Monte Carlo Project Report

1.Introduction

The following report is based on studying the consequences of applying the Ordinary Least Square Estimators (OLS) on the linear model given below:

$$Yi = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i \sqrt{x_{2i}} \qquad (1)$$

i=1, 2,.N

where ei ∼ i.i.d.N (0, σ2). $x_{1i} = x_{2i} + u_i$ where u_i ∼ i.i.d.N (0, 1) and x_{2i} is following a chi squared distribution with 2 degrees of freedom.

To analyse the consequences of applying OLS to the linear model the following points are considered:

- Unbiasedness and Consistency: OLS estimators are unbiased if their expected values are equal to the true population parameters. OLS estimators are consistent if they in converge in probability to the true parameter values as N (sample size) approaches infinity. Assuming the classical linear regression assumptions hold ^β0, ^β1, ^β2 hat are unbiased and consistent estimators of β0, β1 &β2 respectively.
- ^β0, ^β1 and ^β2 is normally distributed regardless of the same size. When ei (the error term) is independently and identically distributed(i.i.d.) normal random variables with constant variance (*insert symbol*), it is true that the OLS estimators β0, β1 and β2 are normally distributed regardless of the sample size under the linear regression theory.
- The variance estimators for √N (ˆβ1−β1),which ignores or take account of heteroskedasticity are asymptotically the same, i.e. in large samples, whether you consider or ignore

heteroskedasticity, the variance estimators for this scaled difference converge to the same limit.-slope

- The variance estimators for $\sqrt{N}(\hat{\beta}_0-\beta_0)$,which ignores or take account of heteroskedasticity are asymptotically different ,i.e. ignoring heteroskedasticity the variance estimators may behave differently compared to when it is considered. -intercept
- The variance estimators for $\sqrt{N}(\hat{\beta}_2-\beta_2)$, which ignores or take account of heteroskedasticity are asymptotically the same-slope.
- The variance estimators, which take account of heteroskedasticity are consistent for the variance of the OLS estimators while the ones that ignore heteroskedasticity are not always consistent.

Section 2 presents the theoretical findings, whereas Section 3 gives the Monte Carlo evidence before the conclusion in section 4.

<u>2.Theoretical Results Related to the Econometric Model</u>

The Gauss Markov Assumptions (1-6) are checked and made sure the following assumptions are satisfied, such as the model being a linear regression model (multiple)and then with Zero Conditional Mean i.e. the Expected value($\varepsilon i/X$) =0, where xi= (1, x1i, x2i, xki), β=(β0, β1,.. βk) and X is the collection of all xis, which implies E(εi) =0 and Cov (εI, xki)=E(εi)E(Xki)=0 for k=1,2,…K.

Next, there is no perfect collinearity between the independent variables i.e. independent or explanatory variable are not perfectly correlated which avoids problems such as estimation of the regression coefficients. The error term in the model $\varepsilon i = ei\sqrt{x2i}$, which is independent and should not provide information about the error term of any other observation. Moreover, normality of errors of terms is also assumed. However, only the assumption about homoscedasticity is not because Var(ei|xi) = σ^2xi, which changes over i, heteroskedasticity exists.

Theorem 1: The OLS estimators for the slope and intercept are defined respectively below:

Theorm 1

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{N}(x_{1i}-\widehat{x}_1)(y_i - Y)}{\sum_{i=1}^{N}(x_1-\widehat{x}_1)^2}$$

$$\widehat{\beta}_2 = \frac{\sum_{i=1}^{N}(x_{1i}-\widehat{x}_2)(y_i - Y)}{\sum_{i=1}^{N}(x_{1i}-\widehat{x}_1)^2}$$

$$\widehat{\beta}_0 = \widehat{Y} - \hat{\beta}_1\widehat{x}_1 - \widehat{\beta}_2\widehat{x}_2$$

Where x bar=summation xi/N and y bar=summation Yi/N. Then E(^β1)=β1 , E(^β0)= β0, E(^β2)= β2,

plim N→∞ βˆ 0 = β0.

plim N→∞ βˆ 1 = β1.

plim N→∞ βˆ 2 = β2.

The theorem above shows the estimators for β0 which is the intercept and β1 and β2 which are the slopes.

β0 is the estimated intercept which is the expected value of y when both x1i and x2i are zero. It is unbiased and consistent under classical linear regression assumptions. Y bar is the mean of the dependent variable y and is average value of the observed outcomes. β1xbar1 is the estimated contribution of the first explanatory variable(x1i) to the intercept. It is the product of the estimated slope (β1^) for x1i and the mean of x1i. β2^X2bar is the estimated contribution of the second explanatory variable x2i to the intercept. β2 is the OLS estimate for the slope associated with x2i, x2bar is the same mean of x2i. it represents the expected value of y when x1i&x2i are at their average values. The equation above (for the intercept is derived from a first order condition where we know all the slopes (1&2).

