IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN:2319-765X. Volume 10, Issue 3 Ver. VI (May-Jun. 2014), PP 10-21
www.iosrjournals.org

Generalization of Consistent Standard Error Estimators under Heteroscedasticity

[1]O. Job and, [2]B. A Oyejola
[1,2]*(Department of Statistics, University of Ilorin, Ilorin. Nigeria)*

Abstract : *In many econometric studies, especially those based on cross-sectional data, the assumption of a constant variance for the disturbance term is unrealistic. For example in consumer budget studies (micro consumption function), the residual variance about the regression function is very likely to increase with income. Also, in cross-sectional studies of firms the residual variance probably increases with the size of the firm. In a simple relation the dependent variable Y is explained by Z. Thus we assume* $y = f(z) + e$*.In this formulation* $\text{var}(y) = \text{var}(e) = \sigma^2 Z$*. Using* $Z_i = X_i$ *just gives us a way of formulating the assumption about* $\text{var}(e_i) = \sigma^2 Z_i$ *in a fairly general manner. To make this assumption operational and general, it is convenient and quite plausible to specify the form of association* $\text{var}(e_i) = \sigma^2 Z_i^g$ *where g is the strength of heteroscedasticity, the lower the magnitude of g, the smaller the difference between the individual variances. When g =0, the model is homoscedastic otherwise* $|g| \le 2$ *generally.*

This paper proposes a generalization of consistent standard error (CSE) estimators denoted by HC5.Comparison of this proposed estimator and other CSE estimators using various strength of heteroscedasticity at sample sizes 25, 30, 35, 40, 45, and 50 was done.

The OLS estimator remains unbiased and the results showed that the developed estimator is indeed a generalization of all and produces a consistent and asymptotic efficient.

Keywords: *Generalization, Heteroscedasticity Consistent Standard Error Estimator, Monte Carlo Simulation, simple linear regression model, Error terms, weighting factor.*

I. Introduction

The usual OLS estimator for the variance of model parameters minimizes the error sum of squares. Heteroscedasticity does not destroy the unbiasedness of the OLS estimators except that it becomes less efficient. Assuming a simple linear regression model,

$$y_i = \beta_0 + \beta_1 x_i + e_i ------------(1) \text{ for } i=1,\ldots,n$$

Where $e_i \sim N(0, \sigma^2)$, i.e. homoscedasticity assumption holds. Then the variance of the model parameters is given by,

$$\text{var}(\beta_j) = \frac{\sum \hat{e}_i^2}{n-k}(X'X)^{-1} = \frac{\sum \hat{e}_i^2}{(n-k)\sum x_i^2} --- (2)$$

This is referred to as ordinary least squares covariance matrix estimator (OLSCME).Building on the works of [1] and [2] using the OLS residuals as estimators of the errors. The sandwich estimator HC0 was derived.

$$HC0 = (X'X)^{-1} X' \Phi X (X'X)^{-1} --------- (3)$$

Where Φ, is a diagonal matrix with the squared OLS residuals on the main diagonal and zero elsewhere. Hence equation (3) becomes

$$HC0 = (X'X)^{-1} X' diag(\hat{e}_i^2) X (X'X)^{-1} ------ (4)$$

When this assumption of homoscedasticity is violated ($\Phi \ne \sigma^2 I$) then the variance of the model parameters is no longer given by equation (2)

[3] modified the OLSCM when he introduced a degree of freedom correction that inflates each residual by a weighting factor, $\sqrt{\frac{n}{n-k}}$ which resulted in

$$HC1 = \frac{n}{n-k}(X'X)^{-1}X'diag(e_i^2)X(X'X)^{-1} = \frac{n}{n-k}HC0 --- (5)$$

Where k=p+1 and p is the number of predictor variable in the regression model.

[4] developed another variation of OLS estimator by scaling the squares of residuals by $(1-h_{ii})^{-1}$, where $h_{ii=}x_i(X'X)^{-1}x'_i$ then $\mathrm{var}(e_i) = \sigma^2(1-h_{ii}) \neq \sigma^2$

That is, $HC2 = (X'X)^{-1}X'diag\left[\frac{e_i^2}{(1-h_{ii})}\right]X(X'X)^{-1} --- (6)$ with $(1-h_{ii})^{-1}$ as a weighting factor.

A third variation by [5] gave an approximation to a more complicated jackknife estimator than the earlier one presented by [4] as

$$HC3 = (X'X)^{-1}X'diag\left[\frac{e_i^2}{(1-h_{ii})^2}\right]X(X'X)^{-1} ----- (7)$$

Such that $0 \leq h_{ii} \leq 1$, with $(1-h_{ii})^{-2}$ as a weighting factor.

This further inflates the squares of residuals.

A fourth variation was derived by [6] as

$$HC4 = (X'X)^{-1}X'diag\left[\frac{e_i^2}{(1-h_{ii})^{\delta i}}\right]X(X'X)^{-1} ---- (9)$$

Where $\delta i = \min(4, \frac{nh_{ii}}{k}) -------------- (10), k = p+1$

With $(1-h_{ii})^{-\delta i}$ as a weighting factor and p is the number of predictors in the model. In the previous studies the existing HSCE were subjected to analysis assuming heteroscedasticity of unknown form. The choice of weighting factors focuses on making OLS squares residuals less bias.

II. Model Specification

We shall consider a classical linear regression model in (1) above

Where β_0 and β_1 are the unknown true parameters, $\beta_0 = 2$, $\beta_1 = 3$

$e_i \sim N(0, x_i^g)$. The model shall be studied under the strength of heteroskedasticity (g) ranging from -2 to +2.

III. Proposed Generalization

The proposed estimator is

$$HC5 = (X'X)^{-1}X'diag\left[\frac{e_i^2}{(1-h_{ii})^g}\right]X(X'X)^{-1} ---- (11)$$

$0 \leq h_{ii} \leq 1$, with $(1-h_{ii})^{-g}$ as a weighting factor. Where "g" is the strength of heteroscedasticity that characterizes the error terms relating to the predictor in simple regression model and is estimable from the bivariate data(x, y). In this study, strength of heteroscedasticity "g" ranges from -2 to +2.

IV. The Design Of Monte Carlo Studies Used For Data Generation Is As Follows:

(I) formulate the statistical model to study

(II) Specify the distribution of the normal error term to obtain values for the unobserved error terms (e_i)

(III) Specify the distribution of the independent variable(s) X

(IV) Specify the values of the unknown true parameters of regression model β_0, β_1

(V) Generate the values of the dependent variable(Y) from the equation which involves using the true values of the model parameters, generated values of independent variable and values of the unobserved error terms.

(VI) Regress Y on X in order to obtain estimates of the residuals $(\hat{e}_i)$.

