Linear IV model

$$y_i = \underset{1\times k_1}{x_{1i}'}\underset{k_1\times 1}{\delta} + \underset{1\times k_2}{x_{2i}'}\underset{k_2\times 1}{\beta} + \underset{1\times 1}{u_i} \qquad (1)$$

$$\underset{1\times k_1}{x_{1i}'} = \underset{1\times \ell}{z_i'}\underset{\ell\times k_1}{\pi_1'} + \underset{1\times k_2}{x_{2i}'}\underset{k_2\times k_1}{\pi_2'} + \underset{1\times k_1}{v_i'} \qquad (2)$$

where

$$E\, z_i u_i = 0$$
$$E\, z_i v_i' = 0$$
$$E\, x_{2i} u_i = 0$$
$$E\, x_{2i} v_i' = 0$$

A) Identification :

$$E\begin{pmatrix} z_i \\ x_{2i} \end{pmatrix} u_i = \begin{pmatrix} 0_{\ell\times 1} \\ 0_{k_2\times 1} \end{pmatrix}$$

Replace the structural equation (1)

$$\begin{pmatrix} E\, z_i(y_i - x_{1i}'\delta - x_{2i}'\beta) \\ E\, x_{2i}(y_i - x_{1i}'\delta - x_{2i}'\beta) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} E\, z_i y_i \\ E\, x_{2i} y_i \end{pmatrix} = \underbrace{\begin{pmatrix} E\, z_i x_{1i}' & E\, z_i x_{2i}' \\ E\, x_{2i} x_{1i}' & E\, x_{2i} x_{2i}' \end{pmatrix}}_{\text{must have full column rank to be able to take a generalized inverse (i.e. rank condition!)}} \begin{pmatrix} \delta \\ \beta \end{pmatrix}$$

We can replace (2) onto the rank condition.

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

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

$$\underbrace{\begin{pmatrix} E\, z_i z_i' & E z_i x_{2i}' \\ E\, x_{2i} z_i' & E\, x_{2i} x_{2i}' \end{pmatrix}}_{\substack{\text{If exogenous variables} \\ \text{are linearly independent} \\ \text{this matrix of 2nd} \\ \text{moments is p.d.}}} \underbrace{\begin{pmatrix} \pi_1' a \\ \pi_2' a + b \end{pmatrix}}_{\text{This must be 0.}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Then
$$\pi_1' a = 0$$
$$\pi_2' a + b = 0$$

CASE 1: $a = 0$, $b \neq 0$

Can't happen due to eq 2.

CASE 2: $a \neq 0$, $b = 0$

Both π_1' π_2 are rank deficient

CASE 3: $a \neq 0$, $b \neq 0$

π_1' is rank deficient.

The rank condition means that π_1' cannot be rank deficient.

B) Estimation Approach

B.1) Fitted Values Approach:

$$\underset{n\times 1}{y} = \underset{n\times k_1}{X_1}\, \underset{k_1\times 1}{\delta} + \underset{n\times k_2}{X_2}\, \underset{k_2\times 1}{\beta} + \underset{n\times 1}{u}$$

$$\underset{n\times k_1}{X_1} = \underset{n\times \ell}{Z}\, \underset{\ell\times k_1}{\pi_1'} + \underset{n\times k_2}{X_2}\, \underset{k_2\times k_1}{\pi_2'} + \underset{n\times k_1}{V}$$

We can vectorize and estimate (see old tutorial) but let's assume $k_1 = 1$.

$$\underset{n\times 1}{X_1} = \underset{n\times l}{Z}\ \underset{l\times 1}{\pi_1} + \underset{n\times k_2}{X_2}\ \underset{k_2\times 1}{\pi_2} + \underset{n\times 1}{V}$$

We define $\tilde{Z}_i = \underset{(l+k_2)\times 1}{\begin{pmatrix} z_i \\ x_{2i} \end{pmatrix}}$ and project x_{1i} onto the subspace of cols $\tilde{z}_i$.

