Linear Model with Endogenous Controls

- $$y_i = \beta_1 \overset{\text{exogenous}}{x_{1i}} + \overset{\text{endogenous}}{x_{i2}'} \beta_2 + u_i$$

$$E\, x_{1i} u_i = 0$$

$$E\, x_{2i} u_i \neq 0$$

Running OLS yields

$\circledast$ $M_2 = I - P_{X_2}$.

$$\begin{aligned}
\hat{\beta}_1 &= (X_1' M_2 X_1)^{-1} X_1' M_2 Y \\
&= (X_1' M_2 X_1)^{-1} X_1' M_2 (X_1 \beta_1 + X_2 \beta_2 + u) \\
&= \beta_1 + \frac{X_1' M_2 u}{X_1' M_2 X_1}
\end{aligned}$$

$\circledast$ Let $\Sigma_{A,B} = E A_i B_i'$.

where

- $$\begin{aligned}
\frac{X_1' M_2 u}{n} &= \frac{X_1' u}{n} - \frac{X_1' P_2 u}{n} \\
&= \underbrace{\Sigma_{x_1,u}}_{=0} - \Sigma_{x_1,x_2} \Sigma_{x_2,x_2}^{-1} \underbrace{\Sigma_{x_2,u}}_{\neq 0} + o_p(1) \\
&= - \Sigma_{x_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,u} + o_p(1)
\end{aligned}$$

- $$\begin{aligned}
\frac{X_1' M_2 X_1}{n} &= \frac{X_1' X_1}{n} - \frac{X_1' P_2 X_1}{n} \\
&= \underbrace{\Sigma_{x_1,x_1} - \Sigma_{x_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,x_1}}_{\text{nonsingular unless } x_1 = x_2 \text{ a.s. which wouldn't make sense.}} + o_p(1)
\end{aligned}$$

Therefore

$$\hat{\beta}_1 = \beta_1 - \frac{\Sigma_{x_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,u}}{\Sigma_{x_1,x_1} - \Sigma_{x_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,x_1}} + o_p(1)$$

$\not\xrightarrow{p}$ β_1, unless $E\, X_{1i} X_{2i}' = 0$ (X_2 does not interfere on X_1).

Linear IV Model

endogenous, exogenous

- $\underset{1\times1}{Y_i} = \underset{1\times1}{\beta_1} \underset{1\times1}{X_{1i}} + \underset{1\times k_2}{X_{i2}'} \underset{k_2\times1}{\beta_2} + \underset{1\times1}{u_i}$ (1)

exogenous

- $\underset{1\times1}{X_{1i}} = \underset{1\times l_1}{Z_{1i}'} \underset{l_1\times1}{\pi_1} + \underset{1\times k_2}{X_{i2}'} \underset{k_2\times1}{\pi_2} + \underset{1\times1}{V_i}$ (2)

$$E Z_{i1} V_i = E Z_{i1} u_i = 0$$

$$E X_{i2} V_i = E X_{i2} u_i = 0$$

We also assume $\pi_1 \neq 0$, $\pi_2 \neq 0$, $E Z_{i1} X_{i2}' \neq 0$.

Suppose the econometrician runs : 1) $X_{i1} = Z_{i1}' \hat{\pi}_1^* + \hat{V}_i^*$,

2) $Y_i = \hat{\beta}_1^* X_{i1}^* + X_{i2}' \hat{\beta}_2^* + \hat{u}_i^*$.

What does the true model imply for this process ?

- $\hat{\pi}_2^* = (Z_1'Z_1)^{-1} Z_1 X_1$

$= (Z_1'Z_1)^{-1} Z_1'Z_1 \pi_1 + (Z_1'Z_1)^{-1} Z_1' X_2 \pi_2 + (Z_1'Z_1)^{-1} Z_1' V$

$= \underbrace{\pi_1 + \rho\pi_2}_{\pi_1^*} + o_p(1).$

- Therefore, 1) $X_{i1} = Z_{i1}'\pi_1 + X_{i2}'\pi_2 + V_i$

$= Z_{i1}'\pi_1 + Z_{i1}'\rho\pi_2 + X_{i2}'\pi_2 - Z_{i1}'\rho\pi_2 + V_i$

$= Z_{i1}'(\pi_1 + \rho\pi_2) + X_{i2}'\pi_2 - Z_{i1}'\rho\pi_2 + V_i$

$= Z_{i1}'\pi_1^* + V_i^*.$

$E Z_{1i}' V_i^* = 0$ by construct.

