Middle-East Journal of Scientific Research 12 (8): 1157-1164, 2012
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2012.12.8.1852

On Finite Sample Distribution of Quasi Test Statistic Based on Heteroskedastic Consistent Covariance Matrix Estimators

Sohail Chand and Nuzhat Aftab

College of Statistical and Actuarial Sciences,
University of the Punjab, Lahore, Pakistan

Abstract: In regression the assumption that the errors are independently and identically distributed(IID) is often violated in practical situations. In such situation, the least square estimates of the regression parameters are still remain unbiased and consistent but no more efficient. Various Heteroskedastic estimators are suggested to deal with this problem. In this paper, through simulations we look at the appropriateness of asymptotic distribution of the test statistic used for the testing the significance of regression coefficients. We consider the quasi test statistic based on various heteroskedastic consistent covariance estimators suggested in the literature.

Key words: Linear regression • Heteroskedasticity • Consistent covariance matrix estimators

INTRODUCTION

Regression analysis is commonly used to check and model the relationship between two or more than two variables. Among the other common assumptions about the error term in the regression model, one assumption is that the error variance should be constant for all the observations. But in many practical applications the error variances are not constant and this condition is known as Heteroskedastic errors. In case of Heteroskedastic errors, the ordinary least square (OLS) estimates of the parameters are still unbiased and consistent, but the variance covariance matrix estimate of the regression model is no more unbiased and reliable. Thus the results of the tests which use these variance estimates may be highly misleading. In this situation, Heteroskedastic consistent covariance matrix estimators (HCCMEs) are used. HCCMEs remain consistent and efficient whether the errors are constant or not. Hypothesis testing and other inferences are then made by using the OLS estimates coupled with the standard errors obtain from these HCCMEs.

There are several consistent covariance matrix estimates for the OLS estimates which remain consistent under the heteroskedasticity of unknown form. The well known and most commonly used HCCME denoted by HC0 was proposed by [1] following the results of [2] and [3]. The Heteroskedasticity consistent standard errors can be obtained by applying the square root on the diagonal values of these HCCMEs. With these estimators the researchers are now able to make any hypothesis testing inference or to compute confidence interval for the parameters of the regression model. Usually it is a general practice to base the inference on these heteroscedasticity consistent standard errors because these are robust of heteroscedasticity and provide accurate inference with minimal model assumption see e.g [4].

A very common and identified flaw of [1], is that it is biased when the sample size is small and leverage points are present in the data, see e.g., [5]. A few alternatives of [1] estimator,HC0, are proposed and studied in the literature. Some of the variants are, the HC1estimator by [6], HC2estimator by [7], HC3 estimator by [8], the HC4estimator by [9], HC4mrecentlyproposed by [10] and HC5proposed by [9]. There are several simulation studies in which the small sample performance of these HCCMEs is investigated and studied that how accurately these estimators estimate the OLS covariance matrix. Some of these simulation studies include, [8, 10, 11-13]. Recently, [14] has suggested a monte carlo algorithm to estimate the covariance matrix of regression coefficients.

[15] studied the performance of the HCCMEs in terms of quasi test statistics and concluded that the HC3 estimator is best and it is also an approximation to the

Corresponding Author: Sohail Chand, College of Statistical and Actuarial Sciences, University of the Punjab, Lahore-Pakistan. E-mail: sohail.stat@gmail.com.

jackknife estimator. Similar results are also reported in [11, 16, 17] studied the amount of bias in HCCMEs while estimating the true variance. [12] has used quasi test statistics based on different HCCMEs with leverage observation in the data to check the performance of these estimators. They suggested that HC4 is best estimator in presence of leverage observation [18] computed the confidence interval using the different HCCMEs and showed that the confidence interval estimation obtained using the HC4estimator is much reliable then any other technique.

More recently, [10] suggested a new estimator, HC4mand showed that this new estimator performed the best among all the HCCMEs.

