Consider the system

$$\underset{m\times m}{\Gamma}\underbrace{\begin{pmatrix} y_{1i} \\ \vdots \\ y_{mi}\end{pmatrix}_{m\times 1}}_{\text{endogenous}} = \underset{m\times \ell}{B}\underbrace{\begin{pmatrix} z_{1i} \\ \vdots \\ z_{\ell i}\end{pmatrix}_{\ell\times 1}}_{\text{exogenous}} + \begin{pmatrix} u_{1i} \\ \vdots \\ u_{mi}\end{pmatrix}_{m\times 1}$$

$$\Gamma \; Y_i = B \; Z_i + u_i$$

$$\text{and} \quad E \underset{\ell\times 1}{Z_i} \underset{1\times m}{u_i'} = \underset{\ell\times m}{0}$$

In this case the first equation is

$$\sum_{j=1}^{m} \Gamma_{1j} \, y_{ji} = \sum_{j=1}^{\ell} B_{1j} \, z_{ji} + u_{1i}$$

Normalize $\Gamma_{jj} = 1$

$$\underbrace{y_{1i}}_{\text{Becomes the dependent variable.}} + \sum_{j=2}^{m} \Gamma_{1j} \, y_{ji} = \sum_{j=1}^{\ell} B_{1j} \, z_{ji} + u_{1i}$$

Provided Γ^{-1} exists we can rewrite

$$\underset{m\times 1}{Y_i} = \underbrace{\Gamma^{-1} B}_{\Pi_{m\times \ell}} \underset{\ell\times 1}{Z_i} + \underbrace{\Gamma^{-1} u_i}_{V_{i\,m\times 1}} \qquad (\text{reduced form})$$

$$\text{and} \quad E \, Z_i V_i' = E \, Z_i u_i' \, (\Gamma^{-1})' = 0$$

Notice that we can estimate Π from an OLS regression.
We can focus without loss of generality on the first equation

$$\underset{m\times m}{\Gamma} \; \underset{m\times \ell}{\Pi} = \underset{m\times \ell}{B}$$

$$\begin{pmatrix} \underset{1\times 1}{1} & \underset{1\times m_1}{-\gamma_1'} & \underset{1\times(m-m_1-1)}{0'} \\ \Gamma_1 & \Gamma_2 & \Gamma_3 \end{pmatrix} \begin{pmatrix} \underset{1\times \ell_1}{\pi_1} & \underset{1\times(\ell-\ell_1)}{\pi_2} \\ \underset{m_1\times \ell_1}{\Pi_1} & \underset{m_1(\ell-\ell_1)}{\Pi_2} \\ \Pi_3 & \Pi_4 \end{pmatrix} = \begin{pmatrix} \underset{1\times \ell_1}{\beta_1'} & \underset{1\times(\ell-\ell_1)}{0'} \\ B_1 & B_2 \end{pmatrix}$$

where ℓ_1 are the number of exogenous regressors and m_1 the number of endogenous regressors included.

$$1\,\pi_1 - \gamma_1' \Pi_1 = \beta_1' \qquad (1)$$

$$1\,\pi_2 - \gamma_1' \Pi_2 = 0' \quad \Rightarrow \quad \pi_2 = \gamma_1' \Pi_2 \qquad (2)$$

We recover γ_1' from (2) which we use in (1) to recover β_1.

Therefore, the first equation looks like

$$\begin{pmatrix}1 & -\gamma_1' & 0'\end{pmatrix}\begin{pmatrix} y_{1i} \\ Y_{1i} \\ Y_{1i}^*\end{pmatrix} = \begin{pmatrix}\beta_1' & 0'\end{pmatrix}\underbrace{\begin{pmatrix} Z_{1i} \\ Z_{1i}^*\end{pmatrix}}_{Z_i} + u_{1i}$$

$$y_{1i} - \gamma_1' Y_{1i} = \beta_1' Z_{1i} + u_{1i}$$

$$\underset{1\times1}{y_{1i}} = \underset{1\times m_1}{\gamma_1'}\,\underset{m_1\times1}{Y_{1i}} + \underset{1\times\ell_1}{\beta_1'}\, Z_{1i} + \underset{1\times1}{u_{1i}}$$

We need at least m_1 instruments for the endogenous regressors included. This is the same as

$$\ell - \ell_1 \geq m_1 \qquad \text{(Necessary Condition)}$$

This relied on the fact that we could use (2) to recover γ_1':

$$\underset{1\times(\ell-\ell_1)}{\pi_2} = \underset{1\times m_1}{\gamma_1'}\,\underset{m_1\times(\ell-\ell_1)}{\Pi_2}$$

$$\Rightarrow \underset{(\ell-\ell_1)\times m_1}{\Pi_2'}\,\underset{m_1\times1}{\gamma_1} = \underset{(\ell-\ell_1)\times1}{\pi_2'}$$ rewritten as a linear system.

