Advanced Econometrics

Lecture 10: Instrumental Variables (Hansen Chapter 11)

Instructor: Ma, Jun

Renmin University of China

December 18, 2018

Introduction

① 先找谁是谁的因.

观测的到 $X_i \longrightarrow Y$

观测不到 e_i

② 构建数据生成模型

$Y_i = X_i'\beta + e_i$

③ 三种情况假设:

1) $E(e_i \mid X_i) = 0$

2) $E(e_i X_i) = 0 \Rightarrow X_i'\beta = P(Y_i \mid X_i)$

3) $E(e_i X_i) \neq 0$

- Endogeneity in the linear model:

$$Y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + e_i$$
$$\mathbb{E}\left(\boldsymbol{X}_i e_i\right) \neq \mathbf{0}.$$

$\Leftarrow$ 该假设发生了变化. $\boxed{\text{内生性假设}}$

误差项与解释变量是相关的.

- Note that the above model is not the linear projection model, since otherwise, if $\boldsymbol{\beta}^* = \mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i\right)^{-1}\mathbb{E}\left(\boldsymbol{X}_i Y_i\right)$, and the linear projection model is

$$Y_i = \boldsymbol{X}_i'\boldsymbol{\beta}^* + e_i^*$$
$$\mathbb{E}\left(\boldsymbol{X}_i e_i^*\right) = \mathbf{0}.$$

Introduction

- Under endogeneity, the projection coefficients $\boldsymbol{\beta}^*$ does not equal the structural parameter $\boldsymbol{\beta}$:

$$
\begin{aligned}
\boldsymbol{\beta}^* &= \left(\mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i'\right)\right)^{-1}\mathbb{E}\left(\boldsymbol{X}_iY_i\right) \\
&= \left(\mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i'\right)\right)^{-1}\mathbb{E}\left(\boldsymbol{X}_i\left(\boldsymbol{X}_i'\boldsymbol{\beta}+e_i\right)\right) \\
&= \boldsymbol{\beta} + \left(\mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i'\right)\right)^{-1}\underline{\mathbb{E}\left(\boldsymbol{X}_ie_i\right)} \neq 0 \\
&\neq \boldsymbol{\beta}.
\end{aligned}
$$

- Endogeneity implies that the LS estimator is inconsistent for the structural parameter $\boldsymbol{\beta}$. The LS estimator is consistent for the projection coefficient $\boldsymbol{\beta}^*$:

$$
\widehat{\boldsymbol{\beta}} \xrightarrow{p} \left(\mathbb{E}\left(\boldsymbol{X}_i\boldsymbol{X}_i'\right)\right)^{-1}\mathbb{E}\left(\boldsymbol{X}_iY_i\right) = \boldsymbol{\beta}^* \neq \boldsymbol{\beta}.
$$

$\hat{\beta}$是 β^*的一致估计量

不是 β的一致估计量.

Examples: Measurement Error

- Suppose in the true model that generates the dependent variable Y:

$$Y_i = \boldsymbol{Z}_i'\boldsymbol{\beta} + e_i$$

  $\boldsymbol{Z}_i$ is not observed. Instead we observe $\boldsymbol{X}_i = \boldsymbol{Z}_i + \boldsymbol{U}_i$, with measurement error $\boldsymbol{U}_i$.
- $\boldsymbol{Z}_i$ and $\boldsymbol{U}_i$ are independent and $\mathbb{E}\boldsymbol{U}_i = \boldsymbol{0}$.
- The model becomes

$$\begin{aligned} Y_i &= \boldsymbol{Z}_i'\boldsymbol{\beta} + e_i \\ &= (\boldsymbol{X}_i - \boldsymbol{u}_i)'\,\boldsymbol{\beta} + e_i \\ &= \boldsymbol{X}_i'\boldsymbol{\beta} + v_i. \end{aligned}$$

$v_i = e_i - u_i'\beta$

Examples: Measurement Error

- In this "fitted" model

$$Y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + v_i,$$

the error v_i is not the one in a projection model since

X_i和 v_i 有内生性.

$$\mathbb{E}\left(\boldsymbol{X}_i v_i\right) = \mathbb{E}\left[\left(\boldsymbol{Z}_i + \boldsymbol{u}_i\right)\left(e_i - \boldsymbol{u}_i'\boldsymbol{\beta}\right)\right] = -\mathbb{E}\left(\boldsymbol{u}_i\boldsymbol{u}_i'\right)\boldsymbol{\beta} \neq \mathbf{0}.$$

$v_i = e_i - u_i'\beta$

- When $k = 1$, we find

$$\boldsymbol{\beta}^* = \boldsymbol{\beta} + \frac{\mathbb{E}\left(X_i v_i\right)}{\mathbb{E}\left(X_i^2\right)} = \boldsymbol{\beta}\left(1 - \frac{\mathbb{E}\left(U_i^2\right)}{\mathbb{E}\left(X_i^2\right)}\right).$$

$$\beta^* = \frac{E(X_i Y_i)}{E(X_i^2)} = \beta + \frac{E(X_i v_i)}{E(X_i^2)}$$

$$= \beta\left(1 - \frac{E(u_i^2)}{E(X_i^2)}\right)$$

The projection coefficient always shrinks towards zero.

