RESEARCHERID
THOMSON REUTERS

[Badawaire* 6(12): December, 2019] ISSN 2349-4506
Impact Factor: 3.799

Global Journal of Engineering Science and Research Management

A COMPARATIVE STUDY OF SOME ROBUST ESTIMATORS

A. B. Badawaire*, M. G. Bukar, M. L. Danyaro, Umar A. Ahmad, A. Ibrahim
* Department of Mathematics & Statistics, Federal University Wukari, Nigeria
Department of Basic Science, MohametLawan College of Agriculture Maiduguri, Nigeria
Department of Basic Science, Yobe State College of Agriculture Gujba, Nigeria
Department of Basic Science, Taraba State College of Agriculture Jalingo, Nigeria

DOI: 10.5281/zenodo.3591013

KEYWORDS: Outlier(s), Robust Estimator, OLS, Robust M, Robust MM, Robust S, Robust LTS, Robust LMS, Robust LAD, MSE.

ABSTRACT
In the presence of outliers, Ordinary Least Square estimator is found to be inefficient. In this paper, some robust estimators and OLS were used to estimate the parameters of linear regression model in the presents and absents of outlier(s). Linear regression models with three and seven predictors (p = 3 and p = 7) each at four levels of percentage of outliers ($n1\% = 0\%,\ 5\%,\ 10\%\ and\ 20\%$), three levels of variance of outliers ($\sigma^2_{outliers} = 50,\ 100\ and\ 250$) and five levels of sample size ($n = 20, 30, 40, 60\ and\ 100$) were considered through Monte Carlo experiments. The experiments were carried out 2000 times, and the performances of these Robust estimators and Ordinary Least Square were investigated and compared using the Mean Square Error (MSE) criterion. Results show that when there is no outlier in the data, at all situations, Ordinary Least Squares (OLS) is the most efficient estimator among the estimators considered. But when outlier(s) exists in the data, Robust MM consistently performed more efficiently than all other methods of parameter estimation of linear regression model considered. It also shows that efficiency of these estimators increased as the sample size increased. Also, variance of outliers and percentage of outliers affect the efficiency of these estimators.

INTRODUCTION
Ordinary Least Squares (OLS) estimator is the most popular estimator used in estimating the parameters of linear regression model. When all the assumptions of Classical Linear Regression Model are met, Ordinary Least Squares (OLS) estimator is the Best Linear Unbiased Estimator (Fonby, 1984; Maddala, 2002).

One of the problems encountered in regression analysis is the existence of outlier(s) in a dataset. Outlier(s) renders the Ordinary Least Squares (OLS) estimator to be inefficient and the estimates obtained from it to be imprecise as a result of the inflated error variance. Various estimation techniques to handle the problems of outliers have been developed, these include; M estimator proposed by Huber (1964), MM estimator proposed by Yohai (1987), S estimator proposed by Rousseuw and Yohai (1984), LAD estimator proposed by Dielman (1984), LTS proposed by Rousseuw, P. J. and Van Driessen, K. (1998). Different researchers such as Neykov & Neytchev (1991), Atkinson & Weisberg (1991), Stronberg (1993), Rousseeuw & Van Driessen (1999), Agullo (2001), Jung (2005), Li (2005), Cizek (2005) and Rousseeuw & Van Driessen (2006) also suggested different algorithms for computing the LTS estimates.

A Monte-Carlo experiment of 2000 trials were carried out for four sample sizes (20, 40, 60 and 100), each with different number of explanatory variables, different percentage and variance of outliers. Six robust estimators (Robust-M, Robust-MM, Robust-S, Robust- Least Trimmed Square, Robust Least Median Square, and Robust-Least Absolute Deviations) and OLS were used to estimate parameters fitted to this simulated data. These estimators were assessed in terms of their mean square errors (MSE).