(VII) Generate the values of the heteroskedastic residuals using the estimated residuals above, as $\hat{e}_i^{/} = \hat{e}_i * x_i^{\frac{g}{2}}$, g=-2, -1, -0.5, 0, 0.5, 1, 2 respectively

(VIII) Generate values of the response variable y_i using the pre-specified values of unknown parameters, independent variable(s) and the heteroskedastic residuals.

(IX) Using data set (y_i, x_i) test for the presence of heteroskedastidity using Goldfeld-Quandt test [7].

(X) If heteroskedasticity is present in (IX) go to step (XI) otherwise go to step V.

(XI) Regress Y on X using the data set and obtain the square of residual $(\hat{e}_i^2)$

(XII) Regress log ($\hat{e}_i^2$) on the log (x_i), carry out the appropriate test of significance in order to obtain the estimates of the strength of heteroskedasticity $(\hat{g})$ in the model [8].

(XIII) If estimated value of the slope $(\hat{g})$ in (XII) is approximately equal to g introduced in the (VII) above and the test of significance supports the estimates in (XII), then subject the data set to the existing HCSE and HC5 estimators. Otherwise restart the processes above.

Using the data generated, estimates were therefore obtained for these estimators with varying sample sizes at replications, 1000.

See appendix"A1" for the R snippet on Monte Carlo simulation experiment used in this research.

For Example

5. Generation Of Sample Values When Sample Size=25, And Parameters $\beta_0 = 2, \beta_1 = 3$, And $g = -2$

The generated error term distributed mean zero and variance=1 follows:

e<-c(0.79259937,0.15159755,0.22236327,4.07372642,-0.66352880,2.73412938, 13339313, 1.06999387, 1.41852392,-0.12010026,-2.39275398,-2.49368351,-0.45371828,0.03439936, -1.97778767, 1.85725137, 2.70763668,-0.25774271, -2.29936163, 3.75785389, 1.67069689, 0.91467720, 1.35848385, 2.66101714,-2.31255552)

The values for the variable X simulated follows:

X<-c(7.227193, 4.850246, 4.624653, 7.142836, 6.245022, 5.349890, 3.538438, 5.432725, 4.447327, 4.657059, 7.345125, 5.589621, 4.050712, 6.692459, 2.492058, 3.683330, 335663, 8.707044, 6.726798, 1.482140, 2.764793, 7.475252, 5.348404, 7.265780, 8.207077)

The generated values of the response variable Y follows:

y=2+3x+e

y<-c(23.779468, 16.447947, 15.769862, 23.984105, 20.577797, 18.461407, 2.381149, 18.400940, 15.493217, 15.795923, 23.701839, 18.237863, 13.832704, 22.050801, .1967150, 13.302403, 18.414223, 28.122153, 21.808233, 8.0030190, 10.484203, 24.544614, 8.199790, 24.153124, 26.357462)

Obtain the estimate of residual by regressing Y on X, and then generate values for heteroscedastic error term as follows:

$$\hat{e}_i^* = \hat{e}_i * x_i^{g/2}$$

$\hat{e}_i^*$ <-c (0.09788777, -0.10278944 -0.10409695, 0.55559569, -0.15726749, 0.41173639, -0.23416495, 0.10276592, 0.15123638, -0.17525494,-0.33353644, -0.53099915, -0.31943055, -0.02657525, -1.27945905, 0.25241322, 0.40723524, 0.00101939,-.37216107, 1.55659729, 0.18982553, 0.11885831, 0.15457802, 0.35578338, -0.26376994)

The values for the heteroscedastic variable Y becomes:

Y=2+3x+ $\hat{e}_i^*$

y <-c(23.646433, 16.587695, 15.917761, 23.395921, 20.729712, 18.071472, 12.692068, 18.317463, 15.391162, 16.013998, 23.996650, 18.783392, 14.213348, 22.058451, 9.584670, 13.122348, 18.029221, 28.041093, 22.160426, 6.585554, 10.394600, 24.383083, 18.067059, 23.761022, 26.556359)

Test for the presence of heteroscedasticity using Goldfeld-Quandt Test:

GQ = 4.2001, df1 = 11, df2 = 10,

P-value = 0.01574. Since the P-value is less than the pre-selected alpha this indicates heteroscedasticity is likely present.

Obtain the strength of heteroscedasticity in order to determine its form

By regressing log ($\hat{\hat{e}}_i^2$) on log (x_i)

coefficient	Estimate	s.e	t-value	P-value
intercept	0.176	1.6173	0.109	0.9140
G	-2.0110	0.9637	-2.087	0.0482

g=-2.0110, $\operatorname{var}(e_i) = \sigma^2 Z_i = \sigma^2 x_i^{-2}$

AVERAGE VALUE OF $\hat{\beta}_0$ =2.006492, BIAS ($\hat{\beta}_0$) =0.0065, ABSOLUTE BIAS ($\hat{\beta}_0$) =0.0065

AVERAGE VALUE OF $\hat{\beta}_1$ =2.099893, BIAS ($\hat{\beta}_1$) =-0.00344, ABSOLUTE BIAS ($\hat{\beta}_1$) =0.00344

USING THE EXISTING HCSE AND HC5 ESTIMATOR, THE ESTIMATES OF COVARIANCES, WHEN g=-2, n=25 ARE GIVEN AS BELOW:

HC0= $\begin{pmatrix} 0.30949320 & -0.048650450 \\ & 0.007751424 \end{pmatrix}$ **HC1=** $\begin{pmatrix} 0.33640565 & -0.052880923 \\ & 0.008425461 \end{pmatrix}$

HC2= $\begin{pmatrix} 0.38641945 & -0.060778546 \\ & 0.009671662 \end{pmatrix}$ HC3= $\begin{pmatrix} 0.48403331 & -0.07617608 \\ & 0.01210936 \end{pmatrix}$

HC4= $\begin{pmatrix} 0.56679233 & -0.08933222 \\ & 0.01419241 \end{pmatrix}$ HC5= $\begin{pmatrix} 0.2005973 & -0.031494896 \\ & 0.005034438 \end{pmatrix}$

The Standard error estimates of the model parameters shall be obtained by taking the square root of the entries on the main diagonals of the co variances matrices. See Appendix"A5" and Appendix"A6"respectively

V. Numerical Results

Table1: Average Estmates Of $\hat{\beta}_0$ And $\hat{\beta}_1$, Bias And Absolute Bias Of Ols For Different Strength Of Heteroscedasticity

