Advanced Econometrics

Chapter 12: Linear Models for Panel Data

In Choi

Sogang University

Useful references:

- Badi Baltagi (2008) *Econometric Analysis of Panel Data*, 4th Edition, John Wiley and Sons.
- Cheng Hsiao (2003) *Analysis of Panel Data*, 2nd Edition, Cambridge University Press.
- Cameron, A.C., and P.K. Trivedi (2005). Microeconometrics: Methods and Applications. Cambridge University Press: New York.
- Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

Basic framework

- A general model for panel data

$$
\begin{aligned}
y_{it} &= \alpha + x_{it}'\beta + u_{it}, \ (i = 1, ..., n, \ t = 1, ..., T) \\
u_{it} &= \mu_i + \nu_{it}.
\end{aligned} \quad (1)
$$

i : households, individuals, firms, countries, etc.
t : time
α : a scalar coefficient
β : $K \times 1$ coefficient vector
x_{it} : the $(i, t)-$th observation on K regressors
μ_i : unobservable individual specific effect
ν_{it} : remainder disturbance term

Basic framework

Example

Earnings equation

y_{it} : earnings of the head of the household

x_{it} : a set of variables affecting earnings (experience, education, gender, race, etc.).

μ_i : individual's unobserved ability

Advantages of panel data

- Large number of data points – better efficiency
- Panel data allow a researcher to study a number of important economic questions that cannot be addressed using cross-sectional or time series data sets.

Advantages of panel data

Example

Consider a simple linear regression model

$$y_{it} = \alpha + \beta' x_{it} + \rho' z_{it} + u_{it}.$$

If z_i is unobservable and related to x_{it}, the OLS regression of y_{it} on x_{it} yields biased estimate of β. However, if $T \geq 2$ (i.e., if panel data are available),

$$y_{it} - y_{i,t-1} = \beta' (x_{it} - x_{i,t-1}) + u_{it} - u_{i,t-1}.$$

Running OLS using this model, we can obtain a consistent estimate of β.

The fixed effects model

- Assumption
  1. μ_i are fixed parameters to be estimated. (It is usually assumed to be a random variable correlated with the regressors.)
  2. $\{x_{it}\}$ and $\{v_{it}\}$ are independent.
  3. $v_{it} \sim iid(0, \sigma_v^2)$.

The fixed effects model

- The LSDV (least squares dummy variables) estimator of β
  Using matrix notation, write model (1) as

$$y = \alpha \iota_{NT} + X\beta + Z_\mu \mu + \nu \tag{2}$$

where

$$\begin{aligned} y &= [y_{11}, \dots, y_{1T}, y_{21}, \dots, y_{2T}, \dots, y_{N1}, \dots, y_{NT}]'; \\ \iota_{NT} &= [1, \dots, 1]'; \\ X &= \begin{bmatrix} x_{11}' \\ \vdots \\ x_{1T}' \\ \vdots \\ x_{N1}' \\ \vdots \\ x_{NT}' \end{bmatrix}; \end{aligned}$$

The fixed effects model

$$Z_\mu = I_N \otimes \iota_T;$$

$$\mu = [\mu_1, ..., \mu_N]'$$

The fixed effects model

Let

$$P = Z_\mu (Z'_\mu Z_\mu)^{-1} Z'_\mu \text{ and } Q = I - P.$$

Premultiply model (2) by Q. The resulting model is

$$Qy = QX\beta + Q\nu \tag{3}$$

since $QZ_\mu = 0$ and $Q\iota_{NT} = 0$. The latter relation holds because

$$\begin{aligned} Q\iota_{NT} &= \iota_{NT} - Z_\mu (Z'_\mu Z_\mu)^{-1} Z'_\mu \iota_{NT} \\ &= \iota_{NT} - \frac{1}{T} Z_\mu Z'_\mu \iota_{NT} \\ &= \iota_{NT} - \iota_{NT} \\ &= 0. \end{aligned}$$

The fixed effects model

- Running OLS on model (3), we obtain the LSDV (least squares dummy variables) estimator of β. This is

$$\tilde{\beta}_{LSDV} = (X'QX)^{-1}X'Qy.$$

This estimator is also called the Within-OLS estimator.

