

Hoja de Ejercicios de Econometría: Heterocedasticidad

Ejercicio 1

1.- Se pretende estudiar los determinantes de las horas trabajadas por las mujeres casadas para lo que se ha propuesto el siguiente modelo:

$$\text{horas} = \beta_0 + \beta_1 \text{hijospeq} + \beta_2 \text{hijosmay} + \beta_3 \text{edad} + \beta_4 \text{edad}^2 + \beta_5 \text{educ} + \beta_6 \log(\text{ingfam}) + \varepsilon$$

donde:

horas: es el nº de horas trabajadas al año por la mujer

hijospeq: es el nº de hijos de menos de 6 años que tiene la mujer

hijosmay: es el nº de hijos de 6 años o más que tiene la mujer

ingfam: son los ingresos familiares anuales en euros

edad: es la edad de la mujer

educ: son los años de estudio de la mujer

ε : es un término de perturbación que cumple $E(\varepsilon | \text{explicativas}) = 0$ y log indica logaritmo neperiano

Empleando datos de 428 mujeres casadas se ha estimado:

Ecuación 1:

Dependent Variable: HORAS

Method: Least Squares

Included observations: 428

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1775.575	1258.090	-1.411327	0.1589
HIJOSPEQ	-280.7094	98.88290	-2.838806	0.0047
HIJOSMAY	-129.2067	30.47934	-4.239157	0.0000
EDAD	8.892394	52.94656	0.167950	0.8667
EDAD^2	-0.237979	0.616153	-0.386233	0.6995
EDUC	-54.43718	16.98832	-3.204390	0.0015
LOG(INGFAM)	404.8818	80.29989	5.042122	0.0000
R-squared	0.118724	Mean dependent var		1302.930
Adjusted R-squared	0.106164	S.D. dependent var		776.2744
S.E. of regression	733.9122	Akaike info criterion		16.05088
Sum squared resid	2.27E+08	Schwarz criterion		16.11726
Log likelihood	-3427.887	F-statistic		9.452736
Durbin-Watson stat	2.036110	Prob(F-statistic)		0.000000

Wald Test:

Equation: Ecuación 1

Null Hypothesis: C(4)=0
C(5)=0

Test Statistic	Value	df	Probability
F-statistic	2.367406	(2, 421)	0.0950
Chi-square	4.734812	2	0.0937

White Heteroskedasticity Test:

Equation: Ecuación 1

F-statistic	1.661648	Prob. F(26,401)	0.023183
Obs*R-squared	41.62697	Prob. Chi-Square(26)	0.026810

Ecuación 2:

Dependent Variable: HORAS

Method: Least Squares

Sample: 1 428

Included observations: 428

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1775.575	1380.349	-1.286323	0.1990
HIJOSPEQ	-280.7094	129.0996	-2.174363	0.0302
HIJOSMAY	-129.2067	28.79240	-4.487527	0.0000
EDAD	8.892394	56.34303	0.157826	0.8747
EDAD^2	-0.237979	0.664306	-0.358236	0.7203
EDUC	-54.43718	16.57487	-3.284321	0.0011
LOG(INGFAM)	404.8818	72.52055	5.582994	0.0000
R-squared	0.118724	Mean dependent var		1302.930
Adjusted R-squared	0.106164	S.D. dependent var		776.2744
S.E. of regression	733.9122	Akaike info criterion		16.05088
Sum squared resid	2.27E+08	Schwarz criterion		16.11726
Log likelihood	-3427.887	F-statistic		9.452736
Durbin-Watson stat	2.036110	Prob(F-statistic)		0.000000

Wald Test:

Equation: Ecuación 2

Null Hypothesis: C(4)=0
C(5)=0

Test Statistic	Value	df	Probability
F-statistic	2.021177	(2, 421)	0.1338
Chi-square	4.042354	2	0.1325

Contraste al 10% si las variaciones en la edad de la mujer afectan a las horas trabajadas. Plantee la hipótesis nula, alternativa y justifique el resultado.

Ejercicio 2

Empleando datos de 500 empresas se ha estimado el modelo lineal que se presenta a continuación:

Salida 1

Dependent Variable: Ventas

Method: Least Squares

Included observations: 500

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	320.0229	0.402607	794.8758	0.0000
Sector	250.7840	0.531199	472.1096	0.0000
Public	0.699146	0.008418	83.04965	0.0000
Public*Sector	-0.021591	0.011506	-1.876562	0.0612
R-squared	0.998491			
Adjusted R-squared	0.998482			
S.E. of regression	4.909086			
Sum squared resid	11953.17			
		F-statistic	109414.6	
		Prob(F-statistic)	0.000000	

