1. Modelo Lineal de Variables Instrumentales

Consideremos el siguiente modelo

$$\underbrace{Y_i}_{1\times 1} = \underbrace{X_{1i}}_{1\times K_1}{}'\delta + \underbrace{X_{2i}{}'}_{1\times K_2}\beta + \underbrace{u_i}_{1\times 1}, \quad (1)$$: Ecuación Estructural

$$\underbrace{X_{1i}{}'}_{1\times K_1} = \underbrace{Z_i{}'}_{1\times \ell}\pi_1{}' + \underbrace{X_{2i}{}'}_{1\times K_2}\pi_2{}' + \underbrace{V_i{}'}_{1\times K_1}, \quad (2)$$: Primera etapa

donde X_{1i} es un vector de K_1 variables endógenas, X_{2i} es un vector de K_2 variables exógenas, y Z_i es un vector de ℓ variables instrumentales. Además, el modelo satisface:

$$\left.\begin{aligned} E\, Z_i\, u_i &= 0 \\ E\, Z_i\, V_i{}' &= 0 \\ E\, X_{2i}\, u_i &= 0 \\ E\, X_{2i}\, V_i{}' &= 0 \end{aligned}\right\}$$

variables exógenas no están correlacionadas con los términos de error en (1) y (2).

(*) La diferencia entre X_{2i} y Z_i es que Z_i no está directamente formando parte de la ecuación (1).

1.1. Identificación

Vamos a explotar los siguientes momentos:

$$E\begin{pmatrix} Z_i \\ X_{2i}\end{pmatrix} u_i = \begin{pmatrix} 0_{\ell\times 1} \\ 0_{K_2\times 1}\end{pmatrix} \quad (3)$$

Reemplazando la ecuación estructural (1) en (3) obtenemos:

$$\begin{pmatrix} E\, Z_i\,(Y_i - X_{1i}{}'\delta - X_{2i}{}'\beta) \\ E\, X_{2i}\,(Y_i - X_{1i}{}'\delta - X_{2i}{}'\beta)\end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} E\, Z_i Y_i \\ E\, X_{2i} Y_i \end{pmatrix} = \begin{pmatrix} E\, Z_i X_{1i}{}' & E\, Z_i X_{2i}{}' \\ E\, X_{2i} X_{1i}{}' & E\, X_{2i} X_{2i}{}' \end{pmatrix} \underbrace{\begin{pmatrix} \delta_{K_1\times 1} \\ \beta_{K_2\times 1} \end{pmatrix}}_{K = K_1+K_2} \quad (4)$$

Definamos $\tilde{Z}_i = \begin{pmatrix} Z_i \\ X_{2i} \end{pmatrix}_{(\ell+K_2)\times 1}$ como nuestro vector de variables exógenas, $X_i = \begin{pmatrix} X_{1i} \\ X_{2i} \end{pmatrix}_{(K_1+K_2)\times 1}$,

y definamos $\tilde{\delta} := \begin{pmatrix} \delta \\ \beta \end{pmatrix}_{(K_1+K_2)\times 1}$ como el vector de coeficientes estructurales.

Finalmente, definimos $\tilde{\ell} := \ell + K_2$, $K = K_1 + K_2$.

Entonces podemos ver a la ecuación 4 como el siguiente sistema lineal:

$$\underbrace{E\,\tilde{Z}_i Y_i}_{\tilde{l}\times 1} = \underbrace{E\,\tilde{Z}_i X_i'}_{\tilde{l}\times K}\ \underbrace{\tilde{\delta}}_{K\times 1}$$

Caso 1: Identificación Exacta

Si $\tilde{l} = K \iff l + K_2 = K_1 + K_2 \iff l = K_1$ $\left(\begin{matrix}\#\text{ instrumentos} \\ = \\ \#\text{ endógenas}\end{matrix}\right)$,

entonces la matriz $E\,\tilde{Z}_i X_i'$ es $K\times K$ (cuadrada). Además, si la matriz es positiva definida, la podemos invertir y solucionar para $\tilde{\delta}$:

$$\underbrace{E\,\tilde{Z}_i Y_i}_{K\times 1} = \underbrace{E\,\tilde{Z}_i X_i'}_{K\times K}\ \underbrace{\tilde{\delta}}_{K\times 1}$$

$$\Rightarrow \left(E\,\tilde{Z}_i X_i'\right)^{-1} E\,\tilde{Z}_i Y_i = \tilde{\delta}$$

$$\Rightarrow \begin{pmatrix} E\,Z_i Z_i' & E\,Z_i X_{2i}' \\ E\,X_{2i} Z_i' & E\,X_{2i} X_{2i}' \end{pmatrix}^{-1} \begin{pmatrix} E\,Z_i Y_i \\ E\,X_{2i} Y_i \end{pmatrix} = \begin{pmatrix} \delta \\ \beta \end{pmatrix} \qquad (5)$$

Cuando reemplazamos esperanzas por promedios en 5 lo llamamos el estimador de variables instrumentales.

