

Solution to Chapter 2 Analytical Exercises

1. For any $\varepsilon > 0$,

$$\text{Prob}(|z_n| > \varepsilon) = \frac{1}{n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

So, $\text{plim } z_n = 0$. On the other hand,

$$E(z_n) = \frac{n-1}{n} \cdot 0 + \frac{1}{n} \cdot n^2 = n,$$

which means that $\lim_{n \rightarrow \infty} E(z_n) = \infty$.

2. As shown in the hint,

$$(\bar{z}_n - \mu)^2 = (\bar{z}_n - E(\bar{z}_n))^2 + 2(\bar{z}_n - E(\bar{z}_n))(E(\bar{z}_n) - \mu) + (E(\bar{z}_n) - \mu)^2.$$

Take the expectation of both sides to obtain

$$\begin{aligned} E[(\bar{z}_n - \mu)^2] &= E[(\bar{z}_n - E(\bar{z}_n))^2] + 2E[\bar{z}_n - E(\bar{z}_n)](E(\bar{z}_n) - \mu) + (E(\bar{z}_n) - \mu)^2 \\ &= \text{Var}(\bar{z}_n) + (E(\bar{z}_n) - \mu)^2 \quad (\text{because } E[\bar{z}_n - E(\bar{z}_n)] = E(\bar{z}_n) - E(\bar{z}_n) = 0). \end{aligned}$$

Take the limit as $n \rightarrow \infty$ of both sides to obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} E[(\bar{z}_n - \mu)^2] &= \lim_{n \rightarrow \infty} \text{Var}(\bar{z}_n) + \lim_{n \rightarrow \infty} (E(\bar{z}_n) - \mu)^2 \\ &= 0 \quad (\text{because } \lim_{n \rightarrow \infty} E(\bar{z}_n) = \mu, \lim_{n \rightarrow \infty} \text{Var}(\bar{z}_n) = 0). \end{aligned}$$

Therefore, $z_n \xrightarrow{\text{m.s.}} \mu$. By Lemma 2.2(a), this implies $z_n \xrightarrow{\text{p}} \mu$.

3. (a) Since an i.i.d. process is ergodic stationary, Assumption 2.2 is implied by Assumption 2.2'. Assumptions 2.1 and 2.2' imply that $\mathbf{g}_i \equiv \mathbf{x}_i \cdot \varepsilon_i$ is i.i.d. Since an i.i.d. process with mean zero is mds (martingale differences), Assumption 2.5 is implied by Assumptions 2.2' and 2.5'.
 (b) Rewrite the OLS estimator as

$$\mathbf{b} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} = \mathbf{S}_{\mathbf{xx}}^{-1}\bar{\mathbf{g}}. \quad (\text{A})$$

Since by Assumption 2.2' $\{\mathbf{x}_i\}$ is i.i.d., $\{\mathbf{x}_i\mathbf{x}_i'\}$ is i.i.d. So by Kolmogorov's Second Strong LLN, we obtain

$$\mathbf{S}_{\mathbf{xx}} \xrightarrow{\text{p}} \boldsymbol{\Sigma}_{\mathbf{xx}}$$

The convergence is actually almost surely, but almost sure convergence implies convergence in probability. Since $\boldsymbol{\Sigma}_{\mathbf{xx}}$ is invertible by Assumption 2.4, by Lemma 2.3(a) we get

$$\mathbf{S}_{\mathbf{xx}}^{-1} \xrightarrow{\text{p}} \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}.$$

Similarly, under Assumption 2.1 and 2.2' $\{\mathbf{g}_i\}$ is i.i.d. By Kolmogorov's Second Strong LLN, we obtain

$$\bar{\mathbf{g}} \xrightarrow{p} \mathbf{E}(\mathbf{g}_i),$$

which is zero by Assumption 2.3. So by Lemma 2.3(a),

$$\mathbf{S}_{\mathbf{xx}}^{-1} \bar{\mathbf{g}} \xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1} \cdot \mathbf{0} = \mathbf{0}.$$

Therefore, $\text{plim}_{n \rightarrow \infty}(\mathbf{b} - \boldsymbol{\beta}) = \mathbf{0}$ which implies that the OLS estimator \mathbf{b} is consistent.

