

Taller #7

Heteroscedasticidad

Respuestas Sugeridas

Econometría 06216

Profesor: Julio César Alonso

**Monitores: Paul Semaán
Francisco Quevedo**

Notas:

- o Recuerde que tres preguntas, seleccionadas al azar, serán calificadas.
- o Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo **Marzo 15 del 2006**.

INSTRUCCIONES:

- Este taller debe ser escrito en computador y entregado en papel.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

Un productor de derivados del Aguacate desea realizar una inversión de planta en un país del caribe y desea conocer la relación existente entre las cantidades demandadas de este tipo de bien en esta nación. Para lo cual considera el comportamiento del nivel de ingreso y de otros productos de la misma gama en 30 regiones de este país, Los datos se encuentran en el archivo T7-01-06.xls. El inversionista ha contratado a un investigador, que se baso en investigaciones anteriores y sugirió el siguiente modelo.

$$Q_A = \beta_0 \cdot P_A^{\beta_1} \cdot P_B^{\beta_2} \cdot P_C^{\beta_3} \cdot INC^{\beta_4} \quad (1)$$

Donde Q_A y P_A son las cantidades demandadas de los derivados del aguacate (medida en decenas de miles de moneda local) y su nivel de precios (medido en miles de moneda local) respectivamente. P_B y P_C son los niveles de precios de los bienes B y C respectivamente que pertenecen a la misma gama de productos (medido en miles de moneda local), e INC es el nivel de ingreso de las familias en las diferentes regiones (medido en millones de moneda local).

Sin embargo el investigador llegó a la conclusión de que existía la posibilidad de la presencia de heteroscedasticidad dado que se emplean datos para 30 regiones del país. Pero, él no sabía como determinar con certeza la presencia de ésta. Así, el presidente de la compañía decide contratarlo a usted para que termine la investigación, y en especial le piden:

1. Estime el modelo (1) y reporte sus resultados en una tabla.
2. Muestre intuitivamente que el problema de heteroscedasticidad puede estar presente.
3. Realice las pruebas que determinen la existencia o no de un problema de heteroscedasticidad

4. En caso de existir heteroscedasticidad, resuelva el problema y estime un modelo que no presente problemas de heteroscedasticidad, Reporte los resultados del modelo corregido y de ser posible demuestre que el método empleado solucionó el problema.
5. Interprete el significado económico de los coeficientes estimados. Comente la significancia de ellos.
6. El inversionista desea comprobar que $-4\beta_2 = \beta_3$ y simultáneamente que $\beta_1 = -1$, para realizar un flujo de caja basándose en esta información, como lo ayudaría a tomar la decisión. (efectúe las pruebas necesarias y explique claramente la interpretación económica de la de la hipótesis probada y de la conclusión)

Taller #7
Heteroscedasticidad
Respuestas Sugeridas
Econometría 06216

Profesor: Julio César Alonso
Monitores: Paul Semaán
Francisco Quevedo

Notas:

- o Recuerde que tres preguntas, seleccionadas al azar, serán calificadas.
- o Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo **Marzo 15 del 2006**.

INSTRUCCIONES:

- Este taller debe ser escrito en computador y entregado en papel.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

Un productor de derivados del Aguacate desea realizar una inversión de planta en un país del caribe y desea conocer la relación existente entre las cantidades demandadas de este tipo de bien en esta nación. Para lo cual considera el comportamiento del nivel de ingreso y de otros productos de la misma gama en 30 regiones de este país, Los datos se encuentran en el archivo T7-01-06.xls. El inversionista ha contratado a un investigador, que se baso en investigaciones anteriores y sugirió el siguiente modelo.

$$Q_A = \beta_0 \cdot P_A^{\beta_1} \cdot P_B^{\beta_2} \cdot P_C^{\beta_3} \cdot INC^{\beta_4} \quad (1)$$

Donde Q_A y P_A son las cantidades demandadas de los derivados del aguacate (medida en decenas de miles de moneda local) y su nivel de precios (medido en miles de moneda local) respectivamente. P_B y P_C son los niveles de precios de los bienes B y C respectivamente que pertenecen a la misma gama de productos (medido en miles de moneda local), e INC es el nivel de ingreso de las familias en las diferentes regiones (medido en millones de moneda local).

