

# Investigating Performances of Some Statistical Tests for Heteroscedasticity Assumption in Generalized Linear Model: A Monte Carlo Simulations Study

Oluwafemi Clement Onifade, Samuel Olayemi Olanrewaju\*

Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria

Email: \*olanrewaju.samuel@uniabuja.edu.ng

**How to cite this paper:** Onifade, O.C. and Olanrewaju, S.O. (2020) Investigating Performances of Some Statistical Tests for Heteroscedasticity Assumption in Generalized Linear Model: A Monte Carlo Simulations Study. *Open Journal of Statistics*, 10, 453-493.

<https://doi.org/10.4236/ojs.2020.103029>

**Received:** April 3, 2020

**Accepted:** May 31, 2020

**Published:** June 3, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In a linear regression model, testing for uniformity of the variance of the residuals is a significant integral part of statistical analysis. This is a crucial assumption that requires statistical confirmation via the use of some statistical tests mostly before carrying out the Analysis of Variance (ANOVA) technique. Many academic researchers have published series of papers (articles) on some tests for detecting variance heterogeneity assumption in multiple linear regression models. So many comparisons on these tests have been made using various statistical techniques like biases, error rates as well as powers. Aside comparisons, modifications of some of these statistical tests for detecting variance heterogeneity have been reported in some literatures in recent years. In a multiple linear regression situation, much work has not been done on comparing some selected statistical tests for homoscedasticity assumption when linear, quadratic, square root, and exponential forms of heteroscedasticity are injected into the residuals. As a result of this fact, the present study intends to work extensively on all these areas of interest with a view to filling the gap. The paper aims at providing a comprehensive comparative analysis of asymptotic behaviour of some selected statistical tests for homoscedasticity assumption in order to hunt for the best statistical test for detecting heteroscedasticity in a multiple linear regression scenario with varying variances and levels of significance. In the literature, several tests for homoscedasticity are available but only nine: Breusch-Godfrey test, studentized Breusch-Pagan test, White's test, Nonconstant Variance Score test, Park test, Spearman Rank, Glejser test, Goldfeld-Quandt test, Harrison-McCabe test were considered for this study; this is with a view to examining, by Monte Carlo simulations, their asymptotic behaviours. However, four different forms of heteroscedastic

---

structures: exponential and linear (generalize of square-root and quadratic structures) were injected into the residual part of the multiple linear regression models at different categories of sample sizes: 30, 50, 100, 200, 500 and 1000. Evaluations of the performances were done within R environment. Among other findings, our investigations revealed that Glejser and Park tests returned the best test to employ to check for heteroscedasticity in EHS and LHS respectively also White and Harrison-McCabe tests returned the best test to employ to check for homoscedasticity in EHS and LHS respectively for sample size less than 50.

## Keywords

Homoscedasticity, Heteroscedasticity, Generalized Linear Model, Monte Carlo

---

## 1. Introduction

One of the crucial assumptions in the multiple linear regression models is that the variance of the errors should be constant [1]. The Ordinary Least Squares (OLS) method is very popular with statistics practitioners as it provides efficient and unbiased estimates of the parameters when the assumptions, especially the assumption of homoscedastic error variances, are met. But in many real-life applications, variances of the errors vary across observations. Since homoscedasticity is often unrealistic assumption, researchers should consider how the results are affected by heteroscedasticity. Even though the OLS estimates retain unbiasedness in the presence of heteroscedasticity, its estimates become inefficient [2] [3].

However, heteroscedasticity yields hypothesis tests that fail to keep false rejections at the nominal level, or estimated standard errors as well as confidence intervals that are either too narrow or too large [4]. Every statistical procedure carries with it certain assumptions that must be at least approximately true before the procedure can produce reliable and accurate results [3]. Researchers often apply a statistical procedure to their data without checking on the validity of the assumptions of the procedure. If one or more of the assumptions of a given statistical procedure are violated, most especially in multiple linear regression analysis, then misleading results will be produced by the procedure [5].

In short, a number of assumptions are associated with the analysis of data using OLS in multiple linear regression but the current study deals with only one of them, that is, homogeneity of variance assumption. Literarily, assumptions refer to basic principles that are accepted on faith, or assumed to be true, without proof or verification. It is frequent and common that a researcher applies a statistical method to a set of data without thoroughly checking that the assumptions of the methods are valid [6]. This may be especially true in multiple linear modelling.

Notwithstanding, it is a known fact that all statistical procedures should have underlying assumptions; some are more stringent than others. In some cases, violation of these assumptions will not change substantive research conclusions. In other cases, violation of assumptions will undermine meaningful research. Establishing that one's data meet the assumptions of the procedure one is using is an expected component of all quantitatively-based theses, journal articles, and dissertations. This assumption practically and usually exists in regression and experimental design but this research discusses its relation with regression analysis.

The homogeneity of variance assumption is one of the critical assumptions underlying most parametric statistical procedures such as the analysis of variance and it is very important for a good researcher to be able to test this assumption before the application of ANOVA technique. Simply, the term "homo" means "the same" while "hetero" means different, therefore variance homogeneity assumption, which is equivalently called "homoscedasticity", means that the variance of each residual should be the same throughout the experiment. If the errors (residuals) fail to possess equal (but sometimes unknown) variances, the reliability of application of analysis of variance technique may be badly affected [3]. Direct opposite in meaning to "homogeneity assumption" is "heterogeneity of error variances", which simply refers to a situation where the variance of the residuals is affected by at least one predictor variable leading to unequal magnitude in spread. Thus, heterogeneity problem may arise in most of the economic (econometric), experimental and agricultural modelling where specifically analysis of variance technique is applied. Thus, homogeneity of variance is a major assumption underlying the validity of many parametric tests. More importantly, it serves as the null hypothesis in substantive studies that focus on cross- or within-group dispersion.

In addition, showing that several samples do not come from populations with the same variance is sometimes of importance *per se*. The statistical validity of many commonly used tests such as the *t*-test and ANOVA depend on the extent to which the data conform to the assumption of homogeneity of variance. When a research design, however, involves groups that have very different variances, the *p*-value accompanying the test statistic, such as *t* and *F* may be too lenient or too harsh. Thus, substantive research often requires investigation of cross- or within-group fluctuation in dispersion. For example, in quality control research, homogeneity of variance tests is often "a useful endpoint in an analysis" [7]. In human performance studies, an increase or decrease in the dispersion of performance scores within the same group of subjects may shed light on how changing condition affect human behaviour. Recent studies on gender-related differences in the dispersion of academic performance have provoked substantive as well as methodological interest in homogeneity of variance [8] [9].

It has been reported in some literatures that the assumption of the error term is such that its probability distribution remains the same over all observations of

the explanatory variables, and in particular the variance of each  $e_i$  is the same for all values of the predictor variables [10]. This assumption is also known as Assumption of Homogeneity of Variances or the Assumption of Constant Variance of the error term. If it is not satisfied in any particular case, we say that the error term ( $e_i$ ) is heteroscedastic. The meaning of the assumption of homoscedasticity is that the variation of each error ( $e_i$ ) around its zero mean does not depend on the values of predictor variables. The variance of each  $e_i$  remains the same irrespective of small or large values of the explanatory variables.

Apparently, the present research intends to investigate the best statistical test, through the computation of the number of time (frequency) each test commits type II error (when  $\sigma = 0$ ) and type I error ( $\sigma \neq 0$ ) for confirming homoscedasticity assumption when different levels of heteroscedasticity are injected into the multiple linear regression models at 30, 50, 100, 200, 500 and 1000 sample sizes.

## 2. Aim and Objectives of the Study

This study aims at providing a comprehensive comparative analysis of asymptotic behaviour of some selected statistical tests for homoscedasticity assumption in order to hunt for the best statistical test for detecting heteroscedasticity in a multiple linear regression scenario with varying variances. In order to achieve this aim, the following objectives are pursued:

- 1) To compare nine different statistical tests under different heteroscedastic conditions such as Exponential and Linear (generalized structure for Quadratic and Square-root) Forms;
- 2) To evaluate, through the computation of the number of time (frequency) each test commits type II error (when  $\sigma = 0$ ) and type I error ( $\sigma \neq 0$ ) as the case may be;
- 3) To investigate the asymptotic behaviour of the selected tests when the variances are varied across all simulations.

## 3. Theoretical Framework and Literature Review

The analysis of variance (ANOVA) is one of the most important and useful techniques for variety of fields such as economics, agriculture, biology and so on with a view to comparing different groups or treatments with respect to their means [3]. Let us consider testing equality of means of  $k$  populations given samples  $\{x_{ij} : i = 1, 2, \dots, k; j = 1, 2, \dots, n_i\}$  from  $i^{\text{th}}$  population with mean  $\mu_i$ , variance  $\sigma_i^2$  and distribution function  $F\{\sigma_i^{-1}(x - \mu_i)\}$ . While doing ANOVA, the null hypothesis to be tested is

$$H_{01} : \mu_1 = \mu_2 = \dots = \mu_k \text{ versus } H_{11} : \mu_i \neq \mu_j \text{ for some } i \neq j \quad (2.1)$$

Hence, a set of assumptions such as normality, homogeneity and independence of observations has to be made in order to employ an  $F$  test for (2.1). As [1] has pointed out, in practice, the assumption of homogeneity of variances is the one most often unmet in ANOVA. In the absence of homogeneity of variances,

the sample means will not necessarily have equal expected standard errors, and an exact solution for comparison among the means requires reference to a compound  $F$  distribution with unknown parameters if the true population variances are unknown. It is now well established that the violation of the assumption of homogeneity of variances can have severe effects on inference about the population means, especially in the case of unequal sample sizes [11] [12].

Furthermore, [13] has demonstrated that the ANOVA  $F$  is not robust to all degrees of variance heterogeneity even when sample sizes are equal. In fact, the conventional ANOVA  $F$  provides generally poor control over both Type I and Type II error rates under a wide variety of variance heterogeneity conditions. Therefore, the problem of homogeneity of variances has to be settled before performing an ANOVA [14] [15] [16].

Obviously, there are a good number of methods available to test for homogeneity of variances for different situations [17] [18] [19]. The most common and popular tests, however, in the case of one-way ANOVA, are Bartlett and Levene tests. The hypothesis to be tested for homogeneity of variances is

$$H_{02} : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 \text{ against } H_{12} : \sigma_i^2 \neq \sigma_j^2 \text{ for some } i \neq j \quad (2.2)$$

Unfortunately, these tests (Bartlett and Levene) are sensitive to the assumption of normality [20]. Specifically, the probability of a Type I error ( $\alpha$ ) is dependent upon the kurtosis of the distribution. However, these alternative tests are affected adversely by non-normality [21]. As [22] pointed out, none of the procedures directly handle the problem of skewness bias. [23] has also argued that failure to consider the impact of the combined violations of variance equality and distribution normality is an important omission of a statistical procedure. Under these circumstances, the development of new alternative tests, namely, trimming, transforming statistics, bootstrapping [24] to deal with unequal variances, and non-normality is worthwhile.

[25] employed four homogeneity tests, named SNHT, Buishand Range test, Pettitt test and Von Neumann Ratio (VNR) test to the European Climate. The results are categorized into three classes, which are useful, doubtful and suspect according to the number of tests rejecting the null hypothesis. Three testing variables were used, each consists of annual values. For temperature, the two testing variables are annual mean of diurnal temperature range and annual mean of the absolute day-to-day differences. Meanwhile for precipitation, the annual number of wet days (threshold 1 mm) is employed.

Tests for equality of variances are of interest in many situations such as analysis of variance or quality control [26]. The classical approach to hypothesis testing begins with the likelihood ratio test under the assumption of normal distributions given by [27]. However, as this test is very sensitive to departures from normality, many alternative tests have been prompted since then. Some of these are modifications of likelihood ratio test. Others are adaptations of the  $F$ -test to test variances rather than means. Against one would hope, recent comparative studies reveal that some of these tests present lack of robustness and have poor

power [28]. This problem was also studied by [7], who performed an increasing study of many of the existing parametric and non-parametric tests. In his last paper Monte Carlo simulations of some distributions for several sample sizes show a few tests that are robust and have good power.

## 4. Methodology

### 4.1. Forms of Heteroscedasticity

The study considers four different heteroscedastic structures coined from [29] additive and multiplicative heteroscedastic model but in our model, we assume that the variance of the error varies as the mean of the responses. The two general forms are:

- 1)  $Var(e'e) = \sigma^2 e^{E(y_i)}$ ;
- 2)  $Var(e'e) = \sigma^2 E(y_i)^g$ ; where  $g \geq 0$ .

Emanating from the two above, four heteroscedastic structures were formulated as follows:

- 1) Exponential Form:  $h_{i1} = \sigma^2 e^{E(y_i)} = \sigma^2 e^{(\beta_1 x_1 + \beta_2 x_2)}$ ;
- 2) Linear Form:  $h_{i2} = \sigma^2 E(y_i)^g = \sigma^2 (\beta_1 x_1 + \beta_2 x_2)^1$ ;
- 3) Square-rooted Form:  $h_{i3} = \sigma^2 E(y_i)^g = \sigma^2 (\beta_1 x_1 + \beta_2 x_2)^{0.5}$ ;
- 4) Quadratic Form:  $h_{i4} = \sigma^2 E(y_i)^g = \sigma^2 (\beta_1 x_1 + \beta_2 x_2)^2$ .

### 4.2. Procedure for the Monte Carlo Simulation Experiment

To investigate the finite sample properties of the test statistics of the presence of heteroscedasticity in any given dataset, we use a Monte Carlo experiment. We simulate a linear multiple regression model with three explanatory variables model using a simple Least square function of the form:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + e \tag{3.1}$$

where  $e$  is a normal error variable.

Following [30], we then simulate the independent variables  $X_i$  as follows:

$X_i$ 's are a set of independent variables that are fixed following  $X_1 = 1, \dots, N$ ,  $X_2 = 1, \dots, \sqrt{N}$ ,  $X_3 = \text{norm}(10, 1)$  and  $Z = \sqrt{i}$ , for  $i = 1, \dots, N$ .

We also generate the two error terms as follows  $v \sim N(0, \sigma^2)$  and  $u \sim N(0, \sigma^2)$

Then the model for the heteroscedasticity function was formed as follows:

$$\sigma_v = \exp(\alpha_0 + \sigma_{\alpha 1} \ln X_{1i} + \sigma_{\alpha 2} \ln X_{2i}) \tag{3.2}$$

The parameters were then set at:

$$\beta_0 = \alpha_0 = \gamma_0 = \alpha_1 = \alpha_2 = \gamma_1 = 1$$

$$\beta_1 = \beta_2 = \beta_3 = 0.5 \text{ (constant return to scale)}$$

The parameter  $\sigma$  measures the degree of heteroscedasticity. We use several degrees of heteroscedasticity by letting  $\sigma$  to vary (0, 0.1, 0.3, 0.7 and 0.9). When  $\sigma = 0$ , we obtain the homoscedastic case. We considered also different sample sizes: 30, 50, 100, 200, 500 and 1000 observations (Hadri and Garry,

1998). To analyze the performance of the heteroscedasticity test statistic as we estimated the OLS model and then tested with the following heteroscedastic test statistics (Breusch-Godfrey test, studentized Breusch-Pagan test, White's Test, Non-constant Variance Score Test, Park test, Spearman Rank, GLEJSER test, Goldfeld-Quandt test, Harrison-McCabe test) for the true presence of the problem using the test frequency of significance at 1%, 5% and 10% respectively. We also set the number of replications to 1000. Refer to **Appendix 1** for the R simulation code.

### 4.3. Tests for Detecting Heteroscedasticity

This study considers nine (9) popularly used heteroscedastic test as evident in past works of literature. These tests include: Breusch-Godfrey test, studentized Breusch-Pagan test, White's Test, Non-constant Variance Score Test, Park test, Spearman Rank, GLEJSER test, Goldfeld-Quandt test, Harrison-McCabe test.

### 4.4. Park Test

Park (1966) formalizes the graphical method by suggesting that  $\sigma_i^2$  is some function of the explanatory variable  $X_i$ . The functional form he suggested was:

$$\sigma_i^2 = \sigma^2 X_i^\beta e^{v_i} \quad (3.3)$$

$$\text{or } \ln \sigma_i^2 = \ln \sigma^2 + \beta \ln X_i + v_i \quad (3.4)$$

where  $v_i$  is the stochastic disturbance term. Since  $\sigma_i^2$  is generally not known, Park suggests using  $\hat{u}_i^2$  as a proxy and running the following regression:

$$\ln \hat{u}_i^2 = \ln \sigma^2 + \beta \ln X_i + v_i = \alpha + \beta \ln X_i + v_i \quad (3.5)$$

If  $\beta$  turns out to be statistically significant, it would suggest that heteroscedasticity is present in the data. If it turns out to be insignificant, we may accept the assumption of variance homogeneity. The Park test is thus a two-stage procedure. In the first stage, we run the OLS regression disregarding the heteroscedasticity question. We obtain  $\hat{u}_i$  from this regression, and then in the second stage, we run the regression (3.5).

Although empirically appealing, the Park test has some problems. [31] has argued that the error term  $v_i$  entering into (3.5) may not satisfy the OLS assumptions and may itself be heteroscedastic. Nonetheless, as a strictly exploratory method, one may still use the Park test.

