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 : +，工資↑，勞動供給↑

β_3 : +，教育↑，勞動供給↑

β_4 : 不一定，因各個年齡願意參與勞動市場時數可能不一致

β_5 : —，有了小孩後，可能會減少時間參與勞動市場

β_6 : —，有 non-wife income 後，可能使 married women 工作意願降低

b. 因 WAGE 是內生變數，可能與誤差項相關，像 Ability 等未納入迴歸模型的變數可能影響 WAGE & HOURS ※

c. 因 EXPER、$EXPER^2$ 與 WAGE 有相關性，即 EXPER↑，WAGE 通常越多

且 EXPER、$EXPER^2$ 不會直接影響 HOURS，只會經由 WAGE 間接產生影響

因此 EXPER、$EXPER^2$ 是好的工具變數，（且 $\text{cov}(EXPER, e_i) = 0$、$\text{cov}(EXPER^2, e_i) = 0$）

d. 因 WAGE 是內生變數，工具變數有 EXPER、$EXPER^2$

因 2>1，因此 supply equation 是 identified

e. 首先是選定工具變數，即 EXPER、$EXPER^2$

接著對 WAGE 迴歸

⇒ $WAGE = \beta_0 + \beta_1 EXPER + \beta_2 EXPER^2 + \beta_3 EDUC + \cdots + u$

然後對原始模型跑迴歸，並將 WAGE 換成上一步的 $\widehat{WAGE}$，

再重新估計並得到 $\hat{\beta}_0$，$\hat{\beta}_1 \cdots$ #

$$Hours = \beta_0 + \beta_1 \widehat{WAGE} + \beta_2 EDUC + \cdots + e$$

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. $X = \gamma_1 + \theta_1 z + v$, $E(X) = \gamma_1 + \theta_1 E(z)$, $X - E(X) = \theta_1 (z - E(z)) + v$

$\Rightarrow X - E(X) = \theta_1 (z - E(z)) + v$.

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

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

$= \theta_1 E[(z - E(z))^2]$, $E(z - E(z))v = 0$

$$\theta_1 = \frac{E[(z - E(z))(X - E(X))]}{E[(z - E(z))^2]} = \frac{\text{Cov}(z, X)}{\text{Var}(z)}$$ #

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

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

$(z - E(z))(y - E(y)) = \pi_1 (z - E(z))^2 + (z - E(z)) \cdot u$, $E(z - E(z))u = 0$

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

$$\pi_1 = \frac{E(z - E(z))(y - E(y))}{E(z - E(z))^2} = \frac{\text{cov}(z, y)}{\text{Var}(z)}$$ #

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

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

$= \pi_0 + \pi_1 z + u$

$\pi_0 = (\beta_1 + \beta_2\gamma_1)$, $\pi_1 = \beta_2\theta_1$, $u = \beta_2 v + e$.

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

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

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

$$\hat{\beta}_2 = \frac{\hat{\pi}_1}{\hat{\theta}_1} = \frac{\frac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})^2}}{\frac{\sum(z_i-\bar{z})(x_i-\bar{x})}{\sum(z_i-\bar{z})^2}} = \frac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})(x_i-\bar{x})} = \frac{\sum(z_i-\bar{z})(y_i-\bar{y})/N}{\sum(z_i-\bar{z})(x_i-\bar{x})/N}$$

$$= \frac{\widehat{cov}(z,y)}{\widehat{cov}(z,x)}$$

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