- $\tilde{\beta}_{LSDV}$ is equivalent to the OLS estimator from the model

$$y_{it} - \bar{y}_{i.} = (x_{it} - \bar{x}_{i.})'\beta + u_{it} - \bar{u}_{i.} \tag{4}$$

- Parameters α and μ_i cannot be estimated separately.
- Only $\alpha + \mu_i$ can be estimated by

$$\bar{y}_{i.} - \tilde{\beta}'_{LSDV}\bar{x}_{i.},$$

where $\bar{z}_{i.} = \frac{1}{T}\sum_{t=1}^{T} z_{it}$.

- If $\sum_{i=1}^{N} \mu_i = 0$ (i.e., individual effects cancel out each other), μ_i can be estimated. Averaging (1) over time gives

$$\bar{y}_{i.} = \alpha + \beta' \bar{x}_{i.} + \mu_i + \bar{v}_{i.} \tag{5}$$

Averaging across all observations in (1) and utilizing the restriction $\sum_{i=1}^{N} \mu_i = 0$, we obtain

$$\bar{y}_{..} = \alpha + \beta' \bar{x}_{..} + \bar{v}_{..}, \tag{6}$$

where $\bar{z}_{..} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} z_{it}$. From (6),

$$\tilde{\alpha} = \bar{y}_{..} - \tilde{\beta}'_{LSDV} \bar{x}_{..}.$$

We obtain from (5)

$$\tilde{\mu}_i = \bar{y}_{i.} - \tilde{\alpha} - \tilde{\beta}'_{LSDV} \bar{x}_{i.}$$

- If there are any time invariant variables in the model (some elements of x_{it} are represented as z_i), their coefficients cannot be estimated because Q wipes out the variables.
- If T is fixed and $N \to \infty$, $\tilde{\beta}_{LSDV}$ is consistent.
- If T is fixed and $N \to \infty$, the OLS estimator of $\alpha + \mu_i$ is inconsistent (the incidental parameter problem). Intuitively, this happens because the number of parameters increases at exactly the same rate as the number of sample increases.
- OLS on model (1) yields biased and inconsistent estimates of the regression coefficients.

The fixed effects model

First-differencing estimator

- The first-differencing gives

$$\Delta y_{it} = \Delta x_{it}'\beta + \Delta v_{it}.$$

  The individual effects variable μ_i is eliminated by the first-differencing.
- Running OLS on this model gives a consistent estimator of β.
- The variance-covariance matrix of the first-differencing estimator is $E\Big[\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta x_{it}\Delta x_{is}'\Delta v_{it}\Delta v_{is}\right) \times \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'\right)^{-1}\Big].$

- The variance-covariance matrix is estimated by
$$\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\Delta x_{it}\Delta x_{it}'\Delta\hat{v}_{it}\Delta\hat{v}_{is}\right)$$
$$\times\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'\right)^{-1},$$
where $\Delta\hat{v}_{it}$ is the residual from the first-differencing estimation.
- When $T = 2$, the Within-OLS and first-differencing estimators are equivalent.

The fixed effects model

- If

$$u_{it} = \mu_i + \lambda_t + \nu_{it},$$

where λ_t denotes the time-specific variable common to every individual, $\tilde{\beta}_{LSDV}$ is equivalent to the OLS estimator from the model

$$y_{it} - \bar{y}_{i.} - \bar{y}_{t.} + \bar{y}_{..} = (x_{it} - \bar{x}_{i.} - \bar{x}_{t.} + \bar{x}_{..})'\beta + u_{it} - \bar{u}_{i.} - \bar{u}_{t.} + \bar{u}_{..},$$

where $\bar{z}_{t.} = \frac{1}{N}\sum_{i=1}^{N} z_{it}$

- Assume $\sum_{i=1}^{N} \mu_i = 0$ and consider the null hypothesis

$$H_0 : \mu_1 = \mu_2 = ... = \mu_{N-1} = 0.$$

Under this null, there are no fixed effects. This can be tested by Chow test. Let
$RRSS$: restricted residual sum of squares from OLS
$URSS$: unrestricted residual sum of squares from LSDV
Then,

$$F = \frac{(RRSS - URSS)/(N-1)}{URSS/(NT - N - K)} \sim F_{N-1,N(T-1)-K}$$

under $v_{it} \sim iidN(0, \sigma^2)$.