$E X_{i2}' V_i^* = \underbrace{E X_{i2}'(X_{i2}' - Z_i'\rho)}_{\text{Covariance between dep. var and error term} \neq 0}\pi_2$

2) $Y_i = \beta_1 X_{1i} + X_{2i}'\beta_2 + u_i$

$= \beta_1 (Z_{i1}'\pi_1^* + V_i^*) + X_{i2}'\beta_2 + u_i$

$= \beta_1 X_{1i}^* + X_{i2}'\beta_2 + u_i + \beta_1 V_i^*$

$= \beta_1 X_{1i}^* + X_{i2}'\beta_2 + u_i^*.$

Even if we knew x_{1i}^* we have an endogenous control in 2nd stage. To see this,

$$E\, x_{1i}^* x_{i2}' = E\, z_{i1}' \pi_1^* x_{i2}'$$

$$= E\, \pi_1^{*\prime} z_{i1} x_{i2}'$$

$$= \pi_1^{*\prime} \underbrace{E\, z_{i1} x_{i2}'}_{\neq 0} .$$

The only detail is that we don't know x_{1i}^* exactly, but we can consistently estimate π_1^*. Write

$$\hat{X}_1^* = P_{Z_1} X_1$$

Then

$$\hat{\beta}_1^* = (\hat{X}_1^{*\prime} M_2 \hat{X}_1^*)^{-1} \hat{X}_1^{*\prime} M_2 Y$$

$$= (\hat{X}_1^{*\prime} M_2 \hat{X}_1^*)^{-1} \hat{X}_1^{*\prime} M_2 (X_1 \beta_1 + X_2 \beta_2 + u)$$

$$= (\hat{X}_1^{\prime *} M_2 \underline{\hat{X}_1^*})^{-1} \hat{X}_1^{\prime *} M_2 \underline{X_1} \; \beta_1 + \frac{\hat{X}_1^{*\prime} M_2 u}{(\hat{X}_1^{*\prime} M_2 \hat{X}_1^*)^{-1}}$$

does not cancel! $\hat{X}_1^* \neq X_1$

- $$\frac{\hat{X}_1^{*\prime} M_2 \hat{X}_1^*}{n} = \frac{X_1' P_{Z_1} M_2 P_{Z_1} X_1}{n}$$

$$= \frac{X_1' P_{Z_1} X_1}{n} - \frac{X_1' P_{Z_1} P_{X_2} P_{Z_1} X_1}{n}$$

$$= \Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_1} - \Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_1} + o_p(1)$$

- $$\frac{\hat{X}_1^{*\prime} M_2 X_1}{n} = \frac{X_1' P_{z_1} M_2 X_1}{n}$$

$$= \frac{X_1' P_{z_1} X_1}{n} - \frac{X_1' P_{z_1} P_{x_2} X_1}{n}$$

$$= \Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_1} - \Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,x_1} + op(1)$$

- $$\frac{\hat{X}_1^{*\prime} M_2 u}{n} = \frac{X_1' P_{z_1} M_2 u}{n}$$

$$= \frac{X_1' P_{z_1} u}{n} - \frac{X_1' P_{z_1} P_{x_2} u}{n}$$

$$= op(1) \quad \text{because} \quad \Sigma_{z_1,u} = 0, \; \Sigma_{x_2,u} = 0.$$

Finally, we showed that

$$\hat{\beta}_1^* = \left\{ \frac{\Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_1} - \Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,x_1}}{\Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_1} - \Sigma_{x_1,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_2} \Sigma_{x_2,x_2}^{-1} \Sigma_{x_2,z_1} \Sigma_{z_1,z_1}^{-1} \Sigma_{z_1,x_1}} \right\} \beta_1$$

$$+ \; op(1)$$

$$\not\xrightarrow{p} \beta_1 \quad \text{unless} \quad \Sigma_{z_1,x_2} := E\, z_{1i} x_{2i}' = 0.$$