

## Trabajo Fin de Grado

### ESTUDIO DE TAMAÑO Y POTENCIA DE ALGUNOS CONTRASTES DE HETEROCEDASTICIDAD

Autor

Alberto Tafalla Sánchez

Director/as

M<sup>a</sup> Teresa Aparicio Aspas  
M<sup>a</sup> Isabel Ayuda Bosque

Facultad de Economía y Empresa de Zaragoza  
Año 2014

# ESTUDIO DE TAMAÑO Y POTENCIA DE ALGUNOS CONTRASTES DE HETEROSCEDASTICIDAD

Alberto Tafalla Sánchez

## RESUMEN

Este trabajo lleva a cabo un análisis de tamaño y potencia de algunos contrastes de heterocedasticidad en el contexto del modelo de regresión lineal. Resulta muy importante que la detección del problema de la heterocedasticidad se realice de manera adecuada, ya que su presencia implica que los estimadores obtenidos dejan de tener propiedades deseables y, adicionalmente, invalida la inferencia habitual sobre los parámetros de los modelos. Por ello se analiza la adecuación de dos contrastes de heterocedasticidad: el contraste específico de Breusch-Pagan y el contraste general de White. El estudio del comportamiento de ambos contrastes se lleva a cabo en términos de tamaño y potencia de los mismos. Para llevar a cabo el análisis, se han diseñado dos experimentos de Monte Carlo en los que se asumen varios supuestos sobre la estructura de la varianza del error, distintos grados de heterocedasticidad y varios tamaños muestrales. A partir de estos experimentos se analiza el tamaño de error tipo I y la potencia calculadas. Se concluye que el contraste de Breusch-Pagan es el más eficaz para detectar la presencia de heterocedasticidad.

Palabras clave: heterocedasticidad; econometría; Monte Carlo; gretl; simulación; White ;Breusch-Pagan

## **ENGLISH ABSTRACT**

### **ANALYSIS OF SIZE AND POWER OF SOME HETEROCEDASTICITY TEST'S**

**Alberto Tafalla Sánchez**

#### **ABSTRACT**

This paper examines the size and the power of some heterocedasticity tests in the context of the linear regression model. It is important that the detection of the problem of the heterocedasticity may be examined in a suitable way because their presence implies that the obtained estimators do not have desirable properties and the usual inference procedures about the parameters are no longer appropriate. Therefore, the Breusch-Pagan specific test and the White general test are analyzed. The study of the behavior of both tests is performed in terms of their size and power. The results show that Breusch-Pagan test has better properties than the White test.

Key words: heterocedasticiy; econometrics; Monte Carlo experiment; gretl; Breusch-Pagan; White.

## ÍNDICE

|   |           |
|---|-----------|
| <b>1. Introducción.....</b>   | <b>5</b>  |
| 1.1. La econometría y su importancia en economía .....                    | 5         |
| 1.2. Líneas generales del trabajo.....                                    | 6         |
| <b>2. Marco teórico.....</b>  | <b>7</b>  |
| 2.1. El modelo lineal general .....                                       | 7         |
| 2.1.1. Supuestos básicos del modelo clásico de regresión lineal.....      | 7         |
| 2.1.2. Propiedades de los estimadores .....                               | 8         |
| 2.2. El problema de la heterocedasticidad: causas y consecuencias .....   | 9         |
| 2.2.1. Causas de la heterocedasticidad .....                              | 9         |
| 2.2.2. Consecuencias de la heterocedasticidad .....                       | 10        |
| 2.3. Contrastes de heterocedasticidad .....                               | 11        |
| 2.3.1. El Contraste de White.....   | 11        |
| 2.3.2. El Contraste de Breusch-Pagan .....                                | 13        |
| <b>3. Análisis del comportamiento de los contrastes .....</b>             | <b>15</b> |
| 3.1. Conceptos de interés .....   | 15        |
| 3.1.1. Conceptos básicos. Tamaño y potencia de un contraste.....          | 15        |
| 3.2. La Metodología de análisis: Estudios de Monte Carlo .....            | 16        |
| 3.2.1. Introducción a los estudios de Monte Carlo .....                   | 16        |
| 3.3. Objetivo y diseño del estudio de Monte Carlo.....                    | 18        |
| <b>4. Resultados de los experimentos de montecarlo .....</b>              | <b>19</b> |
| 4.1. Experimentos de Monte Carlo para el modelo lineal simple ( M1).....  | 19        |
| 4.2. Experimentos de Monte Carlo para el modelo lineal general ( M2)..... | 23        |
| <b>5. Conclusiones .....</b>  | <b>32</b> |
| <b>6. Bibliografía .....</b>  | <b>33</b> |

## ANEXO

1.- Guión de instrucciones modelo M1

2.-Guion de instrucciones modelo M2

# 1. INTRODUCCIÓN

En esta introducción, en primer lugar, se va a realizar una breve exposición del contexto en el que se va a desarrollar el trabajo y de su importancia en economía, con el objeto de explicar las motivaciones por las cuales se ha decidido realizar el mismo. Posteriormente, en segundo lugar, se van a describir las líneas generales del trabajo, los objetivos del mismo y una breve ilustración de cómo está estructurado éste.

## 1.1 . CONTEXTO DEL TRABAJO: LA ECONOMETRIA Y SU IMPORTANCIA EN ECONOMÍA. MOTIVACIONES DEL TRABAJO.

*Goldberger (1964) “La econometría se define como la ciencia social en la cual las herramientas de la teoría económica, las matemáticas y la inferencia estadística se aplican al análisis de los fenómenos económicos.”*

En economía, resulta fundamental el análisis de los fenómenos económicos y el comportamiento de los agentes económicos. Para ello, los economistas utilizan los modelos económicos teóricos y matemáticos como herramienta para darle soporte a tales fenómenos o relaciones que se producen. El grado de aceptación de las teorías deberá evaluarse mediante la confrontación de las implicaciones de estas teorías con la base empírica, es decir, con los datos. En este sentido, es crucial el papel de la econometría, ya que trata de suministrar las técnicas necesarias para llevar a cabo la mencionada confrontación. La dificultad radica en que al ser la economía una ciencia social, no se puede llevar a cabo la obtención de los datos mediante experimentos controlados sino que hay que utilizar los datos provenientes de la observación o de la simulación mediante ordenador.

Las hipótesis a verificar se sitúan en el seno de los modelos económicos, pero la econometría no trabaja con ellos directamente, sino con los denominados modelos econométricos. Estos se definen como aquellos modelos económicos que contienen el conjunto de hipótesis necesarias para su aplicación empírica. Los modelos econométricos constituyen un instrumento que permite la confrontación entre teoría y realidad. No obstante, éstos pueden no estar correctamente especificados, por lo tanto, resulta fundamental elegir las técnicas más adecuadas para la detección de los problemas y que esa detección de los problemas se realice correctamente.

La aplicación de las habilidades adquiridas a lo largo del Grado de Economía en los campos de las matemáticas, la estadística y la econometría para resolver un problema determinado que afecta a las variables económicas, constituye la motivación fundamental del presente trabajo. De ésta manera, queda patente la importancia de otorgarle una mayor visibilidad práctica a un campo de la economía como es la econometría.

Por lo tanto, la línea de este trabajo va dirigida principalmente a ilustrar la importancia de elegir bien las técnicas econométrico-estadísticas, que en nuestro caso serán los contrastes, a la hora de solucionar un problema, *-heterocedasticad-*, en un determinado contexto- *Modelo Lineal General-* con un tipo de datos *-corte transversal-*.

La realización de éste trabajo puede resultar muy útil para entender la capacidad que diferentes contrastes tienen para captar un problema concreto, para analizar su utilidad bajo diferentes supuestos y su validez según la cantidad de información (tamaño muestral) disponible.

## **1.2. LINEAS GENERALES DEL TRABAJO.**

Las propiedades del estimador mínimo cuadrático ordinario de los parámetros de posición de un modelo lineal general dependen de las propiedades del término de error o perturbación aleatoria del modelo. Entre los supuestos asumidos, se asume que las perturbaciones aleatorias tienen una esperanza igual a cero, una varianza constante, que son independientes entre sí y que se distribuyen según una Normal.

En éste trabajo, a partir de la violación del supuesto de varianza constante también conocido como homocedasticidad, realizaremos un análisis de dos contrastes estadísticos (Breusch-Pagan y White) cuya finalidad consiste en detectar éste problema de varianza no constante o heterocedasticidad. Mediante experimentos de Monte Carlo vamos a observar cómo se comportan los distintos contrastes, en términos de tamaño y potencia, para distintos patrones de heterocedasticidad.