NOTE: Y= (Y BAR) in the theorem above.

In the formula above $\beta 1$ shows how much y changes when x1i changes by one unit while keeping all other variables constant. This also unbiased and consistent under the classical linear regression assumptions.

$\beta 2$ formula given above represents the change in y for a one-unit change in x2i, holding other variables constant. Unbiased and consistent under the classical linear regression assumptions.

The next theorem illustrates that the OLS estimators conditional on the explanatory variable are therefore normally distributed since ei~i.i.d. N(0, σ2).

THEOREM 2: Define X to be the collection of xi with i=1, 2,…,N, then

$$\epsilon_i = e_i \sqrt{x_2} \mid x_1 \,,\; x_2 \sim i,i,d. N(0, \sigma^2)$$

$$\hat{\beta}_1 \mid X \sim N(\beta_1 \,,\; \sigma^2 \frac{\sum_{i=1}^{N} (x_1 - \overline{x}_1)^2 x_1}{[\sum_{i=1}^{N} (x_1 - \overline{x}_2)^2]^2})$$

$$\hat{\beta}_2 \mid X \sim N(\beta_2 \,,\; \sigma^2 \frac{\sum_{i=1}^{N} (x_2 - \overline{x}_2)^2 x_2}{[\sum_{i=1}^{N} (x_2 - \overline{x}_2)^2]^2})$$

$$\hat{\beta}_0 \mid X \sim N(\beta_0 \,,\; \sigma^2 \sum_{i=1}^{N} [\frac{1}{N} - \frac{\overline{x}_1 (x_1 - \overline{x}_1)}{\sum_{i=1}^{N} (x_1 - \overline{x}_1)^2}] x_1 + \sum_{i=1}^{N} [\frac{1}{N} - \frac{\overline{x}_2 (x_2 - \overline{x}_2)}{\sum_{i=1}^{N} (x_2 - \overline{x}_2)^2}] x_2)$$

The above results should hold regardless of the sample size N. 'CLT is a statistical premise that, given a sufficiently large sample size from a population with a finite level of variance, the mean of all sampled variables from the same population will be approximately equal to the mean of the whole population. Furthermore, these samples approximate a normal distribution, with their variances being approximately equal to the variance of the population as the sample size gets larger, according to the law of large numbers.' (Ganti 2023)

Theorem 3:

$$plimN-\infty\ NVar(\overline{\beta}_1|X)=\sigma^2\frac{plim\ \frac{1}{N}\sum_{i=1}^{N}(x_i-\overline{x})^2x_i}{[plim\frac{1}{N}\sum_{i=1}^{N}(x_1-\overline{x})^2]^2}\ =\sigma^2\frac{E[(x_1-E(x_1))^2x_1]}{[E(x_1-E(x_1))^2]^2}$$

$$=\sigma^2\frac{E(x_13)-2\ E(x_12)\ E(x_1)+E(x_1)^3}{var(x_2)^2}$$

$$=\frac{\sigma^2\ \sigma-2(4).2+2^3}{-1}$$

$$=2\sigma^2$$

$$plimN-\infty\ NVar(\hat{\beta}_2|\ X)=\sigma^2\frac{plim\frac{1}{N}\sum_{i=1}^{N}(x_2-\overline{x}_2)^2x_i}{[plim\frac{1}{N}\sum_{i=1}^{N}(x_2-\overline{x}_2)^2)^2}\ =\sigma^2\frac{E[(x_2-E(x_2))^2]}{[E(x_2-E(x_2))^2]^2}$$

$$=\sigma^2\frac{\mathrm{E}(\mathrm{x}_2 3)-2\ \mathrm{E}(x_2 2)\mathrm{E}(x_2)+\mathrm{E}(x_2)^3}{Var(x_2)^2}$$

$$=\sigma^2\frac{\sigma-2(4).2+2^3}{-1}$$

$$=2\sigma^2$$

$$plim N-\infty\ Nvar(\hat{\beta}_0|X)=\sum_{i=1}^{N}[\frac{1}{N}-\frac{\bar{x}_1(x_1-\bar{x}_1)}{\sum_{i=1}^{N}(x_1-\bar{x}_1)^2}]^2\ x_1+\sum_{i=1}^{N}[\frac{1}{N}-\frac{\bar{x}_2(x_2-\bar{x}_2)}{\sum_{i=1}^{N}(x_2-\bar{x}_2)^2}]^2\ x_2$$

