International Journal of Mathematics, Statistics, and Computer Science
Vol. 4
eISSN 2704-1069, DOI: https://doi.org/10.59543/wfvbmv96

On the Power of Heteroskedasticity Tests in Regression: A Monte Carlo Simulation Study

Mohamed R. Abonazel*, Salah M. Mohamed, Hadeer K. Mohamed

Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt

**Corresponding author:* mabonazel@cu.edu.eg

Received 05.01.2026, Revised 21.04.2026, Accepted 17.05.2026

ABSTRACT: An important assumption for the classical linear regression model (CLRM) is homoskedasticity, which assumes that variances of the error vector of the regression model are constant. But if the variances are varying (not constant), this means that the problem of heteroskedasticity is present, and as a result there are some consequences to it, such as the Gauss-Markov theorem not being applied, the ordinary least squares (OLS) estimators no longer being BLUE (Best Linear Unbiased Estimator), the regression predictions being inefficient too, and the standard for the least squares estimator being wrong. Several diagnostic tests have been developed to detect the heteroskedasticity problem. The most common ones include Breusch-Pagan test, White test, and Goldfeld-Quandt test. In this paper we focus on four tests for heteroskedasticity, namely Goldfeld-Quandt (GQ), Glesjer, White, and ARCH tests. We explain each test and indicate the difference between each test. After that we used the Monte Carlo simulation study to apply and compare the performance between the four tests in three types of heteroskedasticity to determine which test has the best performance. The Monte Carlo simulation study has been conducted to evaluate these tests. Several scenarios of simulations were used, such as different sample sizes, different values of the coefficient of variation of the variance of the error, and others. The simulation results were close to each other, but for the first type of heteroskedasticity, the Goldfeld-Quandt test has the best performance and most power. While for the second and third types of heteroskedasticity, the Goldfeld-Quandt and ARCH tests have power close to each other.
Keywords: ARCH; Goldfeld-Quandt test; Glesjer test; Heteroskedasticity; White test

1. INTRODUCTION

Regression is a set of statistical processes for estimating the relationship between two or more variables, Y and X, where Y is the dependent variable and X is the independent variable [1], [2].
The model can be written as follows:

$$Y_i=\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\ldots+\beta_k X_{ki}+u_i;\ i=1, 2, \ldots, n \tag{1}$$

Where Y_i is the ith observation of the dependent variable Y, i=1,2, …, n, X_j the independent variables, j= 1,2, …, k, X_{ji} the ith observation of the jth independent, β_0 is the intercept term, β_j the slope coefficients for each of the independent variables, u_i the error term of the ith observation, n the number of observations; and k the number of independent variables.
The assumptions of model (1), Gujarati [1], presented the assumptions of the model as follows:

1- The relationship between the dependent variable Y_i and independent variable X^s is linear (linearity).

2- The expected value of the residuals is zero, i.e., $E(u_i) = 0$ for $i = 1, 2, \cdots, n$.

3- The variance of the residuals is constant, i.e., $Var(u_i) = \sigma^2$ for $i = 1, 2, \cdots, n$ (homoscedasticity assumption).

4- The residuals are not correlated $Cov(u_i, u_j) = 0$, $i \neq j$ (no autocorrelation).
5- There is no exact linear relationship between the independent variables in the model (no multicollinearity).

6- The independent variables X are non-stochastic matrix.

7- The error term follows the normal distribution with zero mean and (constant) variance σ^2. Symbolically, $u_i \sim N(0, \sigma^2)$.

If assumption three of homoskedasticity is violated, the heteroskedasticity problem occurs and the Gauss-Markov theorem does not apply [2], [3], [4].

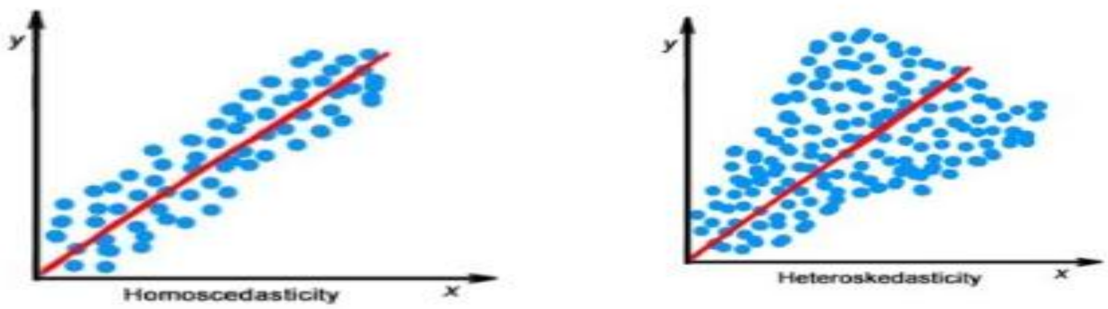
The main contributions of this paper can be summarized as follows:

- Studying the performance of four heteroskedasticity tests (Goldfeld-Quandt (GQ), Glesjer, White, and ARCH tests) in three types of heteroskedasticity.
- Comparing the performance of different tests in different situations.
- Determine the appropriate test for the data.

The following sections of this paper are organized as follows: Section 2 introduces the concepts of heteroscedasticity. Section 3 presents various tests for heteroscedasticity. Section 4 presents a Monte Carlo simulation study. Finally, Section 5 provides the conclusion.

2. Heteroskedasticity Problem

One of the assumptions of the classical linear regression model (CLRM) is that the variance of the residuals is constant; if the variances are not constant, the heteroskedasticity problem occurs, and as a result the ordinary least squares (OLS) estimators are no longer the best linear unbiased estimators (BLUE) [5].



(a) (b)

Fig 1 Homoskedasticity and Heteroskedasticity
Source: [6]

2.1. Heteroskedasticity Types

There are two types of heteroskedasticity as follows [7]:

Conditional heteroskedasticity happens when the error variance is dependent on the values of the independent variables. This creates considerable difficulties for statistical inference. Fortunately, there are many software programs that can identify and fix this issue.

Unconditional heteroskedasticity exists when the heterogeneity variance is not correlated with the values of the independent variables. Even though this goes against the homogeneity assumption, it does not produce major problems concerning statistical inference.

2.2. Reasons for Heteroskedasticity

Gujarati and Porter [8] and Wooldridge [5] introduced heteroskedasticity reasons as follows:

- Outliers.
- Incorrect data transformation.

- Skewness.
- The nature of variables and their relationships can be a reason for appearing heteroskedasticity.
- Incorrect specification.

2.3. Consequences of Heteroskedasticity

- The OLS estimators are no longer BLUE (Best Linear Unbiased Estimators) because they are no longer efficient, so the regression predictions will be inefficient too [5],[9].
- The standard for the least squares estimator is wrong, and this leads to **t** confidence intervals and hypothesis testing that use standard errors, which could lead to misleading conclusions; see[6].