Por lo tanto, el objetivo del trabajo va a consistir en demostrar cuál de los dos contrastes estadísticos tiene una mayor potencia y un menor tamaño, es decir, si detectan mejor la presencia o no de heterocedasticidad bajo una serie de condiciones.

El contenido del presente trabajo, está estructurado en cinco bloques que se detallan a continuación:

En primer lugar se va a realizar un repaso del marco teórico en el que se sitúa de forma específica el trabajo; por un lado, de forma muy breve se va a explicar el modelo lineal general, el método de mínimos cuadrados y las propiedades de los estimadores mínimo cuadrático ordinarios bajo el cumplimiento de las hipótesis ideales. Por otro lado se introducirá al problema de la heterocedasticidad explicando para ello las causas y consecuencias de éste problema; por último se van a desarrollar los contrastes estadísticos objeto del análisis en el presente trabajo.

En segundo lugar, en el apartado tres, para realizar un análisis del comportamiento de los contrastes, se va a repasar algunos conceptos de interés, como el de tamaño y *la potencia* de un contraste estadístico por un lado, y por otro lado se va a explicar la metodología de análisis del trabajo, es decir, los estudios de Monte Carlo.

En el apartado cuatro, se presentarán los resultados de los experimentos de Monte Carlo.

Por último, en el apartado cinco, nos centraremos en las conclusiones del trabajo.

## **2. MARCO TEÓRICO.**

### **2.1. MODELO LINEAL GENERAL. PROPIEDADES DE LOS ESTIMADORES MINIMO CUADRÁTICOS.**

#### **2.1.1. Modelo clásico de regresión lineal: supuestos básicos.**

El modelo de Gauss o modelo clásico de regresión lineal es el cimiento de la mayor parte de la teoría econométrica y plantea siete supuestos. Dado un modelo general:

$$Y = X\beta + u$$

Donde  $Y$  es el vector  $N \times 1$  de observaciones de la variable endógena,  $X$  es una matriz  $N \times k$  que recoge las  $N$  observaciones de las  $k$  variables explicativas,  $\beta$  el vector de parámetros de orden  $k \times 1$  y  $u$  el vector  $N \times 1$  de perturbaciones aleatorias

- 1) El modelo de regresión es **lineal en los parámetros**, aunque puede o no ser lineal en las variables.
- 2) **Valores de  $X$  independientes del término de error:** Lo que supone que la(s) variable(s)  $X$  y el término de error son independientes, esto es,  $Cov(X_i, u_i) = 0$ . Esto puede entenderse como que los valores adoptados por las  $X$ , se consideran fijos en muestras repetidas.
- 3) **La esperanza matemática de las perturbaciones  $u_i$  es igual a cero:** Dado el valor de  $X_i$ , la media o el valor esperado del término de perturbación aleatoria  $u_i$  es cero.
- 4) **Homoscedasticidad o varianza constante de  $u$ :** La varianza del término de error, o de la perturbación, es la misma sin importar el valor de  $X$ .
- 5) **No hay autocorrelación entre las perturbaciones:** Dados dos valores cualesquiera de  $X$ ,  $X_i$  y  $X_j$  ( $i \neq j$ ), la correlación entre dos  $u_i$  y  $u_j$  cualesquiera ( $i \neq j$ ) es cero.
- 6) **El número de observaciones  $N$  debe ser mayor que el número de parámetros por estimar.**
- 7) **Ausencia de multicolinealidad:** las variables explicativas del modelo son linealmente independientes lo que implica que el rango de la matriz  $X$  es  $k$ .

### 2.1.2. Propiedades de los estimadores MCO de los parámetros de posición

Dados los supuestos del modelo clásico de regresión lineal, la estimación de mínimos cuadrados de los parámetros de posición poseen algunas propiedades ideales u óptimas, las cuales están contenidas en el teorema de Gauss-Markov. Primero analizaremos las propiedades para muestras finitas y en segundo lugar las propiedades asintóticas.

#### 2.1.2.1 Propiedades para muestras finitas

- 1) **Linealidad**, es decir, los estimadores son una función lineal de la variable dependiente  $Y$ .
- 2) **Insesgadez**, es decir, el valor esperado del estimador es igual al valor del parámetro objeto de estimación.
- 3) **ELIO (Estimadores lineales insesgados óptimos)**, es decir, de los estimadores lineales e insesgados, además son los de menor varianza.



- 4) **Eficiencia**, es decir, de entre todos los estimadores insesgados, son los estimadores que tienen menor varianza.

#### 2.1.2.2. Propiedades asintóticas

- 1) **Insesgadez asintótica**. Dado que en muestras finitas ya cumplen esta propiedad, esto implica que asintóticamente también van a ser insesgados:

$$\lim_{N \rightarrow \infty} E(\hat{\beta}) = \beta$$

Es decir, la esperanza matemática del estimador, cuando el tamaño muestral tiende a infinito, es el verdadero valor.

- 2) **Consistencia**. Un estimador es consistente cuando converge a su valor verdadero conforme el tamaño muestral tiende a infinito. Las condiciones suficientes para que  $\hat{\beta}$  sea consistente son las siguientes:

a. Insesgadez asintótica.

b.  $\lim_{N \rightarrow \infty} V(\hat{\beta}) = 0$

- 3) **Eficiencia asintótica**. El estimador MCO posee la menor varianza asintótica de entre todos los estimadores consistentes del parámetro desconocido.

## 2.1. EL PROBLEMA DE LA HETEROCEDASTICIDAD. CAUSAS Y CONSECUENCIAS

Un supuesto importante del modelo clásico de regresión lineal es que la varianza de cada término de perturbación  $u_i$  condicionada a los valores seleccionados de las variables explicativas, es un valor constante igual a  $\sigma^2$ , es decir, las perturbaciones aleatorias son homocedasticas. Los problemas de heterocedasticidad se darán cuando no se cumpla éste supuesto de varianza constante.

### 2.2.1. Causas de la heterocedasticidad.

Existen diversas causas de la heterocedasticidad, esto es, razones por las cuales la dispersión del modelo deja de ser invariable. Las más frecuentes son las siguientes:

- 1) La presencia de datos atípicos, entendiendo por tales aquellas observaciones que son muy diferentes (o muy grandes o muy pequeñas) con relación a las demás observaciones de la muestra. De manera más

precisa, un dato atípico es una observación que proviene de una población distinta a la que genera las demás observaciones de la muestra. La inclusión o exclusión de una observación de este tipo, en especial si el tamaño de la muestra es pequeño, puede alterar sustancialmente los resultados del análisis de regresión.

- 2) Errores de especificación en el modelo. La omisión de variables relevantes o una incorrecta forma funcional elegida, constituyen ejemplos en los que se puede ver alterada la varianza.
- 3) El hecho habitual de que cuantos mayores son los valores de alguna de las variables del modelo cabe esperar que mayor será la dispersión absoluta del mismo.

El problema de heterocedasticidad es quizá más común en la información de corte transversal que en la de series de tiempo, por lo tanto, en el presente trabajo solamente vamos a referirnos de manera específica al primer tipo de datos.

### **2.2.2 Consecuencias de la heterocedasticidad.**

Las consecuencias de la presencia de heterocedasticidad en el contexto de un modelo lineal general se pueden resumir en los siguientes puntos:

- 1) Los estimadores MCO de los parámetros de posición seguirán siendo insesgados, pero a pesar de esto ya no serán ELIO ni eficientes.
- 2) En el caso de utilizar la expresión habitual de  $V(\hat{\beta})$  :

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

Obtendremos una mala aproximación de las varianzas y covarianzas de los estimadores de los parámetros de posición, porque la verdadera matriz de varianzas y covarianzas vendrá dada por la siguiente expresión:

$$V(\hat{\beta}) = \sigma^2 (X'X)^{-1} X' \Omega X (X'X)^{-1}$$

Siendo  $\sigma^2 \Omega$  la matriz de varianzas y covarianzas de las perturbaciones del modelo.

- 3) El estimador MCO de  $\sigma^2$  será sesgado.

- 4) Como consecuencia de que el estimador MCO de  $\sigma^2$  será sesgado, obtendremos unos estimadores sesgados de las varianzas y covarianzas de los estimadores MCO de los parámetros de posición.
- 5) La utilización de los estadísticos habituales de la  $t$  y la  $F$  para la estimación por intervalo, región de confianza y contrastes de hipótesis tan importantes como los de significatividad individual y conjunta respectivamente de los parámetros de posición, carecerá de validez debido a que ya no podrá definirse que estos estadísticos se distribuyen como una  $t$  de Student o una  $F$  de Snedecor.

