Pertanika J. Sci. & Technol. 6(1): 23 - 35 (1998)
ISSN: 0128-7680

Robust Estimation of a Linearized Nonlinear Regression Model with Heteroscedastic Errors:A Simulation Study

Habshah Midi
Mathematics Department
Faculty of Science and Environmental Studies
Universiti Putra Malaysia
43400 UPM Serdang, Selangor, Malaysia

Received: 24 Disember 1996

ABSTRAK

Suatu kajian simulasi telah dijalankan untuk memeriksa keteguhan penganggar Kuasadua Terkecil Biasa Dilinearkan (LOLS), penganggar Kuasadua Terkecil Teritlak Terjelma (TGLS), penganggar Kuasadua Terkecil Berpemberat Terjelma (LRLS) dan penganggar Kuasadua Terkecil Berpemberat Terjelma Dilinearkan (TLRLS). Kaedah TLRLS adalah pengubahsuaian kaedah Kuasadua Terkecil Berpemberat (RLS) berasaskan kepada kaedah Median Kuasadua Terkecil (LMS). Kajian berangka menunjukkan bahawa penganggar LOLS, TGLS dan LRLS tidak cukup teguh apabila peratusan titik terpencil di dalam data meningkat. Ini bermakna ketiga-tiga penganggar tersebut tidak mempunyai titik musnah yang tinggi. Keputusan kajian menunjukkan bahawa Kaedah TLRLS mempunyai titik musnah yang tinggi berbanding dengan tiga kaedah yang lain.

ABSTRACT

A simulation study is used to examine the robustness of some estimators on a linearized nonlinear regression model with heteroscedastic errors, namely the Linearized Ordinary Least Squares (LOLS), Transformed Generalized Least Squares (TGLS), Linearized Reweighted Least Squares (LRLS) and Transformed Linearized Reweighted Least Squares (TLRLS). The latter is a modification of Reweighted Least Squares (RLS) based on Least Median of Squares (LMS). The empirical evidence shows that the first three estimators are not sufficiently robust when the percentage of outliers in the data increases. That is, they do not have a high breakdown point. On the other hand, the modified estimator (TLRLS) has a higher breakdown point than the other three estimators.

Keywords: **breakdown point, outliers, generalized least squares, heteroscedasticity, least median of squares, linearized model, log-normal distribution, reweighted least squares**

INTRODUCTION

The difficulty in computations involving a complex nonlinear structure can be solved by some transformation of the data so that it appears to follow a new model which is less complicated. A common type of transformation is to linearize the nonlinear model by means of log transformation. However, it is necessary to be aware of the consequences of using log transformation of a nonlinear model with additive error terms. Such a transformation may alter the error structure unless the error terms are multiplicative. While transformation may solve the computational problem, it may produce another problem in terms of violation of error structure. Thus, a clear advantage of linearization as far as the properties of the estimators are concerned is not gained.

Consider the general nonlinear model with multiplicative error terms

$$Y_i = f(X_i, \beta)\varepsilon_i \tag{1.1}$$

where $\beta^T = (\beta_1, \beta_2, ..., \beta_p)$ is the vector of parameters to be estimated. The regressors X_i i = 1,2 ,..., are, in general, p dimensional vectors whose values are assumed known, and the errors ε_i = 1,2, ...,n are i.i.d log-normal random variables, i.e. $\varepsilon \sim \Lambda(0, \sigma^2\Omega)$ where Λ denotes a log-normal distribution and Ω is a n x n scale matrix.

In many applications, errors which are heteroscedastic, multiplicative and not normally distributed may be encountered. Often the nature of heteroscedasticitys not known. Heteroscedasticity is frequently modelled as a function of covariates, mean response or other functional forms. By erroneously assuming that the model has an error structure which is additive, normally distributed, and Ω is an identity matrix, the least square estimator $\hat{\beta}$ is found by minimizing the sum of squares

$$S(\beta) = \sum_i (Y_i - f(X_i, \beta))^2 \tag{1.2}$$

Using the Gauss-Newton Method (see Ratkowsky 1983), $\hat{\beta}$ are obtained by an iterative process. At the (k + 1)th iteration, we have

$$\hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} + \left[J^T(\hat{\beta}^{(k)})\, J(\hat{\beta}^{(k)})\right]^{-1} J^T(\hat{\beta}^{(k)}) \left[Y - f(\hat{\beta}^{(k)})\right] \tag{1.3}$$

where J (β) is the n x p Jacobian matrix, i.e.

