Panel Data Models

James L. Powell
Department of Economics
University of California, Berkeley

Overview

Like Zellner's seemingly unrelated regression models, the dependent and explanatory variables for *panel data models* (or *pooled cross section and time series models*) are typically denoted using two (or more) subscripts, where the different subscripts indicate different characteristics of the variable – usually indicating both individual and time, but possibly denoting location, group, etc. Some variants of the model (*"fixed effects" models*) can be viewed as special cases of the classical linear regression model, while others (*"random effects" models*) are special cases of the generalized regression model.

The simplest model assumes that the dependent variable y_{it} satisfies a linear model with an intercept that is specific to individual i,

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}, \qquad i = 1, ..., N; \qquad t = 1, ..., T.$$

This is known as a *balanced panel,* since all individual observations are assumed observable for every time period (and vice versa); here, Kronecker product notation will useful when using matrix notation for the data set. In contrast, *clustered* or *grouped data models* typically assume the number of observations per group i can vary across groups, which in this notation would replace the common number of "time periods" T with a groups-specific number T_i of individuals.

As for seemingly unrelated equations, the number of time periods T in most panel data applications is usually small relative to the number of individuals N. However, unlike SUR models, where it is more natural to "stack" observations by the second subscript first – that is, across individuals for each equation, and then stack by equation – for panel data models the usual convention is to stack observations in the opposite order of subscripts, that is, first collecting the observations across time for each individual as

$$\underset{(T\times 1)}{\mathbf{y}_i} = \underset{(T\times K)}{\mathbf{X}_i}\boldsymbol{\beta} + \alpha_i \boldsymbol{\iota}_T + \boldsymbol{\varepsilon}_i$$

for $i = 1, ..., N$, where $\mathbf{y}_j$ and $\boldsymbol{\varepsilon}_j$ are T-vectors and $\mathbf{X}_i$ is a $T \times K$ matrix,

$$\underset{(T\times 1)}{\mathbf{y}_i} = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \cdots \\ y_{iT} \end{pmatrix}, \qquad \underset{(T\times 1)}{\boldsymbol{\varepsilon}_i} = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \cdots \\ \varepsilon_{iT} \end{pmatrix}, \qquad \underset{(T\times K)}{\mathbf{X}_i} = \begin{pmatrix} \mathbf{x}_{i1}' \\ \mathbf{x}_{i2}' \\ \cdots \\ \mathbf{x}_{iT}' \end{pmatrix},$$

and $\boldsymbol{\iota}_T$ is a T-dimensional column vector of ones. Then, stacking the entire data set by individuals,

$$\underset{(NT\times 1)}{\mathbf{y}} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdots \\ \mathbf{y}_N \end{pmatrix}, \qquad \underset{(NT\times 1)}{\boldsymbol{\varepsilon}} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \cdots \\ \boldsymbol{\varepsilon}_N \end{pmatrix}, \qquad \underset{(NT\times K)}{\mathbf{X}} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \cdots \\ \mathbf{X}_N \end{pmatrix},$$

and defining

$$\underset{(N\times 1)}{\boldsymbol{\alpha}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_N \end{pmatrix},$$

the data can be represented by the single (relatively simple) equation

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon},$$

where

$$\underset{(NT\times N)}{\mathbf{D}} \equiv \mathbf{I}_N \otimes \boldsymbol{\iota}_T.$$

In panel data (or clustered data) models, the "individual intercept" α_i is meant to control for the effect of unobservable regressors that are specific to individual i, so that

$$\alpha_i = \mathbf{w}_i'\boldsymbol{\gamma}$$

where the unobservable, individual-specific regressors $\mathbf{w}_i$ might include "ability," "intelligence," "family background," "ambition," etc. Different assumptions on the relationship of the observable regressors $\mathbf{x}_{it}$ to the intercept term α_i, and thus to the unobservable regressors $\mathbf{w}_i$, yield different variations on the classical and generalized regression models. With this notation, it is understood that the matrix $\mathbf{X}$ does *not* include a column vector of ones, since otherwise a linear combination of the matrix $\mathbf{D}$ of individual-specific "dummy variables" would yield this vector of ones,

$$\begin{aligned} \mathbf{D}\cdot\boldsymbol{\iota}_N &= (\mathbf{I}_N \otimes \boldsymbol{\iota}_T)(\boldsymbol{\iota}_N \otimes 1) \\ &= (\mathbf{I}_N \cdot \boldsymbol{\iota}_N \otimes \boldsymbol{\iota}_T \cdot 1) \\ &= \iota_{NT}, \end{aligned}$$

so that the combined matrix $[\mathbf{X}, \mathbf{D}]$ of regressors would not have full column rank.