Como podemos observar, las implicaciones del incumplimiento del supuesto de homocedasticidad son muy importantes. Por lo tanto, resulta fundamental detectar de manera precisa la presencia de heterocedasticidad para actuar en consecuencia y, tratar de que la especificación del modelo se ajuste lo máximo posible a los supuestos básicos del Modelo Clásico de regresión lineal.

## **2.3. DIAGNÓSTICO DE LA HETEROCEDASTICIDAD. CONTRASTES**

Aunque existen diversos contrastes en la literatura como la prueba de Park, la prueba de Glejser, el contraste de Godtfeld-Quandt, el test de Levene , el contraste de White o el contraste de Breusch-Pagan entre otros, solamente vamos a centrarnos en estos dos últimos: el contraste de White y el contraste de Breusch-Pagan. La razón por la cual vamos a centrarnos en estos dos contrastes, reside en que nos interesa comparar un contraste general –*contraste en el cual no se exige determinar a priori las variables que causan la heterocedasticidad*- como es el de White con uno específico- *contraste en el cual se exige determinar las variables que causan la heterocedasticidad*- como es el de Breusch-Pagan.

### **2.3.1. Contraste de White**

White (1980) derivó un contraste asintótico que no precisa especificar las variables que provocan la heterocedasticidad. Como ilustración de la idea básica, se considera el siguiente modelo lineal simple:

$$Y = \beta_1 + \beta_2 X_{2i} + u_i$$

Donde la hipótesis nula y alternativa son las siguientes:

$H_0$ : Homocedasticidad

$H_A$ : Heterocedasticidad

Para realizar la prueba de White se procede de la siguiente forma:

- 1) Se estima el modelo especificado anteriormente por MCO y se obtienen los residuos  $\hat{u}_i$ .
- 2) Con el cuadrado de los residuos de la regresión original, se efectúa una regresión (*auxiliar*) sobre las variables  $X$  originales, sobre sus valores al cuadrado y sobre el (los) producto(s) cruzado(s) de las variables explicativas. Debe de incluirse también un término constante. Obtenemos el  $R^2$  o coeficiente de determinación de la regresión (*auxiliar*).

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{2i}^2 + \varepsilon_i$$

La hipótesis nula y alternativa en términos de la regresión auxiliar quedarán formuladas de la siguiente forma:

$$H_0 : \alpha_2 = \alpha_3 = 0$$

$$H_A : \text{No } H_0$$

- 3) Bajo la hipótesis nula de que no hay heterocedasticidad (homocedasticidad), puede demostrarse que el tamaño de la muestra ( $N$ ) multiplicado por el  $R^2$  obtenido de la regresión auxiliar, asintóticamente sigue la distribución Chi-cuadrado con un numero de grados de libertad ( $p$ ) igual al número de variables explicativas (sin el termino constante) de la regresión auxiliar.

$$W = N \cdot R_a^2 \sim_{as} \chi^2(p)$$

- 4) Fijando un nivel de significación  $\varepsilon$ , la regla de actuación será la siguiente:

$W \leq \chi_{\varepsilon}^2 \Rightarrow$  Aceptación de la hipótesis nula de homocedasticidad.

$W > \chi_{\varepsilon}^2 \Rightarrow$  Rechazo de la hipótesis nula de homocedasticidad. (Aceptación de heterocedasticidad)

En el caso de que la homoscedasticidad resulte rechazada, no existe indicación alguna de la forma de la heterocedasticidad y por consiguiente como reespecificar el modelo para solucionar el problema.

### 2.3.2. Contraste de Breusch-Pagan

En los trabajos de Breusch y Pagan (1979), se construyó una prueba en la que la idea fundamental consiste en comprobar si se puede encontrar una variable o un conjunto de variables  $Z$  que sirvan para explicar la evolución de la varianza de las perturbaciones aleatorias

Para ilustrar esta prueba, se considera un modelo de regresión lineal de  $k$  variables:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

Se supone que la varianza del error  $\sigma_i^2$  se describe como

$$\sigma_i^2 = f(\alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi})$$

Es decir,  $\sigma_i^2$  es algún tipo de función de las variables  $Z$  no estocásticas; algunas de las  $X$  o todas ellas, del modelo especificado, pueden servir como  $Z$ . Específicamente se va a suponer que:

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi}$$

Por tanto, se está suponiendo  $\sigma_i^2$  es una función lineal de las variables  $Z$  (aunque cabe destacar que la forma funcional podría ser distinta a la lineal). Si  $\alpha_2 = \dots = \alpha_m = 0$ ,  $\sigma_i^2 = \alpha_1$  que es una constante. Por consiguiente para probar si  $\sigma_i^2$  es homocedástica, se puede contrastar la hipótesis de que  $\alpha_2 = \dots = \alpha_m = 0$ .

La hipótesis nula y alternativa serán las siguientes:

$$H_0 : \alpha_2 = \dots = \alpha_m = 0 ; \text{ Homocedasticidad}$$

$$H_A : \text{No } H_0 ; \text{ Heterocedasticidad}$$

El procedimiento de la prueba es el siguiente:

- 1) Estimar el modelo especificado por MCO y obtener los residuos.
- 2) Obtener el estimador máximo verosímil del parámetro de dispersión del modelo, es decir  $\tilde{\sigma}^2 = \sum \hat{u}_i^2 / N$
- 3) Construir la variable  $p_i$  definida como:  $p_i = \hat{u}_i^2 / \tilde{\sigma}^2$
- 4) Efectuar la regresión de la  $p_i$  así construida sobre las Z como

$$p_i = \alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi} + v_i$$

Donde  $v_i$  es el término de perturbación para esta regresión.

- 5) Calcular la suma explicada (SE) de la regresión construida y definir el estadístico:

$$BP = \frac{1}{2} SE$$

- 6) Bajo la hipótesis nula de homocedasticidad el estadístico sigue asintóticamente una distribución chi-cuadrado con (m-1) grados de libertad, donde m es el numero de parámetros de posición de la regresión auxiliar.

$$BP = \frac{1}{2} SE \sim_{as} \chi^2(m-1)$$

- 7) Fijado un nivel de significación  $\varepsilon$ , la regla de actuación será la siguiente:

$BP \leq \chi^2_{\varepsilon/2} \Rightarrow$  Aceptación de la hipótesis nula de homocedasticidad.

$BP > \chi^2_{\varepsilon/2} \Rightarrow$  Rechazo de la hipótesis nula de homocedasticidad. (Aceptación de heterocedasticidad)

El contraste de Breusch-Pagan, requiere conocer tanto la forma funcional como las variables que causan la heterocedasticidad. La ventaja del contraste de Breusch-Pagan frente al de White, es que se trata de un contraste constructivo, es decir, en caso de rechazar la hipótesis nula de homocedasticidad, el contraste de Breusch-Pagan sí que indica la variable o variables que están causando la heterocedasticidad .

### **3. ANÁLISIS DEL COMPORTAMIENTO DE LOS CONTRASTES**

#### **3.1. CONCEPTOS DE INTERÉS.**

Dado que nuestro objetivo va a ser el de decidir cuál de los dos contrastes considerados, White o Breusch-Pagan, detecta mejor la presencia o no de heterocedasticidad bajo distintos supuestos, resulta fundamental la utilización de dos conceptos muy importantes como son el tamaño y la potencia de un contraste.

##### **3.1.1. Tamaño y potencia de un contraste.**

Un contraste de hipótesis es un procedimiento estadístico que sirve para decidir si una hipótesis nula se rechaza o no. Puesto que no se sabe si la hipótesis a contrastar es cierta o no, y es por eso precisamente por lo que se contrasta, entonces se pueden cometer dos errores: rechazar la hipótesis cuando es cierta y aceptarla cuando es falsa. Un contraste será bueno si minimiza, en algún sentido, la posibilidad de cometer estos dos errores.

Podemos definir los conceptos de tamaño y potencia de un contraste de la siguiente manera:

**El tamaño de error tipo I de un contraste** es la probabilidad de rechazar la hipótesis nula siendo ésta cierta. Un buen contraste debería rechazar la hipótesis nula cuando ésta es cierta, el menor número de veces posible.

**La potencia de un contraste** es la probabilidad de rechazar la hipótesis nula siendo ésta falsa. Un buen contraste debería rechazar la hipótesis nula cuando ésta sea falsa el mayor número de veces posible.

En nuestro caso en concreto, lo que buscamos es que el tamaño del contraste sea el más próximo posible al tamaño nominal que hemos fijado a priori, y que la potencia de éste sea elevada.

