# A New Heteroskedastic Consistent Covariance Matrix Estimator using Deviance Measure

Nuzhat Aftab College of Statistical and Actuarial Sciences University of the Punjab, Quaid-i-Azam Campus Lahore, Pakistan m.nuzhat.aftab@gmail.com

Sohail Chand College of Statistical and Actuarial Sciences University of the Punjab, Quaid-i-Azam Campus Lahore, Pakistan sohail.stat@gmail.com

#### Abstract

In this article, we propose a new heteroskedastic consistent covariance matrix estimator,  $HC_6$ , which is based on deviance measure. We have studied the finite sample behavior of the test statistic based on this new HC estimator. We compare its performance with other HC estimators namely  $HC_1$ ,  $HC_3$  and  $HC_{4m}$ , which are also used in case of leverage observations. Extensive simulation studies are used to study the effect of various levels of heteroskedasticity on the performance of the quasi tests based on HC estimators. Results showed that the test statistic based on new suggested estimator has better asymptotic approximation and have less size distortion in small samples especially when high level heteroskedasticity is present in the data.

**Keywords:** Regression, Heteroskedasticity, Deviance, Influential points, Covariance matrix estimation.

# 1. Introduction

In regression analysis, the presence of heteroskedasticity in the data leads to inefficient estimates of ordinary least square (OLS) estimates. In this situation the covariance matrix estimate of OLS estimates become biased and does not remain consistent. Thus the inconsistency of the covariance matrix fails to provide the asymptotically valid inference. The problem becomes more severe with the increased level of heteroskedasticity. In regression analysis, it is very common among practitioners to use the point estimates computed from OLS method even if they suspect the presence of heteroskedasticity in the data. However in order to perform inference about the parameters of the regression model, it is important to use a heteroskedasticity consistent estimate of covariance matrix.

Several authors have suggested covariance matrix estimators which are consistent in case of both homoskedastic and heteroskedastic error variances. The most commonly used heteroskedastic consistent estimator was suggested by White (1980), named as  $HC_0$ . This estimator is widely used in literature but various studies showed that  $HC_0$  can be severely biased for small samples. It tends to underestimate the true variance which in turn results in poor performance of associated quasi t-statistic see, e.g MacKinnon and White (1985), Cribari-Neto and Zarkos (1999), Cribari-Neto and Zarkos (2001). MacKinnon and White

(1985), Suggested alternative HCCMEs called  $HC_1$ ,  $HC_2$ . Later Davidson and MacKinnon (1993) suggested another alternative estimator, called as  $HC_3$ , which is an approximation of the jackknife estimator. The simulation results in Long and Ervin (2000) showed that  $HC_3$  performed the best among the other available such estimators. Cribari-Neto and Zarkos (2001) showed that the presence of high leverage observation is more critical for HCCMEs. So Cribari-Neto (2007) proposed a new version of the HCCME called as  $HC_4$ . Their numerical results showed that the inference about regression parameter using  $HC_4$  is more reliable but it showed large amount of bias. Later Cribari-Neto et al., (2007), Cribari-Neto and da Silva (2011) suggested two new versions of HCCMEs denoted as  $HC_5$  and  $HC_{4m}$ , which have lesser bias relative to  $HC_4$ . In this article we propose a new estimator, called as  $HC_6$ . It performs well in case of small sample especially when heteroskedasticity level is high. The simulations results show that quasi t-test for the inference about regression parameters based on new estimator has better approximation of asymptotic distribution when heteroskedasticity is high and sample size is small.

The rest of the paper is organized as follows: we introduce the model and covariance matrix estimators in Section 2. In section 3, we propose a new HCCME,  $HC_6$ , based on measure of deviance. The results and discussion are reported in Section 4. In Section 5, we study the application to real life data. The concluding remarks are given in Section 6.

