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. WAGE β_2 ⊕ wage↑ 激勵更多勞动

EDUC β_3 ⊗ 可能更願意進入職場
也可能效率高 → 勞动減少

AGE β_4 ⊗ 可能因經驗↑ or 因健康… ↓

KIDSL6 β_5 ⊖ 小孩越多 越沒時間工作

NWIFEINC β_6 ⊖ 越多其它收入來源 會工作時間越少

b. endogeneity

$cov(WAGE, e) \neq 0$

c. $\widehat{WAGE} = \beta_1' + \beta_2' EXPER + \beta_3' EXPER^2 + e'$

$EXPER, EXPER^2$ do not direct effect HOURS. & $cov(EXPER, e) = cov(EXPER^2, e) = 0$
& $EXPER, EXPER^2$ are strongly correlated with WAGE

d. endogeneity ⇒ inconsistent

e. ① estimate first stage regression
$\widehat{WAGE} = \hat{\beta}_1 + \hat{\beta}_2 EXPER + \hat{\beta}_3 EXPER^2$

② replace $\widehat{WAGE}$ in the original regression, and apply the OLS estimation

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.

$$x = \gamma_1 + \theta_1 z + v$$

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

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

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

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

$$\theta_1 = \frac{\text{cov}(z, y)}{\text{var}(z)}$$

b.

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

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

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

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

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

$$\pi_1 = \frac{\text{cov}(z, y)}{\text{var}(z)}$$

c.

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

$$\pi_0 = \beta_1 + \beta_2\gamma_1$$
$$\pi_1 = \beta_2\theta_1$$
$$u = \beta_2 v + e$$

d.

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

e.

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

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

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

$\rightarrow$ It is IV estimator

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