Econometrics II TA Session #5*

Makoto SHIMOSHIMIZU†

November 19, 2019

Contents

*All comments welcome!

†E-mail: vgm067sm@student.econ.osaka-u.ac.jp

1 Cross Section Data and Panel Data

1.1 Cross Section Data

Cross section data is a one with individual specified structure at a fixed time point. Therefore, both dependent and independent varialbles are dependent on the individual $i \in \{1, \ldots, n\}$ (if there are $n \in \mathbb{N}_{++} := \{1, 2, \ldots\}$ observations).

1.2 Panel Data

Different from the cross section data, **panel data** is a one with time and individual specified structure. This fact infers that both dependent and independent variables are dependent on time $t \in \{1, \ldots, T\}$ as well as the individual $i \in \{1, \ldots, n\}$.

2 Panel Data: Model

In this section, the structure of panel data is explained. The model for panel data is basically given by

$$y_{it} = X_{it}\beta + \nu_i + u_{it}, \quad i = 1, 2, \ldots, n, \quad t = 1, 2, \ldots, T, \tag{2.1}$$

where i indicates the individual and t denotes time. There are $n \in \{1, 2, \ldots\}$ observations for each $t \in \{1, 2, \ldots, T\}$. Then, u_{it} indicates the error term. Here we have the following assumption:

- $\mathbb{E}[u_{it}] = 0$ for all $i \in \{1, \ldots, n\}$;
- $\mathbb{V}[u_{it}] = \sigma_u^2$ for all $i \in \{1, \ldots, n\}$;
- $\mathrm{Cov}[u_{it}, u_{js}] = 0$ for all $i \neq j$ and $t \neq s$.

ν_i represents the individual effect, which is **fixed** or **random**. The difference is shown in the next definition.

Definition 2.1 (Fixed Effect Model and Random Effect Model)**.** The difference of the two model analysis is described as follows:

- a **random effect analysis** puts ν_i into the error term with ν_i and X_{it} (in addition to u_{it}) being orthogonal, which exploits the serial correlation in the composite error $\nu_i + u_{it}$ in a generalized least squares (GLS) framework;
- a **fixed effect analysis** allows ν_i to be arbitrary correlated with X_{it} (but not with u_{it}, i.e., $\mathbb{E}[u_{it}|X_{it}] = 0$ for all $t \in \{1, \ldots, T\}$ and $i \in \{1, \ldots, n\}$), which is the whole point of using panel data.

The analysis of panel data, regardless of fixed or random effect model, allows the model builder to learn about economic processes while accounting for both heterogeneity across individuals, firms, countries, and so on for dynamic effects that are not invisible in cross section.

2.1 Individual Effect

The individual effect $\nu_i := z_i\alpha$ for $i \in \{1, \dots, n\}$, as its name indicates, can vary across the individual. z_i may be observable (ex. race, sex) or unobservable (ex. skill, preference).

Remark 2.1 (Difference between Fixed Effect Model and Random Effect Model)**.** If

- z_i is correlated with X_{it} and $\nu_i := z_i\alpha$ is constant across $i \in \{1, \dots, n\}$, or $\mathbb{E}[\varepsilon|X] \neq 0$ where $\varepsilon := (1_T \otimes I_n)\boldsymbol{\nu} + u$, then we use fixed effect model;
- z_i is uncorrelated with X_{it} or $\mathbb{E}[\varepsilon|X] = 0$, then we use ramdom effect model.

2.2 Biased Estimator for Panel Data

If the individual effects defined above are all zero, then we can estimate the coefficient β by a usual ordinary least squares (OLS) method. However, if some (or all) individual effects are not zero and correlates with the explanatory variables, the OLS estimator is biased since

$$\begin{aligned}
\mathbb{E}[\beta_{OLS}|X] &= \mathbb{E}[(X'X)^{-1} X'y|X] \\
&= \mathbb{E}[(X'X)^{-1} X' \left(X'\beta + (1_T \otimes I_n)\boldsymbol{\nu} + u\right)|X] \\
&= \beta + \mathbb{E}[(X'X)^{-1} X' \left((1_T \otimes I_n)\boldsymbol{\nu} + u\right)|X] \\
&\neq \beta
\end{aligned}$$

where

$$y = X\beta + (1_T \otimes I_n)\boldsymbol{\nu} + u \tag{2.2}$$

represents the stacked form of Eq. (2.1) (with respect to both time and indivudual). This reveals the necessity for alternative methods to estimate β, one of which is a fixed effect analysis.