## 2. The Model and Estimators

The regression model considered is,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where, **X** is the  $n \times k$  matrix of independent variables, **Y** is  $n \times 1$  vector of dependent variable and  $\mathcal{E}$  is the  $n \times 1$  column vector of error term and  $\beta = (\beta_0, \beta_1, ..., \beta_{k-1})$  is the vector of parameters need to be estimated. We assume that  $\mathcal{E}_i \sim N(0, \sigma_i^2)$ , with  $(0 < \sigma_i^2 < \infty)$ , i = 1, 2, ..., n and  $E(\mathcal{E}_i, \mathcal{E}_j) = 0$ , for all  $i \neq j$ . Thus  $var(\mathcal{E}) = \Omega$ , where,

$$\mathbf{\Omega} = diag\left(\sigma_1^2, \sigma_2^2, ..., \sigma_n^2\right) \tag{2}$$

OLS estimate of  $\beta$  is defined as,

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y},\tag{3}$$

Such that  $E(\hat{\beta}) = \beta$  and var  $(\hat{\beta}) = \psi = \mathbf{P} \Omega \mathbf{P}^T$ , where

$$\mathbf{P} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \tag{4}$$

and  $\Omega$  as defined in (2). When the assumption of constant error variances is satisfied, the variance is defined as  $\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$  and it is estimated as  $\hat{\sigma}^2(\mathbf{X}^T\mathbf{X})^{-1}$ , where  $\hat{\sigma}^2 = \mathbf{\epsilon}^T \mathbf{\epsilon}/(n-k)$  and  $\hat{\mathbf{\epsilon}} = (\mathbf{I}_k - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X})\mathbf{Y} = (I_k - H)\mathbf{Y}$  is the vector of OLS residuals.

The commonly used HCCME called  $HC_0$  was given by Eicker (1963) and (White, 1980), is given as,

$$HC_0 = \mathbf{P}\widehat{\mathbf{\Omega}}\mathbf{P}^{\mathrm{T}} \tag{5}$$

Where,  $\hat{\Omega} = diag(\hat{\varepsilon}_1, \hat{\varepsilon}_2, ..., \hat{\varepsilon}_n)$ , White (1980) proposed this estimator to resolve the problem of estimation and inference in the presence of heteroskedasticity. This estimator proved to be consistent, in various studies, when nothing is known about the form of heteroskedasticity see e.g Arce and Mora (2002).  $HC_0$  as discussed in Section 1 can be seriously biased for small samples. There are some alternatives to the estimator of (White, 1980), available the in literature. These estimators are proposed to control the tendency of underestimation of the variance of the OLS estimates. These alternative estimators are found to be consistent under heteroskedasticity and incorporates small sample adjustment factors see e.g Cribari-Neto and Zarkos (1999), Cribari-Neto and Zarkos (2004), Davidson and MacKinnon (1993), but none of these work well in the scenario discussed in this paper i.e. small sample size with high level of heteroskedasticity.

According to the MacKinnon and White (1985),  $HC_0$  does not take into account the well known fact that the OLS residuals tend to be very small. They used a modified estimator of  $HC_0$  which they obtained by using the degree of freedom correction similar to one conventionally used to obtain unbiased estimate of variance denoted by  $\sigma^2$ . This yields the modified estimator  $HC_1$  suggested by Hinkley (1977) defined as,

$$HC_1 = \mathbf{P}\mathbf{E}_1 \widehat{\mathbf{\Omega}} \mathbf{P}^{\mathrm{T}} \tag{6}$$

Where,  $\mathbf{E}_1 = \frac{n}{n-k}\mathbf{I}$  is called the finite correction factor, where k denotes the number of parameters and  $\mathbf{I}_n$  is  $n \times n$  identity matrix. But according to them degree of freedom adjustment in  $HC_1$  is not the only way to compensate for the fact that the OLS residuals tend to underestimate the true errors. So following the Horn et al., (1975) they proposed another estimator called  $HC_2$  defined as,

$$HC_2 = \mathbf{PE}_2 \widehat{\mathbf{\Omega}} \mathbf{P}^{\mathrm{T}} \tag{7}$$

Where,  $\mathbf{E}_2 = diag \ (1/(1-h_{ii}))$  and  $h_{ii}, i=1,2,...,n$ , where,  $0 < h_{ii} < 1$  denote the  $i^{th}$  diagonal value of the hat matrix  $\mathbf{H}$ . These  $h_{ii}$  values in  $\mathbf{H}$  are called the leverage of the  $i^{th}$   $\mathbf{X}$  observation and indicate whether or not a value in  $\mathbf{X}$  is outlying. The  $h_{ii}$  measures the distance between  $i^{th}$  values of  $\mathbf{X}$  from the mean of all n values. So when  $h_{ii}$  takes a large value closer to 1, it indicates that the  $i^{th}$  value is distant from mean and has large leverage. In general a diagonal value of the hat matrix  $\mathbf{H}$  greater then 2k/n from the other data points in the diagonal is considered as leverage observation see e.g (Montgomery et al., 2001, p.207).