Therefore, the sufficient condition is that Π_2' is full column rank so that

$$\text{rank}(\Pi_2') = m_1 \qquad \text{(Sufficient Condition)}$$

and

$$\hat{\gamma}_1^{ILS} = (\hat{\Pi}_2')^{+}\,\hat{\pi}_2' = (\hat{\Pi}_2\hat{\Pi}_2')^{-1}\hat{\Pi}_2\,\hat{\pi}_2'$$

Moreover, recall the reduced form

$$\begin{pmatrix} y_{1i} \\ Y_{1i} \\ Y_{1i}^*\end{pmatrix} = \begin{pmatrix}\pi_1 & \pi_2 \\ \Pi_1 & \Pi_2 \\ \Pi_3 & \Pi_4\end{pmatrix}\begin{pmatrix} Z_{1i} \\ Z_{1i}^*\end{pmatrix} + \begin{pmatrix} v_{1i} \\ V_{1i} \\ V_{1i}^*\end{pmatrix}$$

$$y_{1i} = \pi_1 Z_{1i} + \pi_2 Z_{1i}^* + v_{1i} \qquad \text{("Reduced Form")}$$

$$Y_{1i} = \Pi_1 Z_{1i} + \Pi_2 Z_{1i}^* + V_{1i} \qquad \text{("First stage")}$$

$$Y_{1i}^* = \Pi_3 Z_{1i} + \Pi_4 Z_{1i}^* + V_{1i}^*$$

(*) That's why Π_2 is very important, since it captures the exogenous variation onto the endogenous regressors.

We can also rewrite the structural equation (for the first eq) as

$$\underset{1\times 1}{y_{1i}} = \underset{1\times m_1}{Y_{1i}'}\,\underset{m_1\times 1}{\gamma_1} + \underset{1\times \ell_1}{Z_{1i}'}\,\underset{\ell_1\times 1}{\beta_1} + \underset{1\times 1}{u_{1i}}$$

$$= \begin{pmatrix} \underset{1\times m_1}{Y_{1i}'} & \underset{1\times \ell_1}{Z_{1i}'} \end{pmatrix} \begin{pmatrix} \underset{m_1\times 1}{\gamma_1} \\ \underset{\ell_1\times 1}{\beta_1} \end{pmatrix} + u_{1i}$$

$$= \underset{1\times k_1}{X_{1i}'}\,\underset{k_1\times 1}{\delta} + \underset{1\times 1}{u_{1i}}$$

And our instruments are $Z_i = \begin{pmatrix} Z_{1i} \\ Z_{1i}^* \end{pmatrix}$, so that $E u_{1i} Z_i = 0$.

This looks like the linear IV models that we're used to. Then, we require

$$E\, Z_i X_{1i}' \text{ is full rank} \Rightarrow \operatorname{rank}(\underbrace{E Z_i X_{1i}'}_{\ell \times (m_1+\ell_1)}) = m_1+\ell_1 = k_1$$

Write

$$E\, Z_i X_{1i}' = E\begin{pmatrix} Z_{1i} \\ Z_{1i}^* \end{pmatrix} \underbrace{\begin{pmatrix} Y_{1i}' & Z_{1i}' \end{pmatrix}}_{\text{plug in the first stage}} = E\begin{pmatrix} Z_{1i} \\ Z_{1i}^* \end{pmatrix}\begin{pmatrix} (Z_{1i}'\Pi_1' + Z_{1i}^{*\prime}\Pi_2' + V_{1i}') & Z_{1i}' \end{pmatrix}$$

$$= \begin{pmatrix} E Z_{1i}Z_{1i}'\Pi_1' + E Z_{1i}Z_{1i}^{*\prime}\Pi_2' & E Z_{1i}Z_{1i}' \\ E Z_{1i}^* Z_{1i}'\Pi_1' + E Z_{1i}^* Z_{1i}^{*\prime}\Pi_2' & E Z_{1i}^* Z_{1i}' \end{pmatrix}$$

If it's not full rank then $\exists\, \theta$ (→ nonzero!) s.t.

