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,

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

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.

4 Generalized Method of Moments (GMM, 一般化積率法)

4.1 Method of Moments (MM, 積率法)

As $n \longrightarrow \infty$, we have the result: $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mathrm{E}(X) = \mu$.

$\Longrightarrow$ **Law of Large Number (大数の法則)**

$X_1, X_2, \cdots, X_n$ are n realizations of X.

[Review] Chebyshev's inequality (チェビシェフの不等式) is given by:

$$P(|X-\mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \quad \text{or} \quad P(|X-\mu| \leq \epsilon) \geq 1 - \frac{\sigma^2}{\epsilon^2},$$

where $\mu = \mathrm{E}(X)$, $\sigma^2 = \mathrm{V}(X)$ and any $\epsilon > 0$.

Note that $P(|X-\mu| > \epsilon) + P(|X-\mu| \leq \epsilon) = 1$.

Replace X, $\mathrm{E}(X)$ and $\mathrm{V}(X)$ by $\overline{X}$, $\mathrm{E}(\overline{X}) = \mu$ and $\mathrm{V}(\overline{X}) = \dfrac{\sigma^2}{n}$.

As $n \longrightarrow \infty$,

$$P(|\overline{X} - \mu| \le \epsilon) \ge 1 - \frac{\sigma^2}{n\epsilon^2} \longrightarrow 1.$$

That is, $\overline{X} \longrightarrow \mu$ as $n \longrightarrow \infty$.

[End of Review]

$\overline{X}$ is an approximation of $\mathrm{E}(X) = \mu$.

Therefore, $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ is taken as an estimator of μ.

$\Longrightarrow$ $\overline{X}$ is MM estimator of $\mathrm{E}(X) = \mu$.

MM is applied to the regression model as follows:

Regression model: $y_i = x_i\beta + u_i$, where x_i and u_i are assumed to be stochastic.

Familiar Assumption: $\mathrm{E}(x'u) = 0$, called the **orthogonality condition (直交条件)**, where x is a $1 \times k$ vector and u is a scalar.

We consider that $(x_1, x_2, \cdots, x_n)$ and $(u_1, u_2, \cdots, u_n)$ are realizations generated from random variables x and u, respectively.

From the law of large number, we have the following:

$$\frac{1}{n}\sum_{i=1}^{n} x_i' u_i = \frac{1}{n}\sum_{i=1}^{n} x_i'(y_i - x_i\beta) \longrightarrow \mathrm{E}(x'u) = 0.$$

Thus, the MM estimator of β, denoted by β_{MM}, satisfies:

$$\frac{1}{n}\sum_{i=1}^{n} x_i'(y_i - x_i\beta_{MM}) = 0.$$

Therefore, β_{MM} is given by:

$$\beta_{MM} = \Big(\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\Big)^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n} x_i' y_i\Big) = (X'X)^{-1}X'y,$$

which is equivalent to OLS and MLE.

Note that X and y are:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$