$$\hat{X}_1 = Z(\tilde{Z}'\tilde{Z})\tilde{Z}'X_1 \qquad , \qquad \hat{V} = X_1 - P_{\tilde{Z}}X_1$$
$$= P_{\tilde{Z}}X_1. \qquad\qquad = M_{\tilde{Z}}X_1.$$

Next we plug this into the first equation and estimate

$$\hat{\delta} = (\hat{X}_1' M_{X_2} \hat{X}_1)^{-1} \hat{X}_1' M_{X_2} y$$
$$= (X_1' P_{\tilde{Z}} M_{X_2} P_{\tilde{Z}} X_1)^{-1} X_1 P_{\tilde{Z}} M_{X_2} y$$
$$= \delta \underbrace{\frac{X_1' P_{\tilde{Z}} M_{X_2} X_2}{X_1' P_{\tilde{Z}} M_{X_2} P_{\tilde{Z}} X_1}}_{\text{Hopefully } = 1 + o_p(1)} + \underbrace{\frac{X_1' P_{\tilde{Z}} M_{X_2} U}{\hat{X}_1' P_{\tilde{Z}} M_{X_2} P_{\tilde{Z}} X_1}}_{\text{Hopefully } = o_p(1)}$$

Take $\dfrac{X_1' P_{\tilde{Z}} M_{X_2} U}{n} = \dfrac{X_1' P_{\tilde{Z}} U}{n} - \dfrac{X_1' P_{\tilde{Z}} P_{X_2} U}{n}$

- $$\underbrace{\frac{X_1' (Z \ \ X_2)}{n}}\, n \underbrace{\begin{pmatrix} Z'Z & Z'X_2 \\ X_2'Z & X_2'X_2 \end{pmatrix}^{-1}} \underbrace{\begin{pmatrix} Z' \\ X_2' \end{pmatrix} \frac{U}{n}}$$

- $\dfrac{X_1'Z}{n} = E x_{1i} z_i' + o_p(1)$
  $= O(1)$
- $\dfrac{X_1'X_2}{n} = E x_{1i} x_{2i}' + o_p(1)$
  $= O(1)$

  if z_i, x_{2i} have finite 2nd moments.

- $\begin{pmatrix} E z_i z_i' & E z_i x_{2i}' \\ E x_{2i} z_i & E x_{2i} x_{2i}' \end{pmatrix}^{-1} + o_p(1)$

  if no collinearity this is bounded

- $\dfrac{Z'U}{n} = E z_i u_i + o_p(1)$
  $= o_p(1)$
- $\dfrac{X_2'U}{n} = E x_{2i} u_i + o_p(1)$
  $= o_p(1)$

- Notice that $P_{X_2} u$ involves $\frac{X_2' u}{n} = E X_{2i} u_i + o_p(1)$
$= o_p(1)$.

We only need to check the other matrices so that they don't explode.

B.2) Control Function Approach:

We once again regress X_{1i} onto cols $\tilde{z}_i$ and keep the residual

$$\hat{V} = M_{\tilde{z}} X_1$$
$$= (I - P_{\tilde{z}}) X_1$$

The second stage controls for this new regressor as well. Notice that

$$M_{\hat{v}} = I - \hat{v} (\hat{V}' \hat{v}) \hat{v}'$$
$$= I - (I - P_{\tilde{z}}) X_1 (X_1' (I - P_{\tilde{z}}) X_1)^{-1} X_1' (I - P_{\tilde{z}})$$

now

$$\delta = (X_1' M_{x_2} M_{\hat{v}} M_{x_2} X_1)^{-1} X_1' M_{x_2} M_{\hat{v}} M_{x_2} y$$

and once again we can define everything in terms of P_{x_2}, $P_{\tilde{z}}$ to analyze. See midterm 2019, which was also covered in a previous tutorial. In that exam the focus is on λ, but we could also focus on δ.

c) <u>Efficiency</u>

Not using all instruments is always (weakly) inefficient.

