MSc in Environmental Economics:
Answer Keys to Examination in Environmental Econometrics (GR03)

May 2007

Section 1 (Compulsory) [40 points]

Question 1

1 7.3% increase of hourly wage from OLS. Both are significant since $t_{OLS} = 0.073/0.004 = 18.25 > t_{\alpha/2}$ and $t_{IV} = 0.132/0.055 = 2.4 > t_{\alpha/2} = 1.96$ [5 points]

2 There exists correlation between them and thus the problem of endogeneity. [5 points]

3 biased.

$$\widehat{\beta}_{1,OLS} = \beta_1 + \frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)u_i}{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}$$

Since $E\left(u_i|X\right) \neq 0$,

$$E\left(\widehat{\beta}_{1,OLS}|X\right) \neq \beta_1$$

[5 points]

4 IV is a variable Z such that

$$Cov\left(Z,X\right) \neq 0 \text{ and } Cov\left(Z,u\right) = 0.$$

As long as unobservable factor like ability does not relate with proximity and proximity is related to education, then it can serve as IV.[5 points]

5 correlation between measured years of schooling and the error tem. Suppose that the true model is

$$Y_i = \beta_0 + \beta_1 X_i + v_i$$

But X_i is measured with error so that $\widetilde{X}_i = X_i + e_i$ is observed. Then the regression equation used in the estimation is

$$Y_i = \beta_0 + \beta_1 \widetilde{X}_i + u_i,$$

where $u_i = v_i - \beta_1 e_i$. Then,

$$E\left(\widetilde{X}_i u_i\right) = -\beta_1 Var\left(e_i\right) \neq 0.$$

[5 points]

6 The OLS estimator is biased toward zero, which is called the attenuation bias. Since IV estimation is consistent, the magnitude of IV estimator may be larger than that of OLS estimator [5 points]

Question 2

1

$$\begin{aligned} Y_i &= \frac{\beta_0 + \beta_1 \gamma_0}{1 - \gamma_1 \beta_1} + \frac{\beta_1 v_i + u_i}{1 - \gamma_1 \beta_1} \\ X_i &= \frac{\gamma_0 + \beta_0 \gamma_1}{1 - \gamma_1 \beta_1} + \frac{\gamma_1 u_i + v_i}{1 - \gamma_1 \beta_1} \end{aligned}$$

Note that

$$Cov\left(X_i, u_i\right) = \frac{\gamma_1}{1 - \gamma_1 \beta_1} Var\left(u_i\right) \neq 0.$$

Hence, the OLS will lead to a biased estimator. [5 points]

2 first, No. second, (β_0, β_1) can be identified due to order condition.[5 points]

Section 2 [60 points]
Question 1 (Ordinary Least Squares (OLS))

1 (a) zero-mean condition of the error term or no correlation between X and u. [2.5 points]

(b) same as above, but also without heteroskedasticity and autocorrelation in error. [2.5 points]

2

$$\begin{aligned} \widehat{\beta}_{1,OLS} &= \frac{\sum_{i=1}^{N} \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^{N} \left(X_i - \overline{X}\right)^2} \\ \widehat{\beta}_{0,OLS} &= \overline{Y} - \widehat{\beta}_{1,OLS} \overline{X} \end{aligned}$$

[5 points]

3

$$t = \frac{\widehat{\beta}_{1,OLS}}{\sqrt{\widehat{Var\left(\widehat{\beta}_{1,OLS}\right)}}}$$

$t-$distribution.[5 points]

4 biased (omitted variable bias). [5 points]

Question 2 (Generalized Linear Regression)

1 the variancesof error term across samples are not identical. unbiased but inefficient. [5 points]

2 First, get the residuals $\widehat{u}_i$ from the OLS. Then plug the residuals into

$$Var\left(\widehat{\widehat{\beta}_{1,OLS}}\right) = \frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2 \widehat{u}_i^2}{\left[\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2\right]^2}$$

[5 points]

3 The hypotheses are

$$H_0 : \delta_1 = 0 \text{ vs. } H_1 : \delta_1 \neq 0$$

Breusch-Pagan test. and t-statistic. [5 points]

4 GLS. [5 points]

Question 3 (Binary Response Model)

1 (a) 0.1 increase in response probability when there is 1% increase in annual income.[2.5 points]

(b) the predicted probability can be outside the interval $[0, 1]$. Heteroskedasticity. Constant marginal effect. [2.5 points]

2

$$\Pr(Y_i = 1) = \frac{\exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})}{1 + \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})}.$$

$$\log\left(\frac{\Pr(Y_i = 1)}{1 - \Pr(Y_i = 1)}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}$$

[5 points]

3 (a) when there is 1% increase in income, there is 48% increase in the log odds-ratio. Significant since $t = 4$ [2.5 points]

(b) log odds ratio $= -5.07 + 0.48 \times 5 + 0.03 \times 40$ [2.5 points]

4 (a)

$$H_0 : \beta_1 = 0 = \beta_2 \text{ vs. } H_1 : \text{ not } H_0$$

[2.5 points]

(b) The likelihood ratio test statistic is $LR = 2 \times (-278.84 + 286.85) = 16.03$, which follows the chi-square with 2 degress of freedom. [2.5 points]

Question 4

1 (a)

$$\begin{aligned} Y_i^* &= \beta X + u_i \\ Y_i &= \begin{cases} 0 & \text{if} \quad Y_i^* < \mu_1 \\ 1 & \text{if} \quad \mu_1 \leq Y_i^* < \mu_2 \\ 2 & \text{if} \quad Y_i^* \geq \mu_2 \end{cases} \end{aligned}$$

[4 points]

(b) threshold values determining the survey answers [2 points]

(c)

$$\Pr(Y_i = 2) = 1 - \Phi(\mu_2 - \beta X).$$

[3 points]

2

$$\frac{\partial \Pr(Y=0)}{\partial X_1} < 0$$

$$\frac{\partial \Pr(Y=1)}{\partial X_1} \begin{matrix} > \\ = \\ < \end{matrix} ?\ 0$$

$$\frac{\partial \Pr(Y=2)}{\partial X_1} > 0,$$

[6 points]

3 multinomial model ignores the information on the ranking of survey response. hence ordered model is appropriate. [5 points]