

## A Robust Modification of the Goldfeld-Quandt Test for the Detection of Heteroscedasticity in the Presence of Outliers

<sup>1</sup>Md. Sohel Rana, <sup>1</sup>Habshah Midi and <sup>2</sup>A.H.M. Rahmatullah Imon

<sup>1</sup>Laboratory of Applied and Computational Statistics, Institute for Mathematical Research, University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

<sup>2</sup>Department of Mathematical Sciences, Ball State University, Muncie, IN 47306, USA

---

**Abstract: Problem statement:** The problem of heteroscedasticity occurs in regression analysis for many practical reasons. It is now evident that the heteroscedastic problem affects both the estimation and test procedure of regression analysis, so it is really important to be able to detect this problem for possible remedy. The existence of a few extreme or unusual observations that we often call outliers is a very common feature in data analysis. In this study we have shown how the existence of outliers makes the detection of heteroscedasticity cumbersome. Often outliers occurring in a homoscedastic model make the model heteroscedastic, on the other hand, outliers may distort the diagnostic tools in such a way that we cannot correctly diagnose the heteroscedastic problem in the presence of outliers. Neither of these situations is desirable. **Approach:** This article introduced a robust test procedure to detect the problem of heteroscedasticity which will be unaffected in the presence of outliers. We have modified one of the most popular and commonly used tests, the Goldfeld-Quandt, by replacing its nonrobust components by robust alternatives. **Results:** The performance of the newly proposed test is investigated extensively by real data sets and Monte Carlo simulations. The results suggest that the robust version of this test offers substantial improvements over the existing tests. **Conclusion/Recommendations:** The proposed robust Goldfeld-Quandt test should be employed instead of the existing tests in order to avoid misleading conclusion.

**Key words:** Heteroscedasticity, outliers, robust test, modified goldfeld-quandt test, Monte Carlo simulation

---

### INTRODUCTION

It is a common practice over the years to use the Ordinary Least Squares (OLS) as the inferential technique in regression. Under the usual assumptions, the OLS possesses some nice and attractive properties. Among them homogeneity of error variances (homoscedasticity) is an important assumption for which the OLS estimators enjoy the minimum variance property. But there are many occasions<sup>[2]</sup> when the assumption of homoscedastic error variance is unreasonable. For example, if one is examining a cross section of firms in one industry, error terms associated with very large firms might have larger variance than those of error terms associated with smaller firms. If the error variance changes we call the error heteroscedastic. Heteroscedasticity often occurs when there is a large difference among the sizes of the observations. It is really important to detect this problem because if this problem is not eliminated the least squares estimators

will still be unbiased, but they will no longer have the minimum variance property. This means that the regression coefficients will have larger standard errors than necessary.

A large number of diagnostic plots are now available in the literature<sup>[3,4,9,11,14]</sup> for detecting heteroscedasticity. But graphical methods are very subjective so we really need analytical methods to detect the problem of heteroscedasticity. Rigorous procedures for testing the homoscedasticity of data are available in the literature<sup>[1,5,15]</sup>. Most of these techniques are based on the least squares residuals but there is evidence that these residuals may not exhibit heteroscedastic pattern if outliers are present in the data. According to Hampel *et al.*<sup>[7]</sup> the existence of 1-10% outliers in a routine data is rather rule than exceptions. We suspect that these analytical tests may suffer from possessing poor power in the presence of outliers. In this study we first investigate how the commonly used heteroscedastic tests perform in the

---

**Corresponding Author:** Md. Sohel Rana, Laboratory of Applied and Computational Statistics, Institute for Mathematical Research, University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

presence of outlier. We observe that all the tests considered in our study suffer in this situation and for this reason we robustify the Goldfeld-Quandt test. Real data sets and simulation experiments show that the proposed modified Goldfeld-Quandt test outperforms other tests in detecting heteroscedasticity in the presence of outliers.