### 4.5. Glejser Test

The [32] is similar in spirit to the Park test. After obtaining the residuals  $\hat{u}_i$  from the OLS regression, Glejser suggests regressing the absolute values of  $\hat{u}_i$  on the  $X$  variable that is thought to be closely associated with  $\sigma_i^2$ . In his experiments, Glejser used the following functional forms:

$$|\hat{u}_i| = \beta_0 + \beta_1 X_i + v_i \quad (3.6)$$

$$|\hat{u}_i| = \beta_0 + \beta_1 \sqrt{X_i} + v_i \quad (3.7)$$

$$|\hat{u}_i| = \beta_0 + \beta_1 \frac{1}{X_i} + v_i \quad (3.8)$$

$$|\hat{u}_i| = \beta_0 + \beta_1 \frac{1}{\sqrt{X_i}} + v_i \quad (3.9)$$

$$|\hat{u}_i| = \sqrt{\beta_0 + \beta_1 X_i} + v_i \quad (3.10)$$

$$|\hat{u}_i| = \sqrt{\beta_0 + \beta_1 X_i^2} + v_i \quad (3.11)$$

where  $v_i$  is the error term.

Again, as an empirical or practical matter, one may use the Glejser approach. But Goldfeld and Quandt pointed out that the error term  $v_i$  has some problems in that its expected value is nonzero, it is serially correlated. Above all, we shall choose the best form of regression which gives the best fit from the viewpoint of correlation coefficient and standard error of the coefficients. It's reported that if  $\beta_0 = 0$  and  $\beta_1 \neq 0$ , pure heteroscedasticity is suggested and, also if both  $\beta_0$  and  $\beta_1$  differ from zero, it is an indication of mixed heteroscedasticity. This could be achieved by conducting statistical significance of both  $\beta_0$  and  $\beta_1$ .

#### 4.6. Goldfeld-Quandt Test

This is a simple and intuitive test. One orders the observations according to  $X_i$  and omits  $c$  central observations. Next, two regressions are run on the two separated sets of observations with  $(n-c)/2$  observations in each. The  $c$  omitted observations separate the low-value  $X$ 's from the high-value  $X$ 's, and if heteroscedasticity exists and is related to  $X_i$ , the estimates of  $\sigma^2$  reported from the two regressions should be different. Hence, the test statistic is  $s_2^2/s_1^2$ , where  $s_1^2$  and  $s_2^2$  are the Mean Square Error of the two regressions, respectively. Their ratio would be the same as that of the two residual sums of squares because the degrees of freedom of the two regressions are the same. This statistic is  $F$ -distributed with  $[(n-c)/2 - k]$  degrees of freedom in the numerator as well as the denominator.

The only remaining question for performing this test is the magnitude of  $c$ . Obviously, the larger  $c$  is, the more central observations are being omitted and the more confident we feel that the two samples are distant from each other. The loss of  $c$  observations should lead to loss of power. However, separating the two samples should give us more confidence that the two variances are in fact the same if we do not reject homoscedasticity. This trade off in power was studied by [32] using Monte Carlo experiments. Their results recommend the use of  $c = 8$  for  $n = 30$  and  $c = 16$  for  $n = 60$ . This is a popular test, but assumes that we know how to order the heteroscedasticity. In this case, we use  $X_i$ . But what if there are more than one regressor on the right-hand side? In that case one can order the observations using  $\hat{Y}_i$ .

#### 4.7. Breusch-Pagan Test

The success of the Goldfeld-Quandt test depends not only on the value of  $c$  (the



number of central observations to be omitted) but also on identifying the correct  $X$  variable with which to order the observations. This limitation of this test can be avoided if we consider the Breusch-Pagan (BP) test (1979). Consider the  $k$ -variable linear regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + e_i \quad (3.12)$$

Assuming that the error variance  $\sigma_i^2$  is described as follows:

$$\sigma_i^2 = f(\alpha_0 + \alpha_1 z_{1i} + \cdots + \alpha_m z_{mi}) \quad (3.13)$$

that is,  $\sigma_i^2$  is some function of the non-stochastic variables  $z$ 's; some or all of the  $x$ 's can serve as  $z$ 's. Specifically, assume that:

$$\sigma_i^2 = \alpha_0 + \alpha_1 z_{1i} + \cdots + \alpha_m z_{mi} \quad (3.14)$$

that is,  $\sigma_i^2$  is a linear function of the  $z$ 's. If  $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ ,  $\sigma_i^2 = \alpha_0$ , which is a constant. Therefore, to test whether  $\sigma_i^2$  is homoscedastic, one can test the hypothesis that  $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ . This is the basic idea behind the Breusch-Pagan test. The actual procedure is tailored as follows:

- 1) Obtain  $y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + e_i$  by OLS and compute the residuals;
- 2) Obtain  $\hat{\sigma}^2 = \frac{1}{n} \sum e_i^2$ , which would be MLE of  $\sigma^2$  under homoscedasticity;

- 3) Obtain another variable such that  $p_i = \frac{e_i^2}{\hat{\sigma}^2}$ ;

- 4) Obtain the regression such that  $p_i = \alpha_0 + \alpha_1 z_{1i} + \cdots + \alpha_m z_{mi} + v_i$ , where  $v_i$  is the residual term of this regression;

- 5) Obtain the statistic:  $BPG = \frac{1}{2}(SSR)$ , where  $SSR$  is the Regression Sum of Squares;

- 6) Assuming  $e_i$  are normally distributed, one can show that if there is homoscedasticity and if the sample size  $n$  increases indefinitely, then:

$$BP \sim \chi_{m-1}^2. \quad (3.15)$$

#### 4.8. White Test

Another general test for homoscedasticity where nothing is known about the form of this heteroscedasticity is suggested by [5]. This test is based on the difference between the variance of the OLS estimates under homoscedasticity and that under heteroscedasticity. This test does not rely on normality assumption making it very easy to implement. Consider the following three-variable regression model

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e_i \quad (3.16)$$

The White test is tailored as follows:

- 1) Given a set of data, obtain the residual,  $e_i$  from (3.16);
- 2) Run the following auxiliary regression:

$$e_i^2 = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_1^2 + \alpha_4 x_2^2 + \alpha_5 x_1 x_2 + v_i \quad (3.17)$$

3) Obtain  $R^2$  from this auxiliary regression;

4) Under  $H_0$  that there is no heteroscedasticity, it can be shown that the sample size ( $n$ ) times  $R^2$  obtained from the auxiliary regression *asymptotically* follows the chi-squared distribution with *degree of freedom* equal to the number of regressors (excluding the constant term) in the auxiliary regression. Mathematically, we have:

$$nR^2 \sim \chi_{k;\alpha}^2 \tag{3.18}$$

5) It is expected that the null hypothesis will be rejected when  $nR^2$  exceeds the critical value obtained from chi-square table at a given level of significance.

It is observed that if a model has several regressors, then introducing all the regressors, their squared (or higher-powered) terms, and their product can quickly consume degree of freedom. Therefore, one must be very cautious of using the test; this is one of the demerits of this test.

### 4.9. Spearman’s Rank Correlation Test

This test ranks the  $x_i$ ’s and the absolute value of OLS residuals, the  $e_i$ ’s. Then it computes the difference between these rankings, that is,  $d_i = \text{rank}(|e_i|) - \text{rank}(x_i)$ . For this simple linear regression model, we obtain the Spearman’s Rank Correlation Coefficient as follows:

$$r_{|e_i|,x_i} = 1 - \left[ \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \right] \tag{3.19}$$

Having obtained (3.19), the next step is to test for the significance of the coefficient using  $t$ -test as follows:

$$T = \frac{r_{|e_i|,x_i} \sqrt{n-2}}{\sqrt{1-r_{|e_i|,x_i}^2}} \tag{3.20}$$

The statistic is  $t$ -distributed with  $(n - 2)$  degree of freedom under any level of significance.

In a situation where the number of regressors are more than one (multiple linear regression case), that is,  $y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + e_i$ , it’s suggested that  $r_{|e_i|,x_1}, r_{|e_i|,x_2}, \dots, r_{|e_i|,x_k}$  should be computed separately and the same  $t$ -test should be used for testing the significance of each of the correlation coefficient. This is the test intended to modify in the present research.

The choice of statistical tests for this study is in connection with the existing literature, most especially in papers entitled “Heteroscedasticity as Basis of Direction Dependence in Reversible Linear Regression Models” authored by [33], it’s reported that Bartlett’s, Goldfeld-Quandt, Breusch-Pagan, White and Koenker-Bassett Tests are the major tests predominantly used in multiple linear regression models. In addition, we decide judgmentally to include Spearman’s Rank Test making a total of six statistical tests for detecting homoscedasticity assumption.

#### 4.10. Breusch-Godfrey Serial Correlation

The Breusch-Godfrey serial correlation LM test is a test for heteroscedasticity in the errors in a regression model. It makes use of the residuals from the model being considered in a regression, and a test statistic is derived from these.

This test is valid with lagged dependent variables and can be used to test for heteroscedasticity

##### Procedure

Step 1. Estimate.

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 Y_{t-1} + U_t \quad (3.21)$$

obtain the residuals ( $e_t$ ).

Step 2. Estimate the following auxiliary regression.

model:

$$e_t = b_1 + b_2 X_{2t} + b_3 X_{3t} + b_4 Y_{t-1} + c_1 e_{t-1} + c_2 e_{t-2} + c_3 e_{t-3} + w_t \quad (3.22)$$

Step 3. For large sample sizes, the test statistic is:

$$(n-p)R^2 \sim \chi_p^2 \quad (3.23)$$

Step 4. If the test statistic exceeds the critical chi-square value we can reject the null hypothesis of no serial correlation in any of the  $\rho$  terms.

Other tests are “Non-constant variance score test” and Harrison-McCabe test.

#### 4.11. Heteroscedasticity Correction

From the section above, a general linear regression model with the assumption of heteroscedasticity can be expressed as follows:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + e \quad (3.24)$$

Letting  $e = \mu_t$

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \mu_t \quad (3.25)$$

$$Var(\mu_t) = E(\mu_t^2) = \sigma_t^2 \text{ for } t = 1, 2, \dots, n$$

where:

Noting that the  $t$  subscript attached to sigma squared indicates that the disturbance for each of the  $n$ -units is drawn from a probability distribution that has a different variance.

Given such a non-constant variance function

$$Var(e_i) = \sigma_i^2 = \sigma_i^2 x_i^\alpha \quad (3.26)$$

where  $\alpha$  is the unknown parameter in the model.

Taking the natural logarithm

$$\ln(\sigma_i^2) = \ln(\sigma_i^2) + \alpha \ln(x_i) \quad (3.27)$$

Then taking exponential of equation

$$\sigma_i^2 = \exp[\ln(\sigma_i^2) + \alpha \ln(x_i)] \quad (3.28)$$

Letting  $\beta_1 = \ln(\sigma_i^2)$ ,  $\beta_2 = \alpha$ ,  $Z_i = \ln(x_i)$

$$\sigma_i^2 = \exp[\beta_1 + \beta_2 Z_i] \tag{3.29}$$

$$\sigma_i^2 = \exp[\beta_1 + \beta_2 Z_{i2} + \dots + \beta_3 Z_{is}] \tag{3.29*}$$

If the variance depends on more than one explanatory variable (a multiple regression case) Taking the exponential function is best because it gives non-negative value of variance  $\sigma_i^2$ .

From Equation (3.27) with  $\beta_1 = \ln(\sigma_i^2)$ ,  $\beta_2 = \alpha$ ,  $Z_i = \ln(x_i)$

Using the OLS technique to estimate the coefficients  $\beta_1, \beta_2, \dots, \beta_s$  of the variance function

$$\ln(\sigma_i^2) = \beta_1 + \beta_2 Z_{i2} + \dots + \beta_s Z_{is} \tag{3.30}$$

where  $Z_{i2} = \ln(x_2)$ ,  $Z_{i3} = \ln(x_3)$  and so on.

We then took the square root of the exponential of the fitted estimate

$$\hat{\sigma}_i = \sqrt{\exp(\hat{\beta}_1 + \hat{\beta}_2 Z_{i2} + \dots + \hat{\beta}_s Z_{is})} \tag{3.31}$$

Then  $\hat{\sigma}_i$  is the weight required to transform the data set by dividing through.

But;

$$\text{Var}\left(\frac{e_i}{\sigma_i}\right) = \frac{1}{\sigma_i^2} \text{Var}(e_i) = \frac{1}{\sigma_i^2} \times \sigma_i^2 = 1 \tag{3.32}$$

Using the estimate of our variance function  $\hat{\sigma}_i^2$  in place of  $\sigma_i^2$  in Equation (3.30) to obtain the Generalized Least Square Estimator of  $\beta_1, \beta_2, \dots, \beta_s$ .

We then defined the transformed variable as

$$y_i^* = \frac{y_i}{\hat{\sigma}_i}, x_{i1}^* = \frac{1}{\hat{\sigma}_i}, x_{i2}^* = \frac{x_i}{\hat{\sigma}_i}, \dots, x_{is}^* = \frac{x_s}{\hat{\sigma}_i} \tag{3.33}$$

Therefore;

$$y_i^* = \beta_1 x_{i1}^* + \beta_2 x_{i2}^* + \dots + \beta_s x_{is}^* + e_i^* \tag{3.34}$$

which is the Weighted Least Squares model with homoscedasticity.

## 5. Analysis and Results

### 5.1. Comparative Analyses of Some Statistical Tests for Homoscedasticity Assumption

As earlier mentioned, nine statistical tests are compared in this study with the use of the number of time (frequency) each test commits type II error and type I error as the case may be, such that the one with the least frequency type II error when sigma = 0 shall be considered as the best among others and the test with the highest frequency (type I error) when sigma ≠ 0 shall be considered as the best among others. The null hypothesis is such that homoscedasticity assumption is upheld.

### 5.2. Performance of the Tests When Error Follows Exponential Heteroscedastic Structure (EHS)

**Table 1** presents the frequency of tests of significance at 1% level after 1000

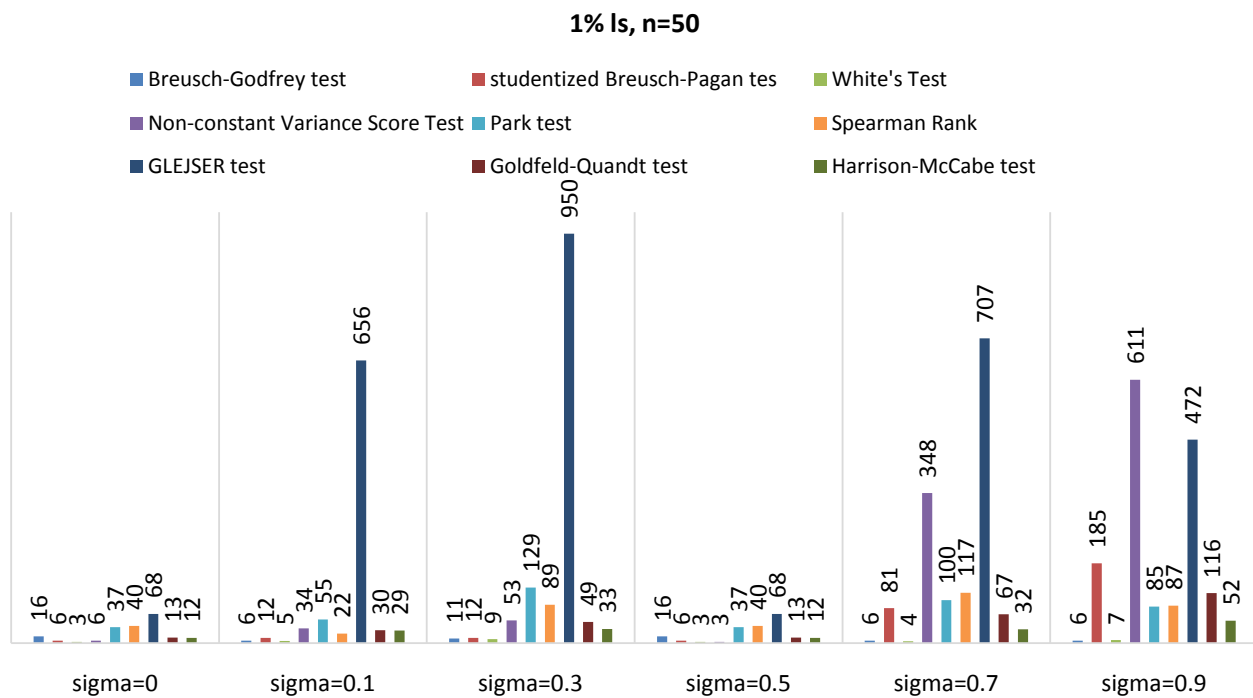
**Table 1.** Performance of the tests when error follows exponential heteroscedastic structure (EHS) at 1%.

