DOI: https://dx.doi.org/10.4314/gjpas.v29i1.10

GLOBAL JOURNAL OF PURE AND APPLIED SCIENCES VOL. 29, 2023: 83-90
COPYRIGHT© BACHUDO SCIENCE CO. LTD PRINTED IN NIGERIA ISSN 1118-0579
www.globaljournalseries.com.ng, Email: globaljournalseries@gmail.com

HETEROSCEDASTICITY OF UNKNOWN FORM: A COMPARISON OF FIVE HETEROSCEDASTICITY-CONSISTENT COVARIANCE MATRIX (HCCM) ESTIMATORS

NWANGBURUKA, C., IJOMAH, M. A. AND NWAKUYA M.T

(Received 22 June 2022; Revision Accepted 5 November 2022)

ABSTRACT

Regression model applications frequently involve violations of the homoscedasticity assumption and the presence of high leverage points (HLPs). The Heteroscedasticity-Consistent Covariance Matrix (HCCM) estimator's impact in the presence of heteroscedasticity of an unknown form was investigated in this study. The effectiveness of five variations of HCCM namely White's estimator (HC0), White-Hinkley (HC1), Mackinnon White (HC2), Davison –Mackinnon (HC3), and Cribari-Neto (HC4) were accessed to identify the optimal Heteroscedasticity-Consistent Covariance Matrix (HCCM) estimator. In the study a simulated dataset was analysed using the Econometric View Software Version 12. The Breush-Pagan Godfery's test for heteroscedasticity was applied and p-value of 0.0123 was obtained indicating presence of heteroscedasticity in the model. Applying the HCCM estimators and comparing the Heteroskedasticity-consistent standard errors estimates showed that HCO had 124.104, HC1 had 1189.222, HC2 had 1175.282, HC3 had 1106.94 and HC4 had 1140.707. These results reveal that HC3 and HC4 produced smaller errors compared to HC0, HC1 and HC2. The study hence comes to the conclusion that when doing inferential tests using OLS regression, the use of HCSE estimator increases the researcher's confidence in the accuracy and potency of those tests. This study therefore suggests that to ensure that findings are not affected by heteroscedasticity; researchers should use HCCM estimator but precisely HC3 and HC4, as the presented better results in comparison to others.

KEY WORDS: Heteroscedasticity, White (HC0), White-Hinkley (HC1), Mackinnon-White (HC2), Davidson-Mackinnon (HC3), and Cribari-Neto (HC4)

INTRODUCTION

It is generally known that ordinary least squares (OLS) deliver accurate and unbiased estimates of the parameters when the linear regression model's underlying assumptions are true. Even though the OLS estimator maintains its objectivity when the errors are heteroscedastic, it becomes inefficient.

More importantly, the traditional methods for testing hypotheses no longer work hence methods that compensate for heteroscedasticity are essential for careful data exploration for it is widespread in cross-sectional data. Many statistical techniques weight each observation by the inverse of the standard deviation of the error to account for heteroscedasticity (Greene, 2002). The resulting coefficient estimates are accurate and unbiased, and the standard errors are also accurately estimated.

According to (Weisberg, 1980) "the generalized least squares makes it relatively easy to use weights to correct for heteroscedasticity when the type and degree of heteroscedasticity are known". Most times, when the heteroscedasticity type is unknown, the weighting strategy becomes useless. "Making variance-stabilizing transformations of the dependent variable or altering both sides can be utilised to address heteroscedasticity that results from an improper functional form" (Carroll and Ruppert, 1988). "While this strategy can effectively and elegantly address the issues brought on by heteroscedasticity, nonparametric techniques may be required when the results must be interpreted in the variables' original scale" (Duan 1983; Carroll and Ruppert, 1988).

Even if its exact form is uncertain, heteroscedasticity must be addressed. The most detrimental effects of

Nwangburuka, C., Department of Mathematics and Statistics, University of Port Harcourt
Ijomah, M. A., Department of Mathematics and Statistics, University of Port Harcourt
Nwakuya M.T., Department of Mathematics and Statistics, University of Port Harcourt

heteroscedasticity is that it leads to biased and inconsistent behaviour in the OLS estimator of the parameter covariance matrix (OLSCM), whose diagonal elements are used to estimate the standard errors of the regression coefficients. Therefore, it is against this background that this study tried to find an alternative variance estimator that remains consistent under heteroscedasticity.

