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.

β1:正，小時不會出現負數

β2（WAGE）：

預期符號正。原因：時薪（WAGE）提高，女性可能會增加工作小時數（HOURS），因為工作變得更有吸引力。

β3（EDUC）：

預期符號正。原因：教育程度（EDUC）越高，通常技能和生產力越高，女性更有可能參與勞動市場，增加工作小時數。

β4（AGE）： 預期符號：負。

原因：但隨著年齡增長（特別是接近退休），工作小時數可能減少。

β5（KIDS6）：

預期符號負。原因 6 歲以下的子女數（KIDS6）增加，女性需要花更多時間照顧孩子，減少工作小時數。

β6（NWIFEINC）：

預期符號負。原因：家庭其他收入（NWIFEINC）增加，家庭財務壓力減少，女性可能減少工作小時數（收入效應，income effect）。

b.

OLS 回歸要得到一致估計（consistent estimates），需要滿足以下假設之一：解釋變量與誤差項 e 不相關(Cov(X,e)=0= 0 Cov(X,e)=0)。在該模型中，可能導致不一致估計的主要問題如下：

內生性（Endogeneity）問題：WAGE 與誤差項相關，因為勞動供給（HOURS）和時薪（WAGE）可能是同時決定的（simultaneity）。

c.

不建議直接使用 EXPER 和 $EXPER^2$ 作為 WAGE 的工具變量。雖然 EXPER 和 $EXPER^2$ 可能與 WAGE 相關（滿足相關性），但它們可能與誤差項 e 相關（不滿

足外生性)。EXPER 可能受到未觀察因素(如能力、健康)的影響，這些因素同時影響 HOURS 和 EXPER，導致內生性問題。

f.

yes

g.

第一階段回歸(First Stage):

將內生變量 WAGE 回歸於工具變量(EXPER, $EXPER^2$)以及所有外生變量(EDUC, AGE, KIDS6, NWIFEINC)。

回歸方程： $WAGE = \pi_0 + \pi_1 EXPER + \pi_2 EXPER^2 + \pi_3 EDUC + \pi_4 AGE + \pi_5 KIDS6 + \pi_6 NWIFEINC + v$

從此回歸中得到 WAGE 的預測值：WAGE^。

第二階段回歸(Second Stage):

將第一階段得到的 WAGE^ 代入原始方程，代替 WAGE。

回歸方程： $HOURS = \beta_1 + \beta_2 \hat{WAGE} + \beta_3 EDUC + \beta_4 AGE + \beta_5 KIDS6 + \beta_6 NWIFEINC + u$

使用普通最小平方法(OLS)估計此方程，得到 β_2 及其他係數的 2SLS 估計值。

調整標準誤(Optional but Recommended):

2SLS 估計的標準誤需要修正，因為第一階段的預測誤差會影響第二階段。可以使用統計軟體自動計算正確的標準誤。

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.

a.

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

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

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

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

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

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

b.

与a.同理

$$\Rightarrow E((y - E(y))(z - E(z))) = \pi E((z - E(z))^2)$$

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

c.

$$y = \beta_1 + \beta_2(v_1 + \theta z + v) + e$$

$$= (\beta_1 + \beta_2 v_1) + \beta_2 \theta z + (\beta_2 v + e)$$

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

$$\Rightarrow \pi_0 = \beta_1 + \beta_2 v_1, \quad \pi_1 = \beta_2 \theta_1, \quad u = \beta_2 v + e$$

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

e.

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

$$\hat{\beta}_2 = \frac{\widehat{cov}(z, y)}{\widehat{cov}(z, x)} \quad (\text{IV estimator})$$

consistent

$$\because \widehat{cov}(z, y) \xrightarrow{p} cov(z, y)$$

$$\widehat{cov}(z, x) \xrightarrow{p} cov(z, x)$$

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