$$=\ \sigma^2E[(1-\frac{E(x_1)(x_1-E(x_1))}{E(x_1-E(x_1))^2})^2]\ x_1+\sigma^2E[(1-\frac{E(x_2)(x_2-E(x_2))}{E(x_2-E(x_2))^2})^2]\ x_2$$

$$=\frac{\sigma^2E(x_13)-2E(x_12)\ E(x_1)+E(x_1)^3}{Var(x_1)^2}+\frac{\sigma^2\ E(x_23)-2E(x_22)E(x_2)+E(x_2)^3}{Var(x_2)^2}$$

$$=\frac{\sigma^2\ \sigma-2(4).2+2^3}{-1}+\sigma^2\frac{\sigma-2(4).2+2^3}{-1}$$

$$=2\sigma^2+2\sigma^2$$

$$=4\sigma^2$$

Theorem 3 shows that in this model the estimated standard errors for the OLS slope estimates ($\beta1$&$\beta2$) which ignore or account for heteroskedasticity are asymptotically same, although the two types of estimated standard errors for the OLS intercept ($\beta0$) estimates are asymptotically not same.

3.Monte Carlo evidence

To check the statements made earlier, Monte Carlo simulations are used in this part. The results are based on 10000 simulations, and the true values are set to $\beta0=1$, $\beta1=2$, $\beta2=3$ and variance, $\sigma^2=1$. The data was generated according to the topic discussed in Section 1. The OLS estimates and their estimated standard errors are then saved in each simulation. Table 1,2 and 3 respectively show the summary statistics for ^ $\beta0$, ^ $\beta1$, ^ $\beta2$ which are in line with the theoretical results in Theorem 1 and 2. ^ $\beta0$, ^ $\beta1$, ^ $\beta2$ is all unbiased and the sample means from simulations are very close to the true values regardless of the sample size N. Next, the estimated standard deviations from simulations are considered which keep going to 0 as N increases. The ranges of the distributions (max-min) also decrease with N which indicates consistency.

Next, the JB probability value is considered-the null hypothesis is set that the betas are normally distributed for different sample sizes. The level is significance is set to 5% and the null hypothesis is rejected if the JB p value is less than 5% and accepted if the JB p value is more than 5%. For beta 0 it is therefore rejected for sample size 30. Beta 1 it is accepted for all sample sizes. For beta 2 it is also again accepted for all sample sizes.

*Table 1: Summary Statistics for(β*0) from simulations

N(β0)	Mean	Median	Max	Min	Std Dev	J.B Value	P-Value
5	0.999652	0.997230	3.633948	-2.126818	0.643861	0.220872	0.895444
30	0.997847	0.995121	2.211812	-0.260106	0.318952	8.540651	0.013977
100	1.001096	1.003967	1.738421	0.208886	0.193727	3.903145	0.142050
500	0.99980	0.999397	1.316292	0.648146	0.089351	1.439424	0.486892

*Table 2: Summary Statistics for(β*1) from simulations

N (β1)	Mean	Median	Max	Min	Std Dev	J.B Value	P-Value
5	1.994146	1.993820	3.776800	0.295258	0.422802	3.111146	0.211068
30	2.001676	2.003354	3.006678	1.120417	0.243253	1.364747	0.505416
100	2.001272	2.001860	2.482698	1.515027	0.130057	0.663462	0.717680
500	1.998406	1.999037	2.209290	1.749954	0.063327	5.453959	0.065417

Table 3: Summary Statistics for(β2) from simulations

N(β2)	Mean	Median	Max	Min	Std Dev	J.B Value	P-Value
5	3.009621	3.014393	5.862988	0.612701	0.690080	2.511062	0.284925
30	2.998123	3.000809	4.103865	1.999492	0.290195	0.962298	0.618073
100	2.997102	2.994799	3.729263	2.270316	0.193121	0.160021	0.923107
500	3.001062	3.000677	3.347382	2.713885	0.079780	1..375002	0.502831

As per theorem 3, consistency means that as the sample size N increases, the standard error estimators converge to the true standard deviation of $\hat{\beta}_0$ given the values of the explanatory variables (X). The robust standard error estimator consistently provides an accurate estimate of the standard deviation, even as the sample size grows.

For ^ $\beta 0$, only the robust standard error estimator will be consistent for NVar d($\hat{\beta}$ 0|X) = while the standard error estimators, which ignore and account for heteroskedasticity are both consistent, which can be seen in tables 4,5&6.The first row in each table represents the true standard deviations calculated based on the theoretical variance formulas. The rows following show the sample means of standard error estimates obtained through simulations for different sample sizes (N).