#### **3.2. LA METODOLOGÍA DE ANÁLISIS: ESTUDIOS DE MONTE CARLO.**

Para llevar a cabo el análisis de los contrastes estadísticos, se van a realizar experimentos de Monte Carlo mediante el software econométrico Gretl. En primer lugar

se va a explicar en qué consisten los experimentos de Monte Carlo y se van a repasar las características de éstos mediante un ejemplo; Posteriormente se concretará la metodología a seguir en el presente trabajo.

### **3.2.1. Introducción a los estudios de Monte Carlo.**

Como se ha destacado anteriormente, en econometría, es de vital importancia la elección de las técnicas más adecuadas para la detección de los problemas y que la detección de estos problemas se realice correctamente.

Una forma de averiguar que una técnica de detección de los problemas como son los contrastes de hipótesis funciona correctamente, es mediante la observación del comportamiento del mismo en diferentes situaciones y bajo determinados supuestos. Esta evaluación del comportamiento de un contraste, requiere una gran cantidad de datos y la realización de cálculos matemáticos que pueden resultar muy costosos. En este contexto la simulación de Monte Carlo ofrece una alternativa al análisis matemático para realizar ésta evaluación y su comportamiento en muestras aleatorias.

El concepto básico de los experimentos de Monte Carlo consiste en realizar sucesivas repeticiones aleatorias de un determinado proceso cuya base está contenido en un modelo con unas condiciones establecidas previamente. El objetivo fundamental de la simulación de Monte Carlo, es la obtención de numerosos resultados de forma aleatoria en base a unos supuestos.

Podemos definir el estudio de Monte Carlo como *“un ejercicio de simulación diseñado para analizar las propiedades en muestras pequeñas de estadísticos alternativos”*. (Kennedy, 2003)

Con carácter general un experimento de Monte Carlo sigue la estrategia siguiente:

- a) Especificar un modelo “verdadero” o Proceso Generador de Datos (PGD).

En el marco del MLG, esto implicaría especificar la distribución del término de perturbación aleatoria, las variables explicativas, los coeficientes del modelo y el tamaño muestral.

- b) Generar un conjunto de datos usando éste PGD.
- c) Calcular el contraste estadístico objeto de análisis con ésta muestra generada artificialmente.



- d) Repetir las etapas b) y c) un elevado número de veces. Cada generación de un nuevo conjunto de datos se denomina réplica.
- e) Evaluar el resultado obtenido con el experimento planteado.

Con el objeto de ilustrar el procedimiento antes descrito para diseñar un experimento de Monte Carlo, se va a proceder mediante un ejemplo.

El primer paso es especificar el proceso generador de datos (PGD), como por ejemplo un modelo lineal simple:

$$Y_i = \beta_1 + \beta_2 X_{2i} + u_i$$

Donde  $Y_i$  es la variable dependiente,  $X_i$  la variable explicativa, y  $u_i$  la perturbación aleatoria del modelo.

Esta especificación requiere:

Primero.

- 1) Seleccionar los valores verdaderos de  $\beta_1$  y  $\beta_2$ .
- 2) Especificar el tamaño muestral y generar los valores de  $X$  para cada observación.
- 3) Especificar la distribución de probabilidad de  $u$  y generar sus valores a partir de tal distribución.

Segundo. Generar los valores de  $Y$  en, cada observación utilizando la expresión del modelo especificado inicialmente y los valores de  $\beta_1, \beta_2, X_{2i}$  y  $u$ .

Tercero. Con esta muestra generada artificialmente, se calcula el estadístico o estadísticos objeto de análisis.

Se repite este procedimiento un gran número de veces y se evalúan los resultados obtenidos.

### 3.3. OBJETIVO Y DISEÑO DEL ESTUDIO DE MONTE CARLO.

El diseño de los experimentos de Monte Carlo, se ha realizado con el objeto de analizar el comportamiento de los contrastes de heterocedasticidad de White y de Breusch-Pagan en términos de tamaño y potencia cuando el PGD corresponde a dos modelos distintos en los cuales se han planteado diferentes estructuras de la varianza del error tanto en muestras pequeñas como en muestras grandes.

- Se asumen dos posibles PGD's:

$$M1: Y_i = \beta_1 + \beta_2 X_{2i} + u_i$$

$$M2: Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

- Los tamaños muestrales considerados son  $N=50, 100$  y  $1000$ .
- Los valores de los coeficientes son:  $\beta_i = 1, \forall i$
- Las variables  $X$  se han generado como una distribución uniforme  $(0,10)$
- La perturbación del modelo se genera como una  $N(0, \sigma_i^2)$ .
- Dado nuestro objetivo, se deberán establecer diferentes supuestos sobre  $\sigma_i^2$ . En concreto:

a)  $\sigma_i^2 = 10 + v \cdot X_{2i} \Rightarrow$  SUPUESTO A-1

b)  $\sigma_i^2 = (10 + v \cdot X_{2i})^2 \Rightarrow$  SUPUESTO B-1

c)  $\sigma_i^2 = \exp(10 + v \cdot X_{2i}) \Rightarrow$  SUPUESTO C-1

- De forma específica para el modelo M2, se considerará también que ambas variables puedan ser las causantes de la heterocedasticidad. Por lo tanto, los supuestos específicos además de los indicados anteriormente, para M2 serán los siguientes:

d)  $\sigma_i^2 = 10 + v \cdot (X_{2i} + X_{3i}) \Rightarrow$  SUPUESTO A-2

e)  $\sigma_i^2 = (10 + v \cdot (X_{2i} + X_{3i}))^2 \Rightarrow$  SUPUESTO B-2

f)  $\sigma_i^2 = \exp(10 + v \cdot (X_{2i} + X_{3i})) \Rightarrow$  SUPUESTO C-2

- Se proponen cinco grados de heterocedasticidad asignando valores a  $v$  (*gravedad de la heterocedasticidad*) desde 1 hasta 5 y el valor  $v=0$  que implicaría homocedasticidad.

- El tamaño de error tipo I, se calcula como el porcentaje de casos en los que se rechaza la hipótesis nula de homocedasticidad siendo ésta cierta ( $v=0$ ).
- La potencia de los contrastes se calcula como el porcentaje de casos en los que se rechaza la hipótesis nula de homocedasticidad siendo ésta falsa ( $v=1, 2, 3, 4$  y  $5$ ).
- 
- Por último, las réplicas (R) realizadas para cada experimento serán 1000.

El soporte en el cual se han desarrollado todos los pasos descritos en el diseño del experimento ha sido la consola de ejecución de comandos que incorpora el software econométrico Gretl. Por ello a través del uso de los “scripts” (guiones de instrucciones) como lenguaje de programación se han podido llevar a cabo los experimentos.

*[En el anexo se incorporan los guiones de instrucciones utilizados en los experimentos]*

Finalmente para que hubiera homogeneidad en la aleatoriedad de los datos para cada experimento, se ha utilizado como semilla de números aleatorios o valor inicial, el número 123.

#### **4. RESULTADOS DE LOS EXPERIMENTOS DE MONTE CARLO.**

Los resultados de los experimentos de Monte Carlo están estructurados en primer lugar en función del PGD seleccionado y en segundo lugar, en función del supuesto sobre  $\sigma^2_i$ , es decir, el patrón de heterocedasticidad considerado en cada caso.

Los resultados se van a presentar en dos formatos: mediante una tabla en la que se podrá observar los resultados de los cálculos de la potencia y tamaño del error tipo I de cada uno de los contrastes para cada tamaño muestral y para cada nivel de heterocedasticidad. Seguidamente, con el objeto de realizar el análisis de una forma más ilustrativa, se van a representar las curvas de potencia de los contrastes, en función del valor de  $v$  (*grado de heterocedasticidad*), y para diferentes tamaños muestrales.

#### 4.1. EXPERIMENTO DE MONTE CARLO PARA EL MODELO LINEAL SIMPLE (M1). $Y_i = \beta_1 + \beta_2 X_{2i} + u_i$

##### 4.1.1. Supuesto A-1. $\sigma^2_i = 10 + v \cdot X_{2i}$

La tabla 1.1 muestra el tamaño y la potencia de los contrastes de Breusch-Pagan (BP) y White (W) para tres tamaños muestrales diferentes (N= 50, 100 y 1000) y para distintos valores de  $v$  (*grado de heterocedasticidad*), suponiendo que el PGD es el modelo M1 y que la estructura de la varianza de las perturbaciones hace referencia al supuesto de que  $\sigma^2_i = 10 + v \cdot X_{2i}$ .