$$J(\beta^*) = \begin{bmatrix} \frac{\partial f(X_1,\beta)}{\partial \beta_1} & \frac{\partial f(X_1,\beta)}{\partial \beta_2} & \mathrm{L} & \frac{\partial f(X_1,\beta)}{\partial \beta_p} \\ \mathrm{M} & & & \\ \frac{\partial f(X_n,\beta)}{\partial \beta_1} & \frac{\partial f(X_n,\beta)}{\partial \beta_2} & \mathrm{L} & \frac{\partial f(X_n,\beta)}{\partial \beta_p} \end{bmatrix}_{\beta = \beta^*}$$

Starting with initial values for $\hat{\beta}$ at k = 0, the process continues until convergence, which occurs when $\left|\beta^{k+1} - \beta^{(k)}\right|$ is smaller than some preselected small quantity. $\hat{\beta}$ obtained from (1.3) is not asymptotically optimal when it violates the assumption of homoscedasticity. Moreover, equation (1.2) is not valid when the error structure is really multiplicative instead of addtive in nature. Consequently, a 'bad' estimator is obtained as a result of employing an incorrect formula on the erroneous assumptions.

There are, however several nonlinear models which can be made linear by some appropriate transformation. It should be noted, however, that the linearization of the non-linear models may require the transformed error terms to be independent and normally distributed with mean 0 and constant variance, σ^2.

Model (1.1) can be linearized by taking a natural logarithm that is,

$$\text{Ln } Y = \text{Ln } f(X,\beta) + \text{Ln } \varepsilon$$

or it can be written as

$$Y^* = X^*\beta^* + \varepsilon^* \tag{1.4}$$

where

$$Y^* = \text{Ln } Y,\ X^*\beta^* = \text{Ln } f(X,\beta),\ \varepsilon^* = \text{Ln } \varepsilon_i$$

Ln ε_i is normally distributed, i.e. Ln $\varepsilon_i \sim N(0,\sigma^2\Omega)$ since ε_i in model (1.1) follows a log-normal distribution, i.e. $\varepsilon_i \sim \Lambda(0,\sigma^2\Omega)$. The transformation enables the use of the standard regression method. It is very important to note that when transformed models are employed, the estimators obtained by least squares have the least squares properties with respect to the transformed observations, not the original ones, However, the ordinary Least Square Estimator of the linearized model (LOLS) in (1.4),

$$\hat{\beta}^* = (X^{*T}X^*)^{-1}X^{*T}Y^* \tag{1.5}$$

is not an optimal estimator because Ω is not an identity matrix. Let $\Omega = PP^T$ where P is an nxn diagonal matrix, i.e.

$$P = \text{diag}\left\{1/\sqrt{p(x_1)},\ 1/\sqrt{p(x_2)},\ \text{L},\ 1/\sqrt{p(x_n)},\right\}$$

The above problem of heteroscedasticity can be removed by means of a suitable transformation.

The transformed model is defined as

$$P^TY^* = P^TX^*\beta^* + P^T\varepsilon^* \tag{1.6}$$

to give

$$Y^{**} = X^{**}\beta^{*} + \varepsilon^{**}$$

where

$$Y^{**} = P^{T}Y^{*},\ X^{**} = P^{T}X^{*} \text{ and } \varepsilon^{**} = P^{T}\varepsilon^{*}$$

This leads to

$$\hat{\beta}^{*} = (X^{**T}X^{**})^{-1}X^{**T}Y^{**}, \tag{1.7}$$

which is the best linear unbiased estimator with respect to the transformed model (1.4), and is called the generalized least square estimator on the transformed variables (TGLS).

