UNIVERSIDAD CARLOS III DE MADRID
ECONOMETRICS I
Academic year 2004/05
FINAL EXAM

September, 1, 2005

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.
    (Siga las instrucciones de la hoja adjunta).
    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 identifiying data (name and surname(s), and DNI or Passport) both in letters and in the corresponding optical reading boxes.
  - 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 tiene 40 preguntas numeradas correlativamente.
  Chcek 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.**
  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 al problema del enunciado el encabezado de la pregunta incluirá entre paréntesis la letra P.
  Se recomienda leer atentamente el enunciado del problema **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 aprobar el examen hay que responder correctamente un mínimo de 24 preguntas.
  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.

- 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 inmediately and his/her overall score will be zero, independently of any other measure that could be undertaken.

- **Fechas de publicación de calificaciones:** Lunes 5 de Septiembre.
  Date of publication of scores: Monday, September, 5th.

- **Fecha de revisión:**
  Date of exam revision:

  - Jueves 8 de Septiembre a las 15 h en las AULAS 15.1.41 y 15.1.43
    Thursday, September, 8th, at 15 h in classrooms 15.1.41 and 15.1.43.

- **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.
      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 grade and group.
    * **Acudir a la revisión con una copia impresa de las soluciones** del examen, que estarán disponibles en Aula Global a partir del martes 6 de septiembre.
      Bring a printed copy of the exam solutions, which will be available in Aula Global from Tuesday, September, 6th.

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.**				

PROBLEM

The specification of Engel curves for food expenditure establishes a relationship between such expense and total expenditure. Let Y be food annual expenditure of a household (in euros) and X its total expenditure (in thousand euros). The most usual specifications are:

$$\ln Y = \beta_0 + \beta_1 \ln X + u \qquad \text{(I)}$$
$$Y = \delta_0 + \delta_1 \ln X + v \qquad \text{(II)}$$
$$\frac{Y}{X} = \gamma_0 + \gamma_1 \ln X + \varepsilon \qquad \text{(III)}$$

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

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

so that the variables determining food expenditure are the logarithm of total expenditure (LGT) and other variables which capture household characteristics.

Furthermore, it is important to remark that $C(\texttt{TAM}, \texttt{LGT}) > 0$.

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

SALIDA 1

Dependent Variable: LAL
Method: Least Squares
Sample: 1 965
Included observations: 965

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.6112	0.0794	83.24	0.000
LGT	0.4917	0.0260	18.93	0.000
TAM	0.1445	0.0178	8.11	0.000
TAM2	−0.0105	0.0025	−4.18	0.000
UH	−0.1286	0.0380	−3.38	0.001
UM	−0.1059	0.0439	−2.41	0.016
MT	−0.0700	0.0294	−2.38	0.017
EDAD	0.0034	0.0013	2.66	0.008

R-squared	0.3939
Adjusted R-squared	0.3895
S.E. of regression	0.3582
Sum squared resid	122.81

SALIDA 1A

Variance-covariance matrix of the estimated coefficients of SALIDA 1

	LGT	TAM	TAM2	UH	UM	MT	EDAD
LGT	0.0007						
TAM	−0.0001	0.0003					
TAM2	0.00003	−0.0004	0.00006				
UH	−0.0002	0.00001	0.00001	0.001443			
UM	−0.0008	0.00007	−0.00008	−0.0007	0.00193		
MT	−0.0002	0.00002	0	0.00004	−0.0003	0.000865	
EDAD	−0.00007	0	0	0	0.00003	0.00001	0.000017

SALIDA 2

Dependent Variable: LAL
Method: Least Squares
Sample: 1 965
Included observations: 965

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.9816	0.0706	98.85	0.000
LGT	0.4774	0.0254	18.78	0.000

R-squared	0.2680
Adjusted R-squared	0.2673
S.E. of regression	0.3925
Sum squared resid	148.33

SALIDA 3				
Dependent Variable: LAL				
Method: Least Squares				
Sample: 1 965				
Included observations: 965				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.7134	0.0816	82.29	0.000
LGT	0.5375	0.0268	20.04	0.000
UH	−0.1350	0.0399	−3.39	0.001
UM	−0.1438	0.0459	−3.13	0.002
MT	−0.0984	0.0307	−3.21	0.001
EDAD	0.0041	0.0014	3.01	0.003
R-squared		0.3308		
Adjusted R-squared		0.3273		
S.E. of regression		0.3761		
Sum squared resid		135.62		

