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.

313707015
陳麗靜

a. β_2(WAGE)：預期＋，時薪越↑，勞動供給↑。

β_3(EDUC)：預期＋，受教育程度↑，工時↑。

β_4(AGE)：預期不確定，年輕時可能工時為正，但年紀大了可能負向影響工時。

β_5(KIDSL6)：預期－，幼兒數量↑，婦女外出工作↓。

β_6(NWIFEINC)：預期－，其他收入↑，經濟壓力較低，勞動供給↓。

b. ∵ WAGE是內生變數，可能和誤差項(e)有關。
① WAGE並非隨機決定，勞動供給(HOURS)同時影響其工資(WAGE)。
② e也可能包含工作意願、個人特質、健康等，這些因素會和WAGE有關。

c. IV必符合以下兩條件：

① IV必和內生變數高度相關
∴ EXPER累積越多會影響個人的WAGE水準；
$EXPER^2$ 隨著年資增加，可能邊際效益下降。

② IV必與e無關
∴ EXPER和$EXPER^2$主要影響WAGE
而不會直接影響HOURS。

d. 該方程式 identified. ∵ 工具變數(2個)數量大於內生變數(1個)
且IV和內生變數強相關、和e無關

e. First Stage：用EXPER和$EXPER^2$迴歸WAGE，得到WAGE的預測值 $\widehat{WAGE}$

Second Stage：將$\widehat{WAGE}$代入原方程式，用$\widehat{WAGE}$、EDUC、AGE、KIDSL6、NWIFEIC迴歸HOURS

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 var(z). Show that cov(z, x)/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 var(z). Show that cov(z, y)/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. $x = r_1 + \theta_1 z + v$

$E(x) = r_1 + \theta_1 E(z)$

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

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

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

Assuming $vE(z - E(z)) = 0$, we get $E[(z - E(z))(x - E(x))] = \theta_1 E[(z - E(z))^2]$

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

This is the OLS estimator of θ_1 in the 1st regression $x = r_1 + \theta_1 z + v$.

b.

$y = \pi_0 + \pi_1 z + u$.

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

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

$(z - E(z))(y - E(y)) = \pi_1 (z - E(z))^2 + u(z - E(z))$

$E[(z - E(z))(y - E(y))] = \pi_1 E[(z - E(z))^2] + uE[(z - E(z))]$

Assuming $uE[(z - E(z))] = 0$, we get $E[(z - E(z))(y - E(y))] = \pi_1 E[(z - E(z))^2]$

Solve $\pi_1 = \dfrac{E[(z - E(z))(y - E(y))]}{E(z - E(z))^2} = \dfrac{\text{cov}(z, y)}{\text{var}(z)}$

This is the OLS estimator of π_1 in the regression $y = \pi_0 + \pi_1 z + u$.

c.

$$y = \beta_1 + \beta_2 x + e$$

x 用 1st regression : $x = r_1 + \theta_1 z + v$. 代入原模型

$$y = \beta_1 + \beta_2 (r_1 + \theta_1 z + v) + e$$

$$= (\beta_1 + \beta_2 r_1) + \beta_2 \theta_1 z + (\beta_2 v + e)$$

對應 $y = \pi_0 + \pi_1 z + u$. 得 : $\begin{cases} \pi_0 = \beta_1 + \beta_2 r_1 \\ \pi_1 = \beta_2 \theta_1 \\ u = \beta_2 v + e \end{cases}$

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

e.

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

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

This estimator is consistent if z is uncorrelated with v.

$$\pi_1 = \frac{E[(z - E(z))(y - E(y))]}{E(z - E(z))^2} = \frac{cov(z, y)}{var(z)}$$

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

This estimator is consistent if z is uncorrelated with u.

Next.

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

$$\hat{\beta}_2 = \frac{\hat{\pi}_1}{\hat{\theta}_1} \quad \frac{\left[\dfrac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})^2}\right]}{\left[\dfrac{\sum(z_i-\bar{z})(x_i-\bar{x})}{\sum(z_i-\bar{z})^2}\right]} = \frac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})(x_i-\bar{x})}$$

$$= \frac{\dfrac{\sum(z_i-\bar{z})(y_i-\bar{y})}{N}}{\dfrac{\sum(z_i-\bar{z})(x_i-\bar{x})}{N}} = \frac{\widehat{cov}(z,y)}{\widehat{cov}(z,x)}$$

當 $N \to \infty$, $\widehat{cov}(z,y) \xrightarrow{p} cov(z,y)$; $\widehat{cov}(z,x) \xrightarrow{p} cov(z,x)$

$$\therefore \hat{\beta}_2 = \frac{\widehat{cov}(z,y)}{\widehat{cov}(z,z)} \xrightarrow{p} \beta_2 = \frac{cov(z,y)}{cov(z,x)}$$