However, the drawback of the LOLS and the TGLS estimator is that they may be affected by outliers, which are observations which significantly deviating from the underlying model governing the bulk of the data. Hampel (1971) pointed out that even one single outlier can have an arbitrarily large effect on the estimates. In this connection, he introduced the so-called breakdown point, ε^{*} as the smallest percentage of contaminated data that can cause an estimator to take an arbitrarily large values. The robustness of each estimator depends on the value of ε^{*}. An estimator becomes more robust as the value of ε^{*} increases. A better approach is to consider a robust method which is much less influenced by the outliers.

Several works on robust estimation have been proposed in the literature. Among them are Barrodale and Roberts (1973) and Armstrong and Kung (1978) who introduce L_1-norm estimators. Huber (1973) proposed M-estimation and this was modified further by Krasker and Welsch (1982) who introduced bounded-influence regression estimator. None of these estimators achieves a breakdown point $\varepsilon^{*} = 30\%$ in the case of simple regression. Rousseeuw (1984) introduced the most robust regression estimator with the highest possible breakdown point, i.e. $\varepsilon^{*} = 50\%$. This method is known as the Reweighted Least Squares (RLS) based on the Least Median of Squares (LMS).

There have been numerous studies concerning the estimation of a linear model with heteroscedastic errors in the literature (Box and Hill 1974; Carroll and Ruppert 1982; Cohen *et al* 1993). However, none of the studies has taken into consideration the estimation of a high breakdown linear regression with heteroscedastic errors.

The focus of this study was to investigate the effect of outliers on the estimates of the linearized regression model when the errors are heteroscedastic. Consider a nonlinear model with multiplicative error terms, log-normally distributed and heteroscedastic; i.e,

$$y_i = \beta_0 \beta_1^{x_i} \varepsilon_i \tag{1.8}$$

i.e. $\varepsilon_i \sim \Lambda(0,x^2)$. The mean and the variance of ε_i are $E(\varepsilon_i) = e^{x_i^2/2}$ and $V(\varepsilon_i) = e^{2x_i^2} - e^{x_i^2}$ respectively (Johnson and Kotz 1970).

By taking logarithms, model (1.8) can be written as a linear model:

$$\text{Ln } y_i = \text{Ln } \beta_0 + x_i \text{Ln } \beta_1 + \text{Ln } \varepsilon_i \qquad (1.9)$$

where

$$\text{Ln } \varepsilon_i \sim N(0, x_i^2)$$

The linearized model (1.9) is an additive model with hetercoscedastic errors. It is very important to note that when heteroscedasticity prevails but other conditions of model (1.9) are satisfied, the estimators obtained by ordinary least squares methods are still unbiased, but they are no longer minimum variance unbiased estimators. The variance of the error terms is proportional to $x^2, V(\varepsilon_i) = kx_i^2, k \geq 1$. For simplicity, we take $k = 1$. The appropriate transformations to obtain minimum variance unbiased estimators for model (1.9) are

$$y' = \frac{\text{Ln } y}{x}, \qquad x' = \frac{1}{x}$$

Model (1.9) then becomes

$$\frac{\text{Ln } y_i}{x_i} = \frac{\text{Ln } \beta_0}{x_i} + \text{In } \beta_1 + \frac{\text{Ln } \varepsilon_i}{x_i} \qquad (1.10)$$

A slight modification of the RLS method based on LMS is proposed. The values of the variables x and y are then substituted with the values of the transformed linearized variables $1/x$ and $\frac{\text{Ln } y}{x}$, respectively.

The modified method, known as the Transformed Linearized Reweighted Least Squares (TLRLS) based on the LMS is performed by implementing the RLS method to the transformed linearized variables. We would expect the modified method to be more robust than the Linearized Ordinary Least Squares (LOLS), the Transformed Generalized Least Squares (TGLS) and the Linearized Reweighted Least Squares (LRLS) procedures would maintain a breakdown point as high as 50%.

ROBUST REGRESSION ESTIMATOR

The Least Median of Squares (LMS) estimator (Rousseeuw 1984; Rousseeuw and Leroy 1987) is defined as the value that minimizes

$$\underset{i}{\text{median}} \; r_i^2 \qquad (2.1)$$

where

$$r_i = y_i - x^T \hat{\beta}, \qquad i = 1, 2, \ldots n$$

for model (1.4).

This estimator is very robust with respect to outliers in y as well as outliers in x and also has the highest possible breakdown point, ε^* 50%.