MATERIALS AND METHODS
In this paper, a linear regression model with a dependent and three independent variables is considered. The model is specified as:

$$y = X\beta + U \qquad (1)$$

Where y is an n by 1 column vector of dependent variable, X is an n by p matrix of independent variables, β is a p by 1 vector of the regression coefficients and $U$is an n by 1 vector of random errors.

Methods of Parameter Estimation Considered

Robust-M Estimator

M-estimator proposed by Huber (1964) perform parameter estimation by minimizing the sum of a less rapidly increasing function of the residuals. It performs better when the outliers are in the y-direction but less robust to leverage. The objective function of M estimate is given by:

$$min \sum_{i=1}^{n} \rho\left(\frac{e_i}{s}\right) = min \sum_{i=1}^{n} \rho\left(\frac{y_i - X'\hat{\beta}_i}{s}\right) \quad (2)$$

Where s is the scale estimate obtained as function of residuals and is estimated by

$$S = \frac{median|e_i - meadian(e_i)|}{h}$$

When n is large, an appropriate choice of h makes S an approximately unbiased estimator ofσ.

To minimize (2), a system of normal equations is solved by taking partial derivative with respect to β and equating them to zero, which gives

$$\sum_{i=1}^{n} X_i \psi\left(\frac{y_i - X'\hat{\beta}_i}{s}\right) = 0 \quad (3)$$

Where ψ is ρ' and X_i is the i^{th} observation. Then Iterative Reweighted Least Squares (IRLS) or nonlinear optimization techniques is used to solve these equations.

MM-Estimator

MM- estimator proposed by Yohai (1987), is the combination of the high breakdown value estimator and M-estimator. By this estimator, parameter estimates are obtained from multiple M-estimation that uses the S-estimation procedure to minimize the scale of the residuals. The procedure involved the following three stages:

1. Initial estimates $\hat{\beta}^{(1)}$ found using a high breakdown point estimator are used to compute the initial residuals$e_i^{(1)}$.
2. M-estimate of the scale of residuals S_n are computed using the initial estimate of residuals$e_i^{(1)}$ in step 1.
3. M-estimates of regression coefficients are obtained from Weighted Least Squares (WLS) whose first iteration uses the residual scale S_n from step 2 and the estimates of residuals $e_i^{(1)}$ from step 1
$$\sum_{i=1}^{n} X_i w_i \left(\frac{e_i^{(1)}}{S_n}\right) = 0 \quad (4)$$
4. Residuals from the initial Weighted Least Squares (WLS) in step 3 are used to construct new weights, which again is used in Weighted Least Squares estimation. The process is continually reiterated until convergence.

S-Estimator

S-estimator proposed by Rousseuw and Yohai (1984), is a high breakdown value that minimize the dispersion of residuals. S-estimator minimize the dispersion in residuals as the solution of

$$\frac{1}{n}\sum_{i=1}^{n} \psi\left(\frac{e_i}{s}\right) = K \quad (5)$$

Where K, a constant, is defined as K=$E_\varphi[\psi(e)]$and φ represents the standard normal distribution. Differentiating equation (5) and solve the results in

$$\frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{e_i}{s}\right) = K \quad (6)$$

Where $\rho$is replaced with an appropriate weight function.

Least Trimmed Squares Regression Estimator

Another estimator proposed by Rousseuw (1984), is the Least Trimmed Squares (LTS) regression. It is also a high breakdown value method that minimizes the sum of the trimmed squared residuals. LTS estimator is

$$\hat{\beta}_{LTS} = argmin\left(\sum_{i=1}^{h} e_i^2\right) \quad (6)$$

Where $h = [n(1-\theta)+1]$is the number of trimmed observations included in estimation, n and θ representing sample size and the proportion of trimming being performed respectively. In this method, the outlier observations are completely excluded in the estimation of model parameters.

