Assumptions

Ass 4.1: Linearity

$$y_{im} = \mathbf{z}_{im}^{'}\delta_m + \varepsilon_{im}$$

$$\begin{bmatrix} \mathbf{y}_1 \\ n \times 1 \\ \vdots \\ \mathbf{y}_M \\ n \times 1 \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 & & \\ n \times L_1 & & \\ & \ddots & \\ & & \mathbf{Z}_M \\ & & n \times L_M \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_1 \\ L_1 \times 1 \\ \vdots \\ \boldsymbol{\delta}_M \\ L_M \times 1 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ n \times 1 \\ \vdots \\ \boldsymbol{\varepsilon}_M \\ n \times 1 \end{bmatrix}$$

Ass 4.2: Jointly Ergodic Stationarity

$$\{\mathbf{w}_i\} = \{y_{i1}, \dots, y_{iM}, \mathbf{z}_{i1}, \dots, \mathbf{z}_{iM}, \mathbf{x}_{i1}, \dots, \mathbf{x}_{iM}\}$$

Ass 4.3: Orthogonality

$$E(\mathbf{x}_{im} \cdot \varepsilon_{im}) = \mathbf{0}$$

We don't assume cross equation orthogonalities.

Ass 4.4: Rank condition
$E(\mathbf{x}_{im}\mathbf{z}_{im}')$ is full column rank.

Ass 4.5: g_i is a m.d.s with finite second moments

$$g_i \sim m.d.s$$

$E(g_i g_i')$ is nonsingular.

Ass 4.6: finite fourth moments
$E[(x_{imk}z_{ihl}')^2]$ exists and finite.

Ass 4.7: conditional homoskedasticity

MEGMM

$$\bar{\mathbf{g}} \equiv \frac{1}{n}\sum_{i=1}^{n}\mathbf{g}_i = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i1}\cdot\varepsilon_{i1} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{iM}\cdot\varepsilon_{iM} \end{bmatrix} = \mathbf{g}_n(\boldsymbol{\delta})$$

population moment:

$$\begin{aligned}
E[\mathbf{g_i}] &\equiv \begin{bmatrix} E[\mathbf{x}_{i1}\cdot(y_{i1}-\mathbf{z}_{i1}^{'}\tilde{\boldsymbol{\delta}}_1)] \\ \vdots \\ E[\mathbf{x}_{iM}\cdot(y_{iM}-\mathbf{z}_{iM}^{'}\tilde{\boldsymbol{\delta}}_M)] \end{bmatrix} \\
&= \begin{bmatrix} E(\mathbf{x}_{i1}\cdot y_{i1}) \\ \vdots \\ E(\mathbf{x}_{iM}\cdot y_{iM}) \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_{i1}\mathbf{z}_{i1}^{'})\tilde{\boldsymbol{\delta}}_1 \\ \vdots \\ E(\mathbf{x}_{iM}\mathbf{z}_{iM}^{'})\tilde{\boldsymbol{\delta}}_M \end{bmatrix} \\
&= \begin{bmatrix} E(\mathbf{x}_{i1}\cdot y_{i1}) \\ \vdots \\ E(\mathbf{x}_{iM}\cdot y_{iM}) \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_{i1}\mathbf{z}_{i1}^{'}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & E(\mathbf{x}_{iM}\mathbf{z}_{iM}^{'}) \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\delta}}_1 \\ \vdots \\ \tilde{\boldsymbol{\delta}}_M \end{bmatrix} \\
&\equiv \boldsymbol{\sigma}_{xy} - \boldsymbol{\Sigma}_{xz}\tilde{\boldsymbol{\delta}}
\end{aligned}$$

sample moment:

