MT 2 Sp. 05

1 (a)

$E[\ln wage \mid usage] - E[\ln wage \mid nonusage]$ for a given level of edu and exp

$= \beta_1 + \beta_5 \text{ Female}$

$\Rightarrow$ the effect is not the same for male and female, unless $\beta_5 = 0$.

(b)

$E(\ln Wage \mid Edu, Exp)$ from (1)

	Usage	Non-usage
Male	$\beta_0 + \beta_1 + \beta_2 Edu + \beta_3 Exp$	$\beta_0 + \beta_2 Edu + \beta_3 Exp$
Female	$\beta_0 + \beta_1 + \beta_2 Edu + \beta_3 Exp + \beta_4 + \beta_5$	$\beta_0 + \beta_2 Edu + \beta_3 Exp + \beta_4$

$E[\ln wage \mid Edu, Exp]$ from (2)

	usage	non-usage
Male	$\alpha_0 + \alpha_1 Exp + \alpha_2 Edu + \alpha_4$	$\alpha_0 + \alpha_1 Exp + \alpha_2 Edu$
Female	$\alpha_0 + \alpha_1 Exp + \alpha_2 Edu + \alpha_3$	$\alpha_0 + \alpha_1 Exp + \alpha_2 Edu + \alpha_5$

Equate the cells. Note the Coe. on edu, exp in (1) and (2) will be the same $\Rightarrow$ $\beta_2 = \alpha_2$, $\beta_3 = \alpha_1$. (An easy way of proving this is to derive (2) from (1)). Alternatively, assume that exp = 0, edu = 0. This gives

(i) $\beta_0 + \beta_1 = \alpha_0 + \alpha_4$, (ii) $\beta_0 = \alpha_0$ $\Rightarrow$ $\beta_1 = \alpha_4$

(iii) $\left.\begin{array}{l}\beta_0 + \beta_1 + \beta_5 \\ + \beta_4\end{array}\right\} = \alpha_0 + \alpha_3$, (iv) $\alpha_0 + \alpha_5 = \beta_0 + \beta_4$ $\Rightarrow$ $\beta_4 = \alpha_5$

and $\left.\begin{array}{l}\beta_5 + \beta_1 \\ + \beta_4\end{array}\right\} = \alpha_3$ $\Rightarrow$ $\beta_5 = \alpha_3 - \alpha_4 - \alpha_5$.

2. $$y = \begin{bmatrix} i_1 & 0 & 0 & 0 \\ 0 & i_2 & 0 & 0 \\ 0 & 0 & i_3 & 0 \\ 0 & 0 & 0 & i_4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} + u \Rightarrow Y = X\alpha + u$$

i_k = Column of ones for the k th quarter ($k = 1, \dots, 4$).
# of elements in i_k depends on # obs for the k th quarter

Assume that n_1 = # first quarter observations, n_2 = ------

$$\Rightarrow \hat{\alpha} = (X'X)^{-1} X'Y = \begin{bmatrix} n_1 & 0 & 0 & 0 \\ 0 & n_2 & 0 & 0 \\ 0 & 0 & n_3 & 0 \\ 0 & 0 & 0 & n_4 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma Y \text{ for 1st quarter} \\ \Sigma Y \text{ --- 2nd --} \\ \Sigma Y \text{ -- 3rd --} \\ \Sigma Y \text{ -- 4th} \end{bmatrix}$$

$$= \begin{matrix} \frac{1}{n_1} \Sigma Y \text{ (first quarter)} = \bar{Y}_1 \\ \frac{1}{n_2} \Sigma Y \text{ 2nd --} = \bar{Y}_2 \\ \frac{1}{n_3} \Sigma Y \text{ 3rd ---} = \bar{Y}_3 \\ \frac{1}{n_4} \Sigma Y \text{ 4th --} = \bar{Y}_4. \end{matrix}$$

