10.3

In the regression model $y = \beta_1 + \beta_2 x + e$, assume x is endogenous and that z is a valid instrument. In Section 10.3.5, we saw that $\beta_2 = \text{cov}(z, y)/\text{cov}(z, x)$.

a. Divide the denominator of $\beta_2 = \text{cov}(z, y)/\text{cov}(z, x)$ by $\text{var}(z)$. Show that $\text{cov}(z, x)/\text{var}(z)$ is the coefficient of the simple regression with dependent variable x and explanatory variable z, $x = \gamma_1 + \theta_1 z + v$. [Hint: See Section 10.2.1.] Note that this is the first-stage equation in two-stage least squares.

Ans.

$(1):\quad x = \gamma_1 + \theta_1 z + v$

對 (1) 取期望值，得 $(2):\quad E(x) = \gamma_1 + \theta_1 E(z)$

$(1)-(2):\quad x - E(x) = \theta_1(z - E(z)) + v$

等式兩邊同乘以 $(z - E(z))$: $\quad (x - E(x))(z - E(z)) = \theta_1(z - E(z))^2 + (z - E(z))v$

等式兩邊取期望值： $\quad E[(x - E(x))(z - E(z))] = \theta_1 E[(z - E(z))^2] + E[(z - E(z))v] \quad ---(*)$

$(*)$ 等式左邊：$E[(x - E(x))(z - E(z))] = \text{cov}(z, x)$

$(*)$ 等式右邊第一項：$E[(z - E(z))^2] = \text{var}(z)$

$(*)$ 等式右邊第二項：$E[(z - E(z))v] = 0$

因為 z 是工具變數，與第一階段的誤差項 v 不相關（2SLS 中工具變數的假設之一：$\text{cov}(z, v) = 0$）

又 $\text{cov}(z, v) = E[(z - E(z))(v - E(v))] = E[(z - E(z))v] \quad (\because E[v] = 0) \quad \therefore \text{cov}(z, v) = E[(z - E(z))v] = 0$

因此，式子 $(*)$ 可簡化為：$\text{cov}(z, x) = \theta_1 \text{var}(z) + 0$

$\theta_1 = \dfrac{\text{cov}(z, x)}{\text{var}(z)}$，也就是簡單回歸 $x = \gamma_1 + \theta_1 z + v$ 的回歸係數。

這也與普通最小平方法（OLS）估計簡單回歸係數的標準公式一致。

b. Divide the numerator of $\beta_2 = \text{cov}(z, y)/\text{cov}(z, x)$ by $\text{var}(z)$. Show that $\text{cov}(z, y)/\text{var}(z)$ is the coefficient of a simple regression with dependent variable y and explanatory variable z, $y = \pi_0 + \pi_1 z + u$. [Hint: See Section 10.2.1.]

Ans.

$(1):\quad y = \pi_0 + \pi_1 z + u$

對 (1) 取期望值，得 $(2):\quad E(y) = \pi_0 + \pi_1 E(z)$

$(1)-(2):\quad y - E(y) = \pi_1(z - E(z)) + u$

等式兩邊同乘以 $(z - E(z))$: $\quad (y - E(y))(z - E(z)) = \pi_1(z - E(z))^2 + (z - E(z))u$

等式兩邊取期望值：

$$\begin{aligned} E[(y - E(y))(z - E(z))] &= \pi_1 E[(z - E(z))^2] + E[(z - E(z))u] \\ &= \pi_1 E[(z - E(z))^2] \end{aligned}$$

$\pi_1 = \dfrac{E[(y - E(y))(z - E(z))]}{E[(z - E(z))^2]} = \dfrac{\text{cov}(z, y)}{\text{var}(z)}$，也就是簡單回歸 $y = \pi_0 + \pi_1 z + u$ 的回歸係數。

c. In the model $y = \beta_1 + \beta_2 x + e$, substitute for x using $x = \gamma_1 + \theta_1 z + v$ and simplify to obtain $y = \pi_0 + \pi_1 z + u$. What are π_0, π_1, and u in terms of the regression model parameters and error and the first-stage parameters and error? The regression you have obtained is a <u>reduced-form</u> equation.

Ans.

