Solution to Chapter 4 Analytical Exercises

1. It should be easy to show that $\widehat{\mathbf{A}}_{mh} = \frac{1}{n}\mathbf{Z}_m'\mathbf{P}\mathbf{Z}_h$ and that $\widehat{\mathbf{c}}_{mh} = \frac{1}{n}\mathbf{Z}_m'\mathbf{P}\mathbf{y}_h$. Going back to the formula (4.5.12) on p. 278 of the book, the first matrix on the RHS (the matrix to be inverted) is a partitioned matrix whose (m,h) block is $\widehat{\mathbf{A}}_{mh}$. It should be easy to see that it equals $\frac{1}{n}[\mathbf{Z}'(\widehat{\mathbf{\Sigma}}^{-1} \otimes \mathbf{P})\mathbf{Z}]$. Similarly, the second matrix on the RHS of (4.5.12) equals $\frac{1}{n}\mathbf{Z}'(\widehat{\mathbf{\Sigma}}^{-1} \otimes \mathbf{P})\mathbf{y}$.

2. The sprinkled hints are as good as the answer.

3. (b) (amplification of the answer given on p. 320) In this part only, for notational brevity, let $\mathbf{z}_i$ be a $\sum_m L_m \times 1$ stacked vector collecting $(\mathbf{z}_{i1}, \ldots, \mathbf{z}_{iM})$.

$$
\begin{aligned}
&\mathrm{E}(\varepsilon_{im} \mid \mathbf{Z}) \\
&= \mathrm{E}(\varepsilon_{im} \mid \mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n) \qquad (\text{since } \mathbf{Z} \text{ collects } \mathbf{z}_i\text{'s}) \\
&= \mathrm{E}(\varepsilon_{im} \mid \mathbf{z}_i) \qquad (\text{since } (\varepsilon_{im}, \mathbf{z}_i) \text{ is independent of } \mathbf{z}_j \ (j \neq i)) \\
&= 0 \qquad (\text{by the strengthened orthogonality conditions}).
\end{aligned}
$$

The (i,j) element of the $n \times n$ matrix $\mathrm{E}(\boldsymbol{\varepsilon}_m \boldsymbol{\varepsilon}_h' \mid \mathbf{Z})$ is $\mathrm{E}(\varepsilon_{im}\,\varepsilon_{jh} \mid \mathbf{Z})$.

$$
\begin{aligned}
&\mathrm{E}(\varepsilon_{im}\,\varepsilon_{jh} \mid \mathbf{Z}) = \mathrm{E}(\varepsilon_{im}\,\varepsilon_{jh} \mid \mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n) \\
&= \mathrm{E}(\varepsilon_{im}\,\varepsilon_{jh} \mid \mathbf{z}_i, \mathbf{z}_j) \qquad (\text{since } (\varepsilon_{im}, \mathbf{z}_i, \varepsilon_{jh}, \mathbf{z}_j) \text{ is independent of } \mathbf{z}_k \ (k \neq i, j)).
\end{aligned}
$$

For $j \neq i$, this becomes

$$
\begin{aligned}
&\mathrm{E}(\varepsilon_{im}\,\varepsilon_{jh} \mid \mathbf{z}_i, \mathbf{z}_j) \\
&= \mathrm{E}\left[\mathrm{E}(\varepsilon_{im}\,\varepsilon_{jh} \mid \mathbf{z}_i, \mathbf{z}_j, \varepsilon_{jh}) \mid \mathbf{z}_i, \mathbf{z}_j\right] \qquad (\text{by the Law of Iterated Expectations}) \\
&= \mathrm{E}\left[\varepsilon_{jh}\,\mathrm{E}(\varepsilon_{im} \mid \mathbf{z}_i, \mathbf{z}_j, \varepsilon_{jh}) \mid \mathbf{z}_i, \mathbf{z}_j\right] \qquad (\text{by linearity of conditional expectations}) \\
&= \mathrm{E}\left[\varepsilon_{jh}\,\mathrm{E}(\varepsilon_{im} \mid \mathbf{z}_i) \mid \mathbf{z}_i, \mathbf{z}_j\right] \qquad (\text{since } (\varepsilon_{im}, \mathbf{z}_i) \text{ is independent of } (\varepsilon_{jh}, \mathbf{z}_j)) \\
&= 0 \qquad (\text{since } \mathrm{E}(\varepsilon_{im} \mid \mathbf{z}_i) = 0).
\end{aligned}
$$