* You can also assume that $n_1 = n_2 = n_3 = n_4$

$$Y = \begin{bmatrix} i_1 & 0 & 0 & 0 \\ i_2 & i_2 & 0 & 0 \\ i_3 & 0 & i_3 & 0 \\ i_4 & 0 & 0 & i_4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + u = X\beta + u$$

$$\hat{\beta} = (X'X)^{-1} X'Y \Rightarrow X'X\hat{\beta} = X'Y$$

$$\Rightarrow \begin{bmatrix} n & n_2 & n_3 & n_4 \\ n_2 & n_2 & 0 & 0 \\ n_3 & 0 & n_3 & 0 \\ n_4 & 0 & 0 & n_4 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{bmatrix} = \begin{pmatrix} \Sigma Y \text{ for all obs} \\ \Sigma Y \text{ for 2nd quarter obs} \\ \Sigma Y \text{ for 3rd ---} \\ \Sigma Y \text{ for 4th --} \end{pmatrix}, \quad n = n_1 + n_2 + n_3 + n_4$$

$$\Rightarrow \quad n\hat{\beta}_1 + n_2\hat{\beta}_2 + n_3\hat{\beta}_3 + n_4\hat{\beta}_4 = n\bar{Y} \quad \text{— (I)}$$

$$n_2\hat{\beta}_1 + n_2\hat{\beta}_2 = n_2\bar{Y}_2 \quad \text{— (II)}$$

$$n_3\hat{\beta}_1 + n_3\hat{\beta}_3 = n_3\bar{Y}_3 \quad \text{— (III)}$$

$$n_4\hat{\beta}_1 + n_4\hat{\beta}_4 = n_4\bar{Y}_4 \quad \text{— (IV)}$$

Write $n\bar{Y}$ as $n\bar{Y} = n_1\bar{Y}_1 + n_2\bar{Y}_2 + n_3\bar{Y}_3 + n_4\bar{Y}_4$. Then

from (I) $(n_1+n_2+n_3+n_4)\hat{\beta}_1 + n_2\hat{\beta}_2 + n_3\hat{\beta}_3 + n_4\hat{\beta}_4 = n_1\bar{Y}_1 + n_2\bar{Y}_2 + n_3\bar{Y}_3 + n_4\bar{Y}_4$

$$\Rightarrow n_1\hat{\beta}_1 = n_1\bar{Y}_1 \Rightarrow \hat{\beta}_1 = \bar{Y}_1$$

From (II) $\hat{\beta}_2 = \bar{Y}_2 - \hat{\beta}_1 = \bar{Y}_2 - \bar{Y}_1$

From (III) $\hat{\beta}_3 = \bar{Y}_3 - \hat{\beta}_1 = \bar{Y}_3 - \bar{Y}_1$

From (IV) $\hat{\beta}_4 = \bar{Y}_4 - \hat{\beta}_1 = \bar{Y}_4 - \bar{Y}_1$.

3. OLS estimates of the true model are unbiased (BLUE) and consistent. We proved it.

$\Rightarrow$ b_1 and b_2 are unbiased and consistent.

Normal equations (first order conditions of minimizing ~~resi~~ sum of squared errors) are (take the one for α_1 and α_0)

$$\Sigma(Y - \alpha_0 - \alpha_1 X_1 - \alpha_2 X_2)(-X_1) = 0.$$

$$\Sigma(Y - \alpha_0 - \alpha_1 X_1 - \alpha_2 X_2)(-1) = 0.$$

$$\Rightarrow \quad \Sigma Y = n\alpha_0 + \alpha_1 \Sigma X_1 + \alpha_2 \Sigma X_2 \quad \text{and}$$

$$\Sigma Y X_1 = \alpha_0 \Sigma X_1 + \alpha_1 \Sigma X_1^2 + \alpha_2 \Sigma X_1 X_2$$

$$\Rightarrow \quad \Sigma y x_1 = \alpha_1 \Sigma x_1^2 + \alpha_2 \Sigma x_1 x_2 \qquad x_1 = X_1 - \bar{X}_1, \text{ etc.}$$

Using b_1, b_2 as the OLS estimators of α_1, α_2

$$\Sigma y x_1 = b_1 \Sigma x_1^2 + b_2 \Sigma x_1 x_2$$

Using matrix notations

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = (x'x)^{-1} x'y$$ in deviation form

$$\Rightarrow \quad (x'x)\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = x'y$$

$$\Rightarrow \Sigma x_1^2 b_1 + \Sigma x_1 x_2 \, b_2 = \Sigma x_1 y$$

and $\Sigma x_1 x_2 \, b_1 + \Sigma x_2^2 b_2 = \Sigma x_2 y$

It does not matter whether y is in deviation form or not. Note that $\Sigma x Y = \Sigma X y = \Sigma x y$ (Proof is Trivial).

$$\Rightarrow \quad b_1 + b_2 \frac{\Sigma x_1 x_2}{\Sigma x_1^2} = \frac{\Sigma x_2 y}{\Sigma x_1^2}$$

$$\Rightarrow \quad b_1 + b_2 \cdot b_{X_2X_1} = b_{YX_1}$$ (Part (b) result).