¿Cuándo se cumple que $E\,\tilde{Z}_i X_i'$ es definida positiva? Debemos mostrar que pasa si no tiene rango completo K:

$$\begin{pmatrix} E\,Z_i X_{1i}' & E\,Z_i X_{2i}' \\ E\,X_{2i} X_{1i}' & E\,X_{2i} X_{2i}' \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{para } \begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Reemplazo 2

$$\Rightarrow \begin{pmatrix} E\,Z_i Z_i' \pi_1' + E\,Z_i X_{2i}' \pi_2' & E\,Z_i X_{2i}' \\ E\,X_{2i} Z_i' \pi_1' + E\,X_{2i} X_{2i}' \pi_2' & E\,X_{2i} X_{2i}' \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} E\,Z_i Z_i' \pi_1' a + E\,Z_i X_{2i}' \pi_2' a + E\,Z_i X_{2i}' b \\ E\,X_{2i} Z_i' \pi_1' a + E\,X_{2i} X_{2i}' \pi_2' a + E\,X_{2i} X_{2i}' b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} E\,Z_i Z_i' \pi_1' a + E\,Z_i X_{2i}' (\pi_2' a + b) \\ E\,X_{2i} Z_i' \pi_1' a + E\,X_{2i} X_{2i}' (\pi_2' a + b) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \underbrace{\begin{pmatrix} E\, z_i z_i' & E\, z_i x_{2i}' \\ E\, x_{2i} z_i' & E\, x_{2i} x_{2i}' \end{pmatrix}}_{\text{Esto es positivo definido siempre que } z_i \text{ y } x_{2i} \text{ sean linealmente independientes.}} \underbrace{\begin{pmatrix} \pi_1' a \\ \pi_2' a + b \end{pmatrix}}_{\text{Esto debe ser } =0 \text{ para cumplir la ecuación.}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Entonces para que $E\, z_i x_i'$ no tenga rango completo debe ser porque

$$\begin{aligned} \pi_1' a &= 0 \\ \pi_2' a + b &= 0 \end{aligned}$$

(a) $a = 0$, $b \neq 0$

No puede ocurrir porque si $a=0$ entonces la segunda ecuación dice que $b=0$, pero $\begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

(b) $a \neq 0$, $b = 0$

Tanto π_1' como π_2' no tienen rango completo.

(c) $a \neq 0$, $b \neq 0$

π_1' debe tener rango incompleto.

Es decir, para que el modelo este identificado debe ocurrir que

$$E\, z_i u_i = 0 \quad \text{(Exogeneidad)}$$

$$\pi_1' \neq 0 \quad \text{(Relevancia, rango completo de } \pi_1' \text{)}$$

Caso 2: Sobreidentificación

Recuerden

$$\underbrace{E\, \tilde{z}_i y_i}_{\tilde{\ell} \times 1} = \underbrace{E\, \tilde{z}_i x_i'}_{\tilde{\ell} \times k} \underbrace{\tilde{\delta}}_{k \times 1},$$

donde $E\, \tilde{z}_i x_i'$ no es cuadrada, por lo que no podemos invertirla incluso si $E\, \tilde{z}_i x_i'$ tiene rango completo ($\neq 0$).

Podemos pre-multiplicar una nueva matriz $W_{\tilde{\ell} \times \tilde{\ell}}$ que es simétrica y positiva definida.

$$\underbrace{W}_{\tilde{\ell} \times \tilde{\ell}} \underbrace{E\, \tilde{z}_i y_i}_{\tilde{\ell} \times 1} = \underbrace{W}_{\tilde{\ell} \times \tilde{\ell}} \underbrace{E\, \tilde{z}_i x_i'}_{\tilde{\ell} \times k} \underbrace{\tilde{\delta}}_{k \times 1}$$