In our study we study the finite sample performance of various HCCMEs and compare asymptotic distributions of these tests by Monte Carlo simulation in the case of normal and heteroscedastic error. The novelty of our study is that we suggest a Monte Carlo method instead of using numerical integration. Our results suggest that performance of these estimators is affected by the distribution of error term but the overall performance HC4m estimator is better. Moreover, the novelty of our work is that we have used the Monte Carol method to study the appropriateness of the asymptotic distribution of the quasi test statistic defined for various HCCMEs. To the best of our knowledge, the newly proposed estimator HC4mhas not been studied and compared under the settings as in this study.

The rest of the paper in organized as follows: we introduce the model and covariance matrix estimators in Section 2. Section 3 provides the simulation design and discussion of results. Finally, the conclusion is given in Section 5.

MATERIALS AND METHODS

The regression model considered is,

$$Y = X\beta + \varepsilon$$

where, X is the $n \times k$ matrix of independent variables, Y is vector of dependent variable with order $n \times 1$ and ϵ is the vector of error term with order $n \times 1$. $\beta(\beta_0,...,\beta_{k-1})$ is the vector of unknown parameters. The ϵ in the model is distributed as $\epsilon \sim N(0, \sigma_i^2)$, $(0 < \sigma_i^2 < \infty)$, $i = 1,...,n$, where n denotes sample size. The error term is independently distributed implies $E(\epsilon_i, \epsilon_j) = 0$, for all $i \neq j$ and its covariance matrix will be a diagonal matrix denoted by Ω and given as $\Omega = diag(\sigma^2_1,...,\sigma^2_n)$.

The ordinary least square (OLS) estimate for the parameters vector β of regression model (\ref{regression model}) can be written as $\hat{\mathbf{B}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$ having its covariance matrix, $\psi = P\Omega P^T$, where,

$$P = (X^TX)^{-1} X^T$$

when the model error term is homoskedastic then it has its variance equals to, $\psi = \sigma^2(X^TX)^{-1}$ and it can be estimated as $\hat{\Psi} = \hat{\sigma}^2(\mathbf{X}^T\mathbf{X})^{-1}$, where $\hat{\sigma}^2 = \mathbf{e}^T\mathbf{e}/(n-k)$ and $\mathbf{e} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})$.

When the model is not homoskedastic it is common practice to use the OLS method to find the estimates of the parameter vector â and then combine it with some heteroscedastic consistent covariance matrix (HCCME) estimator to perform statistical inference. The commonly used HCCME called HC_0 was given by [2] and [1] is given as,

$$HC_O = \mathbf{P}\hat{\Omega}\mathbf{P}^T$$

where, $\hat{\Omega} = diag\{\hat{\epsilon}_1^2,...,\hat{\epsilon}_n^2\}$. This estimator is proved to be consistent in various studies, see e.g [19], when nothing is known about the form of heteroscedasticity. HC_0 can be seriously biased for the small samples. There are some alternatives to the [1] estimator in literature, these estimators are proposed in order to control for the tendency to underestimate the variance of the estimates. These alternative estimators are found to be consistent under heteroscedasticity and incorporates small sample adjustment factors see e.g [8, 11, 12]. In the following paragraphs we shall now discuss some of the variants of the HC_0 estimator.

The HC_1 estimator proposed by [13]is written as,

$$HC_1 = \mathbf{P}\mathbf{E}_1\hat{\Omega}\mathbf{P}^T$$

where $\mathbf{E}_1 = \frac{n}{n-k}I$ is called the finite sample correction factor, where k denotes the number of parameters and I is $n \times n$ identity matrix.

[15] proposed HC_2 estimator given as,

$$HC_2 = \mathbf{P}\mathbf{E}_2\hat{\Omega}\mathbf{P}^T$$

where $E_2 = diag\{1/(1-h_{ii})\}$ and h_{ii}, $i = 1,...,n$ denote the i^{th} diagonal value of the hat matrix, $H = X(X^TX)^{-1}X^T$. These h_{ii} values in H are called the leverage of the i^{th} X observation and indicates whether or not a value in X is outlying. The h_{ii} measures the distance between i^{th} value of X from the mean of all n values. So when h_{ii} approaches 1, it indicates that the i^{th} value is distant from mean and has large leverage. In general a value greater then $2p/n$, where p denotes the number of parameters, is considered as leverage observation.