SALIDA 4				
Dependent Variable: LAL				
Method: Least Squares				
Sample: 1 965				
Included observations: 965				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.6803	0.0793	84.24	0.000
LGT	0.4490	0.0251	17.87	0.000
TAM	0.1530	0.0180	8.51	0.000
TAM2	−0.0112	0.0025	−4.43	0.000
MT	−0.1064	0.0286	−3.72	0.000
EDAD	0.0037	0.0013	2.84	0.005
R-squared		0.3744		
Adjusted R-squared		0.3711		
S.E. of regression		0.3636		
Sum squared resid		126.77		

SALIDA 5

Dependent Variable: LGT
Method: Least Squares
Sample: 1 965
Included observations: 965

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.3080	0.0793	16.49	0.000
LY	0.4803	0.0300	15.98	0.000
TAM	0.0259	0.0199	1.30	0.193
TAM2	0.0005	0.0028	0.18	0.864
UH	0.1491	0.0426	3.50	0.000
UM	−0.0212	0.0492	−0.43	0.667
MT	0.0567	0.0337	1.68	0.093
EDAD	0.0037	0.0015	2.49	0.013
R-squared		0.3699		
Adjusted R-squared		0.3653		
S.E. of regression		0.3961		
Sum squared resid		150.14		

SALIDA 6

Dependent Variable: LAL
Method: Two-Stage Least Squares
Sample: 1 965
Included observations: 965
Instrument list: LY

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.6779	0.1283	52.05	0.000
LGT	0.4584	0.0566	8.09	0.000
TAM	0.1473	0.0183	8.03	0.000
TAM2	−0.0106	0.0025	−4.22	0.000
UH	−0.1172	0.0417	−2.81	0.005
UM	−0.1020	0.0444	−2.30	0.022
MT	−0.0622	0.0317	−1.97	0.050
EDAD	0.0038	0.0014	2.71	0.007
R-squared		0.3929		
Adjusted R-squared		0.3885		
S.E. of regression		0.3585		
Sum squared resid		123.02		

SALIDA 7

Dependent Variable: LAL
Method: Least Squares
Sample: 1 965
Included observations: 965

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.6779	0.1283	52.05	0.000
LGT	0.4584	0.0566	8.09	0.000
TAM	0.1473	0.0183	8.03	0.000
TAM2	−0.0106	0.0025	−4.22	0.000
UH	−0.1172	0.0416	−2.81	0.005
UM	−0.1020	0.0443	−2.30	0.022
MT	−0.0622	0.0317	−1.97	0.049
EDAD	0.0038	0.0014	2.71	0.007
RES5	0.0422	0.0637	0.66	0.508

R-squared	0.3942
Adjusted R-squared	0.3891
S.E. of regression	0.3583
Sum squared resid	122.76

(RES5 are the residuals from SALIDA 5)

1. (P) Considering the original variables Y, X:

   (a) Model (I) is linear in X and Y.

   (b) Model (II) is linear in X and Y.

   (c) Modelo (III) is linear in X and Y.

   (d) None of the three specifications is linear in X and Y.

2. (P) Consider model (I) and assume that it satisfies all the assumptions of the classical linear regression model for the corresponding transformations of the original variables. Given a household with an annual food expenditure of 4000 euros and a total annual expenditure of 20000 euros, a 100 euros increment in total expenditure lead to an average increase in food expenditure of:

   (a) $\left(0.1 \times \beta_1 \times \dfrac{4000}{20}\right)$ euros.

   (b) $\left(\beta_1 \times \dfrac{4000}{20}\right)$ euros.

   (c) β_1 euros.

   (d) β_1 %.

3. (P) Consider model (II) and assume that it satisfies all the assumptions of the classical linear regression model for the corresponding transformations of the original variables. Given a household with an annual food expenditure of 4000 euros and a total annual expenditure of 20000 euros, a 100 euros increment in total expenditure lead to an average increase in food expenditure of:

   (a) $\left(0.1 \times \delta_1 \times \dfrac{1}{20}\right)$ euros.

   (b) $\left(\delta_1 \times \dfrac{4000}{20}\right)$ euros.

   (c) $\dfrac{\delta_1}{20000}$ euros.

   (d) $100 \times \delta_1$ %.

