KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2017/2018

FIRST SEMESTER EXAMINATION FOR THE DEGREE OF MASTER OF ECONOMICS

EES 805: MICROECONOMETRICS

DATE: THURSDAY 8TH FEBRUARY 2018 **TIME: 9.00 A.M. - 12.00 P.M.**

INSTRUCTIONS:

1. Answer any four (4) questions. Each question carries fifteen (15) marks.
2. Show all required solutions.
3. You may use an unprogrammable calculator.

Question (1) [15 Marks]

Consider the standard Tobit model with given specifications and assumptions:

$$y_i^* = x_i'\beta + \varepsilon_i \qquad i = 1, 2, \ldots, N \qquad (1.1)$$

$$y_i = y_i^*, \qquad \text{if } y_i^* > 0$$

$$y_i = 0, \qquad \text{if } y_i^* \leq 0$$

where $\varepsilon_i \sim NID(0, \sigma^2)$ and independent of x_i. This model is also called a censored regression model since some observations are censored (in this case from below).

(a) Using the specification of (1.1), find the probability of $y_i = 0$ given x_i, $P\{y_i = 0 | x_i\}$, and the conditional expectation of y_i given a positive outcome, $E\{y_i | y_i > 0\}$.

(b) Explain why the zero values should be included in the estimation of the model.

*INVOLVEMENT IN ANY **EXAMINATION IRREGULARITY** SHALL LEAD TO **DISCONTINUATION***

(c) The Tobit model may also be written in general form as:

$$y_i^* = x_i'\beta_1 + \varepsilon_i \qquad i = 1, 2, \ldots, N \qquad (1.2)$$

$$y_i = y_i^*, \qquad \text{if } y_i^* > L$$

$$y_i = L, \qquad \text{if } y_i^* \leq L \qquad \text{where } L = \text{lower bound.}$$

If the conditional distribution of the latent variable y^* given x is specified, then the model parameters can be consistently and efficiently estimated by maximum likelihood (ML) estimation based on the conditional distribution of the censored y.

Since y is a transformation of y^*, $y = g(y^*)$, the probability density function (pdf) of the censored variable y may be obtain.

(i) What is the p.d.f. of the censored variable y?

(ii) Derive the first-order conditions for ML estimation of the unknown parameters β and σ^2. Show your work/derivation.

Question (2) [15 marks]

Consider the following panel data model where the dependent variable, y_{it} may be described as a linear function of explanatory variables, $\boldsymbol{x}_{it}$ based on some fundamental economic theory. If the relationship is static, the econometric expression maybe written as:

$$y_{it} = \alpha_i + x_{it}'\beta + u_{it} \qquad i = 1, 2, \ldots, N; t = 1, 2, \ldots, T \qquad (2.1)$$

where: $u_{it} \sim IID(0, \sigma_u^2)$, and $\boldsymbol{x}_{it}$ is uncorrelated with u_{it}.

(a) Explain the following variations of panel data model in (2.1)

(i) Pooled panel model
(ii) Fixed Effect model
(iii) Random Effects model

(b) Suppose α_i is assumed to be random over individuals. It is identically and independently distributed (iid) with zero mean and constant variance, σ_α^2. Show that the error component $v_{it} = \alpha_i + u_{it}$ has a variance-covariance matrix:

$$\mathrm{E}(\boldsymbol{v}\boldsymbol{v}') = \boldsymbol{\Omega}_v = \boldsymbol{I}_N \otimes \mathrm{E}(\boldsymbol{v}_i\boldsymbol{v}_i') = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_v$$

Where $\boldsymbol{\Sigma}_v = \sigma_\alpha^2 \boldsymbol{j}_T \boldsymbol{j}_T' + \sigma_u^2 \boldsymbol{I}_T$

$\boldsymbol{j}_T$ is a column vector of 1's with dimension T.

