Solution to Chapter 3 Analytical Exercises

1. If $\mathbf{A}$ is symmetric and idempotent, then $\mathbf{A}' = \mathbf{A}$ and $\mathbf{AA} = \mathbf{A}$. So $\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{AA}\mathbf{x} = \mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{z}'\mathbf{z} \geq 0$ where $\mathbf{z} \equiv \mathbf{Ax}$.

2. (a) By assumption, $\{x_i, \varepsilon_i\}$ is jointly stationary and ergodic, so by ergodic theorem the first term of $(*)$ converges almost surely to $\mathrm{E}(x_i^2\varepsilon_i^2)$ which exists and is finite by Assumption 3.5.

   (b) $z_i x_i^2 \varepsilon_i$ is the product of $x_i\varepsilon_i$ and $x_i z_i$. By using the Cauchy-Schwarts inequality, we obtain

$$\mathrm{E}(|x_i\varepsilon_i \cdot x_i z_i|) \leq \sqrt{\mathrm{E}(x_i^2\varepsilon_i^2)\,\mathrm{E}(x_i^2 z_i^2)}.$$

   $\mathrm{E}(x_i^2\varepsilon_i^2)$ exists and is finite by Assumption 3.5 and $\mathrm{E}(x_i^2 z_i^2)$ exists and is finite by Assumption 3.6. Therefore, $\mathrm{E}(|x_i z_i \cdot x_i\varepsilon_i|)$ is finite. Hence, $\mathrm{E}(x_i z_i \cdot x_i \varepsilon_i)$ exists and is finite.

   (c) By ergodic stationarity the sample average of $z_i x_i^2 \varepsilon_i$ converges in probability to some finite number. Because $\widehat{\delta}$ is consistent for δ by Proposition 3.1, $\widehat{\delta} - \delta$ converges to 0 in probability. Therefore, the second term of $(*)$ converges to zero in probability.

   (d) By ergodic stationarity and Assumption 3.6 the sample average of $z_i^2 x_i^2$ converges in probability to some finite number. As mentioned in (c) $\widehat{\delta} - \delta$ converges to 0 in probability. Therefore, the last term of $(*)$ vanishes.

3. (a)

$$\begin{aligned}
\mathbf{Q} &\equiv \boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}} - \boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{xz}}(\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{WSW}\boldsymbol{\Sigma}_{\mathbf{xz}})^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{xz}} \\
&= \boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{C}'\mathbf{C}\boldsymbol{\Sigma}_{\mathbf{xz}} - \boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{xz}}(\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}\mathbf{C}^{-1}\mathbf{C}'^{-1}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{xz}})^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{xz}} \\
&= \mathbf{H}'\mathbf{H} - \boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{xz}}(\mathbf{G}'\mathbf{G})^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{xz}} \\
&= \mathbf{H}'\mathbf{H} - \mathbf{H}'\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{H} \\
&= \mathbf{H}'[\mathbf{I}_K - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}']\mathbf{H} \\
&= \mathbf{H}'\mathbf{M}_{\mathbf{G}}\mathbf{H}.
\end{aligned}$$

   (b) First, we show that $\mathbf{M}_{\mathbf{G}}$ is symmetric and idempotent.

$$\begin{aligned}
\mathbf{M}_{\mathbf{G}}' &= \mathbf{I}_K - \mathbf{G}(\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1})' \\
&= \mathbf{I}_K - \mathbf{G}((\mathbf{G}'\mathbf{G})^{-1\prime}\mathbf{G}') \\
&= \mathbf{I}_K - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}' \\
&= \mathbf{M}_{\mathbf{G}}.
\end{aligned}$$

$$\begin{aligned}
\mathbf{M}_{\mathbf{G}}\mathbf{M}_{\mathbf{G}} &= \mathbf{I}_K\mathbf{I}_K - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{I}_K - \mathbf{I}_K\,\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}' + \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}' \\
&= \mathbf{I}_K - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}' \\
&= \mathbf{M}_{\mathbf{G}}.
\end{aligned}$$

   Thus, $\mathbf{M}_{\mathbf{G}}$ is symmetric and idempotent. For any L-dimensional vector $\mathbf{x}$,

$$\begin{aligned}
\mathbf{x}'\mathbf{Q}\mathbf{x} &= \mathbf{x}'\mathbf{H}'\mathbf{M}_{\mathbf{G}}\mathbf{H}\mathbf{x} \\
&= \mathbf{z}'\mathbf{M}_{\mathbf{G}}\mathbf{z} \quad (\text{where } \mathbf{z} \equiv \mathbf{Hx}) \\
&\geq 0 \qquad (\text{since } \mathbf{M}_{\mathbf{G}} \text{ is positive semidefinite}).
\end{aligned}$$

   Therefore, $\mathbf{Q}$ is positive semidefinite.

