otimes M_{\xi_w})X)^{-1} X'(I_N \otimes M_{\xi_w})Y \tag{11}$$

$$\tilde{\xi}_w = (1, \tilde{\theta}_w')' \text{ is an eigenvector of } \sum_{i=1}^{N} (Y_i - X_i\tilde{\beta}_w)(Y_i - X_i\tilde{\beta}_w)'. \tag{12}$$

***Note* 1**

For a matrix A, suppose that λ is an eigenvalue and x is the corresponding eigenvector. Then, $Ax = \lambda x$ and $x' Ax = \lambda$.

***Lemma* 1**

ξ_w is the eigenvector corresponding to the largest eigenvalue.

***Proof*:**

$$\begin{aligned} CSSE &= \sum_{i=1}^{N} (Y_i - X_i\tilde{\beta}_w)' M_\xi (Y_i - X_i\tilde{\beta}_w) \\ &= \sum_{i=1}^{N} (Y_i - X_i\tilde{\beta}_w)'(Y_i - X_i\tilde{\beta}_w) - \frac{1}{\xi_w'\xi_w}\xi_w' \sum_{i=1}^{N} (Y_i - X_i\tilde{\beta}_w)(Y_i - X_i\tilde{\beta}_w)'\xi_w. \end{aligned} \tag{13}$$

By Note 1, (13) can be rewritten as

$$CSSE = \sum_{i=1}^{N}(Y_i - X_i\hat{\beta}_w)(Y_i - X_i\hat{\beta}_w)' - \hat{\lambda}_w \tag{14}$$

where $\hat{\lambda}_w$ is the estimated eigenvalue. We pick the largest eigenvalue since we wish to minimize *CSSE*.

Q.E.D.

The solutions for $\tilde{\beta}_w$ and $\tilde{\theta}_w$ are not closed forms of the data, since the solution for $\tilde{\beta}_w$ depends on $\tilde{\theta}_w$ and vice versa. However, these can be calculated by iteration starting with any initial value of $\tilde{\beta}_w$. The estimate $\tilde{\beta}_w$ from the simple model is a good candidate for the initial value.

For the proof of consistency and asymptotic normality of $\tilde{\beta}_w$ and $\tilde{\theta}_w$, we need two theorems provided by Amemiya (1985).[2]

Theorem 1

Make the Assumptions:

(A) The parameter space Θ is a compact subset of the Euclidean K-space (R^K), and the true value θ_0 is in Θ

(B) $Q_N(y, \theta)$ is continuous in $\theta \in \Theta$ for all y and is a measurable function of y for all $\theta \in \Theta$

(C) $N^{-1}Q_N(\theta)$ converges to a nonstochastic function $Q(\theta)$ in probability uniformly in $\theta \in \Theta$ as N goes to ∞ and $Q(\theta)$ attains an unique global maximum at θ_0.

Define $\hat{\theta}_N$ as a value that satisfies

$$Q_N(\hat{\theta}_N) = \underset{\theta \in \Theta}{\text{Max}}\, Q_N(\theta).$$

Then $\hat{\theta}_N$ converges to θ_0 in probability.

Theorem 2

Assumé:

(α) lim $(1/N)\, \Sigma X_i' X_i$ exists and is finite and nonsingular.

(β) lim $(1/N)\, \Sigma \alpha_i^2$ exists and is finite and nonzero.

Then, $\tilde{\beta}_w$ and $\tilde{\xi}_w$ which satisfy

$$CSSE(\tilde{\beta}_w, \tilde{\xi}_w) = \underset{\beta \ \ \theta}{\text{Min}} \sum_{i=1}^{N}(Y_i - X_i\beta)' M_\xi (Y_i - X_i\beta)$$

[2]See Amemiya (1985), p. 105-14.

are consistent.

Proof: The proof that the assumptions (A) and (B) in Theorem 1 hold in this model is omitted since it is trivial. With (α) and (β), it is shown in Appendix A that

$$\text{p}\lim \frac{1}{N} CSSE(\beta, \xi) = (\beta_0 - \beta)' \mathcal{Q}_X (\beta_0 - \beta) + \mathcal{Q}_\alpha \xi_0' M_\xi \xi_0 + 2(\beta_0 - \beta)' \mathcal{Q}_{X\alpha} \xi_0 + (T-1)\sigma^2 \tag{15}$$

where
$$\mathcal{Q}_X = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i' M_\xi X_i$$
$$\mathcal{Q}_\alpha = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \alpha_{0i}^2$$
$$\mathcal{Q}_{X\alpha} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i' M_\xi \alpha_{0i}.$$

Define a compact parameter space Θ by $\beta' \beta < C_1$ and $\xi' \xi \leq C_2$ where C_1 and C_2 are large positive constants and assume $(\beta_0', \xi_0')'$ is an interior point of Θ. Then N^{-1} $CSSE$ (β, ξ) converges to (15) uniformly in probability.