(c) The Tobit model may also be written in general form as:

$$y_i^* = x_i'\beta_1 + \varepsilon_i \qquad i = 1, 2, \ldots, N \qquad (1.2)$$

$$y_i = y_i^*, \qquad \text{if } y_i^* > L$$

$$y_i = L, \qquad \text{if } y_i^* \leq L \qquad \text{where } L = \text{lower bound.}$$

If the conditional distribution of the latent variable y^* given x is specified, then the model parameters can be consistently and efficiently estimated by maximum likelihood (ML) estimation based on the conditional distribution of the censored y.

Since y is a transformation of y^*, $y = g(y^*)$, the probability density function (pdf) of the censored variable y may be obtain.

(i) What is the p.d.f. of the censored variable y?

(ii) Derive the first-order conditions for ML estimation of the unknown parameters β and σ^2. Show your work/derivation.

Question (2) [15 marks]

Consider the following panel data model where the dependent variable, y_{it} may be described as a linear function of explanatory variables, $\boldsymbol{x}_{it}$ based on some fundamental economic theory. If the relationship is static, the econometric expression maybe written as:

$$y_{it} = \alpha_i + x_{it}'\beta + u_{it} \qquad i = 1, 2, \ldots, N;\ t = 1, 2, \ldots, T \qquad (2.1)$$

where: $u_{it} \sim IID(0, \sigma_u^2)$, and $\boldsymbol{x}_{it}$ is uncorrelated with u_{it}.

(a) Explain the following variations of panel data model in (2.1)

(i) Pooled panel model
(ii) Fixed Effect model
(iii) Random Effects model

(b) Suppose α_i is assumed to be random over individuals. It is identically and independently distributed (iid) with zero mean and constant variance, σ_α^2. Show that the error component $v_{it} = \alpha_i + u_{it}$ has a variance-covariance matrix:

$$\mathrm{E}\,(\boldsymbol{vv}') = \boldsymbol{\Omega}_v = \boldsymbol{I}_N \otimes \mathrm{E}\,(\boldsymbol{v}_i\boldsymbol{v}_i') = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_v$$

Where $\boldsymbol{\Sigma}_v = \sigma_\alpha^2 \boldsymbol{j}_T \boldsymbol{j}_T' + \sigma_u^2 \boldsymbol{I}_T$

$\boldsymbol{j}_T$ is a column vector of 1's with dimension T.

(c) The Tobit model may also be written in general form as:

$$y_i^* = x_i'\beta_1 + \varepsilon_i \qquad i = 1, 2, \ldots, N \qquad (1.2)$$

$$y_i = y_i^*, \qquad \text{if } y_i^* > L$$

$$y_i = L, \qquad \text{if } y_i^* \leq L \qquad \text{where } L = \text{lower bound.}$$

If the conditional distribution of the latent variable y^* given x is specified, then the model parameters can be consistently and efficiently estimated by maximum likelihood (ML) estimation based on the conditional distribution of the censored y.

Since y is a transformation of y^*, $y = g(y^*)$, the probability density function (pdf) of the censored variable y may be obtain.

(i) What is the p.d.f. of the censored variable y?

(ii) Derive the first-order conditions for ML estimation of the unknown parameters β and σ^2. Show your work/derivation.

Question (2) [15 marks]

Consider the following panel data model where the dependent variable, y_{it} may be described as a linear function of explanatory variables, $\boldsymbol{x}_{it}$ based on some fundamental economic theory. If the relationship is static, the econometric expression maybe written as:

$$y_{it} = \alpha_i + x_{it}'\beta + u_{it} \qquad i = 1, 2, \ldots, N; t = 1, 2, \ldots, T \qquad (2.1)$$

where: $u_{it} \sim IID(0, \sigma_u^2)$, and $\boldsymbol{x}_{it}$ is uncorrelated with u_{it}.

