

# ECONOMETRÍA

## APLICADA UTILIZANDO R.

---

PAPIME PE302513 LIBRO ELECTRÓNICO Y COMPLEMENTOS DIDÁCTICOS EN MEDIOS COMPUTACIONALES, PARA EL FORTALECIMIENTO DE LA ENSEÑANZA DE LA ECONOMETRÍA

### Capítulo 7.

Heterocedasticidad

Jorge Feregrino Feregrino



## Objetivo

La detección de la heterocedasticidad juega un papel importante dentro del modelo de regresión lineal, el objetivo de este capítulo es identificar la heterocedasticidad para poder aplicar técnicas de corrección.



## Introducción

- ❑ El origen de la heterocedasticidad, está asociado a la varianza creciente de las perturbaciones aleatorias del comportamiento, de las variables incluidas en la estimación de un modelo.
- ❑ La detección de la heterocedasticidad en la mayoría de los procedimientos es útil para establecer algún tipo de solución que permite corregir este problema.
- ❑ El análisis de regresión condicional implica, obtener un parámetro estable y útil entre ambas variables, la dispersión entre las variables debe ser aleatoria para evitar problemas de estimación e inferencia econométrica.

## Supuestos para la estimación en presencia de heterocedasticidad

- ❑ La presencia de heterocedasticidad en los modelos de regresión lineal, rompe con una restricción importante en econometría, donde la varianza de los errores aleatorios es condicional a los valores de la variable independiente  $X$ , es constante:

$$\text{Var}(v_t|x_t) = \sigma^2$$

- ❑ La restricción implica que los valores de la muestra de la variable dependiente ( $y$ ) son iguales las varianzas de los errores ( $v$ ), para los distintos valores de ( $x$ ), es decir, la dispersión en relación a la minimización de los errores permite representar los valores de ( $y$ ) mediante la variable estimada ( $\hat{y}$ ) de manera eficiente, insesgada y consistente.

## Estimación necesaria

- ❑ La varianza del estimador de Mínimos Cuadrados Ordinarios, no pueden calcularse con la expresión usual cuando se ha detectado heterocedasticidad:

$$v(\beta) = \sigma^2 (X'X)^{-1}$$

- ❑ La expresión anterior es un estimador sesgado de la varianza de los parámetros; alternativamente, debe utilizarse la siguiente expresión:

$$\text{covar-var}(\beta) = \sigma^2 [X'X]^{-1} X' \Sigma X [X'X]^{-1} + \varepsilon_i$$

## Errores en la estimación

- ❑ En presencia de heterocedasticidad se cometerá un error al calcular la varianza, esto implica, básicamente, que el contraste de la “t” de student, no podrá comprarse con los valores de referencia y lo mismo para los contraste que tienen origen en la varianza estimada.
- ❑ Por ejemplo, el contraste “F” ya no se distribuirá como una “F” o los contrastes que se utilizan como referencia a la j-cuadrada.

## Supuesto para la estimación

- ❑ Además, todos aquellos contrastes con base en este estimador serán también incorrectos. Un error frecuente, consiste en suponer errores al calcular la R2.
- ❑ Cuando se utiliza una población para calcular R2, no se utilizarán varianzas condicionales a los valores de las variables explicativas, entonces, no es afectada por la presencia de heterocedasticidad, y puede utilizarse como contraste.

$$\tilde{\sigma}^2 = \frac{e'e}{n - k}$$

## Causas de la heterocedasticidad

- ❑ Una de las causas mas importantes: es la omisión de las variables en la especificación del modelo, ante la imposibilidad de controlar todos los determinantes sobre la variable independiente.
- ❑ Aunque es imposible aseverar que las variables explicativas tengan una varianza constante.
- ❑ La inclusión de variables exógenas en la especificación del modelo cuya varianza crece en el tiempo, puede influir en la varianza de las perturbaciones y perder su condición de aleatoriedad.
- ❑ Cambio estructural: Un cambio de estructural puede provocar un ajuste erróneo de los parámetros en la estimación. Esto se produce en algunas secciones de la muestra y genera diversos desajustes en el modelo.