$$\begin{aligned}
\mathbf{g}_n(\tilde{\boldsymbol{\delta}}) &= \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i1}\cdot(y_{i1}-\mathbf{z}_{i1}^{'}\tilde{\boldsymbol{\delta}}_1) \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{iM}\cdot(y_{iM}-\mathbf{z}_{iM}^{'}\tilde{\boldsymbol{\delta}}_M) \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i1}\cdot y_{i1} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{iM}\cdot y_{iM} \end{bmatrix} - \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i1}\mathbf{z}_{i1}^{'}\tilde{\boldsymbol{\delta}}_1 \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{iM}\mathbf{z}_{iM}^{'}\tilde{\boldsymbol{\delta}}_M \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i1}\cdot y_{i1} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{iM}\cdot y_{iM} \end{bmatrix} - \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i1}\mathbf{z}_{i1}^{'} & & \\ & \ddots & \\ & & \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{iM}\mathbf{z}_{iM}^{'} \end{bmatrix} \begin{bmatrix} \tilde{\delta}_1 \\ \vdots \\ \tilde{\delta}_M \end{bmatrix} \\
&\equiv \mathbf{s}_{xy} - \mathbf{S}_{xz}\tilde{\boldsymbol{\delta}}
\end{aligned}$$

$$\mathbf{S} = E(\mathbf{g}_i\mathbf{g}_i^{'}) = \begin{bmatrix} E(\varepsilon_{i1}\varepsilon_{i1}\mathbf{x}_{i1}\mathbf{x}_{i1}^{'}) & \dots & E(\varepsilon_{i1}\varepsilon_{iM}\mathbf{x}_{i1}\mathbf{x}_{iM}^{'}) \\ \vdots & & \vdots \\ E(\varepsilon_{iM}\varepsilon_{i1}\mathbf{x}_{iM}\mathbf{x}_{i1}^{'}) & \dots & E(\varepsilon_{iM}\varepsilon_{iM}\mathbf{x}_{iM}\mathbf{x}_{iM}^{'}) \end{bmatrix}$$

$$
\widehat{\mathbf{W}} = \begin{bmatrix} \widehat{\mathbf{W}}_{11} & \widehat{\mathbf{W}}_{12} & \dots & \widehat{\mathbf{W}}_{1M} \\ \widehat{\mathbf{W}}'_{12} & \widehat{\mathbf{W}}_{22} & \dots & \widehat{\mathbf{W}}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{W}}'_{1M} & \widehat{\mathbf{W}}'_{2M} & \dots & \widehat{\mathbf{W}}_{MM} \end{bmatrix}
$$

MEGMM estimator:

$$
\begin{aligned}
\tilde{\delta}(\hat{W}) &= \arg\min_{\tilde{\delta}} J(\tilde{\delta}, \hat{W}) = \arg\min_{\tilde{\delta}} n g_n(\tilde{\delta})' \hat{W} g_n(\tilde{\delta}) \\
&= \arg\min_{\tilde{\delta}} n (s_{xy} - S_{xz}\tilde{\delta})' \hat{W} (s_{xy} - S_{xz}\tilde{\delta})
\end{aligned}
$$

FOC:

$$
\tilde{\delta}(\hat{W}) = (S'_{xz}\hat{W}S_{xz})^{-1} S'_{xz}\hat{W}s_{xy} =
$$

$$
\begin{pmatrix} \mathbf{S}'_{x_1z_1}\hat{\mathbf{W}}_{11}\mathbf{S}_{x_1z_1} & \mathbf{S}'_{x_1z_1}\hat{\mathbf{W}}_{12}\mathbf{S}_{x_2z_2} & \dots & \mathbf{S}'_{x_1z_1}\hat{\mathbf{W}}_{1M}\mathbf{S}_{x_Mz_M} \\ \mathbf{S}'_{x_2z_2}\hat{\mathbf{W}}'_{12}\mathbf{S}_{x_1z_1} & \mathbf{S}'_{x_2z_2}\hat{\mathbf{W}}_{22}\mathbf{S}_{x_2z_2} & \dots & \mathbf{S}'_{x_2z_2}\hat{\mathbf{W}}_{2M}\mathbf{S}_{x_Mz_M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}'_{x_Mz_M}\hat{\mathbf{W}}'_{1M}\mathbf{S}_{x_1z_1} & \mathbf{S}'_{x_Mz_M}\hat{\mathbf{W}}'_{2M}\mathbf{S}_{x_2z_2} & \dots & \mathbf{S}'_{x_Mz_M}\hat{\mathbf{W}}_{MM}\mathbf{S}_{x_Mz_M} \end{pmatrix}^{-1} \times \begin{pmatrix} \mathbf{S}'_{x_1z_1}\hat{\mathbf{W}}_{1M}\mathbf{s}_{x_My_M} \\ \mathbf{S}'_{x_2z_2}\hat{\mathbf{W}}_{2M}\mathbf{s}_{x_My_M} \\ \vdots \\ \mathbf{S}'_{x_Mz_M}\hat{\mathbf{W}}_{MM}\mathbf{s}_{x_My_M} \end{pmatrix}
$$