LITERATURE REVIEWS

Astivia *et al* (2019) in their work presented a comprehensible explanation and illustration heteroscedasticity. They clearly showed how to detect it through statistical tests and also how to take care of it by using heteroskedastic-consistent standard errors and the wild bootstrap.

Vynck and Thas (2017) discovered that the conventional method of utilising WLS for dealing with heteroscedasticity had significant flaws. Even though it accounts for heteroscedasticity by weighting the observations in a way that is as appropriate as it is possible, the ideal situation in which the variance function is known is frequently not known. Agunbiade and Adeboye (2012) used the White heteroscedasticity and Newey-West test techniques to examine the presence of heteroscedasticity. Their findings, which show that heteroscedasticity is a built-in characteristic of cross-sectional data, were published in the journal Statistics in Medicine. It was determined that OLS is not appropriate for estimation if heteroscedasticity is present in the research data. It was also determined that the model fitted using WLS is the most appropriate model that is deemed fit for proper review of auditor's remuneration in the banking industry. In their study on "inference when there is heteroscedasticity of unknown form", Munir *et al.* (2011) discovered through the utilisation of an adaptive estimator that the tests that are based on these estimators are not as liberal as the tests that are based on the OLSEs and the adaptive estimators. These weighted HCCMEs demonstrate null rejection rates that are relatively low and it has been reported that WHC3 performs better than WHC2, which performs better than WHC1, which performs better than WHC0. WHC0, performs worse than WHC3.

METHODOLOGY

Data used for this study was a simulated replication of (Venables and Ripley, 2002) simulation of 237 statistics students at the University of Adelaide in Modern Applied Statistics with S-PLUS, with 11 predictor variables with pulse rate as the response variable. The simulation design was based on a number of questions regarding the individual qualities of the students. These individual qualities were student's sex, age, writing hand, span of writing hand, span of non-writing hand, fold arms, pulse rate of the student, clap hands, frequency of exercise, how much the student smokes, and height. Firstly, a Diagnostic test (Breush-Pagan Godfrey test) for detecting heteroscedasticity was performed using the Econometric-View Software (E-views) version 12.0 before comparison of the Heteroscedasticity-Consistent Covariance Matrix (HCCM) Estimators.

Breusch-Pagan Godfrey test

The Breusch-Pagan test was developed in 1979 by Trevor Breusch and Adrian Pagan. The Breusch-Pagan test is employed when it is believed that the variance is some function (though not necessarily multiplicative) of more than one explanatory variable. Considering the following regression model

$$y_i = \beta_1 + \beta_2 X_{i1} + \cdots + \beta_k X_{ik} + \mu_i \quad (1)$$

Equation (1) can be rewritten as;

$y = \boldsymbol{X}\beta + u$, where $u \sim N(0, \sigma^2 I_n)$, $y \sim N(\boldsymbol{X}\beta, \sigma^2 I_n)$ and $\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1}$ and

Var $(\mu_i) = \sigma_i^2$

The Breusch-Pagan LM test involves a series of intermediate stages in detecting heteroscedasticity. First, it implemented the regression of the previous equation and the residuals $\hat{u}_i$ were obtained. Subsequently, the auxiliary regression equation is established, as the following

$$u_i^2 = \beta_1 + \beta_2 Z_{2t} + \beta_3 Z_{3t} + \cdots + \beta_p Z_{pt} + e_t \quad (2)$$

where Z_{pt} is a series of variables established to determine the variance of the error terms. The next step involves setting the null hypothesis of homoskedasticity as

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_n = 0$

In the case that at least one of the β's is different from zero and at least one of the Z's influences the variance of the error terms, the null hypothesis is rejected. The following step is to compute the $LM = nR^2$ statistic, where n is the number of observations established to determine the auxiliary regression and R^2 is the coefficient of determination. The LM-statistic follows the χ^2 distribution characterized by m-1 degrees of freedom. The final step assumes to reject the null hypothesis and to highlight the presence of heteroscedasticity if LM-statistical is higher than the critical value OR when the p-value obtained is less than the critical level of $\alpha = 0.05$, we reject the null hypothesis which says that "There is no heteroscedasticity in the model". Hence, conclude that Heteroscedasticity is present.