3. Heteroskedasticity Tests

There are many tests to test the heteroskedasticity problem. Many researchers introduced comparisons between the tests of heteroskedasticity, such as Griffiths and Surekha [10] presented a set of heteroskedasticity tests (Szroeter's test, Breusch-Pagan test, Goldfeld-Quandt test, and BAMSET) and compared the power of these tests. Lyon and Tsai [11] used a Monte Carlo simulation to compare between heteroskedasticity tests. These tests are the Ordinary likelihood ratio test, Conditional likelihood ratio test, Modified likelihood ratio test, Residual likelihood ratio test, The Breusch and Pagan's score test, Studentized version of Breusch and Pagan test, ans White s score test, and Uyanto [4] presented a power comparison of seven heteroskedasticity tests by Monte Carlo simulation (Breusch-Pagan test, Glesjer test, Goldfeld-Quandt test, Harvey-Godfrey test, Harrison-McCabe test, Park test, and White test).

We presented four tests for heteroskedasticity as follows: the Goldfeld-Quandt, Glesjer, White, and ARCH tests. In the following lines, we will explain the formula and the details of each test separately.

3.1. Goldfeld-Quandt Test

Goldfeld-Quandt [proposed a test to test if the variances of the residuals are the same across all observations (homoscedasticity). This test compares the variances of two groups, one group of high values and one group of low values. There are steps of GQ test as follows:

- Arrange the observations in ascending order (from small to large observations).
- Divided the observations into three groups (small, middle, large).
- Drop c middle observations where c is specified as a priori and divide (n - c) observations that remain into two groups, (n - c) / 2 observations according to small and large values, respectively.
- Fit a separate regression model for small and large groups, and after that, obtain the respective sum of squares RSS_1, RSS_2.
- Calculate the F test statistic as follows:

$$F= \frac{RSS_2/df_2}{RSS_1/df_1} \quad (2)$$

Where RSS_1 belong to the small values group and RSS_2 belong to the large values group with df_1=n_1- k, $df_2 = n_2 - k$ degree of freedom, and n is the sample size of the two groups and k the number of parameters.
The assumptions of the test,

H_0: The variances are not different (Homoskedasticity).

H_1: The variances are different (Heteroskedasticity).

If F-statistical > F-critical, reject the null hypothesis of Homoskedasticity.

GQ test has some problems; it does not consider cases where heteroskedasticity is caused by more than one variable, and it is not always suitable for time series data. However, it is a very popular model for the simple regression case (with only one explanatory variable) [6].

3.2. Glesjer Test

Glejser [13] introduced a new test for heteroskedasticity that regresses the residuals on the independent variable that is thought to be related to the heteroskedastic variance; see also, [1], [6]. The steps of this test are:

1. Run the OLS estimation on the regression model (1) and calculate the residuals $\hat{u}_i$ of OLS.
2. Run the following auxiliary regression:

$$|\hat{u}_i|=\gamma_0+\gamma_1 lnZ_{1i}+\gamma_2 lnZ_{2i}+\ldots+\gamma_k ln\, Z_{ki}+v_i \quad (3)$$

Where Z_{ki} the independent variables of the original regression equation in the model (1) are denoted by X^s.

3. The test hypotheses are as follows:

The null hypothesis is $H_0: \gamma_0 = \gamma_1 = \gamma_2 = \ldots, \gamma_k = 0$. The alternative hypothesis, H_1 is that at least one of the γ^s different from 0.

4. Computing the $LM = nR^2$ statistics, where n is the number of observations and R^2 is the coefficient of determination for this regression in (3).

3.3. White Test

This test is an extremely general test for heteroskedasticity [14]; see also [5], White test is also an LM test [3]. The steps of the White test assumed a model with two explanatory variables as follows:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i \quad (4)$$

- Run the regression model in equation (4) and obtain the residuals $\hat{u}_i$.
- Regress the squared residuals on a constant, the independent variables, the squared terms of these independent variables, and the cross-product term of each independent variable and run the auxiliary regression:

$$\hat{u}_i^2 = \gamma_0 + \gamma_1 X_{1i} + \gamma_2 X_{2i} + \gamma_4 X_{1i}^2 + \gamma_5 X_{2i}^2 + \gamma_6 X_{1i} X_{2i} + v_i \quad (5)$$

  - Formulate the null hypothesis is, $H_0: \gamma_0 = \gamma_1 = \gamma_2 = \ldots = \gamma_6 = 0$, and H_1 is that at least one of the γ^s different from 0.
  - Computing the test statistic $LM = nR^2$, where the R^2 value is from this auxiliary regression in (5) and n is the number of observations.

3.4. ARCH Test

ARCH test is proposed by [15] for testing for the ARCH,], the classical methods for detecting the ARCH effect are constructed based on two tests: the Lagrange multiplier test of [15] which tests for autocorrelation and the portmanteau test which similar to the Ljung-Box test [16]. The ARCH Engle's test supposes that the squared residuals (u_t^2) are autocorrelated, and the portmanteau test supposes that if the residuals (u_t) are heteroscedastic.

We used the "aTSA" package [17] to conduct the ARCH test. The steps for Engle's ARCH test are as follows:

- Estimate the regression equation:

$$y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + u_t,\ t=1,\ldots,n \quad (6)$$

By OLS and obtain the residuals $\hat{u}_t$.

- Regress the squared residual u_t^2 against a contrast $u_{t-1}^2, u_{t-2}^2, \ldots, u_{t-p}^2$, (the value of p will be determined by the order of ARCH (p) being tested for)

$$u_t^2 = \lambda_0 + \lambda_1 u_{t-1}^2 + \cdots + \lambda_p u_{t-p}^2 + \epsilon_t \quad (7)$$

- The null hypothesis is $H_0{:} = \lambda_0 = \lambda_1 = \cdots = \lambda_p = 0$, that there are no ARCH effects, versus the alternative hypothesis: $H_1: \lambda_i \neq 0$ (there are ARCH effects).

The test statistic for Engle's ARCH test follows χ^2distribution with a degree of freedom p.

4. Simulation Study

Abonazel [18] proposed a complete algorithm to make professional Monte Carlo simulation studies using R software. This algorithm is suitable for creating any simulation study in statistical and econometric models and is considered as a practical guide for researchers to conduct simulation studies in statistics and econometrics. In this simulation, we compared the performance of four heteroskedasticity tests (Goldfield-Quandt, Glesjer, White, and ARCH tests) based on the power of each test for the linear regression model as follows [19]:

$$y_i = \beta_0 + \beta_1 x_i + u_i,\ \text{i}=1, 2, \ldots, n \quad (8)$$

Where u_i are unobservable random errors with mean and variances: $E(u_i)=0$, var $(u_i) = \sigma_i^2$, respectively. Heteroskedasticity problem is present, when σ_i^2 are different (i.e., not all equal).