SEGMM versus MEGMM

SEGMM is a special case of MEGMM when:

$$
\hat{\mathbf{W}} = \text{diag}(\hat{\mathbf{W}}_{11}, \dots, \hat{\mathbf{W}}_{MM})
$$

Efficient MEGMM is equivalent to Efficient SEGMM when:

1. Each equation is just identified.
2. At least one equation is over-identified but S is block diagonal.

$$
E[g_{im}g'_{ih}] = E[\varepsilon_{im}\varepsilon_{ih}\, x_{im}x'_{ih}] = 0 \quad \text{for all } m \neq h
$$

$$
\hat{\delta}(\widehat{\mathbf{W}}) = \begin{bmatrix} \hat{\delta}_1(\widehat{\mathbf{W}}) \\ \vdots \\ \hat{\delta}_M(\widehat{\mathbf{W}}) \end{bmatrix} = \left(\begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{i1}\mathbf{x}_{i1}' & & \\ & \ddots & \\ & & \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{iM}\mathbf{x}_{iM}' \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{W}}_{11} \cdots \widehat{\mathbf{W}}_{1M} \\ \vdots \quad\quad \vdots \\ \widehat{\mathbf{W}}_{M1} \cdots \widehat{\mathbf{W}}_{MM} \end{bmatrix} \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{i1}\mathbf{z}_{i1}' & & \\ & \ddots & \\ & & \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{iM}\mathbf{z}_{iM}' \end{bmatrix} \right)^{-1}
$$

$$
\begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{i1}\mathbf{x}_{i1}' & & \\ & \ddots & \\ & & \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{iM}\mathbf{x}_{iM}' \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{W}}_{11} \cdots \widehat{\mathbf{W}}_{1M} \\ \vdots \quad\quad \vdots \\ \widehat{\mathbf{W}}_{M1} \cdots \widehat{\mathbf{W}}_{MM} \end{bmatrix} \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{i1} \cdot y_{i1} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{iM} \cdot y_{iM} \end{bmatrix}
$$

$$
= \begin{bmatrix} \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{i1}\mathbf{x}_{i1}'\right)\widehat{\mathbf{W}}_{11}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{i1}\mathbf{z}_{i1}'\right) & \cdots & \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{i1}\mathbf{x}_{i1}'\right)\widehat{\mathbf{W}}_{1M}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{iM}\mathbf{z}_{iM}'\right) \\ \vdots & & \vdots \\ \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{iM}\mathbf{x}_{iM}'\right)\widehat{\mathbf{W}}_{M1}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{i1}\mathbf{z}_{i1}'\right) & \cdots & \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{iM}\mathbf{x}_{iM}'\right)\widehat{\mathbf{W}}_{MM}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{iM}\mathbf{z}_{iM}'\right) \end{bmatrix}^{-1}
$$

$$
\begin{bmatrix} \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{i1}\mathbf{x}_{i1}'\right)\widehat{\mathbf{W}}_{11}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{i1}y_{i1}'\right) + \cdots + \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{i1}\mathbf{x}_{i1}'\right)\widehat{\mathbf{W}}_{1M}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{iM}y_{iM}'\right) \\ \vdots \\ \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{iM}\mathbf{x}_{iM}'\right)\widehat{\mathbf{W}}_{M1}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{i1}\, y_{i1}\right) + \cdots + \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_{iM}\mathbf{x}_{iM}'\right)\widehat{\mathbf{W}}_{MM}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{iM}\, y_{iM}\right) \end{bmatrix}. \tag{4.2.6}
$$

To go from the second-to-last to the last line, use formulas (A.6) and (A.7) of the appendix on partitioned matrices. If you find the operation too much, just set $M = 2$.

