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):** 正 (+)。工資越高，工作的誘因越大，因此勞動供給 (HOURS) 應該會上升。
- **β_3 (EDUC):** 正 (+)。教育程度高通常意味著更好的就業機會與更高的薪資，因此可能會增加勞動供給。
- **β_4 (AGE):** 不確定。年齡增加可能有兩種方向：一方面隨年齡增加，經驗累積可能增加勞動供給；另一方面，年齡過大可能偏向退休或減少工時。
- **β_5 (KIDS6):** 負 (-)。年幼孩子越多，照顧孩子的需求提高，通常會降低母親的勞動供給。
- **β_6 (NWIFEINC):** 負 (-)。家庭其他來源收入越高，妻子勞動的必要性降低，因此勞動供給應該下降。

(b) 可能有內生性問題，因為WAGE可能與誤差項 e (例如能力,偏好等) 產生內生性問題。

(c)

- **相關性 (Relevance):** EXPER 和 $EXPER^2$ 會影響 WAGE，因為工作經驗通常提高薪資，因此這些工具變數與被內生的解釋變數 (WAGE) 相關。
- **外生性 (Exogeneity):** 在合理假設下，EXPER 和 $EXPER^2$ 只透過影響 WAGE 間接影響 HOURS，本身不直接進入勞動供給方程式，因此可以認為與誤差項 e 不相關。

(d) Yes, 至少2個 I.V. 來對抗1個內生變數, 且這些 I.V. 是和 WAGE 有關, 但和誤差項無關, 因此滿足識別條件。

(e) step1: 用 IV ($EXPER$ & $EXPER^2$) 去回歸 WAGE
得到 $\widehat{WAGE}$
step2: 把 $\widehat{WAGE}$ 代入原本的 equation, 使用 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) $x = \gamma_1 + \theta_1 z + v$

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

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

$[x - E(x)][z - E(z)] = \theta_1 (z - E(z))^2 + v(z - E(z))$

$E[(x - E(x))(z - E(z))] = \theta_1 E[(z - E(z))^2]$

$$\theta_1 = \frac{E[(x - E(x))(z - E(z))]}{E[(z - E(z))^2]} = \frac{\text{cov}(z, x)}{\text{var}(z)}$$

(b) $y = \pi_0 + \pi_1 z + u$

$E(y) = \pi_0 + \pi_1 E(z)$

$y - E(y) = \pi_1(z - E(z))$

同(a) $\Rightarrow \pi_1 = \dfrac{\text{cov}(z, y)}{\text{var}(y)}$

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

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

$$y = \underbrace{(\beta_1 + \beta_2 \gamma_1)}_{\pi_0} + \underbrace{(\beta_2 \theta_1)}_{\pi_1} z + \underbrace{(\beta_2 v + e)}_{u}$$

(d) $\pi_1 = \beta_2 \theta_1$

$\beta_2 = \dfrac{\pi_1}{\theta_1}$

(e) $\hat{\theta}_1 = \dfrac{\widehat{\text{cov}}(z, x)}{\widehat{\text{var}}(z)}$, $\hat{\pi}_1 = \dfrac{\widehat{\text{cov}}(z, y)}{\widehat{\text{var}}(z)}$

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