CARD International Journal of Science and Advanced Innovative Research (IJSAIR)
ISSN: 2536-7315 (Print) 2536-7323 (Online)
Volume 2, Number 2, June 2017
http://www.casirmediapublishing.com

Statistical Comparison of M-Estimators in Multiple Regression Analysis of Nigeria Microfinance Banks Loans and Advances

D. S. A. Wokoma[1], F. N. Nwobi[2], E. M. Eke[3] & A. E. Ukwubuiwe[4]
[1]Department of Statistics, Port Harcourt Polytechnic, Rumuola, Rivers State, Nigeria
[2]Department of Statistics, Imo State University, Owerri, Imo State, 460222, Nigeria
[3]Department of Statistics, Abia State Polytechnic, P.M.B 7166, Aba, Nigeria
[4]Department of Statistics, Abia State College of Health Sciences Management Technology, Aba,
Email: dagywoksy@gmail.com,fnnwobi@yahoo.co.uk,,ekeemenaik@yahoo.com,eluojoa@gmail.com
Corresponding Author: D. S. A. Wokoma

ABSTRACT
Residuals of OLS estimation for the classical linear regression model are heavily affected by outliers in the dataset. This leads to unreliable, inefficient and inaccurate parameter estimates that will not yield a robust predictive model of the phenomena. In this paper, we identified a robust estimator in the class of M-estimators that reduced the RMSE of Nigerian Microfinance Loans and advances data by 45.28% relative to the OLS estimate of the RMSE and also recovered the directional effect of Lending interest rate on Loans/advances that the OLS estimator lost as a result of outliers' effect.
Keywords: Robust, M-estimator, OLS estimator, Outlier, RMSE, Loans and advances, Deposit, Lending interest rate

INTRODCTION
One of the major assumptions of the Ordinary Least Squares (OLS) estimation for the classical linear regression model is that the residuals are normally distributed. This normality assumption of the residuals is greatly affected by outlier(s) in the dataset. The more there are outliers in a dataset, the more the dataset deviates from normality. Outliers have been defined in several ways by authorities in different statistics literature. In data mining, an outlier is defined an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism (Aggarwal, 2013). In econometrics, Greene, (2012) defines an outlier as an observation that appears to be outside the reach of the model, perhaps because it arises from a different data generating process. In multivariate analysis, outliers are identified as unusual observations that do not seem to belong to the pattern of variability produced by

the other observations, (Johnson and Wichern, 2007). In experimental designs, an outlier is viewed as an extreme observation or an extreme residual of observations that are larger in absolute value than four standard deviation from the mean (Inamadar et al., 2015).

Kamile et al., (2012) listed the effects of outliers in regression analysis to include among others: disproportionate influence on the estimated parameters leading to the wrong conclusion of the significance of the parameters, inconsistent and inaccurate predictions and decrease of the efficiency of the estimators. Since not all outliers are wrong data, such outliers must be accommodated in the dataset. Deleting them represents distorting the natural variation in the data. Moreover, every observation in a dataset carries some information that should be exploited (Paul and Bhar, 2011). The way out of this is the use of robust and efficient statistical techniques that will not be unduly affected by outliers or other small departure from the model assumption(s). The application and uses of robust and efficient estimation methods have been richly proposed in the regression literature. They are resistant to errors in the results produced by deviations from normality assumptions, see for example Rousseeuw and Leroy (1987). Robust statistical methods have been developed for many common problems, such as estimating location, scale and regression parameters. Among the several approaches of robust and efficient estimation methods proposed (R-estimators, L-estimators, S-estimators and M-estimators), M-estimators are the most popular and have dominated the field of statistical analysis due to their generality, high breakdown point and efficiency (Huber, 1981).

In regression M-estimation, the objective function to be minimized to get the parameter estimate is weighted according to the residual of each observation. A good number of the objective function to be minimized are non-linear in nature and therefore, normal equations for solving the parameter estimates are also non-linear in parameter. Iteratively Reweighted Least Squares (IRLS) methods are employed to solve these equations (Holland and Welsch, 1977). Many M-estimators are in use but they differ in their efficiency. In this paper, we centre our comparison on the M-estimators' functions in MATLAB 2008: Andrews, Tukey, Cauchy, Fair, Huber, Logistic, Talwar and Welsch.