Table 4.1: Multiple-Equation GMM in the Single-Equation Format

Sample Analogue of Orthogonality Conditions:	$\mathbf{g}_n(\tilde{\delta}) = \mathbf{s}_{\mathbf{xy}} - \mathbf{S}_{\mathbf{xz}}\tilde{\delta} = \mathbf{0}$
GMM Estimator:	$\hat{\delta}(\widehat{\mathbf{W}}) = (\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\mathbf{s}_{\mathbf{xy}}$
Its Sampling Error:	$\hat{\delta}(\widehat{\mathbf{W}}) - \delta = (\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\bar{\mathbf{g}}$
Asymptotic Variance of Optimal GMM:	$\text{Avar}(\hat{\delta}(\widehat{\mathbf{S}}^{-1})) = (\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}})^{-1}$
Its Estimator:	$\widehat{\text{Avar}(\hat{\delta}(\widehat{\mathbf{S}}^{-1}))} = (\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})^{-1}$
J Statistic:	$J(\hat{\delta}(\widehat{\mathbf{S}}^{-1}), \widehat{\mathbf{S}}^{-1}) \equiv n \cdot \mathbf{g}_n(\hat{\delta}(\widehat{\mathbf{S}}^{-1}))'\widehat{\mathbf{S}}^{-1}\mathbf{g}_n(\hat{\delta}(\widehat{\mathbf{S}}^{-1}))$

	Single-Equation GMM applied to the equation in question	Multiple-Equation GMM
$\mathbf{g}_i$	$\mathbf{x}_i \cdot \varepsilon_i$	(4.1.4)
δ	δ	(4.1.6)
$\mathbf{s}_{\mathbf{xy}}$	$\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \cdot y_i$	(4.2.2)
$\mathbf{S}_{\mathbf{xz}}$	$\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i\mathbf{z}'_i$	(4.2.2)
Size of $\mathbf{W}$	$K \times K$	$\sum_m K_m \times \sum_m K_m$
$\boldsymbol{\Sigma}_{\mathbf{xz}}$	$\text{E}(\mathbf{x}_i\mathbf{z}'_i)$	(4.1.9)
$\mathbf{S}$ ($\equiv \text{Avar}(\bar{\mathbf{g}})$)	$\text{E}(\varepsilon_i^2\mathbf{x}_i\mathbf{x}'_i)$	(4.1.11)
$\widehat{\mathbf{S}}$	$\frac{1}{n}\sum_{i=1}^{n} \hat{\varepsilon}_i^2\mathbf{x}_i\mathbf{x}'_i$	(4.3.2)
Estimator consistent under which assumptions?	3.1–3.4	4.1–4.4
Estimator asymptotic normal under which assumptions?	3.1–3.5	4.1–4.5
$\widehat{\mathbf{S}} \to_{\text{p}} \mathbf{S}$ under which assumptions?	3.1, 3.2, 3.6, $\text{E}(\mathbf{g}_i\mathbf{g}'_i)$ finite	4.1, 4.2, 4.6, $\text{E}(\mathbf{g}_i\mathbf{g}'_i)$ finite
d.f. of J	$K - L$	$\sum_m (K_m - L_m)$

