

UNIVERSIDAD CARLOS III DE MADRID  
 ECONOMETRICS I  
 Academic year 2005/06  
 FINAL EXAM

September, 1, 2006

PROBLEM TEXT

PROBLEM 1: INVESTMENT AND MARKET VALUE

We are concerned with the relationship between corporate investment and its market value in the previous year. For this purpose, we use annual data from General Electric (GE) during the years 1935 to 1954 about investment,  $I$ , and market value in the previous year,  $VM(-1)$ , both in million dollars.

---



---

**OUTPUT 1**

---

Dependent Variable:  $I$   
 Method: Least Squares  
 Sample: 1 20  
 Included observations: 20

---



---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	29.658	52.005	0.57	0.576
$VM(-1)$	0.037	0.026	1.42	0.171
Log likelihood	-104.46	F-statistic		2.035
Durbin-Watson stat	0.32	Prob (F-statistic)		0.171

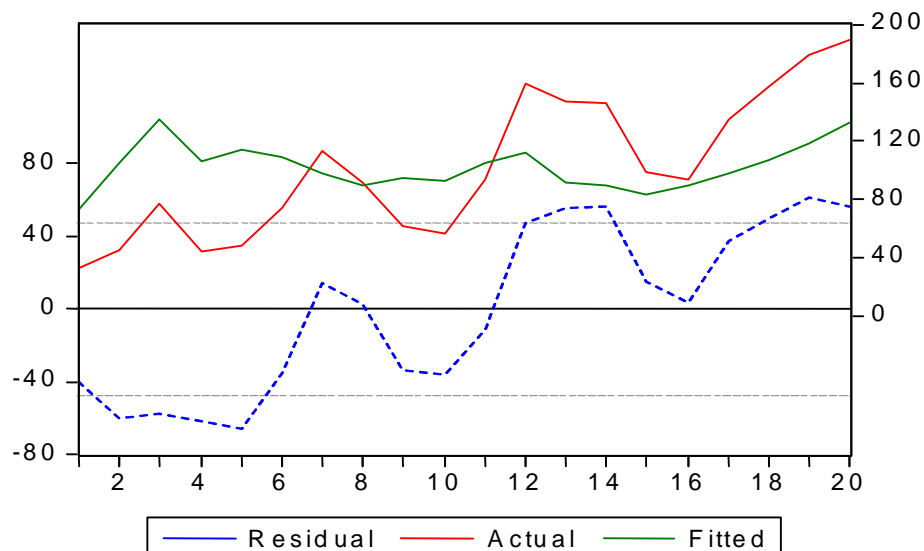
---



---

**FIGURE 1**

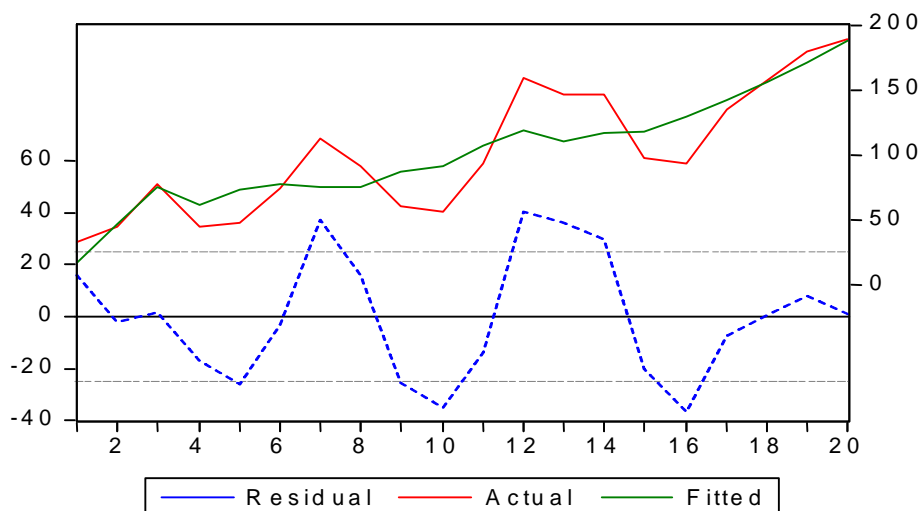
(Los residuos de la OUTPUT 1 son la línea discontinua)



We proceeded to include a trend ( $t$ ) as a further variable in the equation so as to see whether we can capture the investment growth of GE in those years. The results are shown in OUTPUT 2.

<b>OUTPUT 2</b>					
Dependent Variable: I					
Method: Least Squares					
Sample: 1 20					
Included observations: 20					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	
C	-21.723	28.456	-0.76	0.456	
t	6.716	0.974	6.90	0.000	
VM(-1)	0.028	0.014	1.98	0.064	
Log likelihood	-91.11909	F-statistic		27.422	
Durbin-Watson stat	1.18	Prob (F-statistic)		0.000	

**FIGURE 2**



In order to perform a formal test on whether there is second order autocorrelation, we obtained OUTPUT 3, where RES are the residuals from OUTPUT 2.

---

---

**OUTPUT 3**

---

Dependent Variable: RES  
Method: Least Squares  
Sample: 1 20  
Included observations: 18

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.459	3.824	-0.12	0.906
RES(-1)	0.716	0.175	4.09	0.010
RES(-2)	-0.725	0.173	-4.19	0.001

---

---

Log likelihood	-74.042	F-statistic	12.184
Durbin-Watson stat	2.48	Prob (F-statistic)	0.001

---

---

Finally, we estimated the same equation as in OUTPUT 2 but usign the Newey-West variance-covariance estimator. The results are shown in OUTPUT 4.

