3.2.2 Random Effect Model (ランダム効果モデル)

Model:

$$y_{it} = X_{it}\beta + v_i + u_{it}, \qquad i = 1, 2, \cdots, n, \quad t = 1, 2, \cdots, T$$

where i indicates individual and t denotes time.

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

$$\mathrm{E}(v_i|X) = \mathrm{E}(u_{it}|X) = 0 \ \text{ for all } i,$$

$$\mathrm{V}(v_i|X) = \sigma_v^2 \ \text{ for all } i, \qquad \mathrm{V}(u_{it}|X) = \sigma_u^2 \ \text{ for all } i \text{ and } t,$$

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

$$\mathrm{Cov}(v_i, u_{jt}|X) = 0 \ \text{ for all } i,\ j \text{ and } t.$$

Note that X includes X_{it} for $i = 1, 2, \cdots, n$ and $t = 1, 2, \cdots, T$.

In a matrix form with respect to $t = 1, 2, \cdots, T$, we have the following:

$$y_i = X_i\beta + v_i 1_T + u_i, \qquad i = 1, 2, \cdots, n,$$

where $y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix}$, $X_i = \begin{pmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{iT} \end{pmatrix}$ and $u_i = \begin{pmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{iT} \end{pmatrix}$ are $T \times 1$, $T \times k$ and $T \times 1$, respectively.

$$u_i \sim N(0, \sigma_u^2 I_T) \;\text{ and }\; v_i 1_T \sim N(0, \sigma_v^2) \;\;\Longrightarrow\;\; v_i 1_T + u_i \;\sim\; N(0, \sigma_v^2 1_T 1_T' + \sigma_u^2 I_T).$$

Again, in a matrix form with respect to i, we have the following:

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

where $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$, $v = \begin{pmatrix} v_1 1_T \\ v_2 1_T \\ \vdots \\ v_n 1_T \end{pmatrix}$ and $u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ are $nT \times 1$, $nT \times k$, $nT \times 1$ and

$nT \times 1$, respectively.

The distribution of $u + v$ is given by:

$$v + u \sim N\Big(0, I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big)$$

The likelihood function is given by:

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

Remember that $f(x) = (2\pi)^{-k/2}|\Sigma|^{-1/2} \exp\Big(-\frac{1}{2}(x - \mu)'\Sigma^{-1}(x - \mu)\Big)$ when $X \sim N(\mu, \Sigma)$, where X denotes a k-variate random variable.

The estimators of β, σ_v^2 and σ_u^2 are given by maximizing the following log-likelihood function:

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

MLE of β, denoted by $\tilde{\beta}$, is given by:

$$\begin{aligned}\tilde{\beta} &= \Big(X'\Big(I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big)^{-1} X\Big)^{-1} X'\Big(I_n \otimes (\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)\Big)^{-1} y \\ &= \Big(\sum_{i=1}^{n} X_i'(\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)^{-1} X_i\Big)^{-1}\Big(\sum_{i=1}^{n} X_i'(\sigma_v^2 1_T 1_T' + \sigma_u^2 I_T)^{-1} y_i\Big),\end{aligned}$$

which is equivalent to GLS.

Note that $\tilde{\beta}$ is not operational, because $\hat{\beta}$ depends on σ_v^2 and σ_u^2.

3.3 Hausman's Specification Error (特定化誤差) Test

Regression model:

$$y = X\beta + u, \qquad y:\ n\times 1, \quad X:\ n\times k, \quad \beta:\ k\times 1, \quad u:\ n\times 1.$$

Suppose that X is stochastic.

If $\mathrm{E}(u|X) = 0$, OLSE $\hat\beta$ is unbiased because of $\hat\beta = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'u$ and $\mathrm{E}((X'X)^{-1}X'u) = 0$.

However, If $\mathrm{E}(u|X) \neq 0$, OLSE $\hat\beta$ is biased and inconsistent.

Therefore, we need to check if X is correlated with u or not.

$\Longrightarrow$ **Hausman's Specification Error Test**

The null and alternative hypotheses are:

- H_0 : X and u are independent, i.e., $\text{Cov}(X, u) = 0$,

- H_1 : X and u are not independent.

Suppose that we have two estimators $\hat{\beta}_0$ and $\hat{\beta}_1$, which have the following properties:

- $\hat{\beta}_0$ is consistent and efficient under H_0, but is not consistent under H_1,

- $\hat{\beta}_1$ is consistent under both H_0 and H_1, but is not efficient under H_0.

Under the conditions above, we have the following test statistic:

$$(\hat{\beta}_1 - \hat{\beta}_0)'\Big(\text{V}(\hat{\beta}_1) - \text{V}(\hat{\beta}_0)\Big)^{-1}(\hat{\beta}_1 - \hat{\beta}_0) \longrightarrow \chi^2(k).$$

Example: $\hat{\beta}_0$ is OLS, while $\hat{\beta}_1$ is IV such as 2SLS.