The comparison of the four tests for heteroskedasticity will be in three types of heteroskedasticity as follows: $\sigma_i^2 = \sigma^2 x_i^2$, $\sigma_i^2 = \sigma^2 i$, $\sigma_i^2 = \sigma^2 i^2$, see [20] and [21].

Surekha [22] and Kennedy [23], suggested a suitable measure of the degree of heteroskedasticity, which is the coefficient of variation (CV) of the variances [24] later used this measure. Uyanto [4] displayed and applied this measure. The formula of the CV is

$$\text{CV}=\sqrt{\frac{1}{n-1}\sum\left(\sigma_i^2 - \frac{\sum \sigma_i^2}{n}\right)^2} \Big/ \left(\frac{1}{n}\sum \sigma_i^2\right) \quad (9)$$

The coefficient of variation is independent of the unit in which the measurement has been taken, so it is a dimensionless number.

We compared the power of heteroskedasticity tests by using Monte Carlo simulation for three types of heteroskedasticity and different sample sizes n = 15, 30, 60, 90, 120, 150, for replication L = 1000, level of significance $\alpha = 0.05$, $\beta_0 = 2$ and $\beta_1 = 3$, and calculated σ_i^2 for the values of CV. The powers of the tests are calculated by the proportion of rejection for the null hypothesis in the total number of replications (L).

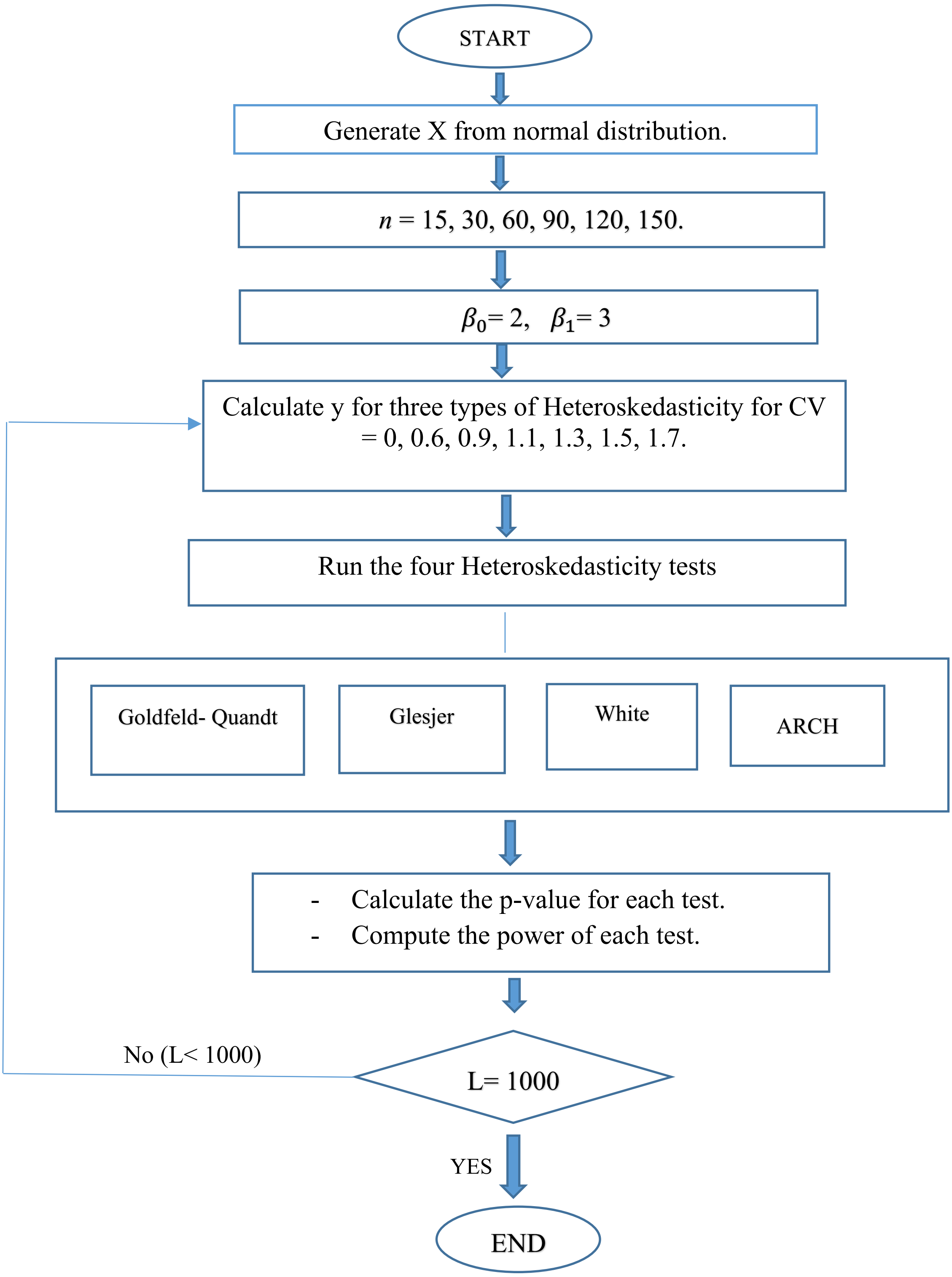


Fig 2. Flow Chart for the Simulation Study

4.1 Simulation Results

For more accessible interpretations of the tables results, the sample sizes were categorized as small, moderate, and large. Also categorized the coefficient of variation to low, moderate, and high levels as follows:

- Small sample sizes (*n*) = 15, 30 Low level (CV) = 0, 0.6
- Moderate sample sizes (*n*) = 60, 90 Moderate level (CV) = 0.9, 1.1
- Large sample sizes (*n*) = 120, 150 High level (CV) = 1.3, 1.5, 1.7

4.1.1 First Type of Heteroskedasticity $\sigma_i^2 = \sigma^2 x_i^2$

The results of the first type are shown in Table 1.

1) Goldfeld-Quandt (GQ) Test

Goldfeld-Quandt test power was good in all cases of different *n*, CV.

(*i*) Small sample sizes (*n*) = 15, 30 low level (CV) = 0, 0.6

The power was 0.053 when *n* = 15 and CV = 0, and by increasing *n* to 30 and CV to 0.6, the power also increased and became 0.83.

(*ii*) Moderate sample sizes (*n*) = 60, 90 moderate level (CV) = 0.9, 1.1

Goldfeld-Quandt test power clearly increased with moderate *n* and CV. For example, when *n* = 60 and CV = 0.9, the power was 0.996, and by increasing *n* to 90 and CV to 1.1, the power clearly changed and was 1, and this was the most powerful.