The  $HC_3$  given by Davidson and MacKinnon (1993) can be written as,

$$HC_3 = \mathbf{PE}_3 \widehat{\mathbf{\Omega}} \mathbf{P}^{\mathrm{T}} \tag{8}$$

where,  $E_3 = diag(1/(1-h_{ii})^2)$ ; i = 1,2,...,n. The estimators,  $HC_2$  and  $HC_3$ , include the finite sample correction factors that are based upon the leverages of different observations,

greater the leverage, more inflated will be the corresponding squared residuals see e.g Cribari-Neto and da Silva (2011). The resulting quasi-t tests tend to be quite liberal when the design matrix includes high leverage observations, thus leading to imprecise inference. So Cribari-Neto and Zarkos (2004), proposed a new estimator denoted by  $HC_4$  that takes into account the impact of high leverage points on the finite-sample behavior of the covariance matrix estimator. The  $HC_4$  estimator is given as,

$$HC_{A} = \mathbf{P}\mathbf{E}_{A}\widehat{\mathbf{\Omega}}\mathbf{P}^{\mathrm{T}} \tag{9}$$

where,  $E_4 = diag (1/(1-h_{ii})^{\delta_i})$  and  $\delta_i = \min (4,(nh_{ii})/k)$ ; i = 1,2,...,n. The exponent  $\delta_i$  controls the level of discounting for  $i^{th}$  observation and is given by the ratio between  $h_{ii}$  and  $\overline{h}$ 

where 
$$\overline{h} = \frac{1}{n} \sum_{i=1}^{n} h_i$$
. Since  $0 < h_{ii} < 1$  and  $\delta_i > 1$  it follows that  $0 < (1 - h_{ii})^{\delta_i} < 1$ . Hence, the  $i^{th}$ 

squared residual will be more strongly inflated when hii is large relative to  $\bar{h}$ .  $HC_4$  aims at discounting for leverage points more heavily than  $HC_2$  and  $HC_3$ . Cribari-Neto and da Silva (2011) showed that the asymptotic approximation of the  $HC_4$  is very poor so that they suggested a modified version of  $HC_4$ , denoted by  $HC_{4m}$ , given as,

$$HC_{4m} = PE_{4m}\widehat{\Omega}P^{T}$$
(10)

Where,  $\mathbf{E}_{4m} = diag \ (1/(1-h_{ii})^{\delta_i})$  and  $\delta_i = \min\{\gamma_1, (nh_{ii})/k\} + \min\{\gamma_2, (nh_{ii})/k\}$  for i=1,2,...,n. The values for  $\gamma_1$  and  $\gamma_2$  are selected in such a way that it will be helpful in reducing the effect of leverage observation. The values suggested by Cribari-Neto and da Silva (2011) are  $\gamma_1 = 1.0$  and  $\gamma_2 = 1.5$  and same values are used in our simulations.

#### 3. New Estimator

As we have discussed in Section 1 and 2 that all the alternative estimators of  $HC_0$  take into account only the effect of leverage observation or the extreme values in the design matrix X. In practice, whenever we have leverage points there also exist some influential observations in the Y variable which affect the results of the variance covariance matrix of the OLS estimates. All the modified versions of  $HC_0$  use leverage measure to rescale the OLS residuals involve in the estimation sandwich estimator to control the underestimation of the covariance matrix. But the leverage measure does not consider the effect of influential observations. Now in order to consider the effect of both leverage and influential observations in the estimation of variance covariance matrix we propose a new estimator denoted by  $HC_6$  given as,

$$HC_6 = \mathbf{PE}_6 \widehat{\mathbf{\Omega}} \mathbf{P}^{\mathrm{T}} \tag{11}$$

