# Weighted fast-trimmed likelihood estimator for mixture regression models

Hassan S. Uraibi[1]* and Mohammed Qasim Waheed[2]

[1]Department of Statistics, University of Al-Qadisiyah, Al Diwaniyah, Iraq, [2]College of Physical Education and Sport Sciences, University of Al-Qadisiyah, Al Diwaniyah, Iraq

The fast-trimmed likelihood estimate is a robust method to estimate the parameters of a mixture regression model. However, this method is vulnerable to the presence of bad leverage points, which are outliers in the direction of independent variables. To address this issue, we propose the weighted fast-trimmed likelihood estimate to mitigate the impact of leverage points. The proposed method applies the weights of the minimum covariance determinant to the rows suspected of containing leverage points. Notably, both real data and simulation studies were considered to determine the efficiency of the proposed method compared to the previous methods. The results reveal that the weighted fast-trimmed estimate method is more robust and reliable than the fast-trimmed likelihood estimate and the expectation−maximization (EM) methods, particularly in cases with small sample sizes.

KEYWORDS

leverage point, mixture regression, RMVN, TLE, EM, outlier

## Introduction

The mixture regression model has been widely used in various scientific fields, such as econometrics, engineering, biology, and others, due to its ability to capture the relationship between independent and dependent variables. As noted by Quandt (1) and Quandt and Ramsay (2), datasets may exhibit multiple data patterns, and variables often cluster according to these patterns. These patterns can be represented by unknown groups of latent variables, leading to the use of mixture regression models. In other words, the regression probabilistic model, in terms of the latent class variable Z, where Z = k, can be expressed as follows:

$$Y = x'\beta_k + e_k, k = 1, 2, \ldots, g, \quad (1)$$

where $\mathbf{x}$ is the design matrix of $(p+1)$ independent variables with constant, and $\beta_k$ is the regression parameters of $g$, which is the number of homogeneous sub-populations (groups). Moreover, $e_k$ is the error term that has to be independent of $\mathbf{x}$ with density $f_k(.)$, which is distributed $N(0, \sigma_k^2)$, and $Y$ is the dependent variable, which has the same distribution of $e_k$ with different parameters.

The maximum likelihood estimation (MLE) method, widely regarded as one of the best methods to estimate regression model parameters when the distribution of errors is normal, becomes challenging when this assumption is violated. Moreover, MLE lacks robustness in the presence of outliers. Alternative methods should prioritize both ease of computation and robustness, such as the M-estimate (3) and the least trimmed squares method (4). These robust approaches address contaminated error distributions through two key procedures: first, fitting a

parametric model to the majority distribution and removing errors following a different distribution; second, assigning reduced weights to outliers so they approximate a normal distribution. In some cases, the random errors are highly heterogeneous, with at least 50% of the errors lacking homogeneity (5).

Dempster et al. (6) introduced the expectation–maximization (EM) algorithm, which simplifies the problem by transforming it into a relatively straightforward single-component MLE. The EM algorithm is robust against outliers and easy to calculate. This issue of robustness in mixture regression models has garnered significant attention in statistical research. Initial efforts to address the problem of outliers in mixture models, particularly within the location-scale family of distributions, were made by Peel and McLachlan (7) and Hennig (8). In a different approach, Markatou (9), Neykov et al. (5), and Shen et al. (10) sought to mitigate the influence of outliers in mixture regression models by assigning reduced weights to individual data points.

Alternatively, Hadi and Luceño (11) and Vandev and Neykov (12) defined the weighted trimmed likelihood estimator (WTLE):

$$\min_{\beta \in \Theta^p} \sum_{i=1}^{h} w_{v(i)} f\left(y_{v(i)}; \beta\right) = \min_{I \in I_h} \min_{\beta \in \Theta^p} \sum_{i=1}^{h} w_i \left[-\log \varphi\left(y_i; \beta\right)\right], \quad (2)$$

where $f\left(y_{v(1)}; \beta\right) \le f\left(y_{v(2)}; \beta\right) \le \cdots \le f\left(y_{v(n)}; \beta\right)$ for fixed $\beta$, $I_h = C_h^n$ is the all-possible subsets of size $h$ that can be fitted by MLE and randomly selected from $n$ observations, $\varphi\left(y_i; \beta\right)$ is the probability density function. $\hat{\beta}_{WTLE}$ is given the subset with the minimal negative log-likelihood. However, Equation 2 accommodates for the number of estimators, MLE when $n = h$, TLE when $w_{v(i)} = 1, i = 1, 2, \ldots, n$, the median likelihood estimator (MedLE) when $w_{v(h)} = 1$ and $w_{v(i)} = 0$ for $i \ne h$. Both MedLE and trimmed likelihood estimator (TLE) can coincide with other estimators depending on the definition of the density function $\varphi\left(y_i; \beta\right)$; when it is multivariate normal, the MedLE and TLE correspond with minimum volume ellipsoid (MVE) and minimum covariance determinant (MCD) estimators (4). When it is normal density, the MedLE and TLE correspond to the least median of squares (LMS) and least trimmed squares (LTS). Neykov et al. (5) highlighted that WTLE is not a practical choice for large datasets and introduced the fast-trimmed likelihood estimate (FTLE) algorithm as a developed version of the approach originally proposed by Müller and Neykov (13). FTLE reduces Fast-LTS and Fast-MCD (14) in computing mixture regression coefficients in two steps, trial and refinement, respectively.