4. (P) Consider model (I) and assume that it satisfies all the assumptions of the classical linear regression model for the corresponding transformations of the original variables. Given a household with an annual food expenditure of 4000 euros and a total annual expenditure of 20000 euros, the elasticity of food expenditure with respect to total expenditure is:

   (a) β_1 %.

   (b) $\dfrac{\beta_1}{20}$ %.

(c) $\beta_1 \times \dfrac{4000}{20}$ %.

(d) $100 \times \beta_1$ %.

5. (P) Consider model (II) and assume that it satisfies all the assumptions of the classical linear regression model for the corresponding transformations of the original variables. Given a household with an annual food expenditure of 4000 euros and a total annual expenditure of 20000 euros, the elasticity of food expenditure with respect to total expenditure is:

   (a) δ_1 %.

   (b) $\dfrac{\delta_1}{20}$ %.

   (c) $\dfrac{-\delta_1}{100 \times 20}$ %.

   (d) $\dfrac{\delta_1}{4000}$ %.

6. (P) Suppose that the model of interest is

$$\texttt{LAL} = \beta_0 + \beta_1\texttt{LGT} + \beta_2\texttt{TAM} + \beta_3\texttt{UH} + \beta_4\texttt{UM} + \beta_5\texttt{MT} + \beta_6\texttt{EDAD} + u,$$

   where it satisfies all the assumptions of the classical linear regression model, and $\beta_2 > 0$. In addition, and for the sake of simplicity, assume that both `LGT` and `TAM` are uncorrelated with the remaining explanatory variables. If we consider the OLS estimates of a model that omits size (`TAM`):

   (a) We will obtain a consistent estimate of β_1.

   (b) We will obtain a inconsistent estimate of β_1, which will tend to overestimate the effect of total expenditure on food expenditure.

   (c) We will obtain a inconsistent estimate of β_1, which will tend to underestimate the effect of total expenditure on food expenditure.

   (d) We will obtain a inconsistent estimate of β_6.

7. (P) Suppose that the model of interest is

$$\texttt{LAL} = \beta_0 + \beta_1\texttt{LGT} + \beta_2\texttt{TAM} + \beta_3\texttt{UH} + \beta_4\texttt{UM} + \beta_5\texttt{MT} + \beta_6\texttt{EDAD} + u$$

   where it satisfies all the assumptions of the classical linear regression model, and $\beta_2 > 0$. In addition, and for the sake of simplicity, assume that both `LGT` and `TAM` are uncorrelated with the remaining explanatory variables. Let d_1 be the OLS estimator of β_1 in the model which omits `TAM`, and b_1 the OLS estimator of β_1 in the model of interest (which does not omit `TAM`).

   (a) The estimated variance of d_1 will be lower than the estimated variance of b_1, and d_1 is still consistent.

(b) The estimated variance of d_1 will be higher than the estimated variance of b_1, but d_1 is inconsistent.

(c) The estimated variance of d_1 will be higher than the estimated variance of b_1, although d_1 is still consistent.

(d) The estimated variance of d_1 will be lower than the estimated variance of b_1, but d_1 is inconsistent.

8. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. Consider the following assertions:
(i) The estimate of $V(\texttt{LAL})$ is equal to $(0.3582)^2$.
(ii) The estimate of $V(\texttt{LAL} \mid \texttt{LGT}, \texttt{TAM}, \texttt{UH}, \texttt{UM}, \texttt{MT}, \texttt{EDAD})$ is equal to $(0.3582)^2$.
(iii) The estimate of $V(\texttt{LAL} \mid \texttt{LGT}, \texttt{TAM}, \texttt{TAM2}, \texttt{UH}, \texttt{UM}, \texttt{MT}, \texttt{EDAD})$ is equal to $(0.3582)^2$.

    (a) Only (ii) and (iii) are true.
    (b) Only (i) and (iii) are true.
    (c) Only (iii) is true.
    (d) Only (i) is true.

9. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model except the assumption of conditional homoskedasticity. Consider the following assertions:
(i) The parameter estimates of SALIDA 1 are not consistent.
(ii) The standard errors of the parameters of SALIDA 1 are not consistent.
(iii) The R^2 of the model has no sense.

    (a) Only (ii) is true.
    (b) Only (i) and (ii) are true.
    (c) The three assertions are true.
    (d) Only (ii) and (iii) are true.

10. (P) Suppose that from model (*) we want to test the null hypothesis that food expenditure is independent of household size.