An excellent review of different analytical tests for the detection of heteroscedasticity is available in Kutner, Nachtsheim and Neter<sup>[10]</sup> and in Chatterjee and Hadi<sup>[3]</sup>. In our study we consider the tests which are very popular and commonly used in econometrics. First we consider the Goldfeld-Quandt test. This test is applicable if one assumes that the heteroscedastic variance  $\sigma_t^2$ 's is positively related to one of the explanatory variables in the model. For simplicity, let us consider the usual two-variable model:

$$Y_t = \alpha + \beta X_t + u_t \tag{1}$$

Suppose  $\sigma_t^2$  is positively related to  $X_t$  as:

$$\sigma_t^2 = \sigma^2 X_t^2 \tag{2}$$

where,  $\sigma^2$  is a constant. Such an assumption has been found quite useful in family budgets. If (2) is appropriate, it would mean  $\sigma_t^2$  would be larger for the larger values of  $X$ . If that turns out to be the case, heteroscedasticity is most likely to be present in the model. To test this explicitly, Goldfeld and Quandt<sup>[9]</sup> suggest ordering the observations according to the values of  $X$ , beginning with the lowest  $X$  value. To possess better power they suggest omitting  $c$  central observations. OLS regressions are fitted separately to the first and last  $(n-c)/2$  observations and the respective residual sum of squares  $RSS_1$  and  $RSS_2$  are obtained. Under normality of errors, each  $RSS$  follows a chi-square distribution with  $(n-c-2k)/2$  degrees of freedom, where  $k$  is the number of parameters to be estimated, including the intercept. Then the ratio:

$$\lambda = \frac{RSS_2 / df_2}{RSS_1 / df_1} = \frac{RSS_2}{RSS_1} \tag{3}$$

Under the assumption of normality and homoscedasticity  $\lambda$  follows an  $F$  distribution with numerator and denominator d.f. each of  $(n-c-2k)/2$ .

The Goldfeld-Quandt test is a natural test to apply when one can order the observations in terms of the increasing variance of the error term (or one independent variable). An alternative test, which does

not require such an ordering and is easy to apply, is the Breusch-Pagan test. To illustrate this test, consider the  $k$ -variable linear regression model:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t \tag{4}$$

Assume that the error variance  $\sigma_t^2$  is described as:

$$\sigma_t^2 = f(\alpha_1 + \alpha_2 Z_{2t} + \dots + \alpha_m Z_{mt}) \tag{5}$$

that is,  $\sigma_t^2$  is some function of the nonstochastic variable  $Z$ 's; some or all of the  $X$ 's can serve as  $Z$ 's. We also assume that:

$$\sigma_t^2 = \alpha_1 + \alpha_2 Z_{2t} + \dots + \alpha_m Z_{mt} \tag{6}$$

that is,  $\sigma_t^2$  is a linear function of the  $Z$ 's. If  $\alpha_2 = \alpha_3 = \dots = \alpha_m = 0$ , we get  $\sigma_t^2 = \alpha_1$ , which is a constant. Therefore, to test whether  $\sigma_t^2$ 's are homoscedastic, one can test the hypothesis that  $\alpha_2 = \alpha_3 = \dots = \alpha_m = 0$ . This is the basic idea behind the Breusch-Pagan test. In this test we estimate the model (6) by the least squares method and obtain the residuals. Then the mean of squared least square residuals  $\bar{\hat{\sigma}}^2 = \sum \hat{u}_t^2 / n$  are computed and the variable  $p_t = \frac{\hat{u}_t^2}{\bar{\hat{\sigma}}^2}$  is constructed. Next we regress  $p_t$  on the  $Z$ 's to obtain the SSE (error sum of squares). Finally we compute the test statistic:

$$T = SSE/2 \tag{7}$$

Under the assumption of normality and homoscedasticity  $T \sim \chi_{m-1}^2$  and if the value of  $T$  exceeds the critical value, we conclude that heteroscedasticity is present in the data.

Unlike the Goldfeld-Quandt test, which requires reordering the observations with respect to the  $X$  variable that supposed to cause heteroscedasticity, or the Breusch-pagan test, which is sensitive to the normality assumption, the general test of heteroscedasticity proposed by White<sup>[20]</sup> does not rely on the normality assumption and is very easy to implement. As an illustration of the basic idea, consider the following three-variable regression model (the generalization to the  $k$ -variable model is straightforward). Given the data, we estimate regression parameters by the OLS method and obtain the residuals.