(*iii*) Large sample sizes (*n*) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

The power still increased with large *n* and high CV. For example, when *n* = 120 and CV = 1.3, the power was 1, and when *n* = 150, CV = 1.7, the power was 1, and this value of 1 was the most power.

The power of Goldfeld-Quandt test was the best for different *n* and CV, especially in the case of small *n* and low CV and as a result.

2) Glesjer Test

The power results of Glesjer test increase by increasing *n* and CV.

(*i*) Small sample sizes (*n*) = 15, 30 low level (CV) = 0, 0.6

The power was 0.049 when *n* = 15 and CV = 0, and increased to 0.682 when *n* was increased to 30 and CV to 0.6.

(*ii*) Moderate sample sizes (*n*) = 60, 90 moderate level (CV) = 0.9, 1.1

With moderate *n* and CV, the power increased; when *n* = 60 and CV = 0.9, the power value was 0.864 and increased to 0.906 when *n* was raised to 90 and CV to 1.1.

(*iii*) Large sample sizes (*n*) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

Still, the Glesjer test power was raised with large *n* and high CV. When *n* = 120 and CV = 1.3, the power was 0.934, and by increasing *n* to 150 and CV to 1.7, the power was 0.956.

The power of Glesjer test was good for different *n* and CV, but not the best compared to Goldfeld-Quandt test power results.

3) White Test

White test power was close to the power of the Glesjer test in the case of small sample size and low-level CV.

(*i*) Small sample sizes (*n*) = 15, 30 low level (CV) = 0, 0.6

When n = 15 and CV = 0, the power was 0.39 and became 0.426, when n = 30 and CV = 0.6.

(*ii*) Moderate sample sizes (*n*) = 60, 90 moderate level (CV) = 0.9, 1.1

With moderate n and CV, the power increased not much; for example, when n = 60 and CV = 0.9, the power value was 0.625, and by increasing n to 90 and CV to 1.1, the power became 0.698**.**

(*iii*) Large sample sizes (*n*) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

Still the power increased. For example, when n = 120 and CV = 1.3, the power was 0.736 and became 0.796 when n = 150 and CV = 1.7.

When we compared White test power with Goldfeld-Quandt, Glesjer power tests, White test has less power than the power of the Goldfeld-Quandt and Glesjer tests, so Goldfeld-Quandt still the best.

4) ARCH Test

***i*) Small sample sizes (*n*) = 15, 30 low level (CV) = 0, 0.6**

The power was 0.007 when n = 15 and CV = 0 and when it increased n to 30 and CV to 0.6, the power value became 0.035.

(*ii*) Moderate sample sizes (*n*) = 60, 90 moderate level (CV) = 0.9, 1.1

The ARCH test power slightly increased by increasing n and CV. For example, when n = 60 and CV = 0.9, the power was 0.216, and became 0.338, when we increased n to 90 and CV to 1.1.

(*iii*) Large sample sizes (*n*) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

With large n and high CV, the power slightly differed; for example, when n = 120 and CV = 1.3, the power value was 0.497, and by increasing n to 150 and CV to 1.7, the power value became 0.588.

When comparing the power of ARCH test with Goldfeld-Quandt test, Glesjer test and the White test, ARCH test has the least power, and Goldfeld-Quandt test still the best.

4.1.2 Second Type of Heteroscedasticity: $\sigma_i^2 = \sigma^2 i$

The results of this type are shown in Table 2.

1) Goldfeld-Quandt Test

Goldfeld-Quandt test power increased by increasing n and CV. We show the examples of the results as follows:

(*i*) Small sample sizes (*n*) = 15, 30 low level (CV) = 0, 0.6

The power was 0.054 when n = 15 and CV = 0, and when n was increased to 30 and CV to 0.6, the power clearly increased and became 0.996.

(*ii*) Moderate sample sizes (*n*) = 60, 90 moderate level (CV) = 0.9, 1.1

The power still clearly increased with moderate n and CV; when n = 60 and CV = 0.9, the power was 1, and by increasing n to 90 and CV to 1.1, the power was 1, and these values were the most power.

(*iii*) Large sample sizes (*n*) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

Goldfeld-Quandt test power also clearly increased with large n and high CV. For example, when n = 120 and CV = 1.3, the power was 1 and when we increased n to 150 and CV to 1.7, the power also was 1, and this was the most power.

We found that Goldfeld-Quandt test has the most power for different n and CV, especially in the case of small n and low CV; therefore Goldfeld-Quandt test was the best and has the best performance.

2) Glesjer Test

(i) Small sample sizes (n) = 15, 30 low level (CV) = 0, 0.6

The power was 0.055 when n = 15 and CV = 0 and the power increased but not much when we increased n to 30 and CV to 0.6 and became 0.056.

(ii) Moderate sample sizes (n) = 60, 90 moderate level (CV) = 0.9, 1.1

In the case of moderate n and CV, the power of the Glesjer test was not differ much. When n = 60 and CV = 0.9, the power was 0.066 and became 0.08 when we increased n to 90 and CV to 1.1.

(iii) Large sample sizes (n) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

Still, the power of Glesjer test slightly differed with large n and high CV; when n = 120 and CV = 1.3, the power was 0.075, and by increasing n to 150 and CV to 1.7, the power did not increase but was least and 0.069.

The results of Glesjer test showed that power was low for different n and CV and when we compared Glesjer test with Goldfeld-Quandt test, Glesjer had the least power, and Goldfeld-Quandt was still the best.

3) White Test

The power of White test was low and decreased although we increased n, CV.

(i) Small sample sizes (n) = 15, 30 low level (CV) = 0, 0.6

The power was 0.047 when n = 15 and CV = 0 and the power became 0.039 when we increased n to 30 and CV to 0.6.

(ii) Moderate sample sizes (n) = 60, 90 moderate level (CV) = 0.9, 1.1

The power did not change much with moderate n and CV; for example, n = 60 and CV = 0.9, the power value was 0.049, and by increasing n to 90 and CV to 1.1, the power value became 0.056.

(iii) Large sample sizes (n) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

When n = 120 and CV = 1.3, the power was 0.056, and when we increased n to 150 and CV to 1.7, the power was 0.056. Noticed that the power is still constant although increasing n and CV.

When compared the power of White test with the power of Goldfeld-Quandt and Glesjer, tests, White test has low power such as Glesjer, so Goldfeld-Quandt has the most power.

4) ARCH Test

ARCH test power increased but not much in the case of small n, low CV, but clearly increased in the two cases of moderate, large n and moderate, high CV. We present examples of the ARCH power as follows:

(i) Small sample sizes (n) = 15, 30 low level (CV) = 0, 0.6

The power was 0.006 when n = 120 and CV = 0 and became 0.206 when n was increased to 30 and CV to 0.6.