For $j = i$,

$$
\mathrm{E}(\varepsilon_{im}\,\varepsilon_{jh} \mid \mathbf{Z}) = \mathrm{E}(\varepsilon_{im}\,\varepsilon_{ih} \mid \mathbf{Z}) = \mathrm{E}(\varepsilon_{im}\,\varepsilon_{ih} \mid \mathbf{z}_i).
$$

Since $\mathbf{x}_{im} = \mathbf{x}_i$ and $\mathbf{x}_i$ is the union of $(\mathbf{z}_{i1}, \ldots, \mathbf{z}_{iM})$ in the SUR model, the conditional homoskedasticity assumption, Assumption 4.7, states that $\mathrm{E}(\varepsilon_{im}\,\varepsilon_{ih} \mid \mathbf{z}_i) = \mathrm{E}(\varepsilon_{im}\,\varepsilon_{ih} \mid \mathbf{x}_i) = \sigma_{mh}$.

(c) (i) We need to show that Assumptions 4.1-4.5, 4.7 and (4.5.18) together imply Assumptions 1.1-1.3 and (1.6.1). Assumption 1.1 (linearity) is obviously satisfied. Assumption 1.2 (strict exogeneity) and (1.6.1) have been verified in 3(b). That leaves Assumption 1.3 (the rank condition that $\mathbf{Z}$ (defined in Analytical Exercise 1) be of full column rank). Since $\mathbf{Z}$ is block diagonal, it suffices to show that $\mathbf{Z}_m$ is of full column rank for $m = 1, 2, \ldots, M$. The proof goes as follows. By Assumption 4.5,

$\mathbf{S}$ is non-singular. By Assumption 4.7 and the condition (implied by (4.5.18)) that the set of instruments be common across equations, we have $\mathbf{S} = \mathbf{\Sigma} \otimes \mathrm{E}(\mathbf{x}_i \mathbf{x}_i')$ (as in (4.5.9)). So the square matrix $\mathrm{E}(\mathbf{x}_i \mathbf{x}_i')$ is non-singular. Since $\frac{1}{n}\mathbf{X}'\mathbf{X}$ (where $\mathbf{X}$ is the $n \times K$ data matrix, as defined in Analytical Exercise 1) converges almost surely to $\mathrm{E}(\mathbf{x}_i \mathbf{x}_i')$, the $n \times K$ data matrix $\mathbf{X}$ is of full column rank for sufficiently large n. Since $\mathbf{Z}_m$ consists of columns selected from the columns of $\mathbf{X}$, $\mathbf{Z}_m$ is of full column rank as well.

(ii) The hint is the answer.

(iii) The unbiasedness of $\widehat{\boldsymbol{\delta}}_{\text{SUR}}$ follows from (i), (ii), and Proposition 1.7(a).

(iv) $\text{Avar}(\widehat{\boldsymbol{\delta}}_{\text{SUR}})$ is (4.5.15) where $\mathbf{A}_{mh}$ is given by (4.5.16′) on p. 280. The hint shows that it equals the plim of $n \cdot \text{Var}(\widehat{\boldsymbol{\delta}}_{\text{SUR}} \mid \mathbf{Z})$.

(d) For the most part, the answer is a straightforward modification of the answer to (c). The only part that is not so straightforward is to show in part (i) that the $Mn \times L$ matrix $\mathbf{Z}$ is of full column rank. Let $\mathbf{D}_m$ be the $\mathbf{D}_m$ matrix introduced in the answer to (c), so $\mathbf{z}_{im} = \mathbf{D}_m' \mathbf{x}_i$ and $\mathbf{Z}_m = \mathbf{X}\mathbf{D}_m$. Since the dimension of $\mathbf{x}_i$ is K and that of $\mathbf{z}_{im}$ is L, the matrix $\mathbf{D}_m$ is $K \times L$. The $\sum_{m=1}^{M} K_m \times L$ matrix $\mathbf{\Sigma}_{\mathbf{xz}}$ in Assumption 4.4′ can be written as