Special Cases of MEGMM: FIVE, 3SLS and SUR

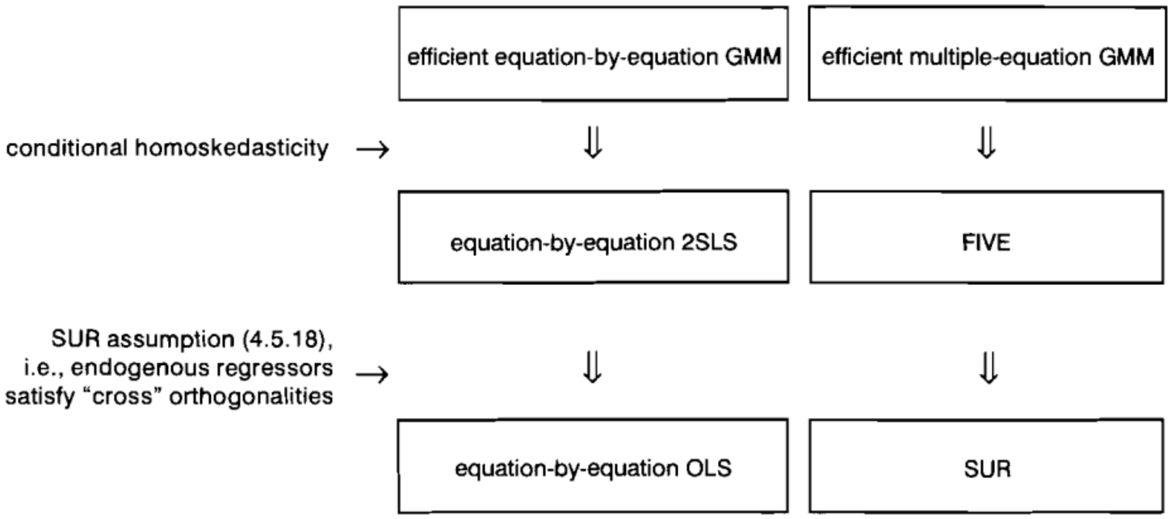

Figure 4.1: OLS and GMM

Full-Information Instrumental Variables Efficient (FIVE)

Under Ass 4.1-4.5 and 4.7,

$$E(\varepsilon_{im}\varepsilon_{ih}|\mathbf{x}_{im},\mathbf{x}_{ih}) = \sigma_{mh}$$

$$\mathbf{S} = E[\mathbf{x}_{im}\mathbf{x}_{ih}^{'}\varepsilon_{im}\varepsilon_{ih}] = \sigma_{mh}E[\mathbf{x}_{im}\mathbf{x}_{ih}^{'}] = \begin{bmatrix} \sigma_{11}E(\mathbf{x}_{i1}\mathbf{x}_{i1}^{'}) & \dots & \sigma_{1M}E(\mathbf{x}_{i1}\mathbf{x}_{iM}^{'}) \\ \vdots & \ddots & \vdots \\ \sigma_{M1}E(\mathbf{x}_{iM}\mathbf{x}_{i1}^{'}) & \dots & \sigma_{MM}E(\mathbf{x}_{iM}\mathbf{x}_{iM}^{'}) \end{bmatrix}$$

$\hat{\sigma}_{mh}^2 \xrightarrow{p} \sigma_{mh}^2 (E(z_{im}z_{ih}')$ exists and finite), $\frac{1}{n}\sum x_{im}x_{ih}' \xrightarrow{p} E(x_{im}x_{ih}')$, $\hat{S} \xrightarrow{p} S$

Large-Sample properties:

1. $\tilde{\delta}_{FIVE} = \tilde{\delta}(\hat{S}^{-1})$ is consistent, asymptotically normal, and efficient with $Avar(\tilde{\delta}(\hat{S}^{-1}))$.

$$Avar(\tilde{\delta}_{FIVE}) = (\Sigma_{xz}'S^{-1}\Sigma_{xz})^{-1}$$

2. $\widehat{Avar}(\tilde{\delta}_{FIVE})$ is consistent for $Avar(\tilde{\delta}_{FIVE})$:

$$\widehat{Avar}(\tilde{\delta}_{FIVE}) = (S_{xz}'\hat{S}^{-1}S_{xz})^{-1}$$

3. Sargan's Statistic:

$$J(\tilde{\delta}_{FIVE},\hat{S}^{-1}) \equiv n\cdot g_n(\tilde{\delta}_{FIVE})'\hat{S}^{-1}g_n(\tilde{\delta}_{FIVE}) \xrightarrow{d} \chi^2(\sum_m(K_m - L_m))$$

Three-Stage Least Squares (3SLS)

Under Ass 4.1-4.5 and 4.7 + same instruments for all equations:
$\mathbf{x}_{i1} = \mathbf{x}_{i2} = \ldots = \mathbf{x}_{im} = \mathbf{x}_i$

moment conditions:

$$\mathbf{g}_i = \begin{bmatrix} \mathbf{x}_i \varepsilon_{i1} \\ \vdots \\ \mathbf{x}_i \varepsilon_{iM} \end{bmatrix} = \boldsymbol{\varepsilon}_i \otimes \mathbf{x}_i, \quad \text{with} \quad \boldsymbol{\varepsilon}_i \equiv \begin{bmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iM} \end{bmatrix}$$