The fixed effects model

- Estimator of σ^2

  Let $\hat{u}_{it}$ be the residual from the regression on (4). Then,

$$\hat{\sigma}^2 = \frac{1}{NT - N - K} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{u}_{it}^2 .$$

  The divisor is chosen to be $NT - N - K$ in order to make $\hat{\sigma}^2$ unbiased.

The random effects model

- Assumption
  1. $\mu_i \sim iid(0, \sigma_\mu^2)$; $v_{it} \sim iid(0, \sigma_v^2)$.
  2. μ_i are independent of v_{it}.
  3. x_{it} are independent of μ_i and v_{it} for all i and t.
- In the random effects model, there is no need for estimating μ_i. Estimating σ_μ^2 is good enough.

- Write the model as

$$\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \alpha \begin{pmatrix} \iota_T \\ \vdots \\ \iota_T \end{pmatrix} + \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \beta + \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix},$$

where $\iota_T = [1, ..., 1]'$. The variance-covariance matrix of u_i is for all i

$$\begin{aligned} V &= E(u_i u_i') = \sigma_\mu^2 \iota_T \iota_T' + \sigma_v^2 I_T \\ &= \sigma_\mu^2 J_T + \sigma_v^2 I_T \\ &= T\sigma_\mu^2 \bar{J}_T + \sigma_v^2 \bar{J}_T + \sigma_v^2 (I_T - \bar{J}_T) \; (\bar{J}_T = J_T / T) \\ &= (T\sigma_\mu^2 + \sigma_v^2) \bar{J}_T + \sigma_v^2 (I_T - \bar{J}_T). \end{aligned}$$

- Note that $\bar{J}_T$ and $(I_T - \bar{J}_T)$ are idempotent matrices and that $\bar{J}_T(I_T - \bar{J}_T) = 0$.
  The inverse of matrix of V is

$$
\begin{aligned}
V^{-1} &= \frac{1}{T\sigma_\mu^2 + \sigma_v^2}\bar{J}_T + \frac{1}{\sigma_v^2}(I_T - \bar{J}_T) \\
&= \frac{1}{\sigma_v^2}\left((I_T - \bar{J}_T) + \psi\bar{J}_T\right) \\
&= \frac{1}{\sigma_v^2}\left(Q + \psi\bar{J}_T\right),
\end{aligned}
$$

  where $\psi = \frac{\sigma_v^2}{T\sigma_\mu^2 + \sigma_v^2}$.

The random effects model

GLS estimation of the random effects model

- Let $\delta = (\alpha, \beta')'$ and $\tilde{X}_i = [\iota_T \; X_i]$. The normal equations for the GLS estimator of δ are written as

$$\left[\sum_{i=1}^{N} \tilde{X}_i' V^{-1} \tilde{X}_i\right] \hat{\delta} = \left[\sum_{i=1}^{N} \tilde{X}_i' V^{-1} y_i\right]. \tag{7}$$

- Since

$$\tilde{X}_i' V^{-1} \tilde{X}_i = \frac{1}{\sigma_v^2} \tilde{X}_i' \left(Q + \psi \bar{J}_T \right) \tilde{X}_i$$
$$= \frac{1}{\sigma_v^2} \left(\tilde{X}_i' \tilde{X}_i - \tilde{X}_i' \bar{J}_T \tilde{X}_i + \psi \tilde{X}_i' \bar{J}_T \tilde{X}_i \right),$$