Next, we prove that the OLS estimator \mathbf{b} is asymptotically normal. Rewrite equation(A) above as

$$\sqrt{n}(\mathbf{b} - \boldsymbol{\beta}) = \mathbf{S}_{\mathbf{xx}}^{-1} \sqrt{n}\bar{\mathbf{g}}.$$

As already observed, $\{\mathbf{g}_i\}$ is i.i.d. with $\mathbf{E}(\mathbf{g}_i) = \mathbf{0}$. The variance of \mathbf{g}_i equals $\mathbf{E}(\mathbf{g}_i \mathbf{g}_i') = \mathbf{S}$ since $\mathbf{E}(\mathbf{g}_i) = \mathbf{0}$ by Assumption 2.3. So by the Lindeberg-Levy CLT,

$$\sqrt{n}\bar{\mathbf{g}} \xrightarrow{d} N(\mathbf{0}, \mathbf{S}).$$

Furthermore, as already noted, $\mathbf{S}_{\mathbf{xx}}^{-1} \xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}$. Thus by Lemma 2.4(c),

$$\sqrt{n}(\mathbf{b} - \boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1} \mathbf{S} \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}).$$

4. The hint is as good as the answer.
5. As shown in the solution to Chapter 1 Analytical Exercise 5, $SSR_R - SSR_U$ can be written as

$$SSR_R - SSR_U = (\mathbf{R}\mathbf{b} - \mathbf{r})' [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} (\mathbf{R}\mathbf{b} - \mathbf{r}).$$

Using the restrictions of the null hypothesis,

$$\begin{aligned} \mathbf{R}\mathbf{b} - \mathbf{r} &= \mathbf{R}(\mathbf{b} - \boldsymbol{\beta}) \\ &= \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} \quad (\text{since } \mathbf{b} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}) \\ &= \mathbf{R}\mathbf{S}_{\mathbf{xx}}^{-1}\bar{\mathbf{g}} \quad (\text{where } \bar{\mathbf{g}} \equiv \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \cdot \varepsilon_i). \end{aligned}$$

Also $[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}]^{-1} = n \cdot [\mathbf{R}\mathbf{S}_{\mathbf{xx}}^{-1}\mathbf{R}]^{-1}$. So

$$SSR_R - SSR_U = (\sqrt{n}\bar{\mathbf{g}})' \mathbf{S}_{\mathbf{xx}}^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_{\mathbf{xx}}^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_{\mathbf{xx}}^{-1} (\sqrt{n}\bar{\mathbf{g}}).$$

Thus

$$\begin{aligned} \frac{SSR_R - SSR_U}{s^2} &= (\sqrt{n}\bar{\mathbf{g}})' \mathbf{S}_{\mathbf{xx}}^{-1} \mathbf{R}' (s^2 \mathbf{R} \mathbf{S}_{\mathbf{xx}}^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_{\mathbf{xx}}^{-1} (\sqrt{n}\bar{\mathbf{g}}) \\ &= \mathbf{z}_n' \mathbf{A}_n^{-1} \mathbf{z}_n, \end{aligned}$$

where

$$\mathbf{z}_n \equiv \mathbf{R} \mathbf{S}_{\mathbf{xx}}^{-1} (\sqrt{n}\bar{\mathbf{g}}), \quad \mathbf{A}_n \equiv s^2 \mathbf{R} \mathbf{S}_{\mathbf{xx}}^{-1} \mathbf{R}'.$$

By Assumption 2.2, $\text{plim } \mathbf{S}_{\mathbf{xx}} = \boldsymbol{\Sigma}_{\mathbf{xx}}$. By Assumption 2.5, $\sqrt{n}\bar{\mathbf{g}} \rightarrow_d N(\mathbf{0}, \mathbf{S})$. So by Lemma 2.4(c), we have:

$$\mathbf{z}_n \xrightarrow{d} N(\mathbf{0}, \mathbf{R}\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\mathbf{S}\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\mathbf{R}').$$

But, as shown in (2.6.4), $\mathbf{S} = \sigma^2\boldsymbol{\Sigma}_{\mathbf{xx}}$ under conditional homoekedasticity (Assumption 2.7). So the expression for the variance of the limiting distribution above becomes

$$\mathbf{R}\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\mathbf{S}\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\mathbf{R}' = \sigma^2\mathbf{R}\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\mathbf{R}' \equiv \mathbf{A}.$$

Thus we have shown:

$$\mathbf{z}_n \xrightarrow{d} \mathbf{z}, \mathbf{z} \sim N(\mathbf{0}, \mathbf{A}).$$