Sample size	Level	$\bar{\hat{\beta}}_0$	$\bar{\hat{\beta}}_1$	Bias($\bar{\hat{\beta}}_0$)	Bias($\bar{\hat{\beta}}_1$)	Absolute Bias($\bar{\hat{\beta}}_0$)	Absolute Bias($\bar{\hat{\beta}}_1$)
25	-2	2.006492	2.99893	0.006492	-0.001070	0.006492	0.001070
	-1	2.003437	2.999502	-0.003437	-0.000498	0.003437	0.000498
	-½	2.011658	2.997627	-0.011658	0.003730	0.011658	0.003730
	0	2.00000	3.00000	0.000000	0.000000	0.000000	0.000000
	½	2.003248	2.999622	-0.003248	0.000378	0.003248	0.000378
	1	2.020864	2.996797	-0.020864	0.003203	0.020864	0.003203
	2	1.994345	3.001383	0.005655	-0.001383	0.005655	0.001383

30	-2	2.014509	2.996844	0.014509	0.003156	0.014509	0.003156
	-1	1.996752	3.00061	0.003248	-0.000610	0.003248	0.000610
	-½	2.002684	2.999539	-0.002684	0.000461	0.002684	0.000461
	0	2.00000	3.00000	0.000000	0.000000	0.000000	0.000000
	½	2.00394	2.999241	-0.00394	0.000759	0.00394	0.000759
	1	1.978417	3.003813	0.021583	-0.003813	0.021583	0.003813
	2	2.00686	2.998429	-0.006860	0.001571	0.006860	0.001571
35	-2	2.019945	2.999136	0.019945	0.000864	0.019945	0.000864
	-1	1.984136	3.003443	0.015864	-0.003443	0.015864	0.003443
	-½	2.002723	2.999497	-0.002723	0.000503	0.002723	0.000503
	0	2.00000	3.00000	0.000000	0.000000	0.000000	0.000000
	½	1.998292	3.00281	0.001708	-0.002810	0.001708	0.002810
	1	1.993081	3.001168	0.006919	-0.001168	0.006919	0.001168
	2	1.991589	3.002033	0.008411	-0.002033	0.008411	0.002033
40	-2	2.005019	2.999136	0.005019	0.000864	0.005019	0.000864
	-1	2.00447	2.999065	-0.00447	0.000935	0.004470	0.000935
	-½	1.99870	3.000302	0.001300	-0.000302	0.001300	0.000302
	0	2.00000	3.00000	0.000000	0.000000	0.000000	0.000000
	½	1.996941	3.000495	0.003059	-0.000495	0.003059	0.000495
	1	1.991806	3.001398	0.008194	-0.001398	0.008194	0.001398
	2	1.989425	3.003047	0.010575	-0.003047	0.010575	0.003047
45	-2	1.998827	3.000183	-0.001173	-0.000183	0.001173	0.000183
	-1	1.984091	3.003301	0.015909	-0.003301	0.015909	0.003301
	-½	2.000513	2.999901	-0.000513	0.000099	0.000513	0.000099
	0	2.00000	3.00000	0.000000	0.000000	0.000000	0.000000
	½	1.990715	3.001506	0.009285	-0.001506	0.009285	0.001506
	1	1.986978	3.002143	0.013022	-0.002143	0.013022	0.002143
	2	2.001857	2.999581	-0.001857	0.000419	0.001857	0.000419
50	-2	1.992516	3.001224	-0.007484	-0.001224	0.007484	0.001224
	-1	2.010046	2.998024	-0.010046	0.001976	0.010046	0.001976
	-½	1.98750	3.002355	0.012500	-0.002355	0.012500	0.002355
	0	2.00000	3.00000	0.000000	0.000000	0.000000	0.000000
	½	1.998667	3.000217	0.001333	-0.000217	0.001333	0.000217
	1	1.998421	3.000265	0.001579	-0.000265	0.001579	0.000265
	2	2.061664	2.985942	-0.061664	0.014058	0.061664	0.014058

From Table1 it can be verified that OLS estimator remains unbiased even when regression error term violates homoscedasticity assumption. The estimates of bias and absolute bias are very close to zero while at g=0 the bias and absolute bias are equal zero exactly. The standard error estimates is a sufficient criterion to judge the performance of the heteroscedasticity consistent covariance matrix estimators (HCCME).

Table 2: The estimates of $s.e(\beta_1)$ from the different Estimators and R^2 when g=-2, at varying sample sizes.

		Strength of heteroscedasticity=-2						
		HC0	**HC1**	**HC2**	**HC3**	**HC4**	**HC5**	
Sample(n)	$\hat{\beta}_1$	S.E	S.E	S.E	S.E	S.E	S.E	R^2
25	**2.9697**	**0.0880**	**0.0918**	**0.0983**	**0.1100**	**0.1191**	**0.07095**	**0.9649**
30	**2.9280**	**0.0835**	**0.0864**	**0.0885**	**0.0939**	**0.0924**	**0.0742**	**0.9881**
35	**2.7961**	**0.1942**	**0.2000**	**0.2191**	**0.2475**	**0.3141**	**0.1533**	**0.9416**
40	**3.0146**	**0.0473**	**0.0486**	**0.0497**	**0.0522**	**0.0522**	**0.0430**	**0.9924**
45	**2.9856**	**0.0361**	**0.0369**	**0.0377**	**0.0395**	**0.0399**	**0.0331**	**0.9954**
50	**2.9906**	**0.0318**	**0.0324**	**0.0327**	**0.0338**	**0.0335**	**0.0299**	**0.9951**

From Table 2, when g=-2, at various sample sizes. HC5 gives the minimum estimates of $s.e(\beta_1)$, While HC0 and HC1 offered almost similar results. HC2, HC3, and HC4 follow in that order. HC4 gives the maximum $s.e(\beta_1)$.

Table 3: The estimates of $s.e(\beta_1)$ from the different Estimators and R^2 when g=-1, at varying sample sizes.

		Strength of heteroscedasticity=-1						
		HC0	**HC1**	**HC2**	**HC3**	**HC4**	**HC5**	
Sample(n)	$\hat{\beta}_1$	S.E	S.E	S.E	S.E	S.E	S.E	R^2
25	**2.9805**	**0.1207**	**0.1258**	**0.1299**	**0.1398**	**0.1378**	**0.1122**	**0.9738**
30	**3.0119**	**0.0659**	**0.0682**	**0.0694**	**0.0733**	**0.0729**	**0.0626**	**0.9828**
35	**2.9191**	**0.2643**	**0.2722**	**0.2846**	**0.3066**	**0.3191**	**0.2456**	**0.8943**
40	**3.0180**	**0.0594**	**0.0610**	**0.0619**	**0.0644**	**0.0639**	**0.0571**	**0.9811**
45	**3.2079**	**0.2219**	**0.2270**	**0.2351**	**0.2494**	**0.2602**	**0.2096**	**0.9223**
50	**3.0505**	**0.1333**	**0.1360**	**0.1392**	**0.1455**	**0.1473**	**0.1276**	**0.9518**

From Table 3, when g=-1, at various sample sizes. HC5 gives the minimum estimates of $s.e(\beta_1)$, While HC0 and HC1 offered almost similar results. HC2, HC3, and HC4 follow in that order. HC4 gives the maximum $s.e(\beta_1)$.

