AFRICAN ECONOMIC RESEARCH CONSORTIUM

Collaborative Masters Programme in Economics for Anglophone Africa (Except Nigeria)

JOINT FACILITY FOR ELECTIVES (JFE) 2017

JUNE – SEPTEMBER

ECONOMETRICS THEORY AND PRACTICE II

Second Semester: Final Examination

Duration: 3 Hours **Date: Thursday, September 21, 2017**

INSTRUCTIONS:

1. This examination is divided into two parts: **Section A** and **Section B**.
2. You are required to answer **ANY TWO QUESTIONS** in **EACH SECTION**.
3. All questions carry 15 marks each.

Section A:

Answer ANY TWO Questions in this Section

Question 1

The following two-way panel data model is considered:

$$y_{it} = \varphi + W_{it}'\beta + u_{it}, \quad i = 1,\ldots,N, \quad t = 1,\ldots,T \tag{1}$$

with

$$u_{it} = \mu_i + \lambda_t + \nu_{it} \tag{2}$$

where i and t denote the cross-section and time-series dimensions respectively, φ is a scalar, β a $K \times 1$ vector of coefficients and W_{it} the related set of explanatory variables.

(i) Based on the panel data terminology, define the four quantities involved in Equation (2). **[2 marks]**

(ii) Why is Equation (2) labeled a two-way error structure? **[1 mark]**

(iii) Is Equation (1) a static or dynamic panel data model? Why? **[1 mark]**

(iv) The researcher finally decides to assume a one-way error structure and goes for the fixed effects estimation approach. After a bit of algebra, he rewrites the model in matrix form as:

$$y = Z\delta + Z_\mu \mu + \nu \text{ with } Z = [\iota_{NT} \quad W]; \delta = \begin{bmatrix} \varphi \\ \beta \end{bmatrix}; Z_\mu = I_N \otimes \iota_T .$$

Derive the fixed effects estimator of the slope coefficient β as an outcome of a GLS approach. **[5 marks]**

(v) Provide a detailed procedure for the estimation of all the parameters (slope coefficients, constant term and individual effects) of the regression equation. **[6 marks]**

Question 2

(i) Let us consider the following Linear Probability Model (LPM, hereafter), $y_i = x_i'\beta + u_i$ for $i = 1,\ldots,n$, where the dependent variable is binary and x_i is a $K \times 1$ vector of individual explanatory factors. Explain why the LPM is not a wise choice of modeling. **[3 marks]**

(ii) In case the error follows a standard normal distribution, which model would you use? **[1 mark]**

(iii) Write down the related likelihood and log-likelihood functions as well as the first order conditions. **[5 marks]**

(iv) We are now interested in estimating a true binary choice model, rather than the LPM. The response variable takes on the values 0 or 1 with $y_i = 1$ if $y_i^* = x_i\beta + \varepsilon_i < \alpha$ where α nonnegative and x_i is a row vector of K columns. The variance of the zero-mean error term ε_i is known to be equal to $\frac{\pi^2}{3}$.

(a) Which model is appropriate? Why? **[1 mark]**

(b) Derive the log-likelihood function and the first order conditions of its maximization. **[5 marks]**

Question 3

Let T be a duration time following the log-logistic distribution of parameters (m,σ^2). By definition, it means that $Z=\dfrac{\ln T-m}{\sigma}$ is distributed according to a logistic law, with its cumulative and probability density functions given by: $F(z)=\dfrac{1}{1+\exp(-z)}$ and $f(z)=F(z)\left[1-F(z)\right]$ respectively.

(i) Show that $\exp\left(-\dfrac{\ln t-m}{\sigma}\right)=\dfrac{1}{ht^{\alpha}}$ with $h=\exp\left(-\dfrac{m}{\sigma}\right)$ and $\alpha=\dfrac{1}{\sigma}$. **[2 marks]**

(ii) Determine the density function g of the log-logistic variable T. **[5 marks]**

(iii) What is the cumulative density function G of T? **[2 marks]**

(iv) Deduce the survival function of T. **[2 marks]**

(v) Find the hazard function of T. What can you make as comments when $\alpha<1$? **[2 marks]**

Section B:

Answer ANY TWO Questions in this Section

Question 4

(i) Show that ordinary least squares, random effects and fixed effects estimators of γ in the dynamic panel data model below are all inconsistent for fixed T as N gets large:

$$y_{it}=\gamma y_{i,t-1}+u_{it},\quad i=1,\ldots,N,\quad t=1,\ldots,T \qquad (1)$$

with $u_{it}=\alpha_i+\eta_{it}$, where α_i and η_{it} are $iid\left(0,\sigma_{\alpha}^2\right)$ and $iid\left(0,\sigma_{\eta}^2\right)$ respectively. **[4 marks]**

(ii) By first-differencing the model, show that the new disturbance $v_{it}=u_{it}-u_{i,t-1}$ is a Moving Average process of order 1, i.e. MA(1), with unit root. **[1 mark]**

(iii) Derive the variance-covariance matrix of the new disturbances for a given individual, and thereafter for the overall system ($i=1,\ldots,N$). **[4 marks]**

(iv) Show that for period t, $\left(y_{i1}, y_{12}, \ldots, y_{i,t-2}\right)$ are valid instruments for the right-hand-side regressor in Equation (1). **[1 mark]**

(v) After defining the matrix of instruments, derive the one-step generalized method of moments (GMM) $\hat{\gamma}_1$ estimator of γ suggested by Arellano and Bond (1991).