significance level	sample size ( <i>n</i> )	level of hetero	Breusch-Godfrey test	studentized Breusch-Pagan Test	White's Test	Non-constant Variance Score Test	Park test	Spearman Rank	GLEJ/SER test	Goldfeld-Quandt test	Harrison-McCabe test
1%	<i>n</i> = 30	sigma = 0	18	3	0	10	34	22	59	6	9
1%	<i>n</i> = 30	sigma = 0.1	13	6	0	27	42	20	417	23	17
1%	<i>n</i> = 30	sigma = 0.3	11	8	1	33	74	54	727	48	20
1%	<i>n</i> = 30	sigma = 0.5	18	3	1	6	34	22	59	6	9
1%	<i>n</i> = 30	sigma = 0.7	6	37	1	178	90	84	413	69	12
1%	<i>n</i> = 30	sigma = 0.9	6	70	2	314	70	70	284	92	11
1%	<i>n</i> = 50	sigma = 0	16	6	3	6	37	40	68	13	12
1%	<i>n</i> = 50	sigma = 0.1	6	12	5	34	55	22	656	30	29
1%	<i>n</i> = 50	sigma = 0.3	11	12	9	53	129	89	950	49	33
1%	<i>n</i> = 50	sigma = 0.5	16	6	3	3	37	40	68	13	12
1%	<i>n</i> = 50	sigma = 0.7	6	81	4	348	100	117	707	67	32
1%	<i>n</i> = 50	sigma = 0.9	6	185	7	611	85	87	472	116	52
1%	<i>n</i> = 100	sigma = 0	9	9	12	25	9	30	6	73	11
1%	<i>n</i> = 100	sigma = 0.1	8	19	12	46	151	28	924	40	34
1%	<i>n</i> = 100	sigma = 0.3	11	14	16	74	350	74	1000	58	52
1%	<i>n</i> = 100	sigma = 0.5	11	9	11	8	30	25	73	9	10
1%	<i>n</i> = 100	sigma = 0.7	8	332	16	728	136	96	984	102	74
1%	<i>n</i> = 100	sigma = 0.9	11	643	16	929	93	55	903	146	119
1%	<i>n</i> = 200	sigma = 0	12	8	19	11	42	36	98	17	16
1%	<i>n</i> = 200	sigma = 0.1	14	46	19	69	460	18	999	37	37
1%	<i>n</i> = 200	sigma = 0.3	12	14	16	76	819	87	1000	52	49
1%	<i>n</i> = 200	sigma = 0.5	12	8	17	6	42	36	98	17	16
1%	<i>n</i> = 200	sigma = 0.7	9	818	18	962	355	71	999	115	94
1%	<i>n</i> = 200	sigma = 0.9	6	991	17	999	173	24	998	167	159
1%	<i>n</i> = 500	sigma = 0	11	8	16	8	31	29	91	11	12
1%	<i>n</i> = 500	sigma = 0.1	9	207	18	93	922	17	1000	39	38
1%	<i>n</i> = 500	sigma = 0.3	6	21	13	85	999	83	1000	60	56
1%	<i>n</i> = 500	sigma = 0.5	11	8	16	12	31	29	91	11	12
1%	<i>n</i> = 500	sigma = 0.7	11	999	22	1000	895	14	1000	140	135
1%	<i>n</i> = 500	sigma = 0.9	13	1000	23	1000	584	0	1000	192	188
1%	<i>n</i> = 1000	sigma = 0	10	15	12	11	29	31	88	7	9
1%	<i>n</i> = 1000	sigma = 0.1	13	595	12	203	999	10	1000	41	39
1%	<i>n</i> = 1000	sigma = 0.3	12	35	21	103	1000	85	1000	52	53
1%	<i>n</i> = 1000	sigma = 0.5	10	15	11	16	29	31	88	7	9
1%	<i>n</i> = 1000	sigma = 0.7	9	1000	19	1000	1000	0	1000	141	134
1%	<i>n</i> = 1000	sigma = 0.9	10	1000	26	1000	963	0	1000	194	189

\*Frequency of test significance after 1000 replications.

replications for errors that follow EHS. As observed from the simulation results in **Table 1** and **Figure 1**, the OLS model was not contaminated with level heteroscedasticity, but yet tested with the nine various tests of heteroscedasticity to detect the true nature of the various test statistics at sample size of  $n = 30$ . At no presence of heteroscedasticity, the following tests returned the following rate in percent of type two error; Goldfeld-Quandt test 0.6%, Studentized Breusch-Pagan test 0.3%, White test with 0%, and so on. Making the White test the best in terms of detecting no presence of heteroscedasticity (*i.e.*  $\sigma = 0$ ). While the model was infused with the level of heteroscedasticity (*i.e.*  $\sigma = 0.1, 0.3, 0.5, 0.7$  &  $0.9$ ) at sample size of 30. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 417, 727, 59 and 413 corrected tests results at 0.1, 0.3, 0.5 & 0.7 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 41.7%, 72.7%, 5.9% and 41.3% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5 or 0.7 at sample size of 30. Howbeit, Non-constant Variance Score Test with 31.4% outperformed the celebrated Glejser test when sigma = 0.9.

Moreover, considering sample size 50, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Studentized Breusch-Pagan test 0.6%, Non-constant Variance Score test 0.6%, White test with 0.3%, and so on (see **Table 1** and **Figure 1**). Hence, White test the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 656, 950, 68 and 707 corrected tests results at 0.1, 0.3, 0.5 & 0.7 sigma levels



**Figure 1.** Sample size 50 results when error follows EHS at 1%.

respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 65.6%, 95.0%, 6.8% and 70.7% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5 or 0.7 at sample size of 50. However, Non-constant Variance Score test with 61.1% outperformed the celebrated Glejser test when sigma = 0.9.

In addition, considering sample size 100, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 0.9%, Studentized Breusch-Pagan test 0.9%, Park test 0.9%, White test with 1.2%, Glejser test 0.6% and so on (see **Table 1** and **Figure 2**). Hence, the celebrated White test at sigma = 0 was displayed by Glejser test being the best in terms of detecting no presence of heteroscedasticity when sample size is 100. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 924, 1000, 73 and 984 corrected tests results at 0.1, 0.3, 0.5 & 0.7 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 92.4%, 100%, 7.3% and 98.4% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5 or 0.7 at sample size of 100. However, Non-constant Variance Score test with 92.9% outperformed the celebrated Glejser test when sigma = 0.9.

Furthermore, considering sample size 200, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two

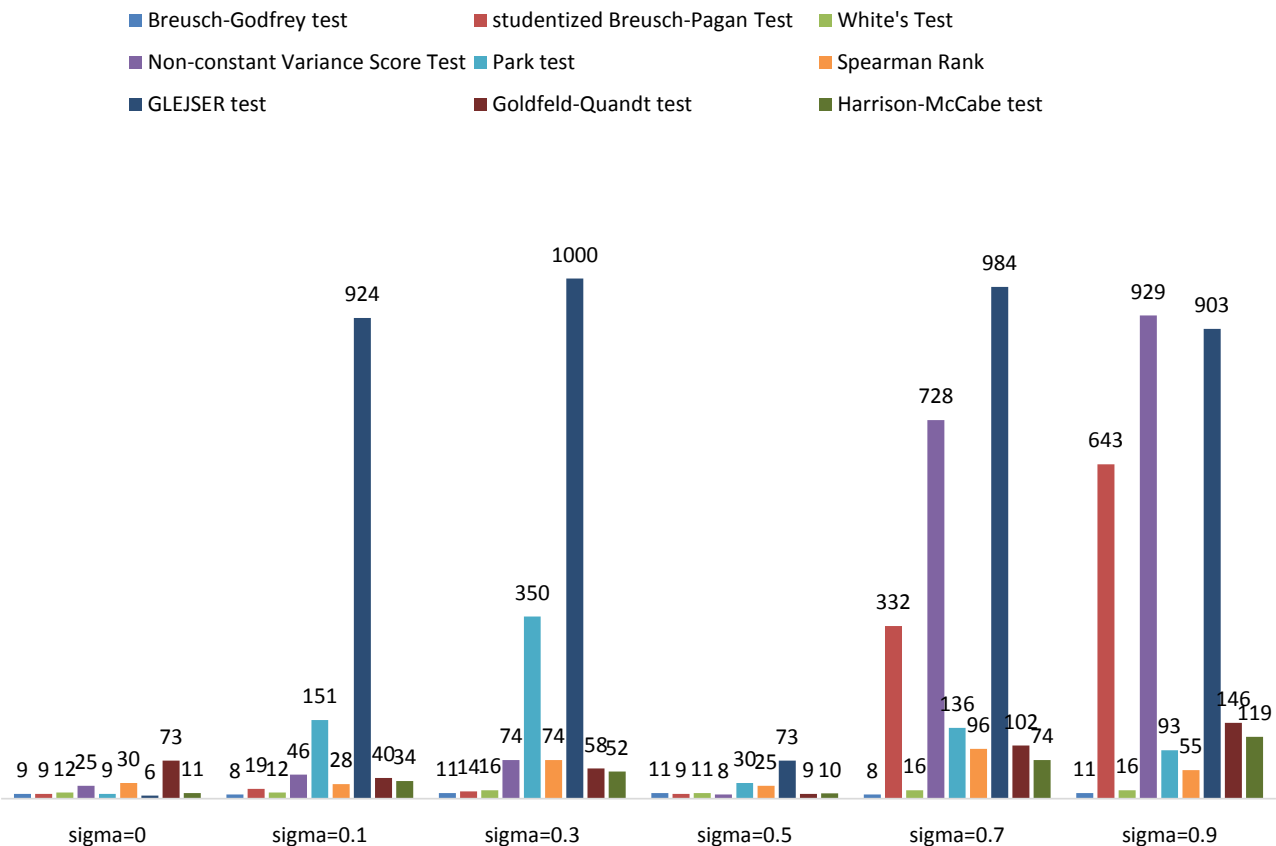
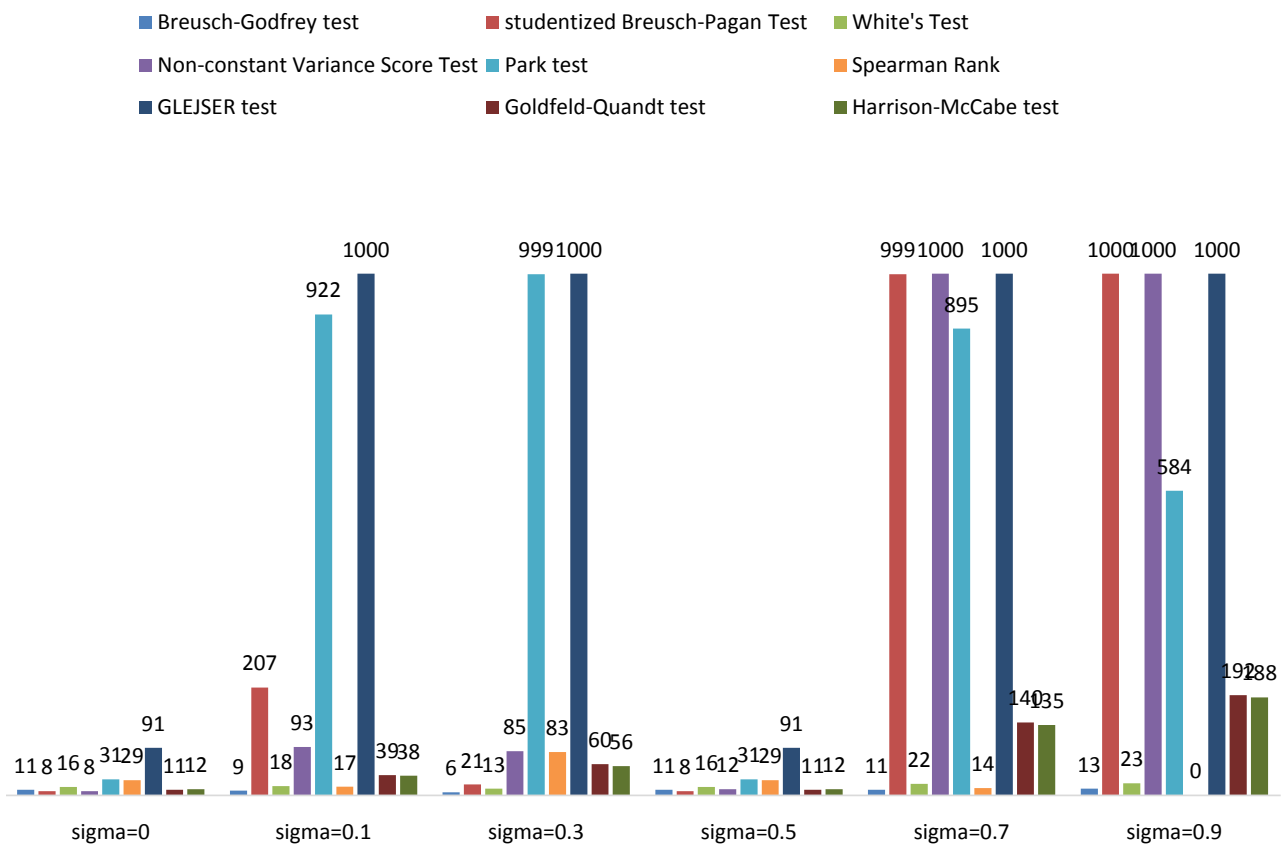


Figure 2. Sample size 100 results when error follows EHS at 1%.

error; Breusch-Godfrey test 1.2%, Studentized Breusch-Pagan test 0.8%, Non-constant Variance Score test 1.1%, White test with 1.9% and so on (see **Table 1**). Hence, Studentized Breusch-Pagan test the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 999, 1000, 98 and 998 corrected tests results at 0.1, 0.3, 0.5 & 0.7 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 99.9%, 100%, 9.8% and 99.8% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5 or 0.7 at sample size of 200. However, Non-constant Variance Score test with 99.9% outperformed the celebrated Glejser test 99.8% when sigma = 0.9.

Additionally, considering sample size 500, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 1.2%, Studentized Breusch-Pagan test 0.8%, Non-constant Variance Score test 0.8%, White test with 1.6% and so on (see **Table 1** and **Figure 3**). Hence, Studentized Breusch-Pagan and Non-constant Variance Score tests are the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 1000, 1000, 91, 1000 and 1000 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying



**Figure 3.** Sample size 500 results when error follows EHS at 1%.



that the Glejser test has the highest rate of type I error of 100%, 100%, 9.1%, 100% and 100% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 500. Interestingly, Studentized Breusch-Pagan recorded 100% performance at sigma = 0.9 also, Non-constant Variance Score test recorded 100% performance at sigma = 0.7 & 0.9. Hence, Non-constant Variance Score test is also best at sigma 0.7 and 0.9 also, Studentized Breusch-Pagan is best at sigma = 0.9.

Lastly, considering sample size 1000, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 1.0%, Studentized Breusch-Pagan test 1.5%, Non-constant Variance Score test 1.1%, White test 1.2%, Goldfeld-Quandt test 0.7% and so on (see Table 1 and Figure 4). Hence, Goldfeld-Quandt test is the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest rate of type I error of 100%, 100%, 88%, 100% and 100% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 100%, 100%, 8.8%, 100% and 100% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 1000. Interestingly, Studentized Breusch-Pagan and Non-constant Variance Score tests recorded 100% performances at sigma = 0.7 & 0.9 also, Park test recorded 100% performance at sigma = 0.7. Hence, Studentized Breusch-Pagan and Non-constant Variance Score tests are also best at sigma 0.7 and 0.9 while Park test is also best at sigma = 0.7.