$|\beta^*| < |\beta|$ 当存在内生性时，低估了真实系数.

Examples: Supply and Demand

- The observed quantity and price q_i and p_i are determined in an equilibrium of the demand equation

$$q_i = -\beta_1 p_i + e_{1i}$$

观测到的 p_i 是均衡价格

and supply equation

$$q_i = -\beta_2 p_i + e_{2i}.$$

- Assume $\boldsymbol{e}_i = \begin{pmatrix} e_{1i} \\ e_{2i} \end{pmatrix}$ are iid, $\mathbb{E}\boldsymbol{e}_i = \mathbf{0}$ and $\mathbb{E}\boldsymbol{e}_i\boldsymbol{e}_i' = \boldsymbol{I}_2$.

Examples: Supply and Demand

- Solve for p_i and q_i:

$$\begin{bmatrix} 1 & \beta_1 \\ 1 & -\beta_2 \end{bmatrix} \begin{pmatrix} q_i \\ p_i \end{pmatrix} = \begin{pmatrix} e_{1i} \\ e_{2i} \end{pmatrix}$$

$$\begin{aligned} \begin{pmatrix} q_i \\ p_i \end{pmatrix} &= \begin{bmatrix} 1 & \beta_1 \\ 1 & -\beta_2 \end{bmatrix}^{-1} \begin{pmatrix} e_{1i} \\ e_{2i} \end{pmatrix} \\ &= \begin{bmatrix} \beta_2 & \beta_1 \\ 1 & -1 \end{bmatrix} \begin{pmatrix} e_{1i} \\ e_{2i} \end{pmatrix} \left(\frac{1}{\beta_1 + \beta_2} \right) \\ &= \begin{pmatrix} (\beta_2 e_{1i} + \beta_1 e_{2i}) \,/\, (\beta_1 + \beta_2) \\ (e_{1i} - e_{2i}) \,/\, (\beta_1 + \beta_2) \end{pmatrix}. \end{aligned}$$

Examples: Supply and Demand

$y_i = x_i'\beta + e_i$ ← 遗漏变量

有一些观测不到但对y有影响的因素进入了e_i，但该因素又与x_i相关 这是一个遗漏变量问题.

- The projection coefficient:

$$q_i = \beta^* p_i + e_i^*$$
$$\mathbb{E}\left(p_i e_i^*\right) = 0$$

$$\beta^* = \frac{\mathbb{E}\left(p_i q_i\right)}{\mathbb{E}\left(p_i^2\right)} = \frac{\beta_2 - \beta_1}{2}.$$

- The projection coefficient equals neither the demand slope β_1 nor the supply slope β_2.

$$p_i = \frac{e_{1i} - e_{2i}}{\beta_1 + \beta_2}$$

一如果用 $q_i = \beta^* p_i + e_i^*$ 估计

发现 p_i 与 e_i 一定是相关的.

Instrumental Variables

- Partition:

$$\begin{array}{c c} \boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{X}_{1i} \\ \boldsymbol{X}_{2i} \end{pmatrix} & \begin{array}{c} k_1 \\ k_2 \end{array} \end{array}$$

and

$$\begin{array}{c c} \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} & \begin{array}{c} k_1 \\ k_2 \end{array} \end{array} .$$

- So the model is:

$$\begin{aligned} Y_i &= \boldsymbol{X}_i' \boldsymbol{\beta} + e_i \\ &= \boldsymbol{X}_{1i}' \boldsymbol{\beta}_1 + \boldsymbol{X}_{2i}' \boldsymbol{\beta}_2 + e_i. \end{aligned}$$

In matrix notation:

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{e} \\ &= \boldsymbol{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{X}_2 \boldsymbol{\beta}_2 + \boldsymbol{e}. \end{aligned}$$

$Y_i = X_{1i}'\beta_1 + X_{2i}'\beta_2 + e_i$ $\quad X_i = \begin{pmatrix} X_{1i} \\ X_{2i} \end{pmatrix}$

X_{1i} 是外生的 $E(X_{1i} e_i) = 0$ k_1个

X_{2i} 是内生的 $E(X_{2i} e_i) \neq 0$ k_2个

z_i: $l \times 1$ $\quad E(z_i' e_i) = 0$

↑

工具变量

Instrumental Variables

工具变量必须满足两个条件：

$cov(X_{1i}', e_i') = 0$

$cov(X_{2i}, e_i') \neq 0$ 且相关性不能太弱

- Assume

Z与外生独立 $$\mathbb{E}\left(\boldsymbol{X}_{1i} e_i\right) = \mathbf{0}$$

与内生相关 $$\mathbb{E}\left(\boldsymbol{X}_{2i} e_i\right) \neq \mathbf{0}$$

Definition

The $l \times 1$ random vector $\boldsymbol{Z}_i$ is an **instrumental variable** if

