Name: _____

Matricola: _____ email: _____

ECONOMETRICS - 19/06/2019 - Time: 2 h

1. Say if the following statements are unambiguously true (True), unambiguously false (False) or impossible to classify the way they are stated (Not necessarily). Write the motivations to your answers **only** in the space provided. A "Not necessarily" answer with no motivations will be considered wrong.

   (a) If $E(\mathbf{x}) = \mathbf{0}$, then $E(\mathbf{x}'\mathbf{x}) = 0$.

   True ◯ False ◯ Not necessarily ◯

   (b) If $E(\mathbf{x}) = \mathbf{0}$, then the matrix $E(\mathbf{x}\mathbf{x}')$ is diagonal.

   True ◯ False ◯ Not necessarily ◯

   (c) In a linear regression model of the type $\log y_i = \beta_0 + \beta_1 x_i + \beta_2 \log x_i + \varepsilon_i$, the coefficient β_2 can be interpreted as the elasticity of y with respect to x.

   True ◯ False ◯ Not necessarily ◯

   (d) Heteroskedasticity makes the OLS estimator inconsistent.

   True ◯ False ◯ Not necessarily ◯

   (e) In a dynamic regression model, the long-run multiplier may be 0 even if all the dynamic multipliers are non-zero.

   True ◯ False ◯ Not necessarily ◯

2. Given a sample of size n, suppose you have the following moment matrices:

$$\mathbf{X}'\mathbf{X} = n \cdot \begin{bmatrix} 1 & q \\ q & q^2+1 \end{bmatrix} \qquad \mathbf{X}'\mathbf{y} = n \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{y}'\mathbf{y} = n \cdot [q(q-4)+9],$$

where q is a positive real number.

(a) Express the OLS statistic as a function of n and q:

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \\ \\ \end{bmatrix}$$

(b) Express the sum of squared residuals as a function of n and q:

$\mathbf{e}'\mathbf{e} =$ _____

(c) Express the variance estimator $\hat{\sigma}^2$ as a function of n and q:

$\hat{\sigma}^2 =$ _____

(d) Find the value of q that minimises the centred R^2 index:

$q^* = \text{Argmin}\, R^2(q) =$ _____

(e) Find the centred R^2 index if $q = q^*$:

$R^2 =$ _____

(f) Test the hypothesis $\beta_2 = 0$ if $n = 100$ and $q = 1$

Test type: _____ Distribution: _____ Test statistic: _____
Decision: ◯ Reject ◯ Don't reject

3. We have a sample of 6750 people from the 2016 SHIW survey (Bank of Italy) who belong to the labour force. We estimate by OLS the following model:

$$y_i = \beta_0 + \beta_1 g_i + \beta_2 c_i + \beta_3 a_1 + \beta_4 a_i^2 + \beta_5 e_1 + \beta_6 e_i^2 + \beta_7 C_1 + \beta_8 S_i + \varepsilon_i$$

where the variables are as follows (descriptive statistics are provided in Table 1):

Variable	**label**	**Description**
y_i	`employed`	employment status: 1 = employed, 0 = unemployed
g_i	`male`	gender: 1 = male, 0 = female
c_i	`italian`	citizenship status: 1 = Italian, 0 = other
a_i	`age`	age (years)
e_i	`educ`	education (years)
C_i	`Centre`	geographical dummy: 1 = lives in Central Italy
S_i	`South`	geographical dummy: 1 = lives in Southern Italy

Note that, since the dependent variable is a binary variable, $E(y_i)$ can also be seen as the probability that $y_i = 1$, that is, the individual being employed. The estimates are provided in Table 2.

Variable	Mean	Median	S.D.	Min	Max
`employed`	0.787	1.00	0.410	0.00	1.00
`male`	0.583	1.00	0.493	0.00	1.00
`italian`	0.932	1.00	0.251	0.00	1.00
`age`	44.7	47.0	12.7	16.0	89.0
`educ`	11.7	13.0	3.48	5.00	20.0
`Centre`	0.211	0.00	0.408	0.00	1.00
`South`	0.371	0.00	0.483	0.00	1.00

Table 1: Descriptive statistics

(a) Describe the effect of age on the probability of being employed

(b) Describe the effect of education on the probability of being employed

(c) Comment on all other aspect of the model that you consider relevant

Dependent variable: employed
Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value
const	−0.505481	0.0735130	−6.876	0.0000
male	0.0268720	0.00900863	2.983	0.0029
italian	−0.0297232	0.0186783	−1.591	0.1116
age	0.0398565	0.00234493	17.00	0.0000
age2	−0.000348403	2.60123e–05	−13.39	0.0000
educ	0.0437600	0.00842887	5.192	0.0000
educ2	−0.00102342	0.000333320	−3.070	0.0021
Centre	−0.0523231	0.0109701	−4.770	0.0000
South	−0.194864	0.0106597	−18.28	0.0000

Mean dependent var	0.786963	S.D. dependent var	0.409484
Sum squared resid	907.5285	S.E. of regression	0.366917
R^2	0.198050	Adjusted R^2	0.197099

White's test for heteroskedasticity: Test statistic: LM = 810.736, p-value = 2.70278e-146

Table 2: OLS estimates