Now, we need show that plim $(1/N)$ $CSSE$ (β, ξ) attains an unique global minimum at (β_0, ξ_0). (15) can be written as

$$\text{p}\lim \frac{1}{N} CSSE(\beta, \xi) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} [X_i(\beta_0 - \beta) + \xi_0 \alpha_{0i}]' M_\xi [X_i(\beta_0 - \beta) + \xi_0 \alpha_{0i}] + (T-1)\sigma^2. \tag{16}$$

Notice that the first term in (16) is positive since it is quadratic and this term is zero at and only at (β_0, ξ_0): plim $(1/N)$ $CSSE$ $(\beta_0, \xi_0) = (T-1)\sigma^2$. Thus plim$(1/N)$ $CSSE$ attains an unique global minimum at (β_0, ξ_0) and assumption (C) holds. Using Theorem 1, $\tilde{\beta}_w$ and $\tilde{\xi}_w$ which minimize the objective function converge to true β_0 and ξ_0 in probability as N goes infinity.

Q.E.D.

Using the following theorem by Amemiya (1985), we may derive the asymptotic normality.

Theorem 3

Make the following assumptions in addition to the assumptions of

Theorem 1.

(AA) $\partial^2 Q_N / \partial\theta\partial\theta'$ exists and is continuous in an open, convex neighborhood of θ_0.

(BB) $N^{-1} (\partial^2 Q_N / \partial\theta\partial\theta')_{\theta^*}$ converges to a finite nonsingular matrix $A(\theta_0) = \lim EN^{-1} (\partial^2 Q_N / \partial\theta\partial\theta')_{\theta_0}$ in probability for any sequence θ_N^* such that plim $\theta_N^* = \theta_0$.

(CC) $N^{-1/2}(\partial Q_N / \partial\theta)_\theta \to N(0, B(\theta_0))$, where $B(\theta_0) = \lim E\, N^{-1}(\partial Q_N / \partial\theta)_{\theta_0} (\partial Q_N / \partial\theta')_{\theta_0}$.

Let $\{\hat{\theta}_N\}$ be a sequence obtained by choosing one element from $\hat{\theta}_N$ defined in Theorem 1 such that plim $\hat{\theta}_N = \theta_0$. (We call $\hat{\theta}_N$ a consistent root).

Then, $\sqrt{N}(\hat{\theta}_N - \theta_0) \to N[0, A(\theta_0)^{-1} B(\theta_0) A(\theta_0)^{-1}]$

Applying Theorem 3, we have

$$A(\lambda_0) = \lim_{N\to\infty} E\frac{1}{N}\left(\frac{\partial CSSE}{\partial\lambda\partial\lambda'}\right)_{\lambda_0}$$

$$= 2\begin{bmatrix} Q_X & \lim\limits_{N\to\infty} E\frac{1}{N}\left(\sum\limits_{i=1}^N X_{i.}'\alpha_{0i} - \frac{1}{\xi_0'\xi_0}\sum\limits_{i=1}^N X_i'\xi_0\theta_0\alpha_{0i}\right) \\ \bullet & Q_\alpha\left(I_{T-1} - \frac{\theta_0\theta_0'}{\xi_0'\xi_0}\right)\end{bmatrix}$$

a finite nonsingular matrix, and

$$B(\lambda_0) = \lim_{N\to\infty} E\frac{1}{N}\left(\frac{\partial CSSE}{\partial\lambda}\right)_{\lambda_0} \times \left(\frac{\partial CSSE}{\partial\lambda'}\right)_{\lambda_0}$$

$$= 4\alpha^2\begin{bmatrix} Q_X & \lim\limits_{N\to\infty} \frac{1}{N}\left(\sum\limits_{i=1}^N X_{i.}'\alpha_{0i} - \frac{1}{\xi_0'\xi_0}\sum\limits_{i=1}^N X_i'\xi_0\theta_0\alpha_{0i}\right) \\ \bullet & \left(\frac{\sigma^2}{\xi_0'\xi_0} + Q_\alpha\right)\left(I_{T-1} - \frac{\theta_0\theta_0'}{\xi_0'\xi_0}\right)\end{bmatrix}$$

where $\lambda = (\beta', \theta')'$ and $\lambda_0 = (\beta_0', \theta_0')'$.

Therefore,

$$\sqrt{N}\begin{bmatrix}\tilde{\beta}_w - \beta_0 \\ \tilde{\theta}_w - \theta_0\end{bmatrix} \to N[0, A(\lambda_0)^{-1}B(\lambda_0)A(\lambda_0)^{-1}].$$

An advantage of the general model over the simple model (1) is the ability to include time-invariant explanatory variables. To see this, consider first the simple model with time-invariant regressors Z_i added:

$$Y = X\beta + (Z \otimes e_T)\gamma + G\alpha + \varepsilon \tag{17}$$

where $Z = (Z_1', Z_2' \dots Z_N')'$, $G = I_N \otimes e_T$.

Premultiplying by $M_G = I_{NT} - (I_N \otimes (e_T e_T')/T)$, the transformed regression model

$$\begin{aligned} M_G Y &= [M_G X \;\; M_G(Z \otimes e_T)]\begin{bmatrix}\beta \\ \gamma\end{bmatrix} + M_G \varepsilon \\ &= [M_G X \;\; 0]\begin{bmatrix}\beta \\ \gamma\end{bmatrix} + M_G \varepsilon = (M_G X)\beta + M_G \varepsilon \end{aligned} \tag{18}$$

since $M_G(Z \otimes e_T) = 0$[3]. This is the reason why we can not incorporate time-invariant explanatory variables into a fixed effects model.

