Econometrics II TA Session #5

Ryo Sakamoto

Nov 17, 2021

Contents

Panel Data Analysis

- Consider the model:

$$y_{it} = X_{it}\beta + v_i + u_{it}, \quad i = 1, \cdots, n, \; t = 1, \cdots, T \tag{1}$$

  where i indicates individual and t denotes time.

- u_{it} is the error term such that

$$\mathbb{E}[u_{it}] = 0, \forall i \text{ and } \forall t,$$
$$V(u_{it}) = \sigma_u^2, \forall i \text{ and } \forall t,$$
$$Cov(u_{it}, u_{js}) = 0, \forall i \neq j \text{ and } \forall t \neq s.$$

Panel Data Analysis

- Consider the model:

$$y_{it} = X_{it}\beta + v_i + u_{it}, \quad i = 1, \cdots, n, \; t = 1, \cdots, T$$

- v_i is the **individual effect**, e.g., IQ, patience and so on.
  - If we treat v_i as a variable that is correlated with the explanatory variables X_{it}, then we choose a fixed effect model to deal with endogeneity.
  - If we treat v_i as a variable that is NOT correlated with the explanatory variables X_{it}, then we CAN choose a random effect model.
- Throughout this slide, X_{it} is assumed to be a **random variable.**

- Suppose that the individual effect v_i is correlated with X_{it}.
- Taking the average of each variable w.r.t. time, we have

$$\overline{y}_i = \overline{X}_i\beta + v_i + \overline{u}_i, \quad i = 1, \cdots, n, \tag{2}$$

where

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

- Note that the individual effect v_i is time invariant.

- Taking a difference between equation (1) and (2),

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

- Note that

$$\begin{aligned} y_{it} - \overline{y}_i &= y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it} \\ &= y_{it} - \frac{1}{T}(y_{i1} + \cdots + y_{iT}) \\ &= y_{it} - \frac{1}{T}(1 \times y_{i1} + \cdots + 1 \times y_{iT}) \\ &= y_{it} - \frac{1}{T}(1, \cdots, 1)'(y_{i1}, \cdots, y_{iT}) \\ &= y_{it} - \frac{1}{T}1_T' y_i. \end{aligned}$$

- where

$$1_T = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^T, \quad y_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix} \in \mathbb{R}^T.$$

- Using this notation, we have

$$\begin{pmatrix} y_{i1} - \overline{y}_i \\ \vdots \\ y_{iT} - \overline{y}_i \end{pmatrix} = \begin{pmatrix} X_{i1} - \overline{X}_i \\ \vdots \\ X_{iT} - \overline{X}_i \end{pmatrix} \beta + \begin{pmatrix} u_{i1} - \overline{u}_i \\ \vdots \\ u_{iT} - \overline{u}_i \end{pmatrix}, \quad i = 1, \cdots, n.$$

- Note that

$$\begin{pmatrix} y_{i1} - \overline{y}_i \\ \vdots \\ y_{iT} - \overline{y}_i \end{pmatrix} = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix} - \begin{pmatrix} \overline{y}_i \\ \vdots \\ \overline{y}_i \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix} - \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \overline{y}_i$$

$$= I_T y_i - 1_T \overline{y}_i$$

$$= I_T y_i - 1_T \frac{1}{T} 1_T' y_i$$

$$= I_T y_i - \frac{1}{T} 1_T 1_T' y_i$$

$$= \left(I_T - \frac{1}{T} 1_T 1_T' \right) y_i, \quad i = 1, \cdots, n,$$

where I_T is an identity matrix that is $T \times T$.

- Using this notation, we have

$$\left(I_T - \frac{1}{T}1_T1_T'\right)y_i = \left(I_T - \frac{1}{T}1_T1_T'\right)X_i\beta + \left(I_T - \frac{1}{T}1_T1_T'\right)u_i, \quad i = 1, \cdots, n,$$

which is expressed as

$$D_Ty_i = D_TX_i\beta + D_Tu_i, \quad i = 1, \cdots, n,$$

where $D_T = \left(I_T - \frac{1}{T}1_T1_T'\right)$ which is $T \times T$.