As already observed, $\mathbf{S}_{\mathbf{xx}} \rightarrow_p \boldsymbol{\Sigma}_{\mathbf{xx}}$. By Assumption 2.7, $\sigma^2 = E(\varepsilon_i^2)$. So by Proposition 2.2, $s^2 \rightarrow_p \sigma^2$. Thus by Lemma 2.3(a) (the ‘‘Continuous Mapping Theorem’’), $\mathbf{A}_n \rightarrow_p \mathbf{A}$. Therefore, by Lemma 2.4(d),

$$\mathbf{z}'_n \mathbf{A}_n^{-1} \mathbf{z}_n \xrightarrow{d} \mathbf{z}' \mathbf{A}^{-1} \mathbf{z}.$$

But since $\text{Var}(\mathbf{z}) = \mathbf{A}$, the distribution of $\mathbf{z}' \mathbf{A}^{-1} \mathbf{z}$ is chi-squared with $\#\mathbf{z}$ degrees of freedom.

6. For simplicity, we assumed in Section 2.8 that $\{y_i, \mathbf{x}_i\}$ is i.i.d. Collecting all the assumptions made in Section 2.8,

- (i) (linearity) $y_i = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i$.
- (ii) (random sample) $\{y_i, \mathbf{x}_i\}$ is i.i.d.
- (iii) (rank condition) $E(\mathbf{x}_i \mathbf{x}'_i)$ is non-singular.
- (iv) $E(\varepsilon_i^2 \mathbf{x}_i \mathbf{x}'_i)$ is non-singular.
- (v) (stronger version of orthogonality) $E(\varepsilon_i | \mathbf{x}_i) = 0$ (see (2.8.5)).
- (vi) (parameterized conditional heteroskedasticity) $E(\varepsilon_i^2 | \mathbf{x}_i) = \mathbf{z}'_i \boldsymbol{\alpha}$.

These conditions together are stronger than Assumptions 2.1-2.5.

(a) We wish to verify Assumptions 2.1-2.3 for the regression equation (2.8.8). Clearly, Assumption 2.1 about the regression equation (2.8.8) is satisfied by (i) about the *original* regression. Assumption 2.2 about (2.8.8) (that $\{\varepsilon_i^2, \mathbf{x}_i\}$ is ergodic stationary) is satisfied by (i) and (ii). To see that Assumption 2.3 about (2.8.8) (that $E(\mathbf{z}_i \eta_i) = \mathbf{0}$) is satisfied, note first that $E(\eta_i | \mathbf{x}_i) = 0$ by construction. Since \mathbf{z}_i is a function of \mathbf{x}_i , we have $E(\eta_i | \mathbf{z}_i) = 0$ by the Law of Iterated Expectation. Therefore, Assumption 2.3 is satisfied.

The additional assumption needed for (2.8.8) is Assumption 2.4 that $E(\mathbf{z}_i \mathbf{z}'_i)$ be non-singular. With Assumptions 2.1-2.4 satisfied for (2.8.8), the OLS estimator $\tilde{\boldsymbol{\alpha}}$ is consistent by Proposition 2.1(a) applied to (2.8.8).

- (b) Note that $\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}} = (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) - (\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha})$ and use the hint.
- (c) Regarding the first term of (**), by Kolmogorov’s LLN, the sample mean in that term converges in probability to $E(x_i \varepsilon_i \mathbf{z}_i)$ provided this population mean exists. But

$$E(x_i \varepsilon_i \mathbf{z}_i) = E[\mathbf{z}_i \cdot x_i \cdot E(\varepsilon_i | \mathbf{z}_i)].$$

By (v) (that $E(\varepsilon_i | \mathbf{x}_i) = 0$) and the Law of Iterated Expectations, $E(\varepsilon_i | \mathbf{z}_i) = 0$. Thus $E(x_i \varepsilon_i \mathbf{z}_i) = \mathbf{0}$. Furthermore, $\text{plim}(b - \beta) = 0$ since b is consistent when Assumptions

2.1-2.4 (which are implied by Assumptions (i)-(vi) above) are satisfied for the original regression. Therefore, the first term of (**) converges in probability to zero.

Regarding the second term of (**), the sample mean in that term converges in probability to $E(x_i^2 \mathbf{z}_i)$ provided this population mean exists. Then the second term converges in probability to zero because $\text{plim}(b - \beta) = 0$.

