A Comparison of Powers of Some Tests for Heteroscedasticity in Non-Linear Models

I. F. Oyenuga[1]; B. A. Oyejola[2]

[1]Department of Mathematics and Statistics,
The Polytechnic, Ibadan,
Ibadan, Nigeria.
e-mail: iyaboadebayo2001@yahoo.com

[2]Department of Statistics,
University of Ilorin,
Ilorin, Nigeria.
e-mail: boyejola2003@yahoo.com

***Abstract*— This paper compares the powers of Glejser, Park and White tests in detecting heteroscedasticity in two non-linear models: Constant Elasticity of Substitution (CES) and Exponential production function. Exponential production function was transformed to an intrinsically linear model through the natural logarithms while CES production function model was transformed using Kmenta (1967) linearization approach. The sample sizes for the simulation were grouped into three categories: 10 and 30, 50 and 100, 150 and 200 for small, medium and large sample sizes respectively with 10,000 replications. The levels of heteroscedasticity introduced were 0.1, 0.5 and 0.9 for mild, moderate and severe heteroscedasticity respectively. The level of significance used was 0.05. Results showed that at all levels of heteroscedasticity for the models considered in this paper, Glejser test is the most powerful test for all the sample sizes.**

Keywords: *Heteroscedasticity, Intrinsically linear model, Production function, Power of the test.*

I. Introduction

It is known that when the assumptions of the classical linear regression model (CLRM) are valid, ordinary least squares (OLS) provides efficient and unbiased estimates of parameters. One of the assumptions of CLRM is that the disturbances u_i in the population regression function (PRF) are homoscedastic; that is, they all have the same variance, σ^2. If the variance of u_i is σ_i^2, $i = 1, \ldots , n$, it indicates that it varies from observation to observation, then there exists heteroscedasticity, or unequal, or non-constant variance. If the errors are heteroscedastic, the OLS estimator remains unbiased but becomes inefficient [1,2].

Generally, non-linear models are used in the estimation of production and demand functions. Even a simple Cobb-Douglass production function cannot be transformed into linearity if the error term is added rather than multiplied [3]. Econometric modeling demands the incorporation of stochastic term (an error term) as well as the specification of its distribution. Violation of homoscedasticity, one of the assumptions of classical linear regression models, leads to bias in estimating regression parameters. Hence, there is need to test for the presence of heteroscedasticity in datasets before modeling.

Examination of the asymptotic behavior of the power of heteroscedasticity tests in two non-linear models for small to large sample sizes will be investigated.[4] Koichi, in his paper comparing the Wald, LR and LM tests for heteroscedasticity in a linear regression model, compare the power of the Wald, Likelihood ratio and Lagrangian multiplier model through the asymptotic expansion of the non-null distribution for the three tests, as a result, he discovered that the powers of the three tests depend on both the nature of the explanatory variables and the direction of the alternative hypothesis and that no one statistic is uniformly superior to the others. [5] in their study titled " estimation parameters of linear econometric model and the power of test in the presence of heteroscedasticity", where they introduced two functional forms of heteroscedasticity into the econometric model, the results showed that Glejser test is more powerful in detecting the presence of heteroscedasticity than Breusch-Pagan and White tests.

Two tests for homoscedasticity that requires little knowledge of the functional relationship for determining the variance of the error term was proposed by [6]. The idea of the first test is to approximate the true relationship by Tailor's series expansion, which is essentially linearizing the function in a neighbourhood.

II. Research Methodology

A. Methodology

In this study we consider the classical regression model which is specified by

$$Y = X\beta + u \qquad 2.1$$

Where,

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \qquad X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nk} \end{pmatrix}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \qquad u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

with the following assumptions

i. X is a non-stochastic matrix of rank $k \le n$ and as the sample size n becomes infinitely large,

$$\lim_{n\to\infty} (X'X/n) = Q \qquad 2.2$$

where Q is a finite and nonsingular matrix

ii. The vector u consists of unobservable random errors which satisfy the properties

$$E(u) = 0 \qquad 2.3$$

$$E(uu') = \sigma^2 I \qquad 2.4$$

The conditional mean of Y given X

$$E(Y/X) = X\beta \qquad 2.5$$

Since by assumption $E(u) = 0$. That is, the expected values of the variables Y is a linear functions of the explanatory variables $X_1, X_2, \ldots, X_k$.