$\tilde{\beta}_n$ using $\tilde{W}_n = \begin{pmatrix} \hat{\Omega}_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$

$\hat{\beta}_n$ using $W_n^* = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1}$

$$Avar(\tilde{\beta}_n) = (Q'WQ)^{-1} Q' W \Omega W Q (Q'WQ)^{-1}$$

$$Avar(\hat{\beta}_n) = (Q'\Omega^{-1}Q)^{-1}$$

$Avar(\tilde{\beta}_n) - Avar(\hat{\beta}_n)$ is p.s.d.

iff (from Assignment 4)

$Avar(\hat{\beta}_n)^{-1} - Avar(\tilde{\beta}_n)^{-1}$ is p.s.d.

$$Q'\Omega^{-1}Q - Q'WQ(Q'W\Omega WQ)^{-1}Q'WQ$$

$$= Q'\Omega^{-1/2}\left(I - \Omega^{1/2}Q'WQ(Q'W\Omega^{1/2}\Omega^{1/2}WQ)^{-1}Q'WQ\Omega^{1/2}\right)\Omega^{-1/2}Q$$

$$= x\, Q'\Omega^{-1/2}(\underbrace{I - H(H'H)^{-1}H'}_{\text{idempotent and symmetric}})\Omega^{-1/2}Q\, x \geq 0$$

$x'x \geq 0$

d) Efficient Instruments

What if we have MEAN INDEP intead of uncorrelatedness?

$E[u_i | z_i] = 0$ instead of $E u_i z_i = 0$

$\Updownarrow$

$E[u_i g(z_i)] = 0$ for every measurable $g(\cdot)$.

$\Downarrow$

$g^{-1}(B) \in \Omega$ for every $B \in \mathcal{B}(\mathbb{R})$.

We can do GMM using $g(z_i)$ instead of z_i!

The efficient $g^*(z_i) = \dfrac{E(x_i | z_i)}{E(u_i^2 | z_i)}$ $\rightarrow$ project endogenous! $\rightarrow$ project squared errors!

You can use the same trick as part c) to show this is more efficient by using Law of Iterated Expectations.

e) <u>Asymptotics</u>

$$\hat{\beta_n}(W_n) = \beta_0 + \left(\frac{X'Z}{n} W_n \frac{Z'X}{n}\right)^{-1} \frac{X'Z}{n} W_n \frac{Z'u}{n}$$

$$\Rightarrow \sqrt{n}(\hat{\beta_n} - \beta_0) = \left(\frac{X'Z}{n} W \frac{Z'X}{n}\right)^{-1} \frac{X'Z}{n} W \frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_i u_i$$

$$= (\Sigma_{x,z} W \Sigma_{z,x})^{-1} \Sigma_{x,z} W \frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_i u_i$$

$$+ \; o_p(1) \underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_i u_i}_{O_p(1)}$$

if $Var(z_i u_i) = O(1)$ by CLT

$$= \underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i + o_p(1)}$$

ECON 626: This is called asymptotic linear representation of an estimator, and ξ_i is called an influence function.

where $\xi_i = (\Sigma_{x,z} W \Sigma_{z,x})^{-1} \Sigma_{x,z} W z_i u_i$

Notice that by CLT

$$\sqrt{n}(\hat{\beta_n} - \beta_0) \xrightarrow{d} N(0, E\xi_i \xi_i').$$

f) Inference

Even though we don't know the distribution of $\hat{\beta}$ because we don't know the joint distribution of the data, we can approximate the distribution of $\sqrt{n}(\hat{\beta}_n - \beta)$!

