Multiple Equation GMM – Analytical Exercises*

Ryan Benschop

1. (Data matrix representation of 3SLS) Consider the 3SLS model. Let $M \in \mathbb{N}$. For $m \in [M]$, let $\{(y_{im}, x_{im})\}_{i \in \mathbb{N}}$ be a stochastic process, where y_{im} is $\mathbb{R}$-valued and x_{im} is $\mathbb{R}^{K_m}$-valued. Similarly, let $\{z_{im}\}_{i \in \mathbb{N}}$ be an $\mathbb{R}^{L_m}$-valued stochastic process for $m \in [M]$. The model assumptions are:

- (Linearity): for each $m \in [M]$, there is $\beta_m \in \mathbb{R}^{K_m}$ and $\{\varepsilon_{im}\}_{i \in \mathbb{N}}$ such that for each $i \in \mathbb{N}$, $y_{im} = x'_{im}\beta_m + \varepsilon_{im}$.
- (Ergodic stationarity): let w_i denote the unique nonconstant elements of $(y_{i1}, \ldots, y_{iM}, x_{i1}, \ldots, x_{iM}, z_{i1}, \ldots, z_{iM})$. Then, $\{w_i\}_{i \in \mathbb{N}}$ is a jointly stationary and ergodic process.
- (Contemporaneous orthogonality): for each $m \in [M]$, $\mathbb{E}[z_{im}\varepsilon_{im}] = 0$.
- (Identification): for each $m \in [M]$, $\mathbb{E}[z'_{im}x_{im}]$ has full column rank.

Suppose that the set of instruments is the same across equations. That is, for $m \in [M]$, $z_{im} = z_i$ for each $i \in \mathbb{N}$. Let $L = L_m$. Define the following matrices, for each $n \in \mathbb{N}$:

$$\underbrace{Z}_{n \times L} = \begin{bmatrix} z'_1 \\ \vdots \\ z'_n \end{bmatrix} \qquad \underbrace{X_m}_{n \times L_m} = \begin{bmatrix} x'_{1m} \\ \vdots \\ x'_{nm} \end{bmatrix} \qquad \underbrace{X}_{Mn \times \sum_m K_m} = \begin{bmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_M \end{bmatrix}$$

$$\underbrace{\beta}_{\sum_m K_m \times 1} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_M \end{bmatrix} \qquad \underbrace{y_m}_{n \times 1} = \begin{bmatrix} y_{1m} \\ \vdots \\ y_{nm} \end{bmatrix} \qquad \underbrace{y}_{Mn \times 1} = \begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix}$$

$$\underbrace{\varepsilon_m}_{n \times 1} = \begin{bmatrix} \varepsilon_{1m} \\ \vdots \\ \varepsilon_{nm} \end{bmatrix} \qquad \underbrace{\varepsilon}_{Mn \times 1} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_M \end{bmatrix}$$

Claim: The linearity assumption can be written as $y = X\beta + \varepsilon$.

---

*Note, the notation used herein switches the meaning of x_i and z_i as used in the textbook. That is, x_i corresponds with the regressors, and z_i corresponds with the instruments. All notation that corresponds specifically with the regressors and instruments are switched accordingly. Similarly, the notation β is used in place of δ.

Proof. Writing out the individual equations of the linearity assumption (across i), we have, for $m \in [M]$:

$$\begin{aligned} y_{1m} &= x_{1m}'\beta_m + \varepsilon_{1m} \\ &\vdots \\ y_{nm} &= x_{nm}'\beta_m + \varepsilon_{nm} \end{aligned}$$

Hence, we have, for $m \in [M]$:

$$\begin{bmatrix} y_{1m} \\ \vdots \\ y_{nm} \end{bmatrix} = \begin{bmatrix} x_{1m}' \\ \vdots \\ x_{nm}' \end{bmatrix} \beta_m + \begin{bmatrix} \varepsilon_{1m} \\ \vdots \\ \varepsilon_{nm} \end{bmatrix}$$

Writing the implied equations across $m \in [M]$, we have:

$$\begin{aligned} y_1 &= X_1\beta_1 + \varepsilon_1 \\ &\vdots \\ y_M &= X_M\beta_M + \varepsilon_M \end{aligned}$$

In matrix notation, this is:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_M \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_M \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_M \end{bmatrix}$$