4. (the answer on p. 254 of the book simplified) If $\mathbf{W}$ is as defined in the hint, then

$$\mathbf{WSW} = \mathbf{W} \quad \text{and} \quad \mathbf{\Sigma}'_{\mathbf{xz}}\mathbf{W}\mathbf{\Sigma}_{\mathbf{xz}} = \mathbf{\Sigma}_{\mathbf{zz}}\mathbf{A}^{-1}\mathbf{\Sigma}_{\mathbf{zz}}.$$

So (3.5.1) reduces to the asymptotic variance of the OLS estimator. By (3.5.11), it is no smaller than $(\mathbf{\Sigma}'_{\mathbf{xz}}\,\mathbf{S}^{-1}\mathbf{\Sigma}_{\mathbf{xz}})^{-1}$, which is the asymptotic variance of the efficient GMM estimator.

5. (a) From the expression for $\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1})$ (given in (3.5.12)) and the expression for $\mathbf{g}_n(\widetilde{\boldsymbol{\delta}})$ (given in (3.4.2)), it is easy to show that $\mathbf{g}_n(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1})) = \widehat{\mathbf{B}}\mathbf{s}_{\mathbf{xy}}$. But $\widehat{\mathbf{B}}\mathbf{s}_{\mathbf{xy}} = \widehat{\mathbf{B}}\overline{\mathbf{g}}$ because

$$\begin{aligned}
\widehat{\mathbf{B}}\mathbf{s}_{\mathbf{xy}} &= (\mathbf{I}_K - \mathbf{S}_{\mathbf{xz}}(\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1})\mathbf{s}_{\mathbf{xy}} \\
&= (\mathbf{I}_K - \mathbf{S}_{\mathbf{xz}}(\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1})(\mathbf{S}_{\mathbf{xz}}\boldsymbol{\delta} + \overline{\mathbf{g}}) \qquad (\text{since } y_i = \mathbf{z}'_i\boldsymbol{\delta} + \varepsilon_i) \\
&= (\mathbf{S}_{\mathbf{xz}} - \mathbf{S}_{\mathbf{xz}}(\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})\boldsymbol{\delta} + (\mathbf{I}_K - \mathbf{S}_{\mathbf{xz}}(\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1})\overline{\mathbf{g}} \\
&= (\mathbf{S}_{\mathbf{xz}} - \mathbf{S}_{\mathbf{xz}})\boldsymbol{\delta} + \widehat{\mathbf{B}}\overline{\mathbf{g}} \\
&= \widehat{\mathbf{B}}\overline{\mathbf{g}}.
\end{aligned}$$

(b) Since $\widehat{\mathbf{S}}^{-1} = \mathbf{C}'\mathbf{C}$, we obtain $\widehat{\mathbf{B}}'\widehat{\mathbf{S}}^{-1}\widehat{\mathbf{B}} = \widehat{\mathbf{B}}'\mathbf{C}'\mathbf{C}\widehat{\mathbf{B}} = (\mathbf{C}\widehat{\mathbf{B}})'(\mathbf{C}\widehat{\mathbf{B}})$. But

$$\begin{aligned}
\mathbf{C}\widehat{\mathbf{B}} &= \mathbf{C}(\mathbf{I}_K - \mathbf{S}_{\mathbf{xz}}(\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\,\widehat{\mathbf{S}}^{-1}) \\
&= \mathbf{C} - \mathbf{C}\mathbf{S}_{\mathbf{xz}}(\mathbf{S}'_{\mathbf{xz}}\,\mathbf{C}'\mathbf{C}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\,\mathbf{C}'\mathbf{C} \\
&= \mathbf{C} - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{C} \qquad (\text{where } \mathbf{A} \equiv \mathbf{C}\mathbf{S}_{\mathbf{xz}}) \\
&= [\mathbf{I}_K - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}']\mathbf{C} \\
&\equiv \mathbf{MC}.
\end{aligned}$$

So $\widehat{\mathbf{B}}'\widehat{\mathbf{S}}^{-1}\widehat{\mathbf{B}} = (\mathbf{MC})'(\mathbf{MC}) = \mathbf{C}'\mathbf{M}'\mathbf{MC}$. It should be routine to show that $\mathbf{M}$ is symmetric and idempotent. Thus $\widehat{\mathbf{B}}'\widehat{\mathbf{S}}^{-1}\widehat{\mathbf{B}} = \mathbf{C}'\mathbf{MC}$.

