PhD in Economics (14th Cycle)
Econometrics test (2013-06-14)

Name: ____

> You are asked to complete **two out of the three** exercises below. Of course, you can attempt all of them if you want, but in that case please indicate which ones you want us to evaluate.

1. Consider a vector of two different estimators for the same parameter

$$m = \begin{bmatrix} \hat{\theta} \\ \tilde{\theta} \end{bmatrix}$$

whose asymptotic properties are

$$m \xrightarrow{\text{p}} \mu = \begin{bmatrix} \theta \\ \theta \end{bmatrix}$$
$$\sqrt{n}(m - \mu) \xrightarrow{\text{d}} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix} \right).$$

We assume $\theta > 0$. We now consider the problem of devising a new estimator of θ by combining, somehow, the ones we already have.

(a) Prove that $\bar{\theta} = \alpha\hat{\theta} + (1-\alpha)\tilde{\theta}$ is consistent for any choice of α;

(b) find the asymptotic distibution of $\bar{\theta}$;

(c) prove that $\alpha = 1$ is the best choice in terms of efficiency of $\bar{\theta}$;

(d) prove that $\bar{\bar{\theta}} = \hat{\theta}^{\alpha} \cdot \tilde{\theta}^{1-\alpha}$ is consistent for any choice of α;

(e) prove that, if we pick the optimal $\bar{\theta}$ and $\bar{\bar{\theta}}$, then $\bar{\theta} = \bar{\bar{\theta}}$.

2. Given the following model for the time series y_t

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \theta\varepsilon_{t-1} + \varepsilon_t,$$

where $0 < \theta \leq 0.5$ and ε_t is Gaussian a white noise with zero mean and variance 1.

(a) calculate $Var(y_t)$;

(b) indicate how you would carry out a test for the hypothesis $H_0 : \phi_2 = 0.5\phi_1$ (indicate the parameters covariance matrix with $\hat{V}$);

(c) show that, under H_0, the model is not necessarily stationary;

(d) prove that, if a real number α exists such that

$$\phi_1 = \alpha - \theta \quad \text{and} \quad \phi_2 = \alpha\theta$$

then the ARMA model has a common factor.

(e) ML estimation provides the following results:

```
Model 1: ARMA, using observations 3-356 (T = 354)
Estimated using BHHH method (conditional ML)
Dependent variable: y

             coeff.    std.err      z      p-value
  -------------------------------------------------------
  phi_1       0.590    0.204      2.892    0.0038  ***
  phi_2       0.215    0.205      1.049    0.2942
  theta_1     0.170    0.085      2.000    0.0455  **
```

```
Mean dependent var  -15.10660   S.D. dependent var   11.24623
Mean of innovations -0.000080   S.D. of innovations  0.967049
Log-likelihood      -490.4432   Akaike criterion     990.8864
Schwarz criterion    1010.233   Hannan-Quinn         998.5838

Parameters covariance matrix (V)
    0.0416    -0.0350    -0.0028
   -0.0350     0.0420     0.0000
   -0.0028     0.0000     0.0072

                           Real  Imaginary   Modulus
     -------------------------------------------------
     AR
       Root  1           0.995     0.000     0.995
       Root  2          -2.413     0.000     2.413
     MA
       Root  1          -5.303     0.000     5.303
     -------------------------------------------------

Test for autocorrelation up to order 4
  Ljung-Box Q' = 3.21353, with p-value = 0.07303

Test for ARCH of order 4
  Test statistic: LM = 1.59911, with p-value = 0.808951
```

Provide some comments to this estimates and calculate the numerical solutions to the above questions (a) and (b).

(a) $Var(y_t) =$ ________________

(b) Test: ________ Distribution: ________ Test stat.: ____________
Result: ACCEPT ◯ REJECT ◯

3. Let y be a binary variable equal to 1 if a woman is employed and 0 otherwise and $\mathbf{x}$ a set of K regressors. Suppose that y follows the model

$$\Pr(y = 1|\mathbf{x}) = 1 - [1 + \exp(\mathbf{x}'\boldsymbol{\beta})]^{-\alpha}, \quad (1)$$

where $\alpha > 0$ is a parameter to be estimated along with the $K \times 1$ vector $\boldsymbol{\beta}$. The model in Equation (1) is called skewed logit model (*scobit*). The $K \times 1$ vector $\mathbf{x}$ contains the constant and a set of characteristics of the sample units: income of other members of the household (NWINC), education (EDUC), labour market experience (EXPER), age (AGE), and a binary indicator for the presence of kids younger that 6 in the household (KIDS6).