Table 4: The estimates of $s.e(\beta_1)$ from the different Estimators and R^2 when g=-½, at varying sample sizes.

		Strength of heteroscedasticity=-½						
		HC0	**HC1**	**HC2**	**HC3**	**HC4**	**HC5**	
Sample(n)	$\hat{\beta}_1$	S.E	S.E	S.E	S.E	S.E	S.E	R^2
25	**3.06**	**0.1900**	**0.1981**	**0.2106**	**0.2339**	**0.2492**	**0.1806**	**0.9355**
30	**2.9964**	**0.1254**	**0.1298**	**0.1327**	**0.1406**	**0.1402**	**0.1219**	**0.9384**
35	**2.9857**	**0.0912**	**0.0939**	**0.0946**	**0.0980**	**0.0956**	**0.0896**	**0.9567**
40	**3.0126**	**0.1647**	**0.1690**	**0.1734**	**0.1827**	**0.1840**	**0.1606**	**0.928**
45	**3.0210**	**0.1049**	**0.1073**	**0.1092**	**0.1138**	**0.1135**	**0.1028**	**0.9417**
50	**2.9213**	**0.1278**	**0.1305**	**0.1361**	**0.1452**	**0.1618**	**0.1239**	**0.9384**

From Table 4, when g=-½, at various sample sizes. HC5 gives the minimum estimates of $s.e(\beta_1)$, While HC0 and HC1 offered almost similar results. This may be as a result of slight heteroscedasticity (i.e. slight deviation from when g=0). HC2, HC3, and HC4 follow in that order. HC4 gives the maximum $s.e(\beta_1)$.

Table5: The estimates of $s.e(\beta_1)$ from the different Estimators and R^2 when, g=0 at varying sample sizes.

		Strength of heteroscedasticity=0						
		HC0	**HC1**	**HC2**	**HC3**	**HC4**	**HC5**	
Sample(n)	$\hat{\beta}_1$	S.E	S.E	S.E	S.E	S.E	S.E	R^2
25	**3.0000**	**0.1914**	**0.1996**	**0.2012**	**0.2119**	**0.2081**	**0.1914**	**0.8695**
30	**3.0000**	**0.2246**	**0.2325**	**0.2380**	**0.2526**	**0.2520**	**0.2246**	**0.8573**
35	**3.0000**	**0.2077**	**0.2139**	**0.2158**	**0.2243**	**0.2194**	**0.2077**	**0.8584**
40	**3.0000**	**0.2173**	**0.2230**	**0.2301**	**0.2438**	**0.2505**	**0.2173**	**0.8559**
45	**3.0000**	**0.1623**	**0.1660**	**0.1680**	**0.1741**	**0.1766**	**0.1623**	**0.8770**
50	**3.0000**	**0.1503**	**0.1534**	**0.1599**	**0.1706**	**0.1893**	**0.1503**	**0.9191**

From Table 5, When g=0, at various sample sizes. HC0 and HC5 give the minimum estimates of $s.e(\beta_1)$, While HC1 and HC2 offered close estimates of $s.e(\beta_1)$. HC2, HC3, and HC4 follow in that order. HC4 offered the maximum $s.e(\beta_1)$.

Table6: the estimates of $s.e(\beta_1)$ from the different Estimators and R^2 when g=½ at varying sample sizes.

		Strength of heteroscedasticity=½						
		HC0	**HC1**	**HC2**	**HC3**	**HC4**	**HC5**	
Sample(n)	$\hat{\beta}_1$	S.E	S.E	S.E	S.E	S.E	S.E	R^2
25	**3.0342**	**0.1727**	**0.1800**	**0.1914**	**0.2145**	**0.2627**	**0.1816**	**0.8773**
30	**3.0489**	**0.3630**	**0.3757**	**0.3823**	**0.4027**	**0.3934**	**0.3725**	**0.7420**
35	**2.9764**	**0.1062**	**0.1093**	**0.1114**	**0.1170**	**0.1174**	**0.1087**	**0.9408**
40	**3.0138**	**0.0866**	**0.0889**	**0.0898**	**0.0932**	**0.0934**	**0.0882**	**0.9544**
45	**2.9873**	**0.1116**	**0.1141**	**0.1174**	**0.1236**	**0.1271**	**0.1144**	**0.9394**
50	**3.0065**	**0.0986**	**0.0950**	**0.0963**	**0.0997**	**0.0996**	**0.0947**	**0.9275**

From Table 6, When g=½, at various sample sizes. HC5 gives estimates of $s.e(\beta_1)$ almost equal to that of HC1 but both may not be significantly different from that of HC2. This may be as a result of slight heteroscedasticity (i.e. deviation from when g=0). HC3 and HC4 follow in that order. HC4 gives the maximum $s.e(\beta_1)$.

Table7: the estimates of $s.e(\beta_1)$ from the different Estimators and R^2 when g=1, at varying sample sizes.

		Strength of heteroscedasticity=1						
		HC0	**HC1**	**HC2**	**HC3**	**HC4**	**HC5**	
Sample(n)	$\hat{\beta}_1$	S.E	S.E	S.E	S.E	S.E	S.E	R^2
25	**3.1301**	**0.3179**	**0.3315**	**0.3865**	**0.4766**	**0.7321**	**0.3865**	**0.7971**
30	**2.9444**	**0.1734**	**0.1795**	**0.1829**	**0.1931**	**0.1902**	**0.1829**	**0.9020**
35	**3.0103**	**0.1080**	**0.1112**	**0.1121**	**0.1166**	**0.1148**	**0.1121**	**0.9309**
40	**3.0055**	**0.2008**	**0.2060**	**0.2103**	**0.2206**	**0.2230**	**0.2103**	**0.9309**
45	**2.9911**	**0.1673**	**0.1711**	**0.1748**	**0.1831**	**0.1942**	**0.1748**	**0.8266**
50	**2.9667**	**0.1551**	**0.1583**	**0.1598**	**0.1648**	**0.1635**	**0.1598**	**0.8443**

From Table 7, When g=1, at various sample sizes. HC2 and HC5 give the same estimates of $s.e(\beta_1)$, HC0 giving the minimum estimates. This may be as a result of under estimation. HC3 and HC4 follow in that order. HC4 gives the maximum $s.e(\beta_1)$

Table8: the estimates of $s.e(\beta_1)$ from the different Estimators and R^2 when, g=2 at varying sample sizes.