“Fixed Effect” Model

When the individual intercepts α_i are treated as fixed constants, which can be arbitrarily related to the regression vectors $\mathbf{x}_{it}$, the resulting model, known as the *fixed effect model*, can be viewed as a special case of the classical linear model under the usual assumptions that $\mathbf{X}$ is nonrandom, $[\mathbf{X}, \mathbf{D}]$ is of full column rank, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$, and $\mathbf{V}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{NT}$. By the usual “residual regression” formulae, the classical LS estimator of the subvector $\boldsymbol{\beta}$ of regression coefficients is

$$\hat{\boldsymbol{\beta}}_{FE} = \left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1} \tilde{\mathbf{X}}'\tilde{\mathbf{y}},$$

where $\tilde{\mathbf{y}} \equiv \left(\mathbf{I}_{NK} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'\right)\mathbf{y}$ are the residuals of a regression of $\mathbf{y}$ on $\mathbf{D}$, with an analogous definition of $\tilde{\mathbf{X}}$. (The subscript “FE” stands for ”fixed effects”.) By the special structure of the matrix $\mathbf{D}$ (which has a lot of ones in it), it is straightforward to show that the subvector of $\tilde{\mathbf{y}}$ corresponding to individual i can be expressed as

$$\begin{aligned}
\tilde{\mathbf{y}}_i &= (\mathbf{I}_T - \boldsymbol{\iota}_T(\boldsymbol{\iota}_T'\boldsymbol{\iota}_T)^{-1}\boldsymbol{\iota}_T')\mathbf{y}_i \\
&= \mathbf{y}_i - y_{i\cdot}\boldsymbol{\iota}_T,
\end{aligned}$$

where $y_{i\cdot}$ is the average value of y_{it} over t,

$$y_{i\cdot} \equiv \frac{\boldsymbol{\iota}_T'\mathbf{y}_i}{\boldsymbol{\iota}_T'\boldsymbol{\iota}_T} = \frac{1}{T}\sum_{t=1}^{T} y_{it}.$$

This result can be interpreted as follows: from the original structural equation

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}$$

it follows that

$$y_{i\cdot} = \mathbf{x}_{i\cdot}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{i\cdot},$$

and subtracting the second equation from the first yields

$$\begin{aligned}
y_{it} - y_{i\cdot} &= (\mathbf{x}_{it} - \mathbf{x}_{i\cdot})'\boldsymbol{\beta} + (\alpha_i - \alpha_i) + (\varepsilon_{it} - \varepsilon_{i\cdot}) \\
&= (\mathbf{x}_{it} - \mathbf{x}_{i\cdot})'\boldsymbol{\beta} + (\varepsilon_{it} - \varepsilon_{i\cdot}),
\end{aligned}$$

so deviating y_{it} and $\mathbf{x}_{it}$ from their time-averages eliminates the “fixed effect” α_i from the structural equation, much as taking deviations from means eliminates the intercept term in a classical regression model (with intercept).

An alternative interpretation of the fixed effect estimator $\hat{\boldsymbol{\beta}}_{FE}$ uses first-differences over time, rather than deviations from time-averages, to eliminate the fixed effect α_i:

$$\begin{aligned}\boldsymbol{\Delta} y_{it} &\equiv y_{it} - y_{i,t-1} \\ &= \boldsymbol{\Delta}\mathbf{x}_{it}'\boldsymbol{\beta} + \boldsymbol{\Delta}\alpha_i + \boldsymbol{\Delta}\varepsilon_{it} \\ &= \boldsymbol{\Delta}\mathbf{x}_{it}'\boldsymbol{\beta} + \boldsymbol{\Delta}\varepsilon_{it},\end{aligned}$$

for $i = 1, ..., N$ and $t = 2, ..., T$. For balanced panels, LS regression of $\boldsymbol{\Delta} y_{it}$ on $\boldsymbol{\Delta}\mathbf{x}_{it}$ gives identical results to LS regression of $y_{it} - y_{i\cdot}$ on $\mathbf{x}_{it} - \mathbf{x}_{i\cdot}$.