- Consider testing $\beta = \beta_0$:

$$W_n = \sqrt{n}(\hat{\beta}_n - \beta)' \widehat{Avar}(\hat{\beta}_n)^{-1} \sqrt{n}(\hat{\beta}_n - \beta)$$

Wald statistic !

$$= \sqrt{n}(\hat{\beta}_n - \beta)' Q' \Omega^{-1} Q \sqrt{n}(\hat{\beta}_n - \beta)$$

assuming we used efficient weights.

$$= \{\underbrace{(Q'\Omega^{-1}Q)^{1/2} \sqrt{n}(\hat{\beta}_n - \beta)}_{\text{If we plug } \beta=\beta_0 \text{ this is } \xrightarrow{d} N(0,I)}\}' \{(Q'\Omega^{-1}Q)^{1/2} \sqrt{n}(\hat{\beta}_n - \beta)\}$$

$$\xrightarrow[d]{H_0} \chi^2_k .$$

Now consider the alternative hypothesis $H_1: \beta = \beta_0 + \delta$

$$W_n \underset{H_1}{=} \underbrace{\{(Q'\Omega^{-1}Q)^{1/2} \sqrt{n}\, \delta\}' \{(Q'\Omega^{-1}Q)^{1/2} \sqrt{n}\, \delta\}}_{\text{explodes to infinity !}} + o_p(1)$$

It's not useful to mimic finite sample behavior.

then $H_1: \beta = \beta_0 + \delta/\sqrt{n}$

$$W_n \underset{H_1}{=} \left\{ (Q'\Omega^{-1}Q)^{1/2} (\sqrt{n}(\hat{\beta}_n - \beta_0) + \delta) \right\}' \left\{ (Q'\Omega^{-1}Q)^{1/2} (\sqrt{n}(\hat{\beta}_n - \beta_0) + \delta) \right\}$$

$$\xrightarrow{d} \chi^2_K \left(((Q'\Omega^{-1}Q)^{1/2}\delta)' ((Q'\Omega^{-1}Q)^{1/2}\delta) \right)$$

$$= \chi^2_K \left(\delta' \underbrace{(Q'\Omega^{-1}Q)}_{\text{This is } AVar(\hat{\beta})^{-1}} \delta \right)$$

This is $AVar(\hat{\beta})^{-1}$, which is p.d. because $AVAR(\hat{\beta}) = O(1)$.

So power increases as δ increases.

- Now consider testing $H_0: E z_i u_i = 0$. We'll use the criterion function itself!

Let $\bar{g}_n(b) = \frac{1}{n}\sum_{i=1} z_i(y_i - x_i'b)$

$$= \frac{1}{n} Z'(Y - Xb)$$

$$J_n(b) = \sqrt{n}\,\bar{g}_n(b)'\,\hat{\Omega}_n^{-1}\,\sqrt{n}\,\bar{g}_n(b).$$

Notice that

$$C'\sqrt{n}\,\bar{g}_n(\hat{\beta}_n) = \Omega^{-1/2}\sqrt{n}\,\bar{g}_n(\hat{\beta}_n)$$

$$= \Omega^{-1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_i(y_i - x_i'\hat{\beta}_n \pm x_i'\beta_0)$$

$$= \Omega^{-1/2} \frac{Z'u}{\sqrt{n}} - \Omega^{-1/2} Z'X \sqrt{n}(\hat{\beta}_n - \beta_0)$$

$$= \Omega^{-1/2} \frac{Z'u}{\sqrt{n}} - \Omega^{-1/2} Z'X \left(X'Z \hat{\Omega}_n^{-1} Z'X\right)^{-1} X'Z \hat{\Omega}_n^{-1} \frac{Z'u}{\sqrt{n}}$$

$$= \underbrace{\left\{ I_\ell - \Omega^{-1/2} Z'X \left(X'Z \hat{\Omega}_n^{-1} Z'X\right)^{-1} X'Z \hat{\Omega}_n^{-1} \Omega^{1/2} \right\}}_{D_n} \Omega^{-1/2} \frac{Z'u}{\sqrt{n}}$$