This problem does not arise in our general model. The equation for the general model corresponding to (17) is

$$\begin{aligned} Y &= X\beta + (Z \otimes e_T)\gamma + (I_N \otimes \xi)\alpha + \varepsilon \\ &= [X \;\; (Z \otimes e_T)]\begin{bmatrix}\beta \\ \gamma\end{bmatrix} + (I_N \otimes \xi)\alpha + \varepsilon. \end{aligned} \tag{19}$$

The within transformation leads (19) to

$$(I_N \otimes M_\xi)Y = [(I_N \otimes M_\xi)X \;\; (Z \otimes M_\xi e_T)]\begin{bmatrix}\beta \\ \gamma\end{bmatrix} + (I_N \otimes \xi)\varepsilon. \tag{20}$$

But $(I_N \otimes M_\xi)(Z \otimes e_T) = Z \otimes M_\xi e_T$ is generally not equal to zero unless $\xi = e_T$. Therefore, the inclusion of time-invariant regressors[4] is allowed, and their coefficients can be estimated consistently. This is an advantage of the general model since time-invariant explanatory variables are often important in many applications. For instance, in a wage equation, years of schooling, race, union status or sex could be important determinants of the wage. Notice that the overall intercept is also identified

[3]
$$\begin{aligned} M_G(Z \otimes e_T) &= [I_{NT} - (I_N \otimes (e_T e_T')/T)](Z \otimes e_T) \\ &= (Z \otimes e_T) - [Z \otimes (e_T e_T' e_T)/T) \\ &= (Z \otimes e_T) - (Z \otimes e_T) = 0. \end{aligned}$$

[4] The simple between estimator has to exclude individual-invariant explanatory variables for the same reason that the simple within estimation cannot include time-invariant regressors. However, those regressors can be included in *the general between estimate.*

in the general model while not in the simple model.

In the simple model with fixed effects, assuming the normality of the ε_{it}, the conditional maximum likelihood estimator (CMLE) is equal to the within estimator. Furthermore, the MLE is the same as the CMLE (or within estimator). Thus the incidental parameters problem is not relevant in the simple model.

The above results do not hold in the general model. The same type of derivation as in the simple model for the CMLE can not be applied in the general model. The individual mean, $\bar{Y}_i$, is a sufficient statistic for α_i in the simple model and the likelihood conditional on $\bar{Y}_i$ does not depend upon incidental parameter α_i. However, $P_\xi Y_i$ in the general model corresponds to $\bar{Y}_i$ in the simple model, and it is not a sufficient statistic since it is not a function of only the data. A parameter ξ is included in $P_\xi Y_i$.

It is impossible to have many different individual effects in the simple model since they would not be identified. However, we may include a number of individual specific components in the general regression mode. For details, see Appendix B.

We have discussed a generalization of the conventional fixed effects model that allows different time-effects of individual specific components on the dependent variable. We derive a consistent estimator of the regression coefficients (β) and of the coefficients of the individual effects (ξ) using the conventional within transformation. We noted that the coefficients of time-invariant explanatory variables, which cannot be estimated in the simple model, may be estimated consistently. The inclusion of several individual specific components in the regression model is also introduced, and the results are similar to those with one individual specific component. Unlike the simple model, the asymptotic theory of this model does not agree with normal likelihood theory. The sufficient statistic for the individual effects depends on other parameters, and so the CMLE cannot be obtained by the usual method (see Chamberlain 1980) by conditioning on a sufficient statistic. The MLE is consistent, but this must be proved directly, and the usual formula for its asymptotic covariance matrix(the inverse of the information matrix) does not apply.

III. Random Effects

An alternative approach in panel data models is to assume that the individual components are random. That is to say, random effects models consider the individual effects to be independently identically distributed and to be independent of the disturbance and the explanatory variables.

Hsiao (1985)[5] mentions the difference between fixed effects models and random effects models. The fixed effects model is regarded as providing inference conditional on the effects in the sample, whereas the random effects model is regarded as providing unconditional inference with respect to the population of effects.

The within estimator does not consider variation between individuals. The GLS estimator used in the random effects model considers both variation between individuals and variation over time within each individual. Therefore, the GLS estimator can be expressed as a combination of the within and the between estimators. The GLS estimator is more efficient than the within estimator because of the utilization of the variation between individuals.

The regression equation (2) is considered as

$$Y_{it} = X_{it}\beta + v_{it} \qquad i = 1,\dots,N, \quad t = 1,\dots,T \tag{21}$$

where $v_{it} = \theta_t\alpha_i + \varepsilon_{it}$.
We let $E(\alpha_i) = \mu$ and assume that $\alpha_i^* = \alpha_i - \mu$ is i.i.d. with $Var(\alpha_i^*) = \sigma_\alpha^2$ and α is uncorrelated with X.

The covariance matrix of the error term is taken into account in GLS estimation. The covariance structure of v is as follows:

$$\Sigma^{-1} = \frac{1}{\sigma^2}(I_{NT} - (1 - q^2)(I_N \otimes P_\xi)) \tag{22}$$

where $q^2 = \dfrac{\sigma^2}{\sigma^2 + \xi'\xi\sigma_\alpha^2}$.

GLS can be obtained by OLS applied to the transformed regression model

$$\Sigma^{-1/2}\, Y = \Sigma^{-1/2}\, X\beta + \Sigma^{-1/2}(e_N \otimes \xi)\mu + \Sigma^{-1/2}v \tag{23}$$

where $\Sigma^{-1/2} = (1/\sigma)(I_{NT} - (1 - q)(I_N \otimes P_\xi)) = (1/\sigma)(I_N \otimes M_\xi) + q(I_N \otimes P_\xi)$ and e_N is a $N \times 1$ vector of ones. This transformation is a combination of the within and the between transformations. For example, $\sigma\Sigma^{-1/2}Y = (I_N \otimes M_\xi)Y + q(I_N \otimes P_\xi)Y$.