**Tabla 1.1**

| SUPUESTO A-1 |    | TAMAÑO | POTENCIA |       |       |       |       |
|--------------|----|--------|----------|-------|-------|-------|-------|
|              |    | $v=0$  | $v=1$    | $v=2$ | $v=3$ | $v=4$ | $v=5$ |
| N=50         | BP | 0,054  | 0,122    | 0,304 | 0,271 | 0,438 | 0,534 |
|              | W  | 0,048  | 0,087    | 0,178 | 0,159 | 0,262 | 0,345 |
| N=100        | BP | 0,059  | 0,24     | 0,503 | 0,606 | 0,803 | 0,794 |
|              | W  | 0,041  | 0,189    | 0,35  | 0,416 | 0,655 | 0,688 |
| N=1000       | BP | 0,052  | 0,987    | 1     | 1     | 1     | 1     |
|              | W  | 0,058  | 0,976    | 1     | 1     | 1     | 1     |

Fuente: Elaboración propia.

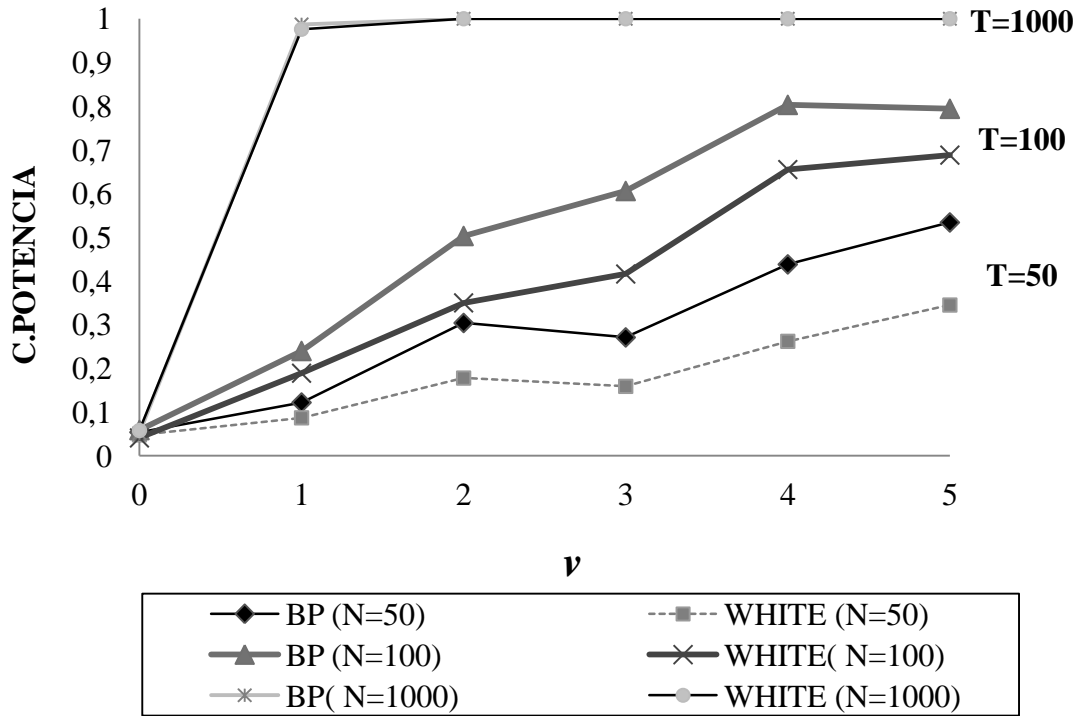
Cabe destacar en primer lugar que para ambos contrastes, el tamaño de error tipo I es aproximadamente el mismo para los tres tamaños muestrales y muy próximo al nivel de significación del 5% seleccionado a priori. No obstante, Breusch-Pagan sobreestima el tamaño nominal mientras que White lo subestima, salvo para T=1000 donde el comportamiento es el contrario. En segundo lugar, la potencia de ambos contrastes, a medida que el tamaño muestral va aumentando, es mayor.

La figura 1.1 muestra la potencia de los contrastes de Breusch-Pagan y White para distintos tamaños muestrales y para cada valor de  $v$  (*grado de heterocedasticidad*).

A la vista de la figura 1.1, para un mismo valor de  $v$ , para tamaños muestrales de N=50 y N=100 la potencia del contraste de Breusch-Pagan es en todo momento superior al de

White y para un tamaño de  $N=1000$  la potencia es máxima e igual 1 para ambos contrastes, es decir, de las 1000 replicas realizadas, ambos contrastes detectan la presencia de heterocedasticidad el 100% de las veces.

**Figura 1.1**



Fuente: Elaboración propia.

#### 4.1.2. Supuesto B-1 $\sigma_i^2 = (10 + \nu \cdot X_{2i})^2$

En la tabla 1.2 se mantiene la misma estructura que la tabla anterior pero bajo el supuesto de que la varianza de las perturbaciones aleatorias es  $\sigma_i^2 = (10 + \nu \cdot X_{2i})^2$ . En ella se puede observar las mismas pautas de comportamiento que en caso anterior. Por otro lado, en comparación con el supuesto anterior la potencia de ambos contrastes es superior para la mayoría de los casos.

#### 4.1.3. Supuesto C-1 $\sigma_i^2 = \exp(10 + \nu \cdot X_{2i})$

La tabla 1.3 muestra los resultados para el supuesto de que la varianza de las perturbaciones aleatorias sigue la forma  $\sigma_i^2 = \exp(10 + \nu \cdot X_{2i})$ .

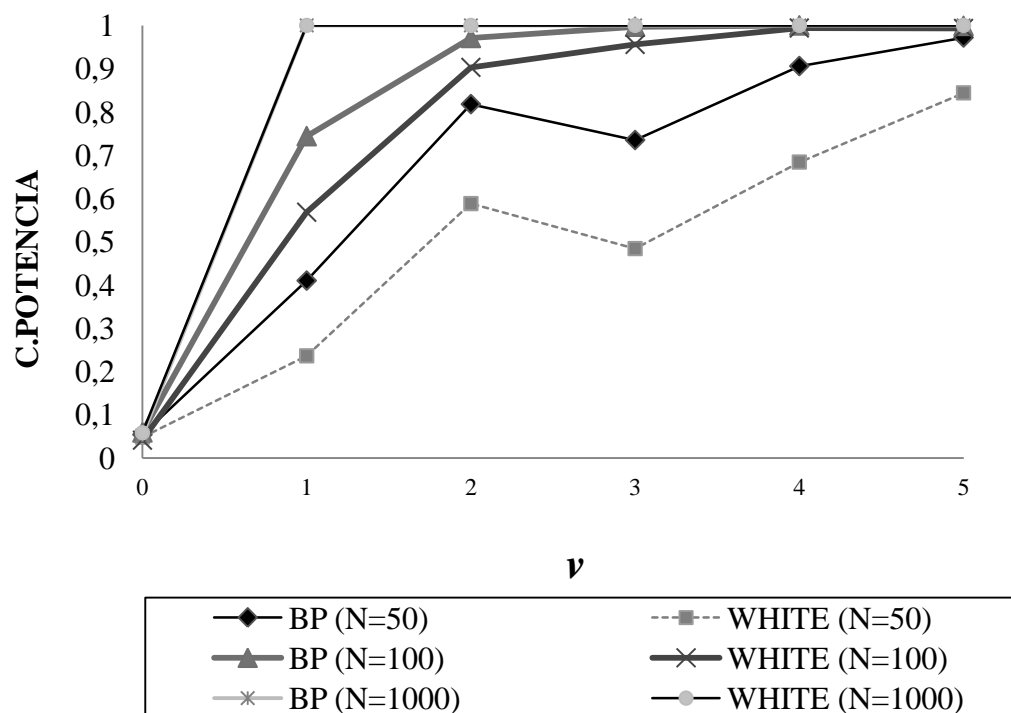
Fijándonos en los resultados de la tabla 1.3, se puede observar que para éste supuesto, la potencia de ambos contrastes de heterocedasticidad es máxima para todo tamaño muestral.

**Tabla 1.2**

| SUPUESTO B-1 |    | TAMAÑO  | POTENCIA |         |         |         |         |
|--------------|----|---------|----------|---------|---------|---------|---------|
|              |    | $\nu=0$ | $\nu=1$  | $\nu=2$ | $\nu=3$ | $\nu=4$ | $\nu=5$ |
| N=50         | BP | 0,054   | 0,41     | 0,818   | 0,735   | 0,906   | 0,972   |
|              | W  | 0,048   | 0,236    | 0,588   | 0,484   | 0,684   | 0,844   |
| N=100        | BP | 0,059   | 0,744    | 0,971   | 0,997   | 1       | 0,99    |
|              | W  | 0,041   | 0,568    | 0,904   | 0,956   | 0,994   | 0,993   |
| N=1000       | BP | 0,052   | 1        | 1       | 1       | 1       | 1       |
|              | W  | 0,058   | 1        | 1       | 1       | 1       | 1       |

Fuente: Elaboración propia.