① $$\mathbb{E}\left(\boldsymbol{Z}_i e_i\right) = \mathbf{0}$$

② $$\mathbb{E}\left(\boldsymbol{Z}_i \boldsymbol{Z}_i'\right) > 0$$

$E(Z_i X_i') \neq 0 \Leftarrow$ ③ $$\operatorname{rank}\left(\mathbb{E}\left(\underbrace{\boldsymbol{Z}_i \boldsymbol{X}_i'}_{\text{列满秩的}}\right)\right) = k$$ $(l \geq k)$

Instrumental Variables

- $\boldsymbol{X}_{1i}$ satisfies $\mathbb{E}\left(\boldsymbol{X}_{1i}e_i\right) = \mathbf{0}$. So it should be included as instrumental variables.

$$\boldsymbol{Z}_i = \begin{pmatrix} \boldsymbol{Z}_{1i} \\ \boldsymbol{Z}_{2i} \end{pmatrix} = \begin{array}{cc} \begin{pmatrix} \boldsymbol{X}_{1i} \\ \boldsymbol{Z}_{2i} \end{pmatrix} & \begin{array}{c} k_1 \\ l_2 \end{array} \end{array}$$

- We say the model is just-identified if $\ell = k$ $(\ell_2 = k_2)$ and over-identified if $\ell > k$ $(\ell_2 > k_2)$.

模型中内生变量的维数一定没有工具变量维数多.

Instrumental Variables Estimator

- The assumption that $\boldsymbol{Z}_i$ is an IV implies

$$
\begin{aligned}
\mathbb{E}\left(\boldsymbol{Z}_i e_i\right) &= \mathbf{0} \\
\mathbb{E}\left(\boldsymbol{Z}_i\left(Y_i - \boldsymbol{X}_i'\boldsymbol{\beta}\right)\right) &= \mathbf{0} \\
\mathbb{E}\left(\boldsymbol{Z}_i Y_i\right) - \mathbb{E}\left(\boldsymbol{Z}_i \boldsymbol{X}_i'\right)\boldsymbol{\beta} &= \mathbf{0}.
\end{aligned}
$$

$\Rightarrow$ 可以把 β 解出来

- If $\ell = k$, solve for $\boldsymbol{\beta}$:

$$
\boldsymbol{\beta} = \left(\mathbb{E}\left(\boldsymbol{Z}_i \boldsymbol{X}_i'\right)\right)^{-1} \mathbb{E}\left(\boldsymbol{Z}_i Y_i\right).
$$

$$\hat{\beta}_{IV} = (Z'X)^{-1} Z'Y$$

Instrumental Variables Estimator

- The IV estimator:

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{\mathrm{iv}} &= \left(\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_i \boldsymbol{X}_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_i Y_i\right) \\ &= \left(\sum_{i=1}^{n} \boldsymbol{Z}_i \boldsymbol{X}_i'\right)^{-1} \left(\sum_{i=1}^{n} \boldsymbol{Z}_i Y_i\right) \\ &= \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1} \left(\boldsymbol{Z}'\boldsymbol{Y}\right).\end{aligned}$$

- The residual satisfies:

$$\begin{aligned}\widehat{\boldsymbol{e}} &= \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\mathrm{iv}} \\ \boldsymbol{Z}'\widehat{\boldsymbol{e}} &= \boldsymbol{Z}'\boldsymbol{Y} - \boldsymbol{Z}'\boldsymbol{X}\left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{Y}\right) = \boldsymbol{0}.\end{aligned}$$

两阶段最小二乘

Two-Stage Least Squares

$l > k$

① X先对Z回归，得到 $\hat{X}$

② Y再对 $\hat{X}$ 回归

- We denote $\overset{l \times k}{\widehat{\boldsymbol{\Gamma}}} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}(\boldsymbol{Z}'\overset{n \times k}{\boldsymbol{X}})$.

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{2\text{sls}} &= \left(\widehat{\boldsymbol{\Gamma}}'\boldsymbol{Z}'\boldsymbol{Z}\widehat{\boldsymbol{\Gamma}}\right)^{-1}\left(\widehat{\boldsymbol{\Gamma}}'\boldsymbol{Z}'\boldsymbol{Y}\right) \\
&= \left(\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{X}\right)^{-1} \\
&\cdot \boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Y} \\
&= \left(\boldsymbol{X}'\underbrace{\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'}_{P_Z}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\underbrace{\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'}_{P_Z}\boldsymbol{Y}.
\end{aligned}
$$