Since $\Sigma^{-1/2}$ includes the parameter vector θ, we cannot simply apply OLS to (23). We will derive the GLS estimator of β and θ which mini-

[5]See Hsiao (1985), p. 131.

mizes the objective function, equal to the error sum of squares of the transformed equation (23).

That is, we wish to minimize

$$SSE = [Y - X\beta - (e_N \otimes \xi)\mu]' \, [I_{NT} - (1 - q^2)(I_N \otimes P_\xi)][Y - X\beta - (e_N \otimes \xi)\mu]. \quad (24)$$

From the derivative of SSE with respect to μ, we obtain

$$\tilde{\mu}_{\mathrm{GLS}} = \frac{1}{N\xi' \xi} (e_N \otimes \xi)' \; (Y - X\beta). \quad (25)$$

Substituting (24) into (25), we obtain the concentrated SSE

$$CSSE = (Y - X\beta)'[(I_N \otimes M_\xi) + q^2 (M_{e_N} \otimes P_\xi)](Y - X\beta) \quad (26)$$

where $M_{e_T} = \underset{\sim}{I_N} - e_N \underset{\sim}{e_N'} / N.$

The values $\tilde{\beta}_{\mathrm{GLS}}$ and $\tilde{\theta}_{\mathrm{GLS}}$ which minimize $CSSE$ are derived by taking derivatives of $CSSE$ with respect to β and θ and setting them to zero. This gives

$$\begin{aligned}\tilde{\beta}_{\mathrm{GLS}} &= (X'[(I_N \otimes M_\xi) + q^2 (M_{e_N} \otimes P_\xi)]X)^{-1} \cdot X'[(I_N \otimes M_\xi) \\ &\quad + q^2 (M_{e_N} \otimes P_\xi)]Y\end{aligned} \quad (27)$$

$$\begin{aligned}\xi'_{\mathrm{GLS}} \text{ is eigenvector of } &\sum_{i=1}^{N} [\sqrt{1-q^2}\, e_i + (1 - \sqrt{1-q^2})\bar{e}] \\ &[\sqrt{1-q^2}\, e_i + (1 - \sqrt{1-q^2})\bar{e}]'\end{aligned} \quad (28)$$

where $e_i = Y_i - X_i \tilde{\beta}_{\mathrm{GLS}}$, $\bar{e} = \frac{1}{N} \sum_{i=1}^{N} (Y_i - X_i \tilde{\beta}_{\mathrm{GLS}})$.

The proof that $\tilde{\xi}_{\mathrm{GLS}}$ is the eigenvector corresponding to the largest eigenvalue is essentially the same as the proof of Lemma 1. Similarly, the asymptotic properties of $\tilde{\beta}_{\mathrm{GLS}}$ and $\tilde{\xi}_{\mathrm{GLS}}$ are derived using Theorem 1 & 3 as before, and we obtain.

$$\sqrt{N} \begin{bmatrix} \tilde{\beta}_{\mathrm{GLS}} - \beta_0 \\ \tilde{\theta}_{\mathrm{GLS}} - \theta_0 \end{bmatrix} \rightarrow N[0, A^{-1} B A^{-1}]. \quad (29)$$

The matrix A comes from the second derivatives of $CSSE$ while B is derived from the cross-products of the first derivatives. These are $(K + T - 1) \times (K + T - 1)$ matrices given by:

$$A = 2\begin{bmatrix} Q_{XX} & \mu(\overline{X}_{\cdot}' - \frac{1}{\xi_0'\xi_0}\overline{X}'\xi_0\theta_0') \\ \bullet & (\mu^2 + (1-q^2)\sigma_\alpha^2)(I_{T-1} - \frac{\theta_0\theta_0'}{\xi_0'\xi_0}) \end{bmatrix}$$

$$B = 4\sigma^2\begin{bmatrix} Q_{XX} & \mu(\overline{X}_{\cdot}' - \frac{1}{\xi_0'\xi_0}\overline{X}'\xi_0\theta_0') \\ \bullet & (\mu^2 + (1-q^2)\sigma_\alpha^2)(I_{T-1} - \frac{\theta_0\theta_0'}{\xi_0'\xi_0}) \end{bmatrix} = 2\sigma^2 A$$

where $Q_{XX} = \lim_{N\to\infty}\frac{1}{N}X'[(I_N \otimes M_\xi) + q^2(M_{e_N} \otimes P_\xi)]X$,

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \quad and \quad \overline{X}_{\cdot} = (\overline{X}_2', \overline{X}_3', \ldots \overline{X}_T')'.$$

The efficiency gain of the GLS estimator compared to the within estimator is shown by the difference of the asymptotic covariance matrices. If $Cov\ (\tilde{\beta}_w, \tilde{\theta}_w) - Cov(\tilde{\beta}_{GLS}, \tilde{\theta}_{GLS})$ is positive semidefinite (PSD), $\tilde{\beta}_{GLS}$ and $\tilde{\theta}_{GLS}$ are more efficient than $\tilde{\beta}_w$ and $\tilde{\theta}_w$.