(ii) Moderate sample sizes (n) = 60, 90 moderate level (CV) = 0.9, 1.1

With moderate n and CV, the power increased much; when n = 60 and CV = 0.9, the power was 0.66, and when we increased n to 90 and CV to 1.1, the power also increased and became 0.861.

(iii) Large sample sizes (n) = 120, 150 high level (CV) = 1.3, 1.5, 1.7

Still, the ARCH test power increased by increasing n and CV; when n = 120 and CV = 1.3, the power was 0.924 and became 0.955 when n =150 and CV = 1.7.

ARCH test power was better than the power of Glesjer and White tests and close to the power of Goldfeld-Quandt test but still Goldfeld-Quandt the best.

4.1.3. Third Type of Heteroscedasticity: $\sigma_i^2 = \sigma^2 i^2$

The results of the third type are shown in Table 3.

The results of the power tests in this type were close to the results of the power tests in the second type of heteroskedasticity. Glesjer, White tests have the least power, while Goldfeld-Quandt test and ARCH test power results were close to each other, but Goldfeld- quandt test was the best and has the most power.

Table 1 Power of Heteroskedasticity Tests for First Type $\sigma_i^2 = \sigma^2 x_i^2$

Test	CV							
	n	**0**	**0.6**	**0.9**	**1.1**	**1.3**	**1.5**	**1.7**
Goldfeld-Quandt	**15**	0.053	0.385	0.368	0.369	0.38	0.364	0.391
	30	0.046	0.83	0.822	0.838	0.833	0.855	0.852
	60	0.049	0.993	0.996	0.997	0.995	0.994	0.999
	90	0.044	1	1	1	1	1	1
	120	0.049	1	1	1	1	1	1
	150	0.045	1	1	1	1	1	1
Glesjer	15	0.049	0.451	0.42	0.442	0.414	0.422	0.449
	30	0.056	0.682	0.662	0.698	0.689	0.671	0.699
	60	0.046	0.842	0.864	0.834	0.843	0.845	0.854
	90	0.061	0.899	0.913	0.906	0.922	0.915	0.912
	120	0.049	0.945	0.944	0.924	0.934	0.936	0.928
	150	0.046	0.948	0.944	0.949	0.939	0.953	0.956
White	**15**	0.39	0.306	0.281	0.271	0.271	0.288	0.0289
	30	0.034	0.426	0.412	0.437	0.431	0.399	0.43
	60	0.041	0.609	0.625	0.612	0.61	0.613	0.595
	90	0.041	0.68	0.696	0.698	0.706	0.701	0.675
	120	0.051	0.776	0.748	0.728	0.736	0.74	0.752
	150	0.047	0.795	0.796	0.783	0.787	0.769	0.796
ARCH	**15**	0.007	0.001	0.003	0.002	0.003	0.005	0.005
	30	0.015	0.035	0.037	0.055	0.038	0.033	0.04
	60	0.033	0.23	0.216	0.221	0.25	0.216	0.23
	90	0.034	0.37	0.396	0.388	0.383	0.375	0.371
	120	0.029	0.518	0.53	0.498	0.497	0.504	0.546
	150	0.036	0.6	0.609	0.591	0.592	0.614	0.588

Table 2 Power of Heteroskedasticity Tests for Second Type $\sigma_i^2 = \sigma^2 i$

Test		CV						
	n	**0**	**0.6**	**0.9**	**1.1**	**1.3**	**1.5**	**1.7**
Goldfeld-Quandt	15	0.054	0.673	0.657	0.654	0.676	0.672	0.651
	30	0.048	0.996	0.997	0.994	0.998	0.993	0.996
	60	0.042	1	1	1	1	1	1
	90	0.058	1	1	1	1	1	1
	120	0.043	1	1	1	1	1	1
	150	0.041	1	1	1	1	1	1
Glesjer	15	0.055	0.056	0.065	0.072	0.064	0.064	0.054
	30	0.054	0.056	0.048	0.064	0.062	0.061	0.064
	60	0.062	0.065	0.066	0.053	0.069	0.059	0.064
	90	0.045	0.064	0.069	0.08	0.073	0.056	0.064
	120	0.051	0.083	0.06	0.056	0.075	0.077	0.07
	150	0.051	0.08	0.08	0.063	0.065	0.067	0.069
White	15	0.047	0.061	0.051	0.055	0.041	0.055	0.045
	30	0.047	0.039	0.045	0.063	0.049	0.054	0.066
	60	0.036	0.053	0.049	0.048	0.055	0.052	0.062
	90	0.06	0.053	0.064	0.056	0.057	0.062	0.056
	120	0.037	0.057	0.054	0.038	0.056	0.062	0.042
	150	0.046	0.082	0.052	0.052	0.052	0.061	0.056
ARCH	15	0.006	0.005	0.007	0.005	0.005	0.002	0.002
	30	0.018	0.206	0.228	0.176	0.228	0.191	0.219
	60	0.023	0.669	0.66	0.667	0.67	0.662	0.675
	90	0.041	0.845	0.832	0.861	0.862	0.839	0.86
	120	0.046	0.917	0.921	0.909	0.924	0.915	0.936
	150	0.039	0.968	0.959	0.954	0.967	0.961	0.955

Table 3 Power of Heteroskedasticity Tests for Third Type $\sigma_i^2 = \sigma^2 i^2$

Test		CV						
	n	0	0.6	0.9	1.1	1.3	1.5	1.7
Goldfeld-Quandt	15	0.05	0.879	0.879	0.866	0.864	0.877	0.893
	30	0.052	0.999	1	0.999	1	1	0.998
	60	0.045	1	1	1	1	1	1
	90	0.038	1	1	1	1	1	1
	120	0.047	1	1	1	1	1	1
	150	0.046	1	1	1	1	1	1
Glesjer	15	0.055	0.068	0.062	0.066	0.065	0.072	0.067
	30	0.046	0.069	0.069	0.067	0.07	0.064	0.069
	60	0.061	0.078	0.076	0.076	0.079	0.064	0.072
	90	0.058	0.067	0.063	0.071	0.06	0.087	0.062
	120	0.041	0.067	0.065	0.083	0.065	0.065	0.063
	150	0.049	0.07	0.062	0.061	0.075	0.065	0.075
White	15	0.041	0.044	0.061	0.047	0.053	0.067	0.071
	30	0.034	0.062	0.063	0.054	0.063	0.051	0.05
	60	0.049	0.063	0.06	0.063	0.069	0.045	0.056
	90	0.051	0.058	0.064	0.053	0.053	0.046	0.062
	120	0.042	0.056	0.052	0.058	0.047	0.049	0.058
	150	0.04	0.06	0.047	0.053	0.048	0.059	0.062
ARCH	15	0.008	0.01	0.006	0.012	0.007	0.014	0.006
	30	0.024	0.254	0.27	0.285	0.259	0.236	0.277
	60	0.029	0.703	0.701	0.685	0.711	0.705	0.711
	90	0.033	0.888	0.859	0.843	0.87	0.863	0.855
	120	0.034	0.927	0.912	0.935	0.926	0.915	0.926
	150	0.043	0.965	0.968	0.959	0.953	0.973	0.958