$\hat{X} = P_Z X$

$\hat{\beta}_{2sls} = (\hat{X}'\hat{X})^{-1}\hat{X}'Y$

- When $k = \ell$, the 2SLS simplifies to IV:

$$
\begin{aligned}
\left(\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{X}\right)^{-1} &= \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\left(\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\right)^{-1}\left(\boldsymbol{X}'\boldsymbol{Z}\right)^{-1} \\
&= \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)\left(\boldsymbol{X}'\boldsymbol{Z}\right)^{-1}
\end{aligned}
$$

Two-Stage Least Squares

- So

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{\text{2sls}} &= \left(\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Y} \\
&= \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)\left(\boldsymbol{X}'\boldsymbol{Z}\right)^{-1}\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Y} \\
&= \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Y} \\
&= \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Y} \\
&= \widehat{\boldsymbol{\beta}}_{\text{iv}}.
\end{aligned}$$

- Define the projection matrix:

$$\boldsymbol{P}_{\boldsymbol{Z}} = \boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'$$

对称且幂等

- We can write

$$\widehat{\boldsymbol{\beta}}_{\text{2sls}} = \left(\boldsymbol{X}'\boldsymbol{P}_{\boldsymbol{Z}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{P}_{\boldsymbol{Z}}\boldsymbol{Y}.$$

Two-Stage Least Squares

- And the fitted values:

$$\widehat{\boldsymbol{X}} = \boldsymbol{P_Z X} = \boldsymbol{Z}\widehat{\boldsymbol{\Gamma}}$$

$\longrightarrow$ X的每一列对Z作回归，得到每个拟合值组成$\widehat{X}$.

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{\text{2sls}} &= \left(\boldsymbol{X}'\boldsymbol{P_Z P_Z X}\right)^{-1} \boldsymbol{X}'\boldsymbol{P_Z Y} \\ &= \left(\widehat{\boldsymbol{X}}'\widehat{\boldsymbol{X}}\right)^{-1} \widehat{\boldsymbol{X}}'\boldsymbol{Y}. \end{aligned}$$

- First regress $\boldsymbol{X}$ on $\boldsymbol{Z}$. Obtain the LS coefficients $\widehat{\boldsymbol{\Gamma}} = \left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{X}\right)$ and the fitted values $\widehat{\boldsymbol{X}} = \boldsymbol{P_Z X} = \boldsymbol{Z}\widehat{\boldsymbol{\Gamma}}$.
- Second regress $\boldsymbol{Y}$ on $\widehat{\boldsymbol{X}}$. Get $\widehat{\boldsymbol{\beta}}_{\text{2sls}} = \left(\widehat{\boldsymbol{X}}'\widehat{\boldsymbol{X}}\right)^{-1} \widehat{\boldsymbol{X}}'\boldsymbol{Y}$.

Two-Stage Least Squares

- Recall $\boldsymbol{X} = [\boldsymbol{X}_1\ \boldsymbol{X}_2]$ and $\boldsymbol{Z} = [\boldsymbol{X}_1\ \boldsymbol{Z}_2]$. Note

  X_1 外生 → $\widehat{\boldsymbol{X}}_1 = \boldsymbol{P_Z X}_1 = \boldsymbol{X}_1$. Then

$$\widehat{\boldsymbol{X}} = \left[\widehat{\boldsymbol{X}}_1, \widehat{\boldsymbol{X}}_2\right] = \left[\boldsymbol{X}_1, \widehat{\boldsymbol{X}}_2\right].$$

$\hat{X} = P_Z X = P_Z [X_1, X_2] = [P_Z X_1, P_Z X_2]$

$= [X_1, \hat{X}_2]$

X_1 在 Z 展开的子空间内. $P_Z X_1 = X_1$.

X_1 在 Z 上的投影就是它本身.

- The 2SLS residuals:

$$\widehat{\boldsymbol{e}} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{2\text{sls}}.$$

$\hat{e}_i = Y_i - X_{1i}' \hat{\beta}_{1,2sls} - X_{2i}' \hat{\beta}_{2,2sls}$

手动两阶段回归得到的 $\hat{e} = Y - \hat{X}\hat{\beta}_{2sls} \neq Y - X\hat{\beta}_{2sls}$

两个残差不同.

- When the model is overidentified, $\boldsymbol{Z}'\widehat{\boldsymbol{e}} \neq \boldsymbol{0}$ but

$$\begin{aligned}
\widehat{\boldsymbol{X}}'\widehat{\boldsymbol{e}} &= \widehat{\boldsymbol{\Gamma}}'\boldsymbol{Z}'\widehat{\boldsymbol{e}} \\
&= \boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\widehat{\boldsymbol{e}} \\
&= \boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{2\text{sls}} \\
&= \boldsymbol{0}.
\end{aligned}$$

reg
reg

ivregress

除系数外都不一样

手动两阶段标准差低估了 ivregress

Consistency of 2SLS

Assumption

1. The observations $(Y_i, \boldsymbol{X_i}, \boldsymbol{Z_i})$, $i = 1, \ldots, n$, *are independent and identically distributed.*

2. $\mathbb{E}\left(Y^2\right) < \infty$.