$y = \beta_1 + \beta_2 x + e = \beta_1 + \beta_2(\gamma_1 + \theta_1 z + v) + e = (\beta_1 + \beta_2\gamma_1) + \beta_2\theta_1 z + (\beta_2 v + e) = \pi_0 + \pi_1 z + u$

$\therefore \pi_0 = (\beta_1 + \beta_2\gamma_1), \quad \pi_1 = \beta_2\theta_1, \quad u = \beta_2 v + e$

d. Show that $\beta_2 = \pi_1/\theta_1$.

Ans. $\because \pi_1 = \beta_2\theta_1 \quad \therefore \beta_2 = \pi_1/\theta_1$

e. If $\hat{\pi}_1$ and $\hat{\theta}_1$ are the OLS estimators of π_1 and θ_1, show that $\hat{\beta}_2 = \hat{\pi}_1/\hat{\theta}_1$ is a consistent estimator of $\beta_2 = \pi_1/\theta_1$. The estimator $\hat{\beta}_2 = \hat{\pi}_1/\hat{\theta}_1$ is an <u>indirect least squares</u> estimator.

Ans.

從 (a)：θ_1 的 OLS 估計為：$\hat{\theta}_1 = \dfrac{\widehat{\mathrm{Cov}}(z,x)}{\widehat{\mathrm{Var}}(z)} = \dfrac{\sum(z_i-\bar{z})(x_i-\bar{x})/N}{\sum(z_i-\bar{z})^2/N} = \dfrac{\sum(z_i-\bar{z})(x_i-\bar{x})}{\sum(z_i-\bar{z})^2}$

只要 z 與 v 不相關，$\hat{\theta}_1$ 就是 θ_1 的一致估計量。

從 (b)：π_1 的 OLS 估計為：$\hat{\pi}_1 = \dfrac{\widehat{\mathrm{Cov}}(z,y)}{\widehat{\mathrm{Var}}(z)} = \dfrac{\sum(z_i-\bar{z})(y_i-\bar{y})/N}{\sum(z_i-\bar{z})^2/N} = \dfrac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})^2}$

只要 z 與 u 不相關，$\hat{\pi}_1$ 就是 π_1 的一致估計量。

IV Estimator: $\hat{\beta}_2 = \dfrac{\hat{\pi}_1}{\hat{\theta}_1} = \dfrac{\frac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})^2}}{\frac{\sum(z_i-\bar{z})(x_i-\bar{x})}{\sum(z_i-\bar{z})^2}} = \dfrac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})(x_i-\bar{x})} = \dfrac{\sum(z_i-\bar{z})(y_i-\bar{y})/N}{\sum(z_i-\bar{z})(x_i-\bar{x})/N} = \dfrac{\widehat{\mathrm{Cov}}(z,y)}{\widehat{\mathrm{Cov}}(z,x)}$

樣本共變異數會以機率趨近（converge in probability）於母體共變異數：

$\because\ \widehat{\mathrm{cov}}(z,y) \xrightarrow{p} \mathrm{cov}(z,y) \quad \widehat{\mathrm{cov}}(z,x) \xrightarrow{p} \mathrm{cov}(z,x)$

$\therefore\ \hat{\beta}_2 = \dfrac{\hat{\pi}_1}{\hat{\theta}_1} = \dfrac{\widehat{\mathrm{Cov}}(z,y)}{\widehat{\mathrm{Cov}}(z,x)} \xrightarrow{p} \dfrac{\mathrm{Cov}(z,y)}{\mathrm{Cov}(z,x)} = \dfrac{\pi_1}{\theta_1} = \beta_2$

$\hat{\beta}_2$ 是 β_2 的一致估計量，且 $\hat{\beta}_2 = \dfrac{\hat{\pi}_1}{\hat{\theta}_1}$ 是 $\beta_2 = \dfrac{\pi_1}{\theta_1}$ 的間接估計，這稱為 indirect least squares (ILS) 估計量。

前提條件：

1. $\theta_1 \neq 0$，即 z 與 x 相關（工具變數的相關性條件）
2. z 與 v 不相關（第一階段外生性條件）
3. z 與 u 不相關（簡化式外生性條件）