10.2 The labor supply of married women has been a subject of a great deal of economic research. Consider the following supply equation specification

$$HOURS = \beta_1 + \beta_2 WAGE + \beta_3 EDUC + \beta_4 AGE + \beta_5 KIDSL6 + \beta_6 NWIFEINC + e$$

where *HOURS* is the supply of labor, *WAGE* is hourly wage, *EDUC* is years of education, *KIDSL6* is the number of children in the household who are less than 6 years old, and *NWIFEINC* is household income from sources other than the wife's employment.

a. Discuss the signs you expect for each of the coefficients.
b. Explain why this supply equation cannot be consistently estimated by OLS regression.
c. Suppose we consider the woman's labor market experience *EXPER* and its square, $EXPER^2$, to be instruments for *WAGE*. Explain how these variables satisfy the logic of instrumental variables.
d. Is the supply equation identified? Explain.
e. Describe the steps [not a computer command] you would take to obtain IV/2SLS estimates.

a. β_2(WAGE) → ⊕，工資提高，激勵更多勞動

β_3 (EDUC) → 不確定，教育程度愈高 → 能力強.有意願進入職場 → 增加勞動供給

→ 效率高，對時間彈性要求高，可能投入其他地方 → 減少勞動供給

β_4(AGE) → 不確定，年齡增加，可能增加經驗(提升供給)，也可能因為健康(降低供給)

β_5 (KIDSLE) → ⊖，家中兒童越多，負擔越重，供給下降

β_6(NWIFEINC) → ⊖，其他來源的收入高，減少女性進入勞動市場的誘因

b. Endogeneity 內生性問題

Hours ⟷ WAGES 都由供需決定，会造成估計失效

c. Instrumental variable

① Relevance：EXPER，$EXPER^2$ 和 WAGE 通常顯著相關

② Exogeneity：EXPER，$EXPER^2$ 透過 WAGE 影響 Hours，所以和誤差不相關

d. Yes，只有一個內生變數 (WAGE) 且至少有一個工具變數可使用

e. $WAGE = \gamma_1 + \gamma_2 EDUC + \gamma_3 AGE + \gamma_4 KIDSLE + \gamma_5 KIDS618 + \gamma_6 NWIFEINC + Q1\, EXPER + Q2\, EXPER^2 + u_e$

先对上面的式子回歸，得到 $\widehat{WAGE}$，將原來的 WAGE 用 $\widehat{WAGE}$ 替換，再用 OLS 估計

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.

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.]

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 **reduced-form** equation.

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

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 **indirect least squares** estimator.

a. $x = \gamma_1 + \theta_1 z + v \Rightarrow$ 取期望值 $E(x) = \gamma_1 + \theta_1 E(z)$

$$x = \gamma_1 + \theta_1 z + v$$
$$-)\ E(x) = \gamma_1 + \theta_1 E(z)$$
$$x - E(x) = \theta_1(z - E(z)) + v$$

同乘 $z - E(z) \Rightarrow (z - E(z))(x - E(x)) = \theta_1(z - E(z))^2 + v(z - E(z))$

同取期望值：$E[(z - E(z))(x - E(x))] = \theta_1 E((z - E(z))^2)$

$$\Rightarrow \theta_1 = \frac{Cov(z, x)}{Var(z)}$$

b. $y = \pi_0 + \pi_1 z + u \rightarrow$ 同 a. $E(y) = \pi_0 + \pi_1 E(z)$

$$\Rightarrow \quad y = \pi_0 + \pi_1 z + u$$
$$-)\ E(y) = \pi_0 + \pi_1 E(z)$$
$$y - E(y) = \pi_1(z - E(z)) + u$$

$\times(z - E(z)) \Rightarrow E[(y - E(y))(z - E(z))] = \pi_1 E[(z - E(z))^2]$
取期望值

$$\Rightarrow \pi_1 = \frac{Cov(z, y)}{Var(z)}$$

c. $x = \gamma_1 + \theta_1 z + v$ 代入 $y = \beta_1 + \beta_2 x + e$

$$y = \beta_1 + \beta_2(\gamma_1 + \theta_1 z + v) + e$$
$$= (\beta_1 + \beta_2\gamma_1) + \beta_2\theta_1 z + (\beta_2 v + e)$$

对照 reduced-form $y = \pi_0 + \pi_1 z + u$

$\Rightarrow \pi_0 = \beta_1 + \beta_2 \gamma_1$

$\pi_1 = \beta_2 \theta_1$

$u = \beta_2 v + e$

d. $\pi_1 = \beta_2 \theta_1 \Rightarrow \beta_2 = \dfrac{\pi_1}{\theta_1}$

e. by (a)

$$\hat{\theta}_1 = \frac{\widehat{cov}(z, x)}{\widehat{var}(z)} = \frac{\sum(z_i - \bar{z})(x_i - \bar{x})/N}{\sum(z_i - \bar{z})^2/N} = \frac{\sum(z_i - \bar{z})(x_i - \bar{x})}{\sum(z_i - \bar{z})^2}$$

by (b)

$$\hat{\pi}_1 = \frac{\widehat{cov}(z, y)}{\widehat{var}(z)} = \frac{\sum(z_i - \bar{z})(y_i - \bar{y})/N}{\sum(z_i - \bar{z})^2/N} = \frac{\sum(z_i - \bar{z})(y_i - \bar{y})}{\sum(z_i - \bar{z})^2}$$

$$\hat{\beta}_2 = \hat{\pi}_1 / \hat{\theta}_1 = \frac{\sum(z_i - \bar{z})(y_i - \bar{y})}{\sum(z_i - \bar{z})(x_i - \bar{x})} = \frac{\sum(z_i - \bar{z})(y_i - \bar{y})/N}{\sum(z_i - \bar{z})(x_i - \bar{x})/N} = \frac{\widehat{cov}(z, y)}{\widehat{cov}(z, x)}$$

$\Rightarrow$ 是 IV estimator

consistent

$\because \widehat{cov}(z, y) \xrightarrow{P} cov(z, y)$

$\widehat{cov}(z, x) \xrightarrow{P} cov(z, x)$

$$\hat{\beta}_2 = \frac{\hat{\pi}_1}{\hat{\theta}_1} = \frac{\widehat{cov}(z, y)}{\widehat{cov}(z, x)} \xrightarrow{P} \frac{cov(z, y)}{cov(z, x)} = \beta_2$$