$$\underset{(KM \times L)}{\mathbf{\Sigma}_{\mathbf{xz}}} = [\mathbf{I}_M \otimes \mathrm{E}(\mathbf{x}_i \mathbf{x}_i')]\mathbf{D} \quad \text{where} \quad \underset{(KM \times L)}{\mathbf{D}} \equiv \begin{bmatrix} \mathbf{D}_1 \\ \vdots \\ \mathbf{D}_M \end{bmatrix}.$$

Since $\mathbf{\Sigma}_{\mathbf{xz}}$ is of full column rank by Assumption 4.4′ and since $\mathrm{E}(\mathbf{x}_i \mathbf{x}_i')$ is non-singular, $\mathbf{D}$ is of full column rank. So $\mathbf{Z} = (\mathbf{I}_M \otimes \mathbf{X})\mathbf{D}$ is of full column rank if $\mathbf{X}$ is of full column rank. $\mathbf{X}$ is of full column rank for sufficiently large n if $\mathrm{E}(\mathbf{x}_i \mathbf{x}_i')$ is non-singular.

4. (a) Assumptions 4.1-4.5 imply that the Avar of the efficient multiple-equation GMM estimator is $(\mathbf{\Sigma}_{\mathbf{xz}}' \mathbf{S}^{-1} \mathbf{\Sigma}_{\mathbf{xz}})^{-1}$. Assumption 4.2 implies that the plim of $\mathbf{S}_{\mathbf{xz}}$ is $\mathbf{\Sigma}_{\mathbf{xz}}$. Under Assumptions 4.1, 4.2, and 4.6, the plim of $\widehat{\mathbf{S}}$ is $\mathbf{S}$.

(b) The claim to be shown is just a restatement of Propositions 3.4 and 3.5.

(c) Use (A9) and (A6) of the book's Appendix A. $\mathbf{S}_{\mathbf{xz}}$ and $\widehat{\mathbf{W}}$ are block diagonal, so $\widehat{\mathbf{W}}\mathbf{S}_{\mathbf{xz}}(\mathbf{S}_{\mathbf{xz}}'\widehat{\mathbf{W}}\mathbf{S}_{\mathbf{xz}})^{-1}$ is block diagonal.

(d) If the same residuals are used in both the efficient equation-by-equation GMM and the efficient multiple-equation GMM, then the $\widehat{\mathbf{S}}$ in (**) and the $\widehat{\mathbf{S}}$ in $(\mathbf{S}_{\mathbf{xz}}'\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})^{-1}$ are numerically the same. The rest follows from the inequality in the question and the hint.

(e) Yes.

(f) The hint is the answer.

5. (a) For the *LW*69 equation, the instruments $(1, \mathit{MED})$ are 2 in number while the number of the regressors is 3. So the order condition is not satisfied for the equation.

(b) (reproducing the answer on pp. 320-321)

$$\begin{bmatrix} 1 & \mathrm{E}(\mathit{S69}) & \mathrm{E}(\mathit{IQ}) \\ 1 & \mathrm{E}(\mathit{S80}) & \mathrm{E}(\mathit{IQ}) \\ \mathrm{E}(\mathit{MED}) & \mathrm{E}(\mathit{S69} \cdot \mathit{MED}) & \mathrm{E}(\mathit{IQ} \cdot \mathit{MED}) \\ \mathrm{E}(\mathit{MED}) & \mathrm{E}(\mathit{S80} \cdot \mathit{MED}) & \mathrm{E}(\mathit{IQ} \cdot \mathit{MED}) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \mathrm{E}(\mathit{LW69}) \\ \mathrm{E}(\mathit{LW80}) \\ \mathrm{E}(\mathit{LW69} \cdot \mathit{MED}) \\ \mathrm{E}(\mathit{LW80} \cdot \mathit{MED}) \end{bmatrix}.$$

The condition for the system to be identified is that the 4×3 coefficient matrix is of full column rank.