		Strength of heteroscedasticity=2						
		HC0	**HC1**	**HC2**	**HC3**	**HC4**	**HC5**	
Sample(n)	$\hat{\beta}_1$	S.E	S.E	S.E	S.E	S.E	S.E	R^2
25	**3.0469**	**0.3850**	**0.4014**	**0.4044**	**0.4250**	**0.4100**	**0.4250**	**0.8380**
30	**3.1386**	**0.4378**	**0.4532**	**0.4722**	**0.5098**	**0.5222**	**0.5098**	**0.6879**
35	**3.1334**	**0.2514**	**0.2589**	**0.2633**	**0.2760**	**0.2742**	**0.2760**	**0.6488**
40	**3.1458**	**0.3926**	**0.4028**	**0.4111**	**0.4308**	**0.4303**	**0.4308**	**0.7986**
45	**3.0108**	**0.3479**	**0.3559**	**0.3622**	**0.3773**	**0.3789**	**0.3773**	**0.6999**
50	**2.9341**	**1.0445**	**1.0661**	**1.1007**	**1.1617**	**1.2447**	**1.1617**	**0.6032**

From Table 8, When g=2, at various sample sizes. HC3 and HC5 give the same estimates of $s.e(\beta_1)$, With HC0 giving the minimum estimates. This may be as a result of under estimation. HC2 and HC4 follow in that order. HC4 gives the maximum $s.e(\beta_1)$

VI. Conclusion

The OLS estimators of model parameters remain unbiased when data is front with heteroscedasticity problem.

For negative strengths of heteroscedasticity HC5 is preferred as it has the minimum variances at g=-2, -1, -0.5 and 0 respectively.

HC5 performs equally as HC0 when g=0(homoscedasticity assumption holds).

HC5 performs equally as HC2 when g=1.

HC5 performs equally as HC3 when g=2.

HC5 performs almost equally as HC0 and HC1.When g=0.5 and HC1= (n/ (n-k))*HC5 when g=0.

Further, It is expected that HC5 will offer similar results to HC4 when g=4.

Also the graphs of $s.e(\beta_j)$ showed consistent patterns in the estimates offered by the existing family of HCSE and HC5 estimator (see appendix"A2" and"A3")

HC5 can be obtained for any strength "g" of heteroscedasticity that may be present in Simple Linear Regression model.

Finally HC5 is indeed a generalization of the existing family of HSCE.

References

[1]. EICKER, F. (1967). Limit theorems for regression with unequal and dependent errors. In L. M. Le Cam &J. Neyman (Eds.), Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Berkeley, CA: University of California Press.

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

[3]. HINKLEY, D. V.(1997). Jacknifing in unbalanced situations. Technometrics, 19, 285-292.

[4]. MACKINNON, J. G., &WHITE, G. P. (1985).Some hetreoskedasticity consistent covariance matrix estimators with improved finite sample properties. Journal of Econometrics, 29, 305-325.

[5]. DAVIDSON, R., &MACKINNON, J. G.(1993). Estimation and inference in econometrics. Oxford:Oxford University Press.

[6]. CRIBARI-NETO, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. Computational Statistics & Data Analysis, 45,215-233.

[7]. GOLDFELD, S. M., &QUANDT, R. E.(1965). Some tests for heteroskedsaticity. Journal of the American Statistical Association, 60, 539-547.

[8]. Johnston, J. Econometrics methods second edition, McGraw-Hill Book Company. New York (1972)

Appendix"A2"