Sin embargo el investigador llegó a la conclusión de que existía la posibilidad de la presencia de heteroscedasticidad dado que se emplean datos para 30 regiones del país. Pero, él no sabía como determinar con certeza la presencia de ésta. Así, el presidente de la compañía decide contratarlo a usted para que termine la investigación, y en especial le piden:

1. Estime el modelo (1) y reporte sus resultados en una tabla.

Primero se debe linealizar el modelo, empleando una transformación logaritmica se obtiene:

$$\ln(Q_A) = \alpha_0 + \beta_1 \ln(P_A) + \beta_2 \ln(P_B) + \beta_3 \ln(P_C) + \beta_4 \ln(INC) + \varepsilon \quad (2)$$

, donde $\alpha_0 = \ln(\beta_0)$

Los resultados de la estimación del modelo (2) se encuentran reportados en la Tabla 1.

Tabla 1.

| | VARIABLE DEPENDIENTE: $\ln(Q_A)$ | |
|-------------------------|----------------------------------|------------------------|
| | Estadísticos t entre paréntesis | |
| | Ecuación 1 | Ecuación 2 |
| | MCO | MCO |
| Constante | 4,2747 (3,55) *** | 4,2747 (4,75) *** |
| $\ln(P_A)$ | -0,8901 (-4,13) *** | -0,8901 (-6,56) *** |
| $\ln(P_B)$ | 0,5542 (4,26) *** | 0,5542 (6,65) *** |
| $\ln(P_C)$ | -0,3490 (-1,91) * | -0,3490 (-2,32) ** |
| $\ln(INC)$ | 0,7440 (5,50) *** | 0,7440 (6,26) *** |
| R ² | 0,66430 | 0,66430 |
| R ² ajustado | 0,61060 | 0,61060 |
| F | 12,37 *** | 12,37 *** |
| BP | 6,78 | 6,78 |
| # de Obs. | 30 | 30 |

(*) nivel de significancia: 10%

(**) nivel de significancia: 5%

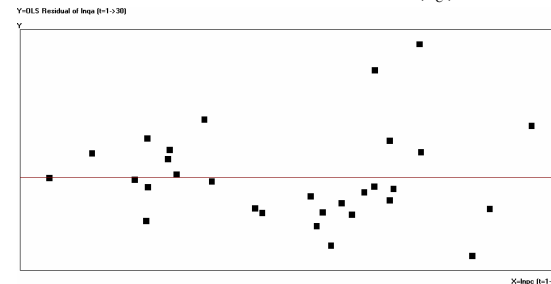
(***) nivel de significancia: 1%

MCO: Mínimos Cuadrados Ordinarios

2. Muestre intuitivamente que el problema de heteroscedasticidad puede estar presente.

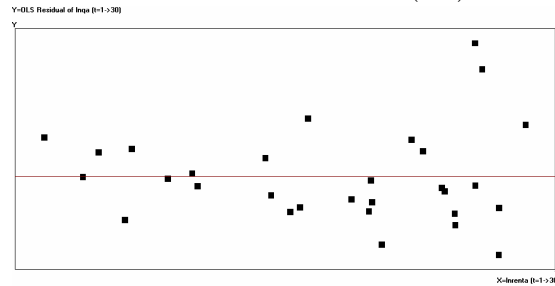
Se grafican los residuos estimados versus la variable $\ln(P_C)$ (Gráfico 1).

Gráfico 1. Residuos estimados versus $\ln(P_C)$.



Y se grafican los residuos estimados versus la variable $\ln(INC)$ (Gráfico 2).

Gráfico 2. Residuos estimados versus $\ln(INC)$.



Se observa una relación positiva entre la variabilidad de los residuos y el logaritmo de la renta por una parte y también se observa una relación positiva entre la variabilidad de los residuos y el logaritmo del precio del bien C . Así se dan los primeros indicios sobre la existencia de heteroscedasticidad en las dos variables.