- Note that $D_T D_T' = D_T$, i.e., D_T is a symmetric and idempotent matrix.

$$D_T = I_T - \frac{1}{T} 1_T 1_T' = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{T} & \cdots & \frac{1}{T} \\ \vdots & \ddots & \vdots \\ \frac{1}{T} & \cdots & \frac{1}{T} \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{T} & \cdots & -\frac{1}{T} \\ \vdots & \ddots & \vdots \\ -\frac{1}{T} & \cdots & 1 - \frac{1}{T} \end{pmatrix}$$

- We have confirmed that D_T is symmetric.
- For D_T to be idempotent, it must hold that the ij element of $D_T D_T'$ is equal to the ij element of D_T for any i and j.

- In case of $i = j$ (diagonal element):

$$\left(D_T D_T'\right)_{ii} = \left(1 - \frac{1}{T}\right)^2 + (T-1)\frac{1}{T^2} = 1 - \frac{1}{T},$$

which is equal to the ii element of D_T.

- In case of $i \neq j$ (off-diagonal element):

$$\left(D_T D_T'\right)_{ij} = 2\left(1 - \frac{1}{T}\right) \cdot \left(-\frac{1}{T}\right) + (T-2)\frac{1}{T^2} = -\frac{1}{T},$$

which is equal to the ij element of D_T.

- Hence, we have confirmed that D_T is idempotent.

- Using the matrix form for $i = 1, \cdots, n$, we have:

$$\begin{pmatrix} D_T y_1 \\ \vdots \\ D_T y_n \end{pmatrix} = \begin{pmatrix} D_T X_1 \\ \vdots \\ D_T X_n \end{pmatrix} \beta + \begin{pmatrix} D_T u_1 \\ \vdots \\ D_T u_n \end{pmatrix}.$$

- Using the Kronecker product, we obtain the following stacked model:

$$(I_n \otimes D_T)y = (I_n \otimes D_T)X\beta + (I_n \otimes D_T)u,$$

where $(I_n \otimes D_T)$ is $nT \times nT$, $y = (y_1, \cdots, y_n)'$ is $nT \times 1$, $X = (X_1, \cdots, X_n)'$ is $nT \times k$ and $u = (u_1, \cdots, u_n)'$ is $nT \times 1$.

- Note that

$$\begin{pmatrix} D_T y_1 \\ \vdots \\ D_T y_n \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 1-\frac{1}{T} & \cdots & -\frac{1}{T} \\ \vdots & \ddots & \vdots \\ -\frac{1}{T} & \cdots & 1-\frac{1}{T} \end{pmatrix} y_1 \\ \vdots \\ \begin{pmatrix} 1-\frac{1}{T} & \cdots & -\frac{1}{T} \\ \vdots & \ddots & \vdots \\ -\frac{1}{T} & \cdots & 1-\frac{1}{T} \end{pmatrix} y_n \end{pmatrix}$$

$$
\begin{pmatrix} D_T y_1 \\ \vdots \\ D_T y_n \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 1-\frac{1}{T} & \cdots & -\frac{1}{T} \\ \vdots & \ddots & \vdots \\ -\frac{1}{T} & \cdots & 1-\frac{1}{T} \end{pmatrix} & \cdots & \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \\ \vdots & \ddots & \vdots \\ \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} & \cdots & \begin{pmatrix} 1-\frac{1}{T} & \cdots & -\frac{1}{T} \\ \vdots & \ddots & \vdots \\ -\frac{1}{T} & \cdots & 1-\frac{1}{T} \end{pmatrix} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} y_{11} \\ \vdots \\ y_{1T} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} y_{n1} \\ \vdots \\ y_{nT} \end{pmatrix} \end{pmatrix}
$$

$$
= (I_n \otimes D_T) y.
$$

Kronecker Product

- For matrices A: $n \times n$ and B: $m \times m$

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

$$(A \otimes B)' = A' \otimes B',$$

$$|A \otimes B| = |A|^m |B|^n,$$

$$tr(A \otimes B) = tr(A)tr(B).$$

- For matrices A, B, C and D such that the products are defined,

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

- Recall that we have the following:

$$(I_n \otimes D_T)y = (I_n \otimes D_T)X\beta + (I_n \otimes D_T)u,$$

- Applying OLS to the above regression model,

$$\begin{aligned}
\hat{\beta} &= \left[((I_n \otimes D_T)X)'((I_n \otimes D_T)X)\right]^{-1}((I_n \otimes D_T)X)'((I_n \otimes D_T)y) \\
&= \left[X'(I_n \otimes D_T)'(I_n \otimes D_T)X\right]^{-1}X'(I_n \otimes D_T)'(I_n \otimes D_T)y \\
&= \left[X'(I_n \otimes D_T'D_T)X\right]^{-1}X'(I_n \otimes D_T'D_T)y \\
&= \left[X'(I_n \otimes D_T)X\right]^{-1}X'(I_n \otimes D_T)y.
\end{aligned}$$