3 Least Squares Dummy Variable (LSDV) Model and Within Model

When we estimate a fixed effect model, there exist two kinds of estimator: least squares dummy variable (LSDV) estimator and Within estimator.

3.1 LSDV model

LSDV model stems from the following regression model:

$$y_{it} = X_{it}\beta + D_i\nu_i + u_{it} \tag{3.1}$$

where D_i is a dummy variable for each individual. LSDV estimation is a normal OLS estimation with dummy variables added and we can estimate the estimator with a simple method. However, when the number of individuals counts so large, it takes time to calculate the computer since the number of explanatory variables also increases.

3.2 Within Model

In order to avoid this problem, **Within model** is an alternative method to estimate. We obtain the Within estimator from the following equation:

$$\tilde{y}_{it} = \tilde{X}_{it}\beta + u_{it} \tag{3.2}$$

where

$$\tilde{y}_{it} := y_{it} - \overline{y}_i; \quad \tilde{X}_{it} := X_{it} - \overline{X}_i, \tag{3.3}$$

and $\overline{y}_i$ and $\overline{X}_i$ are the average with respect to time for each individual.

Remark 3.1 (Estimator of the Two Model)**.** The estimator obtained in Within Model **coincides with** the one in LSDV model.

4 The Fixed Effect Model

In this section, we first review and derive the estimator for fixed effect model. Then, the properties of the estimator are stated with its derivation. The estimator for fixed effect model is given as follows.

Theorem 4.1 (Estimator for Fixed Effect Model)**.** The estimator for the fixed effect model is given by

$$\hat{\beta}_{FE} = \left(X'(I_n \otimes D_T)X\right)^{-1} X'(I_n \otimes D_T)y. \tag{4.1}$$

The next subsection shows how to derive the above estimator.

4.1 Estimator: Derivation

Taking the average of the model:

$$y_{it} = X_{it}\beta + \nu_i + u_{it}$$

with respect to time $t \in \{1, 2, \ldots, T\}$ yields

$$\overline{y}_i = \overline{X}_i\beta + \nu_i + \overline{u}_i, \quad i = 1, 2, \ldots, n, \tag{4.2}$$

where

$$\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}; \quad \overline{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}; \quad \overline{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{it}.$$

Therefore, by subtracting Eq. (4.2) from Eq. (2.1), we obtain

$$y_{it} - \overline{y}_i = (X_{it} - \overline{X}_i)\beta + (u_{it} - \overline{u}_i), \quad i = 1, 2, \ldots, n, \quad t = 1, 2, \ldots, T. \tag{4.3}$$

By the operation above, the individual effect ν_i disappear, which is one of the key point of the fixed effect model.

Here we have the following representation for $\overline{y}_i$, $\overline{X}_i$ and $\overline{u}_i$:

$$y_{it} - \overline{y}_i = y_{it} - \frac{1}{T}1_T' y_i \tag{4.4}$$

where

$$1_T = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},$$

which means $1_T \in \mathbb{R}^{T\times 1}$ (a $T \times 1$ vector), and

$$y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix},$$

which also means $y_i \in \mathbb{R}^{T\times 1}$ (a $T \times 1$ vector). Then,

$$\begin{pmatrix} y_{i1} - \overline{y}_i \\ y_{i2} - \overline{y}_i \\ \vdots \\ y_{iT} - \overline{y}_i \end{pmatrix} = I_T y_i - 1_T \overline{y}_i = I_T y_i - \frac{1}{T}1_T 1_T' y_i = (I_T - \frac{1}{T}1_T 1_T') y_i \tag{4.5}$$

holds. Thus, by defining

$$X_i := \begin{pmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{iT} \end{pmatrix}$$

and applying the same rearrangement above, we have

$$(I_T - \frac{1}{T}1_T 1_T') y_i = (I_T - \frac{1}{T}1_T 1_T') X_i \beta + (I_T - \frac{1}{T}1_T 1_T') u_i, \quad i = 1, 2, \ldots, n. \tag{4.6}$$