Table 2 presents the frequency of tests of significance at 5% level after 1000 replications for errors that follow EHS. As observed from the simulation results



Figure 4. Sample size 1000 results when error follows EHS at 1%

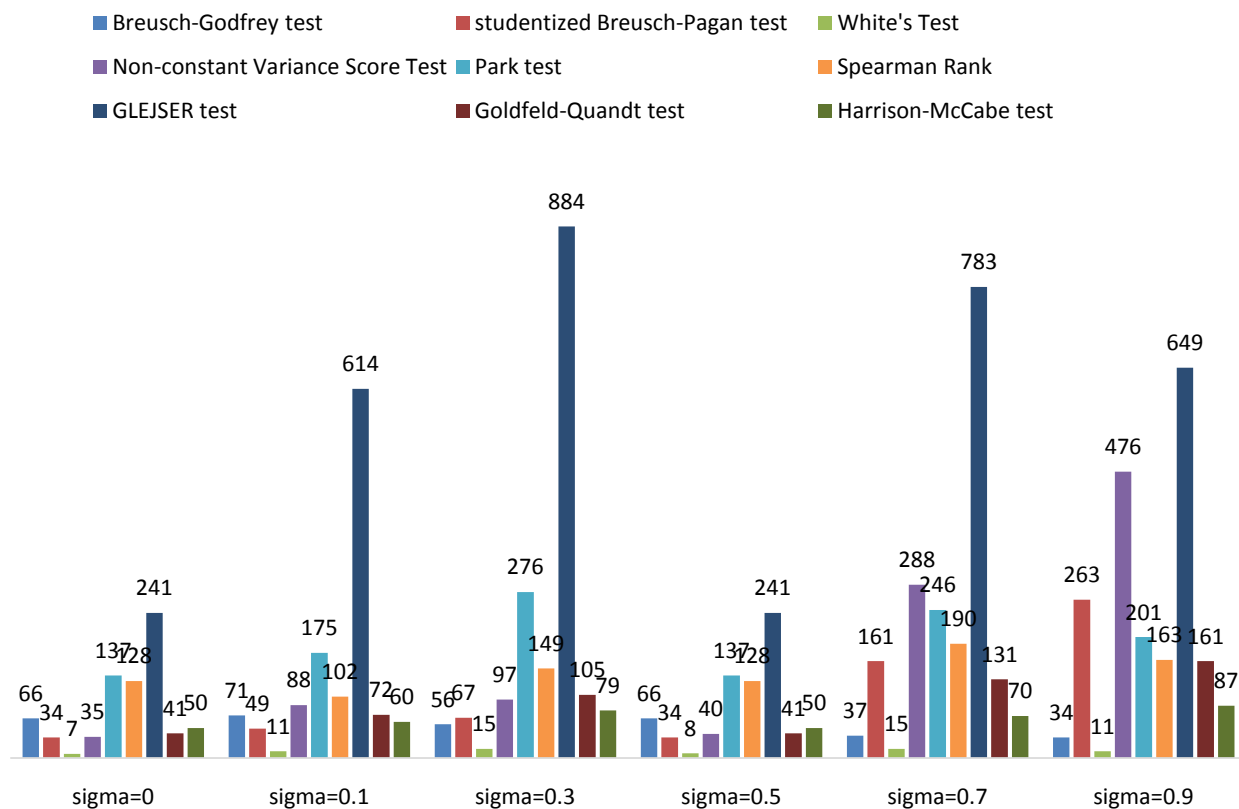
**Table 2.** Performance of the tests when error follows exponential heteroscedastic structure at 5%.

significance level	sample size ( $n$ )	level of hetero	Breusch-Godfrey test	studentized Breusch-Pagan test	White's Test	Non-constant Variance Score Test	Park test	Spearman Rank	GLEJSER test	Goldfeld-Quandt test	Harrison-McCabe test
5%	$n = 30$	sigma = 0	66	34	7	35	137	128	241	41	50
5%	$n = 30$	sigma = 0.1	71	49	11	88	175	102	614	72	60
5%	$n = 30$	sigma = 0.3	56	67	15	97	276	149	884	105	79
5%	$n = 30$	sigma = 0.5	66	34	8	40	137	128	241	41	50
5%	$n = 30$	sigma = 0.7	37	161	15	288	246	190	783	131	70
5%	$n = 30$	sigma = 0.9	34	263	11	476	201	163	649	161	87
5%	$n = 50$	sigma = 0	47	40	28	43	151	139	257	61	59
5%	$n = 50$	sigma = 0.1	57	68	35	82	269	118	819	87	78
5%	$n = 50$	sigma = 0.3	56	68	39	126	410	187	988	103	94
5%	$n = 50$	sigma = 0.5	47	40	33	40	151	139	257	61	59
5%	$n = 50$	sigma = 0.7	47	285	29	502	274	209	964	171	112
5%	$n = 50$	sigma = 0.9	49	493	29	740	231	172	877	199	163
5%	$n = 100$	sigma = 0	57	48	55	128	48	139	40	273	58
5%	$n = 100$	sigma = 0.1	48	94	58	154	455	76	973	103	101
5%	$n = 100$	sigma = 0.3	51	78	61	171	720	177	1000	118	112
5%	$n = 100$	sigma = 0.5	58	48	48	43	139	128	273	57	63
5%	$n = 100$	sigma = 0.7	49	600	63	834	410	156	1000	178	157
5%	$n = 100$	sigma = 0.9	47	888	50	968	269	83	998	230	206
5%	$n = 200$	sigma = 0	55	63	68	63	150	123	301	59	55
5%	$n = 200$	sigma = 0.1	62	183	65	146	732	77	1000	128	123
5%	$n = 200$	sigma = 0.3	51	66	61	181	965	180	1000	137	133
5%	$n = 200$	sigma = 0.5	55	63	63	69	150	123	301	59	55
5%	$n = 200$	sigma = 0.7	49	940	49	986	713	84	1000	198	191
5%	$n = 200$	sigma = 0.9	47	999	48	1000	463	40	1000	254	251
5%	$n = 500$	sigma = 0	50	55	57	55	143	136	293	45	44
5%	$n = 500$	sigma = 0.1	42	484	55	229	976	52	1000	101	98
5%	$n = 500$	sigma = 0.3	50	83	51	184	1000	185	1000	141	136
5%	$n = 500$	sigma = 0.5	50	55	56	49	143	136	293	45	44
5%	$n = 500$	sigma = 0.7	65	999	72	1000	991	15	1000	224	224
5%	$n = 500$	sigma = 0.9	58	1000	71	1000	872	11	1000	281	278
5%	$n = 1000$	sigma = 0	35	67	47	50	142	116	283	49	49
5%	$n = 1000$	sigma = 0.1	40	801	60	342	1000	30	1000	107	108
5%	$n = 1000$	sigma = 0.3	51	124	62	214	1000	181	1000	129	128
5%	$n = 1000$	sigma = 0.5	35	67	50	52	142	116	283	49	49
5%	$n = 1000$	sigma = 0.7	38	1000	66	1000	1000	0	1000	229	223

\*Frequency of test significance after 1000 replications.

in **Table 2** and **Figure 5**, the OLS model was not contaminated with level heteroscedasticity, but yet tested with the nine various tests of heteroscedasticity to detect the true nature of the various test statistics at sample size of  $n = 30$ . At no presence of heteroscedasticity, the following tests returned the following rate in percent of type two error; Goldfeld-Quandt test 4.1%, Studentized Breusch-Pagan test 3.4%, White test with 0.7%, and so on. Making the White test the best in terms of detecting no presence of heteroscedasticity (*i.e.*  $\sigma = 0$ ). While the model was infused with the level of heteroscedasticity (*i.e.*  $\sigma = 0.1, 0.3, 0.5, 0.7$  &  $0.9$ ) at sample size of 30. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 614, 884, 241, 783 and 649 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 61.4%, 88.4%, 24.1%, 78.3% and 64.9% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 30.

Moreover, considering sample size 50, at no presence of heteroscedasticity ( $\sigma = 0$ ), the following tests returned the following rate in percent of type two error; Studentized Breusch-Pagan test 4.7%, Non-constant Variance Score test 4.3%, White test 2.8% and so on (see **Table 2**). Hence, White test the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 819, 988, 257, 964 and 877 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels

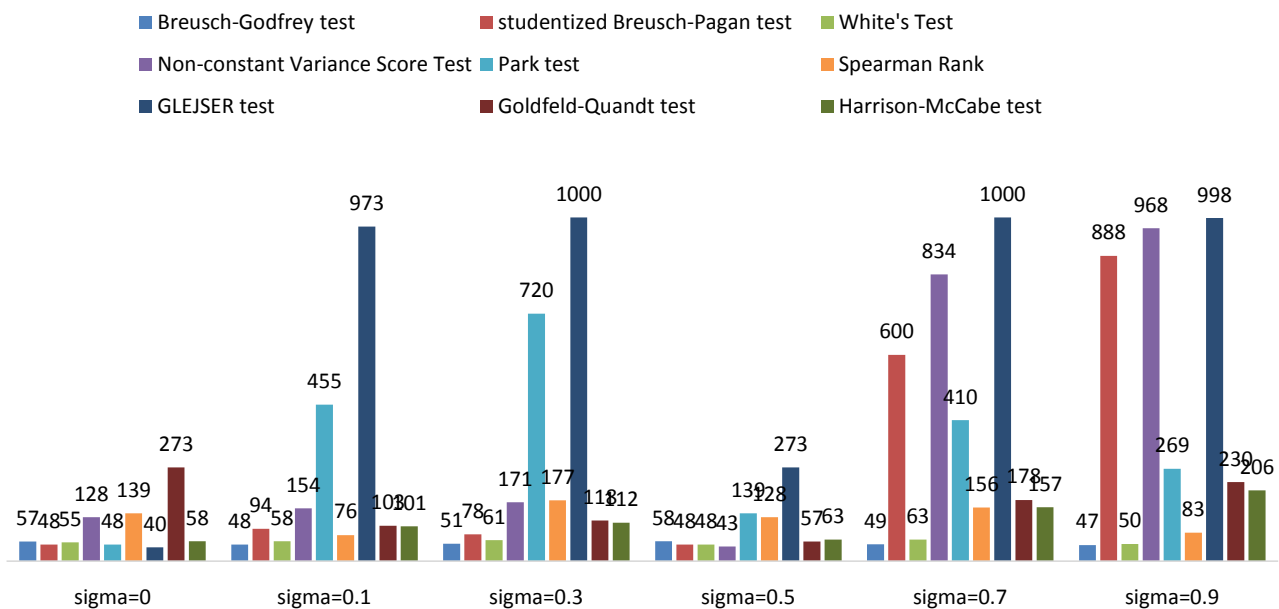


**Figure 5.** Sample size 30 results when error follows EHS at 5%.

respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 81.9%, 98.8%, 25.7%, 96.4% and 87.7% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 50.

In addition, considering sample size 100, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 5.7%, Studentized Breusch-Pagan test 4.8%, Park test 4.8%, White test 5.5%, Glejser test 4.0% and so on (see **Table 2** and **Figure 6**). Hence, the celebrated White test at sigma = 0 was outperformed by Glejser test being the best in terms of detecting no presence of heteroscedasticity when sample size is 100. The Glejser test also returned the highest returned, the highest presence of heteroscedasticity with 973, 1000, 273, 1000 and 998 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 97.3%, 100%, 27.3%, 100% and 99.8% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 100.

Furthermore, considering sample size 200, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 5.5%, Studentized Breusch-Pagan test 6.3%, Non-constant Variance Score test 6.3%, White test 6.8%, Harrison-McCabe test 5.5% and so on (see **Table 2**). Hence, Harrison-McCabe and Breusch-Godfrey tests are the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 1000, 1000, 301, 1000 and 1000 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 100%, 100%, 30.1%, 100%



**Figure 6.** Sample size 100 results when error follows EHS at 5%.

and 100% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 200. Interestingly, Non-constant Variance Score test recorded 100% performance at sigma 0.9. Hence, Non-constant Variance Score test is also best at sigma 0.9.

Additionally, considering sample size 500, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 5.0%, Studentized Breusch-Pagan test 5.5%, Non-constant Variance Score test 5.8%, White test with 5.7%, Harrison-McCabe 4.4% and so on (see Table 2 and Figure 7). Hence, Harrison-McCabe test is the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 1000, 1000, 293, 1000 and 1000 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 100%, 100%, 29.3%, 100% and 100% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 500. Interestingly, Studentized Breusch-Pagan recorded 100% performance at sigma = 0.9 also, Non-constant Variance Score test recorded 100% performance at sigma = 0.7 & 0.9. Hence, Non-constant Variance Score test is also best at sigma 0.7 and 0.9 also, Studentized Breusch-Pagan is best at sigma = 0.9.

Lastly, considering sample size 1000, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 3.5%, Studentized Breusch-Pagan test 6.7%, White test 4.7% and so on (see Table 2 and Figure 8). Hence, Breusch-Godfrey test is the best in terms of detecting no presence of heteroscedasticity. The Glejser test returned the highest returned, the highest presence of heteroscedasticity with 1000, 1000,

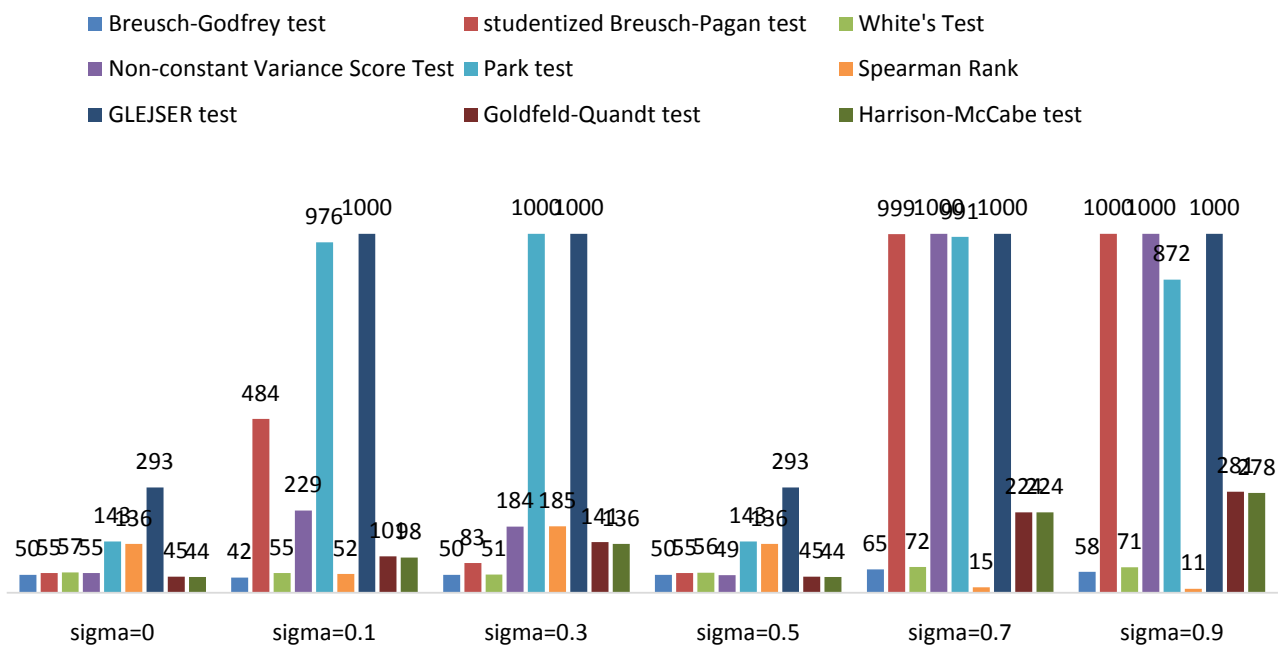


Figure 7. Sample size 500 results when error follows EHS at 5%.

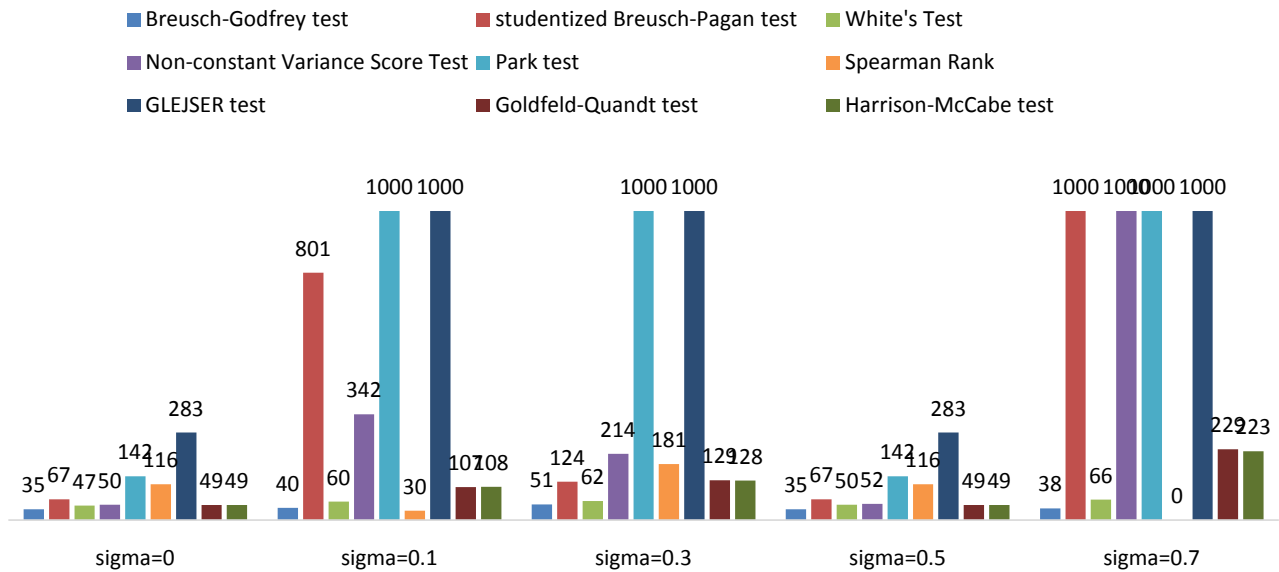


Figure 8. Sample size 1000 results when error follows EHS at 5%.

283, 1000 and 1000 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Glejser test has the highest rate of type I error of 100%, 100%, 28.3%, 100% and 100% which makes Glejser the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 1000. Interestingly, Studentized Breusch-Pagan and Non-constant Variance Score tests recorded 100% performances at sigma = 0.7 & 0.9 also, Park test recorded 100% performance at sigma 0.1, 0.3 and 0.7. Hence, Studentized Breusch-Pagan and Non-constant Variance Score tests are also best at sigma 0.7 and 0.9 while Park test is also best at sigma 0.1, 0.3 and 0.7.