(d) Multiplying both sides of (*) by \sqrt{n} ,

$$\begin{aligned} \sqrt{n}(\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}) &= \left(\frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i' \right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{z}_i \cdot v_i \\ &= \left(\frac{1}{n} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i' \right)^{-1} \left[-2\sqrt{n}(b - \beta) \frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i \mathbf{z}_i + \sqrt{n}(b - \beta) \cdot (b - \beta) \frac{1}{n} \sum_{i=1}^n x_i^2 \mathbf{z}_i \right]. \end{aligned}$$

Under Assumptions 2.1-2.5 for the original regression (which are implied by Assumptions (i)-(vi) above), $\sqrt{n}(b - \beta)$ converges in distribution to a random variable. As shown in (c), $\frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i \mathbf{z}_i \rightarrow_p \mathbf{0}$. So by Lemma 2.4(b) the first term in the brackets vanishes (converges to zero in probability). As shown in (c), $(b - \beta) \frac{1}{n} \sum_{i=1}^n x_i^2 \mathbf{z}_i$ vanishes provided $E(x_i^2 \mathbf{z}_i)$ exists and is finite. So by Lemma 2.4(b) the second term, too, vanishes. Therefore, $\sqrt{n}(\hat{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}})$ vanishes, provided that $E(\mathbf{z}_i \mathbf{z}_i')$ is non-singular.

7. This exercise is about the model in Section 2.8, so we continue to maintain Assumptions (i)-(vi) listed in the solution to the previous exercise. Given the hint, the only thing to show is that the LHS of (**) equals $\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1} \mathbf{S} \boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}$, or more specifically, that $\text{plim} \frac{1}{n} \mathbf{X}' \mathbf{V} \mathbf{X} = \mathbf{S}$. Write \mathbf{S} as

$$\begin{aligned} \mathbf{S} &= E(\varepsilon_i^2 \mathbf{x}_i \mathbf{x}_i') \\ &= E[E(\varepsilon_i^2 | \mathbf{x}_i) \mathbf{x}_i \mathbf{x}_i'] \\ &= E(\mathbf{z}_i' \boldsymbol{\alpha} \mathbf{x}_i \mathbf{x}_i') \quad (\text{since } E(\varepsilon_i^2 | \mathbf{x}_i) = \mathbf{z}_i' \boldsymbol{\alpha} \text{ by (vi)}). \end{aligned}$$

Since \mathbf{x}_i is i.i.d. by (ii) and since \mathbf{z}_i is a function of \mathbf{x}_i , $\mathbf{z}_i' \boldsymbol{\alpha} \mathbf{x}_i \mathbf{x}_i'$ is i.i.d. So its sample mean converges in probability to its population mean $E(\mathbf{z}_i' \boldsymbol{\alpha} \mathbf{x}_i \mathbf{x}_i')$, which equals \mathbf{S} . The sample mean can be written as

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^n \mathbf{z}_i' \boldsymbol{\alpha} \mathbf{x}_i \mathbf{x}_i' \\ &= \frac{1}{n} \sum_{i=1}^n v_i \mathbf{x}_i \mathbf{x}_i' \quad (\text{by the definition of } v_i, \text{ where } v_i \text{ is the } i\text{-th diagonal element of } \mathbf{V}) \\ &= \frac{1}{n} \mathbf{X}' \mathbf{V} \mathbf{X}. \end{aligned}$$

8. See the hint.

9. (a)

$$\begin{aligned} &E(g_t | g_{t-1}, g_{t-2}, \dots, g_2) \\ &= E[E(g_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1) | g_{t-1}, g_{t-2}, \dots, g_2] \quad (\text{by the Law of Iterated Expectations}) \\ &= E[E(\varepsilon_t \cdot \varepsilon_{t-1} | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1) | g_{t-1}, g_{t-2}, \dots, g_2] \\ &= E[\varepsilon_{t-1} E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1) | g_{t-1}, g_{t-2}, \dots, g_2] \quad (\text{by the linearity of conditional expectations}) \\ &= 0 \quad (\text{since } E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1) = 0). \end{aligned}$$

(b)

$$\begin{aligned}
\mathbf{E}(g_t^2) &= \mathbf{E}(\varepsilon_t^2 \cdot \varepsilon_{t-1}^2) \\
&= \mathbf{E}[\mathbf{E}(\varepsilon_t^2 \cdot \varepsilon_{t-1}^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1)] \quad (\text{by the Law of Total Expectations}) \\
&= \mathbf{E}[\mathbf{E}(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1) \varepsilon_{t-1}^2] \quad (\text{by the linearity of conditional expectations}) \\
&= \mathbf{E}(\sigma^2 \varepsilon_{t-1}^2) \quad (\text{since } \mathbf{E}(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1) = \sigma^2) \\
&= \sigma^2 \mathbf{E}(\varepsilon_{t-1}^2).
\end{aligned}$$