where

$$D_n = I_\ell - \Omega^{-1/2} \frac{Z'X}{n} \left(\frac{X'Z}{n} \hat{\Omega}_n^{-1} \frac{Z'X}{n} \right)^{-1} \frac{X'Z}{n} \hat{\Omega}_n^{-1} \Omega^{1/2}$$

$$= I_\ell - \Omega^{-1/2} E Z_i X_i' \left(E X_i Z_i' \hat{\Omega}_n^{-1} E Z_i X_i' \right)^{-1} E X_i Z_i' \hat{\Omega}_n^{-1} \Omega^{1/2} + o_p(1)$$

$$= I_\ell - \Omega^{-1/2} E Z_i X_i' \left[\left(E X_i Z_i' \Omega^{-1} E Z_i X_i' \right)^{-1} + o_p(1) \right] E X_i Z_i' \left[\Omega^{-1} + o_p(1) \right] \Omega^{1/2}$$
$$+ o_p(1)$$

$$= I_\ell - \Omega^{-1/2} E Z_i X_i' \left(E X_i Z_i' \Omega^{-1} E Z_i X_i' \right)^{-1} E X_i Z_i' \Omega^{-1/2} + o_p(1)$$

provided Ω^{-1}, EZ_iX_i' are bounded (i.e., $\mathrm{Var}(u_i Z_i)$ is pos def, $E\|Z_i X_i'\| \le (E\|Z_i\|^2 \, E\|X_i\|^2)^{1/2}$ finite second moments)

$$= I_\ell - \underbrace{\Omega^{-1/2} E Z_i X_i'}_{R} \Big(\underbrace{E X_i Z_i' \Omega^{-1/2}}_{R'} \underbrace{\Omega^{-1/2} E Z_i X_i'}_{R} \Big)^{-1} \underbrace{E X_i Z_i' \Omega^{-1/2}}_{R'} + o_p(1)$$

$$= \underbrace{I_\ell - R(R'R)^{-1} R'}_{} + o_p(1)$$

A projection matrix! Properties: - Eigenvalues are either 0 or 1
- Has orthonormal eigen decomp.

$$= D + o_p(1)$$

therefore

$$J_n(\hat\beta_n) = \left[\Omega^{-1/2}\sqrt{n}\,\bar g_n(\hat\beta_n)\right]' \,\Omega^{1/2}\,\hat\Omega_n^{-1}\,\Omega^{1/2}\left[\Omega^{-1/2}\sqrt{n}\,\bar g_n(\hat\beta_n)\right]$$

$$= \left[D_n\,\Omega^{-1/2}\frac{Z'U}{\sqrt{n}}\right]'\left[I_\ell + o_p(1)\right]\left[D_n\,\Omega^{-1/2}\frac{Z'U}{\sqrt{n}}\right]$$

$$= \left[D\,\underset{\circledast}{\Omega^{-1/2}}\frac{Z'U}{\sqrt{n}} + o_p(1)\,O_p(1)\right]'\left[I_\ell + o_p(1)\right]\left[D\,\Omega^{-1/2}\frac{Z'U}{\sqrt{n}} + o_p(1)\,O_p(1)\right]$$

$$= \left[N'\,D' + o_p(1)\right]\left[I_\ell + o_p(1)\right]\left[D\,N + o_p(1)\right]$$

$$= N'\,D\,N + o_p(1) \qquad \text{because } DN = O_p(1)$$

$$= N'\,C\Lambda C'\,N = \underbrace{\tilde N'}_{N(0,I)}\,\Lambda\,\underbrace{\tilde N}_{N(0,I)} = \sum_{i=1}^{tr(\Lambda)} \tilde n_i'\tilde n_i$$

$$\xrightarrow{d} \chi^2_{\mathrm{rank}(D)} = \chi^2_{tr(D)} = \chi^2_{\ell-K}.$$

For alternative hypothesis is the same idea with non-centered standard normals.

