

Taller #6
Heteroscedasticidad
Econometría I 06169

Profesor: Julio César Alonso
Monitores: Ana María Lotero
Carlos I. Patiño

El ÉXITO quiere planear el número de almacenes a inaugurar en España en los próximos años. Para ello desea conocer la relación que existe entre el consumo final de las familias residentes y la renta bruta de las mismas, y usted ha sido contratado el ÉXITO como consultor. Usted cuenta con una base de datos para 18 regiones (ver archivo "T6I-02-04.xls"). Los datos se encuentran en millones de pesetas.

1. Estime un modelo que explique el consumo final de las familias.
2. Realice la prueba de Goldfeld-Quandt. Señale claramente la hipótesis nula, alterna y su decisión.
3. Realice la prueba de Breusch-Pagan. Señale claramente la hipótesis nula, alterna y su decisión.
4. Realice la prueba de White. Señale claramente la hipótesis nula, alterna y su decisión.
5. Si existe heteroscedasticidad, resuelva el problema y estime un modelo que no presente problemas de heteroscedasticidad. En caso de no existir problema, determine si existe o no multicolinealidad en el modelo.
6. Interprete los coeficientes estimados. Comente significancia individual.

Taller #6
Heteroscedasticidad
Respuestas Sugeridas
Econometría I 06169

Profesor: Julio César Alonso
Monitores: Ana María Lotero
Carlos I. Patiño

El ÉXITO quiere planear el número de almacenes a inaugurar en España en los próximos años. Para ello desea conocer la relación que existe entre el consumo final de las familias residentes y la renta bruta de las mismas, y usted ha sido contratado el ÉXITO como consultor. Usted cuenta con una base de datos para 18 regiones (ver archivo "T6-02-04.xls"). Los datos se encuentran en millones de pesetas.

1. Estime un modelo que explique el consumo final de las familias. El modelo es el siguiente:

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i \quad (1)$$

Donde Y_i representa el consumo final de las familias para la región i , y X_i representa la renta bruta de las familias para la región i . Los resultados de la estimación del modelo (1) se encuentran reportados en la Tabla 1.

Tabla 1.

	VARIABLE DEPENDIENTE: Y_i	
	Estadísticos t entre paréntesis	
	Ecuación 1	Ecuación 2
	MCO	MCP
Constante	-85.184,05 (-2,93) **	-1.054,0525 (-0,19)
X_i	1,1672 (91,52) ***	1,0901 (72,18) ***
R^2	0,99809	0,00221
R^2 ajustado	0,99797	-0,06015
F	8.375,21 ***	NA
BP	5,01 **	0,0015
# de Obs.	18	18

(*) nivel de significancia: 10%

(**) nivel de significancia: 5%

(***) nivel de significancia: 1%

MCO: Mínimos Cuadrados Ordinarios

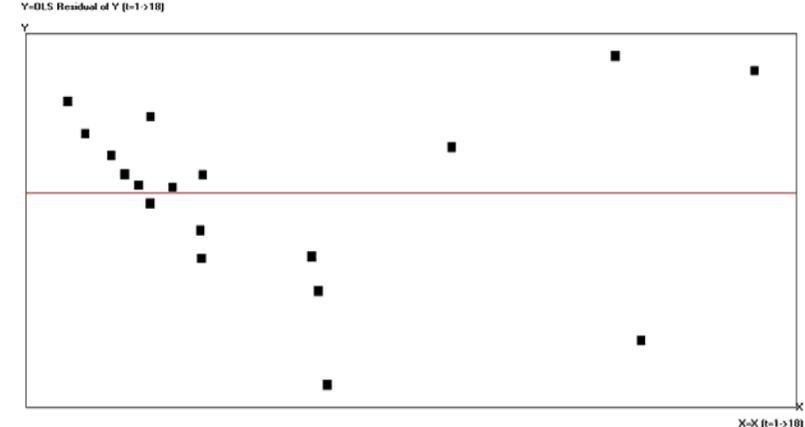
MCP: Mínimos Cuadrados Ponderados

NA: No Aplica

2. Realice la prueba de Goldfeld-Quandt. Señale claramente la hipótesis nula, alterna y su decisión.

Inicialmente se grafican los residuos estimados versus la variable X_i (Gráfico 1).

Gráfico 1. Residuos estimados versus renta bruta.



Se observa una relación positiva entre la variabilidad de los residuos y la renta bruta. De acuerdo a esto se lleva a cabo la prueba de Goldfeld y Quandt (GQ) planteando la siguiente hipótesis nula:

$$H_0 : \sigma_i^2 = \sigma^2$$

$$H_A : \sigma_i^2 = \sigma^2 X_i^2$$

El siguiente paso es ordenar las observaciones de acuerdo a la variable que genera el problema de heteroscedasticidad y posteriormente se indican las observaciones que forman las muestras empleadas para calcular el estadístico GQ. En este caso se eliminan 2 observaciones¹. La primera muestra contiene las observaciones 1 a 8 y la segunda muestra, las observaciones 11 a 18, es decir, se omiten las observaciones 9 y 10. Al estimar las dos regresiones a partir de las dos submuestras, se calcula el estadístico GQ:

$$GQ = \frac{SSE_2}{SSE_1} = \frac{6.8047 \times 10^{10}}{4.2111 \times 10^9} = 16.1572$$

Se compara con $F_{(n-d-2k, n-d-2k), \alpha=0.01} = F_{(12,12), \alpha=0.01} = 4.1552$. Dado que el estadístico GQ es mayor al valor crítico con 12 grados de libertad en el numerador y 12 grados de

¹ $d < \frac{1}{5}n \rightarrow d < 3.6$

libertad en el denominador, se rechaza la hipótesis nula a favor de la hipótesis alterna de heteroscedasticidad provocada por la variable X (renta bruta).