Hence it is feasible to say that for smaller sample sizes, none of the standard error estimates are close to the standard deviations ^ $\beta 0$. But, as the sample size increases, the sample means of both standard error estimates get closer to the true values. Therefore, we can conclude that only the sample means of the robust standard error estimates are consistently close to the true values even for large samples i.e. Even in the face of heteroskedasticity, robust standard error estimates for ^ $\beta 0$ show consistency and provide trustworthy estimates. In example, with smaller sample sizes, non-robust standard error estimations may not reliably reflect the true standard deviation of ^ $\beta 0$, regardless of whether they account for or ignore heteroskedasticity.

Table 4: The Standard Deviation of ^ $\beta 0$ *and the means of Estimates*

N(β0)	True Std Dev	Non-Robust	Robust
5	0.638601	0.886671	0.489902
30	0.319414	0.353401	0.271145
100	0.192536	0.202557	0.181025
500	0.088962	0.093598	0.087675

Table 5: The Standard Deviation of ^ $\beta 1$ *and the means of Estimates*

N(β1)	True Std Dev	Non-Robust	Robust
5	0.419373	0.718345	0.315074
30	0.244529	0.230918	0.225568
100	0.131020	0.124851	0.127130
500	0.064015	0.064313	0.063537

Table 6: The Standard Deviation of ^ β2 and the means of Estimates

N(β2)	True Std Dev	Non-Robust	Robust
5	0.682943	0.880979	0.475044
30	0.289742	0.259560	0.270766
100	0.193509	0.149047	0.178401
500	0.080095	0.071259	0.079069

Table 7: NVar(ˆβ0|X) and the Means of the Estimates

N(β0)	5	30	100	500
True Std Dev	2.039055	3.06075	3.707	3.957
Non-Robust	5.06124	3.86088	4.1413	4.389
Robust	1.606695	2.34912	3.4097	3.886

Table 8: NVar(ˆβ1|X) and the Means of the Estimates

N(β1)	5	30	100	500
True Std Dev	0.87937	1.79385	1.7166	2.049
Non-Robust	3.321995	1.64841	1.5734	2.072
Robust	0.71366	1.6032	1.6513	2.0295

Table 9: NVar(ˆβ2|X) and the Means of the Estimates

N(β2)	5	30	100	500
True Std Dev	2.33206	2.51853	3.7446	3.2075
Non-Robust	4.99647	2.08269	2.2423	2.544
Robust	1.534145	2.33388	3.3637	3.1565

The first rows in Tables 7,8&9 and are obtained by squaring the standard deviations (first rows in tables 4,5, &6) and then multiplying them by N. They represent the variance of the OLS estimators for ^ β0, ^ β1, ^ β2 based on standard error estimates obtained from EViews.

With the variance being set to 1 and the theorem 3 where the plim N→∞ the results for robust and non-robust for$\beta 0$, $\beta 1$ and $\beta 2$ all converges to $\beta 0=4\ \sigma 2$, $\beta 1=2\ \sigma 2$ and $\beta 2=2\ \sigma 2$. As the sample size increases the standard error tends to decrease. According to the theorem 3, when N approaches infinity the estimators converge to certain values. These values match the values i.e. The ^ $\beta 0$, ^ $\beta 1$, ^ $\beta 2$ values are consistent with 4 $\sigma 2$, 2 $\sigma 2$ and 2 $\sigma 2$respectively, which states that all our estimators are consistent.

Since to construct the t test statistics involves using the standard error estimates, the t test statistic for $\beta 0$ will not be valid unless the robust standard errors are used. Table 10 is obtained by setting the level of significance equal to 5%. We can see that the probability of making Type I errors (the size of the test) is close to 5% for large N under the robust standard errors, while the non-robust standard is a little below the level of significance .In table 10, however, the two standard error estimates both yield acceptable test sizes close to the level of significance for Large N. Th results can be confirmed using the curves.

The test statistics using robust standard errors are more powerful than those using non robust standard errors for different values of N.

TABLE 10: Sizes Of T tests with different standard error Estimates for ˆβ0

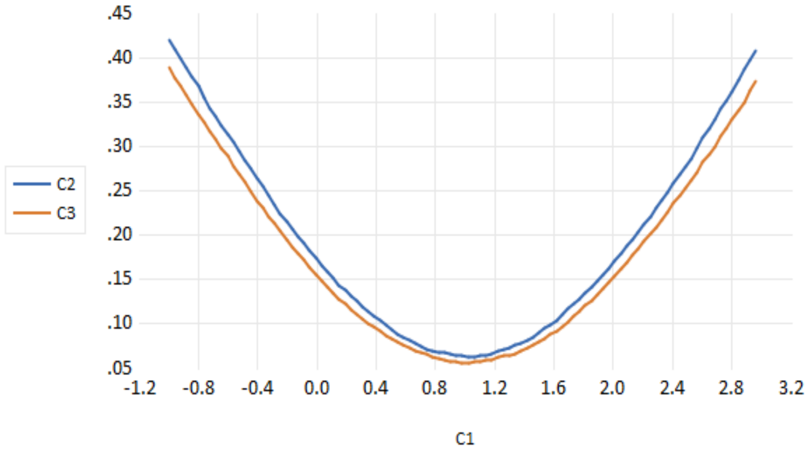
B0	5	30	100	500
Non-Robust	0.017500	0.023100	0.035000	0.040400
Robust	0.061300	0.075200	0.072000	0.057600