### 5.3. Performance of the Tests When Error Follows Linear Heteroscedastic Structure (LHS)

Table 3 and Table 4 present the frequency of tests of significance at 1% and 5% levels respectively after 1000 replications for errors that follow LHS. As observed from the simulation results in Table 4 and Figure 9, the OLS model was not contaminated with level heteroscedasticity, but yet tested with the nine various tests of heteroscedasticity to detect the true nature of the various test statistics at sample size of  $n = 30$ . At no presence of heteroscedasticity, the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 0.3%, Spearman Rank tests 0.6%, White test 2.2%, Harrison-McCabe test 0% and so on. Making the Harrison-McCabe test the best in terms of detecting no presence of heteroscedasticity (*i.e.* sigma = 0). While the model was infused with the level of heteroscedasticity (*i.e.* sigma = 0.1, 0.3, 0.5, 0.7 & 0.9) at sample size of 30. The Park test returned the highest in the highest presence of heteroscedasticity with 295, 353, 353, 353 and 354 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 29.5%, 35.3%, 35.3%, 35.3%

**Table 3.** Performance of the tests when error follows LHS at 1%.

significance level	sample size ( <i>n</i> )	Heteroscedasticity Test									
		level of hetero	Breusch-Godfrey test	Stuldentized Breusch-Pagan test	White's Test	Non-constant Variance Score Test	Park test	Spearman Rank	GLEJSER test	Goldfeld-Quandt test	Harrison-McCabe test
1%	<i>n</i> = 30	sigma = 0	3	34	22	10	59	6	18	9	0
1%	<i>n</i> = 30	sigma = 0.1	22	52	71	115	295	86	8	40	0
1%	<i>n</i> = 30	sigma = 0.3	21	123	106	141	353	116	2	53	1
1%	<i>n</i> = 30	sigma = 0.5	21	125	108	143	353	117	2	56	1
1%	<i>n</i> = 30	sigma = 0.7	23	128	107	141	353	119	2	56	1
1%	<i>n</i> = 30	sigma = 0.9	23	125	104	142	354	118	2	56	1
1%	<i>n</i> = 50	sigma = 0	6	37	40	6	68	13	16	12	3
1%	<i>n</i> = 50	sigma = 0.1	28	87	87	217	352	108	11	71	0
1%	<i>n</i> = 50	sigma = 0.3	30	153	195	259	389	123	8	84	4
1%	<i>n</i> = 50	sigma = 0.5	30	144	187	266	391	123	7	83	4
1%	<i>n</i> = 50	sigma = 0.7	30	146	191	266	392	123	7	83	4
1%	<i>n</i> = 50	sigma = 0.9	30	142	190	266	393	124	7	84	3
1%	<i>n</i> = 100	sigma = 0	9	30	25	6	73	9	11	10	12
1%	<i>n</i> = 100	sigma = 0.1	64	132	140	326	474	114	13	105	25
1%	<i>n</i> = 100	sigma = 0.3	63	193	300	388	522	138	10	114	26
1%	<i>n</i> = 100	sigma = 0.5	64	191	302	391	522	138	10	116	26
1%	<i>n</i> = 100	sigma = 0.7	64	190	304	393	520	138	10	116	26
1%	<i>n</i> = 100	sigma = 0.9	64	191	304	394	520	138	10	116	27
1%	<i>n</i> = 200	sigma = 0	8	42	36	11	98	17	12	16	19
1%	<i>n</i> = 200	sigma = 0.1	117	218	231	461	643	137	9	128	28
1%	<i>n</i> = 200	sigma = 0.3	122	283	467	538	687	154	11	143	29
1%	<i>n</i> = 200	sigma = 0.5	120	288	470	540	690	154	11	144	29
1%	<i>n</i> = 200	sigma = 0.7	120	280	473	542	690	155	11	144	29
1%	<i>n</i> = 200	sigma = 0.9	120	272	474	543	689	155	11	144	29
1%	<i>n</i> = 500	sigma = 0	8	31	29	8	91	11	11	12	16
1%	<i>n</i> = 500	sigma = 0.1	379	434	400	676	887	154	8	154	22
1%	<i>n</i> = 500	sigma = 0.3	378	515	709	745	914	167	9	166	28
1%	<i>n</i> = 500	sigma = 0.5	375	514	722	751	915	166	9	167	29
1%	<i>n</i> = 500	sigma = 0.7	375	524	724	753	914	166	9	166	29
1%	<i>n</i> = 500	sigma = 0.9	375	525	724	753	915	165	9	166	29
1%	<i>n</i> = 1000	sigma = 0	15	29	31	11	88	7	10	9	12
1%	<i>n</i> = 1000	sigma = 0.1	676	732	618	870	977	165	6	162	23
1%	<i>n</i> = 1000	sigma = 0.3	691	742	901	912	984	173	5	174	25
1%	<i>n</i> = 1000	sigma = 0.5	691	749	911	914	984	173	5	171	26
1%	<i>n</i> = 1000	sigma = 0.7	691	745	912	914	983	173	5	171	26
1%	<i>n</i> = 1000	sigma = 0.9	691	747	913	915	983	173	5	171	26

\*Frequency of test significance after 1000 replications.

**Table 4.** Performance of the tests when error follows LHS at 5%.

significance level	sample size ( <i>n</i> )	level of hetero	Heteroscedasticity Test								
			Breusch-Godfrey test	studentized Breusch-Pagan test	White's Test	Non-constant Variance Score Test	Park test	Spearman Rank	GLEJSER test	Goldfeld-Quandt test	Harrison-McCabe test
5%	<i>n</i> = 30	sigma = 0	34	137	128	35	241	41	66	50	7
5%	<i>n</i> = 30	sigma = 0.1	89	228	184	211	516	168	48	122	14
5%	<i>n</i> = 30	sigma = 0.3	93	335	264	253	602	195	46	139	12
5%	<i>n</i> = 30	sigma = 0.5	93	327	255	257	605	196	46	139	10
5%	<i>n</i> = 30	sigma = 0.7	93	335	249	257	608	198	47	139	10
5%	<i>n</i> = 30	sigma = 0.9	93	333	251	260	608	198	47	138	10
5%	<i>n</i> = 50	sigma = 0	40	151	139	43	257	61	47	59	28
5%	<i>n</i> = 50	sigma = 0.1	100	287	235	326	593	187	48	166	33
5%	<i>n</i> = 50	sigma = 0.3	104	374	359	383	662	224	43	183	40
5%	<i>n</i> = 50	sigma = 0.5	103	367	365	382	659	223	43	186	40
5%	<i>n</i> = 50	sigma = 0.7	103	359	368	382	656	222	45	186	40
5%	<i>n</i> = 50	sigma = 0.9	103	363	367	383	655	222	45	187	40
5%	<i>n</i> = 100	sigma = 0	48	139	128	40	273	57	58	63	55
5%	<i>n</i> = 100	sigma = 0.1	164	354	305	446	706	194	55	187	69
5%	<i>n</i> = 100	sigma = 0.3	174	429	471	498	747	218	45	204	69
5%	<i>n</i> = 100	sigma = 0.5	172	436	482	498	748	220	46	206	68
5%	<i>n</i> = 100	sigma = 0.7	172	439	483	498	750	219	47	206	69
5%	<i>n</i> = 100	sigma = 0.9	172	436	479	499	749	219	47	206	68
5%	<i>n</i> = 200	sigma = 0	63	150	123	63	301	59	55	55	68
5%	<i>n</i> = 200	sigma = 0.1	266	491	405	580	858	216	45	212	78
5%	<i>n</i> = 200	sigma = 0.3	278	543	611	635	878	244	45	236	80
5%	<i>n</i> = 200	sigma = 0.5	277	552	622	638	873	248	45	239	80
5%	<i>n</i> = 200	sigma = 0.7	277	547	625	639	873	249	45	243	83
5%	<i>n</i> = 200	sigma = 0.9	278	543	627	639	874	249	45	243	83
5%	<i>n</i> = 500	sigma = 0	55	143	136	55	293	45	50	44	57
5%	<i>n</i> = 500	sigma = 0.1	551	719	595	764	969	215	43	215	66
5%	<i>n</i> = 500	sigma = 0.3	562	750	809	816	980	232	46	231	67
5%	<i>n</i> = 500	sigma = 0.5	564	751	820	825	980	235	47	233	66
5%	<i>n</i> = 500	sigma = 0.7	566	748	822	826	980	235	46	233	65
5%	<i>n</i> = 500	sigma = 0.9	565	752	822	827	980	235	46	233	66
5%	<i>n</i> = 1000	sigma = 0	67	142	116	50	283	49	35	49	47
5%	<i>n</i> = 1000	sigma = 0.1	821	916	782	915	997	253	36	251	66
5%	<i>n</i> = 1000	sigma = 0.3	831	921	946	939	999	263	43	256	70
5%	<i>n</i> = 1000	sigma = 0.5	830	915	951	943	999	261	45	257	70
5%	<i>n</i> = 1000	sigma = 0.7	830	915	951	943	999	262	45	258	70
5%	<i>n</i> = 1000	sigma = 0.9	830	910	952	944	999	262	45	259	71

\*Frequency of test significance after 1000 replications.



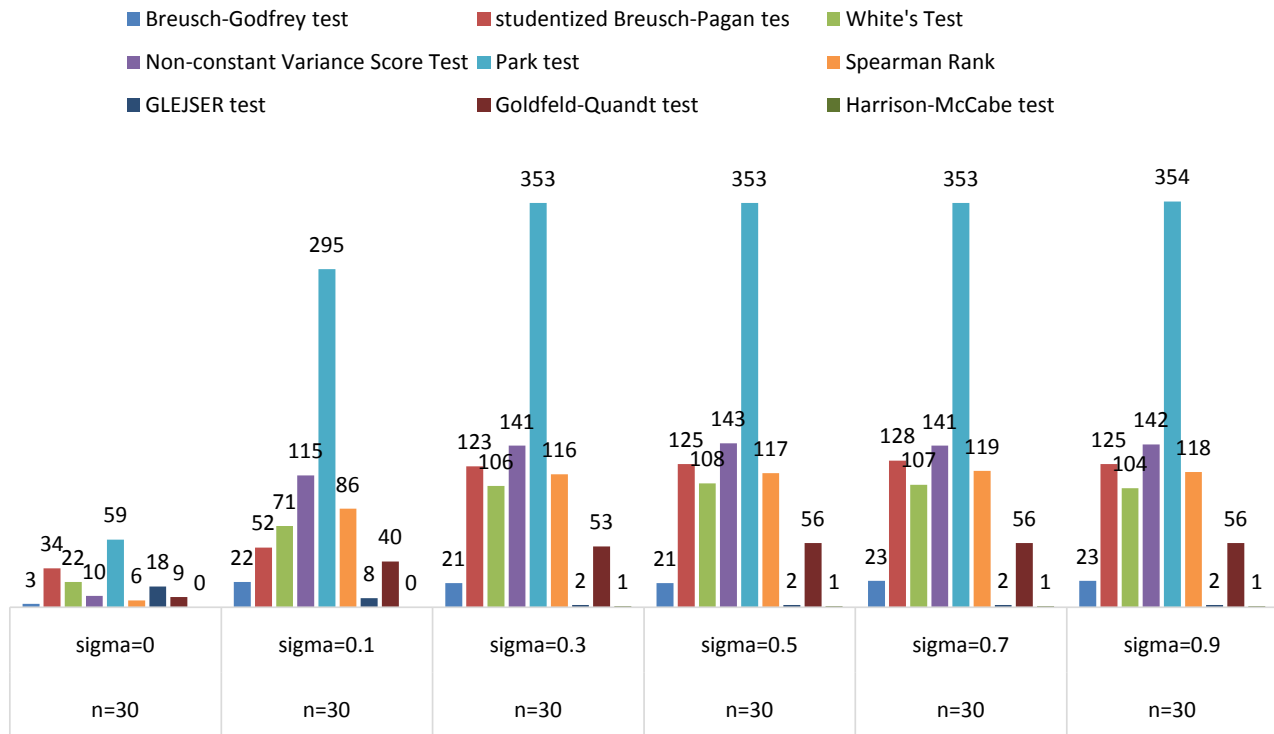


Figure 9. Sample size 30 results when error follows LHS at 1%.

and 35.4% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 30.

Moreover, considering sample size 50, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 0.6%, Spearman Rank tests 1.3%, White test 4.0%, Harrison-McCabe test 0.3% and so on (see Table 3 and Figure 10). Hence, Harrison-McCabe test the best in terms of detecting no presence of heteroscedasticity. The Park test returned the highest returned, the highest presence of heteroscedasticity with 352, 389, 391, 392 and 393 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 35.2%, 38.9%, 39.1%, 39.2% and 39.3% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 50.

In addition, considering sample size 100, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 0.9%, Non-constant Variance Score test 0.6%, Spearman Rank test 1.3%, White test 2.5%, Harrison-McCabe test 1.2% and so on (see Table 3 and Figure 11). Hence, the celebrated Harrison-McCabe test at sigma = 0 was outperformed by Non-constant Variance Score test being the best in terms of detecting no presence of heteroscedasticity when sample size is 100. The Park test returned the highest returned, the highest presence of heteroscedasticity with 474, 522, 522, 520 and 520 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the

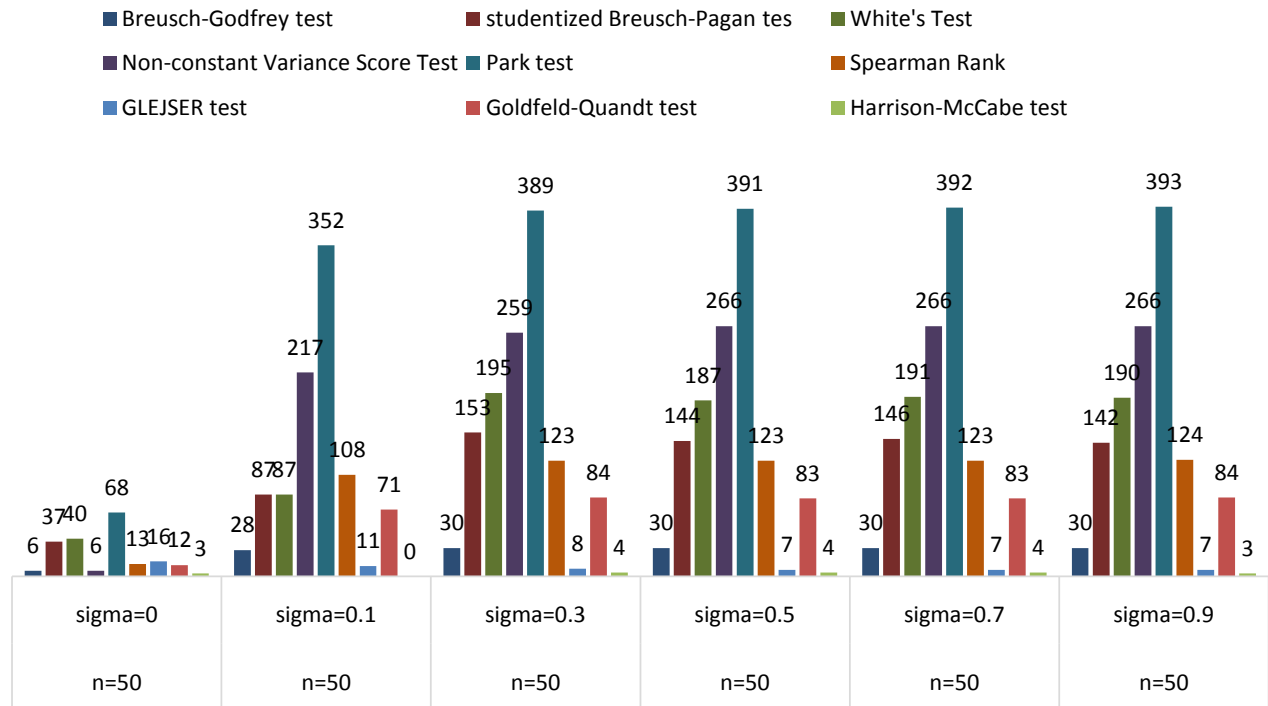


Figure 10. Sample size 50 results when error follows LHS at 1%.

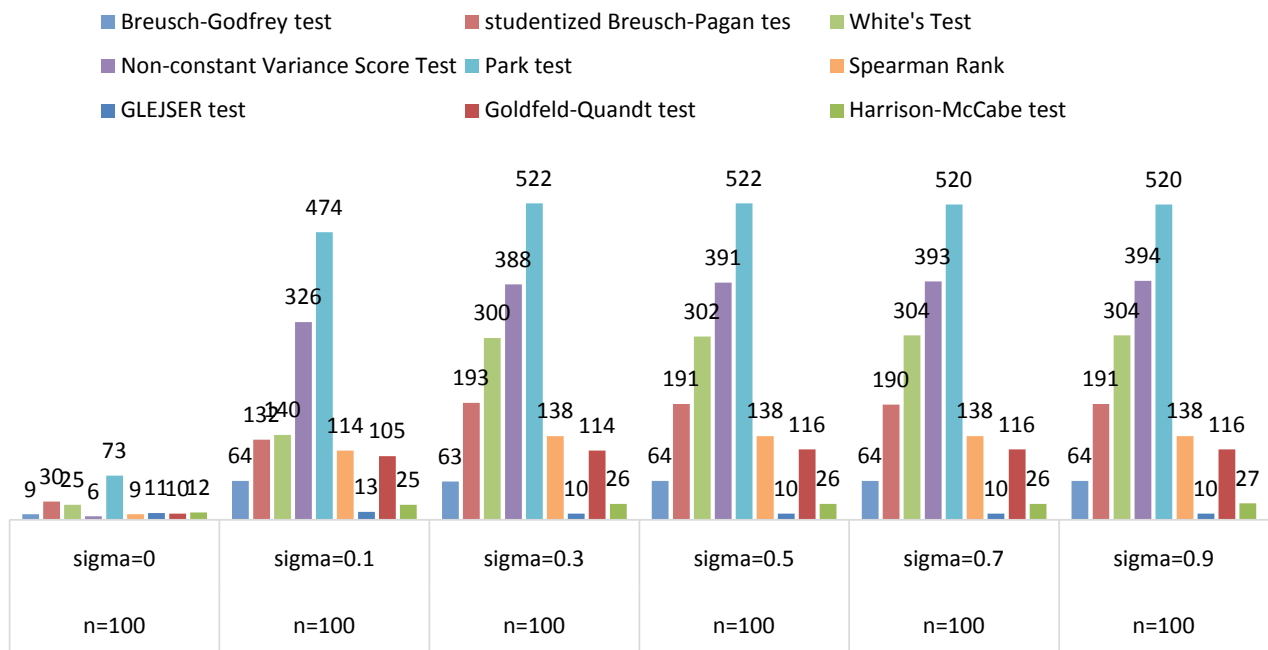


Figure 11. Sample size 100 results when error follows LHS at 1%.