$$E Z_i X_{1i}' \theta = \begin{pmatrix} E Z_{1i}Z_{1i}'\Pi_1'\theta_1 + E Z_{1i}Z_{1i}^{*\prime}\Pi_2'\theta_1 + E Z_{1i}Z_{1i}'\theta_2 \\ E Z_{1i}^* Z_{1i}'\Pi_1'\theta_1 + E Z_{1i}^* Z_{1i}^{*\prime}\Pi_2'\theta_1 + E Z_{1i}^* Z_{1i}'\theta_2 \end{pmatrix} = 0$$

$$= \underbrace{\begin{pmatrix} E Z_{1i}Z_{1i}' & E Z_{1i}Z_{1i}^{*\prime} \\ E Z_{1i}^* Z_{1i}' & E Z_{1i}^* Z_{1i}^{*\prime} \end{pmatrix}}_{\text{p.d. matrix of second moments}} \underbrace{\begin{pmatrix} \Pi_1'\theta_1 + \theta_2 \\ \Pi_2'\theta_1 \end{pmatrix}}_{\text{must be } 0} = 0$$

Then

$$\Pi_1'\theta_1 = -\theta_2$$

$$\underbrace{\Pi_2'\theta_1 = 0}$$

(*) Both θ_1 and $\theta_2 \neq 0$ because if one is 0 the other is also 0.

There's a nonzero vector that'll mapped to 0, i.e. $\operatorname{rank}(\Pi_2') < m_1$.

This implies that rank $(E Z_i X_{1i}') < k_1 \Rightarrow$ rank$(\Pi_2') < m_1$.
Now suppose rank$(\Pi_2') < m_1$. Then such θ_1 exist and we can find a $\theta_2 = -\Pi_1'\theta_1$ so that $E Z_i X_{1i}' = 0$, i.e. rank $(E Z_i X_{1i}') = 0$.

In other words,

$$\text{rank}(E Z_i X_{1i}') = k_1 \iff \text{rank}(\Pi_2') = m_1.$$

Multiple Equation GMM

Suppose M equations are identified:

$$\underset{1\times1}{y_{1i}} = \underset{1\times k_1}{X_{1i}'}\underset{k_1\times1}{\delta_1} + \underset{1\times1}{u_{1i}}$$
$$\vdots$$
$$\underset{1\times1}{y_{Mi}} = \underset{1\times k_M}{X_{Mi}'}\underset{k_M\times1}{\delta_M} + \underset{1\times1}{u_{Mi}} \quad \text{where} \quad E\underset{\ell\times1}{Z_i}\underset{1\times1}{u_{ji}} = 0$$
$$\text{rank}(E\underset{\ell\times1}{Z_i}\underset{1\times k_j}{X_{ji}'}) = k_j$$

In matrix form can be written as

$$\underset{n\times1}{y_1} = \underset{n\times k_1}{X_1}\underset{k_1\times1}{\delta_1} + \underset{n\times1}{u_1}$$
$$\vdots$$
$$\underset{n\times1}{y_M} = \underset{n\times k_M}{X_M}\underset{k_M\times1}{\delta_M} + \underset{n\times1}{u_M} \quad \text{where} \quad X_j = \begin{pmatrix} x_{j1}' \\ \vdots \\ x_{jn}' \end{pmatrix}.$$

$$\Rightarrow \quad y = \underset{nm\times k}{\begin{pmatrix} X_1 & & 0 \\ & \ddots & \\ 0 & & X_M \end{pmatrix}}\underset{k\times1}{\begin{pmatrix} \delta_1 \\ \vdots \\ \delta_M \end{pmatrix}} + u$$

where $k = k_1 + \dots + k_M$

$$\underset{mn\times1}{Y} = \underset{mn\times k}{X}\underset{k\times1}{\delta} + \underset{mn\times1}{u}$$

The population moment conditions are

$$E\begin{pmatrix} \underset{\ell\times1}{Z_i}\underset{1\times1}{u_{1i}} \\ \vdots \\ \underset{\ell\times1}{Z_i}\underset{1\times1}{u_{Mi}} \end{pmatrix} = E(\underset{m\times1}{u_i} \otimes \underset{\ell\times1}{Z_i}) = \underset{m\ell\times1}{0}$$

The sample moment conditions are

$$\begin{pmatrix} \underset{\ell\times n}{Z'}\underset{n\times1}{u_1} \\ \vdots \\ \underset{\ell\times n}{Z'}\underset{n\times1}{u_M} \end{pmatrix} = \begin{pmatrix} Z' & & 0 \\ & \ddots & \\ 0 & & Z' \end{pmatrix}\begin{pmatrix} u_1 \\ \vdots \\ u_M \end{pmatrix}$$

$$= (I_m \otimes Z')\begin{pmatrix} u_1 \\ \vdots \\ u_M \end{pmatrix}$$

$$= \underbrace{(\underset{m\times m}{I_m} \otimes \underset{n\times\ell}{Z})'}_{m\ell\times mn}\underset{mn\times1}{u} = (I_m \otimes Z)'(Y - X\delta) = \underset{m\ell\times1}{0}$$