Note that the covariance matrices of the within and the GLS estimators are block-diagonal when $\mu = 0$. This gives

$$[Var(\tilde{\beta}_{GLS})]^{-1} - [Var(\tilde{\beta}_w)]^{-1} = \frac{q^2}{N}X'\ (I_N \otimes P_\xi)X'\text{, which is PSD.} \tag{30}$$

$$[Var(\tilde{\theta}_{GLS})]^{-1} - [Var(\tilde{\theta}_w)]^{-1} = 0. \tag{31}$$

Thus $\tilde{\beta}_{GLS}$ is more efficient than $\tilde{\beta}_w$ and $\tilde{\theta}_{GLS}$ and $\tilde{\theta}_w$ are equally efficient. However, the efficiency gain of GLS over within disappears as T goes to infinity since $q^2 \to 0$ as $T \to \infty$.

Because of the lack of knowledge of q^2, we need a feasible GLS estimator using a consistent estimator of q^2. We can estimate q^2 from the results of the within and the between regressors: or, for that matter, from the within and between sums of squares evaluated at any consistent estimates. Specifically

$$\text{p}\lim\ \hat{\sigma}^2 = \text{p}\lim\frac{SSE_w}{N(T-1)-K} = \sigma^2 \tag{32}$$

$$\text{p}\lim\ \widehat{\sigma^2 + \xi'\xi\sigma_\alpha^2} = \text{p}\lim\frac{SSE_B}{N-K-1} = \sigma^2 + \xi'\xi\sigma_\alpha^2 \tag{33}$$

$$\text{p}\lim\ \hat{q}^2 = \text{p}\lim\frac{SSE_w}{SSE_B}\frac{N-K-1}{N(T-1)-K} = q^2. \tag{34}$$

Since q^2 is consistently estimated, the asymptotic properties of the feasible GLS estimator are asymptotically equivalent to those of the GLS estimator.

As in the fixed effects model, we can include a finite number of individual specific components in the random effects model. For details, see Appendix B.

We have discussed a generalization of the conventional random effects model that assumes α_i to be i.i.d. and independent of the disturbance and the explanatory variables. We derived the GLS estimator, showed that it is consistent, and derived its asymptotic distribution. The GLS estimator is more efficient than the within estimator, but the efficiency gain disappears as $T \to \infty$.

IV. Test Statistics

It is meaningful to test the hypothesis that θ is a vector of ones. This is the restriction that reduces our general model to the usual simple panel data model. The within estimator and the GLS estimator of the simple model are not consistent if $\theta \doteqdot e_{T-1}$. In the case of the within estimator,

$$\begin{aligned} \operatorname{plim} \hat{\beta}_w &= \operatorname{plim}(X'M_G X)^{-1} X'M_G Y \\ &= \beta + \lim_{N\to\infty}(X'M_G X)^{-1} X'M_G (I_N \otimes \xi)\alpha \neq \beta \end{aligned} \tag{35}$$

since $M_G(I_N \otimes \xi) \neq 0$.[6] This means that the conventional panel data model produces inconsistent estimators (has a specification problem) if $\theta \neq e_{T-1}$.

We may develop test-statistics for the hypothesis $\theta \doteqdot e_{T-1}$ based on the work of Ronald Gallant (1985).[7] Gallant considers estimators derived by minimizing an objective function $S_n(\theta)$, where n = sample size and θ = parameters. Our estimators minimize objective functions and therefore fit his framework. For example, for GLS we have

$$S_N(\beta, \theta) = \frac{1}{N}\sum_{i=1}^{N} S(\beta, \theta)$$

[6] $M_G(I_N \otimes \xi) = [I_{NT} - (I_N \otimes (e_T e_T')/T)]\,(I_N \otimes \xi)$
$= (I_N \otimes \xi) - [I_N \otimes (e_T e_T' \xi)/T)] \neq 0.$

[7] See Gallant (1985), p. 217-20.

$$
\begin{aligned}
&= \frac{1}{2N\tau^2}(Y - X\beta)'[(I_N \otimes M_\xi) + q^2(M_{e_N} \otimes P_\xi)](Y - X\beta) \qquad (36)\\
&= \frac{1}{2N\tau^2} CSSE
\end{aligned}
$$

where a preliminary estimator[8] τ^2 is $\hat{\sigma}^2$ derived from the within estimator (32).

The null hypothesis is considered as

$$
H_0 : h(\beta, \theta) = H\begin{bmatrix} \beta \\ \theta \end{bmatrix} - e_{T-1} = 0 \qquad (37)
$$

where $H = [0 : I_{T-1}]$ is a $(T-1) \times (K + T - 1)$ matrix. Then, the LM statistic given by Gallant (p. 219) is

$$
LM = N\left(\frac{\partial S_N(\hat{\lambda}_{\mathrm{GLS}})}{\partial \lambda}\right)' A^{-1}H'(H\hat{V}H')^{-1}H \ \ A^{-1}\left(\frac{\partial S_N(\hat{\lambda}_{\mathrm{GLS}})}{\partial \lambda}\right) \qquad (38)
$$

where $\hat{\lambda}_{\mathrm{GLS}}$ = restricted estimate of $\begin{bmatrix} \beta \\ \theta \end{bmatrix} = \begin{bmatrix} \hat{\beta}_{\mathrm{GLS}} \\ e_{T-1} \end{bmatrix}$.