Least Absolute Deviation Estimator

Least Absolue Deviation (LAD) regression proposed by Dielman (1984) is very resistant to observations with unusual y values. Estimates are obtained by minimizing the sum of the absolute values of the residuals.

$$min \sum_{i=1}^{n} |e_i| = min \sum_{i=1}^{n} \left|y_i - X_i\hat{\beta}\right| \quad (7)$$

LAD minimized the objective function given by

$min \sum_{i=1}^{n} \rho_\theta(e_i)$,

Where

$$\rho_\theta(e_i) = \begin{cases} \theta e_i & if\ e_i \geq 0 \\ (\theta - 1)e_i & if\ e_i < 0 \end{cases} \quad (8)$$

Here the quantile being estimated is θ.

LAD fails to account for leverage (Mosteller and Turkey 1977), and thus has a breakdown point of zero.

DATA GENERATION AND MODEL FORMULATION PROCEDURE

The response variable is obtained from the relation given by:

$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i,\ i = 1, \ldots, n\ and\ j = 1,2, \ldots, k$

The error term was simulated to have a Gaussian mixture, i.e. $\varepsilon_i \sim (1 - m\%)N(0,1) + m\%N(0, \sigma^2_{outliers})$.

Where $m\% = (0\%,\ 5\%,\ 10\%,\ 20\%)$ were the percentages of outliers infused into the error term, and $\sigma^2_{outliers} = (50,\ 100,\ 250)$ is the variance of outliers considered.

The predictors were simulated to be independent and identically distributed random variables as;

$X_{ij} \sim N(0,1)$, i = 1,2,3,...,n; j = 1,2,...,p.

The number of predictors were varied as p = (3, 7) so as to study the impact of increase in number of predictors in the model. When p = 3, the true parameter values were fixed as: $\beta_0 = 0, \beta_1 = 0.8, \beta_2 = 0.1, \beta_3 = 0.6$, and when p = 7, $\beta_0 = 0,\ \beta_1 = 0.8, \beta_2 = 0.1, \beta_3 = 0.6,\ \beta_4 = 0.2, \beta_5 = 0.25,\ \beta_6 = 0.3$ and $\beta_7 = 0.53$, such that $\beta'\beta = 1$ in both cases.

To investigate the impact of sample size on the performance of these estimators, five sample sizes were considered: n = 20, 30, 40, 60 and 100.

This experiment is replicated 2000 times and at each stage, the MSE is computed. The MSE of all the estimators is average over the number of replications and parameters as:

$$AMSE(\hat{\beta}) = \frac{1}{2000}\sum_{i=1}^{p}\sum_{j=1}^{2000}(\hat{\beta}_{ij} - \beta_i)^2.$$

Where $\hat{\beta}_{ij}$ is the i[th] element in the j[th] replication of β that gives the estimate of β_i. β_i are the true (fixed) values of parameters.

Ranks were then assigned to these estimators, with rank 1 being assigned to the estimator with lowest average value of $AMSE(\hat{\beta})$, and so on, up to rank 7 which was assigned to the estimator with the largest average value of the average MSE. Thereafter, ranks of these estimators were sum over the levels of number of predictors considered. That is, at a particular level of sample size, percentage of outlier and variance of outlier, an estimator with rank 1 when p (number of predictors) = 3, and rank 1 when p = 7 would have sum of ranks as 2 at those levels considered.

RESULTS AND DISCUSSION

The average MSE of OLS and 6 other robust estimators with varying number of predictors at different sample sizes, variance and percentage of outliers are presented in Table1 to Table3. Also, the performances of these estimators using sum of ranks of their average MSE is presented in Table 4.