Hausman, J.A. (1978) "Specification Tests in Econometrics," *Econometrica*, Vol.46, No.6, pp.1251–1271.

3.4 Choice of Fixed Effect Model or Random Effect Model

3.4.1 The Case where X is Correlated with u — Review

The standard regression model is given by:

$$y = X\beta + u, \qquad u \sim N(0, \sigma^2 I_n)$$

OLS is:

$$\hat{\beta} = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'u.$$

If X is not correlated with u, i.e., $\mathrm{E}(X'u) = 0$, we have the result: $\mathrm{E}(\hat{\beta}) = \beta$.

However, if X is correlated with u, i.e., $\mathrm{E}(X'u) \neq 0$, we have the result: $\mathrm{E}(\hat{\beta}) \neq \beta$.
$\Longrightarrow$ $\hat{\beta}$ is biased.

Assume that in the limit we have the followings:

$$\begin{aligned}(\frac{1}{n}X'X)^{-1} &\longrightarrow M_{xx}^{-1}, \\ \frac{1}{n}X'u &\longrightarrow M_{xu} \neq 0 \text{ when } X \text{ is correlated with } u.\end{aligned}$$

Therefore, even in the limit,

$$\text{plim}\,\hat{\beta} = \beta + M_{xx}^{-1}M_{xu} \neq \beta,$$

which implies that $\hat{\beta}$ is not a consistent estimator of β.

Thus, in the case where X is correlated with u, OLSE $\hat{\beta}$ is neither unbiased nor consistent.

3.4.2 Fixed Effect Model or Random Effect Model

Usually, in the random effect model, we can consider that v_i is correlated with X_{it}.

[Reason:]

v_i includes the unobserved variables in the *i*th individual, i.e., ability, intelligence, and so on.

X_{it} represents the observed variables in the *i*th individual, i.e., income, assets, and so on.

The unobserved variables v_i are related to the observed variables X_{it}.

Therefore, we consider that v_i is correlated with X_{it}.

Thus, in the case of the random effect model, usually we cannot use OLS or GLS.

In order to use the random effect model, we need to test whether v_i is uncorrelated with X_{it}.

Apply Hausman's test.

- H_0 : X_{it} and e_{it} are independent ($\longrightarrow$ Use the random effect model),

- H_1 : X_{it} and e_{it} are not independent ($\longrightarrow$ Use the fixed effect model),

where $e_{it} = v_i + u_{it}$.

Note that:

- We can use the random effect model under H_0, but not under H_1.
- We can use the fixed effect model under both H_0 and H_1.
- The random effect model is more efficient than the fixed effect model under H_0.

Therefore, under H_0 we should use the random effect model, rather than the fixed effect model.

3.5 Applications

Example of Panel Data in Section 3: Production Function of Prefectures from 2001 to 2010.

pref: 都道府県（通し番号 1～47）

year: 年度（2001～2010 年）

y : 県内総生産（支出側、実質：固定基準年方式），出所：県民経済計算（平成 13 年度 - 平成 24 年度）（93SNA，平成 17 年基準計数）

k : 都道府県別民間資本ストック（平成 12 暦年価格，年度末，国民経済計算ベース　平成 23 年 3 月時点）一期前（2000～2009 年）

l : 県内就業者数，出所：県民経済計算（平成 13 年度 - 平成 24 年度）（93SNA，平成 17 年基準計数）

```
. tsset pref year
```

```
      panel variable:  pref (strongly balanced)
       time variable:  year, 2001 to 2010
               delta:  1 unit

. gen ly=log(y)

. gen lk=log(k)

. gen ll=log(l)

. reg ly lk ll

      Source |       SS       df       MS              Number of obs =     470
-------------+------------------------------           F(  2,   467) =19374.95
       Model |  316.479302     2  158.239651           Prob > F      =  0.0000
    Residual |  3.81409572   467  .008167229           R-squared     =  0.9881
-------------+------------------------------           Adj R-squared =  0.9880
       Total |  320.293398   469  .682928354           Root MSE      =  .09037

------------------------------------------------------------------------------
          ly |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          lk |   .0941587   .0081273    11.59   0.000     .0781881    .1101294
          ll |   .9976399   .0102641    97.20   0.000     .9774703    1.017809
       _cons |   .5970719   .0773137     7.72   0.000     .4451461    .7489978
------------------------------------------------------------------------------

. xtreg ly lk ll,fe
```

```
Fixed-effects (within) regression               Number of obs      =       470
Group variable: pref                            Number of groups   =        47

R-sq:  within  = 0.1721                         Obs per group: min =        10
       between = 0.9456                                        avg =      10.0
       overall = 0.9439                                        max =        10

                                                F(2,421)           =     43.77
corr(u_i, Xb)  = 0.8803                         Prob > F           =    0.0000

------------------------------------------------------------------------------
          ly |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          lk |   .2329208   .0252321     9.23   0.000     .1833242    .2825175
          ll |   .3268537   .0810662     4.03   0.000     .1675088    .4861987
       _cons |   7.691145   1.376677     5.59   0.000     4.985128    10.39716
-------------+----------------------------------------------------------------
     sigma_u |  .41045507
     sigma_e |  .03561437
         rho |  .99252757   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(46, 421) =    56.22             Prob > F = 0.0000

. est store fixed

. xtreg ly lk ll,re

Random-effects GLS regression                   Number of obs      =       470
Group variable: pref                            Number of groups   =        47
```

```
R-sq:  within  = 0.1058                         Obs per group: min =        10
       between = 0.9805                                        avg =      10.0
       overall = 0.9787                                        max =        10

                                                Wald chi2(2)       =   3875.75
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
          ly |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          lk |   .2457767   .0153094    16.05   0.000     .2157708    .2757827
          ll |   .8105099   .0220256    36.80   0.000     .7673406    .8536793
       _cons |   .8332015   .2411141     3.46   0.001     .3606265    1.305776
-------------+----------------------------------------------------------------
     sigma_u |    .081609
     sigma_e |  .03561437
         rho |   .8400205   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. hausman fixed

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed          .          Difference          S.E.
-------------+----------------------------------------------------------------
          lk |    .2329208     .2457767       -.0128559         .020057
          ll |    .3268537     .8105099       -.4836562        .0780167
------------------------------------------------------------------------------
```

```
                   b = consistent under Ho and Ha; obtained from xtreg
    B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test:  Ho:  difference in coefficients not systematic

              chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                      =      144.66
            Prob>chi2 =      0.0000
```

