Three-Stage Least Squares (3SLS)

$$\underset{T\times 1}{y^j} = \underset{T\times r_j}{Y^j}\underset{r_j\times 1}{\delta^j} + \underset{T\times s_j}{Z^j}\underset{s_j\times 1}{\gamma^j} + \underset{T\times 1}{\varepsilon^j}, \quad j=1,\dots,M$$

$$= \underset{T\times k_j}{[Y^j \;\; Z^j]} \underset{k_j\times 1}{\begin{bmatrix}\delta^j \\ \gamma^j\end{bmatrix}} + \underset{T\times 1}{\varepsilon^j}, \quad k_j = r_j + s_j$$

$$\underset{T\times 1}{y^j} = \underset{T\times k_j}{X^j}\underset{k_j\times 1}{\beta^j} + \underset{T\times 1}{\varepsilon^j}, \quad j=1,\dots,M$$

This is the same setup as the Seemingly Unrelated Regressors model, but now X^j includes a set of variables Y^j that are endogenous...

This is our simultaneous equations system with normalization and exclusion restrictions.

Let's stack the equations, as before:

$$\underset{(TM\times \sum_{j=1}^m k_j)}{\bar{X} = \begin{bmatrix} X^1 & & 0 \\ & X^2 & \\ & & \ddots \\ 0 & & & X^m \end{bmatrix}}, \quad \underset{TM\times 1}{\bar{y} = \begin{bmatrix} y^1 \\ \vdots \\ y^m \end{bmatrix}}, \quad \underset{TM\times 1}{\bar{\varepsilon} = \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^m \end{bmatrix}}, \quad \underset{\sum_{j=1}^m k_j \times 1}{\bar{\beta} = \begin{bmatrix} \beta^1 \\ \vdots \\ \beta^m \end{bmatrix}}.$$

$$\Rightarrow \bar{y} = \bar{X}\bar{\beta} + \bar{\varepsilon}$$

To do system IV, we need an instrumental variable $\underset{TM\times \sum_{j=1}^m k_j}{\bar{W}}$ s.t. $\hat{\bar{\beta}}_{IV} = (\bar{W}'\bar{X})^{-1}(\bar{W}'\bar{y})$

Where does $\bar{W}$ come from? Note that

$$(I_m \otimes Z)'\bar{\varepsilon}/T = \frac{1}{T}\begin{bmatrix} Z' & & 0 \\ & Z' & \\ 0 & \cdots & Z' \end{bmatrix}\begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^m \end{bmatrix} = \begin{bmatrix} \frac{1}{T}Z'\varepsilon^1 \\ \vdots \\ \frac{1}{T}Z'\varepsilon^m \end{bmatrix} \xrightarrow{p} 0$$

under the orthogonality assumption. So,

$\underset{(Tm \times km)}{\underset{m \times m}{I_m} \otimes \underset{T \times k}{Z}}$ is the set of instruments to be used to form $\bar{W}$.

We need $mk \geq \sum_{j=1}^{m} k_j$. This holds if each equation is identified since the order condition for the i^{th} equation is $K \geq q_i$.

$$\underset{Tm \times \sum_{j=1}^m k_j}{\bar{W}} = \underset{Tm \times km}{(I \otimes Z)} \; \underset{km \times \sum_{j=1}^m k_j}{\bar{A}}$$

What is the best $\bar{A}$?

$$\bar{A} = (\hat{\Sigma}^{-1} \otimes I)\cdot(I \otimes (Z'Z)^{-1}Z')\bar{X}$$

where $Var(\bar{\varepsilon}) = (\Sigma \otimes I)$ and $(\hat{\Sigma}^{-1} \otimes I)$ is the inverse of the estimate of $Var(\bar{\varepsilon})$.

$$\bar{W} = (I \otimes Z)\bar{A} = (\hat{\Sigma}^{-1} \otimes Z(Z'Z)^{-1}Z')\bar{X}$$
$$= (\hat{\Sigma}^{-1} \otimes P_Z)\bar{X}$$

(1) $$\hat{\beta}_{3SLS} = [\bar{X}'(\hat{\Sigma}^{-1} \otimes P_Z)\bar{X}]^{-1}[\bar{X}'(\hat{\Sigma}^{-1} \otimes P_Z)\bar{y}]$$

Let's step back a minute and find the 3SLS estimator in a different way, relating it more to SUR.