Table 1: AMSE of Coefficient Estimates When the Error is Distributed as $\boldsymbol{\varepsilon_i \sim (1 - m\%)N(0, 1) + m\%N(0, 50)}$

Sample size (n)	Estimator	Number of Explanatory Variables							
		3				7			
		% Of Contamination (m%)				% of Contamination (m%)			
		0%	5%	10%	20%	0%	5%	10%	20%
20	OLS	0.061423	6.247426	23.24759	30.22334	0.105514	7.190349	18.55769	42.15712
	M	0.06524	0.069705	0.230811	0.898928	0.112633	0.098469	0.280036	8.337681
	MM	0.067963	0.05983	0.124662	0.096024	0.151417	0.086163	0.109721	0.204764
	S	0.187554	0.152182	0.300936	0.161218	0.361859	0.194213	0.23901	0.25874
	LTS	0.272169	0.221025	0.516459	0.2561	0.499447	0.274026	0.364723	0.386735
	LMS	0.345078	0.301813	0.741765	0.358486	1.376671	1.153672	4.080843	15.79526
	LAD	0.094256	0.092072	0.215907	0.215711	0.155275	0.11674	0.268161	4.012679
30	OLS	0.039367	3.004969	8.376903	14.4213	0.054398	4.054638	13.89453	26.02491
	M	0.0415	0.042076	0.056405	0.117887	0.058116	0.061405	0.123279	0.9975
	MM	0.042267	0.038296	0.038866	0.043362	0.062804	0.055556	0.072366	0.091045
	S	0.132157	0.11593	0.103268	0.084857	0.186252	0.163965	0.192417	0.185728
	LTS	0.186604	0.178378	0.165294	0.140961	0.242572	0.219243	0.273968	0.259227
	LMS	0.212268	0.199284	0.188907	0.166249	0.359259	0.299046	0.393588	0.396511
	LAD	0.059645	0.058377	0.065936	0.088697	0.083039	0.080778	0.131074	0.312645
	OLS	0.028477	3.697655	6.434459	14.11909	0.028773	4.042546	8.905901	17.33911

40	M	0.030059	0.036631	0.042755	0.089108	0.030428	0.043438	0.06383	0.213486
	MM	0.030628	0.032267	0.030596	0.038362	0.031326	0.036512	0.04385	0.051215
	S	0.097841	0.095056	0.086321	0.081434	0.102742	0.116488	0.130771	0.125921
	LTS	0.149918	0.146714	0.13664	0.130775	0.13777	0.155815	0.176737	0.175501
	LMS	0.155854	0.155404	0.146878	0.1482	0.156982	0.187527	0.212012	0.22794
	LAD	0.044974	0.050362	0.052032	0.075079	0.043364	0.058049	0.075188	0.112161
60	OLS	0.022894	2.23443	4.106808	8.728432	0.017832	2.390121	4.380885	9.799899
	M	0.024297	0.022794	0.027058	0.04744	0.019021	0.023905	0.030856	0.065162
	MM	0.02448	0.019808	0.019745	0.022984	0.019363	0.02042	0.021878	0.027686
	S	0.085094	0.063025	0.058707	0.056456	0.074309	0.076768	0.078338	0.08868
	LTS	0.138364	0.102712	0.096524	0.093293	0.100203	0.105709	0.104193	0.113549
	LMS	0.135771	0.101106	0.096311	0.095875	0.107141	0.111272	0.112801	0.125426
	LAD	0.0352	0.030791	0.033516	0.044846	0.027591	0.032234	0.037423	0.054901
100	OLS	0.011754	1.28081	2.325764	5.232161	0.010523	1.455848	2.953979	5.326584
	M	0.012474	0.012954	0.014469	0.027555	0.011195	0.014468	0.018455	0.029046
	MM	0.012521	0.011405	0.010846	0.014155	0.01129	0.012521	0.013472	0.015071
	S	0.049318	0.040977	0.036357	0.039522	0.053126	0.057922	0.06253	0.067002
	LTS	0.080809	0.07055	0.061225	0.06188	0.071586	0.076254	0.081886	0.079429
	LMS	0.074652	0.06487	0.058472	0.06239	0.07047	0.074451	0.080013	0.078941
	LAD	0.018238	0.017808	0.018048	0.026759	0.016671	0.019762	0.022758	0.028289

Table 2: AMSE of Coefficient Estimates When the Error is Distributed as $\varepsilon_i \sim (1-m\%)N(0,1) + m\%N(0,100)$