Pre multiplicamos $E\, x_i \tilde{z}_i'$ $\Rightarrow$

$$E\, x_i \tilde{z}_i' \, W \, E \tilde{z}_i y_i = \underbrace{\underbrace{E\, x_i z_i'}_{k\times \ell} \, \underbrace{W}_{\ell\times\ell} \, \underbrace{E \tilde{z}_i x_i'}_{\ell\times k}}_{\text{¡Ahora sí es cuadrada!}} \tilde{\delta}$$

$$\Rightarrow \quad \left(E\, x_i z_i' \, W \, E \tilde{z}_i x_i'\right)^{-1} E\, x_i \tilde{z}_i' \, W \, E \tilde{z}_i y_i = \tilde{\delta} .$$

A esta última forma de resolver el sistema lineal se le llama una inversa generalizada. GMM: método de momentos generalizado

La idea detrás es que tenemos más variables exógenas $\ell + k_2$ que parámetros de interés $k_1 + k_2$. Por ello, podemos usar distintas combinaciones de variables exógenas para identificar el coeficiente $\tilde{\delta}$. La matriz W representa las distintas formas de combinar.

1.2. Estimación 2SLS (Fitted Values approach)

$$\underbrace{y}_{n\times 1} = \underbrace{X_1}_{n\times k_1} \delta + \underbrace{X_2}_{n\times k_2} \beta + \underbrace{u}_{n\times 1} \qquad (6)$$

$$\underbrace{X_1}_{n\times k_1} = \underbrace{Z}_{n\times \ell} \underbrace{\Pi_1'}_{\ell\times k_1} + \underbrace{X_2}_{n\times k_2} \underbrace{\Pi_2'}_{k_2\times k_1} + \underbrace{V}_{n\times k_1} \qquad (7)$$

endógenas

(7): Tenemos una primera etapa para cada variable endógena.

Por simplicidad pongamos $k_1 = 1$ (una sola variable endógena).

Definimos

$$\tilde{z}_i = \begin{pmatrix} z_i \\ x_{2i} \end{pmatrix}_{(\ell+k_2)\times 1} \quad \text{y}$$

$$\tilde{Z} = \begin{pmatrix} \underbrace{\tilde{z}_1'}_{1\times\tilde{\ell}} \\ \tilde{z}_2' \\ \vdots \\ \tilde{z}_n' \end{pmatrix}_{n\times\tilde{\ell}} , \quad \text{donde}$$

$\tilde{\ell} = \ell + k_2$ es el # de variables exógenas.

Podemos estimar

$$\underbrace{\hat{X}_1}_{n\times 1} = \tilde{Z}(\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'X_1 = \underbrace{P_{\tilde{Z}}}_{n\times n} \underbrace{X_1}_{n\times 1} \qquad (8),$$

$$\underbrace{\hat{V}}_{n\times 1} = X_1 - P_{\tilde{Z}} X_1 = \underbrace{M_{\tilde{Z}}}_{n\times n} \underbrace{X_1}_{n\times 1} . \qquad (9)$$

Reemplazando $\hat{X}_1$ por X_1 en la primera etapa, y estimando por OLS obtenemos:

$$\hat{\delta} = (\hat{X}_1' M_{x_2} \hat{X}_1)^{-1} \hat{X}_1' M_{x_2} Y$$

$$= (X_1' P_{\tilde{z}} M_{x_2} P_{\tilde{z}} X_1)^{-1} X_1' P_{\tilde{z}} M_{x_2} Y$$

reemplazo ⑥

$$= \delta \underbrace{\frac{X_1' P_{\tilde{z}} M_{x_2} X_1}{X_1' P_{\tilde{z}} M_{x_2} P_{\tilde{z}} X_1}}_{\text{Ojalá esto sea } 1+o_p(1)} + \underbrace{\frac{X_1' P_{\tilde{z}} M_{x_2} u}{X_1' P_{\tilde{z}} M_{x_2} P_{\tilde{z}} X_1}}_{\text{Ojalá esto sea } o_p(1)}$$

dividimos y multiplicamos n

$$= \delta \frac{X_1' P_{\tilde{z}} M_{x_2} X_1 /n}{X_1' P_{\tilde{z}} M_{x_2} P_{\tilde{z}} X_1 /n} + \frac{X_1' P_{\tilde{z}} M_{x_2} u /n}{X_1' P_{\tilde{z}} M_{x_2} P_{\tilde{z}} X_1 /n}$$