Consider the IV estimator formed from

$$\bar{W} = \hat{\bar{X}} = \begin{bmatrix} Z(Z'Z)^{-1}Z'X^1 & 0 & \cdots & 0 \\ 0 & Z(Z'Z)^{-1}Z'X^2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & Z(Z'Z)^{-1}Z'X^m \end{bmatrix}$$

$$= \begin{bmatrix} \hat{X}^1 & & 0 \\ & \hat{X}^2 & \\ 0 & & \ddots \;\hat{X}^m \end{bmatrix}$$

The IV estimator $\hat{\bar{\beta}}_{IV} = (\hat{\bar{X}}'\bar{X})^{-1}(\hat{\bar{X}}'\bar{y})$ is nothing more than equation by equation 2SLS. By analogy with the SUR model, we would expect this estimator to be less efficient than a GLS estimator. A natural candidate is

$$\hat{\bar{\beta}}_{3SLS,I} = (\hat{\bar{X}}'(\Sigma^{-1}\otimes I)\bar{X})^{-1}(\hat{\bar{X}}'(\Sigma^{-1}\otimes I)\bar{y})$$

Infeasible since Σ not known.

Plugging in an estimate for Σ gives

$$(2) \quad \hat{\bar{\beta}}_{3SLS} = (\hat{\bar{X}}'(\hat{\Sigma}^{-1}\otimes I)\bar{X})^{-1}(\hat{\bar{X}}'(\hat{\Sigma}^{-1}\otimes I)\bar{y})$$

Noting that $\hat{\bar{X}}'(\hat{\Sigma}^{-1} \otimes I) = ((I \otimes P_Z)\bar{X})'(\hat{\Sigma}^{-1} \otimes I)$
$= \bar{X}'(\hat{\Sigma}^{-1} \otimes P_Z') = \bar{X}'(\hat{\Sigma}^{-1} \otimes P_Z)$, the 3SLS estimator in Eqns (1) and (2) are identical.

The three-stages of computing the 3SLS estimator are:

1. Compute $\hat{\bar{X}}$ by regressing each component of $\bar{X}$ on Z:
$$\hat{\bar{X}} = \begin{bmatrix} Z(Z'Z)^{-1}Z'X^1 & & 0 \\ & \ddots & \\ 0 & & Z(Z'Z)^{-1}Z'X^m \end{bmatrix}$$

2. Compute $\hat{\beta}^j_{2SLS}$ for each equation j by doing 2SLS on each equation.
$$\begin{pmatrix} \hat{\beta}^1_{2SLS} \\ \vdots \\ \hat{\beta}^m_{2SLS} \end{pmatrix} = \hat{\bar{\beta}}_{2SLS} = (\hat{\bar{X}}'\bar{X})^{-1}(\hat{\bar{X}}'y)$$
Define $S_{ij} = \dfrac{\hat{\varepsilon}^{i\prime}\hat{\varepsilon}^j}{T}$ where $\hat{\varepsilon}^i$ are the residuals from equation i, $\hat{\varepsilon}^i = y^i - X^i\hat{\beta}^i_{2SLS}$

3. $$\hat{\Sigma} = S = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1m} \\ S_{21} & S_{22} & & \vdots \\ \vdots & & \ddots & \\ S_{m1} & & \cdots & S_{mm} \end{pmatrix}$$
Compute $$\hat{\bar{\beta}}_{3SLS} = [\bar{X}'(S^{-1} \otimes P_Z)\bar{X}]^{-1}[\bar{X}'(S^{-1} \otimes P_Z)\bar{y}]$$