Comparison of Heteroscedasticity-consistent covariance matrices.

Since its inception in econometrics (White, 1980), heteroscedasticity-consistent covariance matrices (HCCMs) have received more attention in the literature on statistics and behavioural sciences. Simply put, this method adjusts the parameter estimates' covariance matrix. The unbiased nature of regression coefficients prevents them from being changed or corrected. There are now five categories of HCCMs: HC0 – HC4. In general, later variations aimed to enhance the original HCCMs.

White (HC0)

Based on the foundation White's (1980) method, HC0 became a popularisation estimator. The formula for the HC0 estimator of $\sum\hat{\beta}$ is given as

$$HCO = (X'X)^{-1}X'diag(e_i^2)X(X'X)^{-1} \quad (3)$$

and the entries on the main diagonal of HC0 are the estimated squared standard errors of the regression coefficients. Dividing the regression coefficients by these standard errors produces a ratio used to derive p values for hypothesis testing. "However, HC0 is a reliable estimator when the errors are heteroskedastic, meaning that the bias gets smaller as the sample size grows. Although they are all asymptotically identical to HC0, HC1, HC2, and HC3 have significantly better small sample properties", (Long & Ervin, 2000; MacKinnon & White, 1985).

White-Hinkley (HC1)

Hinkley (1977) developed the estimator HC1, which is only a degree-of-freedom adjustment to HC0. Every squared OLS residual for HC1 is compounded by n/ (n-p-1). The HC1 estimator of $\sum\hat{\beta}$ is;

$$\text{HC1} = \frac{n}{n-p-1}(X'X)^{-1}X'diag[e_i^2]\,X(X'X)^{-1} \quad (4)$$

Despite being straightforward, HC1 is rarely suggested or employed since it shares many of the same finite sample biases as HC0.

Mackinnon-White (HC2)

The theory of HC2 is similar to HC1, except instead of a correction for degree of freedom, the i^{th} squared OLS residual is weighted by the inverse of ($1-h_{ii}$), where $h_{ii} = X_i(X'X)^{-1}X_i'$

"The "hat" matrix H = $X(X'X)^{-1}X'$ has diagonal elements called $h_{ii's}$ which are also referred to as leverage values". (Cribari-Neto 2004). The high leverage spots in the X matrix are the main cause of the bias. Consequently, the HC2 estimator of $\sum\hat{\beta}$ is defined as;

$$HC2 = (X'X)^{-1}X'diag\left[\frac{e_i^2}{1-h_{ii}}\right]X(X'X)^{-1} \quad (5)$$

Davidson-Mackinnon (HC3)

For historical context, it should be noted that the HC methods have close ties to the jackknife approach. The existence or absence of high leverage points in X does have some bearing on how well HC3 performs (Kauermann & Carroll, 2001; Wilcox, 2001). Furthermore, studies have demonstrated that HC3 can have a liberal bias in relatively small samples (Long & Ervin, 2000).

The HC3 estimator of $\sum\hat{\beta}$ is defined as

$$HC2 = (X'X)^{-1}X'diag\left[\frac{e_i^2}{(1-h_{ii})^2}\right]X(X'X)^{-1} \quad (6)$$

Notice that HC3 weights each squared OLS residual by a factor of $\frac{1}{(1-h_{ii})^2}$ rather than $\frac{1}{1-h_{ii}}$.

Cribari-Neto (HC4)

The HC4 proposed by Cribari-Neto (2004) was built under HC3, and is defined as follows:

$$HC4 = (X'X)^{-1}X_i'\hat{\Phi}_4X(X'X)^{-1} \quad (7)$$

where

$$\hat{\Phi}_4 = diag\left[\frac{e_i^2}{(1-h_i)^{\delta_i}}\right]$$

Similar to this, Cribari-Neto et al. (2007) suggested a different tweak to the exponent (1 – *hi*) of HC4 to determine the degree of maximal leverage.

Wald test

A parametric statistical test called the Wald test can determine whether a group of independent variables is considered to be "significant" for a model or not.