Salida 2

Dependent Variable: Ventas

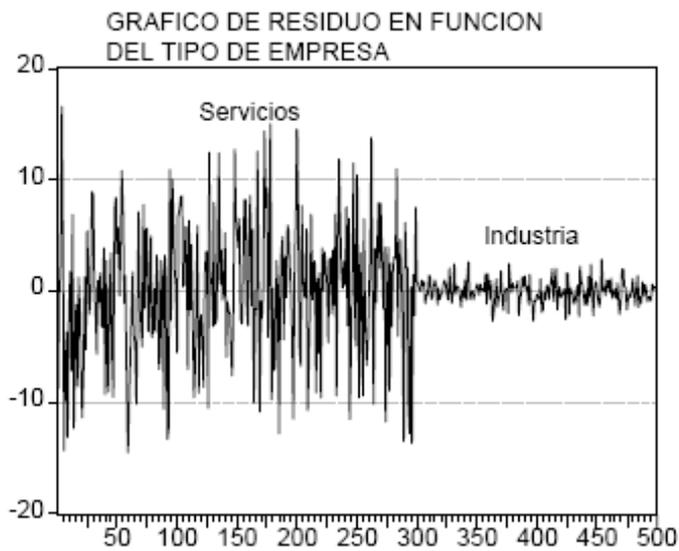
Method: Least Squares

Included observations: 500

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	320.0229	0.086691	3691.521	0.0000
Sector	250.7840	0.455807	550.1974	0.0000
Public	0.699146	0.001846	378.6482	0.0000
Public*Sector	-0.021591	0.009434	-2.288747	0.0225

R-squared	0.998491		
Adjusted R-squared	0.998482		
S.E. of regression	4.909086		
Sum squared resid	11953.17		
		F-statistic	109414.6
		Prob(F-statistic)	0.000000



El gráfico representa los residuos del modelo ordenados en función del tipo de empresa: servicios las 300 primeras observaciones e industria las 200 últimas observaciones. Contraste al 5% si el efecto de la publicidad sobre las ventas depende del sector al que pertenezca la empresa.

Nota: Sector es una variable ficticia que toma valor 1 cuando la empresa pertenece al sector industria y 0 al sector servicios.

Ejercicio 3

Se pretende estimar empleando datos de España el siguiente modelo:

$$Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 X_2 + \beta_3 (X_2 \times X_3) + \varepsilon$$

donde:

Y es el gasto en alimentación, bebidas y tabaco (en euros)

X_1 es el gasto total (en euros)

X_2 es el número de hijos menores de 17 años

X_3 es una variable ficticia que indica la situación laboral de la mujer (1 si trabaja y 0 en caso contrario)

Empleando datos, procedentes de la Encuesta de Presupuestos Familiares de 1990, de 899 familias con dos cónyuges en el hogar se ha obtenido:

Salida 1

Dependent Variable: Y

Method: Least Squares

Included observations: 899

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-17776.84	1424.957	-12.46833	0.0000
LOG(X1)	2116.400	157.2934	13.45511	0.0000
X2	311.4864	47.45287	6.522833	0.0000
X2*X3	246.9344	119.2371	2.070952	0.0387

Salida 2

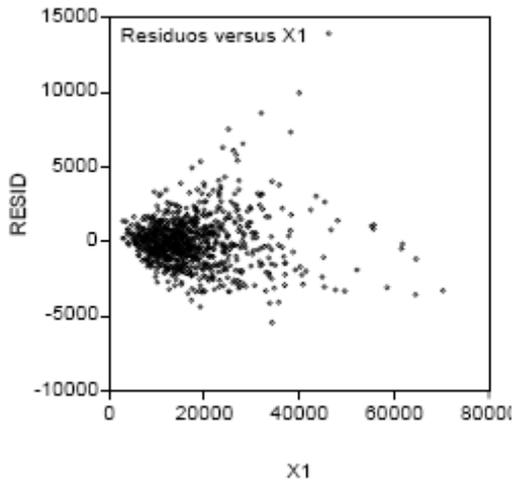
Dependent Variable: Y

Method: Least Squares

Included observations: 899

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-17776.84	1158.827	-15.33174	0.0000
LOG(X1)	2116.400	122.9337	17.21579	0.0000
X2	311.4864	61.10748	5.097352	0.0000
X2*X3	246.9344	155.4135	1.588886	0.1121

El siguiente gráfico representa los residuos de la estimación MCO del modelo frente a la variable X_1 :



- A la vista del gráfico ¿se cumplen todos los supuestos del modelo de regresión clásico?
- Interprete β_3 y su valor estimado. Contraste al 5% su significación individual.

Ejercicio 4

Dado el modelo:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \quad i = 1, \dots, n \quad [1]$$

en el que:

$$E(\varepsilon_i | X_{1i}, X_{2i}) = 0$$

$$V(\varepsilon_i | X_{1i}, X_{2i}) = \sigma^2 X_{1i}$$

Un investigador propone, con objeto de obtener estimaciones eficientes de los parámetros del modelo [1], que se estime por MCO el siguiente modelo:

$$\frac{Y_i}{X_{1i}} = \beta_1 + \beta_0 \frac{1}{X_{1i}} + \beta_2 \frac{X_{2i}}{X_{1i}} + \varepsilon_i^* \quad [2]$$

- ¿Qué propiedades tendrán las estimaciones de MCO obtenidas del modelo [2]?
- ¿Sería válida la inferencia sobre los parámetros realizada a partir de la estimación MCO de [2]?
- ¿Cómo podría obtener estimaciones eficientes de los parámetros del modelo [1]?