The HC_3 given by [8] can be written as,

$$HC_3 = \mathbf{P}\mathbf{E}_3\hat{\Omega}\mathbf{P}^T$$

where $E_3 = diag\{1/(1-h_{ii})^2\}$; $i = 1,\ldots,n$. The estimators, CH_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 [10].

The HC_4 estimator proposed by [9] is,

$$HC_4 = \mathbf{P}\mathbf{E}_4\hat{\Omega}\mathbf{P}^T$$

where, $\mathbf{E}_4 = diag\{1/(1-h_{ii})^{\delta_i}\}$ and $\delta_i = min\{4, (nh_{ii})/k\}$; $i = 1,\ldots,n$.

HC_5 estimator given by [5], is given as,

$$HC_5 = \mathbf{P}\mathbf{E}_5\hat{\Omega}\mathbf{P}^T$$

where $\mathbf{E}_5 = diag\{1/\sqrt{(1-h_{ii})^{\delta_i}}\}$, here $\delta_i = min\{(nh_{ii})/k, max\{4,(nch_{max})/k\}\}$ where $h_{max} = max\{h_1,\ldots,h_n\}$ and c some fixed value in [0 1] interval, see e.g [10].

The HC_{4m} by [5] is,

$$HC_{4m} = \mathbf{P}\mathbf{E}_{4m}\hat{\Omega}\mathbf{P}^T$$

where, $\mathbf{E}_{4m} = diag\{1/(1-h_{ii})^{\delta_i}\}$ and $\delta_i = min\{\gamma_1,(nh_{ii})/k\} + min\{\gamma_2,(nh_{ii})/k\}$; $i = 1,\ldots,n$. The values for γ_1 and γ_2 are selected in such a way that they will reduce the effect of leverage observation. The values suggested by [10] are $\gamma_1 = 1.0$ and $\gamma_2 = 1.5$ and we will also use these values in our simulations.

RESULTS

In this section, we give simulation results regarding the performance of considered HCCMEs, see Section 2 for definitions. We use the following regression model,

$$Y_i = \beta_1 + \beta_2 X_{i1} + \ldots + \beta_k X_{ik-1} + \epsilon_i \; ; i = 1,\ldots,n,$$

where, X_{ij} is the $i^t h$ observation of the $j^t h$ predictor. The error terms in the model are independent of each other and have mean zero with i^{th} error variance $\sigma_i^2 = \exp(\sum_{j=1}^{k-1} \alpha_j X_{ij})$ where $i = 1,\ldots,n$, k is the number of parameters and α_j being a real scalar.

In our simulation study, we use the model given in (10) with three and five regression parameters i.e, k = 3 and k = 5, we mainly follow paper by [10]. We have considered the different sample sizes, n = 25, 50, 100, 500 in order to compare the behavior of HCCMEs for small and large samples.

The interest lies in testing the hypothesis H_0: $\beta_2 = 0$ against the two sided alternative hypothesis. The quasi test statistic used is

$$\tau^2 = \frac{\hat{\beta}_2^2}{\widehat{var}(\hat{\beta}_2)}$$

where $\hat{\beta}_2$ denotes the OLS estimate of β_2 and $\widehat{var}(\hat{\beta}_2)$ is variance estimate of $\hat{\beta}_2$ and is based on the HC_1, HC_3, HC_4 and HC_{4m} estimators. The asymptotic distribution of τ^2 is chi-square with one degree of freedom (χ_1^2), see [20]. The data is generated under the null hypothesis for all considered simulation design. We have calculated the relative probability discrepancies for all the estimators. Relative probability discrepancies (RPD) is computed as,

- Select the exact probabilities from the desired distribution for the desired level of significance γ here in our study we compute the exact probabilities from χ_1^2, i.e $P(\chi_{1,\gamma}^2)$ and γ lies between 0 and 1.
- Compute the test statistic τ^2 for each of the N Monte Carlo runs.
- The relative probability discrepancy is defined as

$$RPD = \frac{\#(\tau^2 < \chi_{1,\gamma}^2)/N - \gamma}{\gamma}$$

We consider the model with unequal disturbances with leverage points and in order to set the level of heteroscedasticity, which is denoted by λ and computed as $\lambda = max(\sigma_i^2/min(\sigma_i^2)), i = 1,\ldots,n$, in simulations we use α = 0.26 by to obtain $\lambda \approx 100$ following [10], when $k = 3$, we set $\alpha_1 = \alpha_2 = 0.26$ and when $k = 5$, we set $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.26$ to obtain $\lambda \approx 100$, where λ is the strength of