Where, 
$$\mathbf{E}_6 = diag \ (\sqrt{d_{ii}})$$
 and  $d_{ii} = \frac{r_i^2 h_{ii}}{k(1 - h_{ii})}$  and  $r_1^2 = \frac{e_i^2}{s^2 (1 - h_{ii})}$  for  $i = 1, 2, ..., n$ . where  $k$ 

is the number of parameters,  $r_i$  is the studentized residual,  $h_{ii}$  is the leverage measure for the  $i^{th}$  observation and  $e_i$  are the OLS residuals. It can be noticed that apart from the factor k,  $d_{ii}$  is the product of  $i^{th}$  squared studentized residual and the factor  $h_{ii}/(1-h_{ii})$ . Thus  $d_{ii}$  is made up of a component that reflects how well the model is fitted to the  $i^{th}$  observation  $y_i$  and a component that measures the distance of the  $i^{th}$  observation from the rest of the data, see e.g. (Montgomery et al., 2001). The diagonal elements of  $\mathbf{E}_6$ ,  $d_{ii}$ , behave like  $h_{ii}$  but unlike  $h_{ii}$ , it can take value greater than one, when both the leverage and the

influential observations are present at same point. That is why,  $d_{ii}$  is used to detect the influential observations in the data. The most interesting feature of  $\mathbf{E}_6$  is that it takes into account the effect of outliers or extreme observations both in x-space and y-space, so when it is used to rescale the residuals it may improve the results of the covariance matrix estimator even when x-space have no extreme values but only y-space have some or more extreme values. The value of  $d_{ii}$  will be larger if there are large values of X or Y or both in the data, so the  $i^{th}$  squared residuals will be more appropriately weighted when there are outliers of any type in the data.

## 4. Results

In order to evaluate the performance of the new suggested estimator, we compute its relative probability discrepancy (RPD), see e.g, Chand and Aftab (2012), Davidson and MacKinnon (1998). Simulation study is performed following design given in Cribari-Neto and da Silva (2011). The numerical results which are stated in this section are obtained using the heteroskedastic regression model given as:

$$Y_{i} = \beta_{1} + \beta_{2} X_{2i} + \beta_{3} X_{3i} + ... + \beta_{k} X_{ki} + \varepsilon_{i} , i = 1, 2, ..., n$$
(12)

Here  $\varepsilon_i \sim N(0, \sigma_i^2)$  and  $E(\varepsilon_i \varepsilon_j) = 0$  for  $i \neq j$ . The error variance  $\sigma_i^2$  is defined as  $\sigma_i^2 = \exp \sum_{j=1}^{k-1} \alpha_j X_{ij}$  for i = 1, 2, ..., n and j = 1, 2, ..., k. Where k is defined as the number of

parameters in the model and  $a_j$  being the real scalar used to control the level of heteroskedasticity. In the simulation study, we used the model defined in (12) with k=3i-e with three regression parameters. Different choices of the sample size are being studied i.e n = 50, 100 and 150 for different values of  $\alpha$  given as,  $\alpha = 0$ ,  $\alpha = 0.26$ ,  $\alpha = 1.5$ ,  $\alpha = 2$  and  $\alpha = 2.5$ . The sample sizes are selected with the intuition that heteroskedasticity is more problematic in small samples. In case of large samples all the sandwich estimators perform equally well, see e.g Cribari- Neto et al., (2007), Cribari-Neto (2004), Cribari-Neto and da Silva (2011). When the sample size is sufficiently large one can easily use the usual  $HC_0$  estimator. The values of  $\alpha$  are selected in order to consider the homoskedasticity ( $\alpha = 0$ ) and various levels of heteroskedasticity, low level heteroskedasticity ( $\alpha = 0.26$ ), high level heteroskedasticity, ( $\alpha = 1.5$ ) and severe level of heteroskedasticity, ( $\alpha = 2.5$ ). The values of the covariates are obtained as random draws from the standard lognormal distribution to make the data heteroskedastic. The level of heteroskedasticity can be measured using  $\lambda = max(\sigma_i^2)/min(\sigma_i^2)$ . When  $\lambda = 1$  it denoted that there is no heteroskedasticity and when  $\lambda > 1$  it implies that the heteroskedasticity is present. The larger values of  $\lambda$  indicate the higher level of heteroskedasticity. When  $\alpha = 0.26$  then  $\lambda$  is approximately equals to 60, for  $\alpha = 1.5$ ,  $\lambda$  is approximately equals to 90, when  $\alpha = 2$ ,  $\lambda$  is approximately equals to 150 and finally when  $\alpha = 2.5$  the value of the  $\lambda$ is greater than or equals to 190.