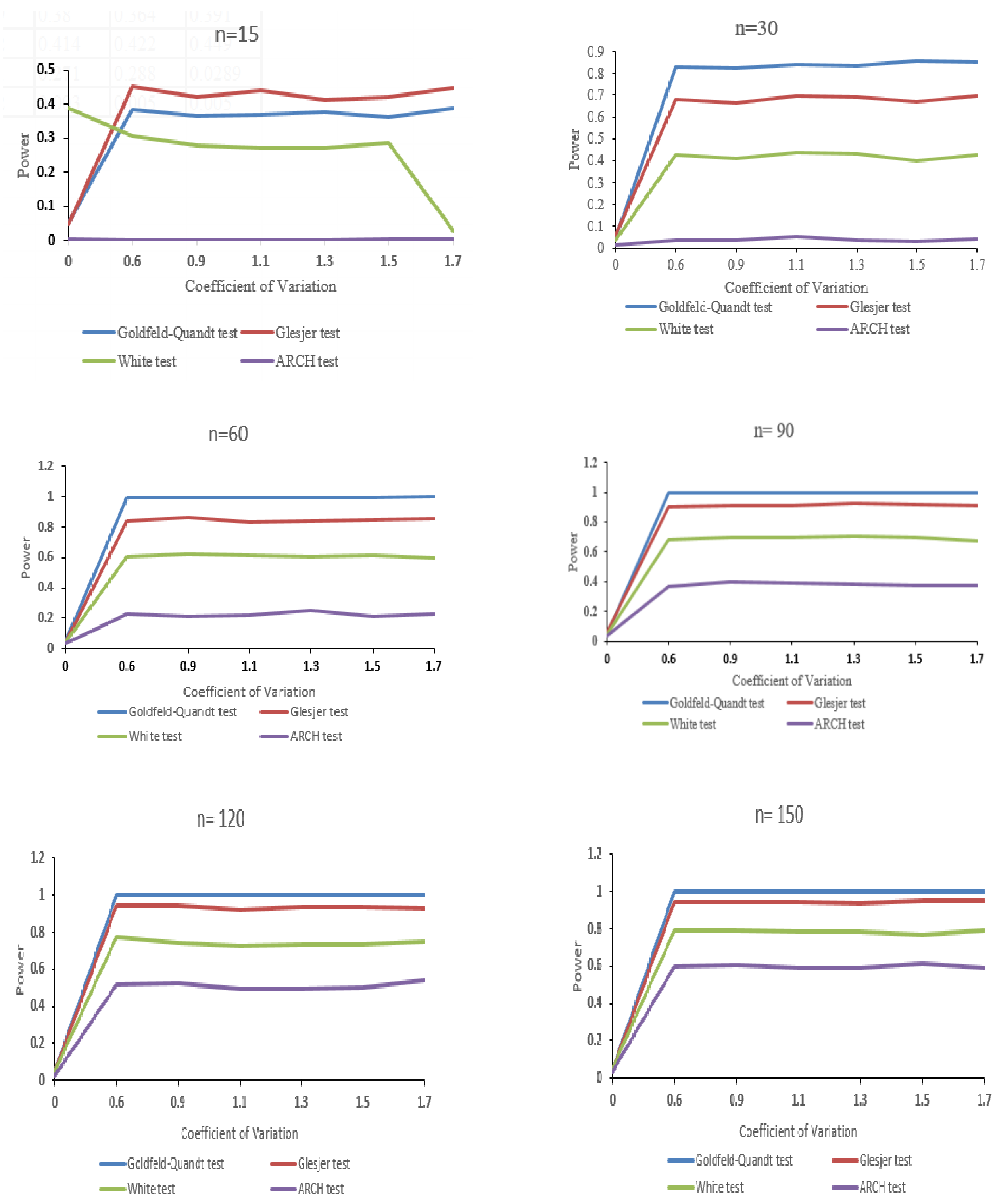


Fig 3: Power curves for Heteroskedasticity Tests for First Heteroskedasticity Type $\sigma_i^2 = \sigma^2 x_i^2$