[5 marks]

Question 5

We consider the choice of transport mode in a given low-income country. The different options available are: own car, bus, or minivan. The dependent variable (the choice) is made based on a set of k explanatory factors in a sample of size N.

(i) A debate raises among some of your classmates about the logit-type model to use. On which basis would you go for a Multinomial Logit model (MNL, hereafter) or for a Conditional Logit model (CL, hereafter)? Clearly explain your argumentation. What is the major common weakness of the two models? How can we test it? **[3 marks]**

(ii) Suggest two explanatory factors for each of the MNL and CL models potentially affecting the choice of transportation means. **[3 marks]**

(iii) You have decided to implement the MNL approach. Write down the probability $p_{ij} = \Pr\left(y_i = j \middle| x_i\right)$ for an individual i to choose the alternative j, with $i = 1, \ldots, N$ and $j = 1, \ldots, J$ where $J = 3$. **[1 mark]**

(iv) Provide the formula of the marginal effect of the lth explanatory variable on the probability of choosing alternative j. What is the difference in its interpretation compared to the standard linear regression model and to the binary regression?

[6 marks]

(v) Let β_j and β_j^* be two different coefficients' vectors related to alternative j. They differ just by a constant. Show that it is impossible to distinguish between these two vectors through the MNL procedure. What does it imply for the practical conduct of the MNL estimation? **[2 marks]**

Question 6

(i) Explain the difference between truncation and censoring. Provide a clear example of each notion. **[3 marks]**

(ii) In the case of a left-censored model, express mathematically the link between the observed dependent variable y and its corresponding latent variable y^*. Write down the probability density function of y in terms of the distribution of y^*. **[3 marks]**

(iii) Deduce the likelihood as well as the log-likelihood functions. Hint: Define an indicator function d that will be helpful. **[4 marks]**

(iv) The table below is the output of a regression done by an econometrician using STATA. The dataset contains 616 workers and the interest is in explaining the variable *pension* measuring their pension benefits in USD. The explanatory variables considered are the number of years of experience (*experience*), the respondent age (*age*), the number of years with current employer (*tenure*), the number of years of schooling (education), the number of dependents (*depends*), being married (*married*), and being member of a union (*union*).

```
                                                Number of obs     =       616
                                                LR chi2(7)        =    216.00
                                                Prob > chi2       =    0.0000
Log likelihood =  -3657.312                     Pseudo R2         =    0.0287
------------------------------------------------------------------------------
     pension |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  experience |   13.43862   5.452726     2.46   0.014     2.730193    24.14705
         age |   -9.46133   5.312918    -1.78   0.075    -19.89519    .9725337
      tenure |    28.6519   4.600697     6.23   0.000     19.61674    37.68705
   education |   111.5236   10.93757    10.20   0.000     90.04367    133.0036
     depends |   43.14056   21.24504     2.03   0.043     1.418127     84.863
     married |   98.33365   68.44222     1.44   0.151    -36.07777    232.7451
       union |   466.3035   63.48515     7.35   0.000     341.6271    590.9799
       _cons |  -1262.156   203.4932    -6.20   0.000     -1661.79   -862.5229
-------------+----------------------------------------------------------------
      /sigma |   666.9027   23.64846                      620.4603    713.3451
------------------------------------------------------------------------------
          172  left-censored observations at pension <= 0
          444     uncensored observations
            0 right-censored observations
```

According to you, which regression model did the econometrician run? Explain your answer by giving a clear description of the data structure. **[2 marks]**

(v) Interpret the marginal effect of the education and the marital status. **[3 marks]**

```
Conditional marginal effects                    Number of obs     =        616
Model VCE     : OIM
Expression    : E(pension*|pension>0), predict(ystar(0,.))
dy/dx w.r.t. : experience age tenure education depends married union
at            : experience      =    18.61688 (mean)
                age               =    37.65422 (mean)
                tenure            =    7.751623 (mean)
                education         =    12.51461 (mean)
                depends           =    1.227273 (mean)
                married           =    .6866883 (mean)
                union             =    .3181818 (mean)
------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  experience |    10.5024   4.264459     2.46   0.014     2.144211    18.86058
         age |   -7.39411   4.154051    -1.78   0.075     -15.5359    .7476795
      tenure |    22.3917   3.597758     6.22   0.000     15.34022    29.44318
   education |   87.15665   8.530963    10.22   0.000     70.43627     103.877
     depends |   33.71472   16.61656     2.03   0.042     1.146868    66.28257
     married |   76.84858   53.47519     1.44   0.151    -27.96087     181.658
       union |   364.4202   49.41403     7.37   0.000     267.5704    461.2699
------------------------------------------------------------------------------
```