ISEG – Lisbon School of Economics and Management
ECONOMETRICS
First Semester 2017/2018
Problem Set I

Question:	**1**	**2**	**3**	**4**	**5**	Total
Points:	4	4	4	4	34	50

Justify all your answers. You are required to show your work on each problem (except for multiple choice questions). **Organize your work**. Work scattered all over the page will receive very little credit. A correct answer in a multiple choice question worths 4 points; an incorrect one worths -1 point. **Delivery date: 10 of October**.

(4) **1**. Consider the following Multiple Linear Regression Model (MLRM)

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + u_i, \quad i = 1, \ldots, n$$

Assume that the assumptions MLR.1 to MLR.5 hold. For the OLS estimator of the unknown coefficients, which of the following statements is **true**?

- ◯ The sum of squared residuals (SSR) is equal to zero.
- ◯ It is proven that $\sum_{i=1}^{n} u_i = 0$.
- ◯ The OLS minimizes the residuals, therefore $\hat{u}_i = 0$.
- √ **The SSR is minimum.**

(4) **2**. Which of the following statements is **true**?

- ◯ The R-squared, R^2, cannot be calculated if the dependent variable is logarithmic.
- ◯ Adding an irrelevant variable to the model may result in a reduction of the R^2.
- √ **The R^2 is always greater or equal than the adjusted *R*-squared, $\bar{R}^2$.**
- ◯ The R^2 can be negative if the explanatory variables of the model are strongly correlated.

(4) **3**. Consider the following MLRM:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i, \quad i = 1, \ldots, n$$

Assume that the assumptions MLR.1 to MLR.4 hold. For the OLS estimator of the unknown coefficients, choose the **correct** statement:

- ◯ The OLS estimator is the Best Linear Unbiased Estimator (BLUE).
- ◯ $\mathrm{E}(y_i \mid x_{i1}, x_{i2}) = 0$.
- ◯ Because the assumption MLR.5 does not hold, the OLS estimator may be biased.
- √ **The error term, u_i, is uncorrelated with x_{i1} and x_{i2}.**

(4) **4**. The omission of a relevant variable in a given model:

- ◯ Is never a problem if one is not interested on estimating the coefficient of that variable.
- √ **Implies that the assumption MLR.4 does not hold, if the omitted variable is correlated with at least one of the explanatory variables included in the model.**
- ◯ Implies that the OLS estimator is necessarily biased.
- ◯ Increases the variance of the OLS estimates.

5. Use the data set apple.WF1, to explain the quantity (in pounds) of ecolabeled apples purchased by a family, *ecolbs*.

(7) (a) Estimate the following regression by OLS:

$$ecolbs_i = \beta_0 + \beta_1 \log(faminc_i) + \beta_2 regprc_i + \beta_3 ecoprc_i + u_i$$

where $faminc$ is the family income (in thousands of dollars), $regprc$ is the price of regular apples (in dollars), $ecoprc$ is the price of ecolabeled apples (in dollars). Write the estimated equation with the corresponding standard errors.

Solution:

Dependent Variable: ECOLBS
Method: Least Squares

Sample: 1 660
Included observations: 660

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.009834	0.691309	1.460756	0.1446
LOG(FAMINC)	0.238804	0.144406	1.653692	0.0987
REGPRC	3.035029	0.709904	4.275265	0.0000
ECOPRC	-2.881176	0.587793	-4.901686	0.0000

R-squared	0.040413	Mean dependent var	1.473990
Adjusted R-squared	0.036025	S.D. dependent var	2.525781
S.E. of regression	2.479868	Akaike info criterion	4.660330
Sum squared resid	4034.233	Schwarz criterion	4.687556
Log likelihood	-1533.909	Hannan-Quinn criter.	4.670883
F-statistic	9.209255	Durbin-Watson stat	2.017676
Prob(F-statistic)	0.000006		

$$\widehat{ecolbs} = 1.0098 + 0.2388 \log(faminc) + 3.0350\, regprc - 2.8812\, egoprc$$

$$se(\hat{\beta}_1) = 0.1444 \qquad se(\hat{\beta}_2) = 0.7099 \qquad se(\hat{\beta}_3) = 0.5878$$

(8) (b) Interpret the estimated coefficients. Discuss the signs of these estimates.

Solution:

$\hat{\beta}_1 = 0.2388$: a raise of 1% in the family income, *ceteris paribus*, makes the estimated quantity of ecolabeled apples bought increase, on average, by $\frac{0.2388}{100}$=0.002388 pounds.

- $\hat{\beta}_1$ is positive, corresponding to an income effect: if a family has more money to spend, their demand for ecolabeled apples will rise.

$\hat{\beta}_2 = 3.0350$: Holding all other factors fixed, if the price of regular apples increases by 1 dollar, families will buy, on average, an estimated more 3.0350 pounds of ecolabeled apples.

• $\hat{\beta}_2$ is positive, once again an expected result due to the substitution effect: a raise in the price of regular apples means that its price will become less competitive - the ecolabeled's price remains constant in this analysis - making families buy more of the ecolabeled type.

$\hat{\beta}_3 = -2.8812$: *ceteris paribus*, if the price of ecolabeled apples increases by 1 dollar, families will buy an estimated less 2.8812 pounds of that product (on average).

• $\hat{\beta}_3$ is negative, which follows the law of demand - if the price of ecolabeled apples increases, their demand is expected to fall.

(7) (c) Estimate the quantity of ecolabeled apples purchased by a family with an income of 45 thousand dollars when the price of both types of apples is equal to 1 dollar.

Solution:

Considering $faminc = 45$ (*faminc* is expressed in thousands of dollars), $regprc = 1$, $ecoprc = 1$, and making the substitution in the estimated equation, we get:

$$\widehat{ecolbs} = 1.0098 + 0.2388 \times \log(45) + 3.0350 \times 1 - 2.8812 \times 1 = 2.0726 \text{ pounds}$$

(6) (d) Suppose that the family referred in part (c) has, in fact, purchased 2 pounds of ecolabeled apples. Calculate the corresponding residual and comment on this result.

Solution:

For this family, we have $\widehat{ecolbs} = 2.0726$ (predicted value) and $ecolbs = 2$ (actual value).

Residual: $\hat{u} = ecolbs - \widehat{ecolbs} = 2 - 2.0726 = -0.0726$

The estimated model predicted a (slightly) higher consumption of ecolabeled apples than the value that actually ocurred, hence the negative value of $\hat{u}$. This family is consuming slightly under the amount estimated for the average consumption of the families facing the same characteristics.

(6) (e) Interpret the value obtained for the R^2 of the regression.

Solution:

The R^2 of the regression is 0.04013, which means that, for this sample, the variables *faminc*, *regprc* and *ecoprc* explain only 4.013% of the total variation in ecolabeled apples purchased.