Optimality of 3SLS

Let's find the variance of the general IV estimator using $\bar{W} = (I \otimes Z)\bar{A}$

$$\sqrt{T}(\hat{\bar{\beta}}_{IV} - \bar{\beta}) = \left[\frac{\bar{A}'(I \otimes Z)'\bar{X}}{T}\right]^{-1}\left[\frac{\bar{A}'(I \otimes Z)'\bar{\varepsilon}}{\sqrt{T}}\right]$$

$$(I \otimes Z)'\bar{\varepsilon}/\sqrt{T} \xrightarrow{d} N\left(0, \operatorname{plim} \frac{(I \otimes Z)'(\Sigma \otimes I)(I \otimes Z)}{T}\right)$$

$$\xrightarrow{d} N\left(0, \operatorname{plim} \Sigma \otimes \left(\frac{Z'Z}{T}\right)\right)$$

$$\xrightarrow{d} N(0, \Sigma \otimes M_Z), \quad M_Z = \operatorname{plim}\left(\frac{Z'Z}{T}\right)$$

$$\frac{\bar{A}'(I \otimes Z)'\bar{X}}{T} = \bar{A}' \begin{bmatrix} \frac{Z'X^1}{T} & & 0 \\ & \frac{Z'X^2}{T} & \\ 0 & & \ddots \quad \frac{Z'X^m}{T} \end{bmatrix}$$

$$\frac{Z'X^j}{T} = \frac{1}{T}Z'[Y^j \; Z^j] = \frac{1}{T}Z'[Z\Pi_j + V_j, \; Z^j]$$

$$\xrightarrow{p} M_Z[\Pi_j \; C_j]$$

where Π_j is the matrix of coefficients given from the reduced form for $Y^j = Z\Pi_j + V_j$ and where C_j is a selection matrix that selects from M_Z the part that represents the plim of $\frac{1}{T}Z'Z^j$.

Optimality of 3SLS (Cont).

Let $D_j = [\pi_j \quad C_j]$, $\quad A = \text{plim}\ \bar{A}$.

Then $\dfrac{\bar{A}'(I \otimes Z)'\bar{X}}{T} \xrightarrow{p} A' \begin{bmatrix} M_Z D_1 & & 0 \\ & M_Z D_2 & \\ 0 & & \ddots & M_Z D_m \end{bmatrix}$

$$= A'(I \otimes M_Z) \begin{bmatrix} D_1 & & 0 \\ & D_2 & \\ 0 & & \ddots & D_m \end{bmatrix}$$

$$= A'(I \otimes M_Z)\bar{D}$$

Putting this together, we have

$$\sqrt{T}(\hat{\bar{\beta}}_{IV} - \bar{\beta}) \xrightarrow{d} (A'(I \otimes M_Z)\bar{D})^{-1} A' N(0, \Sigma \otimes M_Z)$$

so that

$$AVar(\hat{\bar{\beta}}_{IV}) = (A'(I \otimes M_Z)\bar{D})^{-1}[A'(\Sigma \otimes M_Z)A](\bar{D}'(I \otimes M_Z)A)^{-1}$$

For 3SLS,

$$\bar{A} = (\hat{\Sigma}^{-1} \otimes I) \begin{bmatrix} (Z'Z)^{-1}Z'X' & & 0 \\ 0 & \ddots & (Z'Z)^{-1}Z'X^m \end{bmatrix}$$

$$= (\hat{\Sigma}^{-1} \otimes I) \begin{bmatrix} \left(\frac{Z'Z}{T}\right)^{-1}\left(\frac{Z'X'}{T}\right) & & 0 \\ & \ddots & \\ 0 & & \left(\frac{Z'Z}{T}\right)^{-1}\left(\frac{Z'X^m}{T}\right) \end{bmatrix}$$

$$\xrightarrow{p} (\Sigma^{-1} \otimes I) \begin{bmatrix} M_Z^{-1} M_Z D_1 & & 0 \\ & \ddots & \\ 0 & & M_Z^{-1} M_Z D_m \end{bmatrix} = (\Sigma^{-1} \otimes I) \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_m \end{pmatrix}$$

