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. 預期 β_2 為正 工資越高可以提升女性勞動參與

預期 β_3 為正 教育程度越高 通常有更高的工作能力，會有較高的工作意願

預期 β_4 為正，青壯年時期，隨著年齡增加，婦女可能提供更多的勞動時數，中老年時期，可能相反

預期 β_5 為負，家中有小孩的婦女應該會減少工作時數

預期 β_6 為負，家中有其它收入來源，可能會降低婦女的工作意願

b.

WAGE 可能與誤差項相關，因為工資可能會受到婦女未觀察到的個人特徵影響，這些特徵可能同時也影響 HOURS

c.

EXPER 和 EXPER 可作為 WAGE 的工具變數，因為 EXPER 和 WAGE 有正相關性 通常經驗豐富的人薪資更高，使得經驗可作為解釋工資的有效工具，且 EXPER 和 $EXPER^2$ 與誤差項不相關

d.

Yes. EXPER 和 $EXPER^2$ 雖然是有效的工具變數，但它們的關係較單一，可能無法完全捕捉 WAGE 的所有變異性。實際上，工具變數數量必須至少與內生解釋變數的數量相等才能達到識別

e.

1. 檢查 EXPER 和 $EXPER^2$ 是否符合相關性和外生條件
2. 建立第一階段回歸，將 WAGE 設為 dependent variable，EXPER 和 $EXPER^2$ 為 independent variable
3. 將第一階段預測的 WAGE 取代原本的 WAGE，進行第二階段的回歸
   $HOURS = \widehat{WAGE} + EDUC + AGE + KIDSL6 + NWIFEINC$
4. 檢查回歸結果，並驗證是否存在內生性問題，查看工具變數的有效性和檢測估計的有效性
5. 如果模型有效，報告 IV/2SLS 估計結果

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.

$$E(X) = \gamma_1 + \theta_1 E(Z)$$

$$X - E(X) = (\gamma_1 + \theta_1 Z + v) - (\gamma_1 + \theta_1 E(Z))$$

$$X - E(X) = \theta_1 (Z - E(Z)) + v$$

$$(X - E(X))(Z - E(Z)) = \theta_1 (Z - E(Z))^2 + v(Z - E(Z))$$

$$E(X - E(X))(Z - E(Z)) = \theta_1 E(Z - E(Z))^2 + \underbrace{E(Z - E(Z))v}_{设=0}$$

$$E(X - E(X))(Z - E(Z)) = \theta_1 E(Z - E(Z))^2$$

$$\theta_1 = \frac{E(X - E(X))(Z - E(Z))}{E(Z - E(Z))^2} = \frac{\text{cov}(Z, X)}{\text{Var}(Z)}$$

b.

$$Y = \pi_0 + \pi_1 Z + u$$

$$E(Y) = \pi_0 + \pi_1 E(Z)$$

$$Y - E(Y) = \pi_1 (Z - E(Z)) + u$$

$$(Y - E(Y))(Z - E(Z)) = \pi_1 (Z - E(Z))^2 + u(Z - E(Z))$$

$$E(Y - E(Y))(Z - E(Z)) = \pi_1 E(Z - E(Z))^2 + uE(Z - E(Z))$$

$$E(Y - E(Y))(Z - E(Z)) = \pi_1 E(Z - E(Z))^2$$

$$\pi_1 = \frac{E(Y - E(Y))(Z - E(Z))}{E(Z - E(Z))^2} = \frac{\text{cov}(Y, Z)}{\text{Var}(Z)}$$

c.

$$Y = \beta_1 + \beta_2 X + e$$

$$X = \gamma_1 + \theta_1 Z + V \quad (1)$$

$$Y: \pi_0 + \pi_1 Z + u \quad (2)$$

$$\begin{aligned} Y &= \beta_1 + \beta_2 X + e \\ &= \beta_1 + \beta_2(\gamma_1 + \theta_1 Z + V) + e \\ &= \beta_1 \beta_2 \gamma_1 + \beta_2 \theta_1 Z + \beta_2 V + e \\ &= \pi_0 + \pi_1 Z + u \end{aligned}$$

$$\therefore \pi_0 = \beta_1 \beta_2 \gamma_1 \quad \pi_1 = \beta_2 \theta_1 \quad u = \beta_2 V + e$$

d.

$$\pi_1 = \beta_2 \theta_1 \Rightarrow \beta_2 = \frac{\pi_1}{\theta_1}$$

e.

$$\hat{\pi}_1 = \frac{\widehat{\text{cov}}(Y, Z)}{\widehat{\text{var}}(Z)} = \frac{\Sigma(Y - \bar{Y})(Z - \bar{Z})/N}{\Sigma(Z - \bar{Z})^2/N} = \frac{\Sigma(Y - \bar{Y})(Z - \bar{Z})}{\Sigma(Z - \bar{Z})^2}$$

$$\hat{\theta}_1 = \frac{\widehat{\text{cov}}(X, Z)}{\widehat{\text{var}}(Z)} = \frac{\Sigma(X - \bar{X})(Z - \bar{Z})/N}{\Sigma(Z - \bar{Z})^2/N} = \frac{\Sigma(X - \bar{X})(Z - \bar{Z})}{\Sigma(Z - \bar{Z})^2}$$

∵ $\widehat{\text{cov}}(Y, Z)$ 是 $\text{cov}(Y, Z)$ 的 MME
$\widehat{\text{cov}}(X, Z)$ 是 $\text{cov}(X, Z)$ 的 MME
$\widehat{\text{var}}(Z)$ 是 $\text{var}(Z)$ 的 MME
MME 必满足一致性

∴ $\widehat{\text{cov}}(Y, Z) \xrightarrow{p} \text{cov}(Y, Z)$
$\widehat{\text{cov}}(X, Z) \xrightarrow{p} \text{cov}(X, Z)$
$\widehat{\text{var}}(Z) \xrightarrow{p} \text{var}(Z)$

故 $\hat{\pi}_1 \xrightarrow{p} \pi_1 \quad \hat{\theta}_1 \xrightarrow{p} \theta_1$

$$\hat{\beta}_2 = \frac{\hat{\pi}_1}{\hat{\theta}_1} \xrightarrow{p} \frac{\pi_1}{\theta_1} = \beta_2$$