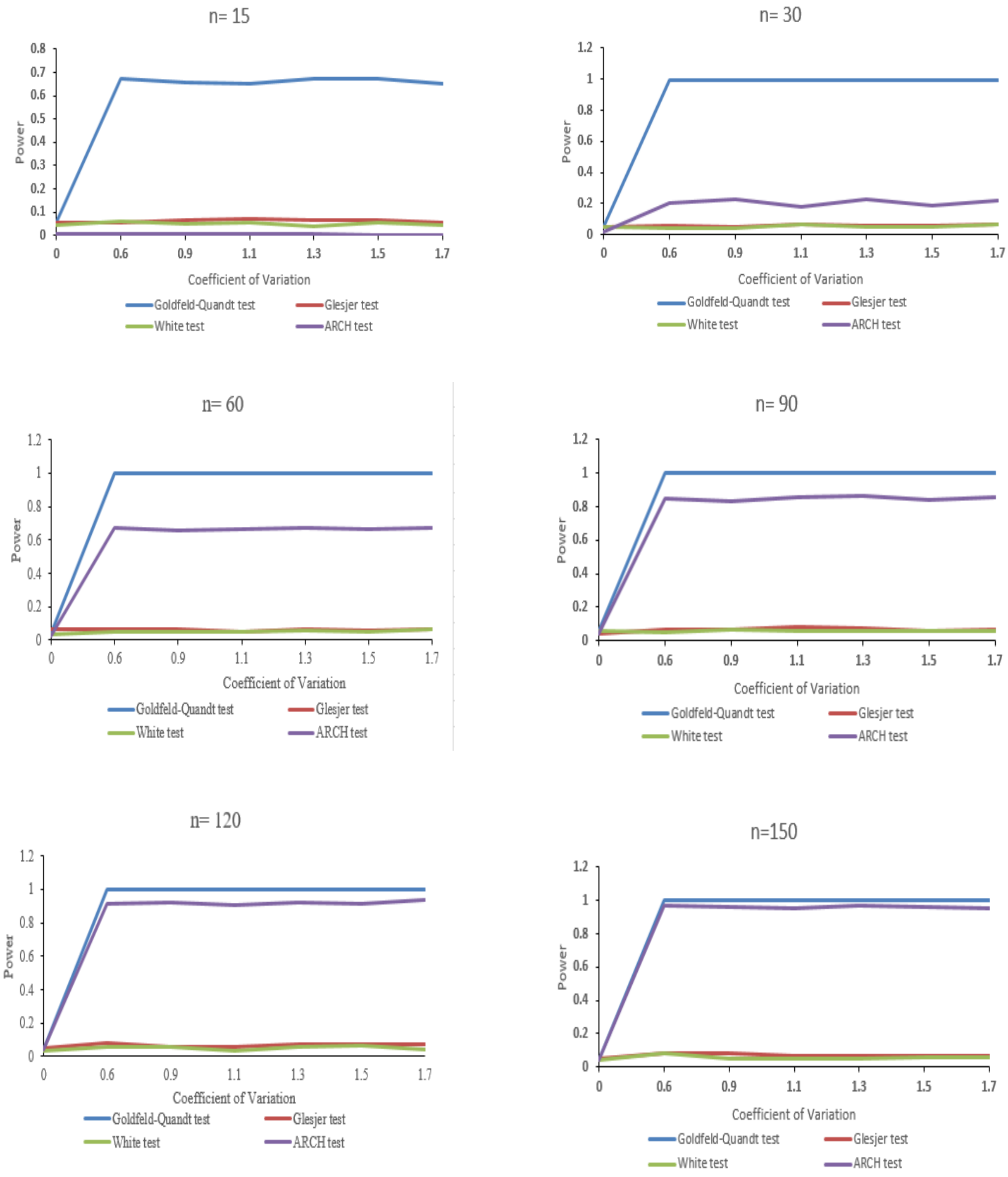


Fig 4: Power curves for Heteroskedasticity Tests for Second Heteroskedasticity Type $\sigma_i^2 = \sigma^2 i$

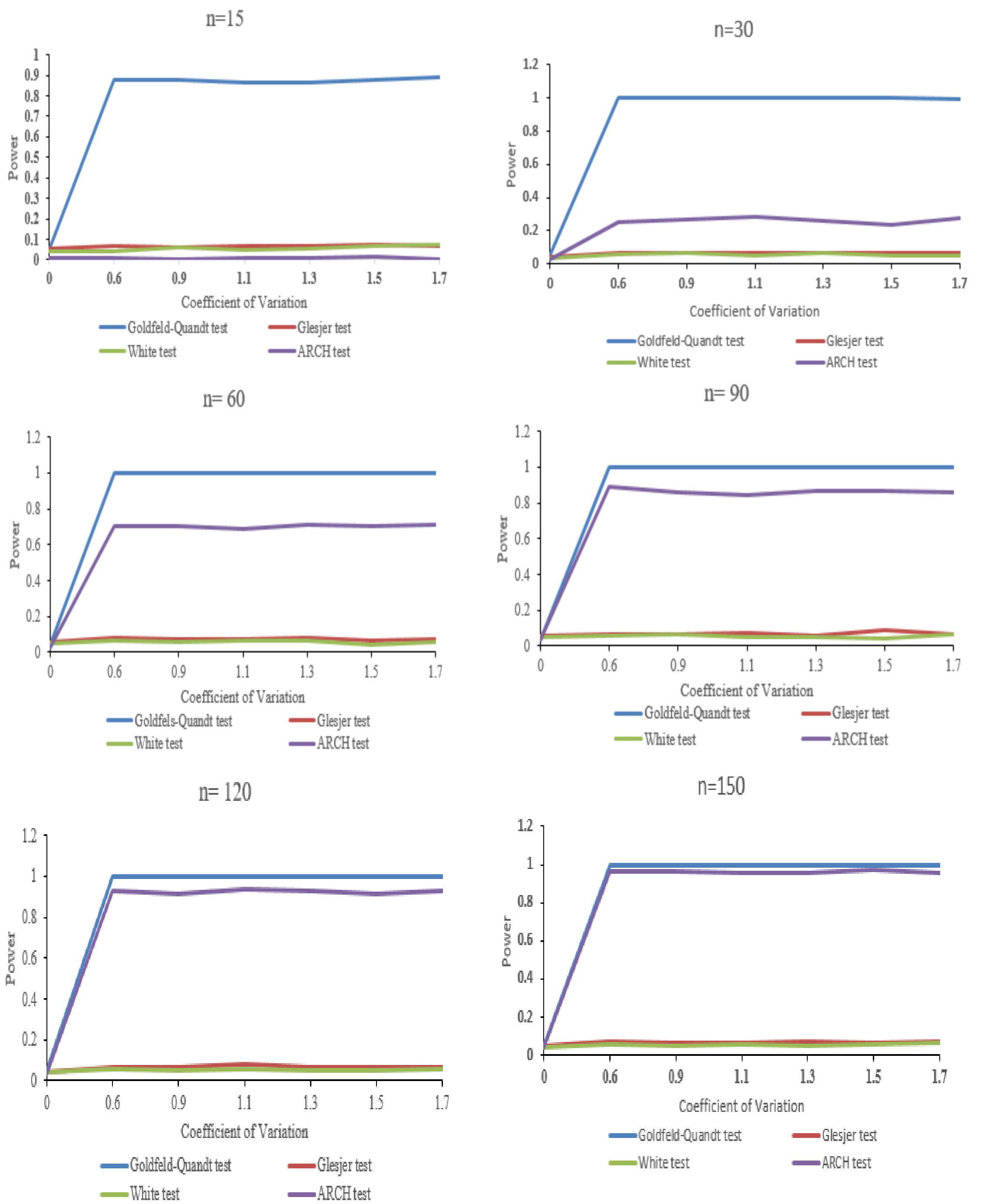


Fig 5: Power curves for Heteroskedasticity Tests for Third Heteroskedasticity Type $\sigma_i^2 = \sigma^2 i^2$.

- For the first type of heteroskedasticity $\sigma_i^2 = \sigma^2 x_i^2$, Goldfeld-Quandt has the most power and the best performance.
- For the second type of heteroskedasticity $\sigma_i^2 = \sigma^2 i$ and the third type of heteroskedasticity $\sigma_i^2 = \sigma^2 i^2$, Goldfeld-Quandt test and ARCH test have better power than Glesjer and White tests, but the best test was Goldfeld-Quandt test.
- Fig 3 shows the power curves for Goldfeld-Quandt, Glesjer, White and ARCH tests for the first type of heteroskedasticity with different *n* = (15, 30, 60, 90, 120, 150) and CV= (0, 0.6, 0.9, 1.1, 1.3, 1.5, 1.7). The vertical axis points to the power of the four mentioned tests, while the horizontal axis points to the value of CV. This Fig contains six plots for six different *n,* and each plot shows the

power curves of the mentioned tests. Notice that the power curve of Goldfeld-Quandt was at the top for different value of *n*, so Goldfeld-Quandt has the best performance.

- Fig 4 and 5 show the power curves for Goldfeld-Quandt, Glesjer, White and ARCH tests for the first type of heteroskedasticity. The curves of Goldfeld-Quandt and ARCH tests were close to each other and were at the top compared to other tests, and this is indicated through their curve for different n, so they have the best performance. Glesjer and White curves were at the bottom and as follows do not have the best performance.