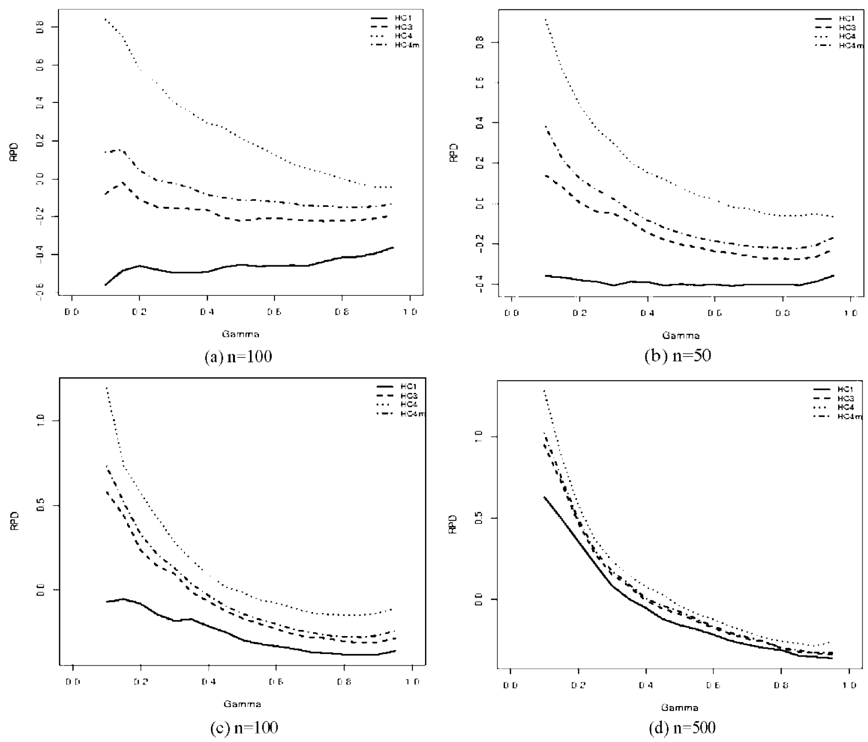

Fig. 1: RPD versus asymptotic probabilities (Gamma); covariates value are selected randomly from LN(0,1), i.e standard lognormal distribution with k = 3 and n = 25, 50, 100, 500

heteroscedasticity. The number of Monte Carlo replications is 1000. All simulations results are carried out using the R programming language see [21].

We use Monte Carlo method to compute different quasi t-test with independent and heteroscedastic errors and with three and five regression parameters for various choices of n. To study the effect of underlying distribution of covariates, we consider the following cases for the regression model given in (10).

Case-I: Lognormal(LN) $X_{ij} \sim e^{Z_{ij}}$, where Z_{ij} ~ (0,1), where $i = 1,...,n$ and $j = 1,...,k-1$,

Case-II: Chi-square $X_{ij} \sim \chi_v^2$, where $v = 1,3,5$ is degree of freedom.

Figure 1 shows the plots of the RPD against the corresponding asymptotic probabilities, γ andfor various choices of n, when the distribution of predictors is standard lognormal distribution. We present the results for test statistics which are based on the variances from HC_1, HC_3, HC_4 and HC_{4m} in heteroscedastic case. We simulate all the values of the covariates using Monte Carlo simulation and predictors are independent and random. To check the performance of the estimators from the graph of RPD, we see that how close the discrepancy lines are to the zero line, the closer the discrepancy lines to the zero line ($RPD = 0$), the more reliable the inference will be, see e.g [12].