## Causas de la heterocedasticidad

- ❑ Errores en la especificación de la forma funcional: la utilización de una forma funcional incorrecta, provoca un ajuste con errores crecientes y alta dispersión. Por ejemplo, la utilización de una función lineal en lugar de una logarítmica potencial, tasa de crecimiento porcentual o una función cuadrática
- ❑ Fallas en el supuesto de normalidad de las variables explicativas: cuando se incluyen variables explicativas cuya distribución incumple normalidad y los regresores estarán asociados a una mayor dispersión en las perturbaciones.
- ❑ La presencia de valores atípicos en la muestra: esto implica desajustes en la varianza de las perturbaciones, por lo regular pertenecen a otro tipo de distribuciones y, por tanto, tienen una varianza diversa.

## Problemas en la detección de heterocedasticidad

- ❑ Es imposible observar directamente la presencia de heterocedasticidad, en la mayoría de los análisis econométricos, solo se dispone de un par de valores para cada valor  $(y,x)$ , entonces, es imposible observar si la varianza de los errores “U” para cada valor de “X” es la misma en toda la muestra.
- ❑ Por tanto, la mayor parte de los métodos se apoyarán en los residuos obtenidos en un modelo previo, estos se utilizarán como una muestra valida de las perturbaciones aleatorias desconocidas.



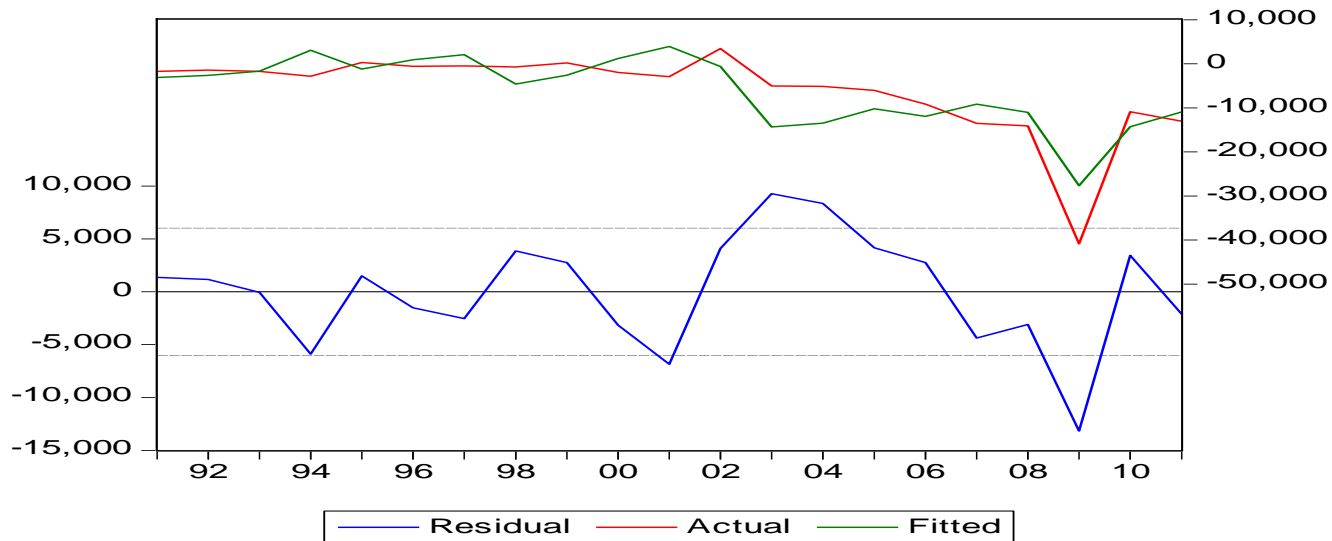
## La eficiencia del modelo y su poder explicativo

- ❑ Antes de utilizar cualquier método debe realizarse un análisis previo para identificar mediante análisis descriptivos y gráficos la naturaleza de los fenómenos.
- ❑ Graficar los errores, permitirá observar una tendencia definida para identificar intuitivamente si los errores crecen en el tiempo y si la varianza es heterogénea, es decir, en el caso de que se presenten valores de los errores mayores en el tiempo.