4 Introduction to Causal Inference (因果推論) — Micro-econometrics —

● Notations:

- i : ith individual,
- Y_i^1: Output of the ith individual for State 1,
- Y_i^0: Output of the ith individual for State 0,
- X_i: Explanatory variables of the ith indvidual (i.e., individual information such as sex, age, region, occupation, income, and etc.),
- D_i: Dummy variable, i.e., $D_i = 1$ for State 1 and $D_i = 0$ for State 0,

for $i = 1, 2, \cdots, n$.

● **Example:**

- What is cause of wage gap? Is undergraduate school graduation a big factor?
- Is the policy effective?

. . .

. . .

. . .

● **Special Words in this Field:**

Treatment group or **Experimental group (処置群，実験群)** $\longrightarrow$ Output 1

Control group (対照群， 比較群) $\longrightarrow$ Output 0

● **Problem:**

We want to know $Y_i^1 - Y_i^0$ for all i, which is called the **treatment effect (処置効果)**.

However, we can observe either Y_i^1 or Y_i^0.

Therefore, we cannot obtain $Y_i^1 - Y_i^0$.

Y_i^1 and Y_i^0 are assumed to be identically distributed.

Instead, consider estimating $\mathrm{E}(Y_i^1 - Y_i^0) = \mathrm{E}(Y^1 - Y^0)$, which is called the **average treatment effect (ATE，平均处置効果)**.

● **Note 1:**

Using Y_i^1, Y_i^0 and D_i, the output of the ith individual is given by:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0.$$

● **Note 2:**

D_i is a discrete random variable.

D_i takes 1 or 0.

The probability of $D_i = 1$ given X_i is denoted by $\mathrm{P}(D_i = 1|X_i)$, and that of $D_i = 0$ given X_i is represented by $\mathrm{P}(D_i = 0|X_i)$.

The conditional expectation of D_i given X_i is:

$$\mathrm{E}(D_i|X_i) = 1 \times \mathrm{P}(D_i = 1|X_i) + 0 \times \mathrm{P}(D_i = 0|X_i) = \mathrm{P}(D_i = 1|X_i),$$

which is derived from the definition of expectation in the case of the discrete random variable.

Therefore, $\mathrm{E}(D_i|X_i)$ is between 0 and 1.

Consider that $\mathrm{E}(D_i|X_i) = \mathrm{P}(D_i = 1|X_i)$ is related to the distribution function $\pi(X_i)$, i.e., $\mathrm{P}(D_i = 1|X_i) = \pi(X_i)$.

$\pi(X_i)$ is called the **propensity score (傾向スコア)**.

● **Example of $\pi(\cdot)$:**

$F(x)$ is denoted by a distribution function, where $F(-\infty) = 0$ and $F(\infty) = 1$.

The relationship between $\pi(X_i)$ and $F(X_i\beta)$ is as follows:

$$\pi(X_i) = F(X_i\beta),$$

where β denotes the parameter to be estimated, and $F(x)$ often takes one of the following two distribution functions:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2)\,\mathrm{d}t \qquad \longrightarrow \qquad \text{Standard normal distribution (\textbf{Probit model})}$$

or

$$F(x) = \frac{1}{1 + \exp(-x)} \qquad \longrightarrow \qquad \text{Logistic distribution (\textbf{Logit model})}.$$