If a variable improves the model in any way, it is seen to be "significant." Variables that don't contribute to the model's value can be removed without the model being significantly affected. The Wald Test statistic formula is:

$$W_T = \frac{[\hat{\theta}-\theta_0]^2}{1/I_n(\hat{\theta})} = I_n(\hat{\theta})[\hat{\theta} - \theta_0]^2 \quad (8)$$

Where $\hat{\theta}$ = Maximum Likelihood Estimator (MLE)

$I_n(\hat{\theta})$ = expected fisher information (evaluated at the MLE)

Basically, the test looks for differences: $\Theta^0 - \Theta$. The general steps are:

1. Find the MLE.
2. Find the expected Fisher information.
3. Evaluate the Fisher information at the MLE.

With the combination of the MLE and Fisher information, the Wald test is very complex to work and is not usually calculated by hand. Many software applications can run the test as in the case of this study.

RESULTS

Table 1: Test for presence of Heteroscedasticity using Breush-Pagan Godfrey test Heteroskedastidty Test: Breusch-Pagan-Godfrey Null hypothesis: Homoskedasticity

F-statistic	2.331532	Prob. F (10,226) 0.0124
Obs*R-squared	22.16362	Prob. Chi-Square (10) **0.0143**
Scaled explained SS	1074.836	Prob. Chi-Square(10) 0.0000

Source: E-view 12 output

Table 1 above demonstrated that there was a presence of heteroskedasticity in the dataset by using the Breush-Pagan Godfrey test. The Prob-Chi-square value obtained for all ten variables were below the critical value of = 0.05. Given this, the "null hypothesis," which asserts that the model does not contain any heteroscedasticity, is shown to be false, which indicates that heteroscedasticity does exist in the data.

Table 2: OLS regression analysis Estimating Pulse rate using standard error estimates.
Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AGE	-0.004655	0.025190	-0.184785	0.8536
CLAP	0.477972	0.199601	2.394642	**0.0175**
EXER	0.008630	0.175477	0.049178	0.9608
FOLD	0.308323	0.255481	1.206832	0.2288
HEIGHT	1.018928	0.092233	11.04733	**0.0000**
NW.HND	0.068170	0.098641	0.691096	0.4902
SEX	0.232397	0.388943	0.597510	0.5508
SMOKE	0.134911	0.191043	0.706181	0.4808
W_HND	-2.682422	0.538150	-4.984525	**0.0000**
WR HND	-0.371974	0.471016	-0.789727	0.4305
C	-95.11930	25.15933	-3.780677	0.0002
R-squared	0.952640	Mean dependent var		73.94093
Adjusted R-squared	0.950544	S. D. dependent var		10.88211
S.E. of regression	2.420033	Akaike into criterion		4.650741
Sum squared resid	1323.582	Schwarz criterion		4.811706
Log-likelihood	-540.1128	Hannan-Quinn criter.		4.715620
F-statistic	454.5945	Durbin-Watson stat		0.698392
Prob(F-statistic)	0.000000		..	

Source: E-view 12 output

Table 2 revealed the OLS regression estimates and the estimated standard errors and p values for every regression coefficient. Pulse rate is significantly related to clap, height and w_hnd with p-values that are lower than the critical level of α = 0.05, which leads to the rejection of the null hypothesis. This agrees with (Berry, 1993) study which revealed that the effects of assumption violations in such a scenario can move standard errors and p-values in unpredictable ways.

Table 3: White (HC0) analysis Estimating Pulse rate using standard error estimates White (HC0) heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std.Error	t-Statistic	Prob.
AGE	-0.004655	0.010208	-0.455982	0.6488
CLAP	0.477972	0.236649	2.019754	**0.0446**
EXER	0.008630	0.132496	0.065131	0.9481
FOLD	0.308323	0.319730	0.964324	0.3359
HEIGHT	1.018928	0.160266	6.3577 14	**0.0000**
NW_HND	0.068 170	0.114769	0.593981	0.5531
SEX	0.232397	0.560364	0.414725	0.6787
SMOKE	0.134911	0.128 123	1.052980	0.2935
W_HND	-2.682422	1.458466	-1.839208	0.0672
WR_HND	-0.371974	0.852508	-0.436329	0.6630
c	-95.11930	45.84260	-2.074911	0.0391
R-squared	0.952640	Mean dependent var		73.94093
Adjusted R-squared	0.950544	S.D. dependent var		10.88211
S.E. of regression	2.420033	Akaike info criterion		4.650741
Sum squared resid	1323.582	Schwarz criterion		4.811706
Log-likelihood	-540.1128	Hannan-Quinn criter.		4.715620
F-statistic	454.5945	Durbin-Watson stat		0.698392
Prob(F-statistic)	0.000000	Wald F-statistic		1247.104
Prob(Wald F-statistic)	0.000000			