letting

$$T_{\bar{x}\bar{x}} = \sum_i^N \tilde{X}_i' \tilde{X}_i; B_{\bar{x}\bar{x}} = \sum_i^N \tilde{X}_i' \bar{J}_T \tilde{X}_i; W_{\bar{x}\bar{x}} = T_{\bar{x}\bar{x}} - B_{\bar{x}\bar{x}}$$

we may write

$$\sum_{i=1}^N \tilde{X}_i' V^{-1} \tilde{X}_i = \frac{1}{\sigma_v^2} \left[\left(T_{\bar{x}\bar{x}} - B_{\bar{x}\bar{x}} \right) + \psi B_{\bar{x}\bar{x}} \right]$$
$$= \frac{1}{\sigma_v^2} \left[W_{\bar{x}\bar{x}} + \psi B_{\bar{x}\bar{x}} \right].$$

- In the same manner, letting

$$T_{\bar{x}y} = \sum_{i}^{N} \tilde{X}_i' y_i;\, B_{\bar{x}y} = \sum_{i}^{N} \tilde{X}_i' \bar{J}_T y_i;\, W_{\bar{x}y} = T_{\bar{x}y} - B_{\bar{x}y},$$

we obtain

$$\begin{aligned} \sum_{i=1}^{N} \tilde{X}_i' V^{-1} y_i &= \frac{1}{\sigma_v^2}\left[\left(T_{\bar{x}y} - B_{\bar{x}y}\right) + \psi B_{\bar{x}y}\right] \\ &= \frac{1}{\sigma_v^2}\left[W_{\bar{x}y} + \psi B_{\bar{x}y}\right]. \end{aligned}$$

Thus, the normal equation (7) becomes

$$\left[W_{\bar{x}\bar{x}} + \psi B_{\bar{x}\bar{x}}\right]\hat{\delta} = \left[W_{\bar{x}y} + \psi B_{\bar{x}y}\right].$$

- Further calculations give, letting $\bar{z}_i = \frac{1}{T}\sum_{t=1}^{T} z_{it}$,

$$\begin{bmatrix} \psi NT & \psi T \sum_{i=1}^{N} \bar{x}_i' \\ \psi T \sum_{i=1}^{N} \bar{x}_i & \sum_{i=1}^{N} X_i' Q X_i + \psi T \sum_{i=1}^{N} \bar{x}_i \bar{x}_i' \end{bmatrix}^{-1} \begin{bmatrix} \hat{\alpha}_{GLS} \\ \hat{\beta}_{GLS} \end{bmatrix}$$
$$= \begin{bmatrix} \psi NT \bar{y} \\ \sum_{i=1}^{N} X_i' Q y_i + \psi T \sum_{i=1}^{N} \bar{x}_i \bar{y}_i' \end{bmatrix},$$

from which we obtain

$$\begin{aligned} \hat{\beta}_{GLS} &= \left[\frac{1}{T} \sum_{i=1}^{N} X_i' Q X_i + \psi T \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1} \\ &\quad \times \left[\frac{1}{T} \sum_{i=1}^{N} X_i' Q y_i + \psi T \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y})' \right] \\ &= \Delta \hat{\beta}_b + (I - \Delta) \tilde{\beta}_{LSDV}, \end{aligned}$$

where

$$\Delta = \psi T \left[\sum_{i=1}^{N} X_i' Q X_i + \psi T \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1} \times \left[\sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right],$$

$$\hat{\beta}_b = \left[\sum_{i=1}^{N} (\bar{x}_{i.} - \bar{x}_{..})(\bar{x}_{i.} - \bar{x}_{..})' \right]^{-1} \left[\sum_{i=1}^{N} (\bar{x}_{i.} - \bar{x}_{..})(\bar{y}_{i.} - \bar{y}_{..})' \right].$$

- The estimator $\hat{\beta}_b$ is called the between-group estimator because it ignores variation within the group. This formula shows that the $\hat{\beta}_{GLS}$ is a weighted average of $\hat{\beta}_b$ and $\tilde{\beta}_{LSDV}$.