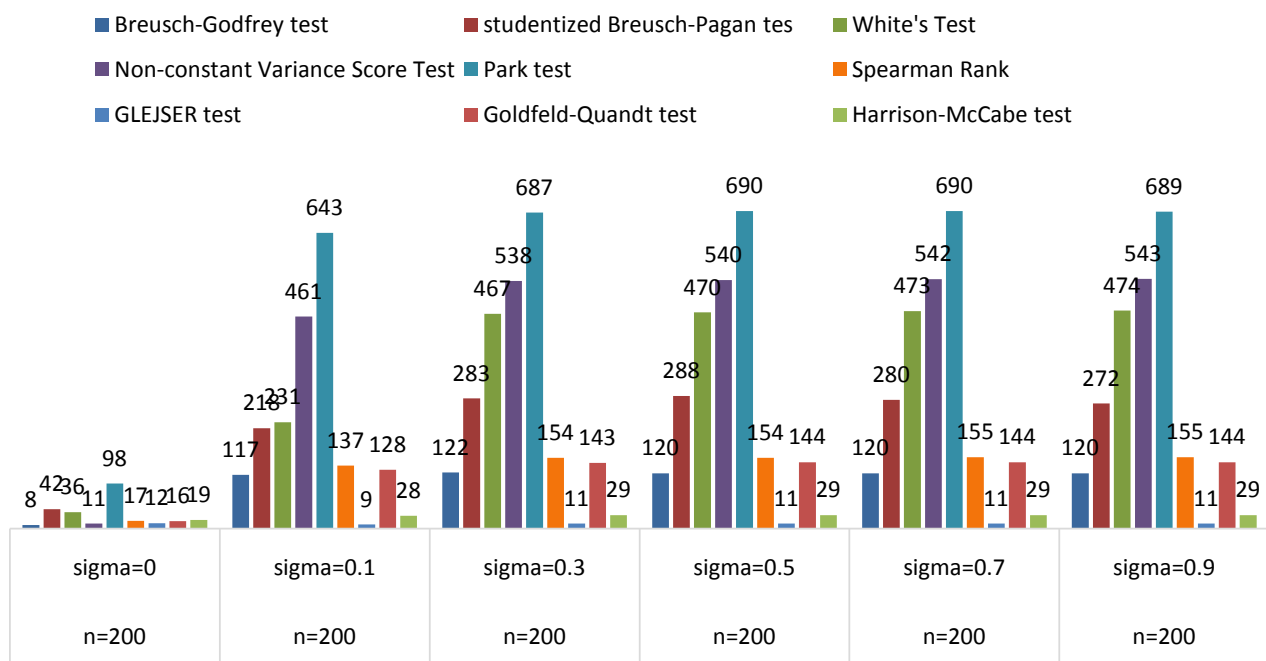
Park test has the highest rate of type I error of 47.4%, 52.2%, 52.2%, 52.0% and 52.0% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 100.

Furthermore, considering sample size 200, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two

error; Breusch-Godfrey test 0.8%, Non-constant Variance Score test 1.1%, Spearman Rank test 1.7%, White test 2.9%, Harrison-McCabe test 1.9% and so on (see **Table 3** and **Figure 12**). Hence, the earlier celebrated Harrison-McCabe and Non-constant Variance Score test at  $\sigma = 0$  were outperformed by Breusch-Godfrey test being the best in terms of detecting no presence of heteroscedasticity when sample size is 200. The Park test returned the highest returned, the highest presence of heteroscedasticity with 643, 687, 690, 690 and 689 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 64.3%, 68.7%, 69.0%, 69.0% and 68.9% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 200.

Additionally, considering sample size 500, at no presence of heteroscedasticity ( $\sigma = 0$ ), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 0.8%, Non-constant Variance Score test 0.8%, Spearman Rank test 1.11%, White test 2.9%, Harrison-McCabe test 1.6% and so on (see **Table 3** and **Figure 13**). Hence, Non-constant Variance Score and Breusch-Godfrey tests are the best in terms of detecting no presence of heteroscedasticity when sample size is 500. The Park test returned the highest returned, the highest presence of heteroscedasticity with 887, 914, 915, 914 and 915 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 88.7%, 91.4%, 91.5%, 91.4% and 91.5% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 500.

Lastly, considering sample size 1000, at no presence of heteroscedasticity ( $\sigma = 0$ ), the following tests returned the following rate in percent of type two



**Figure 12.** Sample size 200 results when error follows LHS at 1%.

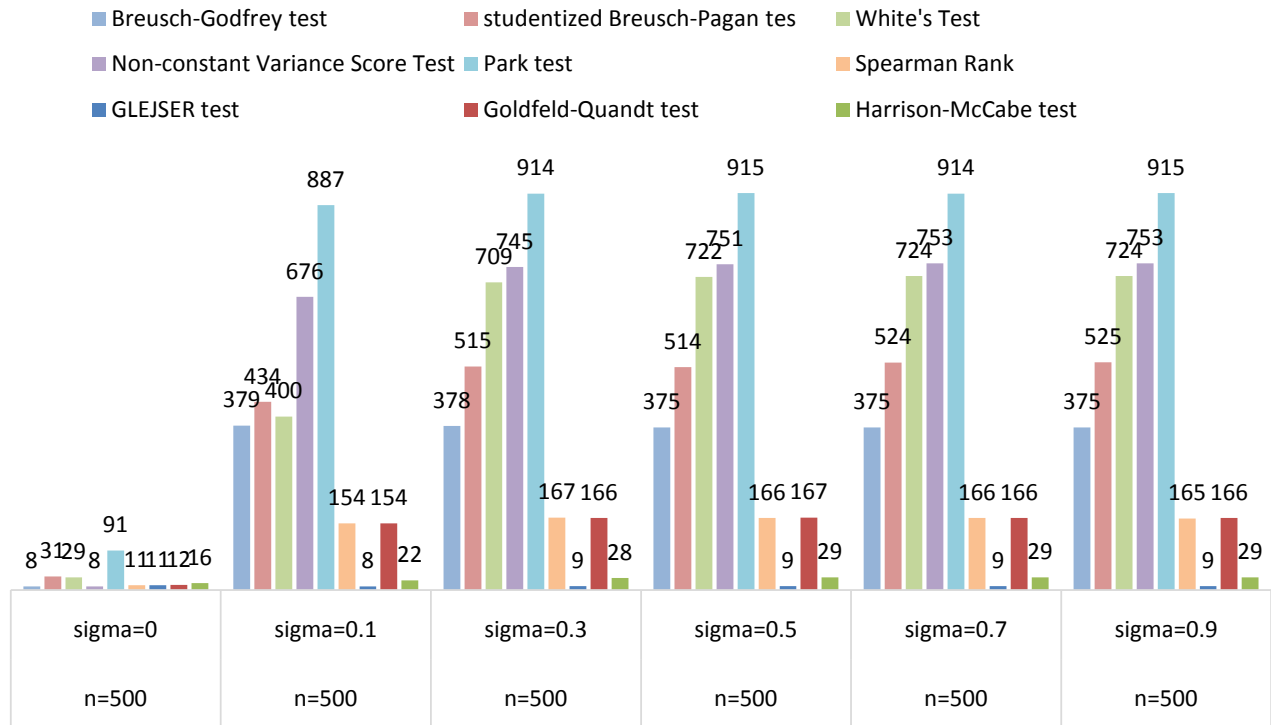


Figure 13. Sample size 500 results when error follows LHS at 1%.

error; Breusch-Godfrey test 1.5%, Non-constant Variance Score test 1.1%, Spearman Rank test 0.7%, White test 3.1%, Harrison-McCabe test 1.2% and so on (see Table 3 and Figure 14). Hence, Spearman rank test is the best in terms of detecting no presence of heteroscedasticity when sample size is 1000. The Park test returned the highest returned, the highest presence of heteroscedasticity with 977, 984, 984, 983 and 983 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 97.7%, 98.4%, 98.4%, 98.3% and 98.3% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 1000.

Table 4 presents the frequency of tests of significance at 5% level after 1000 replications for errors that follow LHS. As observed from the simulation results in Table 4 and Figure 15, the OLS model was not contaminated with level heteroscedasticity, but yet tested with the nine various tests of heteroscedasticity to detect the true nature of the various test statistics at sample size of  $n = 30$ . At no presence of heteroscedasticity, the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 3.4%, Spearman Rank tests 4.1%, White test 12.8% Harrison-McCabe test 0.7% and so on. Making the Harrison-McCabe test the best in terms of detecting no presence of heteroscedasticity (i.e. sigma = 0). While the model was infused with the level of heteroscedasticity (i.e. sigma = 0.1, 0.3, 0.5, 0.7 & 0.9) at sample size of 30. The Park test returned the highest returned, the highest presence of heteroscedasticity with 516, 602, 605, 608 and 608 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9

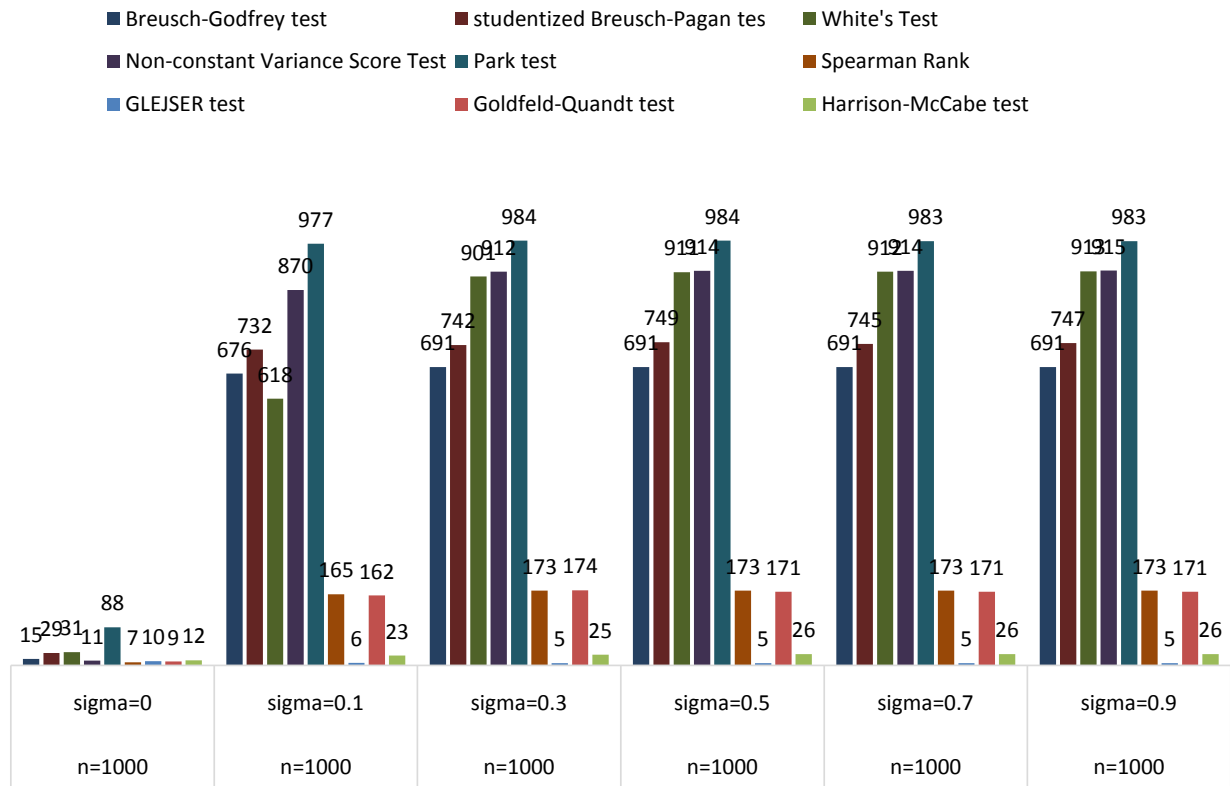


Figure 14. Sample size 1000 results when error follows LHS at 1%.

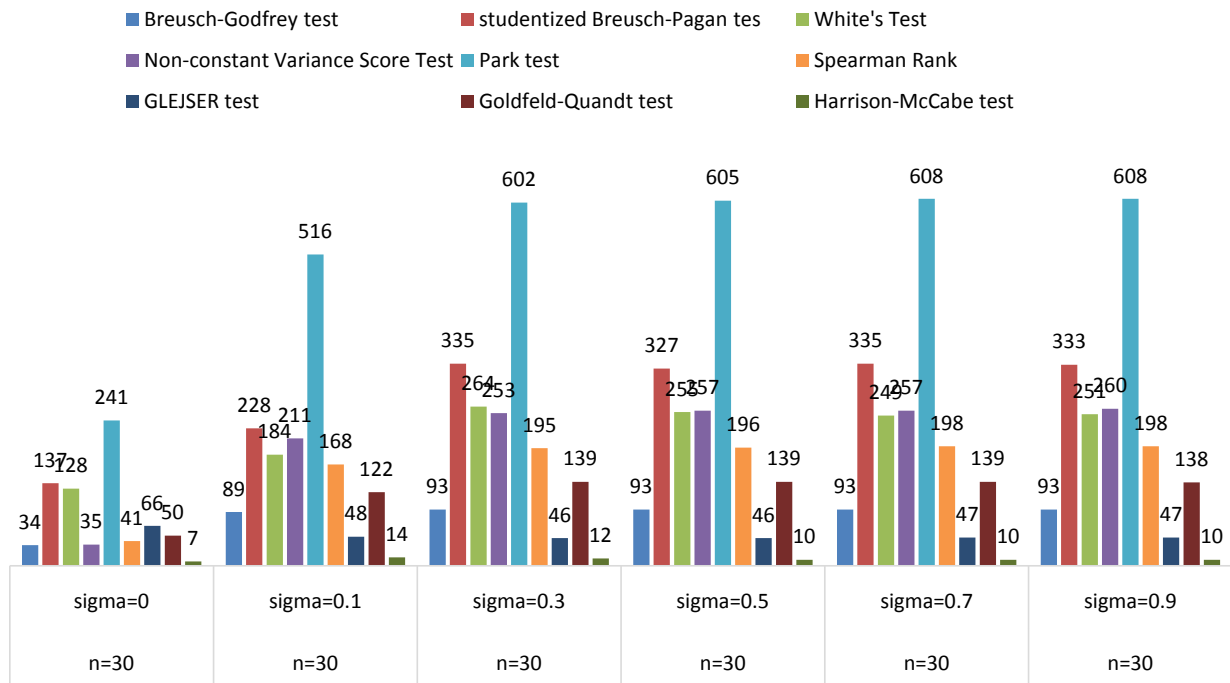


Figure 15. Sample size 30 results when error follows LHS at 5%.

sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 51.6%, 60.2%, 60.5%, 60.8% and

60.8% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 30.

Moreover, considering sample size 50, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 4.0%, Spearman Rank tests 6.1%, White test 13.9%, Harrison-McCabe test 2.8% and so on (see Table 4). Hence, Harrison-McCabe test the best in terms of detecting no presence of heteroscedasticity. The Park test returned the highest returned, the highest presence of heteroscedasticity with 593, 662, 659, 656 and 655 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 59.3%, 66.2%, 65.9%, 65.6% and 65.5% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 50.

In addition, considering sample size 100, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 4.8%, Non-constant Variance Score test 4.0%, Spearman Rank test 5.7%, White test 13.9%, Harrison-McCabe test 5.5% and so on (see Table 4 and Figure 16). Hence, the celebrated Harrison-McCabe test at sigma = 0 was outperformed by Non-constant Variance Score test being the best in terms of detecting no presence of heteroscedasticity when sample size is 100. The Park test returned the highest returned, the highest presence of heteroscedasticity with 706, 747, 748, 750 and 749 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 70.6%, 74.7%, 74.8%, 75.0%

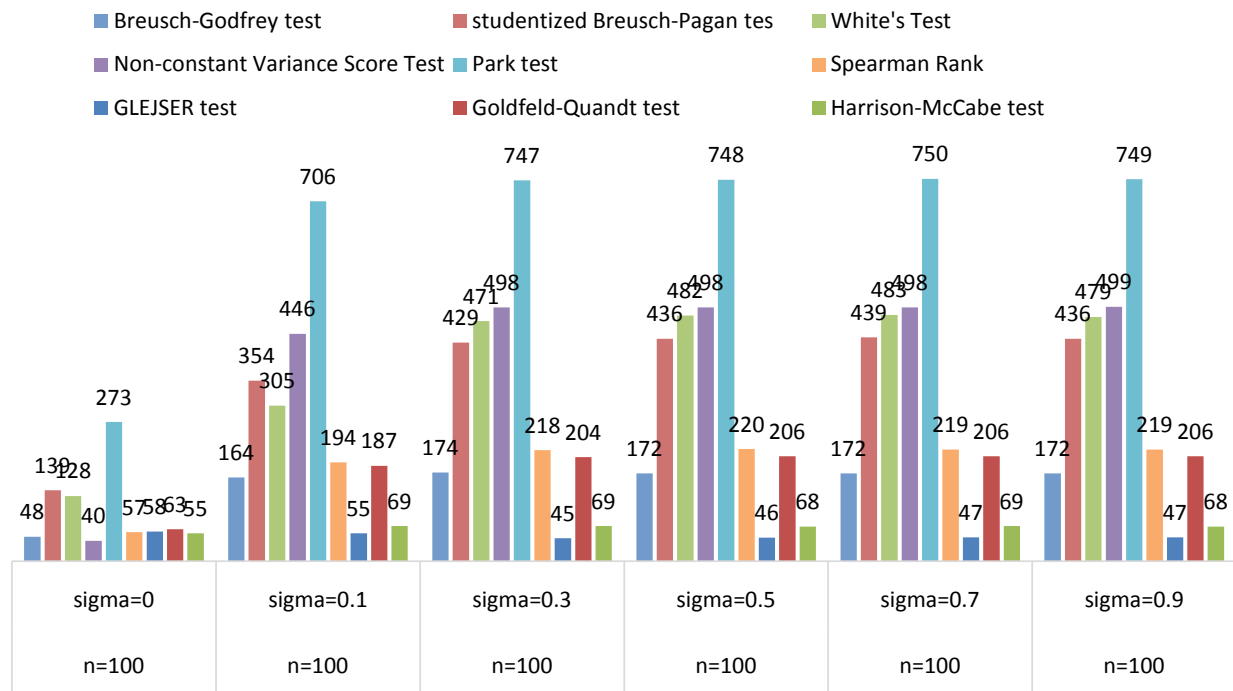


Figure 16. Sample size 100 results when error follows LHS at 5%.

and 74.9% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 100.