Source: E-view 12 output

Table 3 revealed that only clap and height have p-values that are lower than the critical level of = 0.05, indicating that they are significantly related to pulse rate.

Table 4: White-Hinkley (HC1) analysis Estimating Pulse rate using standard error estimates White-Hinkley (HC1) heteroscedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AGE	-0.004655	0.010454	-0.445275	0.6565
CLAP	0.477972	0.242340	1.972325	**0.0498**
EXER	0.008630	0.135682	0.063601	0.9493
FOLD	0.308323	0.327418	0.941679	0.3474
HEIGHT	1.018928	0.164120	6.208420	**0.0000**
NW_HND	0.068170	0.117528	0.580033	0.5625
SEX	0.232397	0.573840	0.404987	0.6859
SMOKE	0.134911	0.131204	1.028253	0.3049
W_HND	-2.682422	1.493538	-1.796018	0.0738
WR_HND	-0.371974	0.873008	-0.426083	0.6705
C	-95.11930	46.94499	-2.026187	0.0439
R-squared	0.952640	Mean dependent var		73.94093
Adjusted R-squared	0.950544	S.D. dependent var		10.88211
S.E. of regression	2.420033	Akaike info criterion		4.650741
Sum squared resid	1323.582	Schwarz criterion		4.811706
Log-likelihood	-540.1128	Hannan-Quinn criter.		4.715620
F-statistic	454.5945	Durbin-Watson stat		0.698392
Prob(F-statistic)	0.000000	Wald F-statistic		1189.222
Prob(Wald F-statistic)	0.000000			

Source: E-view 12 output

Table 4 also revealed that only clap and height have p-values that are lower than the critical level of = 0.05, indicating that they are significantly related to pulse rate.

Table 5: MacKinnon-White (HC2) analysis Estimating Pulse rate using standard error estimates

MacKinnon-White (HC2) heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std.Error	t-Statistic	Prob.
AGE	-0.004655	0.010829	-0.429839	0.6677
CLAP	0.477972	0.246311	1.940523	0.0536
EXER	0.008630	0.137627	0.062702	0.9501
FOLD	0.308323	0.332879	0.926233	0.3553
HEIGHT	1.018928	0.167062	6.099083	**0.0000**
NW_HND	0.068 170	0.119465	0.570629	0.5688
SEX	0.232397	0.583458	0.398310	0.6908
SMOKE	0.1349 11	0.13326 1	1.012379	0.3124
W_HND	-2.682422	1.520296	-1.764408	0.0790
WR_HND	-0.371974	0.888774	-0.418525	0.6760
c	-95.11930	47.79 109	-1.990314	0.0478
R-squared	0.952640	Mean dependent var		73.94093
Adjusted R-squared	0.950544	S.D. dependent var		10.88211
S.E. of regression	2.420033	Akaike info criterion		4.650741
Sum squared resid	1323.582	Schwarz criterion		4.811706
Log-likelihood	-540.1128	Hannan-Quinn criter.		4.715620
F-statistic	454.5945	Durbin-Watson stat		0.698392
Prob(F-statistic)	0.000000	Wald F-statistic		1175.282
Prob(Wald F-statistic)	0.000000			

Source: E-view 12 output

Table 5 shows results obtained using HC2, it displayed the OLS regression estimates and the estimated standard errors and p values for each regression coefficient. In these cases, pulse rate is significantly related to only height with p-value less than the critical level of α = 0.05