The rank of $\mathbf{M}$ equals its trace, which is

$$\begin{aligned}
\text{trace}(\mathbf{M}) &= \text{trace}(\mathbf{I}_K - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}') \\
&= \text{trace}(\mathbf{I}_K) - \text{trace}(\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}') \\
&= \text{trace}(\mathbf{I}_K) - \text{trace}(\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}) \\
&= K - \text{trace}(\mathbf{I}_L) \\
&= K - L.
\end{aligned}$$

(c) As defined in (b), $\mathbf{C}'\mathbf{C} = \widehat{\mathbf{S}}^{-1}$. Let $\mathbf{D}$ be such that $\mathbf{D}'\mathbf{D} = \mathbf{S}^{-1}$. The choice of $\mathbf{C}$ and $\mathbf{D}$ is not unique, but it would be possible to choose $\mathbf{C}$ so that $\text{plim}\,\mathbf{C} = \mathbf{D}$. Now,

$$\mathbf{v} \equiv \sqrt{n}(\mathbf{C}\overline{\mathbf{g}}) = \mathbf{C}(\sqrt{n}\,\overline{\mathbf{g}}).$$

By using the Ergodic Stationary Martingale Differences CLT, we obtain $\sqrt{n}\,\overline{\mathbf{g}} \to_{\text{d}} N(\mathbf{0}, \mathbf{S})$. So

$$\mathbf{v} = \mathbf{C}(\sqrt{n}\,\overline{\mathbf{g}}) \underset{\text{d}}{\to} N(\mathbf{0}, \text{Avar}(\mathbf{v}))$$

where

$$\begin{aligned}
\text{Avar}(\mathbf{v}) &= \mathbf{DSD}' \\
&= \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}' \\
&= \mathbf{D}\mathbf{D}^{-1}\mathbf{D}^{-1\prime}\mathbf{D}' \\
&= \mathbf{I}_K.
\end{aligned}$$

(d)

$$\begin{aligned}
J(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1}),\widehat{\mathbf{S}}^{-1}) &= n\cdot \mathbf{g}_n(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1}))\,\widehat{\mathbf{S}}^{-1}\mathbf{g}_n(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1}))\\
&= n\cdot(\widehat{\mathbf{B}}\overline{\mathbf{g}})'\widehat{\mathbf{S}}^{-1}(\widehat{\mathbf{B}}\overline{\mathbf{g}}) \qquad \text{(by (a))}\\
&= n\cdot\overline{\mathbf{g}}'\widehat{\mathbf{B}}'\widehat{\mathbf{S}}^{-1}\widehat{\mathbf{B}}\overline{\mathbf{g}}\\
&= n\cdot\overline{\mathbf{g}}'\mathbf{C}'\mathbf{M}\mathbf{C}\overline{\mathbf{g}} \qquad \text{(by (b))}\\
&= \mathbf{v}'\mathbf{M}\mathbf{v} \qquad (\text{since } \mathbf{v}\equiv\sqrt{n}\mathbf{C}\overline{\mathbf{g}}).
\end{aligned}$$

Since $\mathbf{v}\to_{\mathrm{d}} N(\mathbf{0},\mathbf{I}_K)$ and $\mathbf{M}$ is idempotent, $\mathbf{v}'\mathbf{M}\mathbf{v}$ is asymptotically chi-squared with degrees of freedom equaling the rank of $\mathbf{M}=K-L$.

6. From Exercise 5, $J=n\cdot\overline{\mathbf{g}}'\widehat{\mathbf{B}}'\widehat{\mathbf{S}}^{-1}\widehat{\mathbf{B}}\overline{\mathbf{g}}$. Also from Exercise 5, $\widehat{\mathbf{B}}\overline{\mathbf{g}}=\widehat{\mathbf{B}}\mathbf{s}_{\mathbf{xy}}$.