We then run the following (auxiliary) regression:

$$\hat{u}_t^2 = \alpha_1 + \alpha_2 X_{2t} + \alpha_3 X_{3t} + \alpha_4 X_{2t}^2 + \alpha_5 X_{3t}^2 + \alpha_6 X_{2t} X_{3t} + e_t \tag{8}$$

where  $e_t$  is the random error term. In other words, the squared residuals from the original regression are regressed on the original X variables or regressor, their squared values and the cross product(s) of the regressors. Higher power of the regressors can also be introduced. It is important here to note that there is a constant term in this equation even though the original regression may or may not contain it. Now we obtain the  $R^2$  from this (auxiliary) regression. Under the null hypothesis that there is no heteroscedasticity, it can be shown that sample size (n) times the  $R^2$  obtained from the auxiliary regression asymptotically follows the chi-square distribution with d.f. equal to the number regressors (excluding the constant term) in the auxiliary regression as given in (8). That is, for a regression model with p regressors:

$$n.R^2 \sim \text{Asymptotic } \chi_a^2 \tag{9}$$

The degrees of freedom  $a = 2^p+p-1$ . If the chi-square value thus obtained exceeds the critical chi-square value at the chosen level of significance, the conclusion is that there exists heteroscedasticity.

**MATERIALS AND METHODS**

**Modified Goldfeld-Quandt test:** we have briefly discussed some popular tests for heteroscedasticity detection. But there is evidence that all these tests suffer a huge setback when outliers are present in the data. So we need to develop a test which is not much affected by outliers. Here we propose a new test which is a modification of the Goldfeld-Quandt test. We first identify which components of the Goldfeld-Quandt test are affected by outliers and then replace these components by robust alternatives. It is worth mentioning that this kind of replacement does not help the other two tests that we consider in the previous section. We call this test the Modified Goldfeld-Quandt (MGQ) test which, we believe, will be more powerful than the existing tests in the presence of outliers.

Here we outline the proposed modified Goldfeld-Quandt test. This test contains the following steps:

- Step 1:** Likewise the classical Goldfeld-Quandt test, we order or rank the observations according to the value of X, beginning with the lowest X values
- Step 2:** We omit central c observations, where c is specified a priori and then we divide the remaining (n-c) observations into two groups each of (n-c)/2 observations
- Step 3:** Check for the outliers by any robust regression technique. We prefer to use the robust Least

Trimmed of Squares (LTS) method suggested by Rousseeuw and Leroy<sup>[13]</sup> to fit the regression line. We compute the deletion residuals<sup>[8]</sup> for the entire data set based on a fit without the points identified as outliers by the LTS fit.

**Step 4:** For both the groups compute the Median of the Squared Deletion Residuals (MSDR) and compute the ratio

$$MGQ = \frac{MSDR_2}{MSDR_1} \tag{10}$$

where,  $MSDR_1$  and  $MSDR_2$  are the median of the squared deletion residuals for the smaller and the larger group variances respectively. Under normality, the MGQ statistic follows an F distribution with numerator and denominator degrees of freedom each of  $(n-c-2k)/2$ .

**RESULTS AND DISCUSSION**

**Numerical examples:** Here we present few examples to show the advantage of using the modified Goldfeld-Quandt test in the detection of heterogeneity of error variances problem.

**Housing expenditures data:** Our first example is the housing expenditures data given by Pindyck and Rubinfeld<sup>[12]</sup>. This single-predictor data set (Table 1) contains 20 observations that give housing expenditure for four different income groups.

As expected, peoples with higher income have relatively more variation in their expenditures on housing. The scatter plot as shown in Fig. 1 clearly exhibits the heteroscedastic pattern of the data.

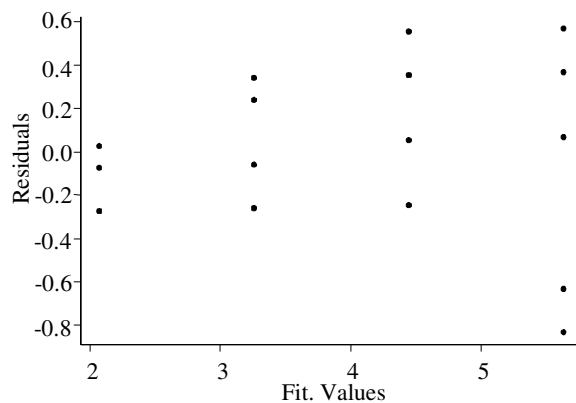


Fig. 1: Residuals vs. fitted plot for original housing expenditures data

Table 1: Original and modified housing expenditures data

Index	Income	Housing Exp.	Index	Income	Housing Exp.
1	5	1.8 (4.9)	11	15	4.2
2	5	2.0	12	15	4.2
3	5	2.0	13	15	4.5
4	5	2.0	14	15	4.8
5	5	2.1	15	15	5.0
6	10	3.1	16	20	4.8
7	10	3.2	17	20	5.0
8	10	3.5	18	20	5.7
9	10	3.5	19	20	6.0
10	10	3.6	20	20	6.2 (2.0)