**Figura 1.2**



Fuente: Elaboración propia.

**Tabla 1.3**

| SUPUESTO C-1 |    | TAMAÑO | POTENCIA |       |       |       |       |
|--------------|----|--------|----------|-------|-------|-------|-------|
|              |    | $v=0$  | $v=1$    | $v=2$ | $v=3$ | $v=4$ | $v=5$ |
| N=50         | BP | 0,054  | 1        | 1     | 1     | 1     | 1     |
|              | W  | 0,048  | 0,998    | 0,984 | 0,99  | 0,99  | 0,99  |
| N=100        | BP | 0,059  | 1        | 1     | 1     | 1     | 1     |
|              | W  | 0,041  | 1        | 1     | 1     | 1     | 1     |
| N=1000       | BP | 0,052  | 1        | 1     | 1     | 1     | 1     |
|              | W  | 0,058  | 1        | 1     | 1     | 1     | 1     |

Fuente: Elaboración propia.

Los resultados más destacados de estos experimentos se pueden resumir en los siguientes puntos:

- 1.- El tamaño de error tipo I es similar al nominal para ambos contrastes.
- 2.- La potencia del contraste de Breusch-Pagan es superior en general, al de White, independientemente del grado de heterocedasticidad asumido y el tamaño muestral.
- 3.- Se verifica que la potencia de los contrastes se incrementa conforme el patrón de heterocedasticidad es más fuerte, lo cual queda reflejado en la forma funcional asumida para  $\sigma_i^2$ .

#### 4.1. EXPERIMENTO DE MONTE CARLO PARA EL MODELO LINEAL GENERAL (M2). $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} u_i$

Tomando como PGD un modelo lineal con dos variables, consideramos distintos supuestos sobre la posible variable o variables causantes de la heterocedasticidad, es decir: una sola variable ( $X_{2i}$  o  $X_{3i}$ ) o conjuntamente  $X_{2i}, X_{3i}$ . En este caso presentaremos los resultados de tres posibles contrastes de Breusch-Pagan dependiendo de la variable que consideremos causante de la heterocedasticidad: BP1 ( $X_{2i}$ ), BP2

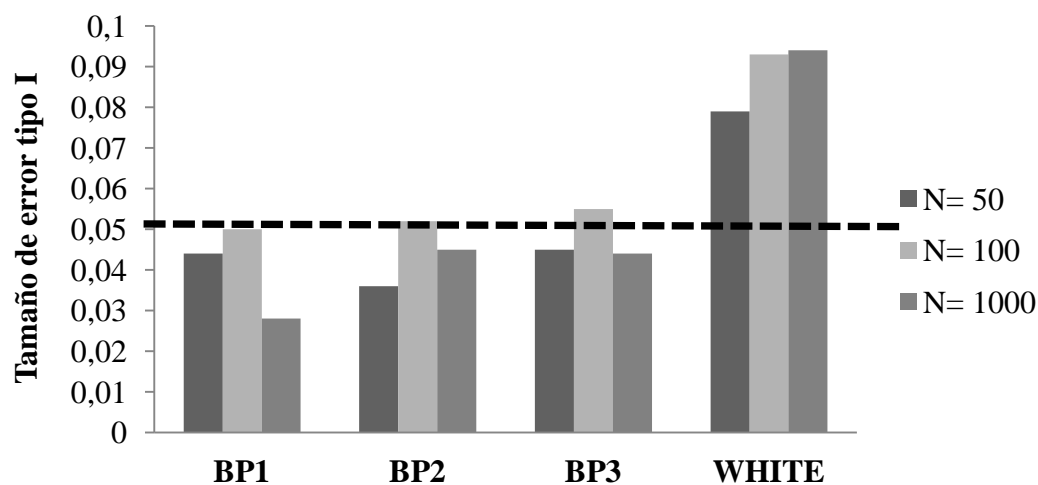
$(X_{3i})$  y  $BP3(X_{2i}, X_{3i})$  [entre paréntesis las variables causantes de la heterocedasticidad]

#### 4.2.1. Supuesto A-1 $\sigma^2_i = 10 + v \cdot X_{2i}$

La figura 2.1, indica el tamaño de error tipo I de cada contraste de heterocedasticidad para cada tamaño muestral. De la misma manera que en el modelo anterior, los resultados en términos de tamaño, son comunes para todos los patrones de heterocedasticidad considerados.

Como puede observarse en la figura 2.1, el contraste de White en términos de error tipo I conduce a sobreeachazar la hipótesis nula cuando es cierta. Por el contrario, el tamaño de los contrastes de Breusch-Pagan de forma general, se aproxima más al tamaño nominal del 5%.

**Figura 2.1**



Fuente: Elaboración propia

Nótese que para un tamaño muestral de  $N=1000$ , el contraste BP1 presenta un tamaño más bajo del 5%. Esto se explica por el componente aleatorio de los datos simulados lo que no significa que para muestras asintóticas no tienda al valor nominal. Para comprobar que éste dato es fruto de la aleatoriedad, se han realizado varios experimentos para tamaños muestrales superiores a  $N = 1000$  y se ha observado que asintóticamente se aproxima al 5%.



En términos de potencia, para el experimento en el que se supone que la heterocedasticidad es causada por la variable  $X_{2i}$ , como se puede observar en la tabla 2.1, el contraste BP1, que tiene en cuenta que la variable  $X_{2i}$  causa la heterocedasticidad, presenta mejores resultados de forma general que el resto.

**Tabla 2.1**

| SUPUESTO A-1  |     | TAMAÑO | POTENCIA |       |       |       |       |
|---------------|-----|--------|----------|-------|-------|-------|-------|
|               |     | $v=0$  | $v=1$    | $v=2$ | $v=3$ | $v=4$ | $v=5$ |
| <b>N=50</b>   | BP1 | 0,044  | 0,128    | 0,201 | 0,381 | 0,406 | 0,605 |
|               | BP2 | 0,036  | 0,042    | 0,068 | 0,042 | 0,07  | 0,085 |
|               | BP3 | 0,045  | 0,087    | 0,17  | 0,3   | 0,318 | 0,482 |
|               | W   | 0,079  | 0,116    | 0,163 | 0,253 | 0,233 | 0,35  |
| <b>N=100</b>  | BP1 | 0,05   | 0,237    | 0,526 | 0,654 | 0,824 | 0,865 |
|               | BP2 | 0,052  | 0,051    | 0,055 | 0,085 | 0,082 | 0,074 |
|               | BP3 | 0,055  | 0,186    | 0,421 | 0,556 | 0,716 | 0,78  |
|               | W   | 0,093  | 0,213    | 0,324 | 0,45  | 0,547 | 0,65  |
| <b>N=1000</b> | BP1 | 0,028  | 0,992    | 1     | 1     | 1     | 1     |
|               | BP2 | 0,045  | 0,056    | 0,071 | 0,075 | 0,069 | 0,079 |
|               | BP3 | 0,044  | 0,981    | 1     | 1     | 1     | 1     |
|               | W   | 0,094  | 0,967    | 1     | 1     | 1     | 1     |

Fuente: Elaboración propia.

#### 4.2.2. Supuesto A-2 $\sigma_i^2 = (10 + v \cdot (X_{2i} + X_{3i}))$

Bajo el supuesto de que la heterocedasticidad está causada conjuntamente por las variables  $X_{2i}$  y  $X_{3i}$  se puede observar en la tabla 2.2 que, para muestras pequeñas, tanto

el contraste de Breusch-Pagan, que tiene en cuenta que ambas variables causan la heterocedasticidad (BP3), como el contraste de White, se comportan bastante bien, lo cual es coherente con el hecho de que ambos contrastes consideren que son las dos variables explicativas del modelo las causantes del problema.