Part (a)

Now $E(b_1) + E(b_2) \cdot \underbrace{b_{X_2X_1}}_{\text{Constant given } X_1, X_2} = E(b_{YX_1})$ (not a random variable)

$$\Rightarrow \quad \alpha_1 + \alpha_2 \cdot b_{X_2X_1} = E(b_{YX_1}) \neq \alpha_1$$ (biased).

Also $\text{Plim}(b_1) + \text{Plim}(b_2) \cdot \text{Plim}(b_{X_2X_1}) = \text{plim}(b_{YX_1})$

$$\Rightarrow \text{Plim}(b_{YX_1}) = \alpha_1 + \alpha_2 \frac{\text{Plim} \frac{1}{n} \Sigma x_1 x_2}{\text{Plim} \frac{1}{n} \Sigma x_1^2}$$

$$= \alpha_1 + \alpha_2 \frac{\sigma_{x_1 x_2}}{\sigma_{x_1}^2}$$ Constant.

$$\neq \alpha_1.$$

(c) From part (b) argue that $b_{X_2X_1}$ = reg. coe. of X_1^2 on X_1 is positive. So if α_2 is positive – bias is positive, If α_2 is negative – bias is negative.

Question 4

$Y_i = \beta + u_i$

$E(u_i|X) = 0 \qquad Var(u_i|X_i) = \sigma^2 X_i^2$

Heteroskedastic model

$$\overset{y^*}{\frac{Y_i}{X_i}} = \beta \cdot \overset{x}{1/X_i} + \frac{u_i}{X_i}$$

$\Omega = diag\{X_1^2, \ldots, X_n^2\}$

$(P')^{-1} P^{-1} = \Omega^{-1}$

$$\hat{\beta}_{GLS} = \frac{\sum Y/X_i \cdot 1/X_i}{\sum (1/X_i)^2} = \frac{\sum_{i=1}^{n} (Y_i/X_i^2)}{\sum (1/X_i^2)}$$

$\Rightarrow P^{-1} = diag\{\frac{1}{X_1}, \ldots \frac{1}{X_n}\}$

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y \qquad \Omega =$$

$X = \underline{i}$

$PP' = \Omega$

or

$$Var(\hat{\beta}_{GLS}) = \frac{\sigma^2}{\sum (1/X_i^2)}$$

$\Rightarrow P^{-1} = diag\{1, 1/2, 1/3, 1/4\}$ 2

$P^{-1}y = P^{-1} i \beta + P^{-1} u$ 2

$Var(P^{-1}u) = \sigma^2, \quad E(P^{-1}u) = 0$

2

$$\hat{\beta}^* = \frac{(3/1 + 5/4 + 10/9 + 19/16)}{(1 + 1/4 + 1/9 + 1/16)} = \frac{943/144}{205/144} = \frac{943}{205} = 4.6$$

$$Var\, \hat{\beta}^* = \frac{\sigma^2}{\sum (1/X_i^2)} = \sigma^2 \frac{144}{205}$$

36 x 16

4 x 9   4 x 4

144

2

$= 0.702\, \sigma^2$

$Y_i = \beta + u_i$, $i=1,\dots,n_1$, $E(u_i)=0$ $V(u_i)=\sigma_1^2$

$Y_i = \beta + u_i$, $i=n_1+1,\dots,n$, $E(u_i)=0$ $Var(u_i)=\sigma_2^2$

two groups $\Rightarrow G=2$

(b) $Var(u) = \begin{bmatrix} \sigma_1^2 I_1 & 0 \\ 0 & \sigma_2^2 I_2 \end{bmatrix} = \Omega$ 2

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y$$ note $X = i$ - column of ones

$$= \left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)^{-1} \left(\frac{1}{\sigma_1^2}\sum_{i=1}^{n_1} y_i + \frac{1}{\sigma_2^2}\sum_{i=n_1+1}^{n} y_i\right)$$ 2

$$Var(\hat{\beta}_{GLS}) = \left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)^{-1}$$ 2