3. $\mathbb{E} \parallel \boldsymbol{X} \parallel^2 < \infty$.

4. $\mathbb{E} \parallel \boldsymbol{Z} \parallel^2 < \infty$.

5. $\mathbb{E}\left(\boldsymbol{Z}'\right)$ *is positive definite.*

6. $\mathbb{E}\left(\boldsymbol{Z}\boldsymbol{X}'\right)$ *has full rank k.*

7. $\mathbb{E}\left(\boldsymbol{Z}\boldsymbol{e}\right) = 0$.

Consistency of 2SLS

在实证研究中，$Y_i = X_i'\beta + e_i$. X中包含很多解释变量. 如果关心的解释变量是外生的，不关心其边际效应的解释变量是内生的，也不可以忽略内生性问题. 因为此时 β_1, β_2 都不是一致的.

- Proof of consistency:

③
$$\hat{\boldsymbol{\beta}}_{2sls} = \left(\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\left(\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{e}\right)$$
$$= \boldsymbol{\beta} + \left(\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Z}\left(\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}'\boldsymbol{e}.$$

- Then

$$\hat{\boldsymbol{\beta}}_{2sls} - \boldsymbol{\beta} = \left(\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}\right)\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{X}\right)\right)^{-1} \cdot \left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}\right)\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{e}\right).$$

① $\hat\beta_2 = (X_2'M_1X_2)^{-1}X_2'M_1Y$

FWL, $M_1 = I_n - X_1(X_1'X_1)^{-1}X_1'$

$\underset{n\times 1}{Y} = X_1\beta_1 + X_2\beta_2 + e$

$\hat\beta_2 = (X_2'M_1X_2)^{-1}X_2'M_1(X_1\beta_1 + X_2\beta_2 + e)$

$= \beta_2 + (\frac{1}{n}X_2'M_1X_2)^{-1}(\frac{1}{n}X_2'M_1e)$

$\frac{1}{n}X_2'M_1X_2 = \frac{1}{n}X_2'X_2 - \frac{1}{n}X_2'X_1(X_1'X_1)^{-1}X_1'X_2$

$\frac{1}{n}X_2'X_2 = \frac{1}{n}\sum_{i=1}^{n}X_{2i}X_{2i}' \xrightarrow{p} E(X_{2i}X_{2i}')$

同理有，$(\frac{1}{n}X_2'M_1X_2)^{-1}(\frac{1}{n}X_2'M_1e)$

$\xrightarrow{p} E(X_{2i}X_{1i}')E(X_{1i}X_{1i}')^{-1}E(X_{1i}X_{2i}')$

Consistency of 2SLS

$\frac{1}{n} X_2' M_1 e = \frac{1}{n} X_2' e - \frac{1}{n} X_2' X_1 (X_1'X_1)^{-1} X_1' e$

$\xrightarrow{p} E(X_{2i} e_i') = 0 \quad \xrightarrow{p} E(X_{2i} X_{1i}') E(X_{1i} X_{1i}')^{-1} \underbrace{E(X_{1i} e_i')}_{\neq 0}$

$\hat{\beta}_2 - \beta_2 \xrightarrow{p} [E(X_{2i} X_{2i}') - E(X_{2i} X_{1i}') E(X_{1i} X_{1i}')^{-1} E(X_{1i} X_{2i}')]$

$\cdot [- E(X_{2i} X_{1i}') E(X_{1i} X_{1i}')^{-1} E(X_{1i} e_i)]$

$\Rightarrow \hat{\beta}_2 \xrightarrow{p} \beta_2 + \boxed{\neq 0}$

- Then,

$$\hat{\boldsymbol{\beta}}_{2\text{sls}} - \boldsymbol{\beta} \xrightarrow{p} \left(\boldsymbol{Q}_{\boldsymbol{XZ}} \boldsymbol{Q}_{\boldsymbol{ZZ}}^{-1} \boldsymbol{Q}_{\boldsymbol{ZX}}\right)^{-1} \boldsymbol{Q}_{\boldsymbol{XZ}} \boldsymbol{Q}_{\boldsymbol{ZZ}}^{-1} \mathbb{E}\left(\boldsymbol{Z}_i e_i\right) = 0,$$

where

$$\boldsymbol{Q}_{\boldsymbol{XZ}} = \mathbb{E}\left(\boldsymbol{X}_i \boldsymbol{Z}_i'\right)$$
$$\boldsymbol{Q}_{\boldsymbol{ZZ}} = \mathbb{E}\left(\boldsymbol{Z}_i \boldsymbol{Z}_i'\right)$$
$$\boldsymbol{Q}_{\boldsymbol{ZX}} = \mathbb{E}\left(\boldsymbol{Z}_i \boldsymbol{X}_i'\right).$$

② $\hat{\beta}_{IV} = (Z'X)^{-1}(Z'Y) = \beta + (\frac{1}{n} Z'X)^{-1} \frac{1}{n} Z'e$

$= \beta + (\frac{1}{n} \Sigma Z_i X_i')^{-1} (\frac{1}{n} \Sigma Z_i e_i')$

$\xrightarrow{p} E(Z_i X_i')^{-1} \quad \xrightarrow{p} 0$

$\hat{\beta}_{IV} \xrightarrow{p} \beta$

Asymptotic Distribution of 2SLS

Assumption

1. $\mathbb{E}\left(Y^4\right) < \infty$.
2. $\mathbb{E} \parallel \boldsymbol{Z} \parallel^4 < \infty$.