Table 6: Davidson-MacKinnon (HC3) analysis Estimating Pulse rate using standard error estimates
Davidson-MacKinnon (HC3) heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std.Error	t-Statistic	Prob.
AGE	-0.004655	0.011566	-0.402455	0.6877
CLAP	0.477972	0.256394	1.864208	0.0636
EXER	0.008630	0.142999	0.060347	0.9519
FOLD	0.308323	0.346595	0.889577	0.3746
HEIGHT	1.018928	0.174157	5.8506 11	**0.0000**
NW_HND	0.068 170	0.124367	0.548137	0.5841
SEX	0.232397	0.607553	0.382514	0.7024
SMOKE	0.1349 11	0.138633	0.973150	0.3315
W_HND	-2.682422	1.584775	-1.692620	0.0919
WR_HND	-0.371974	0.926642	-0.401421	0.6885
c	-95.11930	49.82522	-1.909059	0.0575
R-squared	0.952640	Mean dependent var		73.94093
Adjusted R-squared	0.950544	S.D. dependent var		10.88211
S.E. of regression	2.420033	Akaike info criterion		4.650741
Sum squared resid	1323.582	Schwarz criterion		4.811706
Log-likelihood	-540.1128	Hannan-Quinn criter.		4.715620
F-statistic	454.5945	Durbin-Watson stat		0.698392
Prob(F-statistic)	0.000000	Wald F-statistic		1106.794
Prob(Wald F-statistic)	0.000000			

Source: E-view 12 output

Also table 6 results obtained using HC3 equally displayed the OLS regression estimates and the estimated standard errors and p values for each regression coefficient. This case also show that pulse rate is significantly related to only height with p-value less than the critical level of α = 0.05

Table 7: Cribari-Neto (HC4) analysis Estimating Pulse rate using standard error estimates Cribari-Neto (HC4) heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std.Error	t-Statistic	Prob.
AGE	-0.004655	0.012815	-0.363235	0.7168
CLAP	0.477972	0.253447	1.885885	0.0606
EXER	0.008630	0.141576	0.060953	0.9515
FOLD	0.308323	0.342538	0.900 114	0.3690
HEIGHT	1.018928	0.172353	5.9 11855	**0.0000**
NW_HND	0.068170	0.122955	0.554431	0.5798
SEX	0.232397	0.600566	0.386964	0.699 1
SMOKE	0.1349 11	0.137218	0.983187	0.3266
W_HND	-2.682422	1.566851	-1.711983	0.0883
WR_HND	-0.371974	0.917247	-0.405533	0.6855
C	-95.11930	49.31014	-1.92900 1	0.0550
R-squared	0.952640	Mean dependent var		73.94093
Adjusted R-squared	0.950544	S.D.dependent var		10.88211
S.E. of regression	2.420033	Akaike info criterion		4.650741
Sum squared resid	1323.582	Schwarz criterion		4.811706
Log-likelihood	-540.1128	Hannan-Quinn criter.		4.715620
F-statistic	454.5945	Durbin-Watson stat		0.698392
Prob(F-statistic)	0.000000	Wald F-statistic		1140.707
Prob(Wald F-statistic)	0.000000			

Source: E-view 12 output

Likewise table 7 presents results obtained using HC4 andwhich also displayed the OLS regression estimates and the estimated standard errors and p values for each regression coefficient. In these cases, pulse rate is significantly related to only height with p-value less than the critical level of α = 0.05

Table 8: Comparison of Heteroskedasticity-consistent standard errors estimates

	F_{HC0}	F_{HC1}	F_{HC2}	F_{HC3}	F_{HC4}
Wald F-Statistic	1247.104	1189.222	1175.282	1106.94	1140.707
P	0.000000	0.000000	0.000000	0.000000	0.000000

Table 4.8 revealed results of comparison between the different types of Heteroscedasticity-Consistent Covariance Matrix (HCCM) estimators. The Wald F-statistic was used to determine which estimator performed better. Results disclosed that HC3 and HC4 performed better as they had lower test statistic as compared to HC0, HC1 and HC2. This discovery is consistent with the studies of (Simsek and Orhan, 2016) who all revealed that the use of HC3 and HC4 yielded better results compared to HC0, HC1 and HC2.