(a) Explain the following variations of panel data model in (2.1)

(i) Pooled panel model
(ii) Fixed Effect model
(iii) Random Effects model

(b) Suppose α_i is assumed to be random over individuals. It is identically and independently distributed (iid) with zero mean and constant variance, σ_α^2. Show that the error component $v_{it} = \alpha_i + u_{it}$ has a variance-covariance matrix:

$$\text{E}(\boldsymbol{vv}') = \boldsymbol{\Omega}_v = \boldsymbol{I}_N \otimes \text{E}(\boldsymbol{v}_i\boldsymbol{v}_i') = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_v$$

Where $\boldsymbol{\Sigma}_v = \sigma_\alpha^2 \boldsymbol{j}_T \boldsymbol{j}_T' + \sigma_u^2 \boldsymbol{I}_T$

$\boldsymbol{j}_T$ is a column vector of 1's with dimension T.

(c) Model (2.1) was applied to a Cobb-Douglas production function on a dataset containing observations on 35 African countries, assuming different structure of the individual effects, α_i. Based on the results given on Table 2.1, which of the three models are consistent with the African dataset? Justify your answer and interpret the results from the model of your choice.

Table 2.1: Panel data estimates of the growth equation

Dependent variable: log of production	*Pooled* Coefficients (t-values)	*FE* Coefficients (t-values)	*RE* Coefficients (t-values)
Constant	-0.5330* (-1.77)	2.123*** (9.66)	1.820*** (5.73)
Elasticity of output with respect to capital	0.9144*** (16.08)	0.371*** (8.25)	0.447*** (10.67)
Elasticity of output with respect to labor	0.0856 (1.50)	0.629*** (13.98)	0.553*** (6.29)
Adjusted R-squared	0.72	0.99	0.73
No. of cross-sections	35	35	35
Total no. of observations	140	140	140
H0: all $\alpha_i = \alpha$: $F_{calc} = 796.0$*** *Hausman test*: $\chi^2(2) = 18.70$***			

Notes: Figures in parentheses are *t*-values using robust variance estimate. *** Significance at 1%, ** significance at 5 %; * significant at 10%.

Question (3) [15 Marks]

Greene (2012, pp. 694-696) presented a study on whether a new method of teaching called the Personalized System of Instruction (*PSI*), significantly influenced performance in later economics courses.

The "dependent variable" used in the application is *GRADE*, which indicates whether the student's grade in an intermediate macroeconomics course was higher than that in the principles course.

The explanatory variables are *GPA* denotes the student's grade point average; *TUCE* the score on a pretest that indicates entering knowledge of the material; and *PSI* the binary variable indicator of whether the student was exposed to the new teaching method.

The estimated results using logit and probit and models are presented below:

Variable	*Logistic*				*Probit*			
	Coeffs	*p-value*	$\frac{\partial F(y_i=j)}{\partial x_{ik}}$	*p-value*	*Coeffs*	*p-value*	$\frac{\partial F(y_i=j)}{\partial x_{ik}}$	*p-value*
constant	-13.021	0.008	--	--	-7.452	0.003	--	--
GPA	2.826	0.025	0.534	0.024	1.626	0.019	0.533	0.022
TUCE	0.095	0.501	0.018	0.493	0.052	0.537	0.017	0.531
PSI	2.379	0.025	0.456	0.012	1.426	0.017	0.464	0.006

(a) Compare the estimated results from the logit and probit model. Describe the differences and the similarities between the two sets of results.

(b) Specify the binary probit model and its assumptions used in the estimation.

(c) (i) Derive the probability that a student's grade in an intermediate macroeconomics course was higher than that in the principles course.

(ii) Derive the probability that a student's grade in an intermediate macroeconomics course was higher than that in the principles course

(d) Derive the log-likelihood equation and the first order and second order conditions with respect to the parameters of interest.

Question (4) [15 Marks]

In the context of multinomial models, the individuals may face more than two alternatives to choose from. Depending on the type of observations on the explanatory variables (choice attributes, z_{ij}, or individual characteristics, x_i, where index i refers to ith individual and j stands for a particular choice), the probability of the individual choosing j, j = 1, 2, ..., M may be expressed as:

$$P_{ij} = P\{y_i = j|z_{ij}\} = \frac{\exp\{z'_{ij}\gamma\}}{\sum_{l=1}^{M} \exp\{z'_{il}\gamma\}} \qquad (4.1)$$

for conditional logit, or

INVOLVEMENT IN ANY ***EXAMINATION IRREGULARITY*** *SHALL LEAD TO* ***DISCONTINUATION***

$$P_{ij} = P\{y_i = j|x_i\} = \frac{\exp\{x_i'\beta_j\}}{\sum_{l=1}^{M}\exp\{x_i'\beta_l\}} \qquad (4.2)$$

for multinomial logit.