Table 2: Heteroscedasticity diagnostics for housing expenditures data

Test	Without outliers		With outliers	
	Value of Statistic	p-value	Value of Statistic	p-value
Goldfeld-Quandt	9.3817	0.0077	1.5128	0.3139
Breusch-Pagan	6.8660	0.0088	0.8387	0.3598
White	7.2003	0.0073	0.3416	0.5589
MGQ	5.8450	0.0248	10.8055	0.0034

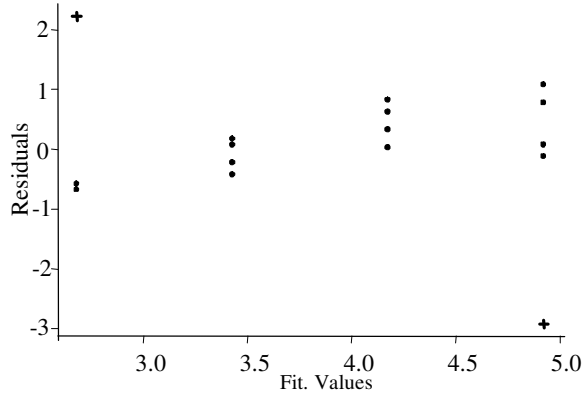


Fig. 2: Residuals vs. fitted plot for modified housing expenditures data

We now deliberately put two outliers into the data set by replacing the housing expenditures of the cases indexed by 1 and 20 (modified values are presented within the parentheses).

Figure 2 shows the residuals-fits plot for the housing expenditures data. This plot is not as clear as Fig. 1 in exhibiting variance heteroscedasticity and we definitely need analytical tests to draw a definite conclusion.

We apply all the conventional tests like the Goldfeld-Quandt, the Breusch-Pagan and the White tests on the original and modified housing expenditures data and the results are presented in Table 2. We observe from this table that all the conventional tests variance when the data set is free from outliers.

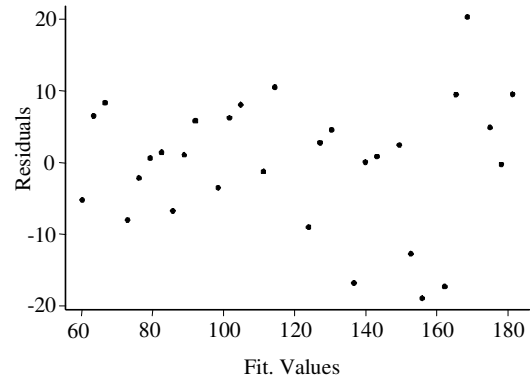


Fig. 3: Residuals vs. fitted values for original consumption expenditure data

But all these tests fail to detect the problem of heteroscedasticity when outliers occur in the data set. We now apply our proposed modified Goldfeld-Quandt test on the original and modified housing expenditures data and these results are presented in Table 2. We observe from this table that the modified Goldfeld-Quandt test performs in a similar way as the Goldfeld-Quandt test when there is no outlier. But unlike the other tests, it can successfully detect the heteroscedasticity in the presence of outliers yielding a highly significant p-value.

**Consumption expenditure data:** Our next example is the consumption expenditure and income data given by Gujarati<sup>[6]</sup>.

This data contains 30 observations and it shows that the expenditure of peoples vary with their income. So we can guess that in this data the variation is not constant. We now deliberately put three outliers into the data set by replacing the housing expenditures of the cases indexed by 1, 2 and 30 (modified values are presented within the parentheses). This data set together with the outliers is shown in Table 3.

Figure 3 and 4 shows the residuals-fits plot for the original and modified consumption data. The variance heterogeneity is clearly visible with the original data but when outliers are present in the data this phenomenon is not clearly visible.

Table 4 offers a comparison between the newly proposed modified Goldfeld-Quandt test and other existing tests in the detection of heteroscedasticity for the consumption data. Table 4 shows that the Goldfeld-Quandt, the Breusch-Pagan and the White tests can correctly identify the heteroscedastic pattern of variance when the data set is free from outliers but they become unsuccessful in the presence of outliers. The modified Goldfeld-Quandt test can successfully detect the heteroscedasticity on both the occasions.