Considering either interpretation, it is clear that the regression coefficients on any component of $\mathbf{x}_{it}$ that is constant over time ($\mathbf{x}_{it} \equiv \mathbf{x}_{is}$) will be unidentified, since that component of either $\boldsymbol{\Delta}\mathbf{x}_{it}$ or $\mathbf{x}_{it} - \mathbf{x}_{i\cdot}$ will be identically zero. It is only variation in the regressors across time for a given individual that allows the corresponding coefficient to be identified relative to the individual-specific, time-invariant intercept α_i.

If the error terms ε_{it} are i.i.d. and normally distributed with zero mean and variance σ^2, then the fixed effect estimator $\hat{\boldsymbol{\beta}}_{FE}$ is also the maximum likelihood (ML) estimator of $\boldsymbol{\beta}$, and the ML estimator of σ^2 is

$$\hat{\sigma}^2_{ML} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - y_{i\cdot} - (\mathbf{x}_{it} - \mathbf{x}_{i\cdot})'\,\hat{\boldsymbol{\beta}}_{FE}\right)^2.$$

This estimator is biased, and, if $N \to \infty$ for fixed T (a reasonable approximation if N is large relative to T), the ML estimator of σ^2 is also inconsistent, with

$$\hat{\sigma}^2_{ML} \xrightarrow{p} \frac{(T-1)}{T}\sigma^2.$$

This inconsistency of the ML estimator of σ^2 is a classic example of the nefarious "incidental parameters problem" described by Neyman and Scott. Because the number N of intercept terms α_i increases to infinity as N increases, and the corresponding ML estimator

$$\begin{aligned}\hat{\alpha}_i &\equiv y_{i\cdot} - \mathbf{x}_{i\cdot}'\hat{\boldsymbol{\beta}}_{FE} \\ &\xrightarrow{p} \alpha_i + \varepsilon_{i\cdot}\end{aligned}$$

is inconsistent for α_i, this inconsistency translates into inconsistency of the ML estimator $\hat{\sigma}^2_{ML}$. Fortunately, this doesn't cause inconsistency of $\hat{\beta}_{FE}$, and an unbiased and consistent estimator

$$s^2_{ML} = \frac{1}{N(T-1)-K}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - y_{i\cdot} - (\mathbf{x}_{it} - \mathbf{x}_{i\cdot})'\,\hat{\boldsymbol{\beta}}_{FE}\right)^2$$

of σ^2 is readily available.

It is worth mentioning that the "fixed effect" label does not mean that the regressors $\mathbf{x}_{it}$ or intercept terms α_i must be viewed as nonrandom, or "fixed" in a statistical sense; they may be indeed be viewed as random and jointly distributed, provided the conditions on ε_{it} are assumed to hold conditional on the realizations of $\mathbf{X}$ and $\boldsymbol{\varepsilon}$. What characterizes a "fixed effect" model is that no structure on the relationship between α_i and $\mathbf{x}_{it}$ is imposed.

"Random Effect" Model

A drawback of the fixed-effect model is its failure to identify any components of $\boldsymbol{\beta}$ corresponding to regressors that are constant over time for a given individual; for such coefficients to be identified, stronger conditions on the relation of the individual-specific intercept α_i to the regressors $\mathbf{x}_{it}$ must be imposed. The *random effects model* uses a simple but very strong assumption to restrict this relationship: namely, that the intercept α_i is a random variable which is not related to $\mathbf{x}_{it}$ and ε_{it}, in the sense that

$$E(\alpha_i) = \alpha,\ Var(\alpha_i) = \boldsymbol{\sigma}_\alpha^2,\ \text{and } Cov(\alpha_i, \varepsilon_{it}) = 0,$$

all assumed independent of $\mathbf{x}_{it.}$ (These moments should be interpreted as conditional on $\mathbf{X}$ if the regressors are viewed as random.) Relabeling the variance of ε_{it} as $Var(\varepsilon_{it}) \equiv \sigma_\varepsilon^2$, the original panel data model can be rewritten as