5. CONCLUSION

In this paper, we examined the performance of four heteroskedasticity tests (Goldfeld-Quandt (GQ), Glesjer, White, and ARCH tests) in three types of heteroskedasticity by a Monte Carlo simulation study. In the first type of heteroskedasticity, Goldfeld-Quandt test has the best performance. While in the second and third types, Goldfeld-Quandt and ARCH tests have the best performance. Therefore, we recommend using Goldfeld-Quandt and ARCH tests to check the heteroskedasticity problem in linear regression. For future work, we can propose to test the heteroskedasticity problem in the presence of outliers and missing data, as an extension of [25], and also test the heteroskedasticity problem in the presence of multicollinearity problem as an extension of [26]. Furthermore, we propose to develop a new estimator for a count (Poisson) regression model in the presence of homoscedasticity with multicollinearity and/or outliers, based on the work of [27, 28]. Test for hetreoskedasticity problem in nonparametric regression, as an extension of Kotekal et al. [29]. Study the performance of heteroskedasticity tests in high dimension linear regression as an extension of [30, 31]. Test for hetreoskedasticity in panel data model, as an extension of [32].

REFERENCES

[1] Gujarati, D. N. (2022). Basic econometrics: Prentice Hall.
[2] Kacapyr, E. (2022). Essential econometric techniques: a guide to concepts and applications. Routledge.
[3] Greene, W. H. (2018). Econometric Analysis (8th ed.). New York: Pearson Education.
[4] Uyanto, S. S. (2019). Monte Carlo power comparison of seven most commonly used heteroscedasticity tests. Communications in Statistics-Simulation and Computation, 1-18, DOI: 10.1080/03610918.2019.1692031.
[5] Wooldridge, J. M. (2025). Introductory Econometrics: A Modern Approach (8th ed.)
[6] Asteriou, D., and Hall, S. G. (2021). Applied econometrics. Bloomsbury Publishing.
[7] Nicolas Joyce (2022). Explain the Types of Heteroskedasticity and How It Affects Statistical Inference.
[8] Gujarati, D. N., and Porter, D. C. (2009). Basic econometrics. McGraw-Hill/Irwin. New York.
[9] Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. Cengage learning.
[10] Griffiths, W. E., and Surekha, K. (1986). A Monte Carlo evaluation of the power of some tests for heteroscedasticity. Journal of Econometrics, 31(2), 219-231, https://doi.org/10.1016/0304-4076(86)90049-7.
[11] Lyon, J. D., and Tsai, C. L. (1996). A comparison of tests for heteroscedasticity. Journal of the Royal Statistical Society: Series D (The Statistician), 45(3), 337-349, https://doi.org/10.2307/2988471.
[12] Goldfeld, S. M., and Quandt, R. E. (1965). Some tests for homoscedasticity. Journal of the American statistical Association, 60(310), 539-547, https://doi.org/10.1080/01621459.1965.10480811.
[13] Glejser, H. (1969). A new test for heteroskedasticity. Journal of the American Statistical Association, 64(325), 316-323.
[14] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: journal of the Econometric Society, 817-838, https://doi.org/10.2307/1912934.
[15] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the econometric society, 987-1007, https://doi.org/10.2307/1912773.
[16] McLeod, A. I., and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual auocorrelations. Journal of time series analysis, 4(4), 269-273, https://doi.org/10.1111/j.1467-9892.1983.tb00373.x.
[17] Qiu, D., & Qiu, M. D. (2015). Package 'aTSA'. Erişim adresi https://cran. r-project. org/web/packages/aTSA/aTSA. pdf.
[18] Abonazel, M. R. (2018). A practical guide for creating Monte Carlo simulation studies using R. International Journal of Mathematics and Computational Science, 4(1), 18-33.
[19] Abonazel, M. R. (2025). A New Biased Estimation Class to Combat the Multicollinearity in Regression Models: Modified Two-Parameter Liu Estimator. Computational Journal of Mathematical and Statistical Sciences, 4(1), 316-347, DOI: 10.21608/cjmss.2025.347818.1096.
[20] Ali, M. M., and Giaccotto, C. (1984). A study of several new and existing tests for heteroscedasticity in the general linear model. Journal of Econometrics, 26(3), 355-373.
[21] Harrison, M. J., and McCabe, B. P. (1979). A test for heteroskedasticity based on ordinary least squares residuals. Journal of the American Statistical Association, 74(366a), 494-499, https://doi.org/10.1080/01621459.1979.10482544
[22] Surekha, K. (1980). Contributions to Bayesian analysis in heteroscedastic models. Unpublished Ph. D. thesis, University of New England, Armidale.
[23] Kennedy, P. E. (1985). A suggestion for measuring heteroskedasticity. Communications in Statistics-Simulation and Computation, 14(4), 845-851.
[24] Evans, M. A., and King, M. L. (1985). A point optimal test for heteroscedastic disturbances. Journal of Econometrics, 27(2), 163-178, https://doi.org/10.1016/0304-4076(85)90085-5.
[25] Abonazel, M. R. (2020). Handling outliers and missing data in regression models using R: Simulation examples. Academic Journal of Applied Mathematical Sciences, 6(8), 187-203, DOI: 10.32861/ajams.68.187.203.
[26]Abonazel, M. R., Dawoud, I., Awwad, F. A., & Tag-Eldin, E. (2023). New estimators for the probit regression model with

multicollinearity. Scientific African, 19, e01565, https://doi.org/10.1016/j.sciaf.2023.e01565.
[27] Abonazel, M. R., & Saber, O. M. (2020). A comparative study of robust estimators for Poisson regression model with outliers. J. Stat. Appl. Probab, 9, 279-286, DOI: 10.18576/jsap/090208.
[28] Algamal, Z. Y., Abonazel, M. R., Awwad, F. A., & Eldin, E. T. (2023). Modified jackknife ridge estimator for the Conway-Maxwell-Poisson model. Scientific African, 19, e01543, https://doi.org/10.1016/j.sciaf.2023.e01543.
[29] Kotekal, S., & Kundu, S. (2025). Optimal heteroskedasticity testing in nonparametric regression. The Annals of Statistics, 53(1), 295-321.
[30] Shinkyu, A. (2025). Testing Heteroskedasticity in High-Dimensional Linear Regression. Econometrics and Statistics, 35, 120-134.
[31] El-Sheikh, A. A., Ali, M. C., & Abonazel, M. R. (2024). Development of two methods for estimating high-dimensional data in the case of multicollinearity and outliers. *International Journal of Analysis and Applications*, *22*, 187-187.
[32] Feng, S., Li, G., Tong, T., & Luo, S. (2020). Testing for heteroskedasticity in two-way fixed effects panel data models. Journal of Applied Statistics, 47(1), 91-116.