$\hat{\beta}_{\mathrm{GLS}}$ = GLS estimator with $\theta = e_{T-1}$ imposed,

$$
\begin{aligned}
\hat{V} &= A^{-1} \ B \ A^{-1}\\
A &= \frac{1}{N}\sum_{t=1}^{N}\frac{\partial S_N(\hat{\lambda}_{\mathrm{GLS}})}{\partial\lambda\partial\lambda'}\\
B &= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial S_N(\hat{\lambda}_{\mathrm{GLS}})}{\partial\lambda}\right)\left(\frac{\partial S_N(\hat{\lambda}_{\mathrm{GLS}})}{\partial\lambda}\right)'.
\end{aligned}
$$

The LM statistic in (38) has asymptotically a Chi-square distribution with $(T-1)$ degrees of freedom. Gallant (p. 220) also provides a test-statistic analogous to the usual likelihood ratio and Wald statistics:

$$
\begin{aligned}
LR &= 2N[S_N(\hat{\lambda}_{\mathrm{GLS}}) - S_N(\tilde{\lambda}_{\mathrm{GLS}})]\\
&= \frac{1}{\hat{\sigma}^2}[CSSE(\hat{\lambda}_{\mathrm{GLS}}) - CSSE(\tilde{\lambda}_{\mathrm{GLS}})]
\end{aligned} \qquad (39)
$$

$$
W = N \cdot h(\tilde{\beta}_{\mathrm{GLS}}, \tilde{\theta}_{\mathrm{GLS}})' (H\tilde{V}H')^{-1} h(\tilde{\beta}_{\mathrm{GLS}}, \tilde{\theta}_{\mathrm{GLS}}) \qquad (40)
$$

[8]Gállant originally defines $\tau^2 = \frac{1}{N}\sum_{t=1}^{N}\hat{e}_t\hat{e}_t'$ and $\hat{e}_t$ as least squares residuals obtained from each univariate model when he discuss multivariate nonlinear least squares. See Gallant (1985), p. 149-50.

where

$$\tilde{\lambda}_{\text{GLS}} = \begin{bmatrix} \tilde{\beta}_{\text{GLS}} \\ \tilde{\theta}_{\text{GLS}} \end{bmatrix} = \text{unrestricted GLS estimate of } \begin{bmatrix} \beta \\ \theta \end{bmatrix}$$

$\tilde{V}$= unrestricted GLS estimate of V.

Under general conditions we have also that LR and W are asymptotically Chi-square with $T-1$ degrees of freedom.

V. Conclusions

The usual motivation for the use of panel data model in labor economics and in related areas is the desire to avoid potential bias caused by the omission of unmeasured individual characteristics from the regression equation. For example, in a wage equation, individual "ability" (or "ambition") is usually unobservable, and may have an effect on wage. If so, the omission of ability from the regression will cause a bias in the estimation of the coefficients of those variables that are correlated with ability. The usual solution to this problem is to assume that ability (or more properly the effect of ability on wage) is time invariant and can therefore be captured by a time-invariant individual-specific effect. However, the assumption that the individual effects are time-invariant is very strong. In this paper we have considered a model that weakens this assumption. In particular, we assume an unobservable time-invariant individual variable (such as ability), but we do not assume that its effect on the dependent variable is time invariant. Rather, we need only to assume that the effect of this variable on the dependent variable has the same temporal pattern for all individuals. Thus, for example, the effect of ability on wage may differ across the business cycle, or may display a trend, so long as it does so for all individuals. We estimate this temporal pattern along with the other parameters of the model.

In the liquidity constraint studies of consumption and investment using panel data, a time dummy is included as a macro shock in Euler's equation. There is also a strong assumption that the responses of all economic agents to a macro shock are identical. This assumption can be relaxed in the sense that each individual has a different response to a macro shock from each other but has the same pattern of response over-time.

We develop fixed-effects and random-effects treatments of our model.

Fixed-effects treatments are relevant when the motivation for the use of panel data is bias reduction, as discussed above, while random-effects treatments are relevant when the motivation is efficiency of estimation. Our model is nonlinear so estimation is more complicated than in the usual simple model. In both the fixed and random effects cases we propose a method of estimation, and we prove the consistency and asymptotic normality of the estimates. This is non-trivial since the standard likelihood theory does not apply, due to the so-called incidental parameters problem (the number of unobservable effects increases with sample size). We also propose asymptotically valid tests of the restrictions that reduce our model to the usual panel data model.

A promising line of future research is to consider models that are intermediate between the simple model, in which individual effects have a time-invariant effect on the dependent variable, and our model, in which the temporal pattern of these effects is completely unrestricted. Kumbhakar (1990) has proposed one such model in the frontier production function setting, and our model can be used to test the specification of his model or of other similar models. It is obviously an empirical question how much flexibility of specification the data will typically support.

Appendix A: Derivation of plim (1/*N*) CSSE

FACT:

$$\begin{aligned}\text{plim } (1/N)\ CSSE(\beta, \xi) = {} & (\beta_0 - \beta)'\, Q_X\, (\beta_0 - \beta) + Q_\alpha \xi_0' M_\xi \xi_0 \\ & + 2(\beta_0 - \beta)'\, Q_{X\alpha}\, \xi_0 + (T-1)\sigma^2\end{aligned}$$

and is finite.