- Write

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{\text{2sls}} - \boldsymbol{\beta}) = \left(\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}\right)\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{X}\right)\right)^{-1} \cdot \left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}\right)\left(\frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}\right)^{-1}\left(\frac{1}{\sqrt{n}}\boldsymbol{Z}'\boldsymbol{e}\right).$$

- CLT:

$$\frac{1}{\sqrt{n}}\boldsymbol{Z}'\boldsymbol{e} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{Z}_i e_i \xrightarrow{d} \mathrm{N}\left(\boldsymbol{0}, \boldsymbol{\Omega}\right).$$

$$\boldsymbol{\Omega} = \mathbb{E}\left(e_i^2 \boldsymbol{Z}_i \boldsymbol{Z}_i'\right).$$

$\sqrt{n}(\hat\beta_{IV}-\beta) = (\frac{1}{n}\sum Z_iX_i')^{-1}(\frac{1}{\sqrt{n}}\sum Z_ie_i)$

$\to_p E(Z_iX_i')^{-1}$

$\to_d N(0, E(e_i^2Z_iZ_i'))$

$\to_d E(Z_iX_i')^{-1}N(0, E(e_i^2Z_iZ_i'))$

$\overset{a}{\sim} N(0, E(Z_iX_i')^{-1}E(e_i^2Z_iZ_i')E(Z_iX_i')^{-1})$

$E(Z_iX_i)^{-1}E(e_i^2Z_iZ_i')E(Z_iX_i')^{-1}$

$=(\frac{1}{n}\sum Z_iX_i)^{-1}(\frac{1}{n}\sum \underbrace{e_i^2Z_iZ_i'})(\frac{1}{n}\sum Z_iX_i')^{-1}$

$(Y_i-X_i'\beta)^2Z_iZ_i'$

β观测不到，用β的一致估计$\hat\beta_{IV}$

$\to_p (Y_i-X_i'\hat\beta_{IV})^2Z_iZ_i'$

$=\hat e^2Z_iZ_i'$

Asymptotic Distribution of 2SLS

- Slutsky's theorem:

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{2\text{sls}}-\boldsymbol{\beta}) \xrightarrow{d} \left(\boldsymbol{Q_{XZ}Q_{ZZ}^{-1}Q_{ZX}}\right)^{\boldsymbol{-1}} \boldsymbol{Q_{XZ}Q_{ZZ}^{-1}}\mathrm{N}\left(\mathbf{0}, \boldsymbol{\Omega}\right) = \mathrm{N}\left(\mathbf{0}, \boldsymbol{V_{\beta}}\right).$$

- We can verify:

$$\begin{aligned}\left(\mathbb{E}\left(e^4\right)\right)^{1/4} &= \left(\mathbb{E}\left(\left(Y - \boldsymbol{X}'\boldsymbol{\beta}\right)^4\right)\right)^{1/4} \\ &\leq \left(\mathbb{E}\left(y^4\right)\right)^{1/4} + \parallel \boldsymbol{\beta} \parallel \left(\mathbb{E} \parallel \boldsymbol{X} \parallel^4\right)^{1/4} < \infty\end{aligned}$$

$$\mathbb{E} \parallel \boldsymbol{Z}e \parallel^2 \leq \left(\mathbb{E} \parallel \boldsymbol{Z} \parallel^4\right)^{1/2} \left(\mathbb{E}\left(e^4\right)\right)^{1/2} < \infty.$$

So the CLT and Slutsky's theorem do apply.

Asymptotic Distribution of 2SLS

Theorem

$$\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{2\mathrm{sls}}-\boldsymbol{\beta}\right)\xrightarrow{d}\mathrm{N}\left(\mathbf{0},\mathbf{V}_{\boldsymbol{\beta}}\right)$$

where

$$\mathbf{V}_{\boldsymbol{\beta}}=\left(\boldsymbol{Q}_{\boldsymbol{XZ}}\boldsymbol{Q}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{Q}_{\boldsymbol{ZX}}\right)^{-1}\left(\boldsymbol{Q}_{\boldsymbol{XZ}}\boldsymbol{Q}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{\Omega}\boldsymbol{Q}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{Q}_{\boldsymbol{ZX}}\right)\cdot\left(\boldsymbol{Q}_{\boldsymbol{XZ}}\boldsymbol{Q}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{Q}_{\boldsymbol{ZX}}\right)^{-1}$$

and

$$\boldsymbol{\Omega}=\mathbb{E}\left(\boldsymbol{Z}_i\boldsymbol{Z}_i'e_i^2\right).$$