    (a) The null hypothesis would be $H_0 : \beta_2 = 0$.
    (b) The null hypothesis would be $H_0 : \beta_2 = \beta_3 = 0$.
    (c) The null hypothesis would be $H_0 : \beta_3 = 0$.
    (d) The null hypothesis would be $H_0 : \beta_2 - \beta_3 = 0$.

11. (P) Suppose that from model (*) we want to test the null hypothesis that food expenditure is independent of whether the wife does work or not.

    (a) The null hypothesis would be $H_0 : \beta_6 = 0$.

(b) The null hypothesis would be $H_0 : \beta_5 = \beta_6 = 0$.

(c) The null hypothesis would be $H_0 : \beta_6 = 1$.

(d) The null hypothesis would be $H_0 : \beta_6 - \beta_5 = 0$.

12. (P) Suppose that from model (*) we want to test the null hypothesis that the effect on food expenditure of the fact that any of the spouses had a university degree is independent of who of the two spouses had such degree.

    (a) The null hypothesis is $H_0 : \beta_4 + \beta_5 = 0$.

    (b) The null hypothesis is $H_0 : \beta_4 - \beta_5 = 0$.

    (c) The null hypothesis is $H_0 : \beta_4 = -\beta_5$.

    (d) The null hypothesis is $H_0 : \beta_4 = \beta_5 = 0$.

13. (P) We are interested in model (*), which satisfies all the assumptions of the classical linear regression model. We want to test that the variables that jointly capture the household characteristics do not affect food expenditure.

    (a) The test statistic, given SALIDA 1 and SALIDA 2, is $W = \frac{(148.33 - 122.81)}{122.81} \times (965 - 8) = 198.87$, which is approximately distributed as a χ^2_6.

    (b) The test statistic, given SALIDA 1 and SALIDA 2, is $W = \frac{(148.33 - 122.81)}{122.81} \times (965 - 8) = 198.87$, which is approximately distributed as a χ^2_8.

    (c) The test statistic, given SALIDA 1 and SALIDA 2, is $W = \frac{(148.33 - 122.81)}{148.33} \times (965 - 8) = 164.65$, which is approximately distributed as a χ^2_6.

    (d) It is not possible to test such a hypothesis with the available information.

14. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. We want to test that the education levels of the spouses do not affect food expenditure. Consider the following assertions:
    (i) The null hypothesis is $H_0 : \beta_4 = \beta_5 = 0$.
    (ii) The test statistic, given SALIDA 1 and SALIDA 4, is $W = \frac{(126.77 - 122.81)}{122.81} \times (965 - 8) = 30.858$, which is approximately distributed as a χ^2_2.
    (iii) At the 1% significance level, we reject the hypothesis that the education levels of the spouses do not affect food expenditure.

    (a) The three assertions are true.

    (b) Only (i) and (iii) are true.

    (c) Only (i) and (ii) are true.

    (d) Only (ii) and (iii) are true.

15. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. We want to test the hypothesis that household size does not affect food expenditure.

    (a) It is not possible to test such a hypothesis with the available information.
    (b) From SALIDA 1, the test statistic is $t = 8.11$, so that we reject such hypothesis at any reasonable significance level.
    (c) Given that `TAM` and `TAM2` are individually significant, we reject such hypothesis.
    (d) Given SALIDA 1 and SALIDA 3, the appropriate test statistic is $W = \frac{(135.62 - 122.81)}{122.81} \times (965 - 8) = 99.822$, which is approximately distributed as a χ^2_2, and therefore we reject such hypothesis at any reasonable significance level.

16. (P) We are interested in model (*), which satisfies all the assumptions of the classical linear regression model. We want to test the null hypothesis that the effect on food expenditure of the fact that any of the spouses had a university degree is independent of what of the two spouses had such degree.

    (a) The test statistic is $t = \frac{-0.1286 - (-0.1059)}{\sqrt{0.001443 + 0.00193 - 2 \times (-0.0007)}} = -0.32857$, so that we do not reject such hypothesis at the 10% significance level.
    (b) The test statistic is $t = \frac{-0.1286 + (-0.1059)}{\sqrt{0.001443 + 0.00193 + 2 \times (-0.0007)}} = -5.2793$, so that we reject such hypothesis at the 1% significance level.
    (c) The test statistic is , given SALIDA 1 and SALIDA 4, $W = \frac{(126.77 - 122.81)}{122.81} \times (965 - 8) = 30.858$, so that we reject such hypothesis at the 1% significance level.
    (d) The test statistic is $t = \frac{-0.1286 - (-0.1059)}{\sqrt{0.001443 + 0.00193 + 2 \times (-0.0007)}} = -0.51105$, so that we do not reject such hypothesis at the 10% significance level.

17. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. The effect, *ceteris paribus*, of a 1% increment in total expenditure implies an average estimated increase in food expenditure of:

    (a) A 49.17 %.
    (b) A 0.4917 %.
    (c) 49.17 euros.
    (d) A $\left(\frac{49.17}{20}\right) \times 100 \;\; \% \simeq 24.59\%$.

18. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. The effect, *ceteris paribus*, of an additional household member in a household initially composed by 5 members (including the spouses) implies an average estimated increase in food expenditure of:

   (a) $(0.1445 - 2 \times 0.0105 \times 3) \times 100$ %. = 8.15%.
   (b) $(0.1445 - 2 \times 0.0105 \times 5) \times 100$ %. = 3.95%.
   (c) $(0.1445 - 2 \times 0.0105 \times 3) \times 100 = 8.15$ euros.
   (d) $(0.1445 - 2 \times 0.0105 \times 5) \times 100 = 3.95$ euros.

19. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. Given the estimated coefficients and the standard errors of `TAM` and `TAM2`, indicate what among the following assertions, regarding the estimated *ceteris paribus* effect of household size on food expenditure, is FALSE:

   (a) Such effect is positive for a household composed by six members (including the spouses).
   (b) Such effect is positive but marginally decreasing, which may be negative for large-sized households.
   (c) Such effect is positive but marginally increasing.
   (d) Such effect is negative for households with more than 9 members (including the spouses).

20. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. The average estimated difference in food expenditure between two households with the same total expenditure, same size, same educational level of the spouses and in which the wife is currently working, but the husband in the first household is 10 years older than the husband in the second one is equal to:

   (a) 34 euros.
   (b) A 0.34%.
   (c) A 3.4%.
   (d) 3.4 euros.

21. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. Comparing two households with the same total expenditure, same size, same husband age and same labour market situation of the wife, the fact that the two spouses in the first household have university degree implies an average estimated difference in food expenditure with respect to the other household in which none of the spouses have a university degree equal to:

   (a) $(-0.1286 - 0.1059) \times 100$ % = 23.45% less.

(b) $-(-0.1286-0.1059)\times 100\ \% = 23.45\%$ more.

(c) $(-0.1286-0.1059)\times 100 = 23.45$ euros less.

(d) $-(-0.1286-0.1059)\times 100 = 23.45$ euros more.

22. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. The fact that the wife works implies, on average, *ceteris paribus*, an estimated difference in food expenditure equal to:

    (a) $-0.07\times 100\ \% = 7\%$ less.

    (b) 7 euros more.

    (c) $0.07\times 100\ \% = 7\%$ more.

    (d) 7 euros less.

23. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. The average estimated difference, *ceteris paribus*, in food expenditure between a household where the wife has a university degree but does not work vs. another household in which the wife does not have a university degree but works is:

    (a) $[-0.1059-(-0.07)]\times 100\ \% = 3.59\%$ less.

    (b) 3.59 euros more.

    (c) $[-0.1059-0.07]\times 100\ \% = 17.59\%$ more.

    (d) $-[-0.1059-(-0.07)]\times 100\ \% = 3.59\%$ more.

24. (P) Assume that model (*) satisfies all the assumptions of the classical linear regression model. The average estimated difference, *ceteris paribus*, in food expenditure between a household where the wife has a university degree but does not work vs. another household in which the wife also has a university degree but works is:

    (a) It cannot be known given the available information: we would need to include the interaction between UM and MT as additional variable.

    (b) $-(-0.07)\times 100\ \% = 7\%$ more.

    (c) $(-0.07)\times 100\ \% = 7\%$ less.

    (d) $[-(-0.07)-(-0.1059)]\times 100\ \% = 17.59\%$ more.

25. (P) In the estimation of model (*), there exists a potential problem regarding LGT, due to the fact that food expenditure is part of total expenditure, and therefore these two variables are determined simultaneously. Given such situation, consider the following assertions:
    (i) LGT is an endogenous variable.
    (ii) $E\left(u \mid \texttt{LGT}, \texttt{TAM}, \texttt{UH}, \texttt{UM}, \texttt{MT}, \texttt{EDAD}\right) \neq 0$.
    (iii) $E\left(u \mid \texttt{LGT}, \texttt{TAM}, \texttt{UH}, \texttt{UM}, \texttt{MT}, \texttt{EDAD}\right) = 0$.