But

$$\mathbf{E}(\varepsilon_{t-1}^2) = \mathbf{E}[\mathbf{E}(\varepsilon_{t-1}^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1)] = \mathbf{E}(\sigma^2) = \sigma^2.$$

- (c) If $\{\varepsilon_t\}$ is ergodic stationary, then $\{\varepsilon_t \cdot \varepsilon_{t-1}\}$ is ergodic stationary (see, e.g., Remark 5.3 on p. 488 of S. Karlin and H. Taylor, *A First Course in Stochastic Processes*, 2nd. ed., Academic Press, 1975, which states that “For any function ϕ , the sequence $Y_n = \phi(X_n, X_{n+1}, \dots)$ generates an ergodic stationary process whenever $\{X_n\}$ is ergodic stationary”). Thus the Billingsley CLT (see p. 106 of the text) is applicable to $\sqrt{n}\hat{\gamma}_1 = \sqrt{n} \frac{1}{n} \sum_{t=j+1}^n g_t$.
- (d) Since ε_t^2 is ergodic stationary, $\hat{\gamma}_0$ converges in probability to $\mathbf{E}(\varepsilon_t^2) = \sigma^2$. As shown in (c), $\sqrt{n}\hat{\gamma}_1 \rightarrow_d N(0, \sigma^4)$. So by Lemma 2.4(c) $\sqrt{n} \frac{\hat{\gamma}_1}{\hat{\gamma}_0} \rightarrow_d N(0, 1)$.

10. (a) Clearly, $\mathbf{E}(y_t) = 0$ for all $t = 1, 2, \dots$

$$\text{Cov}(y_t, y_{t-j}) = \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2 & \text{for } j = 0 \\ (\theta_1 + \theta_1\theta_2)\sigma_\varepsilon^2 & \text{for } j = 1, \\ \theta_2\sigma_\varepsilon^2 & \text{for } j = 2, \\ 0 & \text{for } j > 2, \end{cases}$$

So neither $\mathbf{E}(y_t)$ nor $\text{Cov}(y_t, y_{t-j})$ depends on t .

(b)

$$\begin{aligned}
&\mathbf{E}(y_t | y_{t-j}, y_{t-j-1}, \dots, y_0, y_{-1}) \\
&= \mathbf{E}(y_t | \varepsilon_{t-j}, \varepsilon_{t-j-1}, \dots, \varepsilon_0, \varepsilon_{-1}) \quad (\text{as noted in the hint}) \\
&= \mathbf{E}(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} | \varepsilon_{t-j}, \varepsilon_{t-j-1}, \dots, \varepsilon_0, \varepsilon_{-1}) \\
&= \begin{cases} \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} & \text{for } j = 0, \\ \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} & \text{for } j = 1, \\ \theta_2\varepsilon_{t-2} & \text{for } j = 2, \\ 0 & \text{for } j > 2, \end{cases}
\end{aligned}$$

which gives the desired result.

(c)

$$\begin{aligned}
\text{Var}(\sqrt{n}\bar{y}) &= \frac{1}{n} [\text{Cov}(y_1, y_1 + \cdots + y_n) + \cdots + \text{Cov}(y_n, y_1 + \cdots + y_n)] \\
&= \frac{1}{n} [(\gamma_0 + \gamma_1 + \cdots + \gamma_{n-2} + \gamma_{n-1}) + (\gamma_1 + \gamma_0 + \gamma_1 + \cdots + \gamma_{n-2}) \\
&\quad + \cdots + (\gamma_{n-1} + \gamma_{n-2} + \cdots + \gamma_1 + \gamma_0)] \\
&= \frac{1}{n} [n\gamma_0 + 2(n-1)\gamma_1 + \cdots + 2(n-j)\gamma_j + \cdots + 2\gamma_{n-1}] \\
&= \gamma_0 + 2 \sum_{j=1}^{n-1} \left(1 - \frac{j}{n}\right) \gamma_j.
\end{aligned}$$

(This is just reproducing (6.5.2) of the book.) Since $\gamma_j = 0$ for $j > 2$, one obtains the desired result.

(d) To use Lemma 2.1, one sets $z_n = \sqrt{n}\bar{y}$. However, Lemma 2.1, as stated in the book, inadvertently misses the required condition that there exist an $M > 0$ such that $E(|z_n|^{s+\delta}) < M$ for all n for some $\delta > 0$. Provided this technical condition is satisfied, the variance of the limiting distribution of $\sqrt{n}\bar{y}$ is the limit of $\text{Var}(\sqrt{n}\bar{y})$, which is $\gamma_0 + 2(\gamma_1 + \gamma_2)$.