We can solve the system as with a regular GMM estimator

$$\hat{\delta}_n = \underset{\delta \in \mathbb{R}^K}{\operatorname{argmin}} \ (Y - X\delta)'(I_m \otimes Z) W_n (I_m \otimes Z)'(Y - X\delta)$$

$$= \left[X'(I_m \otimes Z) W_n (I_m \otimes Z)' X \right]^{-1} X'(I_m \otimes Z) W_n (I_m \otimes Z)' Y$$

Recall the sample criterion function

$$\underbrace{(Y - X\delta)'(I_m \otimes Z)'}\ W_n (I_m \otimes Z)'(Y - X\delta)$$

$$= \begin{pmatrix} Z'(Y_1 - X_1\delta_1) \\ \vdots \\ Z'(Y_m - X_m\delta_m) \end{pmatrix}' \begin{pmatrix} \underset{\ell\times\ell}{W_{11}} & \cdots & \underset{\ell\times\ell}{W_{1M}} \\ \vdots & \ddots & \\ \underset{\ell\times\ell}{W_{M1}} & & \underset{\ell\times\ell}{W_{MM}} \end{pmatrix}_{m\ell\times m\ell} \begin{pmatrix} Z'(Y_1 - X_1\delta_1) \\ \vdots \\ Z'(Y_m - X_M\delta_M) \end{pmatrix}$$

$$= \begin{pmatrix} \underset{1\times n}{(Y_1 - X_1\delta_1)'} \underset{n\times\ell}{Z} \\ \vdots \\ \underset{1\times n}{(Y_m - X_m\delta_m)'} \underset{n\times\ell}{Z} \end{pmatrix} \begin{pmatrix} W_{11} & & W_{1m} \\ & \ddots & \\ W_{m1} & & W_{mm} \end{pmatrix} \begin{pmatrix} \underset{\ell\times n}{Z'} (Y_1 - X_1\delta_1)_{n\times 1} \\ \vdots \\ \underset{\ell\times n}{Z'} (Y_m - X_m\delta_m)_{n\times 1} \end{pmatrix}$$

$$= (Y_1 - X_1\delta_1)' Z W_{11} Z'(Y_1 - X_1\delta_1) + (Y_1 - X_1\delta_1)' Z W_{21} Z'(Y_1 - X_1\delta_1) + \ldots + (Y_m - X_m\delta_m)' Z W_{mm} Z'(Y_m - X_m\delta_m)$$

(*) If $W_{ij} = 0$ for $i \neq j$ this becomes the sum of the individual criterion functions !

$$\hat{\delta}_n - \delta = \left[\underline{X'(I_m \otimes Z)}_{n} W_n \underline{(I_m \otimes Z)' X}_{n} \right]^{-1} \underline{X'(I_m \otimes Z)}_{n} W_n \underline{(I \otimes Z)' u}_{n}$$

- $$\frac{(I_m \otimes Z)' X}{n} = \frac{1}{n} \begin{pmatrix} \underset{\ell\times n}{Z'} & & 0 \\ 0 & \ddots & \underset{\ell\times n}{Z'} \end{pmatrix} \begin{pmatrix} X_{1\, n\times k_1} & & 0 \\ 0 & \ddots & X_{m\, n\times k_m} \end{pmatrix} = \begin{pmatrix} \frac{Z'X_1}{n} & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \frac{Z'X_m}{n} \end{pmatrix}$$

$$\xrightarrow{p} \begin{pmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_m \end{pmatrix} := C$$

- $$\frac{(I_m \otimes Z)' u}{n} = \frac{1}{n} \begin{pmatrix} Z'u_1 \\ \vdots \\ Z'u_m \end{pmatrix} = \frac{1}{n} \sum_{i=1}^{n} \begin{pmatrix} Z_i u_{1i} \\ \vdots \\ Z_i u_{mi} \end{pmatrix} = \frac{1}{n} \sum_{i=1}^{n} (u_i \otimes Z_i)$$

Then $\sqrt{n}(\hat{\delta}_n - \delta) = [X'\underbrace{(I_m \otimes Z)}_{n} W_n \underbrace{(I_m \otimes Z)'X}_{n}]^{-1} \underbrace{X'(I_m \otimes Z)}_{n} W_n \underbrace{(I \otimes Z)'u}_{\sqrt{n}}$