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha + u_{it},$$

where

$$u_{it} \equiv (\alpha_i - \alpha) + \varepsilon_{it}$$

has

$$\begin{aligned} E(u_{it}) &= 0, \\ Var(u_{it}) &= \sigma_\alpha^2 + \sigma_\varepsilon^2, \\ Cov(u_{it}, u_{is}) &= \sigma_\alpha^2,\ \text{and} \\ Cov(u_{it}, u_{js}) &= 0 \qquad \text{if } i \neq j. \end{aligned}$$

This implies that the model can be written in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \alpha\boldsymbol{\iota}_{NT} + \mathbf{u},$$

where

$$\begin{aligned} E(\mathbf{u}) &= \mathbf{0}, \\ \mathbf{V}(\mathbf{u}) &= \sigma_\alpha^2 \mathbf{D}\mathbf{D}' + \sigma_\varepsilon^2 \mathbf{I}_{NT} \\ &\equiv \sigma_\varepsilon^2 \boldsymbol{\Omega}, \end{aligned}$$

with

$$\begin{aligned} \boldsymbol{\Omega} &\equiv \mathbf{I}_{NT} + \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2} \mathbf{D}\mathbf{D}' \\ &= \mathbf{I}_{NT} + \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2} \left(\mathbf{I}_N \otimes \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right). \end{aligned}$$

In the unlikely event that the ratio $\theta \equiv \boldsymbol{\sigma}_\alpha^2 / \boldsymbol{\sigma}_\varepsilon^2$ were known – and thus $\boldsymbol{\Omega}$ did not contain unknown parameters – Aitken's GLS estimator of $\boldsymbol{\beta}$ and α would have the usual form

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}_{GLS} \\ \hat{\alpha}_{GLS} \end{pmatrix} = (\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z})^{-1}\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y},$$

where $\mathbf{Z} \equiv [\mathbf{X}, \boldsymbol{\iota}_{NT}]$. Like the fixed effects estimator, there are a couple algebraically-equivalent representations of the GLS estimator $\hat{\boldsymbol{\beta}}_{GLS}$ of the slope coefficients. One interpretation is the coefficients of a LS regression of y_{it}^* on $\mathbf{x}_{it}^*$, where

$$\begin{aligned} y_{it}^* &\equiv y_{it} - y_{i\cdot} + \omega \cdot (y_{i\cdot} - y_{\cdot\cdot}), \\ \mathbf{x}_{it}^* &\equiv \mathbf{x}_{it} - \mathbf{x}_{i\cdot} + \omega \cdot (\mathbf{x}_{i\cdot} - \mathbf{x}_{\cdot\cdot}), \end{aligned}$$

for

$$\begin{aligned} y_{\cdot\cdot} &\equiv \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it}, \\ \mathbf{x}_{\cdot\cdot} &\equiv \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{x}_{it}, \end{aligned}$$

and

$$\begin{aligned} \omega &\equiv \sqrt{\frac{\sigma_\varepsilon^2}{T\sigma_\varepsilon^2 + \sigma_\alpha^2}} \\ &\equiv \sqrt{\frac{1}{T\theta + 1}}, \end{aligned}$$

where as above $\theta \equiv \boldsymbol{\sigma}_\alpha^2 / \boldsymbol{\sigma}_\varepsilon^2$. As the variability of the random effect σ_α^2 declines to zero for σ_ε^2 fixed, $\omega \to 1$, and the GLS estimator reduces to the usual LS regression of the deviations $y_{it} - y_{\cdot\cdot}$ in the dependent variable from its mean value on the corresponding deviations $\mathbf{x}_{it} - \mathbf{x}_{\cdot\cdot}$ in regressors. Using this result, we can obtain another interpretation of the GLS estimator as as a matrix-weighted average

$$\hat{\boldsymbol{\beta}}_{GLS} = \mathbf{A}\hat{\boldsymbol{\beta}}_{FE} + (\mathbf{I} - \mathbf{A})\hat{\boldsymbol{\beta}}_{B}$$

of the fixed effect estimator $\hat{\boldsymbol{\beta}}_{FE}$ defined above and the "between estimator"

$$\hat{\boldsymbol{\beta}}_{B} \equiv \left[\sum_{i=t}^{N}(\mathbf{x}_{i\cdot} - \mathbf{x}_{\cdot\cdot})(\mathbf{x}_{i\cdot} - \mathbf{x}_{\cdot\cdot})'\right]^{-1}\sum_{i=t}^{N}(\mathbf{x}_{i\cdot} - \mathbf{x}_{\cdot\cdot})(y_{i\cdot} - y_{\cdot\cdot})$$

of $\boldsymbol{\beta}$ coming from the LS regression of the time-averages $y_{i\cdot}$ on $\mathbf{x}_{i\cdot}$ and a constant. (The form of the matrix $\mathbf{A} = \mathbf{A}(\theta)$ is given in Ruud's textbook.)