Additionally, considering sample size 500, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 5.5%, Non-constant Variance Score test 5.5%, Spearman Rank test 4.5%, White test 13.6%, Harrison-McCabe test 5.7% and so on (see **Table 4** and **Figure 17**). Hence, Spearman rank test is the best in terms of detecting no presence of heteroscedasticity when sample size is 500. The Park test returned the highest returned, the highest presence of heteroscedasticity with 969, 980, 980, 980 and 980 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the Park test has the highest rate of type I error of 96.9%, 98.0%, 8.0%, 8.0% and 8.0% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 500.

Lastly, considering sample size 1000, at no presence of heteroscedasticity (sigma = 0), the following tests returned the following rate in percent of type two error; Breusch-Godfrey test 6.7%, Non-constant Variance Score test 5.0%, Spearman Rank test 4.9%, White test 11.6%, Harrison-McCabe test 4.7% and so on (see **Table 4** and **Figure 18**). Hence, Spearman rank test is the best in terms of detecting no presence of heteroscedasticity when sample size is 1000. The Park test returned the highest returned, the highest presence of heteroscedasticity with 997, 999, 999, 999 and 999 corrected tests results at 0.1, 0.3, 0.5, 0.7 & 0.9 sigma levels respectively out of every 1000 replications, thus implying that the

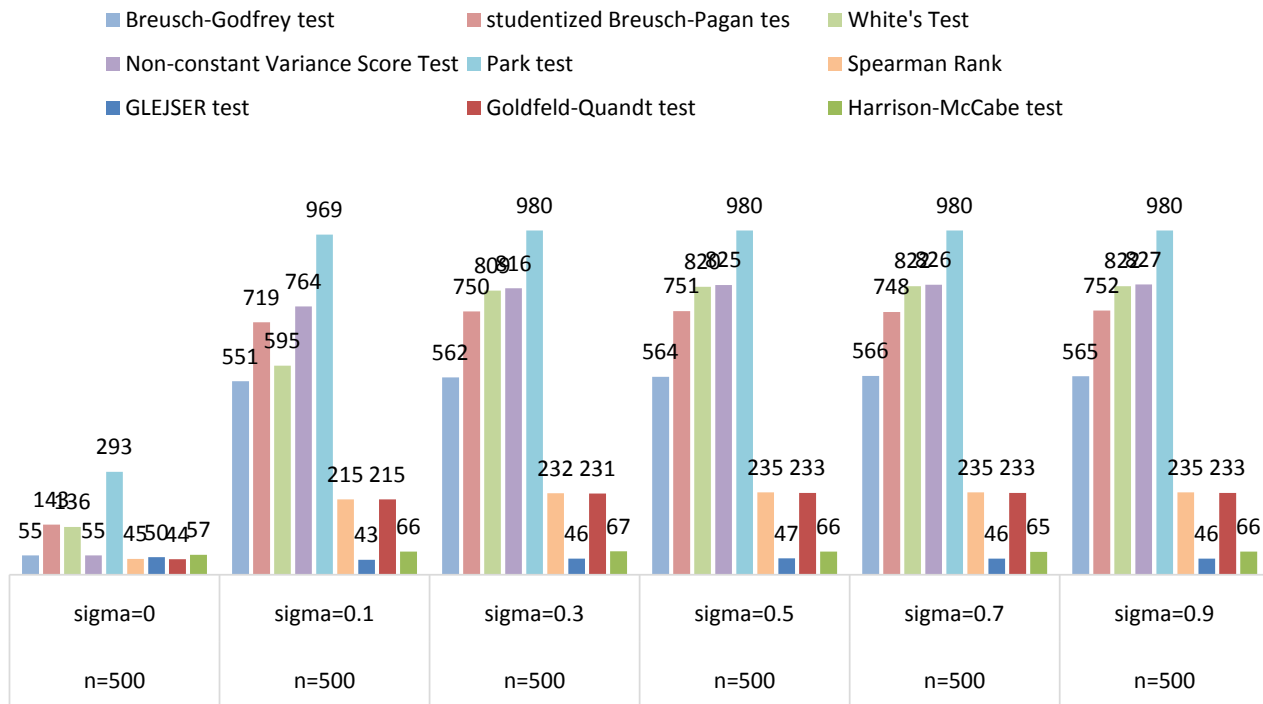


Figure 17. Sample size 500 results when error follows LHS at 5%.

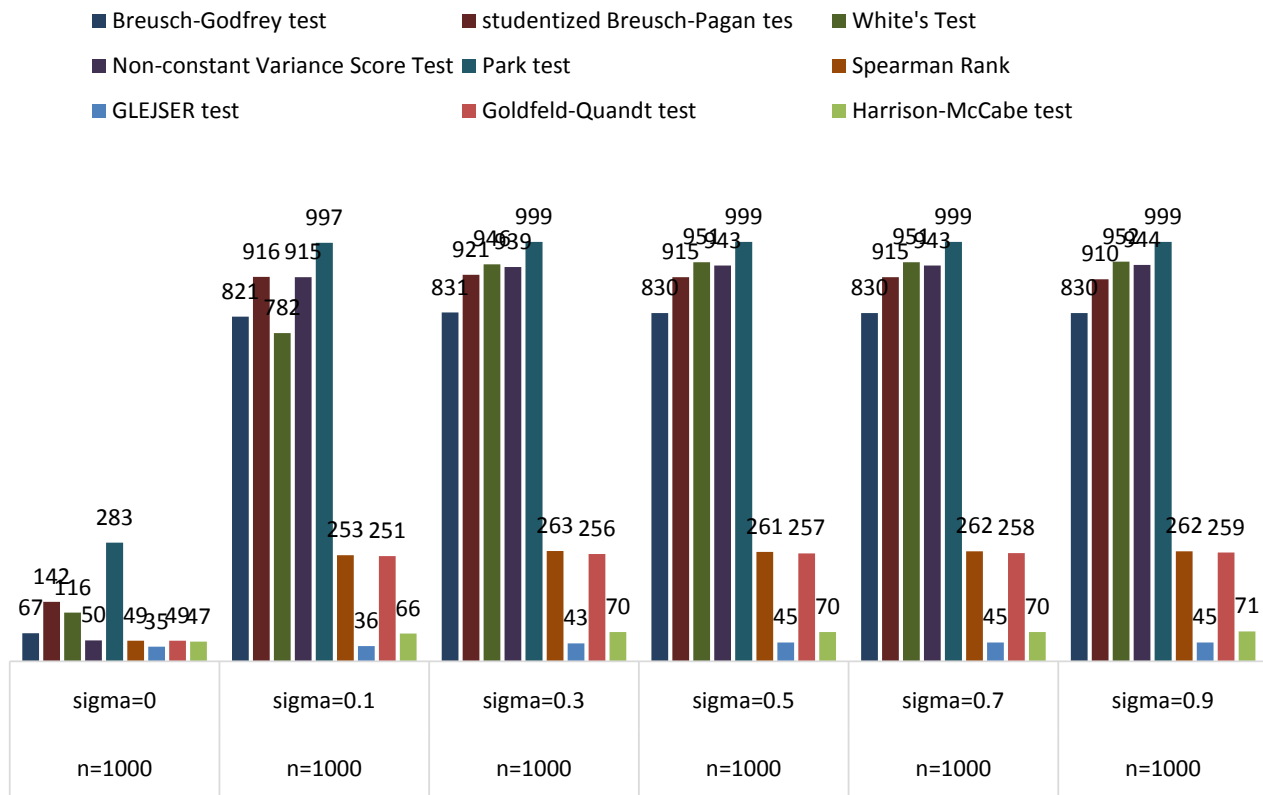


Figure 18. Sample size 1000 results when error follows LHS at 5%.

Park test has the highest rate of type I error of 99.7%, 99.9%, 99.9%, 99.9% and 99.9% which makes Park the best test when sigma is 0.1, 0.3, 0.5, 0.7 or 0.9 at sample size of 1000.

## 6. Summary, Conclusions and Recommendations

### 6.1. Summary

This study focuses on comparative analysis that determines the asymptotic behaviour of some selected statistical tests for homoscedasticity assumption by Monte Carlo simulations, and seeks to recommend the best statistical test for detecting heteroscedasticity in a multiple linear regression scenario with varying variances.

### 6.2. Conclusions

The aim and objectives of the study have been principally accomplished. The analyses show the comparative results of the null hypothesis such that homoscedasticity assumption is upheld under four error structures for sample sizes; 30, 50, 100, 200, 500 and 1000 when  $\sigma^2$  was varied as follows; 0, 0.1, 0.3, 0.5, 0.7 and 0.9.

In the analyses, the homoscedasticity assumption was tested under four different error distributions namely; Exponential Heteroscedastic Structure (EHS), Linear Heteroscedastic Structure (LHS), Square-Root Heteroscedastic Structure



(SHS) and Quadratic Heteroscedastic Structure (QHS), for different sample sizes as the  $\sigma^2$  was varied. However, two major error structures results were reported in course of our study namely Exponential Heteroscedastic Structure (EHS) and Linear Heteroscedastic Structure (LHS). The Linear Heteroscedastic Structure (LHS) results were adopted as it explains the results in Square-Root Heteroscedastic Structure (SHS) and Quadratic Heteroscedastic Structure (QHS). Following our findings, **Table 5** and **Table 6** present the summary of the best tests across all board.

As observed from **Table 5** and **Table 6**:

- when the OLS model was not contaminated with level heteroscedasticity (*i.e.* sigma = 0) White test returned the best test at sample size 30 and 50 for errors following EHS while;
- Harrison-McCabe test returned the best for errors following LHS. Still on sample size 30 and 50;
- when the model was infused with the level of heteroscedasticity (*i.e.* sigma = 0.1, 0.3, 0.5, 0.7 & 0.9), the Glejser test and Park test returned the best test for EHS and LHS respectively at sigma = 0.1, 0.3, 0.5, 0.7 & 0.9 except for EHS at sigma = 0.9 Non-constant Variance Score test returned best (0.01 level only);

**Table 5.** Summary of the best tests (Sample Size 30, 50 and 100).

N	Sigma	EHS			LHS		
		0.01	0.05	0.1	0.01	0.05	0.1
30	0	White	White	White	Harrison-McCabe	Harrison-McCabe	Harrison-McCabe
	0.1	Glejser	Glejser	Glejser	Park	Park	Park
	0.3	Glejser	Glejser	Glejser	Park	Park	Park
	0.5	Glejser	Glejser	Glejser	Park	Park	Park
	0.7	Glejser	Glejser	Glejser	Park	Park	Park
	0.9	Non-constant Variance Score	Glejser	Glejser	Park	Park	Park
50	0	White	White	White	Harrison-McCabe	Harrison-McCabe	Harrison-McCabe
	0.1	Glejser	Glejser	Glejser	Park	Park	Park
	0.3	Glejser	Glejser	Glejser	Park	Park	Park
	0.5	Glejser	Glejser	Glejser	Park	Park	Park
	0.7	Glejser	Glejser	Glejser	Park	Park	Park
	0.9	Non-constant Variance Score	Glejser	Glejser	Park	Park	Park
100	0	Glejser	Glejser	Glejser	Non-constant Variance Score	Non-constant Variance Score	Non-constant Variance Score
	0.1	Glejser	Glejser	Glejser	Park	Park	Park
	0.3	Glejser	Glejser	Glejser	Park	Park	Park
	0.5	Glejser	Glejser	Glejser	Park	Park	Park
	0.7	Glejser	Glejser	Glejser	Park	Park	Park
	0.9	Non-constant Variance Score	Glejser	Glejser	Park	Park	Park

**Table 6.** Summary of the best tests (Sample Size 200, 500 and 1000).

<i>n</i>	Sigma	EHS			LHS		
		0.01	0.05	0.1	0.01	0.05	0.1
200	0	Studentized Breusch-Pagan	Harrison-McCabe/Breusch-Godfrey	Breusch-Godfrey	Breusch-Godfrey	Spearman rank	Spearman rank
	0.1	Glejser	Glejser	Glejser	Park	Park	Park
	0.3	Glejser	Glejser	Glejser	Park	Park	Park
	0.5	Glejser	Glejser	Glejser	Park	Park	Park
	0.7	Glejser	Glejser	Glejser	Park	Park	Park
	0.9	Non-constant Variance Score	Glejser/Non-constant Variance Score	Glejser/Non-constant Variance Score	Park	Park	Park
500	0	Studentized Breusch-Pagan/Non-constant Variance Score	Harrison-McCabe	Harrison-McCabe	Non-constant Variance Score/Breusch-Godfrey	Spearman rank	Spearman rank
	0.1	Glejser	Glejser	Glejser	Park	Park	Park
	0.3	Glejser	Glejser	Glejser/Park	Park	Park	Park
	0.5	Glejser	Glejser	Glejser	Park	Park	Park
	0.7	Glejser/Non-constant Variance Score	Glejser/Non-constant Variance Score	Glejser/Non-constant Variance Score	Park	Park	Park
	0.9	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan	Park	Park	Park
1000	0	Goldfeld-Quandt	Breusch-Godfrey	Breusch-Godfrey	Spearman rank	Spearman rank	Spearman rank
	0.1	Glejser	Glejser/Park	Glejser/Park	Park	Park	Park
	0.3	Glejser/Park	Glejser/Park	Glejser/Park	Park	Park	Park
	0.5	Glejser	Glejser	Glejser	Park	Park	Park
	0.7	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan/Park	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan/Park	Park	Park	Park
	0.9	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan	Glejser/Non-constant Variance Score/Studentized Breusch-Pagan/Park	Park	Park	Park

- Furthermore, at sample size 100 Glejser test returned the best test when sigma is 0, 0.1, 0.3, 0.5 & 0.7 and Non-constant Variance Score test returned best when sigma is 0.9 for EHS. While;
- Park test returned the best test when sigma is 0.1, 0.3, 0.5, 0.7 & 0.9 and Non-constant Variance Score test returned best when sigma is 0 for LHS;
- In addition, at sample size 200 the Glejser test and Park test returned the best test for EHS and LHS respectively when sigma = 0.1, 0.3, 0.5, 0.7 & 0.9 except

for EHS at  $\sigma = 0.9$  for 0.01 level Non-constant Variance Score test returned best;

- However, when  $\sigma = 0$  (no heteroscedasticity) the following tests returned best: Studentized Breusch-Pagan (EHS at 0.01 level); Harrison-McCabe/Breusch-Godfrey (EHS at 0.05 level); Breusch-Godfrey (EHS at 0.1 level); Breusch-Godfrey (LHS at 0.01); and Spearman rank (LHS at 0.05 & 0.1);
- Moreover, at sample size 500 the Glejser test and Park test returned the best test for EHS and LHS respectively when  $\sigma = 0.1, 0.3, 0.5, 0.7$  &  $0.9$  also Non-constant Variance Score at  $\sigma = 0.7$  or  $0.9$  and Studentized Breusch-Pagan returned best when  $\sigma = 0.9$  for EHS;
- However, when  $\sigma = 0$  (no heteroscedasticity) the following tests returned best: Studentized Breusch-Pagan/Non-constant Variance Score (EHS at 0.01 level); Harrison-McCabe (EHS at 0.05 level); Harrison-McCabe (EHS at 0.1 level); Non-constant Variance Score/Breusch-Godfrey (LHS at 0.01); and Spearman rank (LHS at 0.05 & 0.1);
- Lastly, at sample size 1000 the Glejser test and Park test returned the best test for EHS and LHS respectively when  $\sigma = 0.1, 0.3, 0.5, 0.7$  &  $0.9$ ;
- Also the following test returned best: Park (EHS at  $\sigma = 0.3$  and 0.01 level); Non-constant Variance Score/Studentized Breusch-Pagan (EHS;  $\sigma = 0.7$  &  $0.9$  at 0.01 level); Park (EHS;  $\sigma = 0.3$  &  $0.9$  at 0.05 and 0.1 levels); and Non-constant Variance Score/Studentized Breusch-Pagan/Park (EHS;  $\sigma = 0.7$  &  $0.9$  at 0.05 and 0.1 levels);
- However, when  $\sigma = 0$  (no heteroscedasticity) the following tests returned best: Goldfeld-Quandt (EHS at 0.01 level); Breusch-Godfrey (EHS at 0.05 and 0.1 levels); and Spearman rank (LHS at 0.01, 0.05 & 0.1 levels).