11. (a) In the auxiliary regression, the vector of the dependent variable is \mathbf{e} and the matrix of regressors is $[\mathbf{X} \ \dot{\mathbf{E}}]$. Using the OLS formula,

$$\hat{\boldsymbol{\alpha}} = \hat{\mathbf{B}}^{-1} \begin{bmatrix} \frac{1}{n} \mathbf{X}' \mathbf{e} \\ \frac{1}{n} \mathbf{E}' \mathbf{e} \end{bmatrix}.$$

$\mathbf{X}' \mathbf{e} = \mathbf{0}$ by the normal equations for the original regression. The j -th element of $\frac{1}{n} \mathbf{E}' \mathbf{e}$ is

$$\frac{1}{n} (e_{j+1}e_1 + \cdots + e_n e_{n-j}) = \frac{1}{n} \sum_{t=j+1}^n e_t e_{t-j}.$$

which equals $\hat{\gamma}_j$ defined in (2.10.9).

(b) The j -th column of $\frac{1}{n} \mathbf{X}' \mathbf{E}$ is $\frac{1}{n} \sum_{t=j+1}^n \mathbf{x}_t \cdot e_{t-j}$ (which, incidentally, equals $\bar{\mu}_j$ defined on p. 147 of the book). Rewrite it as follows.

$$\begin{aligned}
&\frac{1}{n} \sum_{t=j+1}^n \mathbf{x}_t \cdot e_{t-j} \\
&= \frac{1}{n} \sum_{t=j+1}^n \mathbf{x}_t (\varepsilon_{t-j} - \mathbf{x}'_{t-j} (\mathbf{b} - \boldsymbol{\beta})) \\
&= \frac{1}{n} \sum_{t=j+1}^n \mathbf{x}_t \cdot \varepsilon_{t-j} - \left(\frac{1}{n} \sum_{t=j+1}^n \mathbf{x}_t \mathbf{x}'_{t-j} \right) (\mathbf{b} - \boldsymbol{\beta})
\end{aligned}$$

The last term vanishes because \mathbf{b} is consistent for $\boldsymbol{\beta}$. Thus $\frac{1}{n} \sum_{t=j+1}^n \mathbf{x}_t \cdot e_{t-j}$ converges in probability to $E(\mathbf{x}_t \cdot \varepsilon_{t-j})$.

The (i, j) element of the symmetric matrix $\frac{1}{n} \mathbf{E}' \mathbf{E}$ is, for $i \geq j$,

$$\frac{1}{n} (e_{1+i-j}e_1 + \cdots + e_{n-j}e_{n-i}) = \frac{1}{n} \sum_{t=1+i-j}^{n-j} e_t e_{t-(i-j)}.$$

Using the relation $e_t = \varepsilon_t - \mathbf{x}'_t(\mathbf{b} - \boldsymbol{\beta})$, this can be rewritten as

$$\begin{aligned} \frac{1}{n} \sum_{t=1+i-j}^{n-j} \varepsilon_t \varepsilon_{t-(i-j)} - \frac{1}{n} \sum_{t=1+i-j}^{n-j} (\mathbf{x}_t \varepsilon_{t-(i-j)} + \mathbf{x}_{t-(i-j)} \varepsilon_t)' (\mathbf{b} - \boldsymbol{\beta}) \\ - (\mathbf{b} - \boldsymbol{\beta})' \left(\frac{1}{n} \sum_{t=1+i-j}^{n-j} \mathbf{x}_t \mathbf{x}'_{t-(i-j)} \right) (\mathbf{b} - \boldsymbol{\beta}). \end{aligned}$$

The type of argument that is by now routine (similar to the one used on p. 145 for (2.10.10)) shows that this expression converges in probability to γ_{i-j} , which is σ^2 for $i = j$ and zero for $i \neq j$.

(c) As shown in (b), $\text{plim } \widehat{\mathbf{B}} = \mathbf{B}$. Since $\boldsymbol{\Sigma}_{\mathbf{xx}}$ is non-singular, \mathbf{B} is non-singular. So $\widehat{\mathbf{B}}^{-1}$ converges in probability to \mathbf{B}^{-1} . Also, using an argument similar to the one used in (b) for showing that $\text{plim } \frac{1}{n} \mathbf{E}' \mathbf{E} = \mathbf{I}_p$, we can show that $\text{plim } \widehat{\boldsymbol{\gamma}} = \mathbf{0}$. Thus the formula in (a) shows that $\widehat{\boldsymbol{\alpha}}$ converges in probability to zero.