When the variances σ_α^2 and σ_ε^2 are unknown (as is always the case), an estimator of $\theta \equiv \boldsymbol{\sigma}_\alpha^2/\boldsymbol{\sigma}_\varepsilon^2$ or $\omega \equiv (1 + T\theta)^{-1/2}$ is needed to construct a Feasible GLS estimator. As noted above, the unbiased estimator s_{FE}^2 of σ_ε^2 based upon the fixed-effect estimator $\hat{\boldsymbol{\beta}}_{FE}$ will be consistent; for fixed T, the corresponding variance estimator for the "between" estimator

$$s_B^2 \equiv \frac{1}{N - K - 1}\sum_{i=1}^{N}\left((y_{i\cdot} - y_{\cdot\cdot}) - (\mathbf{x}_{i\cdot} - \mathbf{x}_{\cdot\cdot})'\,\hat{\boldsymbol{\beta}}_B\right)^2$$

will be unbiased and consistent for $Var(\varepsilon_{i\cdot}) = \sigma_\alpha^2 + \left(\sigma_\varepsilon^2/T\right)$. Hence

$$\hat{\omega} \equiv \frac{s_{FE}^2}{Ts_B^2} \xrightarrow{p} \omega,$$

and can be used in place of ω to construct a Feasible GLS estimator. Note that the corresponding estimator of σ_α^2,

$$s_\alpha^2 \equiv s_B^2 - \left(s_{FE}^2/T\right),$$

is not guaranteed to be positive.

Time Effects and "Differences in Differences"

In addition to the assumption that the "intercept term" varies across individuals i, it might also be reasonable to assume that it varies across time t; defining η_t to be the time-specific intercept term, a generalization of the basic panel data model would be

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + \eta_t + \varepsilon_{it}, \qquad i = 1, ..., N; \qquad t = 1, ..., T,$$

which, in matrix form, might be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\alpha} + \mathbf{R}\boldsymbol{\eta} + \boldsymbol{\varepsilon},$$

for

$$\underset{(T\times 1)}{\boldsymbol{\eta}} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \cdots \\ \eta_T \end{pmatrix}$$

and

$$\underset{(NT\times T)}{\mathbf{R}} \equiv \boldsymbol{\iota}_N \otimes \mathbf{I}_T.$$

Since

$$\mathbf{D}\boldsymbol{\iota}_N = \iota_{NT} = \mathbf{R}\boldsymbol{\iota}_T,$$

the columns of the matrix of regressors $[\mathbf{X}, \mathbf{D}, \mathbf{R}]$ would not be linearly independent, and a normalization on $\boldsymbol{\alpha}$ or $\boldsymbol{\eta}$ would need to be imposed, e.g., $\eta_1 \equiv 0$, which would imply that the first column of $\mathbf{R}$ could be dropped, along with the first component of $\boldsymbol{\eta}$. Treating the $\boldsymbol{\alpha}$ and $\boldsymbol{\eta}$ parameters as fixed effects, the corresponding fixed effect estimator of $\boldsymbol{\beta}$ is obtained by a regression of $\tilde{y}_{it}$ on $\tilde{\mathbf{x}}_{it}$, where

$$\tilde{y}_{it} \equiv y_{it} - y_{i\cdot} - y_{\cdot t} + y_{\cdot\cdot},$$

with

$$y_{\cdot t} \equiv \frac{1}{N}\sum_{i=1}^{N} y_{it}$$

and corresponding definitions for $\tilde{\mathbf{x}}_{it}$ and $\mathbf{x}_{\cdot t}$.