- In addition,

$$\hat{\mu}_{GLS} = \bar{y}_{..} - \hat{\beta}'_{GLS}\bar{x}_{..}.$$

  - $Var(\hat{\beta}_{GLS}) = \sigma_v^2 \left[\sum_{i=1}^N X_i' Q X_i + \psi T \sum_{i=1}^N (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1}.$
  - $Var(\tilde{\beta}_{LSDV}) - Var(\hat{\beta}_{GLS}) \geq 0$. (Use the relation

$$A \geq B \text{ implies } B^{-1} \geq A^{-1}$$

    and the fact that $\psi > 0$ to show this.)
  - For fixed N, $\psi \to 0$ as $T \to \infty$. Thus, for large T, $\tilde{\beta}_{LSDV}$ and $\hat{\beta}_{GLS}$ are close to each other.

The random effects model

- Estimating σ_μ^2 and σ_v^2

  Note that

$$\bar{y}_{i.} = \alpha + \beta \bar{x}_{i.} + \mu_i + \bar{v}_{i.}$$

and

$$y_{it} - \bar{y}_{i.} = (x_{it} - \bar{x}_{i.})' \beta + v_{it} - \bar{v}_{i.}.$$

Thus, we can use the LSDV and between group residuals. That is,

$$\hat{\sigma}_v^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T \left[(y_{it} - \bar{y}_{i.}) - \tilde{\beta}'_{LSDV} (x_{it} - \bar{x}_{i.}) \right]^2}{N(T-1) - K}$$

and

$$\hat{\sigma}_\mu^2 = \frac{\sum_{i=1}^N (\bar{y}_{i.} - \hat{\alpha}_b - \hat{\beta}_b \bar{x}_{i.})^2}{N - (K+1)} - \frac{1}{T} \sigma_v^2.$$

Using these, the FGLS estimator can be devised. Note that the divisor for $\hat{\sigma}_v^2$ is chosen to be $NT - N - T$ in order to make it unbiased.

The random effects model

- Let
$$\lambda = 1 - \sqrt{\frac{\sigma_v^2}{T\sigma_\mu^2 + \sigma_v^2}}.$$
- The RE estimator is also obtained by running regression on
$$y_{it} - \lambda \bar{y}_{i.} = \beta'(x_{it} - \lambda \bar{x}_{i.}) + u_{it} - \lambda \bar{u}_{i.}.$$
  1. If $\lambda = 1$, then this is just the fixed effects estimator.
  2. So, the bigger the variance of the unobserved effect, the closer it is to FE.

Comparing the fixed- and random- effects models

- The fixed-effects model do not require assuming that the individual effect variable and the regressors are independent.
- The fixed-effects model has the problem of incidental parameters.
- In the random-effects model, the number of parameters is fixed and efficient estimation methods can be derived.
- In the random-effects model, one has to assume no correlation between the individual effect variable and the regressors.

Comparing the fixed- and random- effects models

- Hausman's test
  - $H_0 : E(\mu_i \mid x_{it}) = 0$
  - Test statistic

$$m = (\tilde{\beta}_{LSDV} - \hat{\beta}_{GLS})' \left(Var(\tilde{\beta}_{LSDV}) - Var(\hat{\beta}_{GLS}) \right)^{-1} (\tilde{\beta}_{LSDV} - \hat{\beta}_{GLS})$$

  - As $N \to \infty$, $m \xrightarrow{d} \chi^2(K)$.

Estimation of models with individual specific variables

- The model and estimation

$$\begin{aligned} y_{it} &= \alpha + \beta' x_{it} + \rho' z_i + u_{it}; \\ u_{it} &= \mu_i + v_{it}. \end{aligned} \qquad (8)$$

The LSDV estimator of β is, as before,

$$\tilde{\beta}_{LSDV} = (X'QX)^{-1} X'Qy.$$

Estimation of models with individual specific variables

- The individual mean over time satisfies

$$\bar{y}_{i.} - \bar{x}_{i.}'\beta = \alpha + \rho' z_i + \mu_i + \bar{v}_{i.}$$

If μ_i are random variables uncorrelated with the regressors, the parameters α and ρ are estimated by running OLS on this model assuming $\mu_i + \bar{v}_{i.}$ is an error term and substituting $\tilde{\beta}_{LSDV}$ for β. These estimators are consistent when $N \to \infty$.