**Tabla 2.2**

| SUPUESTO A-2 |     | TAMAÑO  | POTENCIA |         |         |         |         |
|--------------|-----|---------|----------|---------|---------|---------|---------|
|              |     | $\nu=0$ | $\nu=1$  | $\nu=2$ | $\nu=3$ | $\nu=4$ | $\nu=5$ |
| N=50         | BP1 | 0,044   | 0,104    | 0,094   | 0,176   | 0,13    | 0,208   |
|              | BP2 | 0,036   | 0,09     | 0,104   | 0,221   | 0,201   | 0,153   |
|              | BP3 | 0,045   | 0,092    | 0,128   | 0,264   | 0,248   | 0,251   |
|              | W   | 0,079   | 0,125    | 0,132   | 0,226   | 0,208   | 0,209   |
| N=100        | BP1 | 0,05    | 0,135    | 0,268   | 0,291   | 0,481   | 0,28    |
|              | BP2 | 0,052   | 0,139    | 0,263   | 0,302   | 0,434   | 0,285   |
|              | BP3 | 0,055   | 0,17     | 0,369   | 0,434   | 0,588   | 0,452   |
|              | W   | 0,093   | 0,209    | 0,304   | 0,376   | 0,462   | 0,348   |
| N=1000       | BP1 | 0,028   | 0,924    | 0,994   | 0,996   | 1       | 0,99    |
|              | BP2 | 0,045   | 0,891    | 0,984   | 0,999   | 0,998   | 1       |
|              | BP3 | 0,044   | 0,988    | 1       | 1       | 1       | 1       |
|              | W   | 0,094   | 0,979    | 1       | 1       | 1       | 1       |

Fuente: Elaboración propia.

Como se puede observar en la tabla 2.2, cabe destacar que, para un nivel bajo de heterocedasticidad ( $\nu=1$ ), el contraste de White presenta una potencia superior, mientras que para niveles superiores el contraste BP3 es el de mayor potencia.

Por otro lado, también cabe destacar que los contrastes BP1 y BP2 presentan un comportamiento semejante, puesto que tanto uno como otro reflejan que una de las

variables al menos, está causando la heterocedasticidad a pesar de que la estén causando ambas.

Finalmente, para un tamaño muestral grande ( $N=1000$ ), la potencia de los contrastes BP3 y White es bastante similar, aunque cabe resaltar que de nuevo para niveles bajos de heterocedasticidad, el contraste BP3, se comporta un poco mejor que el de White. Por otro lado, la potencia de los contrastes BP1 y BP2 mejora sustancialmente con respecto a muestras pequeñas.

#### 4.2.3. Supuesto B-1 $\sigma_i^2 = (10 + v \cdot X_{2i})^2$

Suponiendo que la varianza de las perturbaciones aleatorias tiene una forma funcional cuadrática en función de  $X_{2i}$  de acuerdo con los resultados de la tabla 2.3 en primer lugar, para muestras pequeñas y para niveles de heterocedasticidad bajos ( $v=1,2$ ) podemos observar que el contraste que mejores resultados de potencia presenta es el contraste BP1 dado que tiene en cuenta que la heterocedasticidad está causada por la variable  $X_{2i}$ . No obstante también cabe destacar el buen funcionamiento del contraste BP3. Ambos contrastes (BP1 y BP3) presentan mejores resultados que el contraste de White y la divergencia de potencia con éstos es mayor a medida que el valor de  $v$  aumenta. A medida que el tamaño muestral aumenta, se puede observar que estos tres contrastes, alcanzan una potencia máxima.

Por otro lado, el comportamiento del contraste BP2, que supone que la heterocedasticidad está causada por la variable  $X_{3i}$ , es bastante malo tanto para muestras pequeñas como para muestras grandes.

#### 4.2.4. Supuesto B-2 $\sigma_i^2 = (10 + v \cdot (X_{2i} + X_{3i}))^2$

En el caso de que la de que la varianza de las perturbaciones aleatorias sigue la forma  $\sigma_i^2 = (10 + v \cdot (X_{2i} + X_{3i}))^2$  el comportamiento en general de los contrastes en términos de potencia es bastante irregular para muestras pequeñas como puede observarse en la tabla 2.4.

La potencia de los contrastes, a medida que el tamaño muestral aumenta y en concreto para  $N=100$ , es bastante aceptable, aunque continúa presentando un comportamiento irregular a partir de grados de heterocedasticidad superiores a  $\nu=2$ .

Cabe destacar, que el contraste de Breusch-Pagan BP3 se comporta mejor que el contraste de White tanto para muestras pequeñas como para muestras grandes, en términos de potencia.

También se puede observar que la potencia tanto del contraste BP1 como de BP2, que suponen que la heterocedasticidad la causa  $X_{2i}$  y  $X_{3i}$  respectivamente, se comportan de manera similar.

**Tabla 2.3**

| SUPUESTO B-1  |     | TAMAÑO  | POTENCIA |         |         |         |         |
|---------------|-----|---------|----------|---------|---------|---------|---------|
|               |     | $\nu=0$ | $\nu=1$  | $\nu=2$ | $\nu=3$ | $\nu=4$ | $\nu=5$ |
| <b>N=50</b>   | BP1 | 0,044   | 0,407    | 0,591   | 0,877   | 0,895   | 0,981   |
|               | BP2 | 0,036   | 0,053    | 0,082   | 0,059   | 0,114   | 0,133   |
|               | BP3 | 0,045   | 0,32     | 0,509   | 0,794   | 0,825   | 0,964   |
|               | W   | 0,079   | 0,229    | 0,406   | 0,615   | 0,572   | 0,727   |
| <b>N=100</b>  | BP1 | 0,05    | 0,677    | 0,968   | 0,992   | 1       | 1       |
|               | BP2 | 0,052   | 0,06     | 0,084   | 0,139   | 0,167   | 0,143   |
|               | BP3 | 0,055   | 0,566    | 0,952   | 0,981   | 1       | 1       |
|               | W   | 0,093   | 0,5      | 0,833   | 0,942   | 0,967   | 0,989   |
| <b>N=1000</b> | BP1 | 0,028   | 1        | 1       | 1       | 1       | 1       |
|               | BP2 | 0,045   | 0,064    | 0,096   | 0,159   | 0,116   | 0,124   |
|               | BP3 | 0,044   | 1        | 1       | 1       | 1       | 1       |
|               | W   | 0,044   | 0,407    | 0,591   | 0,877   | 0,895   | 0,981   |

Fuente: Elaboración propia.

**Tabla 2.4**

| SUPUESTO B-2  |     | TAMAÑO | POTENCIA |       |       |       |       |
|---------------|-----|--------|----------|-------|-------|-------|-------|
|               |     | $v=0$  | $v=1$    | $v=2$ | $v=3$ | $v=4$ | $v=5$ |
| <b>N=50</b>   | BP1 | 0,044  | 0,303    | 0,275 | 0,429 | 0,321 | 0,617 |
|               | BP2 | 0,036  | 0,237    | 0,267 | 0,621 | 0,502 | 0,444 |
|               | BP3 | 0,045  | 0,344    | 0,388 | 0,753 | 0,663 | 0,717 |
|               | W   | 0,079  | 0,258    | 0,308 | 0,529 | 0,478 | 0,518 |
| <b>N=100</b>  | BP1 | 0,05   | 0,388    | 0,726 | 0,746 | 0,924 | 0,769 |
|               | BP2 | 0,052  | 0,389    | 0,684 | 0,73  | 0,885 | 0,698 |
|               | BP3 | 0,055  | 0,574    | 0,895 | 0,929 | 0,976 | 0,949 |
|               | W   | 0,093  | 0,507    | 0,773 | 0,846 | 0,937 | 0,841 |
| <b>N=1000</b> | BP1 | 0,028  | 1        | 1     | 1     | 1     | 1     |
|               | BP2 | 0,045  | 1        | 1     | 1     | 1     | 1     |
|               | BP3 | 0,044  | 1        | 1     | 1     | 1     | 1     |
|               | W   | 0,094  | 1        | 1     | 1     | 1     | 1     |

Fuente: Elaboración propia.

#### 4.2.5. Supuesto C-1 $\sigma_i^2 = \exp(10 + v \cdot X_{2i})$

Suponiendo que la varianza de las perturbaciones aleatorias siga una función exponencial que depende de  $X_{2i}$ , podemos observar en la tabla 2.5 que, para todo tamaño muestral, los contrastes de Breusch-Pagan BP1 y BP3 presentan un comportamiento excelente siendo la potencia de éstos máxima. Por otro lado, a pesar de que la potencia del contraste de White para muestras pequeñas presenta una cierta irregularidad, conforme el tamaño muestral aumenta, ésta potencia se iguala a la de los

contrastes BP1 y BP3. Por último cabe destacar que el contraste BP3, que supone que la variable que causa la heterocedasticidad es  $X_{3i}$  no presenta una potencia adecuada.