Efficient Weight matrix:

$$\mathbf{S}_{3SLS} = \begin{bmatrix} \sigma_{11} E[\mathbf{x}_i \mathbf{x}_i^{'}] & \sigma_{12} E[\mathbf{x}_i \mathbf{x}_i^{'}] & \ldots & \sigma_{1M} E[\mathbf{x}_i \mathbf{x}_i^{'}] \\ \sigma_{12} E[\mathbf{x}_i \mathbf{x}_i^{'}] & \sigma_{22} E[\mathbf{x}_i \mathbf{x}_i^{'}] & \ldots & \sigma_{2M} E[\mathbf{x}_i \mathbf{x}_i^{'}] \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1M} E[\mathbf{x}_i \mathbf{x}_i^{'}] & \sigma_{2M} E[\mathbf{x}_i \mathbf{x}_i^{'}] & \ldots & \sigma_{MM} E[\mathbf{x}_i \mathbf{x}_i^{'}] \end{bmatrix} = \mathbf{\Sigma} \otimes E[\mathbf{x}_i \mathbf{x}_i^{'}] \quad ,$$

$$\text{with} \quad \mathbf{\Sigma} = E[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^{'}] = \begin{bmatrix} \boldsymbol{\sigma}_{11} & \boldsymbol{\sigma}_{12} & \ldots & \boldsymbol{\sigma}_{1M} \\ \boldsymbol{\sigma}_{12} & \boldsymbol{\sigma}_{22} & \ldots & \boldsymbol{\sigma}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\sigma}_{1M} & \boldsymbol{\sigma}_{2M} & \ldots & \boldsymbol{\sigma}_{MM} \end{bmatrix}$$

$$\widehat{\mathbf{\Sigma}} \equiv \begin{bmatrix} \hat{\boldsymbol{\sigma}}_{11} & \ldots & \hat{\boldsymbol{\sigma}}_{1M} \\ \vdots & & \vdots \\ \hat{\boldsymbol{\sigma}}_{M1} & \ldots & \hat{\boldsymbol{\sigma}}_{MM} \end{bmatrix} = \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\varepsilon}}_i \hat{\boldsymbol{\varepsilon}}_i^{'}, \quad S_{xx} = \frac{1}{n} \mathbf{X}'\mathbf{X}$$

Large-Sample properties:

1. $\tilde{\delta}_{3SLS}$ is consistent, asymptotically normal, and efficient with $Avar(\tilde{\delta}_{3SLS})$.
2. $\widehat{Avar}(\tilde{\delta}_{3SLS})$ is consistent for $Avar(\tilde{\delta}_{3SLS})$.
3. Sargan's Statistic:

$$J(\tilde{\delta}_{\text{3SLS}}, \hat{S}^{-1}) \equiv n \cdot g_n(\tilde{\delta}_{\text{3SLS}})' \hat{S}^{-1} g_n(\tilde{\delta}_{3SLS}) \xrightarrow{d} \chi^2(MK - \sum_m L_m)$$

Seemingly Unrelated Regression (SUR)

Under Ass 4.1-4.5 and 4.7 + same instruments for all equations + the predetermined regressors satisfy "cross" orthogonalities: $E[\mathbf{z}_{im} \varepsilon_{ih}] = 0$ (without endogeneity)