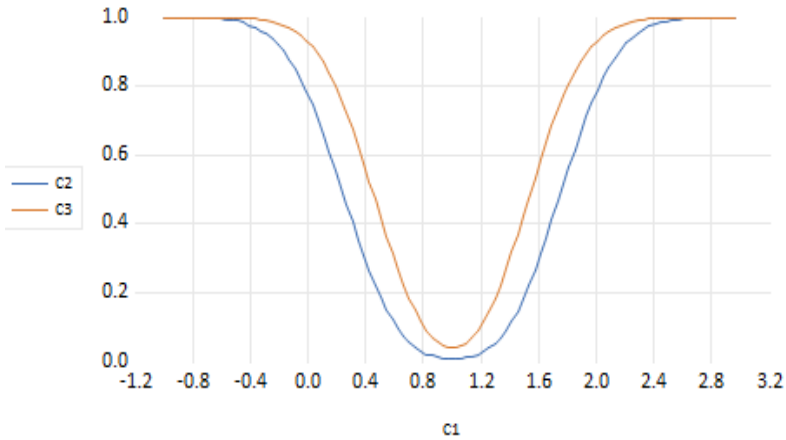
TABLE 11: Sizes Of T tests with different standard error Estimates for ˆβ1

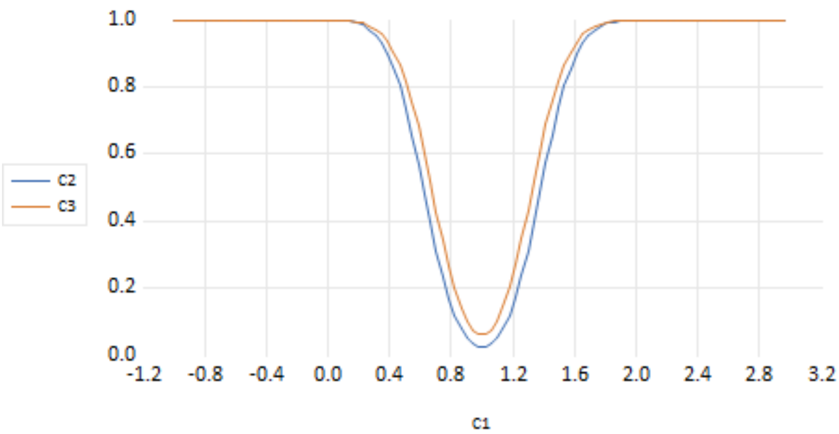
β1	5	30	100	500
Non-Robust	0.013100	0.057500	0.058000	0.044900
Robust	0.093600	0.067100	0.053100	0.048100

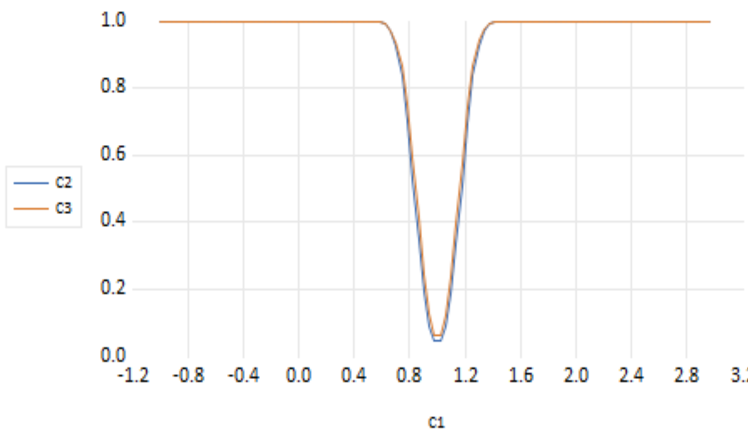
TABLE 12: Sizes Of T tests with different standard error Estimates for ˆβ2

β2	5	30	100	500
Non-Robust	0	0	0	0
Robust	0	0	0	0


C2
C3
.45
.40
.35
.30
.25
.20
.15
.10
.05
-1.2
-0.8
-0.4
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
C1