- Note that the inverse matrix of D_T is not available, because the rank of D_T is $T-1$ (not full rank).
- The rank of a symmetric and idempotent matrix is equal to its trace.

$$
\begin{aligned}
tr(D_T) &= \left(1 - \frac{1}{T}\right) \times T \\
&= \frac{T-1}{T} \times T \\
&= T - 1.
\end{aligned}
$$

- The fixed effect v_i is estimated as:

$$\hat{v}_i = \overline{y}_i - \overline{X}_i\hat{\beta}.$$

- Possibly, we can estimate the following regression:

$$\hat{v}_i = Z_i\alpha + \epsilon_i,$$

- The estimator of $\hat{\sigma}_u^2$ is given by

$$\hat{\sigma}_u^2 = \frac{1}{nT - k - n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - X_{it}\hat{\beta} - \hat{v}_i\right)^2.$$

- Consider the following:

$$y_{it} = X_{it}\beta + v_i + u_{it}, \quad i = 1, \cdots, n, \quad t = 1, \cdots, T.$$

- The assumptions on the error terms v_i and u_{it} are:

$$\mathbb{E}[v_i|X] = 0, \ \forall i, \qquad \mathbb{E}[u_{it}|X] = 0, \forall i, t,$$

$$V(v_i|X) = \sigma_v^2, \ \forall i, \qquad V(u_{it}|X) = \sigma_u^2, \ \forall i, t,$$

$$Cov(v_i, v_j|X) = 0, \ \forall i \neq j, \qquad Cov(u_{it}, u_{js}|X) = 0, \ \forall i \neq j, t \neq s,$$

$$Cov(v_i, u_{jt}|X) = 0, \ \forall i, j, t,$$

$$v_i|X \sim \mathcal{N}_{\mathbb{R}}(0, \sigma_v^2), \qquad u_{it}|X \sim \mathcal{N}_{\mathbb{R}}(0, \sigma_u^2).$$

- In a matrix form w.r.t. $t = 1, \cdots, T$, we have

$$y_i = X_i\beta + v_i 1_T + u_i, \quad i = 1, \cdots, n,$$

where

$$y_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix}, \; X_i = \begin{pmatrix} X_{i1} \\ \vdots \\ X_{iT} \end{pmatrix}, \; u_i = \begin{pmatrix} u_{i1} \\ \vdots \\ u_{iT} \end{pmatrix}.$$

- The distribution of the vectors of error terms are

$$v_i 1_T | X \sim \mathcal{N}_{\mathbb{R}^T}\left(0, \sigma_v^2 1_T 1_T'\right),$$
$$u_i | X \sim \mathcal{N}_{\mathbb{R}^T}\left(0, \sigma_u^2 I_T\right).$$

- Then, by the reproductive property of normal distribution,

$$v_i 1_T + u_i | X \sim \mathcal{N}_{\mathbb{R}^T}\left(0, \sigma_v^2 1_T 1_T' + \sigma_u^2 I_T\right).$$

- Again, in a matrix form w.r.t. $i = 1, \cdots, n$, we have

$$y = X\beta + v + u,$$

where

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \; X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \; v = \begin{pmatrix} v_1 1_T \\ \vdots \\ v_n 1_T \end{pmatrix}, \; u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

- The conditional distribution of $v + u|X$ is

$$v + u|X \sim \mathcal{N}_{\mathbb{R}^{nT}}\Big(0, I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big).$$

- The likelihood function is

$$\begin{aligned} L(\beta, \sigma_v^2, \sigma_u^2) =& (2\pi)^{-\frac{nT}{2}} \Big| I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T) \Big|^{-\frac{1}{2}} \\ & \times \exp\left\{ -\frac{1}{2}(y - X\beta)'\Big(I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big)^{-1}(y - X\beta)\right\}. \end{aligned}$$