When $k = 2$ and the distribution of the predictors is LN(0,1), as we move away from the tail area the first order asymptotic approximation of the HC_4 statistic under the null distribution became very poor and its behavior is approximately same even for large sample sizes (Figure 1).The first order approximation of the HC_{4m} is better. Comparatively HC_4 test is the poor performing test for considered all sample sizes and for the lower values of gamma ($\gamma < 0.7$), but as the value of gamma increases ($\gamma \geq 0.7$) the performance of HC_4 is better than all other estimators especially with the large sample size, see Figure 1 (d). The Figure 1 elaborates that with the increase in the value of γ the value of RPD decreases, implies that, as the value of γ increases the probability of rejecting the null hypothesis decreases. Results of RPD for HC_3 clearly shows that the approximation of this estimator is better for lower values of gamma ($\gamma \leq 0.8$) as compared with HC_4 and HC_{4m} but as value of gamma approaches to one the RPD of HC_3 becomes negative for both $k = 3$ and $k = 5$.

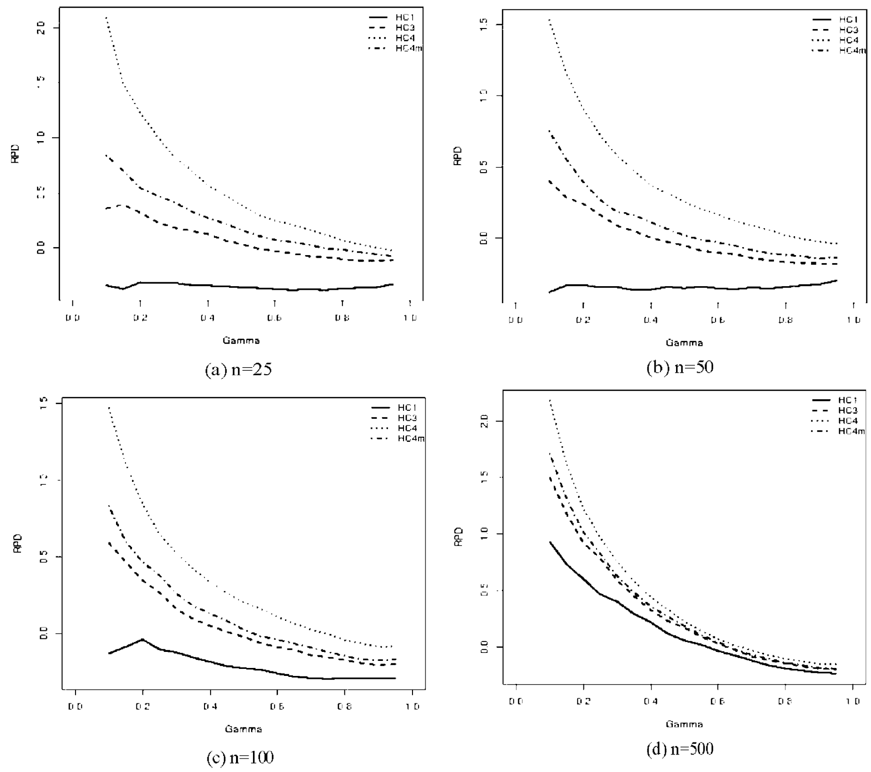

Fig. 2: RPD versus asymptotic probabilities (Gamma); covariates value are selected randomly from LN(0,1), i.e standard lognormal distribution with k = 5 and n = 25, 50, 100, 500.