In this study, we want to test the hypothesis  $H_0$ :  $\beta_2 = 0$  against the two sided alternative hypothesis  $H_1$ :  $\beta_2 \neq 0$ . The test statistic used is

$$\tau^2 = \frac{\hat{\beta}_2^2}{\operatorname{var}(\hat{\beta}_2)} \tag{13}$$

where  $\hat{\beta}_2$  denote the OLS estimate of  $\beta_2$  and  $var(\hat{\beta}_2)$  is variance estimate of  $\hat{\beta}_2$  and it is based on  $HC_1$ ,  $HC_3$ ,  $HC_{4m}$  and  $HC_6$  estimators. The number of Monte Carlo runs are set to 10, 000. All the simulation results are performed using the R programming language, see (R Development Core Team, 2011). In this study, we consider only heteroskedastic errors with high leverage and influential observations.

Table 1 presents the empirical probabilities of quasi t-test based on the considered HCCMEs. We study the effect of level of heteroskedasticity on the approximation of asymptotic distribution of quasi t-test for different choices of sample size, ranging from small to large sample size.

Table 1: Empirical probabilities, when the test statistic is computed using variances of  $HC_1$ ,  $HC_3$ ,  $HC_{4m}$  and  $HC_6$  estimators

$\overline{n}$	n=50			n=100			n=150			
-	0.90	$\frac{n=30}{0.95}$	0.99	0.90	$\frac{n=100}{0.95}$	0.99	0.90	$\frac{n-150}{0.95}$	0.99	
γ	0.90	0.93	0.99			0.99	0.90	0.93	0.99	
α=0										
$HC_1$	0.805	0.869	0.941	0.832	0.891	0.955	0.8416	0.9007	0.9605	
$HC_3$	0.886	0.929	0.971	0.884	0.93	0.975	0.8906	0.9365	0.9777	
$HC_{4m}$	0.902	0.938	0.977	0.896	0.938	0.978	0.9011	0.9437	0.9804	
$HC_6$	0.551	0.619	0.727	0.534	0.61	0.717	0.5191	0.5914	0.6995	
α=0.25										
$HC_1$	0.599	0.683	0.808	0.642	0.745	0.874	0.685	0.797	0.908	
$HC_3$	0.881	0.925	0.968	0.893	0.938	0.978	0.916	0.955	0.984	
$HC_{4m}$	0.909	0.942	0.976	0.914	0.951	0.984	0.932	0.964	0.988	
$HC_6$	0.681	0.758	0.853	0.802	0.861	0.929	0.862	0.909	0.954	
α=1.5										
$HC_1$	0.628	0.747	0.861	0.724	0.838	0.924	0.777	0.881	0.949	
$HC_3$	0.97	0.986	0.996	0.983	0.994	0.999	0.985	0.996	0.999	
$HC_{4m}$	0.98	0.991	0.997	0.988	0.996	0.999	0.989	0.997	0.999	
$HC_6$	0.875	0.92	0.963	0.955	0.975	0.991	0.976	0.988	0.995	
α=2										
$HC_1$	0.647	0.759	0.866	0.747	0.846	0.952	0.8	0.891	0.949	
$HC_3$	0.981	0.991	0.997	0.988	0.995	0.999	0.99	0.997	0.999	
$HC_{4m}$	0.987	0.994	0.998	0.992	0.997	0.999	0.993	0.998	1.000	
$HC_6$	0.895	0.939	0.968	0.961	0.979	0.991	0.982	0.992	0.996	
α=2.5										
$HC_1$	0.649	0.761	0.877	0.75	0.85	0.924	0.807	0.891	0.955	
$HC_3$	0.983	0.994	0.998	0.988	0.996	0.999	0.992	0.998	0.999	
$HC_{4m}$	0.989	0.996	0.999	0.993	0.998	0.999	0.995	0.998	0.999	
$HC_6$	0.896	0.939	0.975	0.967	0.982	0.993	0.986	0.993	0.998	

We have considered asymptotic probabilities  $\gamma = 0.90$ , 0.95 and 0.99 corresponding to 10%, 5% and 1% levels of significance which are common choices in statistical inference. It can be observed from Table 1 that when there is no heteroskedasticity the approximation of  $HC_1$ ,  $HC_3$  and  $HC_{4m}$  is better and  $HC_6$  has relatively poor approximation. The reason behind the poor approximation of the test using  $HC_6$  is due to

the use of deviance measure as correction matrix which is specifically suggested to deal with the influential observation. So it is not recommended to use it in case of homoskedasticity. Next when the heteroskedasticity is present and is mild,  $\alpha = 0.26$ , the quasi t-test based on  $HC_{4m}$  has good approximation of asymptotic distribution. In this scenario, the size of quasi t-test based on new suggested estimator is larger than that of the asymptotic distribution. This distortion of size is larger when the sample size is small. The test based on  $HC_1$  is showing the largest discrepancy in size.