Empecemos por

- $$\frac{X_1 P_{\tilde{z}} M_{x_2} u}{n} = \frac{X_1'\tilde{Z}}{n} \left(\frac{\tilde{Z}'\tilde{Z}}{n}\right)^{-1} \frac{\tilde{Z}'u}{n} - \frac{X_1'\tilde{Z}}{n} \left(\frac{\tilde{Z}'\tilde{Z}}{n}\right)^{-1} \frac{\tilde{Z}'X_2}{n} \left(\frac{X_2'X_2}{n}\right)^{-1} \frac{X_2 u}{n}$$

$$= \Sigma_{x_1,\tilde{z}} \Sigma_{\tilde{z},\tilde{z}}^{-1} \underbrace{\Sigma_{\tilde{z},u}}_{=0} - \Sigma_{x_1,\tilde{z}} \Sigma_{\tilde{z},\tilde{z}}^{-1} \Sigma_{\tilde{z},x_2} \Sigma_{x_2,x_2}^{-1} \underbrace{\Sigma_{x_2,u}}_{=0} + o_p(1),$$

siempre que todas esas matrices sean finitas.

Veamos $$\Sigma_{\tilde{z},\tilde{z}} = E\, \tilde{z}_i \tilde{z}_i' = E \begin{pmatrix} z_i \\ x_{2i} \end{pmatrix} (z_i' \ x_{2i}') = E \begin{pmatrix} z_i z_i' & z_i x_{2i}' \\ x_{2i} z_i' & x_{2i} x_{2i}' \end{pmatrix}$$

es finita si
$E\|z_i z_i'\| \leq E\|z_i\|^2 = O(1)$.
$E\|z_i x_{2i}'\| \leq (E\|z_i\|^2 \, E\|x_{2i}\|^2)^{1/2} = O(1)$
$E\|x_{2i} x_{2i}'\| \leq E\|x_{2i}\|^2 = O(1)$.

Además,

$$\Sigma_{\tilde{z},\tilde{z}}^{-1} = \left[E \begin{pmatrix} z_i z_i' & z_i x_{2i}' \\ x_{2i} z_i' & x_{2i} x_{2i}' \end{pmatrix} \right]^{-1}$$ es finito solo si

la matriz tiene rango completo $\tilde{l}$. Esto ocurre si x_{2i} y z_i no son perfectamente colineales.

Siguiendo la misma lógica

- $$\frac{X_1' P_{\tilde{z}} M_{x_2} P_{\tilde{z}} X_1}{n} = \frac{X_1' P_{\tilde{z}} X_1}{n} - \frac{X_1' \overbrace{P_{\tilde{z}} P_{x_2}}^{P_{x_2}} P_{\tilde{z}} X_1}{n}$$

$$= \frac{X_1'\tilde{Z}}{n} \left(\frac{\tilde{Z}'\tilde{Z}}{n}\right)^{-1} \frac{\tilde{Z}'X_1}{n} - \frac{X_1'X_2}{n} \left(\frac{X_2'X_2}{n}\right)^{-1} \frac{X_2'\tilde{Z}}{n} \left(\frac{\tilde{Z}'\tilde{Z}}{n}\right)^{-1} \frac{\tilde{Z}'X_1}{n}$$

$$= \Sigma_{x_1,\tilde z}\, \Sigma_{\hat z,\tilde z}^{-1}\, \Sigma_{\tilde z,x_1} - \Sigma_{x_1,x_2}\, \Sigma_{x_2,x_2}^{-1}\, \Sigma_{x_2,\tilde z}\, \Sigma_{\tilde z,\tilde z}^{-1}\, \Sigma_{\hat z,x_1} + o_p(1)$$

- $$\frac{X_1' P_{\tilde z} M_{X_2} X_1}{n} = X_1' \tilde z (\hat z' \hat z)^{-1} \tilde z X_1 - \underbrace{X_1' \hat z (\tilde z' \hat z)^{-1} \tilde z\, x_2 (x_2' x_2)^{-1}}_{\text{podemos transponer } 1\times 1} x_2' x_1$$

$$= \frac{X_1' \tilde z}{n} \left(\frac{\tilde z' \hat z}{n}\right)^{-1} \frac{\tilde z' X_1}{n} - \frac{X_1' x_2}{n} \left(\frac{x_2' x_2}{n}\right)^{-1} \frac{x_2' \hat z}{n} \left(\frac{\tilde z' \tilde z}{n}\right)^{-1} \frac{\hat z' X_1}{n}$$

$$= \Sigma_{x_1,\tilde z}\, \Sigma_{\hat z,\tilde z}^{-1}\, \Sigma_{\tilde z,x_1} - \Sigma_{x_1,x_2}\, \Sigma_{x_2,x_2}^{-1}\, \Sigma_{x_2,\tilde z}\, \Sigma_{\tilde z,\tilde z}^{-1}\, \Sigma_{\hat z,x_1} + o_p(1)$$