DISCUSSION AND RESULTS

The study started by carrying out diagnostic test (Breush-Pagan Godfrey test) to check for heteroscedasticity. The result revealed presence of heteroscedasticity, this led to the application of Heteroscedasticity-Consistent Covariance Matrix (HCCM) Estimators. When the Ordinary Least square was applied as shown in table 1, pulse rate was seen to be significantly related to clap, height and w_hnd. This agrees with (Berry, 1993) study which revealed that the effects of assumption violations in such a scenario can move standard errors and p-values in unpredictable ways. By applying the White heteroskedasticity-consistent standard errors & covariance (HC0) and White-Hinkleyheteroscedasticity-consistent standard errors & covariance (HC1) estimators' results revealed that only clap and height significantly related to pulse rate. Application of MacKinnon-White heteroskedasticity-consistent standard errors & covariance (HC2), Davidson-MacKinnon heteroskedasticity-consistent standard errors & covariance (HC3) and Cribari-Neto heteroskedasticity-consistent standard errors & covariance (HC4) showed that only height was significantly related to pulse rate. The results also showed that comparison among the HCCM estimators, the HC3 and HC4 produced better results.

CONCLUSION

In conclusion this study reveals that the researcher's confidence in the validity and power of inferential tests in OLS regression can be increased by utilising an HCSE estimator rather than assuming homoskedasticity. It is also concluded that HC3 and HC4 performed better as they had lower test statistic as compared to HC0, HC1 and HC2.

REFERENCES

Agunbiade, D. A., and Adeboye, N. O., 2012. Estimation of Heteroscedasticity Effects in a Classical Linear Regression Model of a Cross-Sectional Data. Progress in Applied Math-ematics, 4 (2), 18-28.

Astivia, Oscar L. Olvera and Zumbo, Bruno D., 2019. Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it with Applications in R and SPSS. Practical Assessment, Research & Evaluation, 24(1).

Berry, W. D., 1993. Understanding regression assumptions. Newbury Park, C A: Sage.

Carroll, R. J. and D. Ruppert, 1988. Transformation and Weighting in Regression. New York: Chapman and Hall.

Cribari-Neto, F., 2004. Asymptotic inference under heteroskedasticity of unknown form. Computational Statistics and Data Analysis, 45, 215-233.

Cribari-Neto, F. and Zarkos, S. G., 2007. Heteroskedasticity-consistent covariance matrix estimation: White's estimator and the bootstrap. J. Statist. Computat. Simul. 68:391– 411.

Duan, N., 1983. Smearing Estimate: A Nonparametric Retransformation Method. Journal of the American Statistical Association 78:605-610.

Greene, 2002. Econometric Analysis Fifth edition, Pearson Education Inc., New York, USA. 4.

Hinkley, D. V., 1977. Jackknifing in unbalanced situations. Technometrics, 19, 285- 292.

Huber, P. J., 1967. The behavior of maximum likelihood estimation under nonstandard conditions. In L. M. Le Cam and J. Neyman (Eds.), Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Berkeley, CA: University of California Press

Kauermann, G., and Carroll, R. J., 2001. A note on the efficiency of sandwich covariance matrix estimation. Journal of the American Statistical Association, 96, 1387-1396.

Long and Ervin, 2000. Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. The American Statistician. 55(4):374-375.

MacKinnon, J. G., and White, H., 1985. Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties. Journal of Econometrics, 29, 305-325.

Munir A, Muhammad A and G. R. Pasha, 2011. Inference under Heteroscedasticity of Unknown Form Using an Adaptive Estimator, Communications in Statistics - Theory and Methods, 40:24, 4431-4457

Şimşek, E. and Orhan, M., 2016. Heteroskedasticity-Consistent Covariance Matrix Estimators in Small Samples with High Leverage Points. Theoretical Economics Letters, **6**, 658-677.

Venables B. and Ripley B., 2002. Modern Applied Statisitcs with S. Springer. 10.1007/b97626

Vynck and Thas, 2017. Heteroscedasticity in linear models: an empirical comparison of estimation methods. A dissertation submitted to obtain the degree of Master of Science in Statistical Data Analysis. Department of Mathematical Modelling, Statistics and Bioinformatics. Ghent University.

Weisberg, S., 1980. Applied linear regression. New York: Wiley

White, H., 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. Econometrica, 48, 817–838.

Wilcox, R. R., 2001. Comment. The American Statistician, **55**, 374 - 375.