Also, assumption (i) excludes the possibility of "exact" multicollinearity among the regressors because of the rank assumption on X. No column of X can be represented as a linear combination of the remaining columns of X.

Assumption (ii) specifies the probability distribution of the error terms. They are independent and identically distributed with zero mean and variance σ^2.

The models used

1. The Constant Elasticity of Substitution Production Function with multiplicative error term can be represented as

$$Y = \theta_1 \left[\theta_2 K^{-\theta_3} + (1-\theta_2) L^{-\theta_3}\right]^{-\frac{1}{\theta_3}} e^u \qquad 2.6$$

where Y is a vector of dependent variables, θ_1 is the intercept, θ_2 and θ_3 are the regression coefficients, K and L are the capital and labour respectively and u is the error term.

By taking the natural logarithms of both sides of (2.6), we have

$$\ln(Y) = \ln(\theta_1) - \frac{1}{\theta_3} \ln\left[\theta_2 K^{-\theta_3} + (1-\theta_2) L^{-\theta_3}\right] + u \qquad 2.7$$

since taking logarithms will not make the nonlinear function linear in the parameters.

The CES production function can be linearized using Kmenta [12] linearization approach. However, a linear Taylor Series expansion around $\theta_3 = 0$ produced an intrinsically linear model. We have,

$$\ln(Y) = \ln(\theta_1) + \theta_2 \ln(K) + (1-\theta_2)\ln(L)$$
$$+\theta_3\theta_2(1-\theta_2)\left[-\frac{1}{2}(\ln K - \ln L)^2 + u\right] \qquad 2.8$$

2. The Exponential Production Function with multiplicative error term can be expressed as

$$Y = \theta_1 e^{\theta_2 K} e^{\theta_3 L} u \qquad 2.9$$

where

Y_i is a vector of dependent variables, θ_1 is the intercept, θ_2 and θ_3 are the regression coefficients, K and L are the capital and labour respectively and u_i is the error term.

By taking the natural logarithms of both sides of (2.9), we have

$$\ln(Y) = \ln(\theta_1) + K\theta_2 + L\theta_3 + \ln u \qquad 2.10$$

B. Error Variance Structures and Tests for Heteroscedasticity

We considered the multiplicative heteroscedasticity error structure model discussed by Harvey [11]. This is given as follows:

$$Y_i = X_i \beta + u_i$$

with,

$$\sigma_i^2 = \sigma^2 E\left(Y_i\right)^2$$
$$= \sigma^2 \left(\beta_1 X_{i1} + \cdots + \beta_k X_{ik}\right)^2$$
$$= \sigma^2 \exp\left(q_i, \lambda\right) \qquad 2.11$$
$$= \sigma^2 \left(Z_i\right)^\lambda$$

where σ^2 and λ are both unknown real constants, which determines the degree of heteroscedasticity.
The three tests that were considered in this paper are White (1980), Glejser (1969) and Park (1966).

III. Analysis

A sample infected with heteroscedasticity using uniform distribution to generate data for Capital (K), Labour (L) and Output (Y). The study used an arbitrary initial values for θ_1 = 0.2, θ_2 = 0.5 and θ_3 = 0.3 for the models.
The set of parameter estimates obtained were used to compute the residuals which represented the dependent variable for the auxiliary regression. The errors were drawn from a normal distribution with mean zero and variance σ^2.
The sample sizes for the simulation were grouped into three categories: 10 & 30; 50 & 100; and 150 & 200 for small, medium and large sample sizes respectively. Each model fitted was replicated 10,000 times.
The degrees of heteroscedasticity, λ, introduced were 0.1, 0.5 and 0.9 for mild, moderate and severe heteroscedasticity respectively.
.