3. Realice las pruebas que determinen la existencia o no de un problema de heteroscedasticidad

De acuerdo a esto se lleva a cabo la prueba de Goldfeld y Quandt (GQ) planteando la siguiente hipótesis nula en el caso general:

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

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

Los siguientes pasos a realizar para cada una de las variables son:

- Elegir una variable, o una combinación de variables (que generen el problema de heteroscedasticidad) y ordenar las observaciones, en nuestro caso $\ln(P_C)$ y $\ln(INC)$.
- Se indican las observaciones que forman las muestras empleadas para calcular el estadístico GQ, Determinando el numero que se van eliminar, el criterio es $d < \frac{1}{5}n < 6$, es decir, que el numero de observaciones adecuado para omitir es 4. De esta manera, se debe definir dos grupos omitiendo estos datos de la mitad, por lo que cada grupo queda conformado por trece observaciones. Se corre una regresión para cada grupo.
- Calcular el estadístico GQ, teniendo en cuenta los errores obtenidos en cada una de las regresiones del punto anterior, así: $GQ = \frac{SSE_2}{SSE_1}$
- Toma de decisión: Se rechaza la hipótesis nula si $GQ > F_{(n-d-2k, n-d-2k)\alpha=0.01}$, es decir, $F_{(16,16)\alpha=0.01} = 3.44 < GQ$.

El procedimiento debe ser realizado en las variables $\ln(P_C)$ y $\ln(INC)$, los resultados se resumen en la Tabla 2.

Tabla 2.

| | Nombre de la Variable | |
|------------------|-----------------------|------------|
| | $\ln(P_C)$ | $\ln(INC)$ |
| SSE ₁ | 0,022 | 0,022 |
| SSE ₂ | 0,091 | 0,107 |
| F _{GQ} | 4,039 *** | 4,791 *** |

(*) nivel de significancia: 10%
 (**) nivel de significancia: 5%
 (***) nivel de significancia: 1%

Como se observa en la Tabla 2 se puede rechazar la hipótesis nula, favoreciendo la presencia de heteroscedasticidad, del tipo descrito en la hipótesis alterna.

Basándonos en los resultados anteriores se procede a realizar la prueba de Breusch-Pagan, EasyReg genera automáticamente un test de Breusch-Pagan asumiendo que todas las variables independientes afectan la varianza del error. En este caso, dado que en el modelo hay más de una variable independiente, los resultados serán diferentes.

Por lo tanto en esta prueba se considera la hipótesis de una relación entre la varianza del error y un grupo de variables como medida de heteroscedasticidad, versus la hipótesis nula de la no existencia de esta relación como medida de homoscedasticidad, es decir:

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

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

Con el fin de correr la regresión auxiliar se debe hallar $\hat{\sigma}^2$. En este caso, tenemos que: $SSE = \hat{\epsilon}^T \hat{\epsilon} = 0,145117$ y así, $\hat{\sigma}^2 = \frac{\hat{\epsilon}^T \hat{\epsilon}}{n} = \frac{0,145117}{30} = 0,00483723$. La regresión auxiliar será entonces:

$$\frac{\hat{\epsilon}_i^2}{\hat{\sigma}^2} = \gamma + \delta \ln(P_{C_i}) + \gamma \ln(INC_i) + \mu_i$$

A partir de la regresión auxiliar se obtiene:

$$SSR = SST - SSE = 92,85135 - 80,43000 = 12,421350$$

Ahora se construye el estadístico BP:

$$BP = \frac{SSR}{2} = \frac{12,421350}{2} = 6,210675$$

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, dos variables. Los valores críticos obtenidos son 9.21, 5.99 y 4.61 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 las variables $\ln(P_C)$ y $\ln(INC)$. Este resultado es igual al generado automáticamente por EasyReg.

Por último se realiza el Test de White, en el que se desea comprobar:

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

$$H_A : No H_0$$

Para llevar a cabo esta prueba hay que efectuar una regresión auxiliar donde la variable dependiente sea igual a los residuos al cuadrado de la formulación inicial y las variables independientes sean conformadas por regresores del modelo original, a sus cuadrados y a los productos cruzados, después de haber eliminado las posibles redundancias se obtiene:

$$\begin{aligned} \hat{\epsilon}_i^2 = & \gamma + \delta_1 \ln(P_A) + \delta_2 \ln(P_B) + \delta_3 \ln(P_C) + \delta_4 \ln(INC) + \delta_5 \ln(P_A) \cdot \ln(P_B) \\ & + \delta_6 \ln(P_A) \cdot \ln(P_C) + \delta_7 \ln(P_A) \cdot \ln(INC) + \delta_8 \ln(P_A)^2 + \delta_9 \ln(P_B) \cdot \ln(P_C) \\ & + \delta_{10} \ln(P_B) \cdot \ln(INC) + \delta_{11} \ln(P_B)^2 + \delta_{12} \ln(P_C) \cdot \ln(INC) + \delta_{13} \ln(P_C)^2 \\ & + \delta_{14} \ln(INC)^2 + \mu \end{aligned}$$