**Tabla 2.5**

| SUPUESTO C-1  |     | TAMAÑO  | POTENCIA |         |         |         |         |
|---------------|-----|---------|----------|---------|---------|---------|---------|
|               |     | $\nu=0$ | $\nu=1$  | $\nu=2$ | $\nu=3$ | $\nu=4$ | $\nu=5$ |
| <b>N=50</b>   | BP1 | 0,044   | 1        | 1       | 1       | 1       | 1       |
|               | BP2 | 0,036   | 0,413    | 0,53    | 0,768   | 0,997   | 0,181   |
|               | BP3 | 0,045   | 1        | 1       | 1       | 1       | 1       |
|               | W   | 0,079   | 0,974    | 0,992   | 1       | 0,998   | 0,895   |
| <b>N=100</b>  | BP1 | 0,05    | 1        | 1       | 1       | 1       | 1       |
|               | BP2 | 0,052   | 0,385    | 0,528   | 0,544   | 0,727   | 0,649   |
|               | BP3 | 0,055   | 1        | 1       | 1       | 1       | 1       |
|               | W   | 0,093   | 1        | 1       | 1       | 0,996   | 1       |
| <b>N=1000</b> | BP1 | 0,028   | 1        | 1       | 1       | 1       | 1       |
|               | BP2 | 0,045   | 0,715    | 0,505   | 0,689   | 0,618   | 0,72    |
|               | BP3 | 0,044   | 1        | 1       | 1       | 1       | 1       |
|               | W   | 0,094   | 1        | 1       | 1       | 1       | 1       |

Fuente: Elaboración propia.

#### 4.2.6 Supuesto C-2 $\sigma_i^2 = \exp(10 + \nu \cdot (X_{2i} + X_{3i}))$

Finalmente, de acuerdo con que la estructura de la varianza de las perturbaciones aleatorias sigue la forma de una exponencial pero que en este caso depende de las variables  $X_{2i}$  y  $X_{3i}$ , como se puede observar en la tabla 2.6, la potencia de los contrastes de Breusch-Pagan y de White es máxima para todo tamaño muestral y para todos los grados de heterocedasticidad.

**Tabla 2.6**

| SUPUESTO A-1  |     | TAMAÑO | POTENCIA |       |       |       |       |
|---------------|-----|--------|----------|-------|-------|-------|-------|
|               |     | $v=0$  | $v=1$    | $v=2$ | $v=3$ | $v=4$ | $v=5$ |
| <b>N=50</b>   | BP1 | 0,044  | 1        | 1     | 1     | 1     | 1     |
|               | BP2 | 0,036  | 1        | 1     | 1     | 1     | 1     |
|               | BP3 | 0,045  | 1        | 1     | 1     | 1     | 1     |
|               | W   | 0,079  | 0,986    | 0,99  | 1     | 0,898 | 0,988 |
| <b>N=100</b>  | BP1 | 0,05   | 1        | 1     | 1     | 1     | 1     |
|               | BP2 | 0,052  | 1        | 1     | 1     | 1     | 1     |
|               | BP3 | 0,055  | 1        | 1     | 1     | 1     | 1     |
|               | W   | 0,093  | 1        | 1     | 1     | 1     | 1     |
| <b>N=1000</b> | BP1 | 0,028  | 1        | 1     | 1     | 1     | 1     |
|               | BP2 | 0,045  | 1        | 1     | 1     | 1     | 1     |
|               | BP3 | 0,044  | 1        | 1     | 1     | 1     | 1     |
|               | W   | 0,094  | 1        | 1     | 1     | 1     | 1     |

Fuente: Elaboración propia

Los resultados más destacados de éste segundo experimento, se pueden resumir en los siguientes puntos:

- 1.- Los contrastes de Breusch-Pagan (BP1, BP2 y BP3), presentan un tamaño de error tipo I, más próximo al tamaño nominal del 5% para cualquier patrón de heterocedasticidad y para cualquier tamaño muestral, que el contraste de White.
- 2.- Los diferentes contrastes de Breusch-Pagan se comportan adecuadamente, en términos de potencia, siendo capaces de determinar la variable causante de la heterocedasticidad.

3.- De forma general el contraste de Breusch-Pagan BP3 presenta mejores resultados de potencia que el de White independientemente de la forma funcional asumida y el grado de heterocedasticidad, es decir, el contraste BP3 rechaza un mayor número de veces la hipótesis nula cuando esta es falsa.

## 5. CONCLUSIONES

Como se indicaba al principio de éste trabajo, el objetivo del mismo era analizar en términos de tamaño de error tipo I y potencia, cuál de los dos contrastes de heterocedasticidad presentaba mejores resultados de acuerdo con cada supuesto realizado. Para ello mediante experimentos de Monte Carlo, se han generado dos modelos- *un modelo lineal simple y un modelo lineal general con dos variables*-, y para cada modelo, se han hecho distintos supuestos sobre la varianza de las perturbaciones aleatoria con diferentes grados de heterocedasticidad y para distintos tamaños muestrales. Para analizar el comportamiento de los contrastes se ha calculado la potencia y el tamaño de error tipo I de cada contraste en todos los casos.

A continuación se va a exponer las conclusiones generales que se pueden extraer de los resultados de los experimentos llevados a cabo:

- 1.- De forma general cuanto mayor es el grado de heterocedasticidad mayor es la potencia de los contrastes. De la misma manera se observa que, cuanto mayor es el tamaño muestral, mayor es la potencia de los contrastes.
- 2.- Todos los contrastes tienen un nivel de significación próximo al nominal aunque sí que se observa que Breusch-Pagan tiende a rechazar menos de lo que debería y el de White a sobre rechazar.
- 3.- El contraste de Breusch-Pagan BP3, que podría entenderse como un contraste general puesto que tiene en cuenta que todas las variables causan la heterocedasticidad, en comparación con el contraste de White presenta mejores resultados en términos de potencia y en términos de tamaño.
- 4.- Cuando se plantea un contraste de Breusch-Pagan en el que se indica que la variable que causa la heterocedasticidad es distinta a la que realmente la causa, los resultados en términos de potencia son bastante pésimos, debido a que el contraste tiende a no rechazar la hipótesis nula cuando ésta es falsa.



Finalmente, para concluir se pueden hacer las siguientes valoraciones:

- 1.- En el caso de un modelo lineal simple independientemente de la forma funcional de la varianza de las perturbaciones aleatorias, se obtendrán mejores resultados con el contraste de Breusch-Pagan.
- 2.- Si nos encontramos con un modelo lineal general, independientemente de la forma y la gravedad de la heterocedasticidad, en caso de que no se tengan claras las variables que causan la heterocedasticidad, el contraste de Breusch-Pagan, que considera que todos los regresores del modelo causan el problema (BP3) nos dará mejores resultados en términos de tamaño y potencia, y para cualquier tamaño muestral, que el contraste de White.
- 3.- Si la variable causante del problema es distinta a la supuesta en el contraste de Breusch-Pagan, este tendrá una potencia muy baja.

## **6. BIBLIOGRAFÍA**

Este apartado se divide en dos partes; en primer lugar se indican las referencias citadas en el presente trabajo, y en segundo lugar se cita otra bibliografía utilizada para el desarrollo del trabajo.

### **6.1.REFERENCIAS**

BREUSCH, T. S., & PAGAN, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5), 1287.

KENNEDY, P. (2003). *A guide to econometrics*. (5ª ed.) MIT Press

WHITE, H. (1980). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity". *Econometrica*, Vol.48, N°4, páginas 817-838.

### **6.2.OTRA BIBLIOGRAFIA**

ADKINS, L. C. (2011). Using gretl for Monte Carlo experiments. *Journal of Applied Econometrics*, 26(5), 880.

AZNAR GRASA, A. (2009). *Curso de econometría* (Ed. rev 2009 ed.). Copy Center. Zaragoza.

DAVIDSON, R., & MACKINNON, J.G.(1993). *Estimation and inference in econometrics*. Oxford University Press, New York.

DIOS PALOMARES, R., & RODRÍGUEZ FONSECA, C. (1999). Análisis comparativo de estimadores pretest de heterocedasticidad en modelos econométricos: Un estudio Monte Carlo. *Qüestiió*, Vol 23, nº3, páginas 437-464.

DOUGHERTY, C. (1992). *Introduction to econometrics*. Oxford.

GREENE, W. H. (2008). *Análisis econométrico* (3ª ed.) Prentice-Hall, Madrid.

GUJARATI, D. N. (2010). *Econometría* (5ª ed.). México etc.: McGraw-Hill.

MOONEY, C. Z. (1997). “Monte Carlo simulation”. A Sage University Paper series on Quantitative Applications in the Social Sciences Nº 116. Newbury Park, CA