- The log-likelihood function is

$$\log L(\beta, \sigma_v^2, \sigma_u^2) = -\frac{nT}{2}\log(2\pi) - \frac{1}{2}\log\left|I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\right| - \frac{1}{2}(y - X\beta)'\Big(I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big)^{-1}(y - X\beta).$$

- The MLE of β, denoted by $\tilde{\beta}$, is

$$\tilde{\beta} = \left[X'\Big(I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big)^{-1} X\right]^{-1} X'\Big(I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big)^{-1} y = \left[\sum_{i=1}^{n} X_i'(\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)^{-1} X_i\right]^{-1}\left[\sum_{i=1}^{n} X_i'(\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)^{-1} y_i\right],$$

  which is equivalent to GLS.

- Note that $\tilde{\beta}$ is not operational since $\tilde{\beta}$ depends on unknown variables σ_v^2 and σ_u^2.

OLS and GLS

- The estimator of the fixed effect model is considered to be OLS estimator while that of the random effect model is considered to be GLS estimator.
- Why?
- The difference comes from the variance of the error term.
  - In case of the fixed effect model, the variance of the error term depends on σ_u^2 because we remove the individual effect v_i.
  - In case of the random effect model, the variance of the error term depends on σ_v^2 and σ_u^2, which is the source of the **auto correlation**.
- The issue of auto correlation must be dealt with by using GLS to obtain the efficient estimator.

Properties of Fixed/Random Effect Model

- We have derived the estimator of fixed and random effect models.
- The next task is to check the properties of estimators:
  - unbiasedness
  - efficiency
  - consistency
  - asymptotic normality

- Using `nlswork.dta`, we will consider the model that the dependent variable is log wage and the explanatory variables are:
  - completed years of schooling (`grade`)
  - current age, its squared
  - current yeas worked (`exp`), its squared
  - current years of tenure on the current job, its squared
  - whether black (`race`)
  - whether residing in standard metropolitan statistical area (`SMSA`)
  - whether residing in the South.
- Wages are considered to be associated with the individual effect which is correlated with the explanatory variables, so removing it may be important to obtain the estimates that is not biased.

- In using Stata, type `xtreg y x1 x2, fe` to use a fixed effect model.

```
. xtreg ln_w grade age c.age#c.age ttl_exp c.ttl_exp#c.ttl_exp tenure
> c.tenure#c.tenure 2.race not_smsa south, fe
note: grade omitted because of collinearity.
note: 2.race omitted because of collinearity.
```

```
Fixed-effects (within) regression               Number of obs     =     28,091
Group variable: idcode                          Number of groups  =      4,697

R-squared:                                      Obs per group:
     Within  = 0.1727                                         min =          1
     Between = 0.3505                                         avg =        6.0
     Overall = 0.2625                                         max =         15

                                                F(8,23386)        =     610.12
corr(u_i, Xb) = 0.1936                          Prob > F          =     0.0000
```

ln_wage	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
grade	0	(omitted)				
age	.0359987	.0033864	10.63	0.000	.0293611	.0426362
c.age#c.age	-.000723	.0000533	-13.58	0.000	-.0008274	-.0006186
ttl_exp	.0334668	.0029653	11.29	0.000	.0276545	.039279
c.ttl_exp#c.ttl_exp	.0002163	.0001277	1.69	0.090	-.0000341	.0004666
tenure	.0357539	.0018487	19.34	0.000	.0321303	.0393775
c.tenure#c.tenure	-.0019701	.000125	-15.76	0.000	-.0022151	-.0017251
race						
Black	0	(omitted)				
not_smsa	-.0890108	.0095316	-9.34	0.000	-.1076933	-.0703282
south	-.0606309	.0109319	-5.55	0.000	-.0820582	-.0392036
_cons	1.03732	.0485546	21.36	0.000	.9421496	1.13249
sigma_u	.35562203					
sigma_e	.29068923					
rho	.59946283	(fraction of variance due to u_i)				

F test that all u_i=0: F(4696, 23386) = 6.65   Prob > F = 0.0000

Reference

- Stata.com "**xtreg** - Fixed-, between-, and random-effects and population-averaged linear models."