Proof: The true relationship is

$$Y_i = X_i\beta_0 + \xi_0\alpha_{0i} + \varepsilon_i\,. \tag{A 1}$$

By substituting (A1) into *CSSE*.

$$\begin{aligned}CSSE = {} & \sum_{i=1}^{N}(\beta_0 - \beta)'X_i'M_\xi X_i(\beta_0 - \beta) + \sum_{i=1}^{N}\alpha_{0i}^2\xi_0'M_\xi\xi_0 \\ & + \sum_{i=1}^{N}\varepsilon_i'M_\xi\varepsilon_i + 2\sum_{i=1}^{N}(\beta_0 - \beta)'X_i'M_\xi\xi_0\alpha_{0i}^2 \\ & + 2\sum_{i=1}^{N}(\beta_0 - \beta)'X_i'M_\xi\varepsilon_i + 2\sum_{i=1}^{N}\alpha_{0i}\xi_0'M_\xi\varepsilon_i\,.\end{aligned}$$

Using the assumptions (α) and (β),

$$Q_X = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i' M_\xi X_i \quad \text{is finite.} \tag{A 2}$$

$$\begin{aligned} Q_\alpha \xi_0' M_\xi \xi_0 \text{ is finite since } \xi_0' M_\xi \xi_0 = 0 \text{ if } \xi = \xi_0 \text{ and} \\ \xi_0' M_\xi \xi_0 = \xi_0' \xi_0 \text{ if } \xi' \xi_0 = 0. \end{aligned} \tag{A 3}$$

$$\begin{aligned} &\operatorname{plim} \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i' M_\xi \varepsilon_i \\ &\quad = T\sigma^2 - \frac{1}{\xi'\xi} \operatorname{plim} \frac{1}{N} \sum_{i=1}^{N} \operatorname{trace}(\varepsilon_i' \xi \xi' \varepsilon_i) = (T-1)\sigma^2. \end{aligned} \tag{A 4}$$

It is true that

$$\begin{aligned} &\left[\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i' \xi_0 \alpha_{0i} \right]^2 \le \lim_{N\to\infty} \left[\frac{1}{N} \sum_{i=1}^{N} X_i' \xi \xi' X_i \right] \lim_{N\to\infty} \left[\frac{1}{N} \sum_{i=1}^{N} \alpha_{0i}^2 \right] \\ &\text{and } \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i' \xi_0 \xi_0' X_i \text{ and } \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \alpha_{0i}^2 \text{ are finite.} \end{aligned} \tag{A 5}$$

Note that $\sum_{i=1}^{N} X_i' P_\xi \xi_0 \alpha_{0i} = 0$ if $\xi' \xi_0 = 0$ and if $\xi = \xi_0$,

$$\left[\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i' P_\xi \xi_0 \alpha_{0i} \right]^2 = \left[\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i \xi_0 \alpha_{0i} \right]^2.$$

Therefore, $Q_{X_\alpha} \xi_0 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} X_i' M_\xi \xi_0 \alpha_{0i}$ is finite. (A 6)

$$\operatorname{plim} \frac{1}{N} \sum_{i=1}^{N} (\beta_0 - \beta)' X_i M_\xi \varepsilon_i = 0. \tag{A 7}$$

$$\operatorname{plim} \frac{1}{N} \sum_{i=1}^{N} \alpha_{0i} \xi_0' M_\xi \varepsilon_i = 0. \tag{A 8}$$

Using (A 2) – (A 8), we can show

$$\begin{aligned} \operatorname{plim} (1/N)\, CSSE(\beta, \xi) = (\beta_0 - \beta)' Q_X (\beta_0 - \beta) + Q_\alpha \xi_0' M_\xi \xi_0 \\ + 2(\beta_0 - \beta)' Q_{X\alpha} \xi_0 + (T-1)\sigma^2 \end{aligned}$$

and to be finite.

Q.E.D.

Appendix B: General Model with Multiple Components

In the fixed effects model, the regression equation is

$$Y_i = X_i\beta + \xi_1\alpha_{1i} + \ldots + \xi_G\,\alpha_{Gi} + \varepsilon_i\,,\ i = 1,\ldots,N \tag{B 1}$$

where $\xi_g = (1,\ \theta_g')$, $\theta_g = (\theta_{g2},\ \theta_{g2} \ldots \theta_{gT})'$, $g = 1,2,\ldots,G$.
For identification we make the orthorgonality assumption $\xi'_g\ \xi_f = 0,\ g \neq f$.

We can obtain solutions for β in terms of θ and θ in terms of β by the same analysis as the one-component model.

$$\tilde{\beta}_w = (X'\,(I_N \otimes M_\Theta)X)^{-1}\,X'\,(I_N \otimes M_\Theta)Y \tag{B 2}$$

where $M_\Theta = I_T - P_{\xi 1} - P_{\xi 2}\ ,\ldots,\ P_{\xi G}$ and $P_{\xi g} = \xi_g(\xi'_g\xi_g)^{-1}\,\xi'_g$, $g = 1,\ldots,G$. $\tilde{\xi}_{gw}$ is the eigenvector corresponding to the gth largest eigenvalue of

$$\Sigma(Y_i - X_i\tilde{\beta}_w)\,(Y_i - X_i\tilde{\beta}_w)'\,. \tag{B 3}$$

The same asymptotic theory in the one-component model is applied to show that the estimators are consistent and to derive their asymptotic covariance matrix.