OLS Multiple regression analysis and its application is not new in the literature. Ogar et al (2014) investigate the impact of commercial bank loans on manufacturing sector and to establish the relationship between interest rate and manufacturing sector performance. They used OLS multiple regression model to establish the relationship and found that commercial bank credit had a significant relationship on manufacturing sector. Stešević (2008) modelled the effect ofinterest rates on deposits in Montenegro. Nakayiza (2013) studied the contribution of interest rates to loan portfolio performance in commercial banks. The findings revealed that there is lack of effective analysis on the impact of increasing interest rates on loan repayment trends. Ajayi (2007) identified four major determinants of loans/advances in the commercial bank: deposit, liquidity ratio, capital base and lending interest rates. He discovered significant positive effect of deposit and capital base; significant negative effect of lending interest rates and liquidity. Since the bank consolidation of 2005 had solved the capital base and liquidity ratio problemof Nigeria commercial banks and microfinance bank, we consider only deposit and lending interest rate. Imoisi et al (2012) using the Multiple Regression OLS Method found a significant relationship between

Deposit Money Banks loans and advances and agricultural output. Parvesh and Afroze (2014) examined the impact of specific bank performance factors particularly Loan, Asset Quality, Management Efficiency, Liquidity and Sensitivity on capital adequacy requirements among private sector banks of India. The regression results obtained revealed that Loans, Management Efficiency, Liquidity and Sensitivity have statistically significant influence on the capital adequacy of private sector banks. Awoyemi and Jabar (2014) in order to normalized the data applied a log transformation in regressing the commercial prime lending rate and the performance of MFBs: Total assets, total loan to MFBs, total deposits mobilized by

MFBs and Shareholders fund of MFBs. The problem with log-transformation is that the log-transformed data may become normal in certain normality tests but their p-value will still show rejection of the normality tests. The best way out is the use of M-estimation.

M-Estimators have been applied in Regression Analysis in several areas: Shi-Woei (2006) compared the Classical least squares regression with three classes of robust regression Estimators: M-estimators, the bounded influence estimators (GM-estimators) and the high breakdown point estimators and found that both GM-estimators and M-estimators consistently outperform the ordinary least squares method when the normality assumption is violated. High breakdown point estimators, though theoretically robust to the leverage points, cannot achieve the needed stability. Robust regression using M-estimators or GM-estimators can be a viable alternative or a supplement to ordinary least squares method.

Muthukrishnan and Radha (2010) identified three most commonly used M-estimators: Huber M-estimator, Hampel estimator, Tukey estimators and compared them with the least squares estimator when outliers are present in the data. They interestingly found that the M-estimators yield essentially the same results as the least square estimator in normal situation but when outliers are present in the data; least square estimator does not provide useful information for the majority of the data but not in the case of robust estimators. That is, they observed that the M-estimators are not affected by outliers. That is, the performances of M-estimators are almost same as the method of least squares in normal situations and also in the presence of outliers.

The rest of the paper is arranged as follows: In section 2, we explain data sources, sample and sampling techniques. We also discuss the OLS and M-estimations in multiple regression analysis and normality tests. In section 3, we apply these OLS and M-estimation techniques in multiple regression analysis and normality tests to the data. The conclusion is given section 4.

METHODOLOGY

Data Sources

The data for this paper are secondary data obtained from Central Bank of Nigeria Statistical Bulletins Summary of Assets and Liabilities of Microfinance Banks (₦'Million) from 1970 to 2014. The sample for the study covers the loans/advances, deposits and lending interest rate of Microfinance banks from 1970-2014, a period of 45 years.

Ordinary Least Squares (OLS) Estimation of Multiple Regression Parameters

Gujarati (2004) defined Multiple Regression Analysis as a statistical technique for investigating the relationship between one dependent variable Y and two or more independent variables $(X_1, X_2 \ldots, X_k)$. The knowledge of multiple regression enable us to understand the complexity of the interaction among variables in business and other economic activities. One of the variables under study is called the "cause" and the other is called the "effect". It means that one variable depends on the others. In many cases of the relationship, some other factors cause the changes in these variables other than natural, hence the relationship becomes one of association rather than causes and effect. A typical multiple regression model is represented by

$$y = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \ldots + \beta_k X_{kj} + \varepsilon \quad (1)$$