(a) Only (i) is true.
(b) Only (i) and (ii) are true.
(c) Only (i) and (iii) are true.
(d) Only (iii) is true.

26. (P) In the estimation of model (*), there exists a potential problem regarding `LGT`, due to the fact that food expenditure is part of total expenditure, and therefore these two variables are determined simultaneously. Given such situation, and assuming (for the sake of simplicity) that `LGT` is uncorrelated with the remaining explanatory variables, consider the following assertions:
(i) We would expect that OLS estimation of equation (*) will not yield a consistent estimator of β_1.
(ii) The fact that `LGT` and `LAL` are simultaneously determined makes that the model does not satisfy all the assumptions of the classical linear regression model.
(iii) We would expect that the OLS estimation of equation (*) will tend to overestimate β_1.

(a) Only (i) and (ii) are true.
(b) Only (ii) and (iii) are true.
(c) Only (i) and (iii) are true.
(d) The three assertions are true.

27. (P) Suppose that we want to get consistent estimates of the parameters of model (*), and that $E(\texttt{LGT} \times u) \neq 0$, but the remaining explanatory variables that are included in model (*) are uncorrelated with the error term u. Moreover, we know that the variable `LY` is also uncorrelated with u.

(a) Under such conditions, the estimators from SALIDA 1 are consistent.
(b) The reduced form for `LGT` in SALIDA 5 is inappropriate, because it should not include the remaining explanatory variables included in model (*) (it should include only the external instruments, that is, `LY`, and the constant).
(c) The estimators of SALIDA 6 are consistent, since the instrument `LY` fulfills the two conditions needed to be valid as an instrument: being uncorrelated with u and being correlated with the endogenous variable `LGT` (this last condition can be seen because `LY` is a significant variable in the reduced form of SALIDA 5).
(d) The estimators of SALIDA 6 are not consistent, since given SALIDA 5, the condition that the instrument must be correlated with the endogenous variable `LGT` does not hold.

28. (P) Regarding the fact that the variable `LY` is uncorrelated with u, if we want to verifiy that there is an endogeneity problem with the variable `LGT`:

(a) We will use the t test associated with the coefficient of such variable in SALIDA 1.

(b) We will make a Hausman test.

(c) We cannot verify such hypothesis.

(d) We will test the joint significance of the regressors in SALIDA 5.

29. (P) Regarding the fact that the variable `LY` is uncorrelated with u, given the results shown in SALIDA 7:

    (a) Since `RES5` is not statistically significant, we would NOT reject that `LGT` is exogenous.

    (b) It is not correct to test for exogeneity from the significance of `RES5`, because the reduced form in which such residuals are based incorrectly includes the remaining exogenous explanatory variables included in model (*).

    (c) Since `RES5` is statistically significant, we reject that `LGT` is exogenous.

    (d) Since `RES5` is not statistically significant, we would reject that `LGT` is exogenous.

30. (P) Regarding the fact that the variable `LY` is uncorrelated with u, given the results shown in SALIDA 7:

    (a) Both the estimators in SALIDA 1 and SALIDA 6 are inconsistent.

    (b) Both the estimators in SALIDA 1 and SALIDA 6 are consistent, but we would choose those from SALIDA 6 because the instrumental variable estimator is more efficient than the OLS estimator.

    (c) Both the estimators in SALIDA 1 and SALIDA 6 are consistent, but we would choose those from SALIDA 1 because the OLS estimator is more efficient than the instrumental variable estimator.

    (d) The estimators in SALIDA 1 and SALIDA 6 cannot be both consistent, because their numerical values are different.

31. Consider three variables Y, X_1 and X_2, among which the following relations hold: $E(Y|X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$, with $\beta_2 = 0$ and $C(X_1, X_2) \neq 0$. We want to estimate the model $Y = \beta_0 + \beta_1 X_1 + \varepsilon$, and we consider the OLS and the instrumental variable estimators. Given the following assertions:
    (i) The instrumental variable estimator will have minimum variance among the linear and consistent estimators, if we choose the best instrument among the set of ALL the exogenous variables that we have available.
    (ii) The OLS estimator and the instrumental variable estimator that uses $Z = X_2$ as instrument are both consistent, but the OLS estimator has lower variance.
    (iii) The instrumental variable estimator is inconsistent because X_1 is an exogenous variable.