- The asymptotic variance simplifies under a conditional homoskedasticity condition: $\mathbb{E}\left(e_i^2|\boldsymbol{Z}_i\right)=\sigma^2$.
- $\boldsymbol{V}_{\boldsymbol{\beta}}=\boldsymbol{V}_{\boldsymbol{\beta}}^0=\left(\boldsymbol{Q}_{\boldsymbol{XZ}}\boldsymbol{Q}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{Q}_{\boldsymbol{ZX}}\right)^{-1}\sigma^2.$

$Q_{ZZ} = \underset{L\times L}{E(Z_i Z_i')}$

$Q_{XZ} = \underset{K\times L}{E(X_i Z_i')}$

$\Omega = E(e_i^2 Z_i Z_i')$

假设 $E(e_i^2|Z_i)=\sigma^2$ 同方差

$\Rightarrow \Omega = E\,E(e_i^2 Z_i Z_i' | Z_i)$

$= \sigma^2 E(Z_i Z_i') = \sigma^2 Q_{ZZ}$

Covariance Matrix Estimation

- Estimator of the asymptotic variance matrix $\boldsymbol{V}_{\boldsymbol{\beta}}$:

$$\hat{\mathbf{V}}_{\boldsymbol{\beta}} = \left(\hat{\boldsymbol{Q}}_{\boldsymbol{XZ}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZZ}}^{-\mathbf{1}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZX}}\right)^{-\mathbf{1}}\left(\hat{\boldsymbol{Q}}_{\boldsymbol{XZ}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZZ}}^{-\mathbf{1}}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZZ}}^{-\mathbf{1}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZX}}\right) \cdot \left(\hat{\boldsymbol{Q}}_{\boldsymbol{XZ}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZZ}}^{-\mathbf{1}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZX}}\right)^{-\mathbf{1}}$$

where

$$\hat{\boldsymbol{Q}}_{\boldsymbol{ZZ}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Z}_i\boldsymbol{Z}_i' = \frac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}$$

$$\hat{\boldsymbol{Q}}_{\boldsymbol{XZ}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{Z}_i' = \frac{1}{n}\boldsymbol{X}'\boldsymbol{Z}$$

$\Omega = E(e_i^2 Z_i Z_i')$

$$\hat{\boldsymbol{\Omega}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Z}_i\boldsymbol{Z}_i'\hat{e}_i^2$$

$e_i = Y_i - X_i'\beta_{2SL}$

$$\hat{e}_i = Y_i - \boldsymbol{X}_i'\hat{\boldsymbol{\beta}}_{2\text{sls}}.$$

Covariance Matrix Estimation

- The homoskedastic variance matrix can be estimated by

$$\hat{\mathbf{V}}_{\boldsymbol{\beta}}^{0} = \left(\hat{\boldsymbol{Q}}_{\boldsymbol{XZ}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZZ}}^{-\mathbf{1}}\hat{\boldsymbol{Q}}_{\boldsymbol{ZX}}\right)^{-\mathbf{1}}\hat{\sigma}^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{e}_i^2.$$

同方差假设 $E(e_i^2 \mid z_i) = \sigma^2$

$\sigma^2 = E(e_i^2)$

$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{e}_i^2$

Theorem

$$\hat{\mathbf{V}}_{\boldsymbol{\beta}}^{0} \xrightarrow{p} \mathbf{V}_{\boldsymbol{\beta}}^{0}$$

$$\hat{\mathbf{V}}_{\boldsymbol{\beta}} \xrightarrow{p} \mathbf{V}_{\boldsymbol{\beta}}.$$

Covariance Matrix Estimation

- The covariance matrix estimator should be constructed using the correct residual formula: $\widehat{e}_i = Y_i - \boldsymbol{X}_i' \widehat{\boldsymbol{\beta}}_{2\text{sls}}$.
- In the second stage, regress Y_i on $\widehat{\boldsymbol{X}}_i$, $\widehat{\boldsymbol{X}}_i = \widehat{\boldsymbol{\Gamma}}' \boldsymbol{Z}_i$.
- Residuals from the second stage: $Y_i = \widehat{\boldsymbol{X}}_i' \widehat{\boldsymbol{\beta}}_{2\text{sls}} + \hat{v}_i$.
- The standard errors reported by STATA for the second-stage regression use the residual $\hat{v}_i$. The (homoskedastic) formula it uses is

$$\hat{\mathbf{V}}_{\boldsymbol{\beta}} = \left(\frac{1}{n} \widehat{\boldsymbol{X}}' \widehat{\boldsymbol{X}} \right)^{-\mathbf{1}} \hat{\sigma}_v^2 = \left(\hat{\boldsymbol{Q}}_{\boldsymbol{XZ}} \hat{\boldsymbol{Q}}_{\boldsymbol{ZZ}}^{-\mathbf{1}} \hat{\boldsymbol{Q}}_{\boldsymbol{ZX}} \right)^{-\mathbf{1}} \hat{\sigma}_v^2$$