The estimated variance, $\hat{\sigma}$ is given by

$$\hat{\sigma} = \sqrt{\frac{\sum w_i r_i^2}{\sum w_i - p}} \tag{2.2}$$

where

$$w_i = \begin{cases} 1 & \text{if } |r_i / s| \leq 2.5 \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

$$s = 1.4826 \left\{1 + 5/(n-p)\right\} \sqrt{\underset{i}{\text{med}}\, r_i^2} \tag{2.4}$$

In matrix notation, w_i may be written as n x n diagonal matrix:

$$W = \text{diag}\ \{1,0,1,1,...,\underset{\uparrow}{0}, \ .....\ 0,1\}$$

the j^{th} outlier in the data set

Unfortunately, the LMS performs poorly (inefficient) when the errors are actually normally distributed. In order to impove the LMS estimator, Rousseeuw and Leroy (1987) introduced a method called Reweighted Least Squares (RLS) regression based on LMS, which is given by:

$$\underset{\beta}{\text{minimize}} \sum w_i r_i^2 \tag{2.5}$$

The weights w_i are determined from the LMS solutions (2.1), but with the estimated variance of (2.2) instead of (2.4). From (1.4), the Linearized Reweighted Least Square (LRLS) model can be written as: $WY = WX^* + W\varepsilon^*$.

Therefore, the LRLS estimator is given by

$$\hat{\beta} = (X^{*T} WX^*)\ X^{*T}\ WY^* \tag{2.6}$$

From (1.6), the Transformed Linearized Reweighted Least Squares (TLRLS) model then becomes

$$WP^TY = WP^TX^*\beta^* + WP^T\varepsilon^* \tag{2.7}$$

which gives

$$Y^{***} = X^{***}\beta^{*} + \varepsilon^{***}$$

with

$$Y^{***} = WP^{T}Y^{*} \qquad X^{***} = WP^{T}X^{*} \text{ and } \varepsilon^{***} = WP^{T}\varepsilon^{*}$$

The TLRLS estimator is now defined as

$$\hat{\beta}^{*} = (X^{*T}PWP^{T}X^{*})^{-1}X^{*T}\,P\,W^{T}\,P^{T}Y \tag{2.8}$$

Since the same method is used to estimate $\hat{\beta}$, the TLRLS estimates are expected to have the same breakdown point, i.e. $\varepsilon^{*} = 50\%$, as possessed by the RLS estimates in the case of linear model with homoscedastic errors.

SIMULATION STUDY

To illustrate the breakdown properties of the estimator in (1.5), (2.6), (1.7), (2.8), a simulation study was carried out as follows: 100 'good' observations were generated according to nonlinear relation $Y_i = (2)(1.5)^{x_i}\varepsilon_i$ where x_i is uniformly distributed on [1,7]. ε_i is drawn from log-normal distribution, i.e. $\varepsilon_i \sim \Lambda(0,kx_1^2)$ where k = 1 . The above relation was then linearized to produce a linear relation:

$\text{Ln } y_i = \text{Ln } 2 + \text{Ln}(1.5)\ x_i + \text{Ln } \varepsilon_i$ where $\text{Ln } \varepsilon_i \sim N(0,kx_i^2)$ where k = 1. The true values of β_0 and β_1 are Ln 2 ≈ 0.693 and Ln(1.5) ≈ 0.405, respectively. The procedures (1.5), (2.6), (1.7), (2.8) were then applied to these data. The results are as follows:

$$\hat{\beta}_{0(LOLS)} = 1.352,\quad \hat{\beta}_{0(LRLS)} = 1.388,\quad \hat{\beta}_{0(TGLS)} = 0.855,$$
$$\hat{\beta}_{0(TLRLS)} = 0.614,\quad \hat{\beta}_{1(LOLS)} = 0.158,\quad \hat{\beta}_{1(LRLS)} = 0.023,$$
$$\hat{\beta}_{1(TGLS)} = 0.297,\quad \hat{\beta}_{1(TLRLS)} = 0.431.$$

Since the data were uncontaminated, the above estimates were quite close to the true values, especially the TLRLS and the TGLS estimates.