El R^2 obtenido a partir de la regresión auxiliar es igual a 0.4296. Ahora se puede construir el estadístico de White: $W_a = nR^2 = 30 \times 0.4296 = 12.888$ el cual se debe comparar con $\chi_{g(\alpha)}^2 = \chi_{14(\alpha)}^2$. El valor crítico obtenido para un nivel de significancia del 10% es igual a 21.06. De acuerdo a los valores críticos obtenidos, la hipótesis nula no se puede rechazar al nivel de significancia del 10%. Puede ser que dado que esta prueba es más general, su poder de predicción pierde potencia al tratar de identificar el problema de heteroscedasticidad.

Sin embargo se observa que el test de Goldfeld-Quandt y el test de Breusch-Pagan confirman la noción de presencia de heteroscedasticidad en las variables descritas anteriormente, por lo tanto al existir un patrón en estas pruebas se concluye que existen problemas de heteroscedasticidad.

4. En caso de existir heteroscedasticidad, resuelva el problema y estime un modelo que no presente problemas de heteroscedasticidad, Reporte los resultados del modelo corregido y de ser posible demuestre que el método empleado solucionó el problema.

En este caso se hace necesario aplicar la solución de White al problema de heteroscedasticidad, toda vez que los resultados de las pruebas no permiten conocer la naturaleza del problema de heteroscedasticidad, y por lo tanto no es posible conocer el comportamiento de la varianza y así poder aplicar el método de Mínimos Cuadrados Ponderados.

White (1980) sugiere el siguiente estimador consistente para la matriz de varianzas y covarianzas de los β 's estimados en presencia de heteroscedasticidad:

$$Est.Var[\hat{\beta}] = n(X^T X)^{-1} S_0 (X^T X)^{-1},$$

$$\text{donde } S_0 = \frac{1}{n} \sum_{i=1}^n \hat{\epsilon}_i^2 x_i x_i^T \text{ y } x_i^T = (1 \quad x_{1i} \quad x_{2i} \quad \dots \quad x_{ki}).$$

Los resultados la solución de White se resumen en la Tabla 1.

5. Interprete el significado económico de los coeficientes estimados. Comente la significancia de ellos.

$\alpha_0 = 4.2747$: Es la porción de las cantidades demandadas que no depende del nivel de precios y del ingreso ,es significativo al 99%.

$\beta_1 = -0.89012$: Cuando aumenta el precio del bien A en un 1%, las cantidades demandadas del bien A disminuyen en 0.89%. Corresponde al valor de la elasticidad respecto al precio del bien A (a su propio precio) dado el signo negativo se refiere a un bien normal, esta variable es significativa al 99%.

$\beta_2 = 0.55419$: Cuando aumenta el nivel de precios del bien B en un 1%, las cantidades demandadas del bien A aumentan en 0.55%. es decir el bien B es un sustituto del bien A , es significativo al 99%

$\beta_3 = -0.34909$: Cuando aumenta el nivel de precios del bien C en un 1% las cantidades demandadas del bien A disminuyen en 0.34%. es decir el bien B es un bien complementario del bien A , es significativo al 10%.

$\beta_4 = 0.74395$: Cuando aumenta la renta en un 1% las cantidades demandadas del bien A aumentan en 0.74%. Es la elasticidad renta, al ser de signo positivo indica que no es un bien inferior, es significativo al 99%.

6. El inversionista desea comprobar que $-4\beta_2 = \beta_3$ y simultáneamente que $\beta_1 = -1$, para realizar un flujo de caja basándose en esta información, como lo ayudaría a tomar la decisión. (efectúe las pruebas necesarias y explique claramente la interpretación económica de la de la hipótesis probada y de la conclusión)

La prueba de hipótesis a realizar sería:

$$H_0 : 4\beta_2 + \beta_3 = 0 \text{ y } \beta_1 = -1$$

$$H_A : No H_0$$

Note que esta hipótesis se puede escribir de la forma $R\beta = C$, donde

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 1 \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \text{ y } C = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

El estadístico Wald es 11.37, que sigue una distribución chi-cuadrado con 2 grados de libertad, al compararse con los valores críticos iguales a 4.61 y 5.99, con niveles de significancia del 10% y 5% respectivamente es mayor, y por lo tanto podemos rechazar la hipótesis nula.