- $$\frac{(I_m \otimes Z)'u}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \underbrace{(u_i \otimes z_i)}_{\xi_i} \xrightarrow{d} N(0, E\,\xi_i \xi_i')$$

$$= N(0, E[u_i u_i' \otimes z_i z_i'])$$

- $$E[u_i u_i' \otimes z_i z_i'] = E\left[\begin{pmatrix} u_{1i}u_{1i} & \cdots & u_{1i}u_{mi} \\ & \ddots & \\ u_{mi}u_{1i} & \cdots & u_{mi}u_{mi} \end{pmatrix} \otimes z_i z_i' \right]$$

$$= E\begin{bmatrix} u_{1i}^2 z_i z_i' & \cdots & u_{1i}u_{mi} z_i z_i' \\ \vdots & & \\ u_{mi}u_{1i} z_i z_i' & \cdots & u_{mi}^2 z_i z_i' \end{bmatrix}$$

To have Ω finite we need that every sub-matrix is finite.
Then

- $$E\|u_{ji}^2 z_i z_i'\| \underset{\text{Cauchy Schwartz}}{\leq} (E\,u_{ji}^4 \; \underbrace{E\|z_i z_i'\|^2}_{})^{1/2}$$

(*) known as Frobenious Norm

$$= E[\text{trace}(\underbrace{z_i z_i'}_{\ell \times \ell} \underbrace{z_i z_i'}_{\ell \times \ell})]$$

$$= E[\text{trace}(z_i \underbrace{z_i' z_i}_{1 \times 1} z_i')]$$

$$= E[\underbrace{z_i' z_i}_{\text{scalar = trace}} \; \underbrace{\text{trace}(z_i z_i')}_{\text{trace}(A) = \text{trace}(A')}]$$

$$= E[\text{trace}(z_i' z_i) \; \text{trace}(z_i' z_i)]$$

$$= E[(\text{trace}(z_i' z_i)^{1/2})^4] = E[\|z_i\|^4]$$

Then we require $E\,u_{ji}^4 < \infty$ and $E\|z_i\|^4 < \infty$.

- $$E\|u_{ji} u_{si} z_i z_i'\| \leq (\underbrace{E\,u_{ji}^2 u_{si}^2}_{} \; E\|z_i z_i'\|^2)^{1/2}$$

$$\leq (E\,u_{ji}^4 \cdot E\,u_{si}^4)^{1/2}$$

so we don't require further assumptions.

• Alternative way : Consider how $E\, u_{ji}^2 z_i z_i'$ looks

$$E\, u_{ji}^2 z_i z_i' = E\left[\underset{1\times1}{u_{ji}^2} \begin{pmatrix} \underset{1\times1}{z_{1i} z_{1i}} & \cdots & \underset{1\times1}{z_{1i} z_{\ell i}} \\ \vdots & & \\ \underset{1\times1}{z_{\ell i} z_{1i}} & \cdots & \underset{1\times1}{z_{\ell i} z_{\ell i}} \end{pmatrix} \right]$$

Consider an arbitrary element in the (r,s) position

$$E(u_{ji}^2 z_{ri} z_{si}) \le (E\, u_{ji}^4 \; E\, z_{ri}^2 z_{si}^2)^{1/2}$$

$$\le (E\, u_{ji}^4 \; (E z_{ri}^4 \, E z_{si}^4)^{1/2})^{1/2}$$

and we reach the same conditions, since $E\, z_{ri}^4 < \infty$ implies $E\, \|z_i\|^4 < \infty$. However, these are a set of more primitive assumptions.

When $E(u_i u_i' \otimes z_i z_i')$ is block diagonal i.e. $E u_{ji} u_{ri} z_i z_i' = 0_{\ell\times\ell}$. then the efficient weighting matrix W_n is

$$\Omega^{-1} = \begin{bmatrix} E u_{1i}^2 z_i z_i' & & 0 \\ & \ddots & \\ 0 & & E u_{mi}^2 z_i z_i' \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} (E u_{1i} z_i z_i')^{-1} & & 0 \\ & \ddots & \\ 0 & & (E u_{mi}^2 z_i z_i')^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \Omega_{11}^{-1} & & 0 \\ 0 & \ddots & \Omega_{mm}^{-1} \end{bmatrix}$$

So the criterion function will be the sum of the single equation criterion functions since W_n is block diagonal. Intuitively, these are no efficiency gains since the other equations don't provide information to the equation that we're interested.