Finalmente, hemos mostrado que

$$\hat\beta_{2SLS} = \left\{ \frac{\Sigma_{x_1,\tilde z}\, \Sigma_{\hat z,\tilde z}^{-1}\, \Sigma_{\tilde z,x_1} - \Sigma_{x_1,x_2}\, \Sigma_{x_2,x_2}^{-1}\, \Sigma_{x_2,\tilde z}\, \Sigma_{\tilde z,\tilde z}^{-1}\, \Sigma_{\hat z,x_1}}{\Sigma_{x_1,\tilde z}\, \Sigma_{\hat z,\tilde z}^{-1}\, \Sigma_{\tilde z,x_1} - \Sigma_{x_1,x_2}\, \Sigma_{x_2,x_2}^{-1}\, \Sigma_{x_2,\tilde z}\, \Sigma_{\tilde z,\tilde z}^{-1}\, \Sigma_{\hat z,x_1}} \right\} \beta + o_p(1).$$

$$= \beta + o_p(1).$$

1.3 Régimen de varios IV

Sin perder generalidad, consideremos que no hay controles x_{2i}, de manera que nuestro estimador de dos etapas se ve así:

$$\hat\delta = \frac{X_1' P_z Y /n}{X_1' P_z X_1 /n}$$

$$= \delta \frac{X_1' P_z X_1 /n}{X_1' P_z X_1 /n} + \frac{X_1' P_z u /n}{X_1' P_z X_1 /n}$$

$$= \delta + \underbrace{\frac{X_1' P_z u /n}{X_1' P_z X_1 /n}}$$

Esto debe ser $o_p(1)$ para ser consistente. De lo contrario, hay sesgo asintótico.

La manera en que podemos modelar el hecho de que hay varios instrumentos dado el tamaño muestral es poniendo la siguiente condición:

$$\frac{\overset{\text{\# de instrumentos}\to}{\ell}}{n} \to \alpha \in (0,1) \quad \text{cuando} \quad n\to\infty .$$

Consideremos la matriz $\underbrace{P_Z}_{n\times n} = \underbrace{Z}_{n\times\ell}\underbrace{(Z'Z)^{-1}}_{\ell\times\ell}\underbrace{Z'}_{\ell\times n}$.

Es una matriz simétrica y real, por lo que admite una descomposición espectral

$$P_Z = \underbrace{C}_{n\times\ell}\,\underbrace{\Lambda}_{\ell\times\ell}\,\underbrace{C'}_{\ell\times n}$$

, donde Λ es una matriz diagonal cuyos elementos son los autovalores de P_Z.

Además $CC' = I_n$. Es decir, C es una matriz ortonormal donde cada columna es un autovector c_i que cumple $\|c_i\| = 1$.

La matriz P_Z es idempotente, por lo que cumple

$$P_Z \cdot c = \lambda c = P_Z^2 c = \lambda^2 c$$

$$\Rightarrow \lambda^2 - \lambda = 0$$

$$\Rightarrow \underbrace{\lambda(\lambda-1)}_{} = 0$$

Autovalores solo son 0 o 1.

Entonces, la proyección

$$\hat{y} = P_Z y$$

$$d\hat{y} = P_Z (y + dy) - P_Z y$$

$$\Rightarrow d\hat{y} = P_Z \, dy$$

$$\begin{bmatrix} d\hat{y}_1 \\ d\hat{y}_2 \\ \vdots \\ d\hat{y}_n \end{bmatrix} = \underbrace{P_Z}_{n\times n} \begin{bmatrix} dy_1 \\ dy_2 \\ \vdots \\ dy_n \end{bmatrix}$$

Esto implica $\dfrac{d\hat{y}_i}{dy_i} = P_{Z_{ii}}$. ¿Cómo se ve ese elemento?