IV. Results

TABLE 1: Power of the Tests for CES model $\alpha = 0.05$

TEST	λ	SAMPLE SIZE					
		10	30	50	100	1	200
WHITE	0.1	0.2231	0.2873	0.0726	0.5204	0.2255	0.0797
GLEJSER		-	0.9133	0.9989	1.0000	1.0000	1.0000
PARK		-	0.1888	0.5755	0.9999	0.9999	1.0000
WHITE	0.5	0.2231	0.2873	0.3227	0.5206	0.3492	0.0000
GLEJSER		1.0000	0.8851	0.9999	1.0000	1.0000	1.0000
PARK		-	0.1856	0.8898	0.9999	1.0000	1.0000
WHITE	0.9	0.2231	0.2873	0.3234	0.5208	0.3492	0.0000
GLEJSER		-	0.8822	0.9999	1.0000	1.0000	1.0000
PARK		-	0.5123	0.8896	0.9999	1.0000	1.0000

TABLE 2: Power of theTests for Exponential model $\alpha = 0.05$

TEST	λ	SAMPLE SIZE					
		10	30	50	100	150	200
WHITE	0.1	0.1674	0.3932	0.6077	0.3473	0.5778	0.6377
GLEJSER		0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
PARK		-	0.8742	0.9516	0.9999	1.0000	1.0000
WHITE	0.5	0.1892	0.8245	0.4745	0.7264	0.9546	0.4304
GLEJSER		0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
PARK		-	0.0603	0.3044	0.9860	1.0000	1.0000
WHITE	0.9	0.4101	0.9373	0.0250	0.1936	0.6450	0.0000
GLEJSER		0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
PARK		-	0.7296	0.8858	0.9999	1.0000	1.0000

.

V. Discussions

Table 1 shows the power of the tests for CES Model at $\alpha = 0.05$. The result shows that at every level of heteroscedasticity, as the sample size increases, the power of the test for Glejser remains high. The power of the Park test also improves from small to large sample sizes while White test has very low power to detect the presence of heteroscedasticity at every level.

Table 2 shows the power of the tests for Exponential Model at $\alpha = 0.05$. There is no results when the sample is 10 at every level of heteroscedasticity due to insufficient sample size for Park test but the power improves as the sample size increases. The power of Glejser test is high from small to large sample sizes at every level of heteroscedasticity.

VI. Conclusion

In this study, we compared the powers of the three tests in detecting the presence of heteroscedasticity in two nonlinear production function models. In view of the analysis, the results show that Glejser test detects heteroscedasticity more efficiently in all sample sizes and at all levels of heteroscedasticity for the two models. The power of the Park testalso improves from small to large sample size at every level of heteroscedasticity for the two models.

REFERENCES

[1] Draper, Norman Richard and Smith, Harry "Applied Regression Analysis" Second Eition, Wiley, 1981 New York.

[2] Neter, J., Wasserman, W and Kutner, M.H "Applied Linear Regression Models" Irwin, 1985, Homewood, Illinois.

[3] Bodkin, R.G and L.R Klein"Nonlinear Estimation of Aggregate Production Functions", this REVIEW, 49, 1987 , pp 29-45.

[4] Koichi, Maekawa."Comparing the Wald, LR and LM Tests for Heteroscedasticityin a Linear Regression Model"Economics Letters,Vol. 26, 1988, pp 37-41.

[5] Ayoola, J. F., and Olubusoye, O. E "Estimation of Parameters of Linear Econometric Modeland the Power of Test in the Presence ofHeteroscedasticity", Journal of Mathematical Theory and Modelling", Vol. 2 (5), 2012, 27-43

[6] Pegiun-Feissolle A "A Comparison of the Power of Some for Conditional Heteroscedasticity", Economics Letters, Elsevier, vol 63(1), 1999, pp 5-17.

[7] Peter, kennedy "A guide to Econometrics", Third Edition, Blackwell Publishers, 1992, Oxford.

[8] Prais, S. J and Houthakker, H. S "The Analysis of family budget" Cambridge University Press, 1995.

[9] Park, R. E. "Estimation with Heteroscedastic Error Term", Econometrica, *Vol* 34(4), 1966, pp 888

[10] Glejser, H "A New Test of Heteroscedasticity",Journal of American Statistical Association,64, 1969, 314-323

[11] Harvey, A. C., (1976) "Estimating Regression Models with Multiplicative Heteroscedasticity", Econometrica, 44, 461-465.

[12] Kmenta, J. "On Estimation of the CES Production Function", International Economic Review, 8(2), 1967, 180-189.

[13] Kmenta, J. "Elements of Econometrics, Macmillian, New York, 1971.

[14] White, Halbert "A Heteroscedasticity-Consistent Covariance Matrix Estimator and A Direct Test for Heteroscedasticity", *Econometrica,* 48(4), 1980, 817-838.