Hence, we have:

$$y = X\beta + \varepsilon$$

□

Claim: Define $\hat{S}$ by:

$$\hat{S} = \hat{\Sigma} \otimes \left(n^{-1} \sum_{i=1}^{n} z_i z_i' \right)$$

where $\hat{\Sigma} = n^{-1} \sum_{i=1}^{n} \hat{\varepsilon}_i \hat{\varepsilon}_i'$, where the residual $\hat{\varepsilon}_i = \begin{bmatrix} \hat{\varepsilon}_{i1} & \cdots & \hat{\varepsilon}_{iM} \end{bmatrix}'$ is formed from the residuals of the M equations using the corresponding 2SLS estimator $\hat{\beta}_{2SLS,m}$. Then:

$$\hat{S} = \hat{\Sigma} \otimes n^{-1} Z'Z$$

Proof. Clearly, we only need to show $Z'Z = \sum_{i=1}^n z_i z_i'$. Write:

$$Z'Z = \begin{bmatrix} z_1 & \cdots & z_n \end{bmatrix} \begin{bmatrix} z_1' \\ \vdots \\ z_n' \end{bmatrix} = \sum_{i=1}^n z_i z_i'$$

□

Claim:

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

where $P = Z(Z'Z)^{-1}Z'$.

Proof. First, let's clarify the definition of the 3SLS estimator. The 3SLS estimator is defined in the text as the usual multiple equation GMM estimator, in the case where the set of instruments is the same across equations, with the weighting matrix $\hat{W} = \hat{S}^{-1}$, where $\hat{S}$ is defined above. That is:

$$\hat{\beta}_{3SLS} = \left(S_{zx}'\hat{S}^{-1}S_{zx}\right)^{-1} S_{zx}'\hat{S}^{-1}s_{zy}$$

where:

$$S_{zx} = \begin{bmatrix} n^{-1}\sum_{i=1}^n z_{i1}x_{i1}' & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & n^{-1}\sum_{i=1}^n z_{iM}x_{iM}' \end{bmatrix} \qquad s_{zy} = \begin{bmatrix} n^{-1}\sum_{i=1}^n z_{i1}y_{i1} \\ \vdots \\ n^{-1}\sum_{i=1}^n z_{iM}y_{iM} \end{bmatrix}$$

Now, let's work to rewrite this a bit. First, note that:

$$S_{zx} = n^{-1}\begin{bmatrix} Z'X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z'X_M \end{bmatrix} = n^{-1}(I \otimes Z)'X$$

$$s_{zy} = n^{-1}\begin{bmatrix} Z'y_1 \\ \vdots \\ Z'y_M \end{bmatrix} = n^{-1}(I \otimes Z)'y$$

Substituting these into the expression for $\hat{\beta}_{3SLS}$, we have:

$$\begin{aligned}
\hat{\beta}_{3SLS} &= \left[n^{-1}X'(I \otimes Z)\hat{S}^{-1}n^{-1}(I \otimes Z)'X\right]^{-1} n^{-1}X'(I \otimes Z)\hat{S}^{-1}n^{-1}(I \otimes Z)'y \\
&= \left[X'(I \otimes Z)\hat{S}^{-1}(I \otimes Z)'X\right]^{-1} X'(I \otimes Z)\hat{S}^{-1}(I \otimes Z)'y \\
&= \left[X'(I \otimes Z)(Z'Z)^{-1} \otimes \hat{\Sigma}^{-1}(I \otimes Z)'X\right]^{-1} X'(I \otimes Z)(Z'Z)^{-1} \otimes \hat{\Sigma}^{-1}(I \otimes Z)'y
\end{aligned}$$

Thus, we've shown that $\hat{\beta}_{3SLS}$ is the GMM estimator with weighting matrix $(I \otimes Z)(Z'Z)^{-1} \otimes \hat{\Sigma}^{-1}(I \otimes Z)'$. It thus remains to show that this weighting matrix is the same as $\hat{\Sigma}^{-1} \otimes P$. This follows easily from the mixed-product property of the kronecker product.[1]

$$\begin{aligned}
(I \otimes Z)(Z'Z)^{-1} \otimes \hat{\Sigma}^{-1}(I \otimes Z)' &= (I \otimes Z)(Z'Z)^{-1} \otimes \hat{\Sigma}^{-1}(I \otimes Z') \\
&= (I \otimes Z)\hat{\Sigma}^{-1} \otimes (Z'Z)^{-1} Z' \\
&= \hat{\Sigma}^{-1} \otimes Z(Z'Z)^{-1} Z' \\
&= \hat{\Sigma}^{-1} \otimes P
\end{aligned}$$