7. For the most parts, the hints are nearly the answer. Here, we provide answers to (d), (f), (g), (i), and (j).

(d) As shown in (c), $J_1=\mathbf{v}_1'\mathbf{M}_1\mathbf{v}_1$. It suffices to prove that $\mathbf{v}_1=\mathbf{C}_1\mathbf{F}'\mathbf{C}^{-1}\mathbf{v}$.

$$\begin{aligned}
\mathbf{v}_1 &\equiv \sqrt{n}\mathbf{C}_1\overline{\mathbf{g}}_1\\
&= \sqrt{n}\mathbf{C}_1\mathbf{F}'\overline{\mathbf{g}}\\
&= \sqrt{n}\mathbf{C}_1\mathbf{F}'\mathbf{C}^{-1}\mathbf{C}\overline{\mathbf{g}}\\
&= \mathbf{C}_1\mathbf{F}'\mathbf{C}^{-1}\sqrt{n}\mathbf{C}\overline{\mathbf{g}}\\
&= \mathbf{C}_1\mathbf{F}'\mathbf{C}^{-1}\mathbf{v} \qquad (\text{since } \mathbf{v}\equiv\sqrt{n}\mathbf{C}\overline{\mathbf{g}}).
\end{aligned}$$

(f) Use the hint to show that $\mathbf{A}'\mathbf{D}=\mathbf{0}$ if $\mathbf{A}_1'\mathbf{M}_1=\mathbf{0}$. It should be easy to show that $\mathbf{A}_1'\mathbf{M}_1=\mathbf{0}$ from the definition of $\mathbf{M}_1$.

(g) By the definition of $\mathbf{M}$ in Exercise 5, $\mathbf{M}\mathbf{D}=\mathbf{D}-\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{D}$. So $\mathbf{M}\mathbf{D}=\mathbf{D}$ since $\mathbf{A}'\mathbf{D}=\mathbf{0}$ as shown in the previous part. Since both $\mathbf{M}$ and $\mathbf{D}$ are symmetric, $\mathbf{D}\mathbf{M}=\mathbf{D}'\mathbf{M}'=(\mathbf{M}\mathbf{D})'=\mathbf{D}'=\mathbf{D}$. As shown in part (e), $\mathbf{D}$ is idempotent. Also, $\mathbf{M}$ is idempotent as shown in Exercise 5. So $(\mathbf{M}-\mathbf{D})^2=\mathbf{M}^2-\mathbf{D}\mathbf{M}-\mathbf{M}\mathbf{D}+\mathbf{D}^2=\mathbf{M}-\mathbf{D}$. As shown in Exercise 5, the trace of $\mathbf{M}$ is $K-L$. As shown in (e), the trace of D is K_1-L. So the trace of $\mathbf{M}-\mathbf{D}$ is $K-K_1$. The rank of a symmetric and idempotent matrix is its trace.

(i) It has been shown in Exercise 6 that $\overline{\mathbf{g}}'\mathbf{C}'\mathbf{M}\mathbf{C}\overline{\mathbf{g}}=\mathbf{s}_{\mathbf{xy}}'\mathbf{C}'\mathbf{M}\mathbf{C}\mathbf{s}_{\mathbf{xy}}$ since $\mathbf{C}'\mathbf{M}\mathbf{C}=\widehat{\mathbf{B}}'\widehat{\mathbf{S}}^{-1}\widehat{\mathbf{B}}$. Here, we show that $\overline{\mathbf{g}}'\mathbf{C}'\mathbf{D}\mathbf{C}\overline{\mathbf{g}}=\mathbf{s}_{\mathbf{xy}}'\mathbf{C}'\mathbf{D}\mathbf{C}\mathbf{s}_{\mathbf{xy}}$.