Estimation of models with individual specific variables

- More efficient estimators of α, β and ρ can be obtained by GLS (cf. Hsiao, p.53)
- Model (8) can further be generalized by the specification

$$\begin{aligned} y_{it} &= \alpha + \beta' x_{it} + \rho' z_i + \gamma' w_t + u_{it}; \\ u_{it} &= \mu_i + \lambda_t + v_{it}. \end{aligned}$$

See Hsiao for further details.

- Consider the model

$$y_{it} = \alpha + \gamma y_{i,t-1} + \mu_i + v_{it}, \ (t = 1, ..., T; i = 1, ..., N).$$

- Assume

1. $|\gamma| < 1$.
2. y_{i0} are observable.
3. $v_{it} \sim iid(0, \sigma^2)$ for all i.
4. $\{v_{1t}\}, ..., \{v_{Nt}\}$ are independent.
5. $E(v_{it}\mu_i) = 0$ for all i and t.

- The LSDV estimator of γ is

$$\begin{aligned}\hat{\gamma} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\bar{y}_i)(y_{i,t-1}-\bar{y}_{i,-1})}{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1}-\bar{y}_{i,-1})^2} \\ &= \gamma + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1}-\bar{y}_{i,-1})(v_{i,t}-\bar{v}_i)/NT}{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1}-\bar{y}_{i,-1})^2/NT}.\end{aligned} \tag{9}$$

Continuous substitution gives

$$y_{it} = v_{it} + \gamma v_{i,t-1} + \ldots + \gamma^{t-1} v_{i1} + \frac{1-\gamma^t}{1-\gamma}(\alpha + \mu_i) + \gamma^t y_{i0}.$$

Summing $y_{i,t-1}$ over t, we have

$$\begin{aligned} \sum_{t=1}^{T} y_{i,t-1} &= \frac{1-\gamma^T}{1-\gamma} y_{i0} + \frac{(T-1) - T\gamma + \gamma^T}{(1-\gamma)^2}(\alpha + \mu_i) \\ &\quad + \frac{1-\gamma^{T-1}}{1-\gamma} v_{i1} + \frac{1-\gamma^{T-2}}{1-\gamma} v_{i2} + \ldots + v_{i,T-1}. \end{aligned} \tag{10}$$

To analyze the probability limit of the numerator of estimator (9), consider the relations

$$
\begin{aligned}
& p \lim_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{i,t-1} - \bar{y}_{i,-1})(v_{i,t} - \bar{v}_i) \\
= \quad & p \lim_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{i,t-1} - \bar{y}_{i,-1}) v_{i,t} \\
= \quad & -p \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \bar{y}_{i,-1} \bar{v}_i,
\end{aligned}
$$

where the second equality follows since $p \lim_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{i,t-1} v_{i,t} = 0$.

But using (10) and the given assumptions, we find

$$-p \lim_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \bar{y}_{i,-1} \bar{v}_i = -\frac{\sigma^2}{T^2} \frac{(T-1) - T\gamma + \gamma^T}{(1-\gamma)^2}, \tag{11}$$

which is the probability limit of the numerator of estimator (9). Similarly, for the denominator of estimator (9), we obtain

$$\begin{aligned} & p \lim_{N\to\infty} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{i,t-1} - \bar{y}_{i,-1})^2 / NT \\ = \; & \frac{\sigma^2}{1-\gamma^2} \left\{ 1 - \frac{1}{T} - \frac{2\gamma}{(1-\gamma)^2} \frac{(T-1) - T\gamma + \gamma^T}{T^2} \right\}. \end{aligned} \tag{12}$$

Remarks (i) Relations (11) and (12) show that $\hat{\gamma}$ is inconsistent for fixed T.
(ii) This is caused by having to eliminate the unknown individual effects μ_i from each observation, which creates correlation between $y_{i,t-1} - \bar{y}_{i,-1}$ and $v_{i,t} - \bar{v}_i$.
(iii) The inconsistency of the LSDV estimator holds whether μ_i are random or fixed.
(iv) The asymptotic bias will die out as T increases.