Graphs of Standard Error of β_0 Against Strength of Heteroscedasticity

Figures1.0-1.5

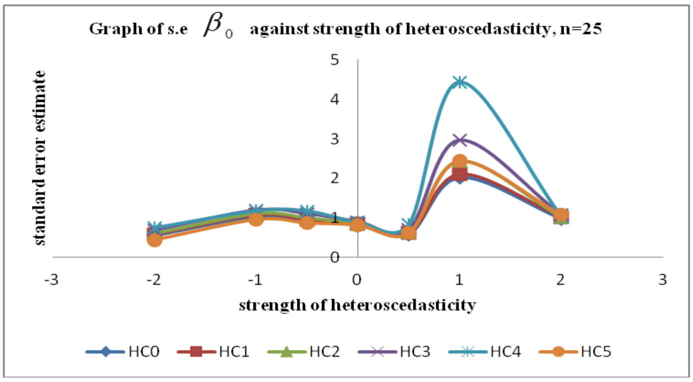

Figure1.0

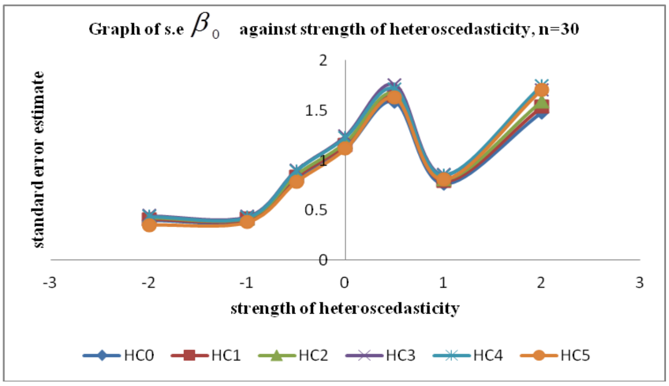

Figure 1.1

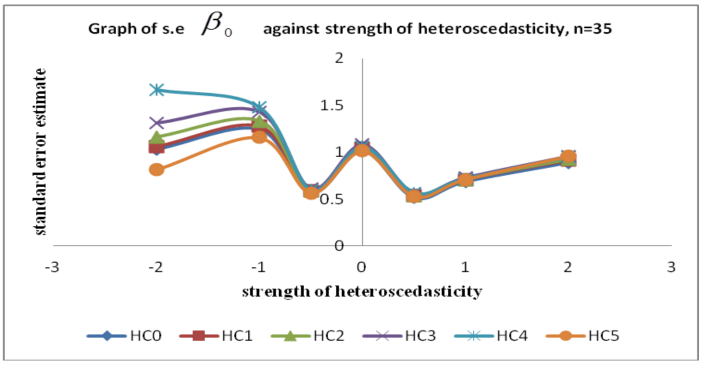

Figure1.2

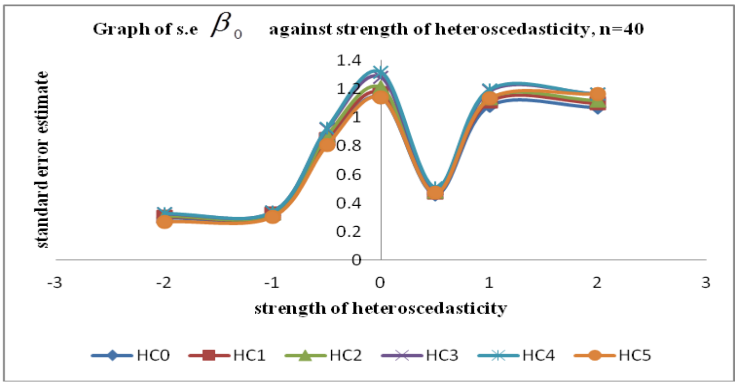

Figure 1.3

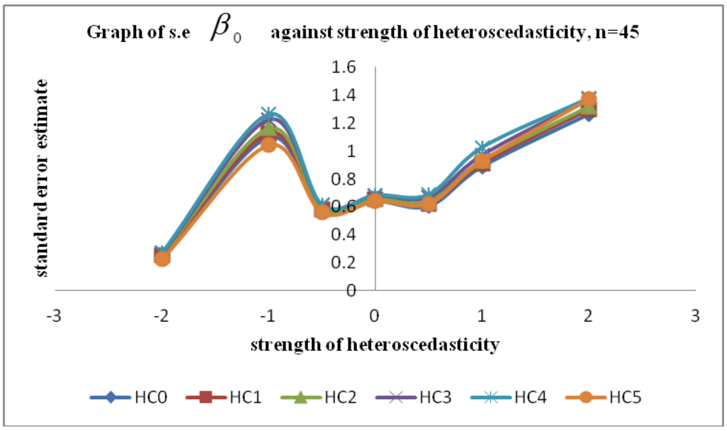

Figure 1.4

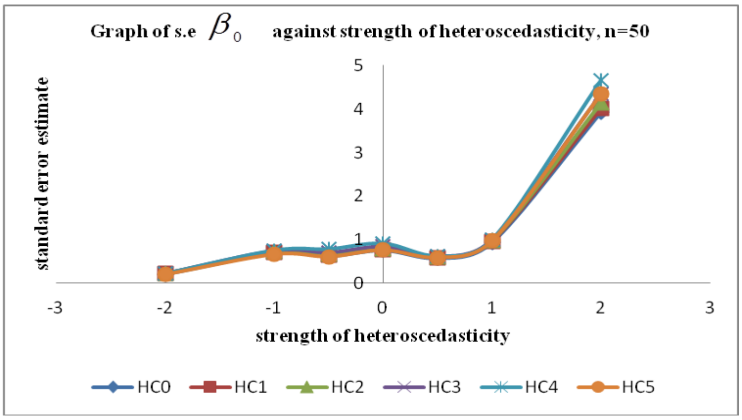

Figure 1.5

Appendix"A3"