## Análisis gráfico

- ❑ La posible heterocedasticidad puede observarse cuando los errores comienzan a superar las bandas de dispersión a partir de 2002 y se acelera en 2008, es decir, la varianza crece en el tiempo.



La evolución en el tiempo está correlacionada con valores de la serie cada vez mayores sobre todo a partir de la crisis de 2008. El cálculo de la varianza por subperíodos, por ejemplo: entre 1992 y 2000 arrojaría valores significativamente diferentes; es decir, el error estimado es heterocedástico.

## Contraste Breusch-Pagan

- Hay procedimientos que permiten cuantificar la heterocedasticidad, y valorar su existencia en términos de la probabilidad, recurriendo a contrastes paramétricos.
- La idea del contraste es comprobar si se puede encontrar un conjunto de variables, que permitan determinar la dinámica de la varianza de las perturbaciones, estimada a partir del cuadrado de los errores del modelo inicial.

## Procedimiento Breusch-Pagan

- 1) Estimar el modelo MCO, para detectar heterocedasticidad, y obtener los errores de la estimación
- 2) Calcular una serie con los errores del modelo anterior al cuadrado estandarizados:

$$\tilde{e}_i^2 = \frac{e_i^2}{\hat{\sigma}^2}$$
$$\hat{\sigma}^2 = \frac{e'e}{n}$$

Criterios de la prueba F

- 3) Estimar una regresión sobre los determinantes de los errores mediante la incorporación de variables independientes (Z), para establecer si este conjunto explica el proceso de heterocedasticidad de las perturbaciones, la estimación propuesta es la siguiente:

$$\tilde{e}_i^2 = \alpha_0 + \alpha_1 z_{1i} + \alpha_2 z_{2i} + \dots + \alpha_p z_{pi} + \varepsilon_t$$

## El contraste Breusch-Pagan (1979)

- ❑ El modelo es ineficiente si la varianza de la variable dependiente estimada y su error es grande. Entonces, el poder explicativo del conjunto de variables Z sobre los errores es limitado. El contraste es el siguiente:

$$\frac{\sum \widehat{\tilde{e}_i^2} * n}{2}$$

- ❑ Breusch y Pagan, mostraron que el contraste se distribuye como una ji-cuadrada, cuando el proceso es homocedástico, se toman en cuenta las siguientes hipótesis:

*H<sub>0</sub>: presencia de homcedastacidad*

*H<sub>a</sub>: se acepta la presencia de heterocedasticdad*

## El contraste de White

- ❑ El contraste de White es una prueba robusta al no requerir supuestos previos como la normalidad de las perturbaciones. Tampoco es necesario determinar a priori las variables explicativas que determinan heterocedasticidad.
- ❑ El objetivo es determinar si las variables explicativas del modelo determinan la evolución de los errores al cuadrado. Es decir; si las variables explicativas en relación a las varianzas y covarianzas son significativas para determinar el valor de la varianza muestral de los errores.



## Aplicación del contraste de White

- ❑ El contraste White es aplicado a los residuos elevados al cuadrado resultados de la estimación de la función objetivo:

$$y_t = c + b_1x_{1t} + b_2x_{2t} + e$$

- ❑ Estimación de White:

$$e_t^2 = \alpha_0 + \alpha_1x_1 + \alpha_2x_2 + \alpha_3x_1^2 + \alpha_4x_2^2 + \alpha_5x_1x_2 + v_t$$

- ❑ La estimación incluye a las dos variables independientes elevadas al cuadrado, la combinación lineal de estas para explicar los residuos al cuadrado.

## Lectura del contraste de White

- ❑ La observación del contraste  $r^2$  en la estimación de White sobre los errores, en presencia de heterocedasticidad debería aproximarse a cero.
- ❑ De igual manera si el contraste de F tiene como resultado un grado de significancia explicativa conjunta del modelo, entonces, las series tienen un comportamiento heterocedástico.