In the random effects model, the regression equation is then

$$Y_i = X_i\beta + v_i,\ \ i = 1,\ldots,N \tag{B 4}$$

where $v_i = \xi_1\alpha_{1i} + \xi_2\alpha_{2i} + \ldots + \xi_G\alpha_{Gi} + \varepsilon_i$.
The assumptions in this regression model are as follows:

(A.1) $E(\alpha_{gi}) = \mu_g$, and $Var(\alpha_{gi}) = \sigma^2_g$. α_{gi} is independent of α_{fj} for all g, f, i, and j except $g = f$ and $i = j$. It is independent of X and ε and we denote $\alpha^*_{gi} = \alpha_{gi} - \mu_g$.

(A.2) The orthogonality conditions hold: $\xi_g\xi_f = 0,\ g \neq f$.
We can derive the solutions for β and each θ_g by minimizing $CSSE$ with respect to β and θ_g. This yields

$$\begin{aligned}\beta_{\text{GLS}} &= (X'[(I_N \otimes M_\Theta) + \sum_{g=1}^{G} q_g^2 (M_e \otimes P_{\xi_g})]X)^{-1} \\ &\cdot X'[(I_N \otimes M_\Theta) + \sum_{g=1}^{G} q_g^2 (M_e \otimes P_{\xi_g})]Y\end{aligned} \tag{B5}$$

where $q_g^2 = \sigma^2/(\sigma^2 + \xi_g'\xi_g\sigma^2_{\alpha_g})$

$\tilde{\xi}_{g\text{GLS}} = (1,\ \tilde{\theta}_{g\text{GLS}}')'$ is the eigenvector corresponding to the largest eigenvalue of

$$\sum_{i=1}^{N}\left[\sqrt{1-q_g^2}e_i+(1-\sqrt{1-q_g^2})\bar{e}\right]\left[\sqrt{1-q_g^2}e_i+(1-\sqrt{1-q_g^2})\bar{e}\right]'. \tag{B6}$$

The estimates in (B5) and (B6) are consistent and asymptotically efficient by the same reasoning in the one-component model.

As in the one-component case, we can get a consistent estimator of q_g^2 using the results of the within and the between regressors. Specifically

$$\text{plim } \hat{\sigma}^2 = \text{plim } \frac{SSE_w}{N(T-G)} = \sigma^2 \tag{B7}$$

$$\text{plim } \widehat{\sigma^2 + \xi_g'\xi_g\sigma_{\alpha_g}^2} = \text{plim } \frac{SSE_{B_g}}{N(T-G)} = \sigma^2 + \xi_g'\xi_g\sigma_{\alpha_g}^2 \tag{B8}$$

$$\text{plim } \hat{q}_g^2 = \text{plim } \frac{SSE_w}{SSE_{B_g}}\frac{N}{N(T-G)} = q_g^2 \tag{B9}$$

where $SSE_{Bg} = (Y-X\beta)'(M_{e_N} \otimes P_{\xi_g})(Y-X\beta)$.

The properties of the feasible GLS estimation using a consistent estimator of q_g^2 are asymptotically equivalent to those of the GLS estimator. (*Manuscript received October, 1993; final revision received January, 1994*)

References

Amemiya, T. *Advanced Econometrics*, Harvard Press, 1985.

Chamberlain, G. "Analysis of Covariance with Qualitative Data." *Review of Economic Studies* 47(1980): 225-38.

Cornwell, C., Schmidt, P., and Sickles, R. "Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels." *Journal of Econometrics* 46(1990): 185-200.

Gallant, R. *Nonlinear Statistical Models*. John Wiley & Sons, 1985.

Hausman, J.A., and Taylor, W. "Panel Data and Unobserable Individual Effects." *Econometrica* 49(1981): 1377-99.

Hamermesh, D.S. "Why Do Individual-Effects Models Perform So Poorly? The Case of Academic Salaries." *Southern Economic Journal* 56(1989): 39-45.

Hsiao, C. "Benefits and Limitations of Panel Data." *Economic Reviews* 4 (1985): 121-74.

__________. *Analysis of Panel Data*, New York: Cambridge University Press, 1986.

Juhn, C., Murphy, K.M., and Pierce, B. "Wage Inequality and the Rise in Returns to Skill." Unpublished Paper, University of Chicago, 1989.

Kumbhakar, S.C. "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency." *Journal of Econometrics* 46(1990): 201-12.

Lee, Y.H. "Panel Data Models with Multiplicative Individual and Time Effects: Applications to Compensation and Frontier Production Functions," Unpublished Ph.D. Dissertation, East Lansing: Department of Economics, Michigan State University, 1991.

____________, and Schmidt, P. "A Production Frontier Model with Flexible Temporal Variation in Technical Efficiency," in Fred, H.O., Lovell, C.A. K., and Schmidt, S.S., ed. *The measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press. 1993.

Maddala, G.S. "The Use of Variance Components Models in Pooling Cross Section and Time Series Data." *Econometrica* 39(1971): 341-58.

Mundlak, Y. "Empirical Production Function Free of Management Bias." *Journal of Farm Economics* 43(1961): 44-56.

____________. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46(1978): 69-85.

Schmidt, P., and Sickles, R. "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics* 2 (1984): 367-74.

Whited, T.M. "Debt, Liquidity Constraints, and Corporate Investment: Evidence from Panel Data." *Journal of Finance* (1992):1425-59.

Zeldes, S.P. "Consumption and Liquidity Constraints: An Empirical Investigation." *Journal of Political Economy* 97(1989): 305-46.