fila y columna # i

$$\underbrace{Z}_{n\times \ell}\underbrace{(Z'Z)^{-1}}_{\ell\times\ell}\underbrace{Z'}_{\ell\times n} = \underbrace{Z}_{n\times \ell}\begin{bmatrix} \underbrace{(Z'Z)^{-1}}_{\ell\times\ell}\underbrace{z_1}_{\ell\times 1} & (Z'Z)^{-1}z_2 & \cdots & (Z'Z)^{-1}z_n \end{bmatrix}_{\ell\times n}$$

$$= \underbrace{\begin{bmatrix} z_1' \\ z_2' \\ \vdots \\ z_n' \end{bmatrix}}_{n\times\ell} \begin{bmatrix} \underbrace{(Z'Z)^{-1}}_{\ell\times\ell}\underbrace{z_1}_{\ell\times 1} & (Z'Z)^{-1}z_2 & \cdots & (Z'Z)^{-1}z_n \end{bmatrix}_{\ell\times n}$$

$$= \begin{bmatrix} z_1'(Z'Z)^{-1}z_1 & z_1'(Z'Z)^{-1}z_2 & \cdots & z_1'(Z'Z)^{-1}z_n \\ z_2'(Z'Z)^{-1}z_1 & z_2'(Z'Z)^{-1}z_2 & \cdots & z_2'(Z'Z)^{-1}z_n \\ \vdots & \vdots & & \vdots \\ z_n'(Z'Z)^{-1}z_1 & z_n'(Z'Z)^{-1}z_2 & \cdots & z_n'(Z'Z)^{-1}z_n \end{bmatrix}$$

Por lo que tenemos

$$PZ_{ii} = z_i(Z'Z)^{-1}z_i' . \longrightarrow$$

A estos se les conoce como statistical leverages y contienen información súper relevante sobre la multicolinearidad. Sus valores están entre 0 y 1.

Ahora estamos listos para entender que

$$\operatorname{rank}(P_Z) \underset{\substack{\text{propiedad}\\ \text{de}\\ \text{proyecciones}}}{=} tr(P_Z) = tr(Z(Z'Z)^{-1}Z')$$

$$\underset{\substack{\text{propiedad}\\ \text{traza}}}{=} tr((Z'Z)^{-1}Z'Z)$$

$$= tr(I_\ell)$$

$$= \ell, \quad (*)$$

Además $tr(Z(Z'Z)^{-1}Z') = \sum_{i=1}^{n} PZ_{ii} = \sum_{i=1}^{n} z_i(Z'Z)^{-1}z_i' \underset{\text{por } (*)}{=} \ell .$

Entonces noten que el leverage promedio

$$\frac{1}{n}\sum_{i=1}^{n} PZ_{ii} = \frac{\ell}{n} \longrightarrow \alpha \in (0,1)$$, en vez de irse hacia cero.

Es decir, normalmente una sola observación no es muy influyente en la proyección $\hat{y}_i$ pero cuando $\ell/n \to \alpha$ entonces sí hay observaciones que se mantienen fuertemente influyentes.

Regresando al modelo ...

$$\hat{\delta}_n = \delta + \frac{X_1' P_z u /n}{X_1' P_z X_1 /n}$$

$$= \delta + \frac{[Z\pi + v]' P_z u /n}{X_1' P_z X_1 /n}$$

$$= \delta + \underbrace{\frac{\pi' Z' P_z u /n}{X_1' P_z X_1 /n}}_{\substack{= o_p(1) \\ \text{no ha cambiado}}} + \frac{v' P_z u /n}{X_1' P_z X_1 /n}$$

Ahora
$$\frac{v' P_z u}{n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \, P_{Z_{ij}} \, u_j$$

$$= \underbrace{\frac{1}{n} \sum_{i=1}^{n} v_i \, P_{Z_{ii}} \, u_i}_{\text{Tomemos esta parte para empezar}} + \frac{1}{n} \sum_{i \neq j} v_i \, P_{Z_{ij}} \, u_j$$

Informalmente,

$$\frac{1}{n} \sum_{i=1}^{n} v_i \, P_{Z_{ii}} \, u_i = E\{v_i z_i' (E z_i z_i')^{-1} z_i u_i\}$$

si z_i es indep de (v_i, u_i)
$$= E\{z_i' (E z_i z_i')^{-1} z_i\} \; E\{v_i u_i\} + o_p(1)$$

$$= \frac{\ell}{n} \sigma_{v,u} + o_p(1)$$

$$\xrightarrow{p} \alpha \cdot \sigma_{v,u} \neq 0 .$$

Es decir, hay un sesgo asintótico que se mantiene por el hecho de que el # de variables instrumentales crezca.