If T is small (relative to N), the time-specific intercepts $\boldsymbol{\eta}$ are typically treated as fixed, which implies that the coefficients of any regressors that are time-specific (i.e., do not vary across individuals) would be unidentified. If such coefficients are of interest, the $\boldsymbol{\eta}$ coefficients can be treated as random effects, with corresponding variance component σ_η^2, along with the individual effects $\boldsymbol{\alpha}$. The corresponding GLS estimator of $\boldsymbol{\beta}$ would combine the fixed effects estimator with two different "between" estimators, one involving the regression of $y_{i\cdot}$ on $\mathbf{x}_{i\cdot}$ and a constant and the other regressing $y_{\cdot t}$ on $\mathbf{x}_{\cdot t}$ and a constant. The details are a bit messy, and can be found in many graduate texts.

A special case of the fixed effects model with individual- and time-specific effects is the so-called "*differences in differences*" (or "diffs in diffs") framework, a name more descriptive of the estimation method than the model itself. The simplest version of this model has $T = 2$ and a single time-varying regressor x_{it} which is binary. Specifically the N individual observations are classified into two groups, the "controls" (for $i = 1, ..., N_c$), for which $x_{it} \equiv 0$, and the "treated" ($i = N_c + 1, ..., N$), for which $x_{i1} = 0$

and $x_{i2} = 1$. The scalar coefficient β is the "treatment effect," i.e., the change in the average value of the dependent variable between the pre-treatment ($t = 1$) and post-treatment ($t = 2$) periods. To repeat,

$$x_{it} = \begin{cases} 0 & \text{if } i = 1, ..., N_c, \\ 0 & \text{if } i = N_c + 1, ..., N, \text{ and } t = 1; \\ 1 & \text{if } i = N_c + 1, ..., N, \text{ and } t = 1. \end{cases}$$

Writing the structural equation

$$y_{it} = x_{it}\boldsymbol{\beta} + \alpha_i + \eta_t + \varepsilon_{it}, \qquad i = 1, ..., N; \qquad t = 1, 2,$$

and taking first differences yields

$$\begin{aligned} \Delta y_{i2} &= y_{i2} - y_{i1} \\ &= \Delta x_{i2}\beta + \Delta\eta_2 + \Delta\varepsilon_{i2} \\ &= \begin{cases} \Delta\eta_2 + \Delta\varepsilon_{i2} & \text{if } i = 1, ..., N_c \\ \beta + \Delta\eta_2 + \Delta\varepsilon_{i2} & \text{if } i = N_c + 1, ..., N. \end{cases} \end{aligned}$$

A classical LS regression of Δy_{i2} on a constant and Δx_{i2} yields

$$\hat{\beta}_{FE} = \Delta\bar{y}^2 - \Delta\bar{y}^1,$$

where

$$\begin{aligned} \Delta\bar{y}^1 &\equiv \frac{1}{N_c}\sum_{i=1}^{N_c}(y_{i2} - y_{i1}) \qquad \text{and} \\ \Delta\bar{y}^2 &\equiv \frac{1}{N - N_c}\sum_{i=N_c+1}^{N}(y_{i2} - y_{i1}) \end{aligned}$$

are the average changes in y_{it} for the control and treatment groups, respectively. So the estimate of the treatment effect is the difference in the average change in the dependent variable across the two groups. As N_c and $N - N_c$ tend to infinity, it is easy to see that

$$\begin{aligned} \Delta\bar{y}^1 &\xrightarrow{p} \Delta\eta_2, \\ \Delta\bar{y}^2 &\xrightarrow{p} \beta + \Delta\eta_2, \end{aligned}$$

so the diffs-in-diffs estimator of the treatment effect is consistent. Note that, with this fixed-effects approach, there is no need to assume the treatment assignment (or "choice") is independent of the individual effect α_i or time effect η_t.

Robust Covariance Estimation

Like other GLS applications, statistical inference with panel data can be sensitive to heteroskedasticity or serial correlation of the error terms. Writing the panel data model as

$$y_{it} = \mathbf{z}'_{it}\boldsymbol{\theta} + u_{it},$$

where $\mathbf{z}'_{it}$ includes the row vector of regressors $\mathbf{x}'_{it}$ and the relevant row of the matrices $\mathbf{D}$ and $\mathbf{R}$ of dummy variables for the fixed effect, we may want to assume that the errors u_{it} are uncorrelated across individuals i but have arbitrary variance-covariance patterns over time t for each individual i. That is, suppose

$$\begin{aligned} E(u_{it}) &= 0, \\ Cov(u_{it}, u_{is}) &= \sigma_{i,ts}, \\ Cov(u_{it}, u_{js}) &= 0 \qquad \text{if } i \neq j. \end{aligned}$$