Sample size (n)	Estimator	Number Of Explanatory Variables							
		3				7			
		% Of Contamination (m%)				% of Contamination (m%)			
		0%	5%	10%	20%	0%	5%	10%	20%
20	OLS	0.052334	29.34554	64.20824	126.5481	0.086021	42.34474	114.8166	178.4473
	M	0.055605	0.08347	0.11823	10.17642	0.092332	0.134523	1.479007	31.49192

	MM	0.05812	0.070827	0.075035	0.106885	0.143079	0.108748	0.147129	0.183252
	S	0.157474	0.176632	0.172091	0.187412	0.309946	0.259832	0.318375	0.25495
	LTS	0.21223	0.270261	0.254621	0.314967	0.444954	0.359082	0.46508	0.40347
	LMS	0.267318	0.349745	0.350236	0.459222	1.437843	1.410487	3.279449	27.64109
	LAD	0.07773	0.110017	0.132293	1.314547	0.127501	0.162941	0.320877	16.00664
30	OLS	0.040054	12.68349	41.74351	86.41252	0.066586	14.39181	43.72525	84.34434
	M	0.042303	0.045676	0.069034	0.191384	0.072012	0.057343	0.084722	1.560415
	MM	0.043526	0.040866	0.04858	0.058136	0.077356	0.052093	0.053283	0.063716
	S	0.137372	0.115901	0.123464	0.117217	0.225143	0.154168	0.141327	0.138154
	LTS	0.211613	0.175411	0.184505	0.181787	0.301908	0.209753	0.189108	0.193559
	LMS	0.242036	0.198587	0.209385	0.223028	0.428712	0.294411	0.271174	0.283755
	LAD	0.061616	0.063446	0.084095	0.120512	0.101556	0.076561	0.095967	0.201144
40	OLS	0.036607	15.38683	26.64707	55.93606	0.032886	19.02597	31.66547	71.12653
	M	0.038699	0.039916	0.045017	0.092344	0.034873	0.051424	0.061097	0.490066
	MM	0.039214	0.033211	0.032231	0.038364	0.036356	0.042549	0.03897	0.050985
	S	0.128232	0.104166	0.088725	0.083155	0.125148	0.138751	0.11705	0.130794
	LTS	0.187369	0.164001	0.140584	0.132703	0.168973	0.192649	0.159038	0.1788
	LMS	0.197956	0.180251	0.152503	0.148945	0.197494	0.228135	0.193892	0.228431
	LAD	0.056889	0.05211	0.054923	0.075529	0.050537	0.067084	0.06979	0.118363
60	OLS	0.015534	8.873185	16.70181	37.88	0.016253	10.19386	19.5704	41.30709
	M	0.016313	0.022692	0.027002	0.050429	0.017347	0.027781	0.034107	0.073917
	MM	0.016473	0.01962	0.019293	0.024354	0.017618	0.023657	0.023054	0.028186
	S	0.058022	0.064178	0.05643	0.058853	0.068431	0.08557	0.085773	0.093531
	LTS	0.093628	0.102573	0.093304	0.097005	0.095728	0.116673	0.116569	0.119982
	LMS	0.093814	0.100723	0.093586	0.101333	0.098025	0.123659	0.131495	0.130761
	LAD	0.024085	0.029959	0.033122	0.048563	0.025054	0.036492	0.040639	0.059107

100	OLS	0.010993	5.41747	10.29576	21.08097	0.010411	5.405534	10.91113	22.21648
	M	0.011588	0.013595	0.016173	0.027719	0.011046	0.013991	0.017406	0.032165
	MM	0.011648	0.011739	0.011952	0.013826	0.011129	0.012057	0.012698	0.015061
	S	0.045986	0.042109	0.040188	0.039703	0.052045	0.057345	0.057673	0.069791
	LTS	0.076883	0.072726	0.06689	0.063935	0.069189	0.074512	0.07449	0.08357
	LMS	0.071316	0.067765	0.064054	0.060865	0.068448	0.071917	0.073883	0.08239
	LAD	0.017284	0.018735	0.020417	0.026595	0.01625	0.01907	0.021625	0.029758