$$\begin{aligned}
\overline{\mathbf{g}}'\mathbf{C}'\mathbf{D}\mathbf{C}\overline{\mathbf{g}} &= \overline{\mathbf{g}}'\mathbf{F}\mathbf{C}_1'\mathbf{M}_1\mathbf{C}_1\mathbf{F}'\overline{\mathbf{g}} \qquad (\mathbf{C}'\mathbf{D}\mathbf{C}=\mathbf{F}\mathbf{C}_1'\mathbf{M}_1\mathbf{C}_1\mathbf{F}' \text{ by the definition of } \mathbf{D} \text{ in (d)})\\
&= \overline{\mathbf{g}}'\mathbf{F}\widehat{\mathbf{B}}_1'(\widehat{\mathbf{S}}_{11})^{-1}\widehat{\mathbf{B}}_1\mathbf{F}'\overline{\mathbf{g}} \qquad (\text{since } \mathbf{C}_1'\mathbf{M}_1\mathbf{C}_1=\widehat{\mathbf{B}}_1'(\widehat{\mathbf{S}}_{11})^{-1}\widehat{\mathbf{B}}_1 \text{ from (a)})\\
&= \overline{\mathbf{g}}_1'\widehat{\mathbf{B}}_1'(\widehat{\mathbf{S}}_{11})^{-1}\widehat{\mathbf{B}}_1\overline{\mathbf{g}}_1 \qquad (\text{since } \overline{\mathbf{g}}_1=\mathbf{F}'\overline{\mathbf{g}}).
\end{aligned}$$

From the definition of $\widehat{\mathbf{B}}_1$ and the fact that $\mathbf{s}_{\mathbf{x}_1\mathbf{y}}=\mathbf{S}_{\mathbf{x}_1\mathbf{z}}\boldsymbol{\delta}+\overline{\mathbf{g}}_1$, it follows that $\widehat{\mathbf{B}}_1\overline{\mathbf{g}}_1=\widehat{\mathbf{B}}_1\mathbf{s}_{\mathbf{x}_1\mathbf{y}}$. So

$$\begin{aligned}
\overline{\mathbf{g}}_1'\widehat{\mathbf{B}}_1'(\widehat{\mathbf{S}}_{11})^{-1}\widehat{\mathbf{B}}_1\overline{\mathbf{g}}_1 &= \mathbf{s}_{\mathbf{x}_1\mathbf{y}}'\widehat{\mathbf{B}}_1'(\widehat{\mathbf{S}}_{11})^{-1}\widehat{\mathbf{B}}_1\mathbf{s}_{\mathbf{x}_1\mathbf{y}}\\
&= \mathbf{s}_{\mathbf{xy}}'\mathbf{F}\widehat{\mathbf{B}}_1'(\widehat{\mathbf{S}}_{11})^{-1}\widehat{\mathbf{B}}_1\mathbf{F}'\mathbf{s}_{\mathbf{xy}} \qquad (\text{since } \mathbf{s}_{\mathbf{x}_1\mathbf{y}}=\mathbf{F}'\mathbf{s}_{\mathbf{xy}})\\
&= \mathbf{s}_{\mathbf{xy}}'\mathbf{F}\mathbf{C}_1'\mathbf{M}_1\mathbf{C}_1\mathbf{F}'\mathbf{s}_{\mathbf{xy}} \qquad (\text{since } \widehat{\mathbf{B}}_1'(\widehat{\mathbf{S}}_{11})^{-1}\widehat{\mathbf{B}}_1=\mathbf{C}_1'\mathbf{M}_1\mathbf{C}_1 \text{ from (a)})\\
&= \mathbf{s}_{\mathbf{xy}}'\mathbf{C}'\mathbf{D}\mathbf{C}\mathbf{s}_{\mathbf{xy}}.
\end{aligned}$$

(j) $\mathbf{M} - \mathbf{D}$ is positive semi-definite because it is symmetric and idempotent.

8. (a) Solve the first-order conditions in the hint for $\overline{\boldsymbol{\delta}}$ to obtain

$$\overline{\boldsymbol{\delta}} = \widehat{\boldsymbol{\delta}}(\widehat{\mathbf{W}}) - \frac{1}{2n}(\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{R}'\boldsymbol{\lambda}.$$

Substitute this into the constraint $\mathbf{R}\overline{\boldsymbol{\delta}} = \mathbf{r}$ to obtain the expression for $\boldsymbol{\lambda}$ in the question. Then substitute this expression for $\boldsymbol{\lambda}$ into the above equation to obtain the expression for $\overline{\boldsymbol{\delta}}$ in the question.

(b) The hint is almost the answer.

(c) What needs to be shown is that $n\cdot(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{W}}) - \overline{\boldsymbol{\delta}})'(\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}\mathbf{S}_{\mathbf{xz}})(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{W}}) - \overline{\boldsymbol{\delta}})$ equals the Wald statistic. But this is immediate from substitution of the expression for $\overline{\boldsymbol{\delta}}$ in (a).

