HW0421

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.

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.

10.2

a.

- β2\beta_2β2（WAGE）：正向，因為工資越高，工作的誘因增加，勞動供給小時數上升。
- β3\beta_3β3（EDUC）：正向，教育程度越高，通常意味著潛在工資越高，勞動參與也越高。
- β4\beta_4β4（AGE）：可能是正向或負向，年齡增加經驗增加，但年紀過大也可能減少工時。
- β5\beta_5β5（KIDSL6）：負向，家中有小孩（尤其是 6 歲以下）會增加照顧責任，可能減少工時。
- β6\beta_6β6（NWIFEINC）：負向，家庭其他收入越高，降低了工作必要性，勞動供給減少。

b.

- WAGE 是內生的（endogenous），因為工資可能受未觀察到的個人特徵（如能力、動機）影響，這些特徵也同時影響工時。這會造成 OLS 估計的偏誤。

c.

- 使用 EXPER（經驗）與 EXPER^2（經驗平方）作為 WAGE 的工具變數是合理的：
  - 相關性（relevance）：工作經驗會影響工資。
  - 外生性（exogeneity）：假設經驗本身不直接影響工時，只是透過工資影響工時。

d.

- 是的，這個模型是可識別的（identified），因為有足夠的工具變數（EXPER 和 EXPER^2）來對應內生變數（WAGE），且工具變數數量不小於內生變數數量（兩個工具對應一個內生變數）。

e.

- 取得 IV/2SLS 估計量的步驟：
  1. 第一階段回歸：用工具變數（EXPER, EXPER^2）去回歸 WAGE，取得 WAGE 的預測值（WAGE^\widehat{WAGE}WAGE）。
  2. 第二階段回歸：以 WAGE^\widehat{WAGE}WAGE、EDUC、AGE、KIDSL6、NWIFEINC 為解釋變數，回歸 HOURS。
  3. 使用這個回歸的結果來估計各係數。

10.3

a.

- 將 β2=cov(z,y)cov(z,x)\beta_2 = \frac{\text{cov}(z, y)}{\text{cov}(z, x)}β2=cov(z,x)cov(z,y)，把分母 cov(z,x)\text{cov}(z, x)cov(z,x) 除以 var(z)\text{var}(z)var(z)，可得：

cov(z,x)var(z)=θ1\frac{\text{cov}(z, x)}{\text{var}(z)} = \theta_1var(z)cov(z,x)=θ1

這是以 z 為解釋變數、x 為被解釋變數的簡單回歸中的係數，這是第一階段回歸。

b.

- 將 cov(z,y)\text{cov}(z, y)cov(z,y) 除以 var(z)\text{var}(z)var(z)，可得：

  cov(z,y)var(z)=π1\frac{\text{cov}(z, y)}{\text{var}(z)} = \pi_1var(z)cov(z,y)=π1

  這是以 z 為解釋變數、y 為被解釋變數的簡單回歸中的係數，這是另一個回歸。

c.

- 用 x=γ1+θ1z+νx = \gamma_1 + \theta_1 z + \nux=γ1+θ1z+ν 代入 y=β1+β2x+ey = \beta_1 + \beta_2 x + ey=β1+β2x+e，得：

  y=β1+β2(γ1+θ1z+ν)+e=(β1+β2γ1)+β2θ1z+(β2ν+e)y = \beta_1 + \beta_2(\gamma_1 + \theta_1 z + \nu) + e = (\beta_1 + \beta_2 \gamma_1) + \beta_2 \theta_1 z + (\beta_2 \nu + e)y=β1+β2(γ1+θ1 z+ν)+e=(β1+β2γ1)+β2θ1z+(β2ν+e)

  設：

  τ0=β1+β2γ1,τ1=β2θ1,u=β2ν+e\tau_0 = \beta_1 + \beta_2 \gamma_1, \quad \tau_1 = \beta_2 \theta_1, \quad u = \beta_2 \nu + eτ0=β1+β2γ1,τ1 =β2θ1,u=β2ν+e

  所以新的「簡化方程式」（reduced-form equation）是：

  y=τ0+τ1z+uy = \tau_0 + \tau_1 z + uy=τ0+τ1z+u

d.

- 因為：

  β2=π1θ1\beta_2 = \frac{\pi_1}{\theta_1}β2=θ1π1

所以用 OLS 分別估計 π1\pi_1π1 和 θ1\theta_1θ1 的估計量 π^1\hat{\pi}_1π^1 和 θ^1\hat{\theta}_1θ^1，那麼：

β^2=π^1θ^1\hat{\beta}_2 = \frac{\hat{\pi}_1}{\hat{\theta}_1}β^2=θ^1π^1

這個估計量是 consistent（一致的），這種方法叫做間接最小平方法（indirect least squares, ILS）。