---

---

**OUTPUT 4**

---

Dependent Variable: I  
Method: Least Squares  
Sample: 1 20  
Included observations: 20  
Newey-West HAC Standard Errors & Covariance (lag truncation=2)

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-21.723	21.543	-1.01	0.327
t	6.716	0.745	9.01	0.000
VM(-1)	0.028	0.011	2.55	0.016

---

---

R-squared	0.763	Mean dependent var	102.29
Adjusted R-squared	0.736	S.D. dependent var	48.58
S.E. of regression	24.985	Akaike info criterion	9.412
Sum squared resid	10612.25	Schwarz criterion	9.561
Log likelihood	-91.119	F-statistic	27.422
Durbin-Watson stat	1.18	Prob (F-statistic)	0.000

---

---

## PROBLEM 2: ENGEL CURVE

In order to estimate the Engel curve for food expenditure, we have available data of Spanish households composed by couples with or without children in which the husband is between 25 and 65 years old, randomly selected among the Encuesta de Presupuestos Familiares (*Household Expenditure Survey*) for 1990-91 with information about the following variables:

**LAL** = natural logarithm of household annual food expenditure in euros;

**LGT** = natural logarithm of household total expenditure in euros;

**LY** = natural logarithm of household disposable income in thousand euros (this variable has a highly positive correlation with **LGT**);

**LY\_UH** =  $LY \times UH$ ;

**TAM** = Number of household members (excluding the spouses, that is, total number of members  $-2$ );

**TAM2** =  $TAM * TAM$  = Squared Number of household members (excluding the spouses);

**EDAD** = Husband age;

**UH** = Binary variable which takes on value 1 if the husband has a university degree and 0 otherwise;

**UM** = Binary variable which takes on value 1 if the wife has a university degree and 0 otherwise;

**MT** = Binary variable which takes on value 1 if the wife is currently working and 0 otherwise.

The empirical model that is used is:

$$\begin{aligned} \mathbf{LAL} = & \beta_0 + \beta_1 \mathbf{LGT} + \beta_2 \mathbf{LGT\_UH} + \beta_3 \mathbf{TAM} + \beta_4 \mathbf{TAM2} & (*) \\ & + \beta_5 \mathbf{UH} + \beta_6 \mathbf{UM} + \beta_7 \mathbf{MT} + \beta_8 \mathbf{EDAD} + u, \end{aligned}$$

that is, the variables determining household food expenditure are the natural logarithm of household total expenditure (**LGT**), as well as  $\mathbf{LGT\_UH} = \mathbf{LGT} \times \mathbf{UH}$ , and other variables which capture household characteristics.

Furthermore, we know that the sum of squares of the observed values of **LAL**, in deviations around its sample mean,  $\sum_{i=1}^n (Y_i - \bar{Y})^2$ , is equal to 202.643.

**Very important:** Notice that some results in the tables have been omitted.

Using a sample of 965 observations, we have obtained the following estimates:

---

---

**OUTPUT 1**

---

---

Dependent Variable: LAL

Method: Least Squares

Sample: 1 965

Included observations: 965

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.955	0.258	11.45	0.000
LGT				
LGT_UH	-0.177		-2.70	0.006
TAM	0.143	0.018	7.94	0.000
TAM2	-0.010	0.002	-5.00	0.000
UH	1.628	0.678	2.40	0.016
UM	-0.105	0.044	-2.39	0.017
MT	-0.071	0.029	-2.45	0.015
EDAD	0.004	0.001	4.00	0.006

---

---

R-squared

Adjusted R-squared

S.E. of regression

Sum squared resid 121.95

---

---

**IMPORTANT:**The confidence interval at the 95% level for the  
LGT coefficient is [0.46312 ; 0.57288]

---

---

---

---

**OUTPUT 2**

---

---

Covariance of the estimators in OUTPUT 1

---

---

	LGT	LGT_UH	TAM	TAM2	UH	UM	MT
LGT_UH	-0.0007						
TAM	-0.00006	0.000028					
TAM2	0.0000035	-0.0000052	-0.000039				
UH	0.0067	-0.046	-0.00028	0.000053			
UM	-0.000074	-0.000035	0.000073	-0.000008	-0.0004		
MT	-0.00016	0.000034	0.000025	0.0000005	-0.00034	-0.00033	
EDAD	-0.0000067	-0.0000037	-0.0000002	-0.0000001	0.000036	0.000003	0.00001

---

---

---

---

**OUTPUT 3**

---

---

Dependent Variable: LAL

Method: Least Squares

Sample: 1 965

Included observations: 965

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.733	0.270	10.12	0.000
LGT	0.565	0.029	19.48	0.000
LGT_UH	-0.181	0.071	-2.55	0.011
UH	1.667	0.711	2.34	0.019
UM	-0.142	0.046	-3.09	0.002
MT	-0.100	0.031	-3.23	0.001
EDAD	0.004	0.001	4.00	0.002

---

---

R-squared	0.335
Adjusted R-squared	
S.E. of regression	0.392
Sum squared resid	

---

---

---

---

**OUTPUT 4**

---

---

Dependent Variable: LAL

Method: Least Squares

Sample: 1 965

Included observations: 965

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.209	0.241	13.32	0.000
LGT	0.491	0.026	18.88	0.000
LGT_UH	-0.017	0.003	-5.67	0.000
TAM	0.148	0.018	8.22	0.000
TAM2	-0.011	0.003	-3.67	0.000
MT	-0.088	0.028	-3.14	0.002
EDAD	0.004	0.001	4.00	0.005

---

---

R-squared	
Adjusted R-squared	0.387
S.E. of regression	0.359
Sum squared resid	

---

---

---

---

**OUTPUT 4A**

---

---

Dependent Variable: LAL

Method: Least Squares

Sample: 1 965

Included observations: 965

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.579	0.233	15.36	0.000
LGT	0.449	0.025	17.96	0.000
TAM	0.153	0.018	8.50	0.000
TAM2	-0.011	0.003	-3.67	0.000
MT	-0.106	0.029	-3.66	0.000
EDAD	0.004	0.001	4.00	0.005

---

---

R-squared	0.374
Adjusted R-squared	
S.E. of regression	
Sum squared resid	

---

---

---

---

**OUTPUT 5**

---

---

Dependent Variable: LGT

Method: Least Squares

Sample: 1 965

Included observations: 965

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.163	0.085	96.04	0.000
LY	0.500	0.032	15.50	0.000
LY_UH	-0.113	0.068	-1.65	0.099
TAM	0.025	0.020	1.28	0.202
TAM2	0.001	0.003	0.21	0.836
UH	0.468	0.198	2.36	0.018
UM	-0.014	0.049	-0.29	0.770
MT	0.057	0.034	1.71	0.088
EDAD	0.004	0.001	2.61	0.009

---

---

R-squared	0.372
Adjusted R-squared	0.366
S.E. of regression	0.396
Sum squared resid	149.72

---

---

(The sum of squares of the residuals in a regression like the one in OUTPUT 5 but that omits LY and LY\_UH is 190.21).