(a) Only (i) is true.
(b) Only (ii) is true.
(c) Only (iii) is true.
(d) Only (i) and (ii) are true.

32. An economist who wants to study the consumption behaviour about an isotonic drink ($Y =$ liters consumed per year) by people who practice and people who do not practice sports ($DEP =$ 1 if the individual practices sports and 0 otherwise), considering also individual income level ($RENTA =$ Income in euros), specifies and estimates the following model:

$$E(Y \mid RENTA, DEP) = \beta_0 + \beta_1 RENTA + \beta_2 DEP + \beta_3 (RENTA \times DEP).$$

Suppose that we consider the variable $NOD = 1 - DEP$ and want to estimate the following model:

$$\begin{aligned} E(Y \mid RENTA, DEP, NOD) &= \delta_2 DEP + \delta_3 (RENTA \times DEP) \\ &\quad + \delta_4 NOD + \delta_5 (RENTA \times NOD). \end{aligned}$$

Indicate what among the following options is correct:

(a) $\delta_2 = \beta_0$; $\delta_4 = \beta_0 + \beta_2$; $\delta_3 = \beta_3$; $\delta_5 = \beta_1 + \beta_3$.
(b) $\delta_2 = \beta_0 + \beta_2$; $\delta_4 = \beta_0$; $\delta_3 = \beta_1 + \beta_3$, $\delta_5 = \beta_1$.
(c) $\delta_2 = \beta_0$; $\delta_4 = \beta_0 + \beta_2$; $\delta_3 = \beta_1 + \beta_3$; $\delta_5 = \beta_1 - \beta_3$.
(d) These two models do not have any relationship between them.

33. An economist who wants to study the consumption behaviour about an isotonic drink ($Y =$ liters consumed per year) by people who practice and people who do not practice sports ($DEP =$ 1 if the individual practices sports and 0 otherwise), considering also individual income level ($RENTA =$ Income in euros), specifies and estimates the following model:

$$E(Y \mid RENTA, DEP) = \beta_0 + \beta_1 RENTA + \beta_2 DEP + \beta_3 (RENTA \times DEP).$$

Suppose that we consider the variable $NOD = 1 - DEP$ and want to estimate the following model:

$$\begin{aligned} E(Y \mid RENTA, DEP, NOD) &= \delta_1 RENTA + \delta_2 DEP + \delta_3 (RENTA \times DEP) \\ &\quad + \delta_4 NOD + \delta_5 (RENTA \times NOD). \end{aligned}$$

Indicate what among the following options is correct:

(a) $\delta_2 = \beta_0$; $\delta_4 = \beta_0 + \beta_2$; $\delta_3 = \beta_3$; $\delta_5 = \beta_1 + \beta_3$.
(b) $\delta_2 = \beta_0 + \beta_2$; $\delta_4 = \beta_0$; $\delta_3 = \beta_1 + \beta_3$, $\delta_5 = \beta_1$.
(c) $\delta_2 = \beta_0$; $\delta_4 = \beta_0 + \beta_2$; $\delta_3 = \beta_1 + \beta_3$; $\delta_5 = \beta_1 - \beta_3$.

(d) We cannot estimate the second model because there is perfect multicollinearity.

34. Suppose that we are interested in the model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon,$$

where $E(\varepsilon | X_1, X_2) = 0$, $V(\varepsilon | X_1, X_2) = \sigma^2$. Furthermore, we know that $\beta_1 > 0$, $\beta_2 < 0$, $C(X_1, X_2) < 0$.
However, we only observe Y and X_1. Let b_1 be the OLS estimator of the slope of the simple regression of Y over X_1. Indicate which of the following assertions is FALSE:

(a) b_1 is an inconsistent estimator of β_1.

(b) b_1 will tend to overestimate β_1.

(c) b_1 will tend to underestimate β_1.

(d) Excluding X_2 makes X_1 is an endogenous variable in the model $Y = \beta_0 + \beta_1 X_1 + u$.