3. Realice la prueba de Breusch-Pagan. Señale la hipótesis nula y alterna.

EasyReg genera automáticamente un test de Breusch-Pagan asumiendo que todas las variables independientes afectan la varianza del error. En este caso, dado que solo se trata de una variable, se debe obtener los mismos resultados. De acuerdo a esto, para comprobar si la variable X_i afecta la varianza del error, se plantea la siguiente hipótesis nula:

$$H_0 : \sigma_i^2 = \sigma^2$$

$$H_A : \sigma_i^2 = f(\gamma + \delta X_i)$$

Con el fin de correr la regresión auxiliar se debe hallar $\hat{\sigma}^2$. En este caso, tenemos que: $SSE = \hat{\varepsilon}^T \hat{\varepsilon} = 1.0755 \times 10^{11}$ y así, $\hat{\sigma}^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{n} = \frac{1.0755 \times 10^{11}}{18} = 5.9752 \times 10^9$. La regresión auxiliar será entonces:

$$\frac{\hat{\varepsilon}_i^2}{\hat{\sigma}^2} = \gamma + \delta X_i + \mu_i.$$

A partir de la regresión auxiliar se obtiene:

$$SSR = SST - SSE = 30.1770 - 20.1613 = 10.0158$$

Ahora se construye el estadístico BP:

$$BP = \frac{SSR}{2} = \frac{10.0158}{2} = 5.0079$$

El cual se compara con $\chi_{g(\alpha)}^2 = \chi_{1(\alpha)}^2$. En este caso g representa el número de variables en Z, es decir, una variables. Los valores críticos obtenidos son 6.63, 3.84 y 2.70 para los niveles de significancia del 1%, 5% y 10% respectivamente. De acuerdo a esto, la hipótesis nula se puede rechazar al 5%, en favor de la presencia de heteroscedasticidad generada por la variable X. Este resultado es igual al generado automáticamente por EasyReg.

4. Realice la prueba de White.

Esta prueba es mucho más general:

$$H_0 : \sigma_i^2 = \sigma^2$$

$$H_A : No H_0$$

Para llevar a cabo esta prueba hay que efectuar una regresión auxiliar donde la variable dependiente sea los residuos al cuadrado de la ecuación inicial y las variables independientes sean la variable original y su cuadrado:

$$\hat{\varepsilon}_i^2 = \gamma + \delta_1 X_i + \delta_2 X_i^2 + \nu_i$$

El R^2 obtenido a partir de la regresión auxiliar es 0.344573. Ahora se puede construir el estadístico de White: $W_a = nR^2 = 18 \times 0.344573 = 6.2023$ el cual se debe comparar con $\chi_{g(\alpha)}^2 = \chi_{2(\alpha)}^2$. Los valores críticos obtenidos son 9.21, 5.99, 4.60 para los niveles de significancia del 1%, 5%, 10% respectivamente. De acuerdo a los valores críticos obtenidos, la hipótesis nula se rechaza al nivel de significancia del 5%. Se concluye entonces que la varianza de los errores no permanece constante a lo largo de la muestra.

5. Si existe heteroscedasticidad, resuelva el problema y estime un modelo que no presente problemas de heteroscedasticidad.

De acuerdo al análisis gráfico y a las pruebas realizadas se determina que la ecuación (1) presenta problemas de heteroscedasticidad. Dado que $Var[\varepsilon_i] = \sigma^2 X_i^2$ se puede emplear el método de Mínimos Cuadrados Ponderados (MCP). Conociendo el comportamiento de la varianza de los errores se puede transformar el modelo (1) de la siguiente forma:

$$\frac{Y_i}{X_i} = \beta_1 \frac{1}{X_i} + \beta_2 + \frac{\varepsilon_i}{X_i}$$

Reordenando se obtiene la ecuación (2)

$$\frac{Y_i}{X_i} = \beta_2 + \beta_1 \frac{1}{X_i} + \frac{\varepsilon_i}{X_i} \tag{2}$$

En esta ecuación el problema de heteroscedaticidad queda resuelto ya que:

$$Var\left[\frac{\varepsilon_i}{X_i}\right] = \frac{1}{X_i^2} Var[\varepsilon_i] = \frac{1}{X_i^2} \sigma^2 X_i^2 = \sigma^2$$

Los resultados de la estimación de la ecuación (2) son reportados en la Tabla 1. Como se observa, de acuerdo al estadístico BP arrojado por EasyReg, el modelo es homoscedástico.

6. Interprete los coeficientes estimados. Comente significancia individual $\beta_1 = -1054.05$: El consumo que no depende de la renta bruta es de -1,054.05 millones de pesetas. No es significativo a ningún nivel.

$\beta_2 = 1.09$: Un aumento de un millón de pesetas en la renta bruta generará un incremento de 1.09 millones de pesetas en el consumo. Significativo al 1%