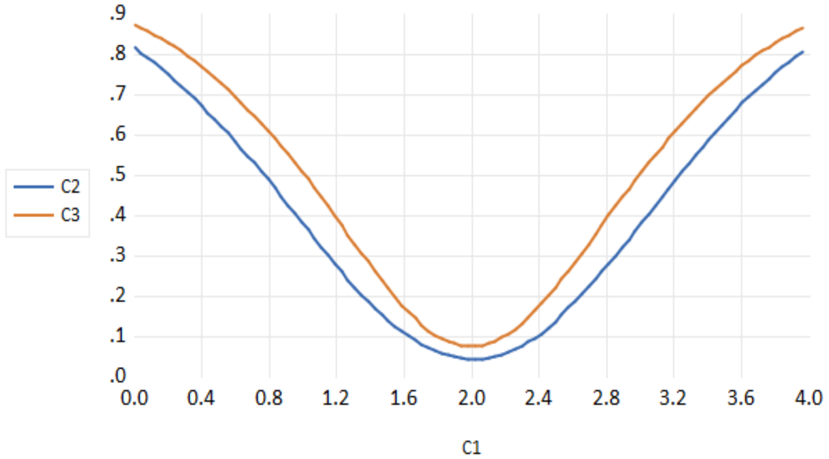

C2
C3
1.0
0.8
0.6
0.4
0.2
0.0
-1.2
-0.8
-0.4
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
C1

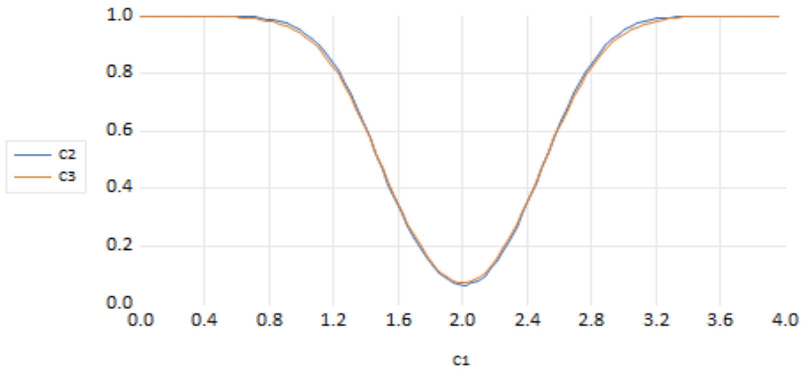


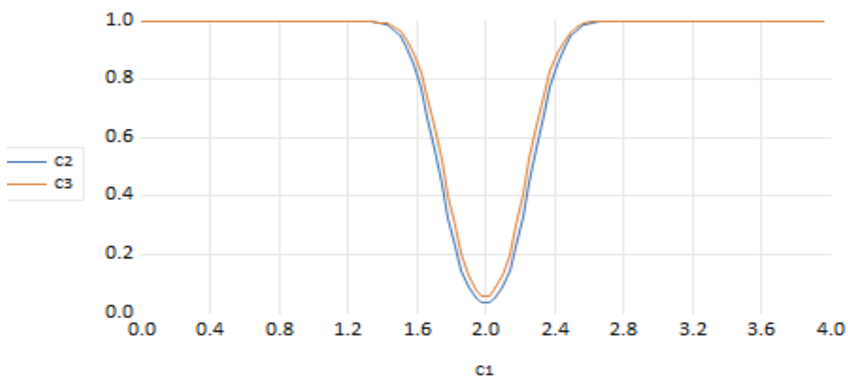


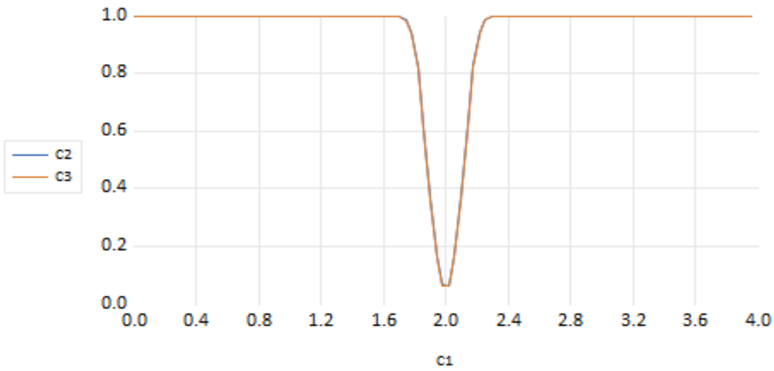
<u>THE CURVE ABOVES ARE POWER CURVESS FOR T TEST FOR θ0 for all the sample sizes</u>