---

---

**OUTPUT 5A**

---

---

Dependent Variable: LGT\_UH

Method: Least Squares

Sample: 1 965

Included observations: 965

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.027	0.035	-0.79	0.427
LY	-0.001	0.013	-0.07	0.940
LY_UH	0.437	0.028	15.83	0.000
TAM	-0.000	0.008	-0.03	0.975
TAM2	0.001	0.001	0.74	0.462
UH	8.728	0.080	108.92	0.000
UM	-0.022	0.020	-1.13	0.261
MT	-0.002	0.014	-0.18	0.856
EDAD	0.001	0.001	1.08	0.281

---

---

R-squared	0.998
Adjusted R-squared	0.998
S.E. of regression	0.160
Sum squared resid	24.50

---

---

(The sum of squares of the residuals in a regression like the one in OUTPUT 5A but that omits LY and LY\_UH is 31.88).

---

---

**OUTPUT 6**

---

---

Dependent Variable: LAL

Method: Two-Stage Least Squares

Sample: 1 965

Included observations: 965

Instrument list: LY LY\_UHT

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.031	0.528	5.74	0.000
LGT	0.508	0.058	8.76	0.000
LGT_UH				
TAM	0.145	0.018	8.06	0.000
TAM2	-0.010	0.003	-3.33	0.000
UH	3.063	1.362	2.25	0.025
UM	-0.100	0.044	-2.27	0.025
MT	-0.065	0.032	-2.03	0.040
EDAD	0.004	0.001	4.00	0.004

---

---

R-squared	0.394
Adjusted R-squared	0.389
S.E. of regression	0.358
Sum squared resid	122.72

---

---

**IMPORTANT:**

The confidence interval at the 95% level for the LGT\_UH coefficient is [-0.58852 ; -0.05148]

The estimated covariance between the coefficients of LGT and LGT\_UH is equal to -0.002.



---

---

**OUTPUT 7**

---

---

Dependent Variable: LAL

Method: Least Squares

Sample: 1 965

Included observations: 965

---

---

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.031	0.528	5.74	0.000
LGT	0.508	0.058	8.76	0.000
LGT_UH				
TAM	0.145	0.018	8.06	0.000
TAM2	-0.010	0.003	-3.33	0.000
UH	3.063	1.362	2.25	0.025
UM	-0.100	0.044	-2.27	0.025
MT	-0.065	0.032	-2.03	0.040
EDAD	0.004	0.001	4.00	0.004
RES5	0.015	0.066	0.23	0.890
RES5A	0.188	0.157	1.20	0.232
R-squared		0.399		
Adjusted R-squared		0.393		
S.E. of regression		0.357		
Sum squared resid		121.74		

---

---

(RES5 and RES5A are the residuals from  
OUTPUTS 5 and 5A, respectively)

### PROBLEM 3: DEMAND FUNCTIONS

The wholesalers of biological apples have implemented a survey among their retail clients, resulting in 660 observations of the amounts in kilograms (*ecocan*) and the corresponding prices in euros (*ecoprc*) at which the apples were sold. They also have available an additional variable, prices in euros of the normal (non-biological) apples (*regprc*). Using both price variables, they have generated a variable reflecting the price differential between biological and non-biological apples, *difprc* = (*ecoprc* – *regprc*). Using these data, the following linear demand functions have been estimated by OLS (standard errors between parentheses):

$$\widehat{ecocan} = 2.388 - 0.845ecoprc \quad (f1)$$

(0.372) (0.332)

$$n = 660; R^2 = 0.00098; \sigma^2 = 2.5153$$

$$\widehat{ecocan} = 1.965 - 2.927ecoprc + 3.029regprc \quad (f2)$$

(0.380) (0.588) (0.711)

$$n = 660; R^2 = 0.0364; \sigma^2 = 2.4831$$

$$\widehat{ecocan} = 2.056 - 2.926difprc \quad (f3)$$

(0.152) (0.588)

$$n = 660; R^2 = 0.0363; \sigma^2 = 2.4814$$

Universidad Carlos III de Madrid  
**ECONOMETRICS I**  
 Academic year 2005/06  
**FINAL EXAM(Convocatoria extraordinaria)**  
 September 1, 2006

Exam's type: 1

**TIEMPO: 2 HORAS 30 MINUTOS**

*TIME: 2 HOURS 30 MINUTES*

*Instrucciones:*

- ANTES DE EMPEZAR A RESPONDER EL EXAMEN:
- *BEFORE YOU START TO ANSWER THE EXAM:*
  - Rellene sus datos personales en el **impreso de lectura óptica**, que será el único documento válido de respuesta. Recuerde que tiene que completar sus datos identificativos (Nombre y apellidos y NIE) tanto en letra como en las casillas correspondientes de lectura óptica.  
**Muy importante:** El número de identificación que debe rellenar es su **NIU** (NO el DNI o el Pasaporte), que tiene 9 dígitos y empieza siempre por 1000.
  - *Fill in your personal data in the optical reading form, which will be the only valid answering document. Remember that you must complete all your identifying data (name and surname(s), and NIU) both in letters and in the corresponding optical reading boxes.*  
**Very important:** The identification number that you must fill is your **NIU** (NOT the DNI or the Passport), which has 9 digit and always begins by 1000.
  - Rellene, tanto en letra como en las correspondientes casillas de lectura óptica, el código de la asignatura y su grupo, de acuerdo con la siguiente tabla:
  - *Fill in, both in letters and in the corresponding optical reading boxes, the course code and your group, according with the following table:*