Define

$$D_T := I_T - \frac{1}{T}1_T 1_T', \tag{4.7}$$

which is a $T \times T$ matrix. Then, the above equation is rewritten as

$$D_T y_i = D_T X_i \beta + D_T u_i, \quad i = 1, 2, \ldots, n. \tag{4.8}$$

Note that $D_T' D_T = D_T$, id est, D_T is a symmetric and idempotent matrix. Using the following notations:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} (\in \mathbb{R}^{TN\times 1}); \quad X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} (\in \mathbb{R}^{TN\times k}); \quad u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} (\in \mathbb{R}^{TN\times 1}),$$

we obtain the following stacked model for fixed effect model:

$$\begin{pmatrix} D_T & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & D_T & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & D_T \end{pmatrix} y = \begin{pmatrix} D_T & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & D_T & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & D_T \end{pmatrix} X\beta + \begin{pmatrix} D_T & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & D_T & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & D_T \end{pmatrix} u.$$

Here we review some concepts related to the **Kronecker product**.

Review (Kronecker Product: Definition and Properties)**.** Let the two matrix A and B be $n \times m$ and $T \times k$, respectively. Then the **Kronecker product** is defined as

$$A \otimes B := \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{pmatrix} \in \mathbb{R}^{nT \times mk}.$$

The Kronecker Product has, by its definition, some important properties. (In the following, assume $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$).

(i) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$;

(ii) $|A \otimes B| = |A|^m |B|^n$:

(iii) $(A \otimes B)' = A' \otimes B'$;

(iv) $Tr(A \otimes B) = Tr(A)Tr(B)$.

Moreover, for A, B, C and D such that the products (of the following equations) are defined, then

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

Using the **Kronecker product**, we obtain the following expression:

$$\underbrace{(I_n \otimes D_T)}_{\in \mathbb{R}^{Tn \times Tn}} \underbrace{y}_{\in \mathbb{R}^{Tn \times 1}} = \underbrace{(I_n \otimes D_T)}_{\in \mathbb{R}^{Tn \times Tn}} \underbrace{X}_{\in \mathbb{R}^{Tn \times k}} \underbrace{\beta}_{\in \mathbb{R}^{k \times 1}} + \underbrace{(I_n \otimes D_T)}_{\in \mathbb{R}^{Tn \times Tn}} \underbrace{u}_{\in \mathbb{R}^{Tn \times 1}} . \tag{4.9}$$

Note that the inverse matrix of D_T is **not available**, as the rank of D_T is $T-1$, not T (full rank). Thus, applying the **OLS method** to the above regression model, we obtain

$$\begin{aligned} \hat{\beta}_{FE} &= \left(((I_n \otimes D_T)X)'(I_n \otimes D_T)X\right)^{-1} ((I_n \otimes D_T)X)'(I_n \otimes D_T)y \\ &= \left(X'(I_n \otimes D_T' D_T)X\right)^{-1} X'(I_n \otimes D_T' D_T)y \\ &= \left(X'(I_n \otimes D_T)X\right)^{-1} X'(I_n \otimes D_T)y, \end{aligned} \tag{4.10}$$

which yields Eq. (4.1). □

4.2 Estimator: Properties

Here we show some properties of $\hat{\beta}_{FE}$. (For more detail, see [2])

Theorem 4.2 (Properties of the Estimator for the Fixed Effect Model)**.** Under the assumptions we set above and some additional appropriate assumptions, the estimator $\hat{\beta}_{FE}$ obtained above has the following properties.

(i) **Unbiasedness** estimator $\hat{\beta}_{FE}$ becomes an unbiased estimator:

$$\mathbb{E}[\hat{\beta}_{FE}] = \beta. \tag{4.11}$$

(ii) **Consistency** the estimator $\hat{\beta}_{FE} = (X'(I_n \otimes D_T)X)^{-1} X'(I_n \otimes D_T)y$ satisfies

$$\hat{\beta}_{FE} \xrightarrow{p} \beta \quad \text{or} \quad \operatorname*{plim}_{n\to\infty} \hat{\beta}_{FE} = \beta. \tag{4.12}$$

Proof. As for the unbiasedness, we can prove Eq. (4.11) directly as follows.