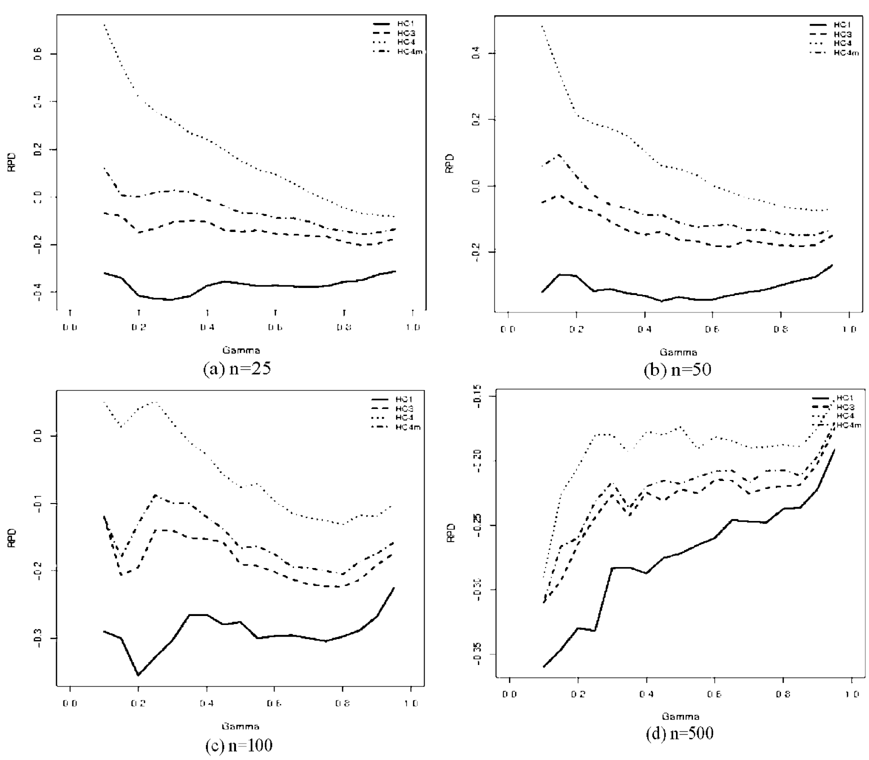

Fig. 3: RPD versus asymptotic probabilities (Gamma); covariates value are selected randomly from $\chi 2$ (df=1) distribution with k = 3 and n = 25, 50, 100, 500

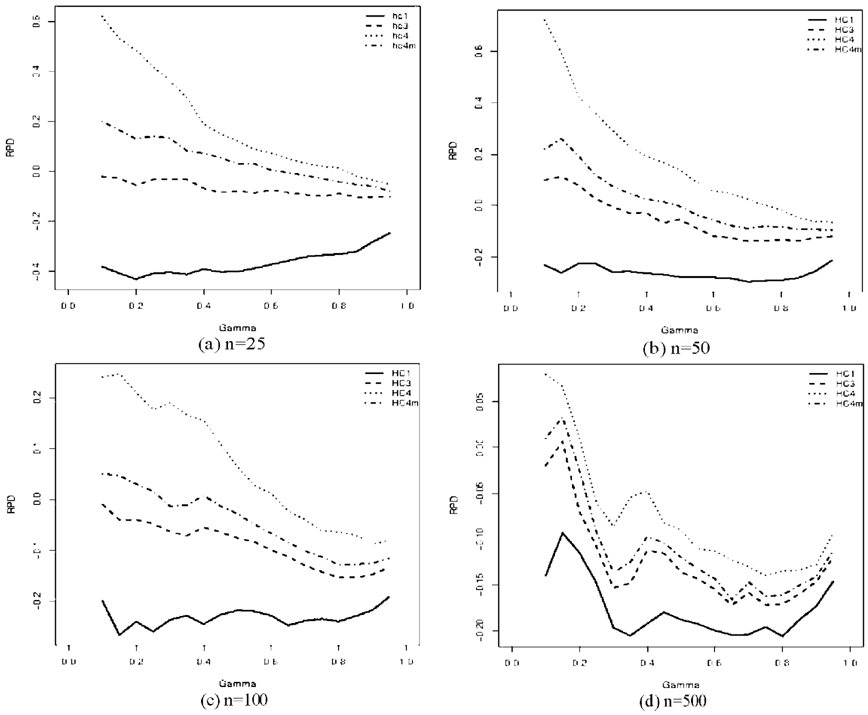

Fig. 4: RPD versus asymptotic probabilities(γ); covariates value are selected randomly from χ2 (df=1)distribution with k = 5 and n = 25, 50, 100, 500

Table 1: Tests of heteroscedasticity

Test	LM	df	p-value
Breusch-Pagan test	30.02	3	0.000
Goldfield-Quandt test	1.99	33	0.026

For $k = 5$, the estimators exhibit a similar kind of behavior as shown in Figure 2, as was for $k = 3$. From the above discussion and the results obtained it can be concluded that the overall performance of HC_{4m} is comparatively better than all other estimators. These results are in agreement with those obtained by [5] who showed that the HC_{4m} outperformed other HCCMEs when there are extreme values in the data and the distribution of errors is non-normal.