TITULACION	GRUPOS					CODIGO DE ASIGNATURA
Economía	61	62	63	64	65*	10188
ADE	71	72	73	74	75*	10188
– ADE (Colmenarejo)	71					10188
Sim. Eco-Dcho.	69					42020
Sim. ADE-Dcho.	77	78				43020
Sim. ADE-Dcho (Colmenarejo)	17					43020

\*Grupos bilingües

- Compruebe que este cuadernillo tiene 5 problemas y que el cuestionario de preguntas tiene 40 preguntas numeradas correlativamente.
- *Check that this document contains 40 questions sequentially numbered.*
- **Compruebe que el número de tipo de examen que aparece en el cuestionario de preguntas coincide con el señalado en el impreso de lectura óptica.**

Exam's type: 1

- *Check that the number of exam type that appears in the questionnaire matches the number indicated in the optical reading form.*
- Lea las preguntas detenidamente.  
**Cuando una pregunta se refiera a algún problema de los enunciados, el encabezado de la pregunta incluirá entre paréntesis el número de problema a que corresponde.** Se recomienda leer atentamente dicho enunciado **antes** de contestar las preguntas relacionadas.
- *Read the questions carefully.*  
*Whenever a question is referred to the Problem included in the enclosed document, the question will include within parentheses at the beginning of the question the letter P.*  
*It is advised to read carefully the text of the problem before answering its corresponding questions.*
- Para la fila correspondiente al número de cada una de las preguntas, rellene la casilla correspondiente a la respuesta escogida en el impreso de lectura óptica (A, B, C ó D).
- *For each row regarding the number of each question, fill the box which corresponds with your chosen option in the optical reading form (A, B, C or D).*
- **Cada pregunta tiene una única respuesta correcta.**  
Cualquier pregunta en la que se seleccione más de una opción será considerada nula y su puntuación será cero.
- *Each question only has one correct answer.*  
*Any question in which more than one answer is selected will be considered incorrect and its score will be zero.*
- Todas las preguntas respondidas correctamente tienen idéntica puntuación. Las respuestas incorrectas tendrán una puntuación de cero. Para superar la asignatura hay que responder correctamente un mínimo de 24 preguntas. (Tenga en cuenta que la calificación complementaria de clase no se tiene en cuenta en la convocatoria extraordinaria).
- *All the questions correctly answered has the same score. Any incorrect answer will score as zero. To pass the exam, you must correctly answer a minimum of 24 questions. (The complementary class grade will not consider in this exam)*
- Si lo desea, puede utilizar la plantilla de respuestas que aparece a continuación como borrador, si bien dicha plantilla carece por completo de validez oficial.
- *If you want, you may use the answer table as a draft, although such table does not have any official validity.*
- Puede utilizar el reverso de las hojas como borrador (no se facilitará más papel).
- *You can use the back side of the sheets as a draft (no additional sheets will be handed out).*
- Al final de este documento, se adjuntan tablas estadísticas.
- *Statistical tables are enclosed at the end of the document.*
- **Cualquier alumno que sea sorprendido hablando o intercambiando cualquier tipo de material en el examen será expulsado en el acto y su calificación será de cero, sin perjuicio de otras medidas que se puedan adoptar.**

- *Any student who were found talking or sharing any sort of material during the exam will be expelled out immediately and his/her overall score will be zero, independently of any other measure that could be undertaken.*
  
- **Fechas de publicación de calificaciones:** Martes 5 de Septiembre.
- *Date of grades publication: Tuesday, September 5th.*
- **Fecha de revisión:**
- *Date of exam revision:*
  - Grupos del Campus de Getafe: Jueves 7 de Septiembre a las 15 h (las aulas se comunicarán en Aula Global)
  - *Groups at Getafe: Thursday, September, 7th, at 15 h (The place will be published in Aula Global).*
  - Grupos del Campus de Colmenarejo: el profesor responsable lo comunicará en Aula Global.
  - *Groups at Colmenarejo: The instructor in charge will publish it in Aula Global.*
  
- **Normas para la revisión:**
- *Rules for exam revision:*
  - La revisión sólo tendrá por objeto comprobar el número de respuestas correctas del examen.
  - *Its only purpose will be to check that the number of correct answers is right.*
  - Para tener derecho a revisión, el alumno deberá:
  - *To be entitled for revision, the student should:*
    - \* *Solicitarlo por escrito, apuntándose en la lista situada en el Tablón de Información del departamento de Economía (junto al despacho 15.2.22), indicando titulación y grupo. Los alumnos de los grupos del Campus de Colmenarejo deberán apuntarse en la lista situada en la puerta del despacho 1.2.B11.*
    - \* *Apply in writing, enrolling in a list located in the Tablón de Información of the Department of Economics (close to room 15.2.22), establishing your titulation and group. Students at Colmenarejo should enroll in a list located at the door of room 1.2.B11.*
    - \* *Acudir a la revisión con una copia impresa de las soluciones del examen, que estarán disponibles en Aula Global desde el día de publicación de las calificaciones.*
    - \* *Bring a printed copy of the exam solutions, which will be available in Aula Global from the date of grades publication.*

Draft of  
ANSWERS

QUESTION	(a)	(b)	(c)	(d)	QUESTION	(a)	(b)	(c)	(d)
1.					21.				
2.					22.				
3.					23.				
4.					24.				
5.					25.				
6.					26.				
7.					27.				
8.					28.				
9.					29.				
10.					30.				
11.					31.				
12.					32.				
13.					33.				
14.					34.				
15.					35.				
16.					36.				
17.					37.				
18.					38.				
19.					39.				
20.					40.				