(c) $u_i \sim N(0, \sigma_1^2)$ $i=1,\dots,n_1$

$u_i \sim N(0, \sigma_2^2)$ $i=n_1+1,\dots,n$

The log-likelihood function is

$$\ln L = -\frac{n}{2}\ln 2\pi - \frac{1}{2}\left[n_1 \ln \sigma_1^2 + n_2 \ln \sigma_2^2\right] - \frac{1}{2}\sum_{g=1}^{2}\sum_{i=1}^{n_g}\left(\frac{u_{ig}^2}{\sigma_g^2}\right)$$ 2/

MLE estimators

$$\hat{\sigma}_1^2 = \frac{\sum \hat{u}_1^2}{n_1}$$ 2/ $$\hat{\sigma}_2^2 = \frac{\sum \hat{u}_2^2}{n_2}$$ 2/ where $\hat{u}_i^2 = (y_i - \hat{\beta}_{MLE})^2$

$\beta_{MLE} =$

$$\ln L = -\frac{n}{2}\ln 2\pi - \tfrac{1}{2}\left[n_1 \ln \sigma_1^2 + n_2 \ln \sigma_2^2\right]$$

$$-\tfrac{1}{2}\left[\frac{1}{\sigma_1^2}\sum_{i=1}^{n_1} u_{i1}^2 + \frac{1}{\sigma_2^2}\sum_{i=n_1+1}^{n_2} u_{i2}^2\right]$$

$$= -\frac{n}{2}\ln 2\pi - \frac{1}{2}\left[n_1 \ln \sigma_1^2 + n_2 \ln \sigma_2^2\right]$$

$$-\tfrac{1}{2}\left[\frac{1}{\sigma_1^2}\sum_{i=1}^{n_1}(y_{i1}-\beta)^2 + \frac{1}{\sigma_2^2}\sum_{i=n_1+1}^{n_2}(y_{i2}-\beta)^2\right]$$

$$\frac{\partial \ln L}{\partial \beta} = 0$$

$$\Rightarrow \hat{\beta}_{MLE} = \left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)^{-1}\left(\frac{1}{\sigma_1^2}\sum_{i=1}^{n_1} y_i + \frac{1}{\sigma_2^2}\sum_{i=n_1+1}^{n_2} y_i\right)$$

3/

4) Derive LR for $H_0: \sigma_1^2 = \sigma_2^2$

$$LR = n \ln \hat{\sigma}^2 - \left(n_1 \ln \hat{\sigma}_1^2 + n_2 \ln \hat{\sigma}_2^2\right)$$

$$LR = -2\left[\ln L_R - \ln L_{UR}\right]$$

For restricted model $Var(u) = \sigma^2 I$ $\qquad \sigma_1^2 = \sigma_2^2 = \sigma^2$

$$\ln L_R = -\frac{n}{2}\ln 2\pi - \tfrac{1}{2} n \ln \sigma^2 - \frac{1}{2\sigma^2} u'u$$

(4) $H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

Under ~~the~~ null, $Var(u) = \sigma^2 I$ where $\sigma_1^2 = \sigma_2^2 = \sigma^2$

then $\ln L = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} u'u$

Now, $\hat{\sigma}^2_{ML} = \frac{\Sigma(y-\hat{\beta})^2}{n}$. Thus

$$\ln L_R = -\frac{n}{2}\left[\ln(2\pi)+1\right] - \frac{n}{2}\ln\hat{\sigma}^2 \quad 2/$$

Also $\ln L_{UR} = -\frac{n}{2}\left[\ln(2\pi+1)\right] - \frac{1}{2}\left(n_1 \ln\hat{\sigma}_1^2 + n_2 \ln\hat{\sigma}_2^2\right) \; 2/$

$\Rightarrow LR = n\ln\hat{\sigma}^2 - \left(n_1 \ln\hat{\sigma}_1^2 + n_2\ln\hat{\sigma}_2^2\right) \sim \chi^2_{(1)} \; 4/$

If $LR > \chi^2_{(1)}$ at 5% $\Rightarrow$ reject H_0 in favor of heteroskedasticity