(2 step GMM as efficient dimension reduction)

Consider the IV model

$$y_i = X_i'\beta + u_i \quad , \quad \beta \in \mathbb{R}^k \text{ and } Z_i \text{ is } \ell\times 1 \text{ with } \ell > k$$

$$E Z_i u_i = 0$$

suppose we use a matrix $B_{k\times \ell}$ that has rank k to linearly transform Z_i into a vector of k instruments $\underset{k\times 1}{W_i} := \underset{k\times\ell}{B}\ \underset{\ell\times 1}{Z_i}$. We then run IV with W_i.

$$\tilde{\beta}_n(B) = \left(\sum_{i=1}^n W_i X_i'\right)^{-1} \sum_{i=1}^n W_i y_i = (W'X)^{-1} W'Y$$

⊛ $W = \begin{pmatrix} W_1' \\ \vdots \\ W_n' \end{pmatrix} = \begin{pmatrix} Z_1' \\ \vdots \\ Z_n' \end{pmatrix} B' = Z\, \underset{k\times\ell}{B'}$

$$\tilde{\beta}_n(B) - \beta = \left(\frac{W'X}{n}\right)^{-1} \frac{W'U}{n}$$

$$\sqrt{n}\left(\tilde{\beta}_n(B) - \beta\right) = \left(\frac{W'X}{n}\right)^{-1} \frac{W'U}{\sqrt{n}} = \left(B \frac{Z'X}{n}\right)^{-1} B \frac{Z'U}{\sqrt{n}}$$

- $\dfrac{Z'U}{\sqrt{n}} = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{i=1}^n u_i Z_i \xrightarrow[d]{} N(0, E u_i^2 Z_i Z_i')$

- $\dfrac{Z'X}{n} \xrightarrow[p]{} E Z_i X_i' := Q$

  by Slutsky's theorem $\left(B \dfrac{Z'X}{n}\right)^{-1} \xrightarrow[p]{} (BQ)^{-1}$

Then

$$\sqrt{n}\left(\tilde{\beta}_n(B) - \beta\right) \xrightarrow[d]{} N\left(0, (BQ)^{-1} B \Omega B' (Q'B')^{-1}\right)$$

$$= N\left(0, \left(Q'B'(B\Omega B')^{-1} BQ\right)^{-1}\right)$$

⊛ $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$

Recall that the Asy-var of the efficient GMM is $V^* = (Q'\Omega^{-1}Q)^{-1}$.
Then we want to show

$$V(B) - V^* \text{ p.s.d} \iff V^{*-1} - V(B)^{-1} \text{ p.s.d}$$

$$= Q'\Omega^{-1}Q - Q'B'(B\Omega B')^{-1}BQ$$

$$= \underset{k\times\ell}{Q'}\ \underset{\ell\times\ell}{\Omega^{-1/2}} \underbrace{\left(I_\ell - \underbrace{\Omega^{1/2}B'}_{H} \underbrace{(B\Omega B')^{-1}}_{H'H} \underbrace{B\Omega^{1/2}}_{H'}\right)}_{\text{projection matrix}} \underset{\ell\times\ell}{\Omega^{-1/2}}\ \underset{\ell\times k}{Q}$$

Therefore, it is p.s.d.

Notice that we can find the B such that it attains the lower bound V^*.

$$\underbrace{Q' B' (B \Omega B')^{-1}}_{\text{we want this to cancel each other}} \underbrace{B Q}_{\text{this should be this}} = Q' \Omega^{-1} Q$$

$$\Rightarrow \boxed{B^* = Q' \Omega^{-1}} \quad \text{and plug it to confirm :}$$

$$Q' \cancel{\Omega^{-1}} Q \, (\cancel{Q' \Omega^{-1} \Omega \Omega^{-1}} Q)^{-1} Q' \Omega^{-1} Q$$
$$= V^* \text{ as desired.}$$

We can estimate it as

$$B_n = \frac{X'Z}{n} \hat{\Omega}_n^{-1} = \sum_{i=1}^{n} \frac{x_i z_i'}{n} \hat{\Omega}_n^{-1}$$

$$\text{where} \quad \hat{\Omega}_n = \frac{1}{n} \sum_{i=1}^{n} (Y_i - X_i' \hat{\beta}_n)^2 Z_i Z_i'$$

$\hat{\beta}_n$ ↳ some consistent estimator of β

Finally, we can write

$$\hat{\beta}_n (B_n) = \left(B_n \sum_{i=1}^{n} Z_i X_i' \right)^{-1} B_n \sum_{i=1}^{n} Z_i Y_i$$

$$= \left(\sum_{i=1}^{n} X_i Z_i' \hat{\Omega}_n^{-1} \sum_{i=1}^{n} Z_i X_i' \right)^{-1} \sum_{i=1}^{n} X_i Z_i' \hat{\Omega}_n^{-1} \sum_{i=1}^{n} Z_i Y_i = \hat{\beta}^{2step}.$$