In Case II, it is found that the performance of the estimators with the smaller degree of freedom i.e, 1 and with smaller sample sizes, $n = 25$ and $n = 50$ is similar to that of standard lognormal distribution and as the sample size increases the RPD of all the estimators decreases gradually and become negative and when $n = 500$ all the estimators have RPD below 0 and highly underestimate the null hypothesis see

Figure 3. When $k = 5$ and the sample size is small there is no change in the behavior of the estimators see Figure 4 (a) and (b)for n = 100 HC_4 perform better for $\gamma <$ 0.6 and with 500 sample size the situation is again similar to that of $k = 3$. We also check the performance of considered HCCMEs by increasing the degree of freedom of chi square up to 2 and 5 in Case II and find that the as the degree of freedom increases all the estimators shows very poor performance and under reject the null hypothesis for all the sample sizes.

Real Example: In this section, we The data set used contains the information regarding the house price of sample of 88 London houses together with some characteristic regarding those houses given in [3] chapter. 7. The regression model according to the data is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i, i = 1,\ldots,88,$$

where, the dependent variable Y, is the price of the house, X_1 is the number of bedrooms in the house, X_2 is the lot size and X_3 is the size of house in square fit. In order to test the presence of heteroscedasticity in the data we apply the Breusch-Pagan test and Goldfield-Quandt test and the results are given in Table 1. Both the tests reject the null hypothesis of homoskedasticity at the 5% level of significance, which implies that the heteroscedasticity is present in the data.

Table 2: Standard deviations (S.D), test-statistic (t) and their p-values for significance testing of regression coefficients

	β_1			β_2			β_3		
Test	S.D	t	p-value	S.D	t	p-value	S.D	t	p-value
HC_1	8479	1.633	0.106	1.251	1.652	0.102	17.725	6.926	0.000
HC_3	11562	1.198	0.234	7.148	0.289	0.713	40.732	3.014	0.003
HC_4	43551	0.318	0.751	45.326	0.045	0.964	231.658	0.529	0.598
HC_{4m}	14395	0.962	0.338	11.332	0.182	0.856	60.694	20.023	0.046
OLS	9010	1.537	0.128	06.42	3.220	0.001	13.24	9.275	0.000

The null hypothesis under consideration is H_0: $\beta_0 + \beta_1 = \beta_2 = \beta_3 = 0$ against the alternative hypothesis H_1: $\beta_j \neq 0$ for j=1,2,3. The results for the inference for the model (13) according to the considered null hypothesis and the p-values are given in Table 2. From the table we can see that for β_1 OLS accept the null hypothesis at 5% level of significance.

The HCCMEs shows the same conclusion about the β_1 and do not reject the null hypothesis. For β_2 the the OLS reject the null hypothesis while all the HCCMEs accept the null hypothesis which means that there is no relation between X_2 and Y. Similarly for β_3 only CH_4 accept the null hypothesis and the OLS and all other HCCMEs reject the null hypothesis at 5% level of significance.

From the above results we can conclude that among the HCCMEs HC_1 estimator is providing the precise inference for the considered data set. It has small standard deviation for all the parameters and also provide reliable inference as compared with the other HCCMEs. The test statistics which is based on HC_4 estimator has the largest p-value as compared with other estimator for all the parameters. Thus HC_4 test is the test which has the smallest amount of evidence against H_0.

CONCLUSION

The asymptotic distribution of quasi test statistic based on HC estimators is more appropriate, in general, when variance is estimated using HC_3 and HC_{4m}. In general, when sample size is large and the predictors are normally distributed then relative probability discrepancy for the all four considered HC estimators is closer to each other. For small sample size, the approximation of asymptotic chi-square distribution of the quasi test statistic is poor especially at the tail for HC_4 and HC_1. The quasi test statistic has heavy left tail when defined on HC_4, while the situation is opposite for HC_1. Interestingly, HC_1, especially for small sample size, has generally negative relative probability discrepancy and the amount of relative probability discrepancy does not seem to depend on the nominal size of the asymptotic distribution.