Then contamination of the data was commenced. At each step, one 'good' observation was deleted and replaced with a bad data point. The contaminated data points were generated according to the non-linear relation, $Y_i = (20)(1.5)^{x_i}\varepsilon_i$ where x_i is uniformly distributed on [0,1]. ε_i is drawn from $\Lambda(20,kx_i^2)$ where k = 1. The above relation was also linearized to produce $\text{Ln } y_i = \text{Ln } 20 + x_i\text{Ln}(1.5) + \text{Ln}\varepsilon_i$ where Ln $\text{Ln } \varepsilon_i \sim N(20,kx_i^2)$.

The above process was repeated until only 50 'good' observations remained. Table 1 presents the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ for the four methods, when good observations are replaced by certain percentages of outliers. OT noted in Table 1 indicates outlier.

The breakdown plots that illustrate the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ as a function of the percentage of outliers are shown in Fig. 1, 2. These figures show that the

TABLE 1
The values of $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ for n = 100

	$\hat{\beta}_0$ Method				$\hat{\beta}_1$ Method			
% OT	LOLS	LRLS	TGLS	TLRLS	LOLS	LRLS	TGLS	TLRLS
0	1.352	1.388	0.855	0.614	0.158	0.023	0.297	0.431
10	9.864	0.508	23.957	0.463	–1.515	0.517	–6.620	0.532
20	14.569	0.506	23.114	0.398	–2.396	0.538	–5.555	0.569
30	17.928	19.460	23.085	0.464	–3.032	–3.299	–4.717	0.607
40	20.136	23.930	23.081	0.250	–3.558	–3.914	–3.960	0.654
45	21.834	24.678	23.082	0.108	–3.695	–4.121	–3.713	0.711
50	21.691	24.169	23.057	23.675	–3.879	–2.750	–3.300	–3.373

estimated values $\hat{\beta}_0$ and $\hat{\beta}_1$, which are based on the LOLS and TGLS were immediately affectd by the outliers. As can be expected, when there is no outlier, the TGLS estimator performs better than the LOLS and the LRLS estimators. But as the percentage of outliers increases, the TGLS estimates move away from the true values drastically, followed by the LOLS and the LRLS estimates. Furthermore, the increase in the percentage of outliers from 0% up to 45% changed not only in the values but also the signs of $\hat{\beta}_1$ of TGLS, LOLS and LRLS, i.e. from positive to negative values. The results also point out that the LRLS estimates can tolerate slightly over 20% of outliers. The values of the TLRLS estimates seem to be consistent, before breaking down at 50% of outliers. This implies that The TGLS and the LOLS estimates break down first, followed by the LRLS and TLRLS estimates.

The breakdown properties of these estimates were investigated further by considering three samples of size 20, 50 and 100 observations. Simulation studies were carried out in the manner described earlier. Tables 2, 3 and 4 show the summary statistics, such as bias, standard error (SE) and root mean square error (RMSE) of the estimators. These results are very useful in assessing the breakdown properties of the estimates.

These tables show that the TLRLS estimates are almost as good as the LOLS estimates in the normal error situation. As was to be expected, the TGLS estimates give the best results followed by the TLRLS, LOLS and LRLS