(i) **Unbiasedness** In the following, define

$$\mathbf{y} := (I_n \otimes D_T)y; \quad \mathbf{X} := (I_n \otimes D_T)X; \quad \mathbf{u} := (I_n \otimes D_T)u. \tag{4.13}$$

Then, Eq. (4.9) is rewritten as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u},$$

and accordingly the obtained estimator:

$$\begin{aligned} \hat{\beta}_{FE} &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}\mathbf{y} \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}\mathbf{u}. \end{aligned} \tag{4.14}$$

Thus, under the assumption that $\underline{\mathbb{E}[\mathbf{u}|\mathbf{X}] = \mathbf{0}}$,

$$\begin{aligned} \mathbb{E}[\hat{\beta}_{FE}|\mathbf{X}] &= \mathbb{E}[\beta + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}\mathbf{u}|\mathbf{X}] \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}\, \underbrace{\mathbb{E}[\mathbf{u}|\mathbf{X}]}_{=\mathbf{0}} \\ &= \beta, \end{aligned}$$

which yields

$$\mathbb{E}[\hat{\beta}_{FE}] = \mathbb{E}[\mathbb{E}[\hat{\beta}_{FE}|\mathbf{X}]] = \beta. \tag{4.15}$$

Note that if $\mathbb{E}[\mathbf{u}|\mathbf{X}] = \mathbf{0}$, then $\mathbb{E}[\mathbf{u}] = \mathbf{0}$, which meets the requirements mentioned above.

(ii) **Consistency** From Eq. (4.14), we have

$$\hat{\beta}_{FE} = \beta + \left(\frac{1}{Tn}\sum_{i=1}^{Tn} \mathbf{X}_i'\mathbf{X}_i\right)^{-1} \left(\frac{1}{Tn}\sum_{i=1}^{Tn} \mathbf{X}_i'\mathbf{u}_i\right).$$

where

$$\mathbf{X} := \begin{pmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_{Tn} \end{pmatrix}; \quad \mathbf{u} := \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{Tn} \end{pmatrix}, \tag{4.16}$$

and $\mathbf{X}_i \in \mathbb{R}^{1\times k}$ for all $i \in \{1, \dots, Tn\}$. By taking the probability limit on both sides, we have

$$\operatorname*{plim}_{Tn\to\infty} \beta_{OLS} = \beta + \operatorname*{plim}_{Tn\to\infty} \left(\frac{1}{n}\sum_{i=1}^{Tn} \mathbf{X}_i'\mathbf{X}_i\right)^{-1} \operatorname*{plim}_{Tn\to\infty} \left(\frac{1}{Tn}\sum_{i=1}^{Tn} \mathbf{X}_i'\mathbf{u}_i\right). \tag{4.17}$$

Here we apply the **convergence of the product of random variables in probability**. If we assume $\mathbb{E}\left[\mathbf{X}_i'\mathbf{X}_i\right] < \infty$ for all $i \in \{1, \dots, Tn\}$, from the **weak law of large numbers** (WLLN),

$$\frac{1}{Tn}\sum_{i=1}^{Tn} \mathbf{X}_i'\mathbf{X}_i \xrightarrow{p} \mathbb{E}\left[\mathbf{X}_i'\mathbf{X}_i\right] < \infty; \tag{4.18}$$

$$\frac{1}{Tn}\sum_{i=1}^{Tn} \mathbf{X}_i'\mathbf{u}_i \xrightarrow{p} \mathbb{E}\left[\mathbf{X}_i'\mathbf{u}_i\right] = \mathbf{0}(\in \mathbb{R}^k). \tag{4.19}$$

Eq. (4.19) holds from the **orthogonal condition** with respect to $\mathbf{X}$ and $\mathbf{u}$: $\mathbb{E}[\mathbf{u}|\mathbf{X}] = \mathbf{0}$. In addition,

$$\operatorname*{plim}_{Tn\to\infty} \left(\frac{1}{Tn}\sum_{i=1}^{Tn} \mathbf{X}_i'\mathbf{X}_i\right)^{-1} = \mathbb{E}\left[\mathbf{X}_i'\mathbf{X}_i\right]^{-1} \tag{4.20}$$

holds from the **continuous mapping theorem**. Thus, substituting Eq. (4.18) and Eq. (4.19) into Eq. (4.17) results in

$$\operatorname*{plim}_{Tn\to\infty} \hat{\beta}_{FE} = \beta + \mathbb{E}\left[\mathbf{X}_i'\mathbf{X}_i\right]^{-1}\mathbf{0} = \beta,$$

which indicates that $\hat{\beta}_{FE} \xrightarrow{p} \beta$.

