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.

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

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:

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

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:

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

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

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

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., $\mathrm{Cov}(X, u) = 0$,
- 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:

$$(\hat\beta_1 - \hat\beta_0)'\Big(\mathrm{V}(\hat\beta_1) - \mathrm{V}(\hat\beta_0)\Big)^{-1}(\hat\beta_1 - \hat\beta_0) \longrightarrow \chi^2(k).$$

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,

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

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 ith individual, i.e., ability, intelligence, and so on.

X_{it} represents the observed variables in the ith 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.