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.

WAGE ⇒ +

EDUC ⇒ +

AGE ⇒ + or −

KIDSL6 ⇒ −

NWIFEINC ⇒ − ※

b. Explain why this supply equation cannot be consistently estimated by OLS regression.

Because WAGE is endogenous and likely correlated with the error term, the OLS estimator will be biased and inconsistent. This violets the exogeneity assumption required for consistent OLS estimation.

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.

EXPER and $EXPER^2$ are valid instruments for WAGE because they are correlated with WAGE (relevance) but do not directly affect HOURS (exogeneity), only through their impact on wages. ※

d. Is the supply equation identified? Explain.

Yes, the supply equation is identified.

Since there are valid instruments for the endogenous variable WAGE — specially EXPER and $EXPER^2$ — and there are more instruments than endogenous regressors, the equation satisfies the order condition for identification. ※

e. Describe the steps [not a computer command] you would take to obtain IV/2SLS estimates.

1. First stage: Regress WAGE on the instruments (EXPER and $EXPER^2$) and other exogenous variables to obtain the fitted values of WAGE.

2. Second stage: Regress HOURS on the fitted values of WAGE and the other exogenous variables to estimate the supply equation.

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.

① $x = \gamma_1 + \theta_1 z + v$

take $E[\ ]$ ② $E[x] = \gamma_1 + \theta_1 E[z]$

① − ② $\Rightarrow$ $x - E[x] = \theta_1 (z - E[z]) + v$

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

take $E[\ ]$ $\Rightarrow E[(z-E[z]) \cdot (x - E[x])] = \theta_1 E[(z - E[z])^2] + \underline{E[v(z - E[z])]} = E[(v - E[v]) \cdot (z - E[z])]$

$(\because E[v] = 0)$

$\Rightarrow \text{cov}(z, x) = \theta_1 \text{var}(z) + \underbrace{\text{cov}(v, z)}_{=0}$

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

This is the OLS estimator of θ_1 in the regression : $x = \gamma_1 + \theta_1 z + v$

*

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.]

① $y = \pi_0 + \pi_1 z + u$

take $E[\ ]$ ② $E[y] = \pi_0 + \pi_1 E[z]$

$\vdots$ same step as (a)

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

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

*

c. 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.

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

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

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

$$\Rightarrow \pi_0 = \beta_1 + \beta_2 \gamma, \quad \pi_1 = \beta_2 \theta_1, \quad u = \beta_2 v + e$$

✱

d. Show that $\beta_2 = \pi_1/\theta_1$.

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

✱

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.

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

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

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

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

✱