Where y is the dependent variable from population of interest, $\beta_0, \beta_1, \ldots, \beta_n$ are population regression parameters, $X_{1j}, X_{2j}, \ldots, X_{kj}$ are observed values of the independent variables $X_1, X_2,, \ldots, X_n$ respectively. The estimated multiple regression model is given as

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_{ij} + \hat{\beta}_2 X_{2j} + \ldots + \beta_n X_{nj} \quad (2)$$

Where $\hat{\beta}_0$ is the amount of change in y that is not attributable to the independent variables, $\hat{\beta}_1$ is the amount of change in y for a unit change in X_1 when holdingX_2 constant and $\hat{\beta}_2$ is the amount changes in y for a unit change in X_2 when X_1 is held constant. (1) can be written in matrix as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad (3)$$

Where $\boldsymbol{y}$ is $n \times 1$ vector of observation, $\boldsymbol{X}$is an $n \times p$ matrix of rank p and $\boldsymbol{\beta}$is a $p \times 1$vector of parameters.The Ordinary Least Squares

(OLS) analysis of the relationship existing between variables that minimizes the sum of squared residuals is given as

$$\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})$$
$$= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta} \quad (4)$$

Differentiating with respect to β and equality to zero, we have

$$\frac{\partial \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{\partial \beta} = -2\boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X\beta} = 0 \quad (5)$$

Which gives the estimates of the regression parameters as

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{y}) \quad (6)$$

Where

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} n & \sum X_1 & \sum X_2 \ldots & \sum X_k \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 \ldots & \sum X_1X_k \\ \sum X_2 & \sum X_1X_2 & \sum X_2^2 \ldots & \sum X_2X_k \\ \sum X_k & \sum X_1X_k & \sum X_2X_k \ldots & \sum X_k^2 \end{bmatrix},$$

$$\boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} \sum y \\ \sum X_1\, y \\ \sum X_2\, y \\ \vdots \\ \sum X_k\, y \end{bmatrix}$$

$$Var - Cov(\hat{\beta}) = \hat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}} \quad (7)$$

The residual variance $\hat{\sigma}^2 = \dfrac{\boldsymbol{y}'\boldsymbol{y} - \widehat{\boldsymbol{\beta}}(\boldsymbol{X}'\boldsymbol{Y})}{n - k - 1}$

M-estimation in Multiple Regression Analysis

For the model defined in (3), Paul and Bhar (2011) defines a class of robust estimators that minimize a function ρ of the errors, i.e.,

$$\underset{\beta}{Minimize}\sum_{i=1}^{n}\rho\,(e_i)$$

$$= \underset{\beta}{Minimize}\sum_{i=1}^{n}\boldsymbol{\rho}\,(y_i - \mathbf{x}_i'\boldsymbol{\beta}) \quad (8)$$

Where $\mathbf{x}_i'$ denotes the i^th^ row of X.

An estimator of $\boldsymbol{\beta}$ from this set is called an M-estimator. If the method of OLS is used (implying the error distribution is normal), then $\rho(e_i) = \left(1/_2\right)e_i^2$. Generally, instead of $\rho(e_i)$, the function $\rho(e_i/\sigma)$ is minimized, where σ is a scale parameter. Iteratively Reweighted Least Squares (IRLS) method is used to obtain the parameter estimates. Suppose that an initial estimate $\widehat{\boldsymbol{\beta}}_0$ is available and that s is an estimate of scale. Then the equations for solving for parameter estimates are given as

$$\sum_{i=1}^{n}\boldsymbol{x_{ij}}\,\boldsymbol{\psi}\left(\frac{y_i - \mathbf{x}_i'\boldsymbol{\beta}}{s}\right) = \sum_{i=1}^{n}\frac{\boldsymbol{x_{ij}}\{\boldsymbol{\psi}[(y_i - \mathbf{x}_i'\boldsymbol{\beta})/s]\}/\{(y_i - \mathbf{x}_i'\boldsymbol{\beta})/s\}}{(y_i - \mathbf{x}_i'\boldsymbol{\beta})/s} = \mathbf{0} \quad (9)$$

$$\boldsymbol{or} \quad \sum \boldsymbol{x_{ij}}\boldsymbol{w_{i0}}(y_i - \mathbf{x}_i'\boldsymbol{\beta})/s = \mathbf{0} \quad (10)$$