Graphs of Standard Error of β_1 Against Strength of Heteroscedasticity

Figures 2.0-2.5

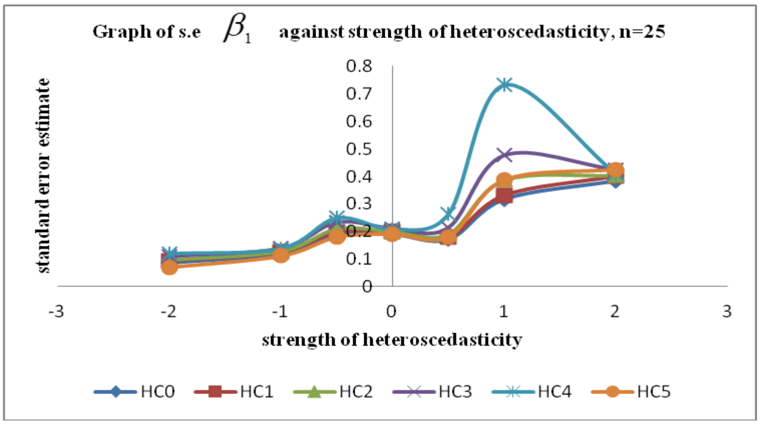

Figure 2.0

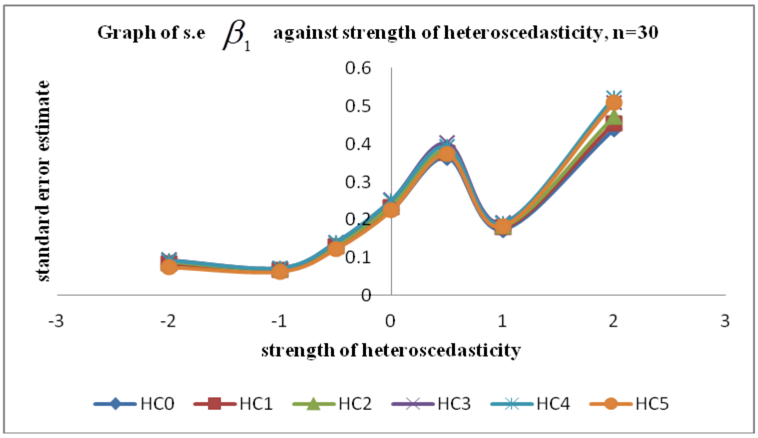

Figure 2.1

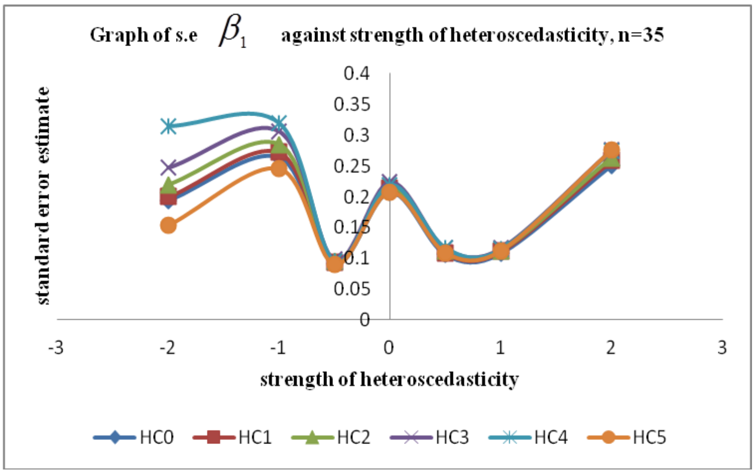

Figure 2.2

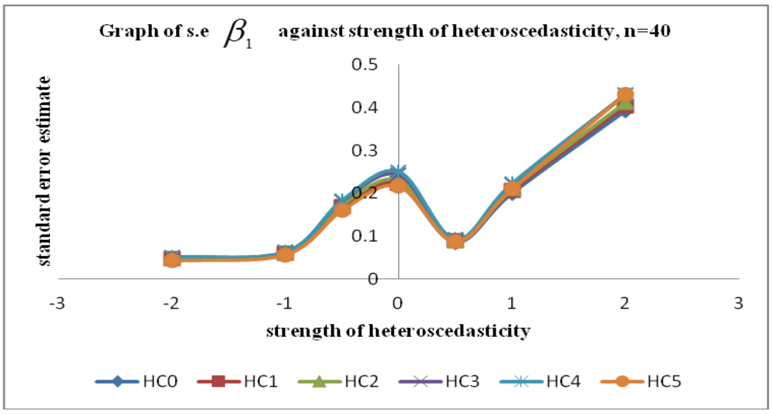

Figure 2.3

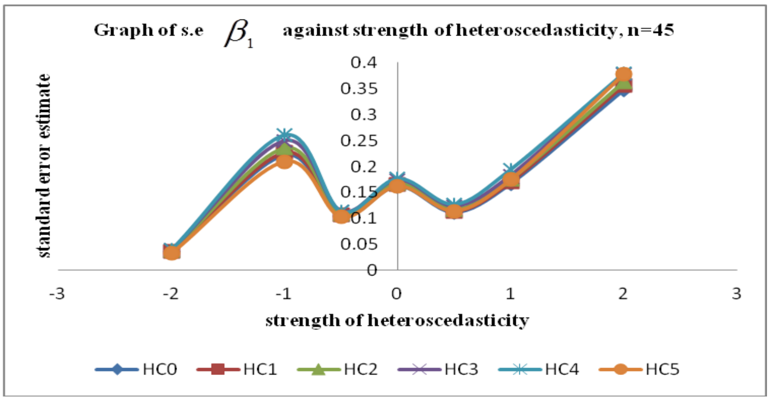

Figure 2.4

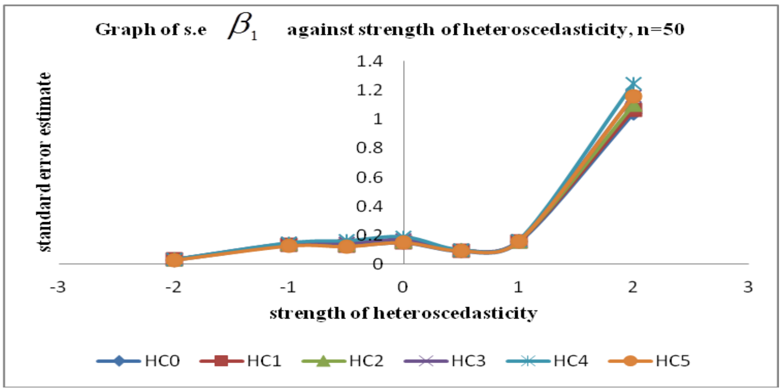

Figure 2.5

APPENDIX"A5"

Table9: Showing $\text{var}(\beta_0)$ for different strengths of heteroscedasticity and sample sizes using the existing HSCE and HC5 estimators.

		STRENGTH OF HETEROSKEDASTICITY						
Sample size	Estimator	-2	-1	-½	0	½	1	2
25	HC0	**0.5563211**	**1.035072**	**0.918468**	**0.821779**	**0.591899**	**2.02742**	**0.985563**
	HC1	**0.5800049**	**1.079137**	**0.957569**	**0.856764**	**0.617097**	**2.113732**	**1.027521**
	HC2	**0.6216265**	**1.113811**	**1.009143**	**0.862423**	**0.646476**	**2.435896**	**1.031884**
	HC3	**0.695725**	**1.19929**	**1.112407**	**0.907213**	**0.713384**	**2.969556**	**1.081199**
	HC4	**0.7528561**	**1.18237**	**1.178001**	**0.897411**	**0.840795**	**4.437786**	**1.046338**
	HC5	**0.4478809**	**0.962537**	**0.877364**	**0.821779**	**0.617813**	**2.435896**	**1.081199**
30	HC0	**0.3980349**	**0.399925**	**0.808479**	**1.117747**	**1.592767**	**0.765355**	**1.482506**
	HC1	**0.412049**	**0.413962**	**0.836856**	**1.156978**	**1.64867**	**0.792218**	**1.534539**
	HC2	**0.4218287**	**0.418871**	**0.853973**	**1.177594**	**1.67193**	**0.810058**	**1.590764**
	HC3	**0.4471221**	**0.439054**	**0.903011**	**1.242048**	**1.755833**	**0.85822**	**1.709419**
	HC4	**0.4391035**	**0.431005**	**0.898114**	**1.231055**	**1.715428**	**0.85314**	**1.744768**
	HC5	**0.3545886**	**0.382114**	**0.786951**	**1.117747**	**1.631774**	**0.810058**	**1.709419**
	HC0	**1.0312684**	**1.240713**	**0.572673**	**1.014122**	**0.518898**	**0.688361**	**0.890327**

35	HC1	**1.0620594**	**1.277758**	**0.589771**	**1.044401**	**0.534391**	**0.708913**	**0.916909**
	HC2	**1.162775**	**1.331594**	**0.592426**	**1.047607**	**0.541842**	**0.709732**	**0.923821**
	HC3	**1.3127832**	**1.429958**	**0.612981**	**1.082529**	**0.566577**	**0.732105**	**0.959496**
	HC4	**1.6643906**	**1.477448**	**0.597427**	**1.056739**	**0.566996**	**0.716518**	**0.948983**
	HC5	**0.815022**	**1.156745**	**0.563086**	**1.014122**	**0.530156**	**0.709732**	**0.959496**
40	HC0	**0.2968187**	**0.316718**	**0.829541**	**1.141438**	**0.463849**	**1.083649**	**1.072228**
	HC1	**0.3045295**	**0.324946**	**0.851091**	**1.171091**	**0.475899**	**1.1118**	**1.100082**
	HC2	**0.311022**	**0.328728**	**0.873799**	**1.20906**	**0.482102**	**1.132815**	**1.117641**
	HC3	**0.3260455**	**0.341388**	**0.9207586**	**1.281772**	**0.501762**	**1.185506**	**1.165778**
	HC4	**0.3255925**	**0.338015**	**0.9267027**	**1.31811**	**0.508691**	**1.196716**	**1.162757**
	HC5	**0.270693**	**0.30532**	**0.8083742**	**1.141438**	**0.47281**	**1.132815**	**1.165778**
45	HC0	**0.247479**	**1.099726**	**0.570311**	**0.641874**	**0.60727**	**0.891327**	**1.266627**
	HC1	**0.2531689**	**1.12501**	**0.583424**	**0.656632**	**0.621232**	**0.91182**	**1.295749**
	HC2	**0.2582951**	**1.160713**	**0.592063**	**0.661466**	**0.638232**	**0.930093**	**1.317767**
	HC3	**0.2697435**	**1.225834**	**0.614939**	**0.682256**	**0.671297**	**0.972726**	**1.371859**
	HC4	**0.2716967**	**1.261305**	**0.612799**	**0.683161**	**0.688452**	**1.02579**	**1.377201**
	HC5	**0.2275922**	**1.042554**	**0.559839**	**0.641874**	**0.622497**	**0.930093**	**1.371859**
50	HC0	**0.2139438**	**0.692306**	**0.629736**	**0.762294**	**0.5807763**	**0.946478**	**3.9279155**
	HC1	**0.2183555**	**0.706582**	**0.642722**	**0.778013**	**0.5927524**	**0.965995**	**4.0089126**
	HC2	**0.2203181**	**0.721148**	**0.668647**	**0.803041**	**0.601495**	**0.977394**	**4.1334864**
	HC3	**0.2269386**	**0.751454**	**0.711214**	**0.847922**	**0.6232285**	**1.009593**	**4.3570139**
	HC4	**0.2245636**	**0.757036**	**0.786666**	**0.914542**	**0.62359127**	**1.002021**	**4.6650113**
	HC5	**0.2018907**	**0.664854**	**0.611554**	**0.762294**	**0.59101211**	**0.977394**	**4.3570139**