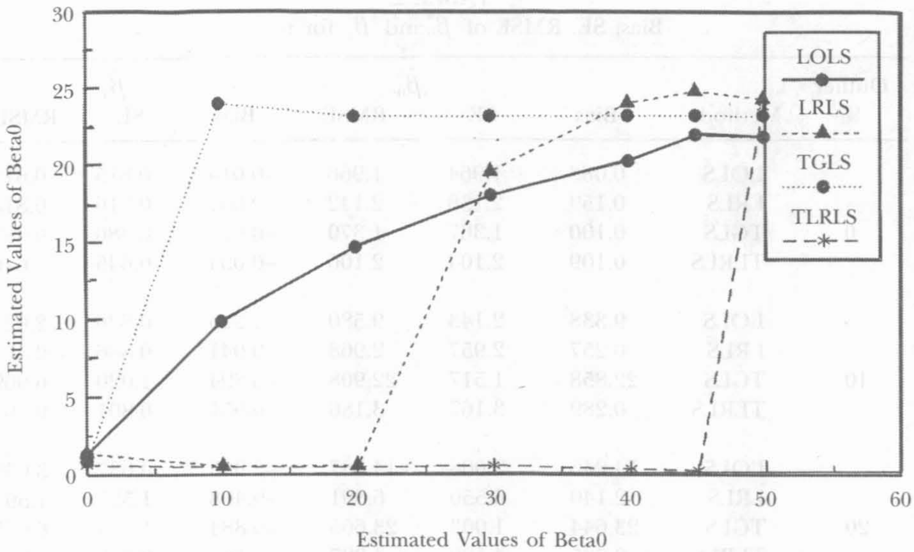

Fig. 1. The breakdown plots of β_o

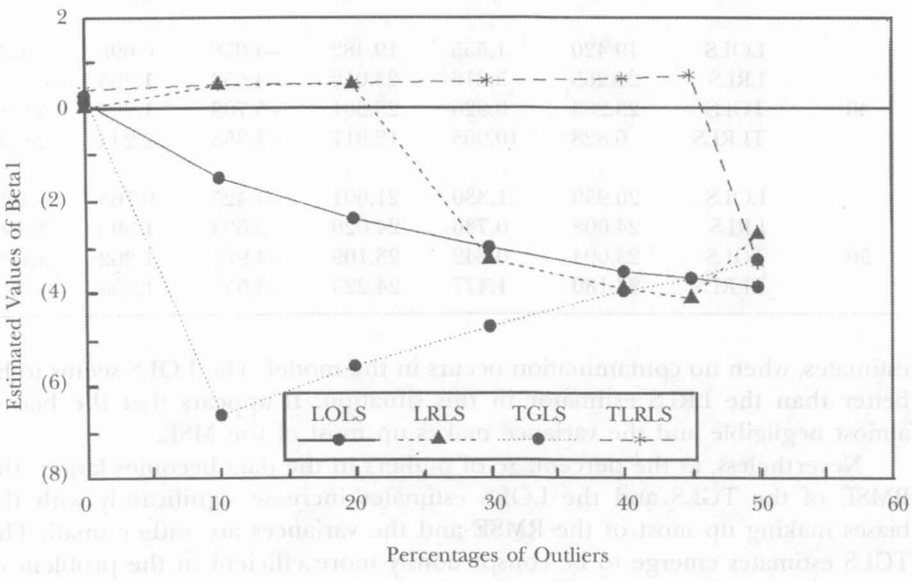

Fig. 2. The breakdown plots of β_1

TABLE 2
Bias, SE, RMSE of $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ for n = 20

Outlier %	Method	$\hat{\beta}_0^*$ Bias	SE	RMSE	$\hat{\beta}_1^*$ Bias	SE	RMSE
	LOLS	0.083	1.964	1.966	–0.014	0.615	0.615
	LRLS	0.159	2.136	2.142	–0.051	0.816	0.818
0	TGLS	0.100	1.367	1.370	–0.021	0.480	0.480
	TLRLS	0.109	2.104	2.106	–0.034	0.645	0.646
	LOLS	9.338	2.143	9.580	–1.925	0.626	2.024
	LRLS	0.257	2.957	2.968	–0.041	0.886	0.881
10	TGLS	22.858	1.517	22.908	–6.894	1.020	6.969
	TLRLS	0.289	3.167	3.180	–0.051	0.801	0.801
	LOLS	14.246	2.008	14.387	–2.965	0.637	3.032
	LRLS	2.140	6.550	6.891	–0.468	1.527	1.597
20	TGLS	23.644	1.003	23.665	–6.884	1.094	6.971
	TLRLS	0.635	4.180	4.227	–0.136	0.957	0.967
	LOLS	17.339	1.811	17.434	–3.622	0.655	3.681
	LRLS	10.711	11.488	15.706	–2.320	2.656	3.526
30	TGLS	23.482	0.984	23.503	–6.358	1.198	6.470
	TLRLS	2.018	6.963	7.249	–0.416	1.509	1.565
	LOLS	19.420	1.555	19.482	–4.078	0.698	4.137
	LRLS	22.265	5.816	23.012	–4.653	1.795	4.987
40	TGLS	23.283	0.926	23.301	–5.709	1.227	5.839
	TLRLS	6.828	10.965	12.917	–1.355	2.244	2.621
	LOLS	20.959	1.330	21.001	–4.425	0.765	4.490
	LRLS	24.008	0.736	24.020	–3.693	1.390	3.946
50	TGLS	23.094	0.842	23.109	–4.961	1.262	5.119
	TLRLS	24.180	1.477	24.225	–4.551	1.556	4.809

estimates, when no contamination occurs in the model. The LOLS seems to be better than the LRLS estimator in this situation. It appears that the bias is almost negligible and the variance makes up most of the MSE.