Efficient Weight Matrix:

$$
\mathbf{S}_{SUR} = \begin{bmatrix} \sigma_{11}E[\mathbf{z}_i\mathbf{z}_i^{'}] & \sigma_{12}E[\mathbf{z}_i\mathbf{z}_i^{'}] & \dots & \sigma_{1M}E[\mathbf{z}_i\mathbf{z}_i^{'}] \\ \sigma_{12}E[\mathbf{z}_i\mathbf{z}_i^{'}] & \sigma_{22}E[\mathbf{z}_i\mathbf{z}_i^{'}] & \dots & \sigma_{2M}E[\mathbf{z}_i\mathbf{z}_i^{'}] \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1M}E[\mathbf{z}_i\mathbf{z}_i^{'}] & \sigma_{2M}E[\mathbf{z}_i\mathbf{z}_i^{'}] & \dots & \sigma_{MM}E[\mathbf{z}_i\mathbf{z}_i^{'}] \end{bmatrix} = \mathbf{\Sigma} \otimes E[\mathbf{z}_i\mathbf{z}_i^{'}]
$$

Large-Sample properties:

1. $\tilde{\delta}_{SUR}$ is consistent, asymptotically normal, and efficient with $Avar(\tilde{\delta}_{SUR})$.
2. $\widehat{Avar}(\tilde{\delta}_{SUR})$ is consistent for $Avar(\tilde{\delta}_{SUR})$.
3. Sargan's Statistic:

$$
J(\tilde{\delta}_{SUR}, \hat{S}^{-1}) \equiv n \cdot g_n(\tilde{\delta}_{SUR})' \hat{S}^{-1} g_n(\tilde{\delta}_{SUR}) \xrightarrow{d} \chi^2(MK - \sum_m L_m)
$$

SUR versus OLS

the same as SEGMM versus MEGMM:

1. Each equation is just identified.
2. At least one equation is over-identified. Need assumption that

$$
\sigma_{mh}E(x_i x_i') = 0 \text{ for all m} \neq \text{h}
$$

Table 4.2: Relationship between Multiple-Equation Estimators

	Multiple-equation GMM	FIVE	3SLS	SUR	Multivariate regression
The model	Assumptions 4.1–4.6	Assumptions 4.1–4.5, Assumption 4.7 $\mathrm{E}(\mathbf{z}_{im}\mathbf{z}'_{ih})$ finite	Assumptions 4.1–4.5, Assumption 4.7 $\mathrm{E}(\mathbf{z}_{im}\mathbf{z}'_{ih})$ finite $\mathbf{x}_{im} = \mathbf{x}_i$ for all m	Assumptions 4.1–4.5, Assumption 4.7 $\mathbf{x}_{im} = \mathbf{x}_i$ for all m $\mathbf{x}_i =$ union of $\mathbf{z}_{i1}, \dots, \mathbf{z}_{iM}$	Assumptions 4.1–4.5, Assumption 4.7 $\mathbf{x}_{im} = \mathbf{x}_i$ for all m $\mathbf{z}_{im} = \mathbf{x}_i$ for all m
$\mathbf{S}$ ($\equiv \mathrm{Avar}(\bar{\mathbf{g}})$)	(4.1.11)	(4.5.2)	$\boldsymbol{\Sigma} \otimes \mathrm{E}(\mathbf{x}_i\mathbf{x}'_i)$	$\boldsymbol{\Sigma} \otimes \mathrm{E}(\mathbf{x}_i\mathbf{x}'_i)$	irrelevant
$\widehat{\mathbf{S}}$	(4.3.2)	(4.5.3)	$\widehat{\boldsymbol{\Sigma}} \otimes (n^{-1}\Sigma_i\mathbf{x}_i\mathbf{x}'_i)$ $\widehat{\boldsymbol{\Sigma}}$ from 2SLS residuals	$\widehat{\boldsymbol{\Sigma}} \otimes (n^{-1}\Sigma_i\mathbf{x}_i\mathbf{x}'_i)$ $\widehat{\boldsymbol{\Sigma}}$ from OLS residuals	irrelevant
$\hat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1})$	no simplification	no simplification	(4.5.12) with (4.5.13), (4.5.14)	(4.5.12) with (4.5.13′), (4.5.14′)	equation-by-equation OLS
$\mathrm{Avar}(\hat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1}))$	(4.3.3)	(4.3.3)	(4.5.15) with (4.5.16)	(4.5.15) with (4.5.16′)	OLS formula
$\widehat{\mathrm{Avar}}(\hat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1}))$	(4.3.4)	(4.3.4)	(4.5.17) with (4.5.13)	(4.5.17) with (4.5.13′)	OLS formula