Where $\boldsymbol{\psi} = \boldsymbol{\rho}'$, the first derivative function of $\boldsymbol{\rho}$ and

$$\boldsymbol{w_{i0}} = \frac{\boldsymbol{\psi}\left[\frac{(y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}_0)}{s}\right]}{\frac{(y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}_0)}{s}} \; if\; y_i \neq \mathbf{x}_i'\widehat{\boldsymbol{\beta}}_0 \quad (11)$$

$$= \mathbf{1} \qquad if\; y_i = \mathbf{x}_i'\widehat{\boldsymbol{\beta}}_0$$

In matrix notation equation (11) is written and solved as

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\boldsymbol{W}^{(0)}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{W}^{(0)}\boldsymbol{y} \quad (12)$$

Where $\boldsymbol{W_0}$ is a $\times\, n$ diagonal matrix of "weights" with diagonal elements $w_{10}, w_{20}, \ldots, w_{n0}$. Equation (12) is very similar to the solution for the OLS estimator, but with the introduction of a weight matrix to reduce the influence of outliers. Stuart (2011) listed the steps of IRLS which is used to solve (12) and the convergence criterion:

1. With the iteration counter, l set to 0, the OLS method is used to fit an initial model to the data, yielding the initial estimates of the regression coefficients, $\hat{\boldsymbol{\beta}}^{(\mathbf{0})}$.
2. Initial residuals $r_i^{(0)}$ are found using $\hat{\boldsymbol{\beta}}^{(\mathbf{0})}$ and used to calculate $s^{(\mathbf{0})}$.
3. A weight function $w(z)$ is chosen and applied to $\frac{r_i^{(0)}}{s^{(\mathbf{0})}}$ to obtain preliminary weights $w_i^{(0)}$. These give the value of $\boldsymbol{W}^{(\mathbf{0})}$ for $\hat{\boldsymbol{\beta}}^{(\mathbf{0})}$.
4. Set $l = 1$. Using $\boldsymbol{W}^{(\mathbf{0})}$, one obtains the estimate $\hat{\boldsymbol{\beta}}^{(\mathbf{1})} = \left(\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{0})}\boldsymbol{X}\right)^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{W}^{(\mathbf{0})}\boldsymbol{y}$.
5. Using $\hat{\boldsymbol{\beta}}^{(\mathbf{1})}$ new residuals, $r_i^{(1)}$ can be found, which, via calculation of $s^{(\mathbf{1})}$ and application of the weight function yield $\boldsymbol{W}^{(\mathbf{1})}$.
6. Set $l = 2$. A new estimate for $\boldsymbol{\beta}$ is found using $\boldsymbol{W}^{(\mathbf{1})}$. This is $\hat{\boldsymbol{\beta}}^{(\mathbf{2})}, r_i^{(2)}$ *and* $s^{(\mathbf{2})}$, and in turn the next weight matrix, $\boldsymbol{W}^{(\mathbf{2})}$ are then found.
7. This iteration process is continued until $l = q$

$$\hat{\boldsymbol{\beta}}^{(q)} = \left(\boldsymbol{X}'\boldsymbol{W}^{(q)}\boldsymbol{X}\right)^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{W}^{(q)}\boldsymbol{y}$$

Until the estimates of $\boldsymbol{\beta}$ converge, at which point the final M-estimate has been found.

Convergence tends to be reached quickly, and the procedure is usually stopped once the estimate changes by less than a selected percentage between iterations, or after a fixed number of iterations have been carried out. The convergence criterion is of the form

$$\frac{\left\|\hat{\beta}^{(q+1)} - \hat{\beta}^{(q)}\right\|}{\left\|\hat{\beta}^{(q+1)}\right\|} < \varepsilon = 0.0001 \quad or \quad \frac{\left\|r^{(q+1)} - r^{(q)}\right\|}{\left\|r^{(q+1)}\right\|} < \varepsilon = 0.0001 \qquad (13)$$

ε is a small positive number, often fixed at 0.0001. This is slightly different to the convergence criterion used by various statistical software, which iterates until the percentage change in the size of the residuals between iterations is smaller than ε, Stuart (2011).

The MATLAB robust it function handles the iteration process and outputs the computation results using various weighting function.