Table 3: AMSE of Coefficient Estimates When the Error is Distributed as $\varepsilon_i \sim (1-m\%)N(0,1) + m\%N(0,250)$

Sample size (n)	Estimator	Number of Explanatory Variables							
		3				7			
		% Of Contamination (m%)				% of Contamination (m%)			
		0%	5%	10%	20%	0%	5%	10%	20%
20	OLS	0.091318	237.1946	349.4772	1186.332	0.084089	227.7056	505.2116	1110.545
	M	0.097732	0.112396	0.146472	25.98181	0.090909	0.120173	11.89643	180.263
	MM	0.105496	0.093694	0.067856	0.131413	0.121189	0.095938	0.110816	0.154341
	S	0.270351	0.24996	0.159402	0.228573	0.273649	0.222946	0.247738	0.248049
	LTS	0.401907	0.385558	0.263652	0.358918	0.396827	0.307696	0.368246	0.352653
	LMS	0.646601	0.501229	0.38368	0.515444	1.171048	1.08217	1.794332	17.79259
	LAD	0.136801	0.143434	0.128079	0.35754	0.126544	0.142613	0.879781	73.8002
30	OLS	0.043671	74.35169	237.2043	490.0008	0.040533	107.6604	268.3947	537.8574
	M	0.046686	0.041978	0.06592	0.374431	0.043254	0.068095	0.097815	8.881622
	MM	0.047909	0.038464	0.043587	0.054029	0.048049	0.059658	0.052833	0.061764
	S	0.147508	0.112175	0.113705	0.112197	0.150288	0.189038	0.149043	0.133283
	LTS	0.216891	0.167664	0.178287	0.184412	0.203396	0.265473	0.215821	0.184434
	LMS	0.239285	0.186781	0.212433	0.215016	0.297986	0.349206	0.320515	0.274583
	LAD	0.06664	0.057788	0.077513	0.118872	0.061312	0.088527	0.101831	0.218969
	OLS	0.024199	87.39178	174.5112	405.8239	0.030101	86.16474	186.479	468.2405

40	M	0.02574	0.036193	0.048174	0.448816	0.032071	0.039646	0.056085	1.588143
	MM	0.02627	0.030386	0.034788	0.042499	0.033101	0.033221	0.036157	0.050858
	S	0.085774	0.091697	0.096855	0.096763	0.111914	0.107228	0.111339	0.134178
	LTS	0.126773	0.146003	0.1498	0.158871	0.14993	0.148103	0.150934	0.18002
	LMS	0.138537	0.154355	0.161997	0.179501	0.1861	0.179426	0.182563	0.225992
	LAD	0.037726	0.04795	0.05875	0.092139	0.045981	0.051912	0.065545	0.122549
60	OLS	0.016429	59.09012	103.552	192.0193	0.02003	60.38009	118.5997	239.6983
	M	0.017376	0.024383	0.028254	0.044851	0.021267	0.025906	0.031961	0.07058
	MM	0.01756	0.020972	0.019803	0.020068	0.021604	0.022054	0.022346	0.025949
	S	0.064105	0.069275	0.05689	0.049184	0.084604	0.079535	0.081117	0.087696
	LTS	0.096648	0.111246	0.094126	0.081778	0.114364	0.107063	0.108184	0.112205
	LMS	0.097358	0.11435	0.09381	0.08339	0.120679	0.116159	0.113125	0.119896
	LAD	0.026012	0.033528	0.034428	0.040259	0.031198	0.034034	0.038422	0.056042
100	OLS	0.013352	31.84357	64.41349	129.1402	0.010942	32.2586	67.91249	127.7668
	M	0.014167	0.013465	0.016973	0.026138	0.011502	0.013688	0.017427	0.03009
	MM	0.014232	0.011655	0.012376	0.013482	0.011581	0.011723	0.012291	0.013828
	S	0.05523	0.041145	0.041454	0.038048	0.055153	0.054464	0.058536	0.065924
	LTS	0.092572	0.068572	0.068086	0.059966	0.07421	0.072228	0.074605	0.076938
	LMS	0.084665	0.063245	0.063708	0.05936	0.07057	0.072125	0.074296	0.075164
	LAD	0.021085	0.018194	0.020876	0.025739	0.016726	0.018547	0.021214	0.02778