Moreover, when the level of heteroskedasticity increases and the sample size is small, the test based on  $HC_6$  outperforms the other estimators. While for large sample size and high level of heteroskedasticity all the estimators, except  $HC_1$ , have shown similar behavior. The empirical distribution of these tests has shown heavier right tail as compared to that of the asymptotic distribution. These results show that for small sample size the quasi-t test based on  $HC_1$ ,  $HC_3$  and  $HC_{4m}$  under the high level of heteroskedasticity could be substantially unreliable and misleading. The quasi t-test for the inference about regression coefficient based on the  $HC_6$  estimator has better asymptotic approximation for high level heteroskedasticity with small sample size.

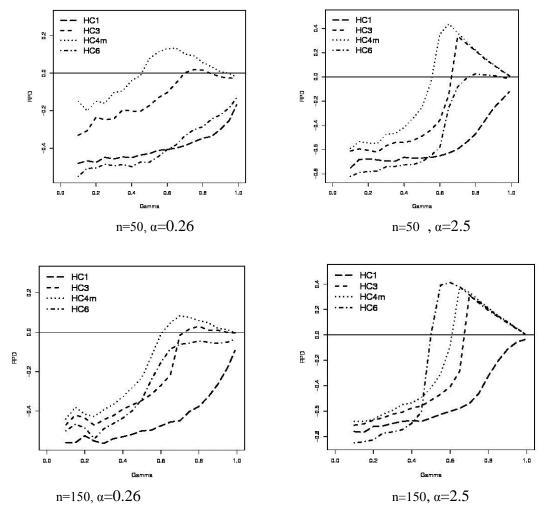


Figure 1: Relative probability discrepancy (RPD) versus asymptotic probabilities( $\gamma$ ); With k=3

Figure 1 showed the plots of relative probability discrepancy (RPD) against the asymptotic probability of the asymptotic distribution of quasi t-test based on  $HC_1$ ,  $HC_3$ ,  $HC_{4m}$  and  $HC_6$ . Usually, we are interested in studying the relative probability discrepancy when  $\gamma > 0.8$ . So we will discuss the results specifically for this situation. It can be observed that for sample size when the level of heteroskedasticity is low,  $HC_3$  and  $HC_{4m}$  have smaller relative probability discrepancy. The situation totally changes when the level of heteroskedasticity is high. The relative probability discrepancy of test based on  $HC_6$  decreases in presence of high level heteroskedasticity. While this is the case which adversely affects the asymptotic approximation of  $HC_3$  and  $HC_{4m}$ . For large sample size, all the three tests,  $HC_3$ ,  $HC_{4m}$  and  $HC_6$ , have shown same amount of relative probability discrepancy especially when  $\gamma > 0.7$ . The behavior of  $HC_1$  is generally poor in all the considered scenarios.

# 5. Application to Real data

In this section, we apply the quasi t-test for the hypothesis testing of the significance of regression coefficient. The test is applied under the new suggested estimator  $HC_6$  and comparison has been made with  $HC_1$ ,  $HC_3$ ,  $HC_{4m}$ . The data has been taken from Greene (1997, p.541). This same data set has already been studied in literature see e.g. Cribari-Neto (2004), Cribari-Neto et al., (2007) and Cribari-Neto and Zarkos (2004). In this data, the per capita spending on public schools (y) is the response variable while the per capita income of the state (x) is independent variable. The data set consist of 50 observations excluding missing observations.

The regression model which is estimated for this data is,

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{1i}^2 + \varepsilon_i$$
,  $i = 1, 2, ..., 50$  (14)

This is the same model studied by Cribari- Neto et al.,(2007). The OLS estimates of the model  $\hat{\beta}_o = 832.9$ ,  $\hat{\beta}_1 = -1834.2$ ,  $\hat{\beta}_2 = 1587.0$ . The Breusch-pagan test (LM = 18.9035, p.value < 0.001) shows the presence of heteroskedasticity. We wish to test the null hypothesis  $H_0$ :  $\beta_2 = 0$  against  $H_1$ :  $\beta \neq 0$ .