Optimality of 3SLS (Cont)

For 3SLS,

$$\bar{A} \xrightarrow{p} (\Sigma^{-1} \otimes I)\bar{D}$$

Plugging in this plim for $AVar(\hat{\beta}_{IV})$ gives

$$AVar(\hat{\beta}_{3SLS}) = (\bar{D}'(\Sigma^{-1} \otimes Q)\bar{D})^{-1}(\bar{D}'(\Sigma^{-1} \otimes Q)\bar{D})(\bar{D}'(\Sigma^{-1} \otimes Q)\bar{D})^{-1}$$
$$= [\bar{D}'(\Sigma^{-1} \otimes Q)\bar{D}]^{-1}$$

The final step to showing optimality is to show that

$$AVar(\hat{\beta}_{IV}) - AVar(\hat{\beta}_{3SLS}) \text{ is p.s.d.}$$

This is left as an (unnecessary) excercise.

Comparing 2SLS, 3SLS

- When all equations are exactly identified, 2SLS ≡ 3SLS. To see this, note:

$$\hat{\beta}_{3SLS} = [\underbrace{\bar{X}'(\hat{\Sigma}^{-1} \otimes P_Z)\bar{X}}]^{-1}[\bar{X}'(\hat{\Sigma}^{-1} \otimes P_Z)\bar{y}]$$

$$\downarrow$$

$$\bar{X}'(I \otimes Z)(\hat{\Sigma}^{-1} \otimes I)(I \otimes (Z'Z)^{-1})(I \otimes Z')\bar{X}$$

$$= \underbrace{\begin{bmatrix} X^{1\prime}Z & & 0 \\ & X^{2\prime}Z & \\ 0 & & \ddots \\ & & & X^{M\prime}Z \end{bmatrix}}(\hat{\Sigma}^{-1} \otimes I)(I \otimes (Z'Z)^{-1}) \underbrace{\begin{bmatrix} Z'X^1 & & 0 \\ & Z'X^2 & \\ 0 & & \ddots \\ & & & Z'X^M \end{bmatrix}}$$

Each of these blocks is square and invertible.

Comparing 2SLS, 3SLS (Cont)

So, when all equations are exactly identified,

$$\hat{\beta}_{3SLS} = \begin{bmatrix} Z'X^1 & & 0 \\ & \ddots & \\ 0 & & Z'X^m \end{bmatrix}^{-1} (I \otimes (Z'Z))(\hat{\Sigma} \otimes I) \begin{bmatrix} X^{1\prime}Z & & 0 \\ & \ddots & \\ 0 & & X^{m\prime}Z \end{bmatrix}^{-1}$$

$$\cdot \begin{bmatrix} X^{1\prime}Z & & 0 \\ & \ddots & \\ 0 & & X^{m\prime}Z \end{bmatrix} (\hat{\Sigma}^{-1} \otimes I)(I \otimes (Z'Z)^{-1}) \begin{bmatrix} Z'y^1 \\ \vdots \\ Z'y^m \end{bmatrix}$$

$$= \begin{bmatrix} (Z'X^1)^{-1} & & 0 \\ & \ddots & \\ 0 & & (Z'X^m)^{-1} \end{bmatrix} \begin{bmatrix} Z'y^1 \\ \vdots \\ Z'y^m \end{bmatrix}$$

$$= \begin{bmatrix} (Z'X^1)^{-1} Z'y^1 \\ \vdots \\ (Z'X^m)^{-1} Z'y^m \end{bmatrix} = \text{2SLS eqn. by eqn.}$$

Also, as in the SUR model, you can see that misspecification in some equation effects in general the consistency in all equations. To see this, expand the term

$$\bar{X}'(\hat{\Sigma}^{-1} \otimes P_Z)\bar{\varepsilon}$$

and argue as we did for the SUR estimator.