Table 4: Sum of Ranks of the Estimators (over the levels of number of predictors) at Different levels of Sample size, variance and Percentage of Outlier

Sample size (n)	Estimators	Distribution of Error Term											
		$\varepsilon_i \sim (1-m\%)N(0,1) + m\%N(0,50)$				$\varepsilon_i \sim (1-m\%)N(0,1) + m\%N(0,100)$				$\varepsilon_i \sim (1-m\%)N(0,1) + m\%N(0,250)$			
		% Of Contamination (m%)				% Of Contamination (m%)				% Of Contamination (m%)			
		0%	5%	10%	20%	0%	5%	10%	20%	0%	5%	10%	20%
	OLS	2	14	14	14	2	14	14	14	2	14	14	14
	M	4	4	7	11	4	4	7	12	4	4	9	12
	MM	6	2	2	2	7	2	2	2	6	2	2	2

20	S	10	8	6	4	10	8	6	4	10	8	6	4
	LTS	12	10	10	7	12	10	9	6	12	10	8	7
	LMS	14	12	12	11	14	12	12	9	14	12	11	9
	LAD	8	6	5	7	7	6	6	9	8	6	6	8
30	OLS	2	14	14	14	2	14	14	14	2	14	14	14
	M	4	4	4	10	4	4	4	11	4	4	4	12
	MM	6	2	2	2	6	2	2	2	6	2	2	2
	S	10	8	8	4	10	8	8	4	10	8	8	4
	LTS	12	10	10	8	12	10	10	7	12	10	10	7
	LMS	14	12	12	11	14	12	12	11	14	12	12	10
	LAD	8	6	6	7	8	6	6	7	8	6	6	7
40	OLS	2	14	14	14	2	14	14	14	2	14	14	14
	M	4	4	4	9	4	4	4	10	4	4	4	12
	MM	6	2	2	2	6	2	2	2	6	2	2	2
	S	10	8	8	6	10	8	8	6	10	8	8	6
	LTS	12	10	10	9	12	10	10	9	12	10	10	8
	LMS	14	12	12	12	14	12	12	11	14	12	12	10
	LAD	8	6	6	4	8	6	6	4	8	6	6	4
60	OLS	2	14	14	14	2	14	14	14	2	14	14	14
	M	4	4	4	6	4	4	4	6	4	4	4	6
	MM	6	2	2	2	6	2	2	2	6	2	2	2
	S	10	8	8	8	10	8	8	8	10	8	8	8
	LTS	13	11	11	10	12	11	10	10	12	10	11	10
	LMS	13	11	11	12	14	11	12	12	14	12	11	12
	LAD	8	6	6	4	8	6	6	4	8	6	6	4
100	OLS	2	14	14	14	2	14	14	14	2	14	14	14
	M	4	4	4	6	4	4	4	6	4	4	4	6
	MM	6	2	2	2	6	2	2	2	6	2	2	2
	S	10	8	8	8	10	8	8	8	10	8	8	8
	LTS	14	12	12	11	14	12	12	12	14	12	12	12
	LMS	12	10	10	11	12	10	10	10	12	10	10	10
	LAD	8	6	6	4	8	6	6	4	8	6	6	4

From Table1 to Table3 above, it can be observed that, at fixed Variance of outliers, number of explanatory variables and percentage of outliers, as sample size increase, the AMSE of the estimator's decreases.