MATLAB (2008a) showed the weighting function as presented in Table 2.1:

Table 2.1: Weighting functions and their tuning constants

Weight Function	Equation	Default Tuning Constant
'andrews'	w = (abs(r)<pi) .* sin(r) ./ r	1.339
'bisquare' (default)	w = (abs(r)<1) .* (1 - r.^2).^2	4.685
'cauchy'	w = 1 ./ (1 + r.^2)	2.385
'fair'	w = 1 ./ (1 + abs(r))	1.400
'huber'	w = 1 ./ max(1, abs(r))	1.345
'logistic'	w = tanh(r) ./ r	1.205
'ols'	OLS (no weighting function)	None
'talwar'	w = 1 * (abs(r)<1)	2.795
'welsch'	w = exp(-(r.^2))	2.985

If tune is unspecified, the default value in the table is used. Default tuning constants give coefficient estimates that are approximately 95% as statistically efficient as the ordinary least-squares estimates, provided the response has a normal distribution with no outliers. Decreasing the tuning constant increases the downweight assigned to large residuals; increasing the tuning constant decreases the down weight assigned to large residuals.

The value r in the weight functions is

r = resid/(tune*s*sqrt(1-h))

Where resid is the vector of residuals from the previous iteration, h is the vector of leverage values from a least-square fit, and s is an estimate of the standard deviation of the error term given by

s = MAD/0.6745

MAD is the median Absolute Deviation of the residuals from their median. The constant 0.6745 makes the estimate unbiased for the normal distribution. If there are p columns in X, the smallest p absolute deviations are excluded when computing the median.

Normality Test

Normality tests that are mostly used for the univariate dataset are: Shapiro–Wilk test (Shapiro and Wilk, 1965), Jarque–Bera test (Jarque and Bera, 1987), Anderson–Darling test (Anderson and Darling, 1954), Lilliefors test (Lilliefors, 1967), Kolmogorov–Smirnov test (Massey, 1951), Razali and Wah (2011). In this study, we shall use the Shapiro–Wilk and Lilliefors tests.The null-hypothesis of this test is the error term is normally distributed. Thus, if the p-value is less than 0.05, then the null hypothesis is rejected and there is evidence that the data tested are not from a normally distributed population; otherwise, the null hypothesis is accepted.

MATLAB software will be used for the OLS estimation, M-estimation and Lilliefors normality tests while R software will be used for the Shapiro-Wilk normality test.

DATA ANALYSIS AND DISCUSSION OF RESULTS

$\mathbf{y}$ is 45×1 vector of loans and advances data, $\boldsymbol{X}$is an 45×3 matrix of rank 3 and $\boldsymbol{\beta}$is a 3×1vector of parameters. OLS method was used to regress $\boldsymbol{X}$on$\mathbf{y}$and the residuals were obtained. The normality test was carried on the residuals. The result of Shapiro-Wilk test and the Lilliefors test are shown in Table 3.1 below:

Table 3.1: Comparison of normality tests for Data

Normality test	Test statistic	p-value	Statistical decision
Shapiro-Wilk	0.9332	0.0121	Reject Ho
Lilliefors	0.1724	0.0018	Reject Ho

Results in Table 3.1 show the normality test of the residuals using Shapiro-Wilk and Lilliefors normality tests. Since the p-value of both methods are each < 0.05, we reject the null hypothesis that the error term of the model is normally distributed. This confirms that the OLS estimation will not yield good estimates, thereby making the application of M-estimation very necessary.