- Consider the random effects model

$$y_{it} = \gamma y_{i,t-1} + \mu_i + v_{it}, (t = 1, ..., T; i = 1, ..., N), \tag{13}$$

where μ_i is a random variable.

- The OLS estimator of γ is

$$\begin{aligned}\hat{\gamma} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it} y_{i,t-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t-1}^2} \\ &= \gamma + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}(\mu_i + v_{it})}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t-1}^2}.\end{aligned}$$

Using the same methods as in the previous subsection, we have

$$\begin{aligned}& p\lim_{N\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}(\mu_i + v_{it}) \\ = \; & \frac{1}{T}\frac{1-\gamma^T}{1-\gamma} Cov(y_{i0}, \mu_i) + \frac{1}{T}\frac{\sigma^2}{(1-\gamma)^2}\left[(T-1) - T\gamma + \gamma^T\right]\end{aligned}$$

and

$$
\begin{aligned}
p\lim_{N\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t-1}^2 \quad = \quad & \frac{1-\gamma^{2T}}{T(1-\gamma^2)}\frac{\sum_{i=1}^{N} y_{i0}^2}{N} \\
& +\frac{\sigma^2}{(1-\gamma)^2}\frac{1}{T}\left(T-2\frac{1-\gamma^T}{1-\gamma}+\frac{1-\gamma^{2T}}{1-\gamma^2}\right) \\
& +\frac{2}{T(1-\gamma)}\left(\frac{1-\gamma^T}{1-\gamma}-\frac{1-\gamma^{2T}}{1-\gamma^2}\right) Cov(\mu_i, y_{i0} \\
& +\frac{\sigma^2}{T(1-\gamma^2)^2}\left[(T-1)-T\gamma^2+\gamma^{2T}\right].
\end{aligned}
$$

Thus, the OLS estimator is not consistent.

- Taking the difference of model (13), we obtain

$$y_{it} - y_{i,t-1} = \gamma \left(y_{i,t-1} - y_{i,t-2} \right) + v_{it} - v_{i,t-1}.$$

  Since $y_{i,t-2} - y_{i,t-3}$ is uncorrelated with $v_{it} - v_{i,t-1}$ and correlated with $y_{i,t-1} - y_{i,t-2}$, it can be used as an instrument (cf. Anderson and Hsiao, 1981, JASA).

- For $t = 3$, we have

$$y_{i3} - y_{i2} = \gamma \left(y_{i,2} - y_{i,1} \right) + v_{i3} - v_{i,2}.$$

  Thus, y_{i1} is a valid instrument.

- For $t = 4$, we have

$$y_{i4} - y_{i3} = \gamma \left(y_{i3} - y_{i2} \right) + v_{i4} - v_{i3}$$

  In this case, y_{i1} and y_{i2} are valid instruments.
- For period T, the set of instruments becomes $\left(y_{i1}, y_{i2}, ..., y_{i.T-2} \right)$.

- Letting $\Delta v_i = (v_{i3} - v_{i2}, ..., v_{iT} - v_{i,T-1})'$, we find

$$E\Delta v_i \Delta v_i' = \sigma_v^2 G,$$

where $G = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$

- Define

$$W_i = \begin{pmatrix} [y_{i1}] & & & 0 \\ & [y_{i1}, y_{i2}] & & \\ & & \ddots & \\ 0 & & & [y_{i1}, y_{i2}, ..., y_{i.T-2}] \end{pmatrix}$$

and $W = [W_1', ..., W_N']'$. Premultiplying the differenced equation by W, we get

$$W'\Delta y = W'\Delta y_{-1}\gamma + W'\Delta v.$$

Performing GLS on this model, Arellano and Bond (1991, RES) obtains

$$\begin{aligned} \hat{\gamma} &= \left[(\Delta y_{-1})' W(W'(I_N \otimes G)W)^{-1} W' (\Delta y_{-1})\right]^{-1} \\ &\quad \times \left[(\Delta y_{-1})' W(W'(I_N \otimes G)W)^{-1} W' (\Delta y)\right]. \end{aligned}$$