9. (a) By applying (3.4.11), we obtain

$$\begin{bmatrix} \sqrt{n}(\widehat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}) \\ \sqrt{n}(\widehat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}) \end{bmatrix} = \begin{bmatrix} (\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_1\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_1 \\ (\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_2\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_2 \end{bmatrix} \sqrt{n}\overline{\mathbf{g}}.$$

By using Billingsley CLT, we have

$$\sqrt{n}\overline{\mathbf{g}} \underset{\mathrm{d}}{\rightarrow} N(\mathbf{0}, \mathbf{S}).$$

Also, we have

$$\begin{bmatrix} (\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_1\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_1 \\ (\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_2\mathbf{S}_{\mathbf{xz}})^{-1}\mathbf{S}'_{\mathbf{xz}}\widehat{\mathbf{W}}_2 \end{bmatrix} \underset{\mathrm{p}}{\rightarrow} \begin{bmatrix} \mathbf{Q}_1^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}_1 \\ \mathbf{Q}_2^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}_2 \end{bmatrix}.$$

Therefore, by Lemma 2.4(c),

$$\begin{aligned} \begin{bmatrix} \sqrt{n}(\widehat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}) \\ \sqrt{n}(\widehat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}) \end{bmatrix} &\rightarrow_{\mathrm{d}} N\left(\mathbf{0},\ \begin{bmatrix} \mathbf{Q}_1^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}_1 \\ \mathbf{Q}_2^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}_2 \end{bmatrix} \mathbf{S}\left(\mathbf{W}_1\boldsymbol{\Sigma}_{\mathbf{xz}}\mathbf{Q}_1^{-1} \ \vdots\ \mathbf{W}_2\boldsymbol{\Sigma}_{\mathbf{xz}}\mathbf{Q}_2^{-1}\right)\right) \\ &= N\left(\mathbf{0},\ \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}\right). \end{aligned}$$

(b) $\sqrt{n}\mathbf{q}$ can be rewritten as

$$\sqrt{n}\mathbf{q} = \sqrt{n}(\widehat{\boldsymbol{\delta}}_1 - \widehat{\boldsymbol{\delta}}_2) = \sqrt{n}(\widehat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}) - \sqrt{n}(\widehat{\boldsymbol{\delta}}_2 - \boldsymbol{\delta}) = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{n}(\widehat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}) \\ \sqrt{n}(\widehat{\boldsymbol{\delta}}_2 - \boldsymbol{\delta}) \end{bmatrix}.$$

Therefore, we obtain

$$\sqrt{n}\mathbf{q} \underset{\mathrm{d}}{\rightarrow} N(\mathbf{0}, \mathrm{Avar}(\mathbf{q})).$$

where

$$\mathrm{Avar}(\mathbf{q}) = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{A}_{11} + \mathbf{A}_{22} - \mathbf{A}_{12} - \mathbf{A}_{21}.$$

(c) Since $\mathbf{W}_2 = \mathbf{S}^{-1}$, $\mathbf{Q}_2$, $\mathbf{A}_{12}$, $\mathbf{A}_{21}$, and $\mathbf{A}_{22}$ can be rewritten as follows:

$$\begin{aligned}
\mathbf{Q}_2 &= \boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}_2\boldsymbol{\Sigma}_{\mathbf{xz}} \\
&= \boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}},
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{12} &= \mathbf{Q}_1^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}_1\mathbf{S}\,\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}}\mathbf{Q}_2^{-1} \\
&= \mathbf{Q}_1^{-1}(\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{W}_1\boldsymbol{\Sigma}_{\mathbf{xz}})\mathbf{Q}_2^{-1} \\
&= \mathbf{Q}_1^{-1}\mathbf{Q}_1\mathbf{Q}_2^{-1} \\
&= \mathbf{Q}_2^{-1},
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{21} &= \mathbf{Q}_2^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\mathbf{S}\mathbf{W}_1\boldsymbol{\Sigma}_{\mathbf{xz}}\mathbf{Q}_1^{-1} \\
&= \mathbf{Q}_2^{-1},
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{22} &= (\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}})^{-1}\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\mathbf{S}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}}(\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}})^{-1} \\
&= (\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}})^{-1} \\
&= \mathbf{Q}_2^{-1}.
\end{aligned}$$