Our simulation results confirm the numerical results of the [10] that the asymptotic approximation of the HC_{1m} is better than others. Particularly this deficiency looks more intense in the HC_4 estimator based test statistic. But according to our results the performance of the HC_3 is also efficient especially for small sample size. [16] also suggested the use of HC_3 when sample size is less then or equal to 250,. The results obtained clearly favor the use of newly purposed HC_{4M} in hypothesis testing inference. It is also concluded that when the sample size is large, $n = 500$ the performance of all the estimators becomes approximately same.

ACKNOWLEDGEMENT

We thank the Editor and the referees for their helpful comments and suggestions that greatly improved the previous version of the paper.

REFERENCES

1. 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.
2. 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.
3. Huber, P.J., 1967. The behavior of maximum likelihood estimates under nonstandard conditions. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, University of California Press, 1: 221-233.
4. Angrist, J.D. and J.S. Pischke, 2009. Mostly harmless econometrics: An empiricist's companion. Princeton University Press.
5. Chesher, A. and I. Jewitt, 1987. The bias of a heteroskedasticity consistent covariance matrix estimator. Econometrica: Journal of the Econometric Society, 55(5): 1217-1222.

6. Hinkley, D.V., 1977. Jackknifing in unbalanced situations. Technometrics, 19: 285-292.
7. Horn, S.D., R.A. Horn and D.B. Duncan, 1975. Estimating heteroscedastic variances in linear models. Journal of the American Statistical Association, 70: 380-385.
8. Davidson, R. and J.G. MacKinnon, 1993. Estimation and inference in econometrics. Oxford University Press, USA.
9. Cribari-Neto, F., T.C. Souza and K.L.P. Vasconcellos, 2007. Inference under heteroscedasticity and leveraged data. Communications in Statistics-Theory and Methods, 36(10): 1877-1888.
10. Cribari-Neto, F. and W.B. da Silva. 2011. A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model. AStA Advances in Statistical Analysis, 95(2): 129-146.
11. Cribari-Neto, F. and S.G. Zarkos, 1999. Bootstrap methods for heteroskedastic regression models: evidence on estimation and testing. Econometric Reviews, 18(2): 211-228.
12. Cribari-Neto, F. and S.G. Zarkos, 2004. Leverage-adjusted heteroskedastic bootstrap methods. Journal of Statistical Computation and Simulation, 74(3): 215-232.
13. Hodoshima, J. and M. Ando, 2006. The effect of non independence of explanatory variables and error term and heteroskedasticity in stochastic regression models. Communications in Statistics-Simulation and Computation, 35(2): 361-405.
14. Vajargah, K.F., F. Mehrdoust and F. Kamalzadeh, 2012. Variance estimation of linear regression coefficients using markov chain monte carlo simulation. Middle-East Journal of Scientific Research, 11(9): 1317-1322.
15. MacKinnon, J.G. and H. White, 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. Journal of Econometrics, 29(3): 305-325.
16. Long, J.S. and L.H. Ervin, 2000. Using heteroscedasticity consistent standard errors in the linear regression model. The American Statistician, 54(3): 217-224.
17. Cribari-Neto, F. and S.G. Zarkos, 2001. Heteroskedasticity-consistent covariance matrix estimation: white's estimator and the bootstrap. Journal of Statistical Computation and Simulation, 68(4): 391-411.
18. Cribari-Neto, F. and A.L. Maria da Glória, 2009. Heteroskedasticity-consistent interval estimators. Journal of Statistical Computation and Simulation, 79(6): 787-803.
19. Arce, M. and A. Mora, 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.
20. Johnson, N.L., S. Kotz and N. Balakrishnan, 1994. Continuous univariate distributions, vol. 1. John Wiley & Sons: New York, USA.
21. Development, R., 2011. Core Team. R: A Language and Environment for Statistical Computing. RFoundation for Statistical Computing, Vienna, Austria.
22. Asteriou, D. and S.G. Hall, 2007. Applied econometrics: a modern approach using EViews and microfit. Palgrave Macmillan.