(c) (reproducing the answer on p. 321) If IQ and MED are uncorrelated, then $\mathrm{E}(IQ \cdot MED) = \mathrm{E}(IQ) \cdot \mathrm{E}(MED)$ and the third column of the coefficient matrix is $\mathrm{E}(IQ)$ times the first column. So the matrix cannot be of full column rank.

6. (reproducing the answer on p. 321) $\widehat{\varepsilon}_{im} = y_{im} - \mathbf{z}'_{im}\widehat{\boldsymbol{\delta}}_m = \varepsilon_{im} - \mathbf{z}'_{im}(\widehat{\boldsymbol{\delta}}_m - \boldsymbol{\delta}_m)$. So

$$\frac{1}{n}\sum_{i=1}^{n}[\varepsilon_{im} - \mathbf{z}'_{im}(\widehat{\boldsymbol{\delta}}_m - \boldsymbol{\delta}_m)][\varepsilon_{ih} - \mathbf{z}'_{ih}(\widehat{\boldsymbol{\delta}}_h - \boldsymbol{\delta}_h)] = (1) + (2) + (3) + (4),$$

where

$$(1) = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_{im}\,\varepsilon_{ih},$$
$$(2) = -(\widehat{\boldsymbol{\delta}}_m - \boldsymbol{\delta}_m)'\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{im} \cdot \varepsilon_{ih},$$
$$(3) = -(\widehat{\boldsymbol{\delta}}_h - \boldsymbol{\delta}_h)'\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{ih} \cdot \varepsilon_{im},$$
$$(4) = (\widehat{\boldsymbol{\delta}}_m - \boldsymbol{\delta}_m)'\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{im}\mathbf{z}'_{ih}\Big)(\widehat{\boldsymbol{\delta}}_h - \boldsymbol{\delta}_h).$$

As usual, under Assumption 4.1 and 4.2, $(1) \to_{\mathrm{p}} \sigma_{mh}$ $(\equiv \mathrm{E}(\varepsilon_{im}\,\varepsilon_{ih}))$.

For (4), by Assumption 4.2 and the assumption that $\mathrm{E}(\mathbf{z}_{im}\mathbf{z}'_{ih})$ is finite, $\frac{1}{n}\sum_i \cdot \mathbf{z}_{im}\mathbf{z}'_{ih}$ converges in probability to a (finite) matrix. So $(4) \to_{\mathrm{p}} 0$.

Regarding (2), by Cauchy-Schwartz,

$$\mathrm{E}(|z_{imj} \cdot \varepsilon_{ih}|) \le \sqrt{\mathrm{E}(z_{imj}^2) \cdot \mathrm{E}(\varepsilon_{ih}^2)},$$

where z_{imj} is the j-th element of $\mathbf{z}_{im}$. So $\mathrm{E}(\mathbf{z}_{im} \cdot \varepsilon_{ih})$ is finite and $(2) \to_{\mathrm{p}} 0$ because $\widehat{\boldsymbol{\delta}}_m - \boldsymbol{\delta}_m \to_{\mathrm{p}} \mathbf{0}$. Similarly, $(3) \to_{\mathrm{p}} 0$.

7. (a) Let $\mathbf{B}$, $\mathbf{S_{xz}}$, and $\widehat{\mathbf{W}}$ be as defined in the hint. Also let

$$\underset{(MK\times 1)}{\mathbf{s_{xy}}} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \cdot y_{i1} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \cdot y_{iM} \end{bmatrix}.$$