Table 3.2: Comparison of Methods

Function	$\hat{\beta}_i$	Estimates	Standard error	Test	p-value	RMSE
OLS	$\hat{\beta}_0$	3.6785	128.8256	0.0286	0.9774	452.5592
	$\hat{\beta}_1$	0.1892	0.0399	4.7440	0.00002	
	$\hat{\beta}_2$	0.6942	0.4665	1.4880	0.1442	
Talwar	$\hat{\beta}_0$	**155.2714**	**70.4880(45.28)%**	**2.2028**	**0.0331(96.61)%**	**247.6216 (45.28)%**
	$\hat{\beta}_1$	**0.3323**	**0.0218(45.36)%**	**15.2278**	**1.05E-**	
	$\hat{\beta}_2$	**-0.1975**	**0.2553(45.27)%**	**-0.7736**	**0.4435(-207.56)%**	
Andrews	$\hat{\beta}_0$	142.8709	76.7381(40.43)%	1.8618	0.0696(92.88)%	269.5777 (40.43)%
	$\hat{\beta}_1$	0.3322	0.0238(40.35)%	13.9812	2.13E-	
	$\hat{\beta}_2$	-0.1503	0.2779(40.43)%	-0.5407	0.5916(-310.26)%	
Tukey	$\hat{\beta}_0$	142.8748	76.9040(40.30)%	1.8578	0.0702(92.82)%	270.1607 (40.30)%
	$\hat{\beta}_1$	0.3321	0.0238(40.43)%	13.9484	2.31E-	
	$\hat{\beta}_2$	-0.1501	0.2785(40.30)%	-0.5388	0.5929(-311.17)%	
Welsch	$\hat{\beta}_0$	135.7899	77.7119(39.68)%	1.7474	0.0879(91.01)%	272.9987 (39.68)%
	$\hat{\beta}_1$	0.3294	0.0241(39.60)%	13.6913	4.40E-	
	$\hat{\beta}_2$	-0.1267	0.2814(39.68)%	-0.4503	0.6548(-354.09)%	
Cauchy	$\hat{\beta}_0$	93.7697	85.9665(33.27)%	1.0908	0.2816(71.19)%	301.9967 (33.27)%
	$\hat{\beta}_1$	0.2932	0.0266(33.33)%	11.0140	5.81E-	
	$\hat{\beta}_2$	0.0783	0.3113(33.27)%	0.2514	0.8027(-	
Huber	$\hat{\beta}_0$	89.6522	91.4373(29.02)%	0.9805	0.3325(65.98)%	321.2156 (29.02)%
	$\hat{\beta}_1$	0.2763	0.0283(29.07)%	9.7586	2.32E-	
	$\hat{\beta}_2$	0.1349	0.3311(29.02)%	0.4074	0.6858(-375.59)%	
Logistic	$\hat{\beta}_0$	73.2524	93.1541(27.69)%	0.7864	0.4361(55.38)%	327.2467 (27.69)%
	$\hat{\beta}_1$	0.2693	0.0288(27.82)%	9.3371	8.35E-	
	$\hat{\beta}_2$	0.1845	0.3374(27.67)%	0.5468	0.5874(-307.35)%	
Fair	$\hat{\beta}_0$	49.0342	101.8392(20.95)%	0.4815	0.6327(35.27)%	357.7572 (20.95)%
	$\hat{\beta}_1$	0.2464	0.0315(21.05)%	7.8155	1.01E-	
	$\hat{\beta}_2$	0.3029	0.3688(20.94)%	0.8213	0.4161(-188.56)%	

Results in Table 3.2 reveals that the OLS estimates inflates the RMSE and p-value as the RMSE and p-values of the M-estimators are smaller (the smaller, the better) than those of the OLS estimator. Using the OLS estimator as benchmark, Talwar M-estimator improves the RMSE by 45.28%; the standard error of $\hat{\beta}_i$: $\hat{\beta}_0$ by 45.28%, $\hat{\beta}_1$ by 45.36% and $\hat{\beta}_2$ by 45.27%; and the p-value of the test statistic of the $\hat{\beta}_i$: $\hat{\beta}_0$ by 96.61%, $\hat{\beta}_1$ by 100.00% and $\hat{\beta}_2$ by -207.56%.

Andrews M-estimator improves the RMSE by 40.43%; the standard error of $\hat{\beta}_i$: $\hat{\beta}_0$ by 40.43%, $\hat{\beta}_1$ by 40.35% and $\hat{\beta}_2$ by 40.43%; and the p-value of the test statistic of the $\hat{\beta}_i$: $\hat{\beta}_0$ by 92.88%, $\hat{\beta}_1$ by 100.00% and $\hat{\beta}_2$ by -310.26%, Tukey M-estimator improves the RMSE by 40.30%; the standard error of $\hat{\beta}_i$: $\hat{\beta}_0$ by 40.30%, $\hat{\beta}_1$ by 40.43% and $\hat{\beta}_2$ by 4030%; and the p-value of the test statistic of the $\hat{\beta}_i$: $\hat{\beta}_0$ by 92.82%, $\hat{\beta}_1$ by 100.00% and $\hat{\beta}_2$ by -311.17% while other M-estimators had less than 40% improvement on the RMSE. Therefore, the Talwar M-estimator has performed better than other M-estimators and the regression model from the Talwar M-estimator is $\hat{y} = 155.2714 + 0.3323X_1 - 0.1975X_2$. That is predicted Loans/advances = 155.2714 + 0.3323Deposit – 0.1975Lending interest rates with standard errors of 70.4880, 0.0218 and 0.2553; test statistic values of 2.2028, 15.2278 and -0.7736; p-values of 0.0331, 1.05E-18 and 0.4435 respectively.