(d) (The hint should have been: " $\frac{1}{n} \mathbf{E}' \mathbf{e} = \widehat{\boldsymbol{\gamma}}$. Show that $\frac{SSR}{n} = \frac{1}{n} \mathbf{e}' \mathbf{e} - \widehat{\boldsymbol{\alpha}}' \begin{bmatrix} \mathbf{0} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix}$." The SSR from the auxiliary regression can be written as

$$\begin{aligned} \frac{1}{n} SSR &= \frac{1}{n} (\mathbf{e} - [\mathbf{X} : \mathbf{E}] \widehat{\boldsymbol{\alpha}})' (\mathbf{e} - [\mathbf{X} : \mathbf{E}] \widehat{\boldsymbol{\alpha}}) \\ &= \frac{1}{n} (\mathbf{e} - [\mathbf{X} : \mathbf{E}] \widehat{\boldsymbol{\alpha}})' \mathbf{e} \quad (\text{by the normal equation for the auxiliary regression}) \\ &= \frac{1}{n} \mathbf{e}' \mathbf{e} - \frac{1}{n} \widehat{\boldsymbol{\alpha}}' [\mathbf{X} : \mathbf{E}]' \mathbf{e} \\ &= \frac{1}{n} \mathbf{e}' \mathbf{e} - \widehat{\boldsymbol{\alpha}}' \begin{bmatrix} \frac{1}{n} \mathbf{X}' \mathbf{e} \\ \frac{1}{n} \mathbf{E}' \mathbf{e} \end{bmatrix} \\ &= \frac{1}{n} \mathbf{e}' \mathbf{e} - \widehat{\boldsymbol{\alpha}}' \begin{bmatrix} \mathbf{0} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix} \quad (\text{since } \mathbf{X}' \mathbf{e} = \mathbf{0} \text{ and } \frac{1}{n} \mathbf{E}' \mathbf{e} = \widehat{\boldsymbol{\gamma}}). \end{aligned}$$

As shown in (c), $\text{plim } \widehat{\boldsymbol{\alpha}} = \mathbf{0}$ and $\text{plim } \widehat{\boldsymbol{\gamma}} = \mathbf{0}$. By Proposition 2.2, we have $\text{plim } \frac{1}{n} \mathbf{e}' \mathbf{e} = \sigma^2$. Hence SSR/n (and therefore $SSR/(n - K - p)$) converges to σ^2 in probability.

(e) Let

$$\mathbf{R} \equiv \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{I}_p \\ (p \times K) & & \end{bmatrix}, \quad \mathbf{V} \equiv [\mathbf{X} : \mathbf{E}].$$

The F -ratio is for the hypothesis that $\mathbf{R}\boldsymbol{\alpha} = \mathbf{0}$. The F -ratio can be written as

$$F = \frac{(\mathbf{R}\widehat{\boldsymbol{\alpha}})' [\mathbf{R}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{R}]^{-1} (\mathbf{R}\widehat{\boldsymbol{\alpha}})/p}{SSR/(n - K - p)}. \quad (*)$$

Using the expression for $\hat{\boldsymbol{\alpha}}$ in (a) above, $\mathbf{R}\hat{\boldsymbol{\alpha}}$ can be written as

$$\begin{aligned}\mathbf{R}\hat{\boldsymbol{\alpha}} &= \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{I}_p \\ (p \times K) & & \end{bmatrix} \hat{\mathbf{B}}^{-1} \begin{bmatrix} \mathbf{0} \\ (K \times 1) \\ \hat{\boldsymbol{\gamma}} \\ (p \times 1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{I}_p \\ (p \times K) & & \end{bmatrix} \begin{bmatrix} \hat{\mathbf{B}}^{11} & \hat{\mathbf{B}}^{12} \\ (K \times K) & (K \times p) \\ \hat{\mathbf{B}}^{21} & \hat{\mathbf{B}}^{22} \\ (p \times K) & (p \times p) \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ (K \times 1) \\ \hat{\boldsymbol{\gamma}} \\ (p \times 1) \end{bmatrix} \\ &= \hat{\mathbf{B}}^{22} \hat{\boldsymbol{\gamma}}. \end{aligned} \quad (**)$$