(a) Write down the contribution to the log-likelihood function of random draw i.

(b) Find the partial effect of a continuous regressor x_k on $\Pr(y = 1|\mathbf{x})$.

(c) Show that if $\alpha = 1$ the scobit model corresponds to the standard logit model.

(d) Table 1 reports the estimation results of the scobit and logit models. Which one would you prefer: logit or scobit? Why? Comment on the impact of the regressors on the probability of being employed.

(e) Table 2 reports the estimation results when we include as a further regressor the square of experience (EXPERSQ). Which model would you choose among the 4 models presented in Tables 1 and 2? Why?

Table 1: Estimation results: scobit vs logit

Scobit model: ML, using observations 1-753

	estimate	std. error	z	p-value	
Constant	0.196962	1.32343	0.1488	0.8817	
NWINC	−0.0173858	0.00757124	−2.296	0.0217	**
EDUC	0.191607	0.0503793	3.803	0.0001	***
EXPER	0.0949536	0.0256899	3.696	0.0002	***
AGE	−0.0812164	0.0178236	−4.557	5.20e-06	***
KIDS6	−1.25947	0.270005	−4.665	3.09e-06	***
ln(alpha)	0.623186	0.887749	0.7020	0.4827	

Log-likelihood	−406.1494	Akaike criterion	826.2988
Schwarz criterion	858.6673	Hannan-Quinn	838.7687

****

Logit model: Logit, using observations 1-753

	coefficient	std. error	z	p-value	
Constant	1.15322	0.742071	1.554	0.1202	
NWINC	−0.0198998	0.00826779	−2.407	0.0161	**
EDUC	0.223366	0.0429694	5.198	2.01e-07	***
EXPER	0.117887	0.0133856	8.807	1.29e-018	***
AGE	−0.0951414	0.0134388	−7.080	1.45e-012	***
KIDS6	−1.46358	0.200354	−7.305	2.77e-013	***

Log-likelihood	−406.4583	Akaike criterion	824.9165
Schwarz criterion	852.6609	Hannan-Quinn	835.6051

Table 2: Estimation results when squared experience is included in the model specification

Scobit model: ML, using observations 1-753

	estimate	std. error	z	p-value	
Constant	−7.98380	1090.54	−0.007321	0.9942	
NWINC	−0.0147901	0.00630920	−2.344	0.0191	**
EDUC	0.149703	0.0396536	3.775	0.0002	***
EXPER	0.138242	0.0323753	4.270	1.95e-05	***
AGE	−0.0601462	0.0153812	−3.910	9.22e-05	***
KIDS6	−1.01165	0.233575	−4.331	1.48e-05	***
EXPERSQ	−0.00226147	0.000725623	−3.117	0.0018	***
ln(alpha)	7.94985	1090.27	0.007292	0.9942	

Log-likelihood	−399.6862	Akaike criterion	815.3723
Schwarz criterion	852.3649	Hannan-Quinn	829.6237

****

Logit model: Logit, using observations 1-753

	coefficient	std. error	z	p-value	
Constant	0.750220	0.760372	0.9866	0.3238	
NWINC	−0.0210297	0.00842723	−2.495	0.0126	**
EDUC	0.217591	0.0431275	5.045	4.53e-07	***
EXPER	0.203687	0.0318759	6.390	1.66e-010	***
AGE	−0.0922022	0.0136754	−6.742	1.56e-011	***
KIDS6	−1.46815	0.202396	−7.254	4.05e-013	***
EXPERSQ	−0.00314252	0.00101454	−3.097	0.0020	***

Log-likelihood	−402.0892	Akaike criterion	818.1783
Schwarz criterion	850.5468	Hannan-Quinn	830.6483