<u>(n=5,30,100,500)</u>


.9
.8
.7
.6
.5
.4
.3
.2
.1
.0
C2
C3
0.0 0.4 0.8 1.2 1.6 2.0 2.4 2.8 3.2 3.6 4.0
C1

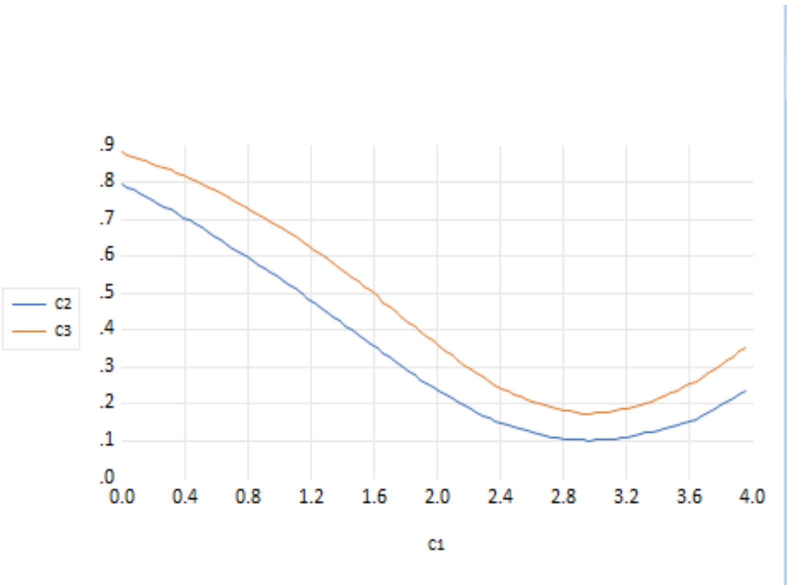

1.0
0.8
0.6
0.4
0.2
0.0
C2
C3
0.0 0.4 0.8 1.2 1.6 2.0 2.4 2.8 3.2 3.6 4.0
C1

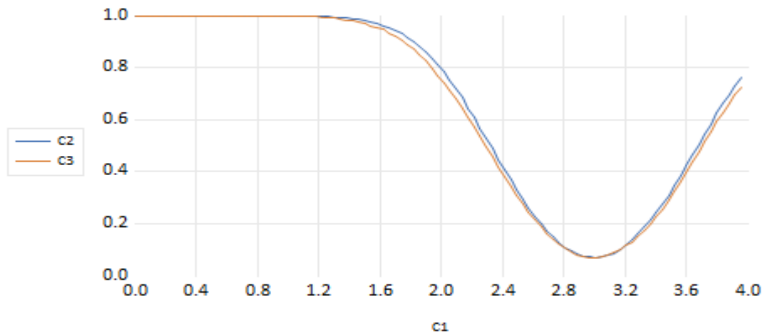

1.0
0.8
0.6
0.4
0.2
0.0
C2
C3
0.0 0.4 0.8 1.2 1.6 2.0 2.4 2.8 3.2 3.6 4.0
C1

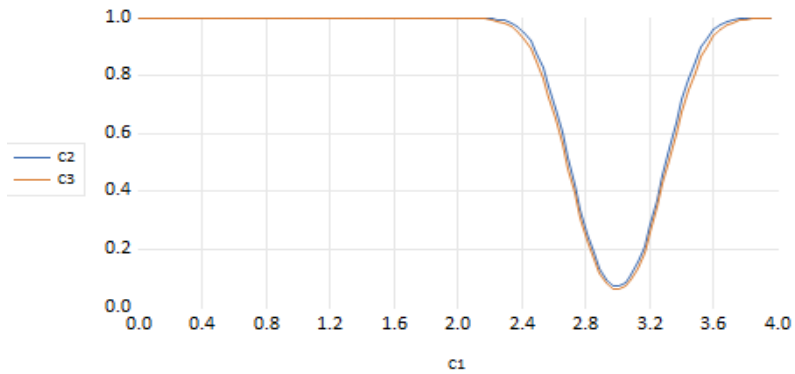


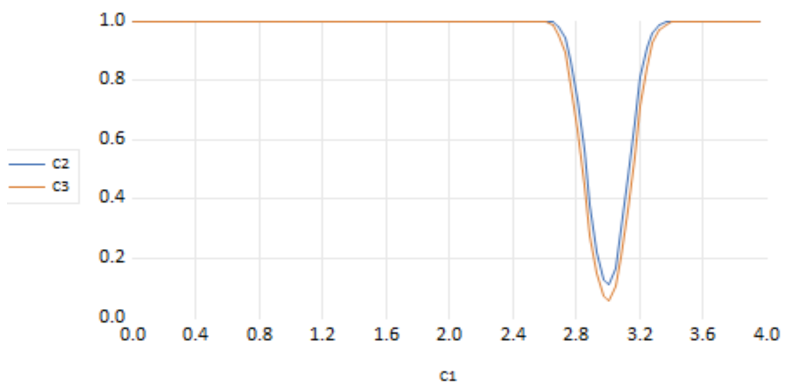
THE CURVE ABOVE ARE POWER CURVES FOR T TEST FOR β1 for all the sample sizes