Table 2: Test of heteroskedasticity  $H_0$ :  $\beta_2 = 0$  Breusch-Pagan test of heteroskedasticity

Test	LM	d.f	p-value					
Breusch Pagan	18.903	2	0:000					
Testing of $H_0$ : $\beta_2 = 0$ against $H_1$ : $\beta_2 \neq 0$								
Test	S.E	T	p-value					
OLS	519.1	3.06	0.0036					
$HC_1$	856.1	1.85	0.0700					
$HC_3$	1995.2	0.8	0.4303					
$\mathrm{HC}_{\mathrm{4m}}$	2553.3	0.62	0.5372					
HC <sub>6</sub>	1146.2	1.38	0.1727					

From the results given in Table 2 it can be noticed that the test on the OLS standard error rejects the null hypothesis even at 1% level of significance. Same is the case for  $HC_1$ 

which rejects the null hypothesis at 10% level of significance. While the test based on other HCCMEs, i.e.  $HC_3$ ,  $HC_{4m}$  and  $HC_6$ , we are unable to reject the null hypothesis even at 10% level of significance. Hence from these results it can be concluded that, the test statistic based on the  $HC_6$  estimator will give reliable inference in real life data.

# 6. Conclusion

In this article we propose a new HC estimator, called  $HC_6$ , which used the deviance measure to rescale the residuals. The numerical results suggest that, when the level of heteroskedasticity is very high and sample size is small, the test based on  $HC_6$  estimator have better approximation of asymptotic distribution. We recommend using newly suggested estimator instead to other HCCMEs, especially when the sample size is small. We study the estimator only under normal distribution with heteroskedastic disturbances it can also be studied assuming heteroskedasticity under some non normal distribution of the error term.

# 7. References

- 1. Arce, M. and Mora, A. (2002). Empirical evidence of the effect of European accounting differences on the stock market valuation of earnings and book value. *European Accounting Review*, 11(3): 573–599.
- 2. Chand, S. and Aftab, N. (2012). On finite sample distribution of quasi test statistic based on heteroskedastic consistent covariance matrix estimators. *Middle-East Journal of Scientific Research*, 12(8): 1157–1164.
- 3. Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics & Data Analysis*, 45(2): 215–233.
- 4. Cribari-Neto, F. and da Silva, W. (2011). A new heteroskedasticity consistent covariance matrix estimator for the linear regression model. *AStA Advances in Statistical Analysis*, 95(2): 129–146.
- 5. Cribari-Neto, F., Souza, T., and Vasconcellos, K. (2007). Inference under heteroskedasticity and leveraged data. *Communications in Statistics-Theory and Methods*, 36(10): 1877–1888.
- 6. Cribari-Neto, F. and Zarkos, S. (1999). Bootstrap methods for heteroskedastic regression models: evidence on estimation and testing. *Econometric Reviews*, 18(2): 211–228.
- 7. Cribari-Neto, F. and Zarkos, S. (2001). Heteroskedasticity consistent covariance matrix estimation: White's estimator and the bootstrap. *Journal of Statistical Computation and Simulation*, 68(4): 391–411.
- 8. Cribari-Neto, F. and Zarkos, S. G. (2004). Leverage-adjusted heteroskedastic bootstrap methods. *Journal of Statistical Computation and Simulation*, 74(3): 215–232.
- 9. Davidson, R. and MacKinnon, J. (1993). *Estimation and inference in econometrics*. Oxford University Press, USA.

- 10. Davidson, R. and MacKinnon, J. (1998). Graphical methods for investigating the size and power of hypothesis tests. *The Manchester School*, 66(1):1–26.
- 11. Eicker, F. (1963). Asymptotic normality and consistency of the least squares estimators for families of linear regressions. *The Annals of Mathematical Statistics*, 34(2): 447–456.
- 12. Greene, W. H. (1997). *Econometric analysis*. Pearson Education India, 3rd edition.
- 13. Hinkley, D. (1977). Jackknifing in unbalanced situations. *Technometrics*, 19: 285–292.
- 14. Horn, S., Horn, R., and Duncan, D. (1975). Estimating heteroskedastic variances in linear models. *Journal of the American Statistical Association*, 70:380–385.
- 15. Long, J. and Ervin, L. (2000). Using heteroskedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3):217–224.
- 16. MacKinnon, J. and White, H. (1985). Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3): 305–325.
- 17. Montgomery, D., Peck, E., Vining, G., and Vining, J. (2001). *Introduction to linear regression analysis*, volume 3.
- 18. Wiley New York. R Development Core Team (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- 19. White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 48(4): 817–838.