Appendix"A6"

Table9: Showing $\mathrm{var}(\beta_1)$ For Different Strengths of Heteroscedasticity And Sample Sizes Using The Existing Hsce And Hc5 Estimators.

		STRENGTH OF HETEROSKEDASTICITY						
Sample size	Estimator	-2	-1	-½	0	½	1	2
25	HC0	**0.0880422**	**0.120685**	**0.189992**	**0.191438**	**0.172678**	**0.31792**	**0.384985**
	HC1	**0.0917903**	**0.125822**	**0.19808**	**0.199587**	**0.180029**	**0.331455**	**0.401375**
	HC2	**0.09834461**	**0.12988**	**0.210592**	**0.201219**	**0.191432**	**0.386526**	**0.404424**
	HC3	**0.1100425**	**0.139857**	**0.233942**	**0.211869**	**0.214537**	**0.476623**	**0.425018**
	HC4	**0.1191319**	**0.137824**	**0.249211**	**0.208093**	**0.262712**	**0.732112**	**0.410023**
	HC5	**0.0709538**	**0.112207**	**0.180621**	**0.191438**	**0.181566**	**0.386526**	**0.425018**
30	HC0	**0.0834623**	**0.065911**	**0.125399**	**0.224585**	**0.363001**	**0.173395**	**0.437835**
	HC1	**0.0863917**	**0.068224**	**0.1298**	**0.232467**	**0.375742**	**0.179481**	**0.453202**
	HC2	**0.08853937**	**0.069444**	**0.132729**	**0.238036**	**0.382309**	**0.182921**	**0.472175**
	HC3	**0.0939435**	**0.073256**	**0.140637**	**0.252571**	**0.402745**	**0.193098**	**0.509755**
	HC4	**0.09241004**	**0.072852**	**0.140196**	**0.251993**	**0.393385**	**0.190222**	**0.522233**
	HC5	**0.0742088**	**0.062628**	**0.121932**	**0.224585**	**0.372518**	**0.182921**	**0.509755**
35	HC0	**0.1942126**	**0.264326**	**0.091224**	**0.207746**	**0.106169**	**0.107974**	**0.251405**
	HC1	**0.2000113**	**0.272218**	**0.093948**	**0.213949**	**0.109339**	**0.111198**	**0.258911**
	HC2	**0.2190869**	**0.284602**	**0.094553**	**0.215829**	**0.111408**	**0.112147**	**0.263312**
	HC3	**0.2474589**	**0.306582**	**0.098021**	**0.224297**	**0.117046**	**0.116559**	**0.276004**
	HC4	**0.3140991**	**0.319054**	**0.095563**	**0.219428**	**0.11735**	**0.114756**	**0.274242**
	HC5	**0.1533048**	**0.245616**	**0.08961**	**0.207746**	**0.108741**	**0.112147**	**0.276004**
40	HC0	**0.0473215**	**0.059425**	**0.164721**	**0.217342**	**0.086603**	**0.200768**	**0.392581**
	HC1	**0.0485509**	**0.060968**	**0.169000**	**0.222988**	**0.088852**	**0.205983**	**0.40278**
	HC2	**0.0496761**	**0.061852**	**0.173440**	**0.230111**	**0.089816**	**0.210338**	**0.411137**
	HC3	**0.0521664**	**0.064413**	**0.182698**	**0.243835**	**0.093236**	**0.220584**	**0.430764**
	HC4	**0.0521792**	**0.063909**	**0.183976**	**0.250514**	**0.093372**	**0.22299**	**0.430323**
	HC5	**0.0429512**	**0.057123**	**0.160552**	**0.217342**	**0.088185**	**0.210338**	**0.430764**
45	HC0	**0.0361061**	**0.221871**	**0.104899**	**0.162276**	**0.111582**	**0.167262**	**0.347905**
	HC1	**0.0369362**	**0.226972**	**0.107311**	**0.166007**	**0.114148**	**0.171107**	**0.355904**
	HC2	**0.0377308**	**0.235105**	**0.10923**	**0.167996**	**0.117381**	**0.1747797**	**0.362205**
	HC3	**0.0394531**	**0.24935**	**0.113778**	**0.17413**	**0.123582**	**0.183069**	**0.377324**
	HC4	**0.0398531**	**0.260214**	**0.113486**	**0.176567**	**0.127133**	**0.194167**	**0.378854**
	HC5	**0.0331252**	**0.209554**	**0.102812**	**0.162276**	**0.114433**	**0.1747787**	**0.377324**
50	HC0	**0.0317609**	**0.13325**	**0.127821**	**0.150327**	**0.09856**	**0.155084**	**1.044531**
	HC1	**0.0324158**	**0.135998**	**0.130457**	**0.153427**	**0.0949714**	**0.158282**	**1.0660699**
	HC2	**0.0327446**	**0.139193**	**0.136137**	**0.159928**	**0.0962898**	**0.159837**	**1.1007125**
	HC3	**0.0337681**	**0.14545**	**0.145227**	**0.170556**	**0.0996829**	**0.16479**	**1.1616914**
	HC4	**0.0334787**	**0.147303**	**0.161776**	**0.189279**	**0.0996062**	**0.163468**	**1.2446819**
	HC5	**0.0299055**	**0.127604**	**0.123933**	**0.150327**	**0.0946522**	**0.159837**	**1.1616914**