$\square$

Remark 4.1 (Efficiency)**.** As for the efficiency, the OLS estimator $\hat{\beta}_{FE}$ is NOT eficient unless $\mathbb{E}[X'\left((I_T \otimes 1_n)\boldsymbol{\nu} + u\right)] = \mathbf{0}$ since

$$\mathbb{V}[\beta_{GLS}] \leq \mathbb{V}[\beta_{OLS}] \tag{4.21}$$

always holds (in the fixed effect model).

4.3 Individual Effect of LSDV Model

In the LSDV model, we can recover the individual effect as follows:

$$\begin{aligned} &\hat{\nu}_i = \overline{y}_i - \overline{X}_i\hat{\beta}_{FE} \\ \iff &\hat{\nu}_i = Z_i\alpha + \epsilon_i, \end{aligned}$$

where it is assumed that the individual–specific effect depends on Z_i. The estimator of σ_u^2 is given by

$$\hat{\sigma}_u^2 = \frac{1}{Tn - k - n}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - X_{it}\hat{\beta} - \hat{\nu}_i)^2. \tag{4.22}$$

Appendix

A Properties of D_T

Here we see and prove some properties of $D_T \in \mathbb{R}^{T\times T}$.

Properity (1) $D_T' = D_T$ (symmetric matrix).

Proof. Direct culculations yield

$$D_T' = (I_T - \frac{1}{T}1_T 1_T')' = I_T - \frac{1}{T}(1_T 1_T')' = I_T - \frac{1}{T}1_T 1_T' = D_T.$$

$\square$

Properity (2) $D_T D_T = D_T$ (idempotent matirx).

Proof. Direct culculations yield

$$\begin{aligned}
D_T D_T &= (I_T - \frac{1}{T}1_T 1_T')(I_T - \frac{1}{T}1_T 1_T') \\
&= I_T - \frac{1}{T}1_T 1_T' - \frac{1}{T}1_T 1_T' + \frac{1}{T^2}1_T \underbrace{1_T' 1_T}_{=T} 1_T' \\
&= I_T - \frac{1}{T}1_T 1_T' - \frac{1}{T}1_T 1_T' + \frac{1}{T}1_T 1_T' \\
&= I_T - \frac{1}{T}1_T 1_T' \\
&= D_T.
\end{aligned}$$

$\square$

Properity (3) Rank of D_T is $T-1$, not T.

Proof. To show the result, we use the following theorem.

Theorem A.1 (Symmetric and Idempotent Matrix)**.** Assume that $A \in \mathbb{R}^{n\times n}$ is symmetric and idempotent. Then $Rank(A) = Tr(A)$.

You can show this theorem by using the following facts:

(i) The eigenvalues of an idempotent matrix are zeros or ones;

(ii) Any symmetric matrix has a **spectral decomposition** with an orthogonal matrix (which consists of eigen vectors of A).

Since D_T is a symmetric and idempotent matrix,

$$\begin{aligned}
Rank(D_T) &= Tr(D_T) \\
&= Tr(I_T - \frac{1}{T} 1_T 1_T') \\
&= Tr \underbrace{\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}}_{\in \mathbb{R}^{T \times T}} - \frac{1}{T} Tr \underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}}_{\in \mathbb{R}^{T \times T}} \\
&= T - \frac{1}{T} T \\
&= T - 1 < T.
\end{aligned}$$

Therefore, Rank of D_T is $T - 1$, not T. $\square$

References

[1] Greene, W. H., *Econometric analysis Seventh Edition*. 2012, Pearson.

[2] Hausman, J. A. and Taylor, W. E., Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric Society*, **1377–1398**, 1981.