Nevertheless, as the percentage of outliers in the data becomes larger, the RMSE of the TGLS and the LOLS estimates increase significantly with the biases making up most of the RMSE and the variances are rather small. The TGLS estimates emerge to be conspicuously more efficient in the problem of heteroscedasticity with no contamination in the model. The LRLS estimator can be considered a good alternative for the case with slightly above 20% outliers.

On the other hand, the RMSE of the TLRLS estimates is consistently small as the percentage of outliers becomes larger. Nonetheless, its values changed

TABLE 3
Bias, SE, RMSE of $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ for n = 50

Outlier %	Method	$\hat{\beta}_0^*$ Bias	SE	RMSE	$\hat{\beta}_1^*$ Bias	SE	RMSE
	LOLS	0.073	1.222	1.224	–0.008	0.386	0.386
	LRLS	0.126	1.375	1.381	–0.031	0.5430	0.543
0	TGLS	0.015	0.783	0.783	0.008	0.285	0.285
	TLRLS	0.032	0.978	0.978	0.005	0.339	0.339
	LOLS	9.164	1.275	9.253	–1.887	0.382	1.926
	LRLS	0.057	1.666	1.667	0.005	0.559	0.559
10	TGLS	22.913	0.578	22.921	–6.896	0.653	6.927
	TLRLS	0.041	1.055	1.056	0.000	0.357	0.357
	LOLS	14.076	1.215	14.129	–2.907	0.372	2.930
	LRLS	0.667	3.852	3.909	–0.127	0.905	0.914
20	TGLS	23.064	0.628	23.073	–6.480	0.766	6.626
	TLRLS	0.127	1.564	1.5669	–0.017	0.444	0.444
	LOLS	17.130	1.106	17.166	–3.543	0.385	3.564
	LRLS	9.210	10.914	14.281	–1.892	2.321	2.995
30	TGLS	22.896	0.579	22.903	–5.828	0.817	5.885
	TLRLS	0.423	3.165	3.193	–0.076	0.706	0.710
	LOLS	19.241	0.954	19.265	–3.989	0.389	4.008
	LRLS	23.713	2.050	23.801	–4.727	0.969	4.825
40	TGLS	22.738	0.457	22.743	–5.090	0.774	5.149
	TLRLS	2.416	7.278	7.669	–0.464	1.442	1.515
	LOLS	20.805	0.835	20.821	–4.320	0.432	4.342
	LRLS	23.810	0.534	23.816	–3.210	1.103	3.395
50	TGLS	22.618	0.349	22.621	–4.307	0.738	4.370
	TLRLS	23.949	1.153	23.976	–4.212	1.210	4.383

dramatically at 50% of outliers. The results seem to be consistent in all 500 trials and for each sample, size n = 20, 50, 100. The RMSE of the TLRLS estimates are relatively smaller than the other three estimates. Summarizing the findings from Tables 2, 3 and 4, it can be concluded that the TGLS and the LOLS estimates break down first and are then followed by the LRLS and TLRLS estimates. Thus, it can be concluded that the TLRLS estimates are the best method for handling the problem of outliers in the linearized model when the error terms are heteroscedastic.