(a) Explain the difference between the conditional logit and the multinomial logit models.

(b) From multinomial logit, derive the marginal effects, that is the change in the probability of $y_i = j$, with respect to a given change in in x_{ik}.

(c) The two models given in (4.1) and (4.2) can be combined into what some authors called a "mixed logit model". For this model, write the probability that $y_i = j$ given x_i and z_{ij}.

(d) What is the main assumption in the multinomial logit? How does one test this assumption?

(e) If this assumption fails, what alternative models may be estimated? Briefly describe each model.

Question (5) [15 marks]

A common assumption in count data models is that, for given x_i, the count variable y_i has a Poisson distribution with expectation $\lambda_i \equiv \exp\{x_i'\beta\}$. Therefore, it follows that the density of the Poisson regression model for a single observation is given by

$$f(y_i \mid x_i) = \frac{e^{-\exp(x_i'\beta)}\exp(x_i'\beta)^{y_i}}{y_i!} \qquad y = 0, 1, 2, \ldots .$$

(a) Give some characteristics of count data.

(b) Derive a log-likelihood function, which the Poisson ML estimator maximizes.

(c) Derive the expression for the marginal effects, that is, the impact of a change in x_{ik} on the expected value of y_i. [where x_{ik} is a continuous variable.]

(d) What are the major deficiencies of the Poisson model?

(e) Explain the negative binomial regression model as an alternative to the Poisson model.

Annex I. Relevant Formulae and Probability Distributions

Normal Distribution

If x is a normal random variable with mean $= \mu$, and variance $= \sigma^2$, then

p.d.f. $\phi(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \dfrac{-1}{2\sigma^2}(x-\mu)^2 \right\}$

Standard Normal Distribution

p.d.f. $\phi(x) = \dfrac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}$

c.d.f. $\Phi(x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} \exp\{-\dfrac{t^2}{2}\} dt$

Logistic distribution

c.d.f. $F(z) = \Lambda(z) = \dfrac{e^z}{1+e^z}$

p.d.f. $f(z) = \Lambda(z)[1-\Lambda(z)]$

Truncated Moments of the Standard Normal:

Suppose the truncated variable, $z \sim N[0, 1]$, then the left-truncated moments of z are:

$E[z|z > c] = \dfrac{\phi(c)}{[1-\Phi(c)]}$

$E[z|z > -c] = \dfrac{\phi(c)}{\Phi(c)}$

****

*INVOLVEMENT IN ANY **EXAMINATION IRREGULARITY** SHALL LEAD TO **DISCONTINUATION***

Annex I. Relevant Formulae and Probability Distributions

Normal Distribution

If x is a normal random variable with mean $= \mu$, and variance $= \sigma^2$, then

p.d.f. $\phi(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \dfrac{-1}{2\sigma^2}(x-\mu)^2 \right\}$

Standard Normal Distribution

p.d.f. $\phi(x) = \dfrac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}$

c.d.f. $\Phi(x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} \exp\{-\dfrac{t^2}{2}\} dt$

Logistic distribution

c.d.f. $F(z) = \Lambda(z) = \dfrac{e^z}{1+e^z}$

p.d.f. $f(z) = \Lambda(z)[1-\Lambda(z)]$

Truncated Moments of the Standard Normal:

Suppose the truncated variable, $z \sim N[0, 1]$, then the left-truncated moments of z are:

$E[z|z > c] = \dfrac{\phi(c)}{[1-\Phi(c)]}$

$E[z|z > -c] = \dfrac{\phi(c)}{\Phi(c)}$

****