Stacking the observations in the usual matrix form

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\theta} + \mathbf{u},$$

it follows that

$$\mathbf{V}(\mathbf{u}) \equiv \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \boldsymbol{\Sigma}_N \end{bmatrix}$$

for

$$\underset{(T\times T)}{\boldsymbol{\Sigma}_j} \equiv [\sigma_{j,ts}].$$

In the absence of a parametric form for $\boldsymbol{\Sigma}_j$, it would be reasonable to use the classical least squares estimator

$$\hat{\boldsymbol{\theta}}_{LS} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$$

of the $\boldsymbol{\theta}$ coefficients (i.e., the slope coefficients $\boldsymbol{\beta}$ and any individual or time fixed effects), which should be consistent and asymptotically normally distributed under fairly general conditions on $\mathbf{u}$. As in other GLS applications, the trick is to find a consistent estimator for

$$\text{plim}\ \frac{1}{N}\mathbf{V}(\hat{\boldsymbol{\theta}}_{LS}) = \text{plim}\ \left(\frac{1}{N}\mathbf{Z}'\mathbf{Z}\right)^{-1}\left(\frac{1}{N}\mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z}\right)\left(\frac{1}{N}\mathbf{Z}'\mathbf{Z}\right)^{-1},$$

the asymptotic covariance matrix of the LS estimator. The middle matrix is the tricky one; since

$$\frac{1}{N}\mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z} \equiv \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sigma_{i,ts}\mathbf{z}_{it}\mathbf{z}_{is}',$$

the same reasoning that led to the Huber-Eicker-White robust covariance matrix estimator yields

$$\frac{1}{N}\mathbf{Z}'\hat{\boldsymbol{\Sigma}}\mathbf{Z} \equiv \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{u}_{it}\hat{u}_{is}\mathbf{z}_{it}\mathbf{z}_{is}'$$

as a consistent estimator of the middle matrix of $plim\ \frac{1}{N}\mathbf{V}(\hat{\boldsymbol{\theta}}_{LS})$, where $\hat{u}_{it}$ are the LS residuals

$$\hat{u}_{it} \equiv y_{it} - \mathbf{z}_{it}'\hat{\boldsymbol{\theta}}_{LS}.$$

This estimator will be consistent as $N \to \infty$ under similar conditions as for consistency of the Huber-Eicker-White heteroskedasticity-robust covariance estimator. This robust covariance matrix estimator immediately extends to clustered data problems, where the number of "time periods" or "group members" T_i depends upon the group index i; the formulae above are easily extended by changing "T" to "T_i" throughout.

Lagged Dependent Variables in Panel Data

A more serious inference problem arises when the regressors include a lagged dependent variable in models with fixed effects. Consider the special case with $T = 2$ and

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it},$$

where the $\{\alpha_i\}$ are considered fixed effects and ε_{it} satisfies the usual Gauss-Markov assumptions. Assuming y_{i0} is observable for all i and considered a nonrandom "starting value", the fixed-effect estimator of $\boldsymbol{\beta}$ and γ is obtained by a LS regression on the differenced model

$$\begin{aligned}\Delta y_{i2} &= y_{i2} - y_{i1} \\ &= \Delta\mathbf{x}_{i2}'\boldsymbol{\beta} + \gamma\Delta y_{i1} + \Delta\varepsilon_{i2}.\end{aligned}$$

But now

$$\begin{aligned}Cov(\Delta y_{i1}, \Delta\varepsilon_{i2}) &= Cov\left((y_{i1} - y_{i0}), (\varepsilon_{i2} - \varepsilon_{i1})\right) \\ &= -Cov(y_{i1}, \varepsilon_{i1}) \\ &= -Var(\varepsilon_{i1}) \\ &\neq 0,\end{aligned}$$

so the LS regression of Δy_{i2} on $\Delta \mathbf{x}_{i2}$ and Δy_{i1} will yield biased and inconsistent estimators of $\boldsymbol{\beta}$ and γ in general. This problem is the same as the problem of inconsistency of LS with lagged dependent variables and serially-correlated errors; elimination of the fixed effect by differencing (or deviating from individual means) yields a differenced error term which is serially correlated, and thus related to the lagged dependent variable. A solution to the inconsistency of the LS estimator for this problem is the "instrumental variables" method to be discussed next.