At a particular level of sample size, number of explanatory variables and percentage of outliers, AMSE of the estimators mostly increase with increase in variance of outliers. It can as well be deduced from table1 to table3 that mostly, AMSE increases as number of explanatory variables increase, at fixed level of sample size, variance and percentage of outliers.

From Table 4 above, it can be deduced that at all levels of sample size and variances of outlier considered, when the percentage of outlier is zero (i.e no outlier), OLS has the minimum sum of ranks. But when outlier exists, irrespective of the level of sample size and level of variance of the outlier(s), Robust MM has the minimum sum of ranks.

CONCLUSSION

In this study, the performances of OLS and some robust estimators that handled the problem of outliers were evaluated and compared using the MSE criterion.

From the simulation study, where the sample size, variance of outliers, number of explanatory variables and percentage of outliers were varied, it's found that if outlier exists, Robust MM estimator outperformed all other methods of parameter estimation, and is therefore, recommended as the most suitable estimator among the various estimators considered, when faced with the problem of outliers. But when there is no outlier(s) in the data, OLS is found to be the most efficient estimator among all the estimators considered in this study.

REFERENCES

1. Agullo, J. (2001), 'New algorithms for computing the least trimmed squares regression estimator', *Computational Statistics and Data Analysis* 36(4), 425–439.
2. Atkinson, A. & Weisberg, S. (1991), simulated annealing for the detection of multiple outliers using least squares and least median of squares fitting, *in* W. Stahel & S. Weisberg, eds, '*Directions in Robust Statistics and Diagnostics'*, Springer-Verlag, New York.
3. Cizek, P. (2005), 'Least trimmed squares in nonlinear regression under dependence', *Journal of Statistical Planning and Inference* 136(11), 3967–3988.
4. Fomby, T. B., Hill, R. C. and Johnson, S. R. 1984. Advance Econometric Methods. Springer-Verlag, New York, Berlin, Heidelberg, London, Paris,Tokyo.
5. Huber, P. H. 1964. ''Robust estimation of location parameter.'' The Annals of Mathematical Statistics, 35 7 101.
6. Jung, K. (2005), 'Multivariate least-trimmed squares regression estimator', *Computational Statistics and Data Analysis* 48(2), 307–316.
7. Li, L. (2005), 'An algorithm for computing exact least-trimmed squares estimate of simple linear regression with constraints', *Computational Statistics and Data Analysis*, 48(4), 717–734.
8. Maddala, G. S. 2002. Introduction to Econometrics. 3rd Edition, John Willey and Sons Limited, England.
9. Neykov, N. & Neytchev, P. (1991), 'Least median of squares, least trimmed squares and S estimations by means of BMDP3R and BMDPAR', *Computational Statistics Quarterly* 4, 281–293.
10. Rousseeuw, P. J. & van Driessen, K. (1999), 'A fast algorithm for the minimum covariance determinant estimator', *Technometrics* 41(3), 212–223.
11. Rousseeuw, P. & van Driessen, K. (2006), 'Computing LTS regression for large data sets', *Data Mining and Knowledge Discovery* 12(1), 29–45.
12. Rousseeuw p. j., and Yohai, 1984. ''Robust regression by means of S estimators in Robust and Nonlinear Time Series Analysis'', Franke J. Hardle, W. and Martin, R. D. Lecture Notes in Statistics, 26, New York: Springer-Verlag 256 - 274.
13. Rousseeuw, P. J. and Van Driessen K., ''Computing LTS for Large Data Sets.'' Technical Report, University of Antwerp, submitted, 1998.
14. Samkar, H. and Alpu, O. (2010). Ridge regression based on some robust estimators. Journal of Modern Applied Statistical Methods: 9(2), 495-501.
15. Stromberg, A. (1993), 'Computing the exact least median of squares estimate and stability diagnostics in multiple linear regression', *SIAM Journal on Scientific Computing* 14(6), 1289–1299.
16. Yohai, V. J. ''High breakdown point and high efficiency robust estimates for regression.'' The Annals of Statistics, 15, 642 – 656, 1987.