TABLE 4
Bias, SE, RMSE of $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ for n = 100

Outlier %	Method	$\hat{\beta}_0^*$ Bias	SE	RMSE	$\hat{\beta}_1^*$ Bias	SE	RMSE
	LOLS	0.027	0.822	0.823	−0.002	0.264	0.264
	LRLS	0.009	1.009	1.009	0.005	0.411	0.411
0	TGLS	0.009	0.506	0.506	0.005	0.193	1.193
	TLRLS	0.040	0.613	0.614	−0.004	0.227	0.227
	LOLS	9.069	0.829	9.107	−1.863	0.265	1.882
	LRLS	−0.026	0.949	0.949	0.021	0.373	0.374
10	TGLS	22.836	0.368	22.839	−6.838	0.432	6.852
	TLRLS	−0.002	0.576	0.576	0.012	0.222	0.223
	LOLS	13.971	0.794	13.994	−2.874	0.269	2.887
	LRLS	0.045	1.306	1.307	0.000	0.413	0.413
20	TGLS	22.791	0.477	22.796	−6.265	0.512	6.286
	TLRLS	0.009	0.595	0.5959	0.010	0.230	0.230
	LOLS	17.059	0.720	17.074	−3.508	0.273	3.518
	LRLS	7.639	10.244	12.779	−1.569	2.146	2.658
30	TGLS	22.632	0.359	22.634	−5.533	0.537	5.559
	TLRLS	0.026	0.624	0.624	0.010	0.239	0.239
	LOLS	19.180	0.618	19.190	−3.941	0.281	3.951
	LRLS	23.787	0.612	23.795	−4.646	0.608	4.686
40	TGLS	22.551	0.275	22.553	−4.805	0.515	4.833
	TLRLS	0.405	3.179	3.205	−0.973	0.675	0.679
	LOLS	20.720	0.530	20.727	−4.246	0.304	4.257
	LRLS	23.643	0.478	23.648	−2.853	1.014	3.028
50	TGLS	22.489	0.210	22.490	−4.055	0.083	4.084
	TLRLS	23.787	0.991	23.808	−4.071	1.157	4.232

CONCLUSION

The TGLS estimator of the linearized model is a better choice than the other three estimators in eliminating the problem of heteroscedasticity. Nevertheless, its performance was inferior to LOLS, LRLS and TLRLS estimators when contamination occurred in the data. The LRLS estimator has a breakdown point slightly over 20% with the presence of outliers. It cannot provide a robust alternative to the TGLS and LOLS since it is not sufficiently robust when the percentage of outliers increases. The simulation studies clearly shows that the TLRLS estimator is definitely the best because it is able to withstand a large amount of outliers and has a highest breakdown point (up to 50% of outliers). Hence, it should provide a robust alternative to the well -known TGLS estimators.

REFERENCES

ARMSTRONG, R.D. and M.T. KUNG. 1978. Least absolute values estimates for a simple linear regression problem. *Applied Statistics* **27**: 363-366.

BARRODALE, I and F.D.K. ROBERTS. 1973. An improved algorithm for discrete L-1 linear approximation. *Siam Journal of Numerical Analysis* **10**: 839-848.

BOX, G.E.P. and W.J. HILL. 1974. Correcting inhomogeneity of variance with power transformation weighting. *Technometrics* **16**: 385-389.

CARROLL, R.J. and D. Ruppert. 1982. Robust estimation in heteroscedastic linear models. *Ann. Stat.* **10**: 429-441.

COHEN, M., S.R. DALAL and J.W. TUKEY. 1993. Robust smoothly heterogeneous variance regression. *Appl. Statist.* **42(2)**: 339-353.

HAMPEL, F.R. 1971. A general qualitative definition of robustness. *Annals of Mathematical Statistics* **42**: 1887-1896.

HUBER, P.J. 1973. Robust estimation of a location parameter. *Annals of Mathematical Statistics* **35**: 73-101.

JOHNSON, N.L. and S. KOTZ. 1970. *Distributions in Statistics-Continuous Univariate Distributions.* Boston : Houghton Mifflin.

KRASKER, W.S. and R.E. WELSCH. 1982. Efficient bounded-influence regression estimation. *Journal of the American Statistical Association.* **77**: 595-605.

RATKOWSKY, D.A. 1983. *Non-Linear Regression Modelling.* New York: Marcel Dekker.

ROUSSEEUW, P.J. 1984. Least median of squares regression. Journal of the *American Statistical Association.* **79**: 871-879.

ROUSSEEUW, P.J. and A. LEROY. 1987. *Robust Regression* and *Outlier Detection.* New York: Wiley.