Then

$$\begin{aligned}
\widehat{\boldsymbol{\delta}}_{3\mathrm{SLS}} &= \left(\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\mathbf{S_{xz}}\right)^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\mathbf{s_{xy}} \\
&= \left[(\mathbf{I}\otimes\mathbf{B}')(\widehat{\boldsymbol{\Sigma}}^{-1}\otimes\mathbf{S}_{\mathbf{xx}}^{-1})(\mathbf{I}\otimes\mathbf{B})\right]^{-1}(\mathbf{I}\otimes\mathbf{B}')(\widehat{\boldsymbol{\Sigma}}^{-1}\otimes\mathbf{S}_{\mathbf{xx}}^{-1})\mathbf{s_{xy}} \\
&= \left(\widehat{\boldsymbol{\Sigma}}^{-1}\otimes\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\mathbf{B}\right)^{-1}\left(\widehat{\boldsymbol{\Sigma}}^{-1}\otimes\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\right)\mathbf{s_{xy}} \\
&= \left(\widehat{\boldsymbol{\Sigma}}\otimes(\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\mathbf{B})^{-1}\right)\left(\widehat{\boldsymbol{\Sigma}}^{-1}\otimes\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\right)\mathbf{s_{xy}} \\
&= \left(\mathbf{I}_M\otimes(\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\right)\mathbf{s_{xy}} \\
&= \begin{bmatrix} (\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \cdot y_{i1} \\ \vdots \\ (\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{S}_{\mathbf{xx}}^{-1}\ \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \cdot y_{iM} \end{bmatrix},
\end{aligned}$$

which is a stacked vector of 2SLS estimators.

(b) The hint is the answer.

8. (a) The efficient multiple-equation GMM estimator is

$$\left(\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}}\right)^{-1}\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{s}_{\mathbf{xy}},$$

where $\mathbf{S}_{\mathbf{xz}}$ and $\mathbf{s}_{\mathbf{xy}}$ are as defined in (4.2.2) on p. 266 and $\widehat{\mathbf{S}}^{-1}$ is a consistent estimator of $\mathbf{S}$. Since $\mathbf{x}_{im} = \mathbf{z}_{im}$ here, $\mathbf{S}_{\mathbf{xz}}$ is square. So the above formula becomes

$$\mathbf{S}_{\mathbf{xz}}^{-1}\widehat{\mathbf{S}}\,\mathbf{S}'_{\mathbf{xz}}{}^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{S}}^{-1}\mathbf{s}_{\mathbf{xy}} = \mathbf{S}_{\mathbf{xz}}^{-1}\mathbf{s}_{\mathbf{xy}},$$

which is a stacked vector of OLS estimators.

(b) The SUR is efficient multiple-equation GMM under conditional homoskedasticity when the set of orthogonality conditions is $\mathrm{E}(\mathbf{z}_{im}\cdot\varepsilon_{ih}) = 0$ for all m, h. The OLS estimator derived above is (trivially) efficient multiple-equation GMM under conditional homoskedasticity when the set of orthogonality conditions is $\mathrm{E}(\mathbf{z}_{im}\cdot\varepsilon_{im}) = 0$ for all m. Since the sets of orthogonality conditions differ, the efficient GMM estimators differ.

9. The hint is the answer (to derive the formula in (b) of the hint, use the SUR formula you derived in Analytical Exercise 2(b)).

10. (a) $\mathrm{Avar}(\widehat{\boldsymbol{\delta}}_{1,\mathrm{2SLS}}) = \sigma_{11}\mathbf{A}_{11}^{-1}$.

(b) $\mathrm{Avar}(\widehat{\boldsymbol{\delta}}_{1,\mathrm{3SLS}})$ equals $\mathbf{G}^{-1}$. The hint shows that $\mathbf{G} = \frac{1}{\sigma_{11}}\mathbf{A}_{11}$.

11. Because there are as many orthogonality conditions as there are coefficients to be estimated, it is possible to choose $\widetilde{\boldsymbol{\delta}}$ so that $\mathbf{g}_n(\widetilde{\boldsymbol{\delta}})$ defined in the hint is a zero vector. Solving

$$\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{i1}\cdot y_{i1} + \cdots + \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{iM}\cdot y_{iM}\Big) - \Big(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{i1}\mathbf{z}'_{i1} + \cdots + \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{iM}\mathbf{z}'_{iM}\Big)\widetilde{\boldsymbol{\delta}} = \mathbf{0}$$

for $\widetilde{\boldsymbol{\delta}}$, we obtain

$$\widetilde{\boldsymbol{\delta}} = \Big(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{i1}\mathbf{z}'_{i1} + \cdots + \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{iM}\mathbf{z}'_{iM}\Big)^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{i1}\cdot y_{i1} + \cdots + \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_{iM}\cdot y_{iM}\Big),$$

which is none other than the pooled OLS estimator.