CONCLUSION AND RECOMMENDATION

We have assessed the m-estimators in predicting the multiple regression model for Nigeria microfinance loans/advances data using the deposit and lending interest rates as explanatory variables. The residuals of the dataset were subjected to the normality test using Shapiro-Wilk and Lilliefors normality tests and the null hypothesis of the error term of the model being normally distributed was rejected at 0.05 significance level by the two normality tests. Comparing the performances of the eight M-estimators in improving the RMSE, standard error of the regression estimates and the p-value of the test statistic for the regression estimates, relative to the OLS estimator, Talwar M-estimator performed better than others with about 45.28% improvement of the RMSE and standard error of the regression estimates, Andrews M-estimator came second with about 40.43% improvement of the RMSE and standard error of the regression estimatesand Tukey M-estimator came third with about 40.30% improvement of the RMSE and standard error of the regression estimates. Estimates of the Talwar M-estimator were used to formulate the robust predictive regression model: $\hat{y} = 155.2714 + 0.3323X_1 - 0.1975X_2$. That is predicted Loans/advances = 155.2714 + 0.3323Deposit – 0.1975Lending interest rates.

Therefore, we recommend the Talwar M-estimator for estimating and predicting such economic variables with outliers and could not satisfy the normality assumption.

REFERENCES

Aggarwal C. C. (2013). *Outlier Analysis*. Kluwer Academic Publishers, Boston.

Ali, A. and Qadir, M.F. (2005). A Modified M-estimator for the Detection of Outliers. *Pakistan Journal of Statistics and Operation Research,* **1,** 49-64.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. Journal of American association, **49**(268), 765-769

Awoyemi, B. O. and Jabar, A. A. (2014). *European Journal of Business and Management*. **6**(12), 131-136

Greene W. H. (2012). *Econometric Analysis,* Seventh Edition, Pearson Education Inc, London.

Holland, P. W., and Welsch, R. E. (1977). Robust Regression Using Iteratively Reweighted Least-Squares. *Communications in Statistics: Theory and Methods,* **A**(6), 813–827.

Inamadar, D., Vengopalan, R., and Padmini K. (2015). Robust ANOVA: An Illustrative Study in Horticultural Crop Research. *Internal Journal of Mathematical computational, Natural and physical Engineering,***9**(2), 85-89.

Jarque, C. M., and Bera, A. K. (1987) A test for normality of observations and regression residuals. International Statistical Review. **55**(2), 163–172.

Johnson R.A and Wichern D.W (2007). *Applied Multivariate Statistical Analysis*. Pearson Education, Inc, London.

Kamile Ş. K., Fatih T. and Türkan E. D. (2012) A Study on Fuzzy Robust Regression and Its Application to Insurance. *Mathematical and Computational Applications*, **17**(3), 223-234.

Lilliefors, H. W. (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*. **62**, 399–402

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*. (46)253, 68–78.

MATLAB version R2008a, (2008), *computer software*, The MathWorks Inc.

Muthukrishnan R. and Radha M. (2010). M-Estimators in Regression Models. *Journal of Mathematics Research*.**2**(4), 23-27.

Paul R. K. and Bhar L. M. (2011). M-estimation in Block designs, *Journal of Indian Society of Agricultural Statistics*, **65**(3), 323-330.

R version 3.3.3 (2017-03-06). The R Foundation for Statistical Computing Platform.

Razali, N. M. and Wah, Y. B. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. *Journal of Statistical Modeling and Analytics*. **2** (1): 21–33.

Rousseeuw, P. J.; Leroy, A. M. (1987), *Robust Regression and Outlier Detection*, John Wiley & Sons, Inc., New York.

Shapiro, S. S.; Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika. 52 (3–4): 591–611.

Shawn P. N. (1996) *Robust Analysis of M-estimators of Nonlinear Models,* Unpublished M.Sc. thesis.

Shi-Woei L. (2006). A Comparative Study of Robust Estimates in Regression. *Proceedings of the 11th Annual Conference of Asia Pacific Decision Sciences Institute,* 569-572.

Stuart C. (2011). Robust Regression. An Unpublished Project.