$$\hat{\sigma}_v^2 = \frac{1}{n} \sum_{i=1}^{n} \hat{v}_i^2.$$

- However,

$$\begin{aligned} \hat{v}_i &= Y_i - \boldsymbol{X}_i' \hat{\boldsymbol{\beta}}_{2\text{sls}} + \left(\boldsymbol{X}_i - \hat{\boldsymbol{X}}_i \right)' \hat{\boldsymbol{\beta}}_{2\text{sls}} \\ &\neq \hat{e}_i. \end{aligned}$$

$\hat{V}_\beta = (\hat{Q}_{xz} \hat{Q}_{zz}^{-1} \hat{Q}_{zx})^{-1} \hat{\sigma}^2$

$\hat{\sigma}^2 = \frac{1}{n} \sum \hat{e}_i^2$, $\hat{e}_i = Y_i - X_i' \hat{\beta}_{2sls}$

↑

正确版本

$s^2 = \frac{1}{n} \sum \hat{v}_i^2$, $\hat{v}_i = Y_i - \hat{X}_i' \hat{\beta}_{2sls}$

$\hat{X} = P_Z X = (\frac{1}{n} X'Z(Z'Z)^{-1} Z'X)^{-1} s^2$

$= (\frac{1}{n} X'Z (\frac{1}{n} Z'Z)^{-1} \frac{1}{n} Z'X)^{-1} s^2$

$= (\hat{Q}_{xz} \hat{Q}_{zz}^{-1} \hat{Q}_{zx})^{-1} s^2$

s^2 一般是低估了 $\hat{\sigma}^2$

Functions of Parameters

- Given $\boldsymbol{r} : \mathbb{R}^k \to \Theta \subset \mathbb{R}^q$, the parameter of interest is $\boldsymbol{\theta} = \boldsymbol{r}(\boldsymbol{\beta})$.
- A natural estimator is $\widehat{\boldsymbol{\theta}}_{\text{2sls}} = \boldsymbol{r}\left(\widehat{\boldsymbol{\beta}}_{\text{2sls}}\right)$.

Theorem
r is continuous at $\boldsymbol{\beta}$, then $\hat{\boldsymbol{\theta}}_{\text{2sls}} \xrightarrow{p} \boldsymbol{\theta}$ as $n \longrightarrow \infty$.

- Estimator of the asymptotic variance matrix:

$$\hat{\mathbf{V}}_\theta = \hat{\boldsymbol{R}}' \hat{\mathbf{V}}_{\boldsymbol{\beta}} \hat{\boldsymbol{R}}$$

$$\hat{\boldsymbol{R}} = \frac{\partial}{\partial \boldsymbol{\beta}} r\left(\hat{\boldsymbol{\beta}}_{\text{2sls}}\right)'$$

$\sqrt{n}(\hat{\beta}_{2sls} - \beta) \xrightarrow{d} N(0, V_\beta)$

$\theta = r(\beta)$. $\hat{\theta}_{2sls} = r(\hat{\beta}_{2sls})$

$\sqrt{n}(\hat{\theta}_{2sls} - \theta) \xrightarrow{d} N(0, R' V_\beta R)$

$R = \left.\frac{\partial r(b)}{\partial b'}\right|_{b=\beta}$

$\downarrow$ 一致估计

$\hat{R} = \left.\frac{\partial r(b)}{\partial b'}\right|_{b=\hat{\beta}_{2sls}}$

Functions of Parameters

Theorem
If r is continuously differentiable at $\boldsymbol{\beta}$,

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}_{2\text{sls}} - \boldsymbol{\theta}\right) \xrightarrow{d} \mathrm{N}\left(\mathbf{0}, \mathbf{V}_{\theta}\right)$$

where

$$\mathbf{V}_{\theta} = \boldsymbol{R}'\mathbf{V}_{\boldsymbol{\beta}}\boldsymbol{R}$$
$$\boldsymbol{R} = \frac{\partial}{\partial \boldsymbol{\beta}} r\left(\boldsymbol{\beta}\right)'$$

and $\hat{\mathbf{V}}_{\theta} \xrightarrow{p} \mathbf{V}_{\theta}$.

Hypothesis Tests

- We are interested in testing

$$\mathbb{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$$
$$\mathbb{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0.$$

- The Wald statistic:

$$W = n\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)' \hat{\mathbf{V}}_{\hat{\theta}}^{-1}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right).$$

Theorem

$$W \xrightarrow{d} \chi^2_q.$$

For c satisfying $\alpha = 1 - G_q(c)$,

$$\Pr\left(W > c \mid \mathbb{H}_0\right) \longrightarrow \alpha$$

so the test "Reject $\mathbb{H}_0$ if $W > c$" has asymptotic size α.