- The optimal GMM estimator is

$$\tilde{\gamma} = \left[(\Delta y_{-1})' W (\sum_{i=1}^{N} W_i' (\Delta v_i)(\Delta v_i)' W)^{-1} W' (\Delta y_{-1}) \right]^{-1} \times \left[(\Delta y_{-1})' W (\sum_{i=1}^{N} W_i' (\Delta v_i)(\Delta v_i)' W)^{-1} W' (\Delta y) \right].$$

To make this estimator operational, replace Δv_i with $\Delta \hat{v}_i$ obtained from $\hat{\gamma}$.

- See also Arellano and Bover (1995, JoE), Blundell and Bond (1998, JoE) for related research.

Exercises

1. Consider the panel data model

$$y_{it} = \mu + \beta x_{it} + u_{it}, (i = 1, ..., N; t = 1, ..., T) \tag{14}$$

where $x_{it} \sim iid(0, \sigma_x^2)$, $u_{it} = \mu_i + v_{it}$, $v_{it} \sim iid(0, \sigma_v^2)$, and x_{it} is independent of v_{js} for all i, t, j and s.

a. Assuming that $N \to \infty$ and that T is fixed, calculate the variance of $\sqrt{N}(\hat{\beta}_{FE} - \beta)$, where $\hat{\beta}_{FE}$ is the fixed effect estimator of β.

b. Consider the differenced model

$$\Delta y_{it} = \beta \Delta x_{it} + \Delta v_{it}.$$

Under the same assumptions as in part (a), calculate the variance $\sqrt{N}(\hat{\beta}_{OLS} - \beta)$, where $\hat{\beta}_{OLS}$ is the OLS estimator of β using the differenced model.

c. Which estimator is more efficient?

Exercises

2. Consider the following two-period fixed effects model with a single regressor, x_{it}

$$y_{it} = \lambda_i + \alpha x_{it} + u_{it}, \ (i = 1, \ldots, n; \ t = 1, 2), \tag{15}$$

where

$$x_{it} = z_i + a_{it} \text{ and } u_{it} = v_i + b_{it} \tag{16}$$

and z_i and v_i are random variables. As usual, λ_i is an individual effects variable correlated with x_{it}. Observed data are $\{y_{it}\}$ and $\{x_{it}\}$. But $\{z_i\}$ and $\{a_{it}\}$ are not separately observed. Random variables $\{a_{it}\}$ and $\{b_{it}\}$ bring time series variations to the observed data. For $\{a_{it}\}$ and $\{b_{it}\}$, assume

$\begin{pmatrix} n^{\beta} a_{it} \\ n^{\gamma} b_{it} \end{pmatrix} \sim iid \left(0, \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{bmatrix} \right)$ for every n, i and t.

a. Show that the Within-OLS and first-differenced estimators are identical for this two-period fixed effects model.

Exercises

(continued)

b. The first-differenced estimator is written as

$$\hat{\alpha}_d = \frac{\sum_{i=1}^{n} \Delta x_i \Delta y_i}{\sum_{i=1}^{n} (\Delta x_i)^2} = \alpha + \frac{\sum_{i=1}^{n} \Delta a_i \Delta b_i}{\sum_{i=1}^{n} (\Delta a_i)^2}$$

where $\Delta w_i = w_{i2} - w_{i1}$. Find the limiting distribution of $\frac{n^{\beta+\gamma}}{\sqrt{n}} \sum_{i=1}^{n} \Delta a_i \Delta b_i$ and the probability limit of $\frac{n^{2\beta}}{n} \sum_{i=1}^{n} (\Delta a_i)^2$ when $n \to \infty$. Using these results, find the limiting distribution of $n^{\frac{1}{2}-\beta+\gamma}(\hat{\alpha}_d - \alpha)$.

c. When is the estimator $\hat{\alpha}_d$ consistent?