$$J_n(\hat\beta_n) \xrightarrow[d]{H_1} \chi^2_{\ell-K}\left(\delta'\,\Omega^{-1/2}\,D\,\Omega^{-1/2}\,\delta\right)$$

$\circledast$ Notice up top where this noncentrality appears

3) Multiple Equations (Extra :))

Suppose M equations are identified:

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

In matrix form can be written as

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

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

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

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

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

$$= \underbrace{(\underset{M\times M}{I_M} \otimes \underset{n\times\ell}{Z})'}_{M\ell\times Mn}\underset{Mn\times1}{u}$$

$$= (I_M \otimes Z)'(Y - X\delta)$$

$$= (I_M \otimes Z')(Y - X\delta)$$

$$= \underset{M\ell\times1}{0}$$

We can solve the system as with a regular GMM estimator

$$\hat{\delta}_n = \underset{d \in \mathbb{R}^K}{\arg\min} \; \| (I_m \otimes Z')(Y - X\delta) \|_{W_n}$$

$$= \underset{d \in \mathbb{R}^K}{\arg\min} \; [(I_m \otimes Z')(Y - X\delta)]' \, W_n \, [(I_m \otimes Z')(Y - X\delta)]$$

Where we can expand the criterion function.

$$= \begin{pmatrix} \underset{l \times n}{Z'}(\underset{n \times 1}{Y_1 - X_1\delta_1}) \\ \vdots \\ Z'(Y_m - X_M\delta_M) \end{pmatrix}' \underset{ml \times ml}{\begin{pmatrix} \underset{l \times l}{W_{11}} & \cdots & \underset{l \times l}{W_{1M}} \\ \vdots & \ddots & \\ \underset{l \times l}{W_{M1}} & & \underset{l \times l}{W_{MM}} \end{pmatrix}} \begin{pmatrix} Z'(Y_1 - X_1\delta_1) \\ \vdots \\ Z'(Y_m - X_M\delta_M) \end{pmatrix}$$

$$= \begin{pmatrix} (Y_1 - X_1\delta_1)'Z & \cdots & (Y_M - X_M\delta_M)'Z \end{pmatrix} \underset{ml \times ml}{\begin{pmatrix} \underset{l \times l}{W_{11}} & \cdots & \underset{l \times l}{W_{1M}} \\ \vdots & \ddots & \\ \underset{l \times l}{W_{M1}} & & \underset{l \times l}{W_{MM}} \end{pmatrix}} \begin{pmatrix} Z'(Y_1 - X_1\delta_1) \\ \vdots \\ Z'(Y_m - X_M\delta_M) \end{pmatrix}$$

consider this

$$= \begin{pmatrix} u_1'ZW_{11} + \dots + u_M'ZW_{M1} & \cdots & u_1'ZW_{1M} + \dots + u_M'ZW_{MM} \end{pmatrix} \begin{pmatrix} Z'(Y_1 - X_1\delta_1) \\ \vdots \\ Z'(Y_m - X_M\delta_M) \end{pmatrix}$$

$$= u_1'ZW_{11}Z'u_1 + \dots + u_M'ZW_{M1}Z'u_1 + \dots + u_1'ZW_{1M}Z'u_M + \dots + u_M'ZW_{MM}Z'u_M$$

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

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

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

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

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

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

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

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

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

provided

$$E \| u_{ji} u_{ri} z_i z_i' \| \le [E \| u_{ji} u_{ri} \|^2 \; E \| z_i z_i' \|^2]^{1/2}$$
$$\le [(E |u_{ji}|^4 E |u_{ri}|^4)^{1/2} \; E \| z_i \|^4]^{1/2}$$

so it suffices to assume fourth moments.

When $E(u_i u_i' \otimes z_i z_i')$ is block diagonal i.e. $E u_{ji} u_{ri} z_i z_i' = 0_{l \times e}$.
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, there are no efficiency gains since the other equations don't provide information to the equation that we're interested.