(Testing on structural Parameters without assuming identification)

$$\underset{1\times1}{y_i} = \underset{1\times k}{x_i'}\,\underset{k\times1}{\gamma} + \underset{1\times1}{u_i}$$

$$\underset{k\times1}{x_i} = \underset{k\times l}{\Pi'}\,\underset{l\times1}{z_i} + \underset{k\times1}{v_i}$$

which can be rewritten into matrix form by defining

$$\Pi = \underset{l\times k}{\begin{pmatrix} \underset{l\times1}{\pi_1} & \underset{l\times1}{\pi_2} & \cdots & \underset{l\times1}{\pi_k} \end{pmatrix}}, \quad X = \underset{n\times k}{\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}}$$

Structural Eq : $\underset{n\times1}{y} = \underset{n\times k}{X}\,\underset{k\times1}{\gamma} + \underset{n\times1}{u}$

First stage Eq : $\underset{n\times k}{X} = \underset{n\times l}{Z}\,\underset{l\times k}{\Pi} + \underset{n\times k}{V}$

Reduced Form Eq : $\underset{n\times1}{y} = \underset{n\times l}{Z}\,\underbrace{\Pi\gamma}_{\underset{l\times1}{\pi}} + \underbrace{V\gamma + u}_{\underset{n\times1}{W}}$

Vectorize the first stage equation

$$\mathrm{vec}(X) = \mathrm{vec}(Z\Pi I_k) + \mathrm{vec}(V)$$

$$\underset{nk\times1}{\begin{pmatrix} \underset{n\times1}{x_1} \\ \vdots \\ \underset{n\times1}{x_k} \end{pmatrix}} = \underbrace{(I_k \otimes Z)}_{nk\times kl}\,\underbrace{\mathrm{vec}(\Pi)}_{lk\times1} + \underset{nk\times1}{\begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}}$$

$$\underset{nk\times1}{\tilde{X}} = \underset{nk\times kl}{\tilde{Z}}\,\underset{kl\times1}{\mathrm{vec}(\Pi)} + \underset{nk\times1}{\tilde{V}}$$

Then we get

- $\hat{\pi} = \pi + (Z'Z)^{-1}Z'W$

  $$\sqrt{n}(\tilde{\pi} - \pi) = \left(\frac{Z'Z}{n}\right)^{-1}\frac{Z'W}{\sqrt{n}} = \left(\frac{Z'Z}{n}\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\underset{l\times1}{z_i}\,\underset{1\times1}{W_i}$$

- $\mathrm{vec}(\hat{\Pi}) = \mathrm{vec}(\Pi) + (\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'V$

  $$\sqrt{n}\left(\mathrm{vec}(\hat{\Pi}) - \mathrm{vec}(\Pi)\right) = \left(\frac{\tilde{Z}'\tilde{Z}}{n}\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\underset{k\times1}{\tilde{v}_i} \otimes \underset{l\times1}{z_i}\right)$$

Consider them jointly

$$\sqrt{n}\begin{pmatrix}\hat{\pi}-\pi \\ vec(\hat{\Pi})-vec(\Pi)\end{pmatrix} = \begin{pmatrix}\left(\frac{Z'Z}{n}\right)^{-1} & \\ & \left(\frac{\tilde{Z}'\tilde{Z}}{n}\right)^{-1}\end{pmatrix}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\underbrace{\begin{pmatrix}Z_i W_i \\ \tilde{V}_i \otimes Z_i\end{pmatrix}}_{\xi_i}$$

$$\xrightarrow{d} N\left(0, \begin{pmatrix}Q_Z^{-1} & \\ & Q_{\tilde{Z}}^{-1}\end{pmatrix}\underbrace{E\,\xi_i\xi_i'}_{\Omega}\begin{pmatrix}Q_Z^{-1} & Q_{\tilde{Z}}^{-1}\end{pmatrix}\right)$$

$$= N\left(0, \begin{pmatrix}\underset{\ell\times\ell}{Q_Z^{-1}} & \\ & \underset{k\ell\times k\ell}{Q_{\tilde{Z}}^{-1}}\end{pmatrix}\begin{pmatrix}\underset{\ell\times\ell}{\Omega_{11}} & \underset{\ell\times k\ell}{\Omega_{12}} \\ \underset{k\ell\times\ell}{\Omega_{21}} & \underset{k\ell\times k\ell}{\Omega_{22}}\end{pmatrix}\begin{pmatrix}\underset{\ell\times\ell}{Q_Z^{-1}} & \underset{k\ell\times k\ell}{Q_{\tilde{Z}}^{-1}}\end{pmatrix}\right)$$

$$= N\left(0, \begin{pmatrix}\underset{\ell\times\ell}{\Sigma_{11}} & \underset{\ell\times k\ell}{\Sigma_{12}} \\ \underset{k\ell\times\ell}{\Sigma_{21}} & \underset{k\ell\times k\ell}{\Sigma_{22}}\end{pmatrix}\right) = N(0, \Sigma)$$

Something really cool that we can do is to test $H_0: \gamma = \gamma_0$ using the equation $\Pi\gamma = \pi$ and our derived distribution.