1. (**Problem 2**) Suppose that in model (\*), all assumptions about the classical regression model hold. Consider the following statements:
- (i) An appropriate estimation of  $V(\text{LAL})$  would be  $(121.95/(965 - 9)) \simeq 0.128$ .
  - (ii) An appropriate estimation of  $V(\text{LAL} \mid \text{LGT, TAM, UH, UM, MT, EDAD})$  would be  $(121.95/(965 - 9)) \simeq 0.128$ .
  - (iii) An appropriate estimation of  $V(\text{LAL} \mid \text{LGT, LGT\_UH, TAM, TAM2, UH, UM, MT, EDAD})$  would be  $(121.95/(965 - 9)) \simeq 0.128$ .
- (a) Only (ii) and (iii) are true.
  - (b) The three statements are false.
  - (c) Only (iii) is true.
  - (d) The three statements are true.
2. (**Problem 2**) Suppose that in model (\*), all assumptions about the classical regression model hold. Consider the following statements:
- (i) An appropriate estimation of  $V(u)$  would be  $(121.95/(965 - 9)) \simeq 0.128$ .
  - (ii) An appropriate estimation of  $V(u \mid \text{LGT, TAM, UH, UM, MT, EDAD})$  would be  $(121.95/(965 - 9)) \simeq 0.128$ .
  - (iii) An appropriate estimation of  $V(u \mid \text{LGT, LGT\_UH, TAM, TAM2, UH, UM, MT, EDAD})$  would be  $(121.95/(965 - 9)) \simeq 0.128$ .
- (a) Only (ii) and (iii) are true.
  - (b) The three statements are false.
  - (c) Only (iii) is true.
  - (d) The three statements are true.
3. (**Problem 2**) Assume that given model (\*) we want to test the hypothesis that the average household annual food expenditure is independent of the fact that both spouses have university degrees.
- (a) The null hypothesis would be  $H_0 : \beta_5 = \beta_6 = 0$ .
  - (b) The null hypothesis would be  $H_0 : \beta_2 = \beta_5 = \beta_6$ .
  - (c) The null hypothesis would be  $H_0 : \beta_2 = \beta_5 = \beta_6 = 0$ .
  - (d) The null hypothesis would be  $H_0 : \beta_5 = \beta_6$ .
4. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model and we want to test the null hypothesis that the average household food expenditure is independent of the fact that both spouses have university degrees.
- (a) From OUTPUT 1 and OUTPUT 4, the statistic test is  $W = \frac{(0.359^2 \times (965 - 7) - 121.95)}{121.95} \times (965 - 9) \simeq 11.9$ , which is approximately distributed as an  $\chi_2^2$ .
  - (b) We do not have enough information to perform this test.
  - (c) From OUTPUT 1 and OUTPUT 4, the statistic test is  $W = \frac{(0.359^2 \times (965 - 7) - 121.95)}{121.95} \times (965 - 9) \simeq 11.9$ , which is approximately distributed as an  $\chi_3^2$ .
  - (d) From OUTPUT 1 and OUTPUT 4A, the statistic test is  $W = \frac{((1 - 0.374) \times 202.643 - 121.95)}{121.95} \times (965 - 9) \simeq 38.45$ , which is approximately distributed as an  $\chi_3^2$ .

5. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model and we want to test the null hypothesis that the food expenditure elasticity with respect to total expenditure is independent on the fact that the husband has an university degree.
- The null hypothesis would be  $H_0 : \beta_2 = \beta_5 = 0$ .
  - The null hypothesis would be  $H_0 : \beta_1 = \beta_2$ .
  - The null hypothesis would be  $H_0 : \beta_2 = 0$ .
  - The null hypothesis would be  $H_0 : \beta_2 = \beta_5$ .
6. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model and we want to test the null hypothesis that the food expenditure elasticity with respect to total expenditure is independent on the fact that the husband has an university degree.
- We do not have enough information to test such hypothesis.
  - The statistic test is  $t \simeq \frac{0.518 - (-0.177)}{\sqrt{0.028^2 + 0.066^2 - 2 \times (-0.0007)}} \simeq 8.42$ , which is asymptotically distributed as  $N(0, 1)$ .
  - The statistic test is  $t \simeq -2.70$ , that is asymptotically distributed as  $N(0, 1)$ .
  - The statistic test is  $t \simeq \frac{-0.177 - 1.628}{\sqrt{0.066^2 + 0.678^2 - 2 \times (-0.046)}} \simeq -2.42$ , which is asymptotically distributed as  $N(0, 1)$ .
7. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model and we want to test the null hypothesis that the food expenditure elasticity with respect to total expenditure is independent on the fact that the husband has an university degree. With the available information:
- We do not have enough information to test such hypothesis.
  - We would reject the null hypothesis at the 0.1% level.
  - We would reject the null hypothesis at the 1% but not at the 0.1% level.
  - We would reject the null hypothesis at the 5% but not at the 1% level.
8. (**Problema 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. Comparing two households with a total expenditure of 20000 euros, same size, same husband's age and same wife's labor status BUT whereas in the first household both spouses have a university degree, both spouses do not have it in the second household, implies an estimated average difference in the food expenditure equals to:
- $(1.628 - 0.105) \% \simeq 1.52\%$  more.
  - $[(-0.177 \times \ln(20000) + (1.628 - 0.105))] \times 1000 \simeq 229.9$  more euros.
  - $[(-0.177 \times \ln(20000) + (1.628 - 0.105))] \times 100\% \simeq 22.99\%$  less.
  - We do not have enough information.



9. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. Using the estimated regressions, consider the following statements about the average food expenditure:
- (i) Considering households with a total expenditure of 8000 euros, those households that have a husband with university degree will have on average higher food expenditure than those households with a husband without university degree (*ceteris paribus*).
  - (ii) Considering households with a total expenditure of 8000 euros, those households that have a husband with university degree will have always higher food expenditure than those households with a husband without university degree (*ceteris paribus*).
  - (iii) It is possible that there are households with husbands with a university degree whose current food expenditure is higher than other households with similar characteristics but husbands without a university degree and viceversa.
- (a) Only (i) and (ii) are true.
  - (b) The three statements are true.
  - (c) Only (ii) and (iii) are true.
  - (d) Only (i) and (iii) are true.
10. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. We want to test the null hypothesis that the average elasticity of food expenditure with respect to total expenditure is 57%. Then:
- (a) We do not reject the null hypothesis at the 5% level if the husband has a university degree.
  - (b) We do not reject the null hypothesis at the 10% level if the husband has not a university degree.
  - (c) We do not have information enough to evaluate this hypothesis.
  - (d) We do not reject the null hypothesis at the 5% level if the husband has not a university degree.
11. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. We want to test the hypothesis that the number of members in the household does not affect food expenditure.
- (a) The available information does not provide sufficient evidence to evaluate such hypothesis.
  - (b) From OUTPUT 1 and OUTPUT 3, the appropriate test statistic is  $W \simeq \frac{(0.392^2 \times (965 - 7) - 121.95)}{121.95} \times (965 - 9) \simeq 198.02$ , which is asymptotically distributed as a  $\chi_2^2$ , so that we reject the null hypothesis at any reasonable significance level.
  - (c) Since TAM and TAM2 are individually significant, we reject the null hypothesis.
  - (d) The statistic test is  $t = 7.94$ , which is asymptotically distributed as a standard normal, so that we reject the null hypothesis at any reasonable significance level.

12. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. The *ceteris paribus* effect of an additional household member in a household with 5 members (including spouses) entails an estimated average increase in food expenditure of:
- (a)  $(0.143 - 2 \times 0.01 \times 3) \times 100 \%$ .  $\simeq 8.3\%$
  - (b)  $(0.143 - 2 \times 0.01 \times 5) \times 10$   $\simeq 43$  euros.
  - (c)  $(0.143 - 2 \times 0.01 \times 3) \times 10$   $\simeq 83$  euros.
  - (d)  $(0.143 - 2 \times 0.01 \times 5) \times 100 \%$ .  $\simeq 4.3\%$ .
13. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. Given the coefficients and standard errors of TAM and TAM2, indicate what of the following statements about the estimated *ceteris paribus* effect of size on food expenditure is FALSE:
- (a) Such effect is positive in a household with less than 8 members (including the spouses).
  - (b) Such effect is negative for households with 10 members (including spouses).
  - (c) Such effect is positive but marginally decreasing.
  - (d) Such effect is positive for households of small size, but marginally decreasing with size, and could be negative for households of large size.
14. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. The estimate of the average difference in food expenditure between two households with the same total expenditure, same size, same educational level of spouses, with the wife working in both households, but in which the husband is 10 years older in the first household than in the second one is equal to:
- (a) 40 euros.
  - (b) 4 euros.
  - (c) 4%.
  - (d) 0.04%.
15. (**Problem 2**) Assume that model (\*) fulfills all the assumptions of the classical regression model. The estimate of the average difference, *ceteris paribus*, in food expenditure between a household with a total expenditure of 29700 euros in which both spouses have university degrees and the wife does not work, in regard to another household with the same total expenditure in which both spouses have not university degrees and the wife works is approximately equal to:
- (a)  $(-0.177 \times 10.3 + 1.628 - 0.105) \times 100 = -30.01\%$  less.
  - (b)  $(-0.177 \times 10.3 + 1.628 - 0.105 - (-0.071)) \times 10 \simeq 229.1$  euros less.
  - (c)  $-(-0.177 \times 10.3 + 1.628 - 0.105 - (-0.071)) \times 100 \simeq 22.91\%$  more.
  - (d)  $(-0.177 \times 10.3 + 1.628 - 0.105 - (-0.071)) \times 100 \simeq 22.91\%$  less.

16. (**Problem 2**) When estimating model (\*), there exists a potential problem regarding LGT, due to the fact that food expenditure is a component of total expenditure, and therefore they are simultaneously determined. Given this, consider the following statements:
- (i) LGT is an endogenous variable.
  - (ii)  $E(u | \text{LGT, TAM, UH, UM, MT, EDAD}) \neq 0$ .
  - (iii)  $E(u | \text{LGT, TAM, UH, UM, MT, EDAD}) = 0$ .
- (a) Only (i) is true.
  - (b) Only (iii) is true.
  - (c) Only (i) and (iii) are true.
  - (d) Only (i) and (ii) are true.
17. (**Problem 2**) In the estimation of model (\*), if LGT is endogenous, consider the following statements:
- (i) The OLS estimation of equation (\*) does not yield a consistent estimator of  $\beta_1$ .
  - (ii) Asymptotically, the OLS estimator of equation (\*) underestimates  $\beta_2$ .
  - (iii) Asymptotically, the OLS estimator of equation (\*) overestimates  $\beta_1$ .
- (a) The three statements are true.
  - (b) Only (i) and (iii) are true.
  - (c) The three statements are false.
  - (d) Only (i) is true.
18. (**Problem 2**) Suppose that we want to obtain consistent estimates of the coefficients of model (\*), and  $E(\text{LGT} \times u) \neq 0$ ,  $E(\text{LGT\_UH} \times u) \neq 0$ , but the remaining explanatory variables included in model (\*) are uncorrelated with the error term  $u$ . Moreover, we know that the variable LY is not correlated with  $u$  either.
- (a) Under such conditions, the estimators in OUTPUT 1 are consistent.
  - (b) The estimators in OUTPUT 6 are not consistent, since even though LY is exogenous, LY\_UH is not.
  - (c) The estimators in OUTPUT 6 are consistent, since the instruments LY, LY\_UH fulfill the two conditions that are needed to be valid instruments: being uncorrelated with  $u$  but correlated with the endogenous explanatory variables LGT, LGT\_UH.
  - (d) The reduced form for LGT in OUTPUT 5 is inappropriate, since it should include LY but not LY\_UH among the explanatory variables.
19. (**Problem 2**) Since the variable LY is uncorrelated with  $u$ , if we want to verify empirically whether there is an endogeneity problem associated to the variable LGT:
- (a) We will use the  $t$  test of the coefficient of such variable in OUTPUT 1.
  - (b) We will test the joint significance of the regressors in OUTPUT 5.
  - (c) We will test the joint significance of the regressors both in OUTPUT 5 and OUTPUT 5A.
  - (d) We will implement a Hausman test.