35. Suppose that we are interested in the model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon,$$

where $E(\varepsilon | X_1, X_2) = 0$, $V(\varepsilon | X_1, X_2) = \sigma^2$. Moreover, we know that $\beta_1 \neq 0$, $\beta_2 \neq 0$, but we do not know anything about $C(X_1, X_2)$.
We only observe Y , X_1 and other variable X_3 not included in the former model, X_3 being independent of both ε and X_2, but correlated with X_1.
Let b_1 be the OLS estimator of the slope of the simple regression of Y on X_1.
Consider the following assertions:
(i) We can verify if the omission of X_2 actually induces a consistency problem with b_1, by means of a Hausman test, thanks to the fact that X_3 is a valid instrument.
(ii) We cannot test whether the omission of X_2 actually induces a consistency problem with b_1 using X_3 as instrument because X_1 es exogenous and therefore we have a problem of omitted variable but not of endogeneity.
(iii) The relation between X_3 and X_2 is irrelevant to ensure the validity of X_3 as instrument in the simple regression of Y on X_1: it suffices that X_3 be independent of ε and correlated with X_1.

(a) Only (i) is true.

(b) Only (ii) is true.

(c) Only (iii) is true.

(d) Only (ii) and (iii) are true.

36. To ensure the consistency of the OLS estimators of the parameters of a linear regression model, indicate what of the following usual assumptions is NOT needed.

    (a) Conditional expectation of the error term (conditional on the explanatory variables) equal to zero.
    (b) Conditional homoskedasticity (conditional on the explanatory variables).
    (c) No correlation betwen each regressor and the error term.
    (d) Linearity in parameters.

37. Given the simple linear regression model

$$Y = \beta_0 + \beta_1 X + \varepsilon, \tag{D}$$

where $E(\varepsilon|\,X) = 0$, $V(\varepsilon|\,X) = \sigma^2$, $\beta_1 \neq 0$, we want to estimate the inverse model

$$X = \gamma_0 + \gamma_1 Y + v, \tag{C}$$

so that, $\gamma_0 = \dfrac{-\beta_0}{\beta_1}$, $\gamma_1 = \dfrac{1}{\beta_1}$ and $v = \dfrac{-\varepsilon}{\beta_1}$. Let b_1 be the OLS estimator of the slope in model (D), g_1 the OLS estimator of the slope in model (C), and $\widehat{g}_1$ an instrumental variable estimator of the slope in model (C), which uses X as instrument.
Indicate what of the following assertions is FALSE:

    (a) b_1 is a consistent estimator of β_1.
    (b) g_1 is a consistent estimator of γ_1.
    (c) $\widehat{g}_1$ is a consistent estimator of γ_1.
    (d) $\dfrac{1}{b_1}$ is a consistent estimator of γ_1.

38. Consider the model

$$Y = \beta_0 + \beta_1 X + \beta_2 \ln(X) + \varepsilon,$$

where $E(\varepsilon|X) = 0$, $\beta_1 \neq 0$, $\beta_2 \neq 0$. We can assert that:

    (a) $E(Y|X) = PLO(Y|X)$
    (b) $E(Y|X) = PLO(Y|X, \ln(X))$
    (c) The conditional expectation function of Y given X is linear in X.
    (d) $\beta_2 = C(\ln(X), Y)/V(\ln(X))$

39. Consider the model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon,$$

where both X_1 and X_2 are endogenous. Let Z be an additional variable that satisfies $C(Z, \varepsilon) = 0$, $C(Z, X_1) \neq 0$, $C(Z, X_2) \neq 0$. Then:

(a) Since Z is a valid instrument both for X_1 and X_2, we can estimate consistently the model parameters by means of an instrumental variable estimator that uses Z as the only instrument.

(b) Since there are two endogenous explanatory variables, we must use a two stage least squares estimator using Z as the only instrument for each endogenous variable in order to get consistent estimates of the model parameters.

(c) Even though β_1 y β_2 cannot be consistently estimated with the available information, if we omit X_2 we could consistently estimate β_1 by means of an instrumental variable estimator that uses Z as instrument. (so that we would be estimating the model $Y = \beta_0 + \beta_1 X_1 + v$, where $v = \varepsilon + \beta_2 X_2$).

(d) None of the other answers is true..

40. If we include an additional irrelevant variable in a linear multiple regression under the classical assumptions:

(a) The properties of the OLS estimator are unaffected, because the additional variable is irrelevant.

(b) It implies a specification error that implies inconsistency of the OLS estimator.

(c) The OLS estimator is still consistent although less efficient.

(d) None of the other answers is true..