(n=5,30,100,500)




c2
c3
c1


c2
c3
c1


c2
c3
c1

THE CURVE ABOVE ARE POWER CURVES FOR T TEST FOR β2 for all the sample sizes

(n=5,30,100,500)

4.Conculsion

To conclude we can say that we have studied the properties of the OLS estimators under heteroskedasticity, are unbiased, consistent, and normally distributed. We analysed the consequences of applying the OLS estimator to the multiple linear regression model. We can say that the estimated standard errors of both types whether taking into account of heteroskedasticity or not for the slope are consistent but only consistent for the intercept when the sample size is large.

APPENDIX

A.1 PROOF FOR THEOREM 1

$$plim-\infty\hat{\beta}_1=\beta_1+\frac{(\frac{1}{n}\sum_{i=1}^{N}(x_1-\overline{x}_1)\in_i)}{(\frac{1}{n}\sum_{i=1}^{N}(x_1-\overline{x}_1)^2)}$$

$$plim-\infty\widehat{\beta}_2=\beta_2+\frac{(\frac{1}{n}\sum_{i=1}^{n}(x_2-\overline{x}_2)\in_i)}{(\frac{1}{n}\sum_{i=1}^{n}(x_2-\overline{x}_2)^2)}$$

$$\mathrm{plim}-\infty\widehat{\beta}_0=\beta_0+\overline{x}_1\beta_1+\overline{x}_2+\beta_2-\overline{\in}-\overline{x_1}\hat{\beta}_1-\overline{x}_2\hat{\beta}_2$$

$$\hat{\beta}_0=\beta_0+\sum_{\mathrm{i}=1}^{\mathrm{N}}[\frac{1}{\mathrm{N}}-\frac{\overline{x}_{1\,(x_1-\overline{x}_1)}}{\sum_{\mathrm{i}=1}^{\mathrm{N}}(x_1-\overline{x}_1)^2}]\in_1+\sum_{\mathrm{i}=1}^{\mathrm{N}}[\frac{1}{\mathrm{N}}-\frac{\overline{x}_2(x_2-\overline{x}_2)}{\sum_{\mathrm{i}=1}^{\mathrm{N}}(x_2-\overline{x}_2)^2}]\in_2$$

$$\mathrm{E}(\in_i|x_1)=E(e\sqrt{x_{2i}}|x_1)=\sqrt{x_2}\,E(e_i|x_i)=0$$

$$E(\hat{\beta}_1|X)=E(\hat{\beta}_1)=\beta_1$$

$$E(\hat{\beta}_0)=\beta_0$$

$$plimN-\infty\hat{\beta}_1=\beta_1+\frac{E[(x_1-E(x_1))\sqrt{x_2e_i}}{E[(x_i-E(x_i))^2]}$$

$$=\beta_1+\frac{E[(x_1-E(x_1))\sqrt{x_1}]Ee_i}{var(x_1)}=\beta_1$$

$plim$ expansion

$$plimN-\infty\hat{\beta}_2=\beta_2+\frac{E[(x_2-E(x_2))\sqrt{x_2}\,e]}{E[(x_2-E(x_2))^2]}$$

$$=\beta_2+\frac{E[(x_2-E(x_2))\sqrt{x_2}]Ee}{var(x_2)}$$

$$plimN-\infty\hat{\beta}_2=\beta_2+\frac{E[(x_2-E(x_2))\sqrt{x_2}\,e]}{E[(x_2-E(x_2))^2]}$$

$$=\beta_2+\frac{E[(x_2-E(x_2))\sqrt{x_2}]Ee}{var(x_2)}$$

$$plimN-\infty\hat{\beta}_0=\beta_0+E(x_1)\beta_1+E(x_2)\beta_2-E(x_1)plim\hat{\beta}_1-E(x_2)plim\hat{\beta}_2+E(e\sqrt{x_2})=\beta_0$$

Formula used

$$plim\hat{\beta}_k=\frac{\beta_k+cov(x_i\,,\in_i)}{var(x_i)}$$

$k=1,2....$

$i=1,2....$

A.2 PROOF FOR THEOREM 2

Since ei is independent of xi and independently, identically and normally distributed with the mean 0 and variance we get the error term, εi = ei√x2i.

REFERENCES

https://www.investopedia.com/terms/c/central_limit_theorem.asp