Table 3: Original and modified consumption expenditure data

Index	Expenditure	Income	Index	Expenditure	Income	Index	Expenditure	Income
1	55 (10)	80	11	74	105	21	152	220
2	65(10)	100	12	110	160	22	144	210
3	70	85	13	113	150	23	175	245
4	80	110	14	125	165	24	180	260
5	79	120	15	108	145	25	135	190
6	84	115	16	115	180	26	140	205
7	98	130	17	140	225	27	178	265
8	95	140	18	120	200	28	191	270
9	90	125	19	145	240	29	137	230
10	75	90	20	130	185	30	189 (100)	250

Table 4: Heteroscedasticity diagnostics for consumption expenditure data

Test	Without outliers		With outliers	
	Value of Statistic	p-value	Value of Statistic	p-value
Goldfeld-Quandt	3.8895	0.0215	1.0149	0.4909
Breusch-Pagan	5.214	0.0224	0.1856	0.6666
White	5.2722	0.0217	0.0697	0.7917
MGQ	2.9066	0.0537	6.2216	0.0039

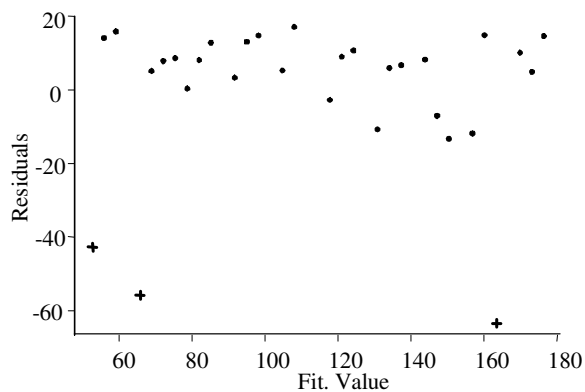


Fig. 4: Residuals vs. fitted values for modified consumption expenditure data

**Restaurant food sales data:** Finally we consider restaurant food sales data given by Montgomery *et al.*<sup>[11]</sup>

In this data set there is a relation of income with advertising expense. Again we deliberately put three outliers into the data set by replacing the income of the cases indexed by 1, 27 and 30 (modified values are presented within the parentheses). The original and the modified data are shown in Table 5.

Figure 5 and 6 show the residuals-fits plot for the original and modified consumption data. The variance heterogeneity is clearly visible with the original data but when outliers are present in the data this phenomenon is not clearly visible.

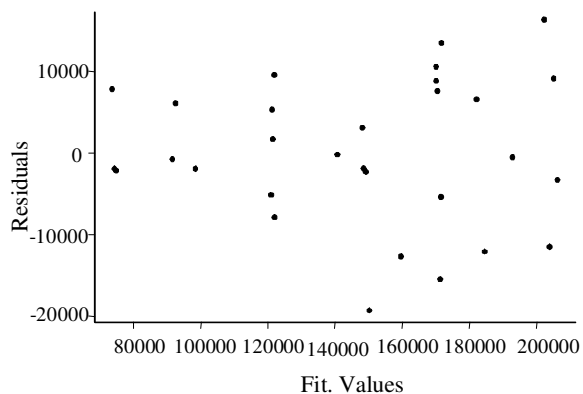


Fig. 5: Residuals vs. fitted plot for original restaurants food sales data

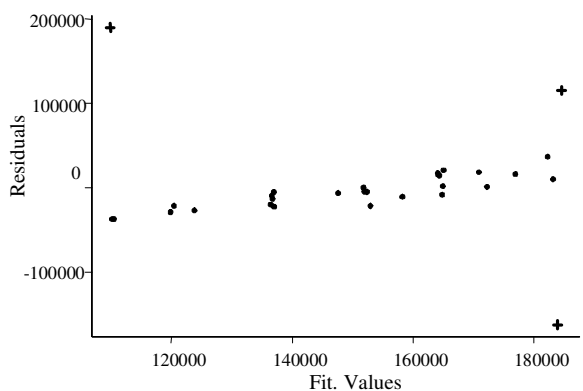


Fig. 6: Residuals vs. fitted plot original restaurants food sales data

Likewise the previous examples we employ the Goldfeld-Quandt, the Breusch-Pagan, White and modified Goldfeld-Quandt test to the restaurants food sales data and obtain similar results that we got earlier. Test results as shown in Table 6 shows that the three conventional tests perform well in detection of heteroscedasticity but their performances become poor when outliers are present in the data.