Also, $\mathbf{R}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{R}'$ in the expression for F can be written as

$$\begin{aligned}\mathbf{R}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{R}' &= \frac{1}{n} \mathbf{R} \hat{\mathbf{B}}^{-1} \mathbf{R}' \quad (\text{since } \frac{1}{n} \mathbf{V}'\mathbf{V} = \hat{\mathbf{B}}) \\ &= \frac{1}{n} \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{I}_p \\ (p \times K) & & \end{bmatrix} \begin{bmatrix} \hat{\mathbf{B}}^{11} & \hat{\mathbf{B}}^{12} \\ (K \times K) & (K \times p) \\ \hat{\mathbf{B}}^{21} & \hat{\mathbf{B}}^{22} \\ (p \times K) & (p \times p) \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ (K \times p) \\ \mathbf{I}_p \end{bmatrix} \\ &= \frac{1}{n} \hat{\mathbf{B}}^{22}. \end{aligned} \quad (***)$$

Substitution of (***) and (**) into (*) produces the desired result.

- (f) Just apply the formula for partitioned inverses.
- (g) Since $\sqrt{n}\hat{\boldsymbol{\rho}} - \sqrt{n}\hat{\boldsymbol{\gamma}}/\sigma^2 \rightarrow_p \mathbf{0}$ and $\hat{\boldsymbol{\Phi}} \rightarrow_p \boldsymbol{\Phi}$, it should be clear that the modified Box-Pierce $Q (= n \cdot \hat{\boldsymbol{\rho}}'(\mathbf{I}_p - \hat{\boldsymbol{\Phi}})^{-1}\hat{\boldsymbol{\rho}})$ is asymptotically equivalent to $n\hat{\boldsymbol{\gamma}}'(\mathbf{I}_p - \boldsymbol{\Phi})^{-1}\hat{\boldsymbol{\gamma}}/\sigma^4$. Regarding the pF statistic given in (e) above, consider the expression for $\hat{\mathbf{B}}^{22}$ given in (f) above. Since the j -th element of $\frac{1}{n}\mathbf{X}'\mathbf{E}$ is $\bar{\boldsymbol{\mu}}_j$ defined right below (2.10.19) on p. 147, we have

$$s^2 \hat{\boldsymbol{\Phi}} = \left(\frac{1}{n} \mathbf{E}'\mathbf{X} \right) \mathbf{S}_{\mathbf{xx}}^{-1} \left(\frac{1}{n} \mathbf{X}'\mathbf{E} \right),$$

so

$$\hat{\mathbf{B}}^{22} = \left[\frac{1}{n} \mathbf{E}'\mathbf{E} - s^2 \hat{\boldsymbol{\Phi}} \right]^{-1}.$$

As shown in (b), $\frac{1}{n}\mathbf{E}'\mathbf{E} \rightarrow_p \sigma^2 \mathbf{I}_p$. Therefore, $\hat{\mathbf{B}}^{22} \rightarrow_p \frac{1}{\sigma^2}(\mathbf{I}_p - \boldsymbol{\Phi})^{-1}$, and pF is asymptotically equivalent to $n\hat{\boldsymbol{\gamma}}'(\mathbf{I}_p - \boldsymbol{\Phi})^{-1}\hat{\boldsymbol{\gamma}}/\sigma^4$.

12. The hints are almost as good as the answer. Here, we give solutions to (b) and (c) only.

- (b) We only prove the first convergence result.

$$\frac{1}{n} \sum_{t=1}^r \mathbf{x}_t \mathbf{x}_t' = \frac{r}{n} \left(\frac{1}{r} \sum_{t=1}^r \mathbf{x}_t \mathbf{x}_t' \right) = \lambda \left(\frac{1}{r} \sum_{t=1}^r \mathbf{x}_t \mathbf{x}_t' \right).$$

The term in parentheses converges in probability to $\boldsymbol{\Sigma}_{\mathbf{xx}}$ as n (and hence r) goes to infinity.

- (c) We only prove the first convergence result.

$$\frac{1}{\sqrt{n}} \sum_{t=1}^r \mathbf{x}_t \cdot \varepsilon_t = \sqrt{\frac{r}{n}} \left(\frac{1}{\sqrt{r}} \sum_{t=1}^r \mathbf{x}_t \cdot \varepsilon_t \right) = \sqrt{\lambda} \left(\frac{1}{\sqrt{r}} \sum_{t=1}^r \mathbf{x}_t \cdot \varepsilon_t \right).$$

The term in parentheses converges in distribution to $N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}_{\mathbf{xx}})$ as n (and hence r) goes to infinity. So the whole expression converges in distribution to $N(\mathbf{0}, \lambda \sigma^2 \boldsymbol{\Sigma}_{\mathbf{xx}})$.