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 → + ⇒ 較高工資誘使勞動供給增加

EDUC → + ⇒ 優秀的工作者可能傾向工作更多

or → − ⇒ 優秀(教育程度愈高)的工作者可能效率愈高, 勞動供給減少

AGE 可能為 + or − ⇒ 預期在中年之前逐漸增加, 然後之後隨之漸減

KIDSL6 → − ⇒ 預期小於 6 歲子女愈多, 勞動供給減少

NWIFEINC → − ⇒ 家庭收入增加會減少對妻子收入的需求

#

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

① 因工時與工資是由供給與需求共同決定的, 此種情況下工資(WAGE)是內生變數 → 這會導致 OLS 估計量不是 consistent

② 該方程式中未包含對能力的衡量, 能力偏誤是一種遺漏變數偏誤, 其來源是个人的能力未被衡量, 被含在誤差項中, 由於一个人的能力常與其教育以及工資有關, 因此 e 可能與 EDUC 及 WAGE 有關, 這種內生性將使 OLS 估計量不是 consistent

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.

由於 wage 與 EXPER 及 $EXPER^2$ 的關聯是由於需求因素而非供給因素, 因此預期 EXPER 及 $EXPER^2$ 與 HOURS (供給) 不相關且與供給方程式的誤差不相關。

Suppose that there is another variable, z_i, such that

1. z_i does not have a direct effect on y_i, and thus it does not belong on the right-hand side of the model as an explanatory variable.
2. z_i is not correlated with the regression error e. (i.e., z_i is exogenous, $corr(z_i, e_i) = 0$)
3. z_i is strongly (or at least not weekly) correlated with x_i, the endogenous variable (i.e., $corr(z_i, x_i) \neq 0$).

預期工資 wage 與 EXPER 及 $EXPER^2$ 之間存在相關性, 因具有更多工作經驗的勞工可要求更高工資。

d. Is the supply equation identified? Explain.

這个供給方程式是 identified，因為只指定了一個內生變數(WAGE)，故至少需有一個工具變數，而在此 Case 有2个工具變數(EXPER、$EXPER^2$)，因此滿足了 $L \geq B$ 的條件。

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

1. 第一階段迴歸：作法用 WAGE (左式) 對所有 (右式) 外生變數(EDUC, AGE, KIDSL6, NWIFEINC) 及工具變數(EXPER、$EXPER^2$) 執行 OLS
⇒ 得出擬適值 $\widehat{WAGE}$

2. 第二階段迴歸：作法根據原始設定，但利用第一階段估出的擬適值 $\widehat{WAGE}$ 取代原始供給模型中內生的WAGE，再執行OLS
⇒ 此步驟得出的估計參數就會是 IV/2SLS 的估計量。

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.

方程式 $x = \gamma_1 + \theta_1 z + v$ — ①

兩邊同取期望值 $\Rightarrow E(x) = \gamma_1 + \theta_1 E(z)$ — ②

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

同 $\times (z - E(z))$

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

再同取期望值

$$E[(z-E(z))(x-E(x))] = \theta_1 E[(z-E(z))^2] + \underbrace{E[v(z-E(z))]}_{E[(v-\overset{0}{E(v)})(z-E(z))]}$$

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

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

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.]

方程式: $y = \pi_0 + \pi_1 z + u$ — ①

兩邊同取期望值 $\Rightarrow E(y) = \pi_0 + \pi_1 E(z)$ — ②

① − ②: $y - E(y) = \pi_1 [z - E(z)] + u$

再同 $\times (z - E(z))$

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

再同取期望值

$$E[(z-E(z))(y-E(y))] = \pi_1 E[(z-E(z))^2] + \underbrace{E[u(z-E(z))]}_{E[(u-\overset{0}{E(u)})(z-E(z))]}$$

$$\Rightarrow \text{cov}(z,y) = \pi_1 \text{Var}(z) + \underbrace{\text{cov}(u,y)}_{\text{Assume } = 0}$$

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

This is the OLS estimater 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.

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

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

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

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

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

$\pi_1 = \beta_2\theta_1 \Rightarrow \beta_2 = \pi_1/\theta_1$ #

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.

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

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

因此

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

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