Table 5: Original and modified restaurant food sales data

Index	Income	Ad. Exp.	Index	Income	Ad. Exp.	Index	Income	Ad. Exp.
1	81464 (300000)	3000	11	131434	9000	21	178187	15050
2	72661	3150	12	140564	11345	22	185304	15200
3	72344	3085	13	151352	12275	23	155931	15150
4	90743	5225	14	146926	12400	24	172579	16800
5	98588	5350	15	130963	12525	25	188851	16500
6	96507	6090	16	146630	12310	26	192424	17830
7	126574	8925	17	147041	13700	27	203112 (300000)	19500
8	114133	9015	18	179021	15000	28	192482	19200
9	115814	8885	19	166200	15175	29	218715	19000
10	123181	8950	20	180732	14995	30	214317 (21431)	19350

Table 6: Heteroscedasticity diagnostics for restaurants food sales data

Test	Without outliers		With outliers	
	Value of statistic	p-value	Value of statistic	p-value
Goldfeld-Quandt	4.03671	0.019	1.074	0.4563
Breusch-Pagan	3.1787	0.0746	0.3799	0.5376
White	4.3575	0.0368	0.0963	0.7562
MGQ	4.9917	0.0090	10.4566	0.0005

The modified Goldfeld-Quandt test performs best. Irrespective of the presence of outliers it can successfully detect the heteroscedastic error variance in the data.

**Simulation Study:** Now from our experience with individual data sets we want to confirm our results by reporting a Monte Carlo simulation experiment. In our simulation experiment, we consider a design of 5 and 10% outliers in heteroscedastic data. Here we consider a simple but interesting heteroscedastic variance problem where the variance is the square of the mean of the response variable.

Let us consider a simple two variable linear model:

$$Y = 4 + 5 X + \epsilon \tag{11}$$

In our simulation study, all the values of X are being taken equally spaced such as 1, 2,..., 10 and these values are replicated several times to get higher sample sizes. We generate the random errors from Normal distributions with mean 0 and standard deviations X, 2X and 3X. We put outliers in the error term in every 20th or 10th position to generate 5 and 10% outliers respectively. The magnitude of the

outlier is 5 times the standard deviation of the original errors. The Y values are obtained from the Eq. 11. We run this simulation experiment for five different sample sizes n = 20, 30, 40, 60 and 100. To assess which of the tests does the best in detecting heteroscedasticity in the presence of outliers we consider powers of four tests, the conventional Goldfeld-Quandt, Breusch-Pagan and White tests and the newly proposed modified Goldfeld-Quandt test. For each test we set the level of significance 0.05 and the results of this experiment are shown in Table 7-9 each of which is based on the average of 10,000 simulations.

Table 7-9 offer comparisons between the newly proposed modified Goldfeld-Quandt test and conventional Goldfeld-Quandt, Breusch-Pagan and White tests in the detection of heteroscedasticity for the 5 and 10% outlier data. All the three conventional tests perform very poorly in simulation. The Goldfeld-Quandt test performs relatively well for 5% outliers. The Bruesch-Pagan test performs relatively well for 10% outlier cases but its performance tends to deteriorate with the increase in sample size. The White test performs worst in every situation. Throughout the simulation experiment each of the conventional tests shows inconsistency pattern for the sample size n = 30. But the newly proposed modified Goldfeld-Quandt test performs superbly throughout. For small sample size (n = 20) and lower contamination (5%) its performance is similar to the Goldfeld-Quandt test, but its power tends to increase with the increase in sample size. We also observe that the test is robust in the sense that it performs exactly in the same way when outliers occur in a data with different levels of error variances. Thus the modified Goldfeld-Quandt test outperforms the conventional tests in every respect and is proved to be the best overall.