Under H_0 we have $\Pi\gamma_0 = \pi$. Therefore,

$$\sqrt{n}(\hat{\Pi}\gamma_0 - \hat{\pi}) = \sqrt{n}(\hat{\Pi}\gamma_0 - \hat{\pi} - \Pi\gamma_0 + \pi)$$

$$= \sqrt{n}\left((\hat{\Pi}-\Pi)\gamma_0 - (\hat{\pi}-\pi)\right)$$

$\circledast$ $vec(\Pi\gamma) = (\gamma' \otimes I_\ell)\, vec(\Pi)$

$$= \sqrt{n}\left((\gamma_0' \otimes I_\ell)\, vec(\hat{\Pi}-\Pi) - (\hat{\pi}-\pi)\right)$$

$$= \left[-I_\ell \quad (\gamma_0' \otimes I_\ell)\right]\sqrt{n}\begin{pmatrix}\hat{\pi}-\pi \\ vec(\hat{\Pi}-\Pi)\end{pmatrix}$$

$$\xrightarrow{d} N\left(0, \left(-I_\ell \quad (\gamma_0' \otimes I_\ell)\right)\begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\begin{pmatrix}-I_\ell \\ (\gamma_0 \otimes I_\ell)\end{pmatrix}\right)$$

$$= N\Big(0, \underbrace{\Sigma_{11} - (\gamma_0' \otimes I_\ell)\overset{\Sigma_{12}'}{\Sigma_{21}} - \Sigma_{12}(\gamma_0 \otimes I_\ell) + (\gamma_0' \otimes I_\ell)\Sigma_{22}(\gamma_0 \otimes I_\ell)}_{\tilde{\Omega}}\Big)$$

We can get the following consistent estimator for $\tilde{\Omega}$

$$\hat{\tilde{\Omega}} = \hat{\Sigma}_{11} - (\gamma_0' \otimes I_\ell)\hat{\Sigma}_{21} - \hat{\Sigma}_{12}(\gamma_0 \otimes I_\ell) + (\gamma_0' \otimes I_\ell)\hat{\Sigma}_{22}(\gamma_0 \otimes I_\ell)$$

And we can construct the following Wald test using the fact that

$$\hat{\tilde{\Omega}}^{-1/2}\sqrt{n}(\hat{\Pi}\gamma_0 - \hat{\pi}) \xrightarrow{d} N(0, I_\ell)$$

$$\begin{aligned} \Rightarrow \quad W(\gamma_0) &= n(\hat{\Pi}\gamma_0 - \hat{\pi})'\hat{\tilde{\Omega}}^{-1}(\hat{\Pi}\gamma_0 - \hat{\pi}) \\ &= \sqrt{n}(\hat{\Pi}\gamma_0 - \hat{\pi})'\hat{\tilde{\Omega}}^{-1/2}\hat{\tilde{\Omega}}^{-1/2}\sqrt{n}(\hat{\Pi}\gamma_0 - \hat{\pi}) \\ &= \left[\hat{\tilde{\Omega}}^{-1/2}\sqrt{n}(\hat{\Pi}\gamma_0 - \hat{\pi})\right]'\left[\hat{\tilde{\Omega}}^{-1/2}\sqrt{n}(\hat{\Pi}\gamma_0 - \hat{\pi})\right] \\ &\xrightarrow{d} \chi^2_\ell \end{aligned}$$

so we reject when $W(\gamma_0) > \chi^2_{\ell, 1-\alpha}$.

Notice what happens to the asymptotic distribution under local alternatives $\gamma = \gamma_0 + \frac{\delta}{\sqrt{n}}$

$$\sqrt{n}(\hat{\Pi}\gamma - \hat{\pi}) = \sqrt{n}\left((\hat{\Pi} - \Pi)\gamma_0 - (\hat{\pi} - \pi)\right) - \Pi\delta$$

$$\xrightarrow{d} N(-\Pi\delta, \hat{\tilde{\Omega}})$$

Hence $W(\gamma_0) \xrightarrow{d} \chi^2_\ell(\delta'\Pi'\Pi\delta)$.

Finally, if $\delta \in N(\Pi)$ then the test will have trivial power:

$$\Pr_{\delta \in N(\Pi)}\left(W(\gamma_0) > \chi^2_{\ell, 1-\alpha}\right) \to \alpha, \text{ as } n \to \infty.$$