### 6.3. Recommendations

From the aforementioned, the following are recommended:

1) White and Harrison-McCabe tests should be employed to check for homoscedasticity in EHS and LHS respectively for sample size 30 and 50 (**Small samples**).

2) others can be employed as follows Glejser (EHS at  $n = 100$ ), Non-constant Variance Score (LHS at  $n = 100$ ), Studentized Breusch-Pagan (EHS at  $n = 200$ , 0.01 level), Harrison-McCabe/Breusch-Godfrey (EHS at  $n = 200$ , 0.05 level), Breusch-Godfrey (EHS at  $n = 200$ , 0.1 level), Breusch-Godfrey (LHS at  $n = 200$ , 0.01), Spearman rank (LHS at  $n = 200$ , 0.05 & 0.1), Studentized Breusch-Pagan/Non-constant Variance Score (EHS at  $n = 500$ , 0.01 level) (**Moderate samples**).

3) Harrison-McCabe (EHS at  $n = 500$ , 0.05 level), Harrison-McCabe (EHS at  $n = 500$ , 0.1 level), Non-constant Variance Score/Breusch-Godfrey (LHS at  $n = 500$ , 0.01), Spearman rank (LHS at  $n = 500$ , 0.05 & 0.1), Goldfeld-Quandt (EHS at  $n = 1000$ , 0.01 level), Breusch-Godfrey (EHS at  $n = 1000$ , 0.05 and 0.1 levels), Spearman rank (LHS at  $n = 1000$ , 0.01, 0.05 & 0.1 levels) (**Large samples**).

4) Glejser and Park tests should be employed to check for heteroscedasticity in

EHS and LHS respectively.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Bartlett, M.S. and Kendall, D.G. (1946) The Statistical Analysis of Variance Heterogeneity and the Logarithmic Transformation. *Supplement to the Journal of the Royal Statistical Society*, **8**, 128-138. <https://doi.org/10.2307/2983618>
- [2] Weerahandi, S. (1995) ANOVA under Unequal Error Variances. *Biometrics*, **51**, 589-599. <https://doi.org/10.2307/2532947>
- [3] Ogunleye, T.A., Olaleye, M.O. and Solomon, A.Z. (2014) Econometric Modelling of Commercial Banks' Expenditure on the Sources of Profit Maximization in Nigeria. *Scholars Journal of Economics, Business and Management*, **1**, 276-290.
- [4] Lim, T.S. and Loh, W.Y. (1996) A Comparison of Tests of Equality of Variances. *Computational Statistics and Data Analysis*, **22**, 287-301. [https://doi.org/10.1016/0167-9473\(95\)00054-2](https://doi.org/10.1016/0167-9473(95)00054-2)
- [5] White, H. (1980) A Heteroscedasticity Consistent Covariance Matrix and Direct Test for Heteroscedasticity. *Econometrica*, **48**, 817-838. <https://doi.org/10.2307/1912934>
- [6] Graybill, F.A. (1976) *The Theory and Applications of the Linear Model*. Duxbury Press, London.
- [7] Conover, W.J., Johnson, M.E. and Johnsons, M.M. (1981) A Comparative Study of Tests for Homogeneity of Variances with Applications to the Outer Continental Shelf Bidding Data. *Technometrics*, **23**, 351-361. <https://doi.org/10.1080/00401706.1981.10487680>
- [8] Feingold, A. (1992) Sex Differences in Variability in Intellectual Abilities: A New Look at an Old Controversy. *Review of Educational Research*, **62**, 61-84. <https://doi.org/10.3102/00346543062001061>
- [9] Hedges, L.V. and Friedman, L. (1993) Gender Differences in Variability in Intellectual Abilities: A Re-Analysis of Feingold's Results. *Review of Educational Research*, **63**, 94-105. <https://doi.org/10.3102/00346543063001094>
- [10] Carroll, R.J. (1982) Tests for Regression Parameters in Power Transformation Models. *Scandinavian Journal of Statistics*, **9**, 217-222.
- [11] Mendes, M. (2003) The Comparison of Levene, Bartlett, Neymann-Pearson and Bartlett 2 Tests in Terms of Actual Type I Error Rates. *Journal of Agriculture Sciences*, **9**, 143-146. [https://doi.org/10.1501/Tarimbil\\_0000000782](https://doi.org/10.1501/Tarimbil_0000000782)
- [12] Camdeviren, H. and Mendes, M. (2005) A Simulation Study for Type III Error Rates of Some Variance Homogeneity Tests. *Pakistan Journal of Statistics*, **21**, 223-234.
- [13] Rogan, J.C. and Keselman, H.J. (1977) Is the ANOVA F-Test Robust to Variance Heterogeneity When Sample Sizes Are Equal? An Investigation via a Coefficient of Variation. *American Educational Research Journal*, **14**, 493-498. <https://doi.org/10.3102/00028312014004493>
- [14] Cochran, W.G. and Cox, G.M. (1957) *Experimental Design*. John Wiley and Sons Inc., New York.
- [15] Brown, M.B. and Forsythe, A.B. (1974) The Small Sample Behavior of Some Statis-

- tics Which Test the Equality of Several Means. *Technometrics*, **16**, 129-132.  
<https://doi.org/10.1080/00401706.1974.10489158>
- [16] Wilcox, R.R., Charlin, V.L. and Thomson, K.L. (1986) New Monte Carlo Results on the Robustness of the Anova F, W and F\*Statistics. *Communications in Statistics—Simulation and Computation*, **15**, 933-943.  
<https://doi.org/10.1080/03610918608812553>
- [17] Nelson, L.S. (2000) Comparing Two Variances from Normal Populations. *Journal of Quality Technology*, **32**, 79-80. <https://doi.org/10.1080/00224065.2000.11979974>
- [18] Wilcox, R.R. (2002) Comparing Variances of Two Independent Groups. *British Journal of Mathematical and Statistical Psychology*, **55**, 169-175.  
<https://doi.org/10.1348/000711002159635>
- [19] Gupta, A.K., Harrar, S. and Pardo, L. (2004) On Testing Homogeneity of Variances for Non-Normal Models Using Entropy. Department of Mathematics and Statistics, Bowling Green State University, Technical Report, No. 04-11.
- [20] Zar, J.H. (1999) Biostatistical Analysis. Prentice-Hall Inc. Simon and Schuster/A Viacom Company, New Jersey, USA.
- [21] Hsiung, T.C. and Olejnik, S. (1996) Type I Error Rates and Statistical Power for James Second-Order Test and the Univariate F Test in Two-Way ANOVA Models under Heteroscedasticity and/or Non-Normality. *Journal of Experimental Education*, **65**, 57-71. <https://doi.org/10.1080/00220973.1996.9943463>
- [22] Wilcox, R.R. (1995) ANOVA: A Paradigm for Low Power and Misleading Measures of Effect Size. *Review of Educational Research*, **65**, 51-77.  
<https://doi.org/10.3102/00346543065001051>
- [23] Oshima, T.C. and Algina, J. (1992) A SAS Program for Testing the Hypothesis of Equal Means under Heteroscedasticity: James's Second-Order Test. *Educational and Psychological Measurement*, **52**, 117-118.  
<https://doi.org/10.1177/001316449205200116>
- [24] Keselman, H.J., Wilcox, R.R., Othman, A.R. and Fradette, K. (2002) Trimming, Transforming Statistics and Bootstrapping: Circumventing the Biasing Effects of Heteroscedasticity and Non-Normality. *Journal of Modern Applied Statistical Methods*, **1**, 288-309. <https://doi.org/10.22237/jmasm/1036109820>
- [25] Wijngaad, J.B., Kleink, A.M.G., Tank, G. and Konnen, P. (2003) Homogeneity of 20th Century European Daily Temperature and Precipitation Series. *International Journal of Climatology*, **23**, 679-692. <https://doi.org/10.1002/joc.906>
- [26] Pardo, J.A., Pardo, M.C., Vicente, M.L. and Esteban, M.D. (1997) A Statistical Information Theory Approach to Compare the Homogeneity of Several Variances. *Journal of Computational Statistics and Data Analysis*, **24**, 411-416.  
[https://doi.org/10.1016/S0167-9473\(96\)00080-1](https://doi.org/10.1016/S0167-9473(96)00080-1)
- [27] Neyman, J. and Pearson, E.S. (1931) On the Problem of  $k$  Samples. *Bulletin de L'academie Polonaise des Sciences-Serie des Sciences de la Terre*, **Ser. A**, 460-481.
- [28] Layard, M.W.J. (1973) Robust Large Sample Tests for Homogeneity of Variances. *Journal of the American Statistical Association*, **68**, 195-198.  
<https://doi.org/10.1080/01621459.1973.10481363>
- [29] Harvey, A.C. (1976) Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica*, **44**, 461-465. <https://doi.org/10.2307/1913974>
- [30] Goldfeld, S.M. and Quandt, R.E. (1972) Nonlinear Methods in Econometrics. North-Holland Publishing Company, Amsterdam.
- [31] Goldfeld, S.M. and Quandt, R.E. (1965) Some Tests for Heteroscedasticity. *Journal*

*of the American Statistical Association*, **60**, 539-547.

<https://doi.org/10.1080/01621459.1965.10480811>

- [32] Glejser, H. (1969) A New Test for Heteroscedasticity. *Journal of the American Statistical Association*, **64**, 316-323. <https://doi.org/10.1080/01621459.1969.10500976>
- [33] Wiedermann, W., Artner, R. and Eye, A. (2017) Heteroscedasticity as a Basis of Direction Dependence in Reversible Linear Models. *Multivariate Behavioural Research*, **52**, 222-241. <https://doi.org/10.1080/00273171.2016.1275498>

## Appendix 1

### Performance of the Tests when Error follows Quadratic Structure (QHS) at 1%

significance level	sample size (n)	level of hetero	Heteroscedasticity Test								
			Breusch-Godfrey test	studentized Breusch-Pagan test	White's Test	Non-constant Variance Score Test	Park test	Spearman Rank	GLEJSER test	Goldfeld-Quandt test	Harrison-McCabe test
1%	n = 30	sigma = 0	3	34	22	10	59	6	18	9	0
1%	n = 30	sigma = 0.1	22	52	71	115	295	86	8	40	0
1%	n = 30	sigma = 0.3	21	123	106	141	353	116	2	53	1
1%	n = 30	sigma = 0.5	21	125	108	143	353	117	2	56	1
1%	n = 30	sigma = 0.7	23	128	107	141	353	119	2	56	1
1%	n = 30	sigma = 0.9	23	125	104	142	354	118	2	56	1
1%	n = 50	sigma = 0	6	37	40	6	68	13	16	12	3
1%	n = 50	sigma = 0.1	28	87	87	217	352	108	11	71	0
1%	n = 50	sigma = 0.3	30	153	195	259	389	123	8	84	4
1%	n = 50	sigma = 0.5	30	144	187	266	391	123	7	83	4
1%	n = 50	sigma = 0.7	30	146	191	266	392	123	7	83	4
1%	n = 50	sigma = 0.9	30	142	190	266	393	124	7	84	3
1%	n = 100	sigma = 0	9	30	25	6	73	9	11	10	12
1%	n = 100	sigma = 0.1	64	132	140	326	474	114	13	105	25
1%	n = 100	sigma = 0.3	63	193	300	388	522	138	10	114	26
1%	n = 100	sigma = 0.5	64	191	302	391	522	138	10	116	26
1%	n = 100	sigma = 0.7	64	190	304	393	520	138	10	116	26
1%	n = 100	sigma = 0.9	64	191	304	394	520	138	10	116	27
1%	n = 200	sigma = 0	8	42	36	11	98	17	12	16	19
1%	n = 200	sigma = 0.1	117	218	231	461	643	137	9	128	28
1%	n = 200	sigma = 0.3	122	283	467	538	687	154	11	143	29
1%	n = 200	sigma = 0.5	120	288	470	540	690	154	11	144	29
1%	n = 200	sigma = 0.7	120	280	473	542	690	155	11	144	29
1%	n = 200	sigma = 0.9	120	272	474	543	689	155	11	144	29
1%	n = 500	sigma = 0	8	31	29	8	91	11	11	12	16
1%	n = 500	sigma = 0.1	379	434	400	676	887	154	8	154	22
1%	n = 500	sigma = 0.3	378	515	709	745	914	167	9	166	28
1%	n = 500	sigma = 0.5	375	514	722	751	915	166	9	167	29
1%	n = 500	sigma = 0.7	375	524	724	753	914	166	9	166	29
1%	n = 500	sigma = 0.9	375	525	724	753	915	165	9	166	29
1%	n = 1000	sigma = 0	15	29	31	11	88	7	10	9	12

**Continued**

1%	$n = 1000$	$\sigma = 0.1$	676	732	618	870	977	165	6	162	23
1%	$n = 1000$	$\sigma = 0.3$	691	742	901	912	984	173	5	174	25
1%	$n = 1000$	$\sigma = 0.5$	691	749	911	914	984	173	5	171	26
1%	$n = 1000$	$\sigma = 0.7$	691	745	912	914	983	173	5	171	26
1%	$n = 1000$	$\sigma = 0.9$	691	747	913	915	983	173	5	171	26

\*Frequency of test significance after 1000 replications.

**Performance of the Tests when Error follows Quadratic Structure (QHS) at 5%**

significance level	sample size ( $n$ )	level of hetero	Heteroscedasticity Test								
			Breusch-Godfrey test	studentized Breusch-Pagan tes	White's Test	Non-constant Variance Score Test	Park test	Spearman Rank	GLEJSER test	Goldfeld-Quandt test	Harrison-McCabe test
5%	$n = 30$	$\sigma = 0$	34	137	128	35	241	41	66	50	7
5%	$n = 30$	$\sigma = 0.1$	89	228	184	211	516	168	48	122	14
5%	$n = 30$	$\sigma = 0.3$	93	335	264	253	602	195	46	139	12
5%	$n = 30$	$\sigma = 0.5$	93	327	255	257	605	196	46	139	10
5%	$n = 30$	$\sigma = 0.7$	93	335	249	257	608	198	47	139	10
5%	$n = 30$	$\sigma = 0.9$	93	333	251	260	608	198	47	138	10
5%	$n = 50$	$\sigma = 0$	40	151	139	43	257	61	47	59	28
5%	$n = 50$	$\sigma = 0.1$	100	287	235	326	593	187	48	166	33
5%	$n = 50$	$\sigma = 0.3$	104	374	359	383	662	224	43	183	40
5%	$n = 50$	$\sigma = 0.5$	103	367	365	382	659	223	43	186	40
5%	$n = 50$	$\sigma = 0.7$	103	359	368	382	656	222	45	186	40
5%	$n = 50$	$\sigma = 0.9$	103	363	367	383	655	222	45	187	40
5%	$n = 100$	$\sigma = 0$	48	139	128	40	273	57	58	63	55
5%	$n = 100$	$\sigma = 0.1$	164	354	305	446	706	194	55	187	69
5%	$n = 100$	$\sigma = 0.3$	174	429	471	498	747	218	45	204	69
5%	$n = 100$	$\sigma = 0.5$	172	436	482	498	748	220	46	206	68
5%	$n = 100$	$\sigma = 0.7$	172	439	483	498	750	219	47	206	69
5%	$n = 100$	$\sigma = 0.9$	172	436	479	499	749	219	47	206	68
5%	$n = 200$	$\sigma = 0$	63	150	123	63	301	59	55	55	68
5%	$n = 200$	$\sigma = 0.1$	266	491	405	580	858	216	45	212	78
5%	$n = 200$	$\sigma = 0.3$	278	543	611	635	878	244	45	236	80
5%	$n = 200$	$\sigma = 0.5$	277	552	622	638	873	248	45	239	80
5%	$n = 200$	$\sigma = 0.7$	277	547	625	639	873	249	45	243	83
5%	$n = 200$	$\sigma = 0.9$	278	543	627	639	874	249	45	243	83



**Continued**

5%	$n = 500$	$\sigma = 0$	55	143	136	55	293	45	50	44	57
5%	$n = 500$	$\sigma = 0.1$	551	719	595	764	969	215	43	215	66
5%	$n = 500$	$\sigma = 0.3$	562	750	809	816	980	232	46	231	67
5%	$n = 500$	$\sigma = 0.5$	564	751	820	825	980	235	47	233	66
5%	$n = 500$	$\sigma = 0.7$	566	748	822	826	980	235	46	233	65
5%	$n = 500$	$\sigma = 0.9$	565	752	822	827	980	235	46	233	66
5%	$n = 1000$	$\sigma = 0$	67	142	116	50	283	49	35	49	47
5%	$n = 1000$	$\sigma = 0.1$	821	916	782	915	997	253	36	251	66
5%	$n = 1000$	$\sigma = 0.3$	831	921	946	939	999	263	43	256	70
5%	$n = 1000$	$\sigma = 0.5$	830	915	951	943	999	261	45	257	70
5%	$n = 1000$	$\sigma = 0.7$	830	915	951	943	999	262	45	258	70
5%	$n = 1000$	$\sigma = 0.9$	830	910	952	944	999	262	45	259	71

\*Frequency of test significance after 1000 replications.