20. (**Problem 2**) Since the variable LY is uncorrelated with  $u$ , and given the results reported in OUTPUT 7, consider the following statements:
- Since RES5 is not statistically significant, we would NOT reject that LGT is exogenous.
  - Since RES5A is not statistically significant, we would NOT reject that LGT\_UH is exogenous.
  - The Hausman test is implemented through the statistic  $W = \frac{121.95 - 121.74}{121.74} \times (965 - 11) \simeq 1.65$ , so that we reject exogeneity of both LGT and LGT\_UH at any reasonable significance level.
- Only (i) and (ii) are true.
  - The three statements are false.
  - Only (iii) is true.
  - Only (i) is true.
21. (**Problem 2**) Given the reported results in OUTPUT 7, taking the most appropriate estimates of model (\*), at the 95% confidence level, the elasticity of food expenditure with respect to total expenditure in those households in which the husband has a university degree:
- Would be approximately between  $-7.6\%$  and  $45.2\%$ .
  - Cannot be calculated with the available information.
  - Would be approximately between  $27.5\%$  and  $40.7\%$ .
  - Would be approximately between  $-58.9\%$  and  $-5.1\%$ .
22. (**Problem 2**) Given the reported results in OUTPUT 7, taking the most appropriate estimates of model (\*), consider the following statements about the difference in the elasticity of food expenditure with respect to total expenditure between households in which the husband has a university degree and those households in which the husband does not have a university degree:
- It is statistically different from zero at the 1% significance level.
  - It is statistically different from zero at the 5% significance level.
  - It is statistically different from zero at the 10% significance level.
- The three statements are true.
  - Only (iii) is true.
  - Only (ii) and (iii) are true.
  - The three statements are false.
23. (**Problem 2**) Given the model of OUTPUT 1, the estimate of the coefficient of LGT is:
- Statistically significant and positive with a confidence level of 90%.
  - Equal to 0.518.
  - We cannot calculate the exact value with the information at hand.
- Considering the three statements:
- Only (i) is true.
  - Only (iii) is true.
  - Only (i) and (ii) are true.
  - Only (i) and (iii) are true.

24. Let the model

$$Y = \alpha + \beta X + u,$$

where  $u = X\varepsilon$ . We know that  $E(\varepsilon|X) = 0$  and  $V(\varepsilon|X) = V(\varepsilon) = \sigma^2$ . Consider the following statements:

- (i) The OLS estimator of  $\beta$  is inconsistent since  $u$  depends on  $X$ .
  - (ii) The assumption that  $E(\varepsilon|X) = 0$  suffices to ensure that  $X$  is mean independent of  $u$ .
  - (iii) Since  $V(\varepsilon|X) = V(\varepsilon) = \sigma^2$ , we can ensure that there is NOT a heteroskedasticity problem.
- (a) Only (i) is true.
  - (b) Only (ii) and (iii) are true.
  - (c) Only (iii) is true.
  - (d) Only (ii) is true.

25. Let the model

$$Y = \alpha + \beta X + u,$$

where  $u = X\varepsilon$ . We know that  $E(\varepsilon|X) = 0$ . Consider the following statements:

- (i) We cannot know whether there is a heteroskedasticity problem, because we have not been told that  $V(\varepsilon|X) = V(\varepsilon) = \sigma^2$ .
  - (ii) In this case, the usual approach to compute the variance of the OLS estimator is incorrect, because with the information at hand we know that the errors will be in general heteroskedastic.
  - (iii) Independently on whether there is heteroskedasticity or not, the OLS estimator of  $\beta$  in this model is inconsistent anyway.
- (a) Only (i) is true.
  - (b) Only (ii) and (iii) are true.
  - (c) Only (i) and (iii) are true.
  - (d) Only (ii) is true.

26. **(Problem 1)** Using OUTPUT 1 and Figure 1:

- (a) Clearly there is NOT a residual autocorrelation problem.
- (b) The first order autocorrelation coefficient is approximately 0.32.
- (c) The first order autocorrelation coefficient is approximately 0.84.
- (d) Clearly there is NOT a residual positive autocorrelation problem.

27. **(Problem 1)** Using OUTPUT 1 and Figure 1:

- (a) The variable VM(-1) play none role explaining investment in the estimated model from OUTPUT 1.
- (b) A Durbin-Watson close to 0 reveals that there is not a residual autocorrelation problem.
- (c) Assuming normality, the test of joint significance (F-statistic) from OUTPUT 1 is still valid.
- (d) We cannot be sure in the estimated model from OUTPUT 1 if the variable VM(-1) is or is not statistically significant.