Table 7: Simulations results of heteroscedasticity tests for error variance =  $X^2$

Test	5% Outliers					10% Outliers				
	20	30	40	60	100	20	30	40	60	100
GQ	84.50	0	57.83	94.58	100.00	100.00	5.80	0.94	1.20	0.00
BP	0.00	0	0.00	22.90	94.90	79.64	18.26	38.34	7.12	5.30
White	0.00	0	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
MGQ	82.58	95	97.58	99.52	100.00	88.16	97.20	99.30	99.88	100.00

Table 8: Simulations results of heteroscedasticity tests for error variance =  $4X^2$

Test	5% Outliers					10% Outliers				
	20	30	40	60	100	20	30	40	60	100
GQ	85.00	0.00	58.36	94.56	100.00	100.00	0.80	5.78	1.22	0.00
BP	0.02	0.00	0.00	22.95	90.18	79.74	38.42	18.96	7.52	4.50
White	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00
MGQ	83.12	95.34	97.50	99.70	100.00	86.98	97.14	99.10	99.92	100.00

Table 9: Simulations results of heteroscedasticity tests for error variance =  $9X^2$

Test	5% Outliers					10% Outliers				
	20	30	40	60	100	20	30	40	60	100
GQ	86.28	0.00	57.12	95.06	100.00	100.00	0.96	6.30	1.21	0.00
BP	0.00	0.00	0.00	24.56	90.32	78.82	36.80	18.08	7.96	4.46
White	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00
MGQ	83.40	96.40	97.72	99.66	100.00	87.36	97.46	99.12	99.94	100.00

### CONCLUSION

In this research we show that all commonly used tests for detecting heteroscedasticity fail when outliers are present in the data. We develop a new test in this regard which is a simple but robust modification of the Goldfeld-Quandt test. The real data sets and Monte Carlo simulations show that modified Goldfeld-Quandt test offers substantial improvements over the existing tests and performs superbly in the detection of heteroscedasticity in the presence of outliers.

### REFERENCES

1. Breusch, T. and A. Pagan, 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47: 1287-1294. <http://ideas.repec.org/a/econ/emetrp/v47y1979i5p1287-94.html>
2. Carroll, R.J. and D. Ruppert, 1988. *Transformations and Weighting in Regression*. 2nd Edn., Wiley, New York, pp: 249. ISBN: 0412014211.
3. Chatterjee, S. and A.S. Hadi, 2006. *Regression Analysis by Examples*. 4th Edn., Wiley, New York, pp: 375. ISBN: 9780471746966.
4. Draper, N.R. and H. Smith, 1998. *Applied Regression Analysis*. 3rd Edn., Wiley, New York, pp: 736. ISBN: 9780471170822.
5. Goldfeld, S.M. and R.E. Quandt, 1965. Some tests for homoskedasticity. *J. Am. Stat. Assoc.*, 60: 539-547. <http://www.belkcollege.uncc.edu/cdepken/econ6090/readings/goldfeld-quandt-1965.pdf>
6. Gujarati, D., 2002. *Basic Econometrics*. 4th Edn., McGraw-Hill, New York, pp: 1002. ISBN: 0071123431.
7. Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw and W. Stahel, 1986. *Robust Statistics: The Approach Based on Influence Function*. 1st Edn., Wiley, New York, pp: 536. ISBN: 0471735779.
8. Imon, A.H.M.R., 2003. Residuals from deletion in added variable plots. *J. Applied Stat.*, 30: 827-841. DOI: 10.1080/0266476032000076083
9. Greene, W.H., 2008. *Econometric Analysis*. 6th Edn., Pearson Education, Inc., pp: 1178. ISBN: 0135137403.
10. Kutner, M.H., C.J. Nachtsheim and J. Neter, 2004. *Applied Linear Regression Models*. 4th Edn., McGraw-Hill/Irwin, New York, pp: 701. ISBN: 0-07-301344-7.
11. Montgomery, D.C., E.A. Peck and G.G. Vining, 2001. *Introduction to Linear Regression Analysis*. 3rd Edn., Wiley, New York, pp: 641. ISBN: 0-471-31565-6.
12. Pindyck, S.R. and L.D. Rubinfeld, 1998. *Econometric Models and Econometric Forecasts*. 4th Edn., Irwin/McGraw-Hill, New York, pp: 634. ISBN: 0071158367.
13. Rousseeuw, P.J. and A. Leroy, 1987. *Robust Regression and Outlier Detection*, 1st Edn., Wiley, New York, pp: 329. ISBN: 0471852333.
14. Ryan, T.P., 2008. *Modern Regression Methods*. 2nd Edn., Wiley, New York. pp: 664. ISBN: 0470081864.
15. White, H., 1980. Heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, 48: 817-838.