□

Claim: Define $Asy.\widehat{\mathbb{V}[\hat{\beta}_{3SLS}]}$ by:

$$Asy.\widehat{\mathbb{V}[\hat{\beta}_{3SLS}]} = \begin{bmatrix} \hat{\sigma}^{11}\hat{A}_{11} & \cdots & \hat{\sigma}^{1M}\hat{A}_{1M} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}^{M1}\hat{A}_{M1} & \cdots & \hat{\sigma}^{MM}\hat{A}_{MM} \end{bmatrix}^{-1}$$

where:

$$\hat{A}_{jk} = \left(n^{-1} \sum_{i=1}^{n} x_{ij} z_i' \right) \left(n^{-1} \sum_{i=1}^{n} z_i z_i' \right)^{-1} \left(n^{-1} \sum_{i=1}^{n} z_i x_{ik}' \right)$$

and $\hat{\sigma}^{jk}$ is the (j,k)th element of $\hat{\Sigma}^{-1}$. Then:

$$Asy.\widehat{\mathbb{V}[\hat{\beta}_{3SLS}]} = n[X'(\hat{\Sigma}^{-1} \otimes P)X]^{-1}$$

Proof. Notice, for $j,k \in [M]$:

$$\begin{aligned}
\hat{A}_{jk} &= n^{-1} X_j' Z \left(n^{-1} Z'Z \right)^{-1} n^{-1} Z' X_k \\
&= n^{-1} X_j' Z (Z'Z)^{-1} Z' X_k
\end{aligned}$$

Writing out the full expression for $\text{Asy.}\widehat{\mathbb{V}[\hat{\beta}_{3SLS}]}$, we have:

$$\text{Asy.}\widehat{\mathbb{V}[\hat{\beta}_{3SLS}]} = n \begin{bmatrix} \hat{\sigma}^{11} X_1' Z(Z'Z)^{-1} Z' X_1 & \cdots & \hat{\sigma}^{1M} X_1' Z(Z'Z)^{-1} Z' X_M \\ \vdots & \ddots & \vdots \\ \hat{\sigma}^{M1} X_M' Z(Z'Z)^{-1} Z' X_1 & \cdots & \hat{\sigma}^{MM} X_M' Z(Z'Z)^{-1} Z' X_M \end{bmatrix}^{-1}$$

---

[1] $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, where the dimensions of A, B, C, and D are appropriate.

Note:

$$n[X'(\hat{\Sigma}^{-1} \otimes P)X]^{-1} = n[X'(I \otimes Z)\hat{\Sigma}^{-1} \otimes (Z'Z)^{-1}(I \otimes Z')X]^{-1}$$

Let's consider the terms $X'(I \otimes Z)$, $\hat{\Sigma}^{-1} \otimes (Z'Z)^{-1}$, and $(I \otimes Z')X$ separately.

$$X'(I \otimes Z) = \begin{bmatrix} X_1' & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_M' \end{bmatrix} \begin{bmatrix} Z & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z \end{bmatrix} = \begin{bmatrix} X_1'Z & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_M'Z \end{bmatrix}$$

$$\hat{\Sigma}^{-1} \otimes (Z'Z)^{-1} = \begin{bmatrix} \hat{\sigma}^{11}(Z'Z)^{-1} & \cdots & \hat{\sigma}^{1M}(Z'Z)^{-1} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}^{M1}(Z'Z)^{-1} & \cdots & \hat{\sigma}^{MM}(Z'Z)^{-1} \end{bmatrix}$$

$$(I \otimes Z')X = \begin{bmatrix} Z' & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z' \end{bmatrix} \begin{bmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_M \end{bmatrix} = \begin{bmatrix} Z'X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z'X_M \end{bmatrix}$$