Substitution of these into the expression for $\text{Avar}(\mathbf{q})$ in (b), we obtain

$$\begin{aligned}
\text{Avar}(\mathbf{q}) &= \mathbf{A}_{11} - \mathbf{Q}_2^{-1} \\
&= \mathbf{A}_{11} - (\boldsymbol{\Sigma}'_{\mathbf{xz}}\mathbf{S}^{-1}\boldsymbol{\Sigma}_{\mathbf{xz}})^{-1} \\
&= \text{Avar}(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{W}}_1)) - \text{Avar}(\widehat{\boldsymbol{\delta}}(\widehat{\mathbf{S}}^{-1})).
\end{aligned}$$

10. (a)

$$\begin{aligned}
\sigma_{xz} \equiv \text{E}(x_i z_i) &= \text{E}(x_i(x_i\beta + v_i)) \\
&= \beta\,\text{E}(x_i^2) + \text{E}(x_i v_i) \\
&= \beta\sigma_x^2 \neq 0 \qquad \text{(by assumptions (2), (3), and (4))}.
\end{aligned}$$

(b) From the definition of $\widehat{\delta}$,

$$\widehat{\delta} - \delta = \left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} x_i\varepsilon_i = s_{xz}^{-1}\ \frac{1}{n}\sum_{i=1}^{n} x_i\varepsilon_i.$$

We have $x_i z_i = x_i(x_i\beta + v_i) = x_i^2\beta + x_i v_i$, which, being a function of $(x_i, \boldsymbol{\eta}_i)$, is ergodic stationary by assumption (1). So by the Ergodic theorem, $s_{xz} \to_{\text{p}} \sigma_{xz}$. Since $\sigma_{xz} \neq 0$ by (a), we have $s_{xz}^{-1} \to_{\text{p}} \sigma_{xz}^{-1}$. By assumption (2), $\text{E}(x_i\varepsilon_i) = 0$. So by assumption (1), we have $\frac{1}{n}\sum_{i=1}^{n} x_i\varepsilon_i \to_{\text{p}} 0$. Thus $\widehat{\delta} - \delta \to_{\text{p}} 0$.

(c)

$$\begin{aligned}
s_{xz} &\equiv \frac{1}{n}\sum_{i=1}^{n} x_i z_i \\
&= \frac{1}{n}\sum_{i=1}^{n} (x_i^2\beta + x_i v_i) \\
&= \frac{1}{\sqrt{n}}\frac{1}{n}\sum_{i=1}^{n} x_i^2 + \frac{1}{n}\sum_{i=1}^{n} x_i v_i \qquad (\text{since } \beta = \frac{1}{\sqrt{n}}) \\
&\to_{\text{p}} 0\cdot\text{E}(x_i^2) + \text{E}(x_i v_i) \\
&= 0
\end{aligned}$$

(d)

$$\sqrt{n}s_{xz} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i v_i.$$

By assumption (1) and the Ergodic Theorem, the first term of RHS converges in probability to $\mathrm{E}(x_i^2) = \sigma_x^2 > 0$. Assumption (2) and the Martingale Differences CLT imply that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i v_i \underset{\mathrm{d}}{\rightarrow} a \sim N(0, s_{22}).$$

Therefore, by Lemma 2.4(a), we obtain

$$\sqrt{n}s_{xz} \underset{\mathrm{d}}{\rightarrow} \sigma_x^2 + a.$$

(e) $\widehat{\delta} - \delta$ can be rewritten as

$$\widehat{\delta} - \delta = (\sqrt{n}s_{xz})^{-1}\sqrt{n}\overline{g}_1.$$

From assumption (2) and the Martingale Differences CLT, we obtain

$$\sqrt{n}\overline{g}_1 \underset{\mathrm{d}}{\rightarrow} b \sim N(0, s_{11}).$$

where s_{11} is the $(1, 1)$ element of $\mathbf{S}$. By using the result of (d) and Lemma 2.3(b),

$$\widehat{\delta} - \delta \underset{\mathrm{d}}{\rightarrow} (\sigma_x^2 + a)^{-1}b.$$

(a, b) are jointly normal because the joint distribution is the limiting distribution of

$$\sqrt{n}\,\overline{\mathbf{g}} = \begin{bmatrix} \sqrt{n}\overline{g}_1 \\ \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} x_i v_i) \end{bmatrix}.$$

(f) Because $\widehat{\delta} - \delta$ converges in distribution to $(\sigma_x^2 + a)^{-1}b$ which is not zero, the answer is No.