28. **(Problem 1)** Using OUTPUT 2 and Figure 2:
- The residual autocorrelation is clearly negative.
  - The first order autocorrelation coefficient is approximately 0.41.
  - In this case, the residuals are independent since they are centered on 0.
  - The residual autocorrelation problem in OUTPUT 2 is higher than the one found in OUTPUT 1.
29. **(Problem 1)** Using OUTPUT 2 and Figure 2:
- There is not residual autocorrelation in OUTPUT 2. In OUTPUT 1, omitting the variable  $\tau$  produces the WHOLE residual autocorrelation.
  - The coefficients estimated in OUTPUT 2 are not reliable because there is a problem of autocorrelation.
  - The standard errors in OUTPUT 2 are correctly calculated.
  - The coefficients estimated in OUTPUT 2, are reliable if residual autocorrelation is the only problem we are facing.
30. **(Problem 1)** Using OUTPUT 3:
- The first and second lags are not statistically significant.
  - There is no residual autocorrelation of first and second order.
  - Using the specification from OUTPUT 3, we can only know if the residual lags (first and second) are individually significant but we do not have enough information to know whether they are jointly significant or not.
  - The test of joint significance for the first and second residual lags shows that they are statistically significant. Therefore, there is residual autocorrelation of first and second order in the specification used in OUTPUT 2.
31. **(Problem 1)** Using OUTPUT 4:
- The lagged firm's market value,  $VM(-1)$ , is statistically significant at the 5% level.
  - OUTPUT 4 is not valid because the estimated coefficients are identical to the ones in OUTPUT 2.
  - The  $R^2$  shows that the model explain less than 1% of the firm's investment variance.
  - The trend variable,  $\tau$ , is NOT statistically significant.
32. **(Problem 1)** Using OUTPUT 4:
- It is NOT usual that the standard deviation of the dependent variable in the regression (48.58) to be higher than the standard error of the regression (24.98).
  - Given the residual evolution in Figure 1, it is surprising that the trend variable,  $\tau$ , has a positive sign.
  - The Durbin-Watson in the regression in which the variance and covariance matrix has been estimated by Newey-West should be the same one than the one found in the regression using Ordinary Least Squares (OLS).
  - The adjusted- $R^2$  is usually higher than the  $R^2$  without adjustment. We do not observe this in OUTPUT 4 because of residual autocorrelation.

33. **(Problem 1)** If the specification used in OUTPUT 4 were estimated using quarterly data, instead of annual data, the maximum lag used to estimate the standard errors by Newey-West should:
- Decrease.
  - The maximum lag used to estimate the standard errors by Newey-West is independent on the data frequency because the same lag is always used.
  - Increase.
  - Stay the Same.
34. **(Problem 1)** Using the correct residuals, facing an increase of one million dollars in the previous year firm's market value, with a confidence level of 10%, the investment should vary on average:
- Approximately between 4900 and 51100 dollars.
  - Approximately between 6550 and 49450 dollars.
  - Approximately between 9850 and 46150 dollars.
  - Approximately between  $-5900$  and 79900 dollars.
35. **(Problem 1)** Using the appropriate results, consider the following statements about the impact of previous year firm's market value on firm's investment:
- It is statistically significant at the 1% level.
  - It is statistically significant at the 5% level.
  - It is statistically significant at the 10% level.
- Only (i) is true.
  - The three statements are false.
  - Only (ii) and (iii) are true.
  - Only (i) and (ii) are true.
36. **(Problem 3)** Using the correlation, the sample prices of non-organic and organic apples:
- Have a positive relation.
  - We cannot conclude anything about the correlation between these two prices with the information available in the text and in the estimations (f1), (f2) and (f3).
  - Are uncorrelated.
  - Have a negative relation.
37. **(Problem 3)** If in addition to the information in the text and in the estimations (f1), (f2) and (f3), we would also know that  $\widehat{V}(ecoprc) = 2$  (using the symbol  $\widehat{\cdot}$  to denote estimated moments), then:
- $\widehat{C}(ecoprc, regprc) = 1.3747$ .
  - None of the other answers is true.
  - The additional information provided here is not necessary to calculate  $\widehat{C}(ecoprc, regprc) = 0.2376$ .
  - We need the variance-covariance matrix of the coefficient estimates of (f2) in order to calculate  $\widehat{C}(ecoprc, regprc)$ .

38. **(Problem 3)** Although the variables *ecoprc* and *regprc* are significant in equation (f2), they have not been included in equation (f3). This can be due to the fact that:
- difprc* would be endogenous with respect to these two variables, so that if they were included we would need to apply instrumental variables estimation.
  - None of the other answers is true.
  - Since, by definition, the correlation between *difprc* and *ecoprc* is exactly equal to 1 and the correlation between *difprc* and *regprc* is exactly equal to  $-1$ , we do not have to include them in the regression, since their estimated coefficients will necessarily be equal to  $-2.926$  and  $2.926$ .
  - A demand function including the three variables would be incorrectly specified.
39. **(Problem 3)** Given the reported estimates, we want to test the following null hypothesis: if the wholesalers of non-organic apples reduce the price by one euro, the demand of organic apples will keep constant if the wholesalers of organic apples also reduce the price by one euro. Consider the following statements:
- The null and the alternative hypotheses will be defined on the basis of the marginal effects of *ecoprc* and *regprc* which arise by interpreting the regression (f2), so that in order to implement the test, we will need the variance-covariance matrix of the estimated coefficients of (f2).
  - The alternative hypothesis can be defined on the basis of the demand function that was estimated in (f2) ( $ecocan = \beta_0 + \beta_1 ecoprc + \beta_2 regprc + \varepsilon$ ) as  $H_1 : \beta_1 + \beta_2 > 0$ .
  - The null hypothesis can be defined on the basis of the demand function that was estimated in (f2) ( $ecocan = \beta_0 + \beta_1 ecoprc + \beta_2 regprc + \varepsilon$ ) as  $H_1 : \beta_1 + \beta_2 = 0$ .
- Only (i) is true.
  - Only (iii) is true.
  - Only (ii) and (iii) are true.
  - Only (ii) is true.
40. **(Problem 3)** Given the reported estimates, we want to test the following null hypothesis: if the wholesalers of non-organic apples reduce the price by one euro, the demand of organic apples will keep constant if the wholesalers of organic apples also reduce the price by one euro. Consider the following statements:
- We can implement the test, whose conclusion is that we cannot reject the null hypothesis at the 5%.
  - We can implement the test by means of the F statistic, whose approximate value is 0.058.
  - We can implement the test, whose conclusion is that we reject the null hypothesis at the 5%.
- Only (i) is true.
  - None of the statements is true, since we need more information in order to implement the test.
  - Only (iii) is true.
  - Only (i) and (ii) are true.