Thus:

$$\begin{aligned} & X'(I \otimes Z)\hat{\Sigma}^{-1} \otimes (Z'Z)^{-1}(I \otimes Z')X \\ &= \begin{bmatrix} X_1'Z & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_M'Z \end{bmatrix} \begin{bmatrix} \hat{\sigma}^{11}(Z'Z)^{-1} & \cdots & \hat{\sigma}^{1M}(Z'Z)^{-1} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}^{M1}(Z'Z)^{-1} & \cdots & \hat{\sigma}^{MM}(Z'Z)^{-1} \end{bmatrix} \begin{bmatrix} Z'X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z'X_M \end{bmatrix} \\ &= \begin{bmatrix} \hat{\sigma}^{11}X_1'Z(Z'Z)^{-1} & \cdots & \hat{\sigma}^{1M}X_1'Z(Z'Z)^{-1} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}^{11}X_M'Z(Z'Z)^{-1} & \cdots & \hat{\sigma}^{MM}X_M'Z(Z'Z)^{-1} \end{bmatrix} \begin{bmatrix} Z'X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Z'X_M \end{bmatrix} \\ &= \begin{bmatrix} \hat{\sigma}^{11}X_1'Z(Z'Z)^{-1}Z'X_1 & \cdots & \hat{\sigma}^{1M}X_1'Z(Z'Z)^{-1}Z'X_M \\ \vdots & \ddots & \vdots \\ \hat{\sigma}^{M1}X_M'Z(Z'Z)^{-1}Z'X_1 & \cdots & \hat{\sigma}^{MM}X_M'Z(Z'Z)^{-1}Z'X_M \end{bmatrix} \end{aligned}$$

Inverting the matrix and multiplying by n, the desired result is shown. □

Claim: For $j, k \in [M]$, let $\hat{A}_{jk}$ be defined as above. Then:

$$\hat{A}_{jk} = n^{-1}X_j'PX_k$$

Proof. This follows immediately from the definition of P, and the expression for $\hat{A}_{jk}$ derived in the proof above. □

Claim: *For $j, k \in [M]$, let:*

$$\hat{c}_{jk} = \left(n^{-1}\sum_{i=1}^{n} x_{ij} z_i'\right)\left(n^{-1}\sum_{i=1}^{n} z_i z_i'\right)^{-1}\left(n^{-1}\sum_{i=1}^{n} z_i y_{ik}\right)$$

Then:

$$\hat{c}_{jk} = n^{-1} X_j' P y_k$$

Proof.

$$\begin{aligned}
\hat{c}_{jk} &= \left(n^{-1}\sum_{i=1}^{n} x_{ij} z_i'\right)\left(n^{-1}\sum_{i=1}^{n} z_i z_i'\right)^{-1}\left(n^{-1}\sum_{i=1}^{n} z_i y_{ik}\right) \\
&= n^{-1} X_j' Z \big(n^{-1} Z'Z\big)^{-1} n^{-1} Z' y_k \\
&= n^{-1} X_j' Z (Z'Z)^{-1} Z' y_k \\
&= n^{-1} X_j' P y_k
\end{aligned}$$

□

Claim: *Define $J(\hat{\beta}_{3SLS}, \hat{S}^{-1})$ (Sargan's statistic) by:*

$$J(\hat{\beta}_{3SLS}, \hat{S}^{-1}) = n(s_{zy} - S_{zx}\hat{\beta}_{3SLS})'\hat{S}^{-1}(s_{zy} - S_{zx}\hat{\beta}_{3SLS})$$

Then:

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

Proof. Consider the term $s_{zy} - S_{zx}\hat{\beta}_{3SLS}$. Substituting the expressions for s_{zy} and S_{zx} from above, we can write this as:

$$\begin{aligned}
s_{zy} - S_{zx}\hat{\beta}_{3SLS} &= n^{-1}(I \otimes Z')y - n^{-1}(I \otimes Z')X\hat{\beta}_{3SLS} \\
&= n^{-1}(I \otimes Z')(y - X\hat{\beta}_{3SLS}) \\
\iff (s_{zy} - S_{zx}\hat{\beta}_{3SLS})' &= n^{-1}(y - X\hat{\beta}_{3SLS})'(I \otimes Z)
\end{aligned}$$

Substituting this and the expression for $\hat{S}^{-1}$ from above into the definition of Sargan's statistic, we have:

$$J(\hat{\beta}_{3SLS}, \hat{S}^{-1}) = n\Big(n^{-1}(y - X\hat{\beta}_{3SLS})'(I \otimes Z)\Big)\hat{\Sigma}^{-1}(n^{-1}Z'Z)^{-1}\Big(n^{-1}(I \otimes Z')(y - X\hat{\beta}_{3SLS})\Big)$$

Finally, using the fact that $(I \otimes Z)\hat{\Sigma}^{-1}(I \otimes Z') = \hat{\Sigma}^{-1} \otimes P$, we have:

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

□