However, FTLE is not resistant in the presence of bad leverage points, which are outliers in the direction of independent variables. Garcia-Escudero et al. (15) proposed an adaptive procedure to identify the real cluster regressions without the influence of bad leverage points. Despite the increase in the second trimming in the breakdown point, it does not necessarily have high efficiency due to the high percentage of trimming, which reduces the degrees of freedom and, consequently, results in a loss of efficiency. Uraibi (16) employed the weights that are derived from re-weighted multivariate normal (RMVN), which is a robust location and scatter matrix proposed by Olive and Hawkins (17), instead of removing the outliers. This study suggests weighting the leverage points in the design matrix X using RMVN prior to using it with the TLE algorithm that deals with outliers in the y-direction.

# Weighted fast-trimmed likelihood estimator

The FTLE method is not a feasible choice with moderate and small sizes due to the twice trimming for outliers and leverage points. This procedure leads to a loss of more information and consequently reduces the degrees of freedom. Thus, assigning down weights for leverage points instead of removing them might improve the performance of the TLE method. Moreover, identifying the percentage of outliers will contribute to the fast computation of TLE in the framework of the C-steps concentrated algorithm.

Uraibi (16) noted that the Fast MVE and MCD methods proposed by Rousseeuw and Driessen (14) are time-consuming and susceptible to masking and swamping phenomena when high leverage points are present. Consequently, using robust Mahalanobis distance based on MVE or MCD to identify leverage points may not be effective, as it can lead to inaccurate diagnostics regarding the exact number of leverage points. Uraibi (16) suggested using RMVN location and scale estimators (17) with Mahalanobis distance instead of MVE and MCD, as RMVN is a more concentrated algorithm and provides higher accuracy. Moreover, it is more resistant to high leverage points than other methods Uraibi and Haraj (18). Additionally, Uraibi (16) proposed an alternative weighting method that adjusts the chi-square distribution's percentage point to address the masking and swamping problems caused by high leverage points. This approach is essential in the proposed algorithm for assigning high-accuracy weights to mitigate the issue of high leverage points.

Consider $F(e)$ as a scale mixture normal distribution of random errors term of Equation 1, where $F(e) = (1 - \epsilon) N\left(0, \sigma^2\right) + \epsilon\, G$. In this study, $G$ is another distribution different from $N\left(0, \sigma^2\right)$ and $\epsilon \ge 0.25$. Assume that $(n \times \alpha)$ leverage points are present in each independent variable, where $\alpha$ is the percentage of leverage points. The WTLE algorithm combines the weights algorithm of Uraibi (16) but uses RMVN (17) and the FTLE method in one algorithmic frame that can be written as follows:

1 Identifying leverage points and giving them down weights as follows:

For $j = 1, 2$ Do {

a) Compute $\hat{\mu}_{RMVN}$ and $\hat{\sum}_{RMVN}$, the location and scatter estimators of x, and consider that the critical value of the desired upper bound of Mahalanobis distance is $\chi^2_{(0.975, p)}$.

b) Calculate Robust Mahalanobis distance, which is denoted as $RD\left(x_i\right)$ based on $\hat{\mu}_{RMVN}$ and $\hat{\sum}_{RMVN}$,
$$RD\left(x_i\right) = \left(x_i - \hat{\mu}_{RMVN}\right)' \hat{\sum}_{RMVN}^{-1} \left(x_i - \hat{\mu}_{RMVN}\right)$$

c) Let $S = \sum_i Rd\left(x_i\right)$ where $Rd\left(x_i\right) = \left\{ RD\left(x_i\right) : RD\left(x_i\right) \le \chi^2_{(0.975, p)} \right\}$.

d) Suppose that $\delta = \dfrac{(0.975 \times n)}{2 \times S}$ and $q = \min\left(\delta, 0.995\right)$, and then

$$\hat{\sum}_{RM(j)} = \left[ Median\left(RD\left(x_i\right)\right) / \chi^2_{(\delta, p)} \right] \times \hat{\sum}_{RMVN}.$$

e) Let $\hat{\sum}_{RMVN} = \hat{\sum}_{RM(j)}$

f) Next }.

g) Finally, $w_i = \min\left[1, \dfrac{\chi^2_{(q,p)}}{RD(x_i)}\right]$ and then $xw_i = x_i \times w_i$,

$Y_{Wi} = Y_i \times w_i$.

2 Identifying the number of outliers in $Y_W$ using robust three-sigma rules to determine the percentage of outliers $(\alpha)$ that so-called the trimming value, the C-steps can be written as follows:

Random select $h$ out of $n$ such that $h = n(1 - \alpha)$.

a) Given $H^{old} := \left\{Y_{W(1\times h)} \subset Y_{W(1\times n)}\right\}$ with corresponding $:= \left\{xw_{(h\times p+1)} \subset xw_{(n\times p+1)}\right\}$ and then compute $\hat{\beta}_w^{old}$ using multiple runs of MLE for mixture regression based on $H^{old}$.

b) Let $Q^{old} := \sum_{i=1}^{h} f\left(y_i, \hat{\beta}_w^{old}\right)$ and then sort.

c) $f\left(y_i, \hat{\beta}_w^{old}\right)$ in ascending order $f\left(y_{\pi(i)}, \hat{\beta}_w^{old}\right) \leq f\left(y_{\pi(i+1)}, \hat{\beta}_w^{old}\right)$

d) Get the permutation $\pi = \left(\pi(1), \pi(2), \ldots, \pi(n)\right)$ t.

e) Obtain $H^{new} := \left\{y_{\pi(1)}, y_{\pi(2)}, \ldots y_{\pi(h)}\right\}$.

f) $\hat{\beta}_w^{new} := MLE$ based on $H^{new}$.

g) $Q^{new} := \sum_{i=1}^{h} f\left(y_{\pi(i)}, \hat{\beta}_w^{new}\right)$.

h) If $\dfrac{\left|Q^{old} - Q^{new}\right|}{Q^{new}} < \delta$, stop; otherwise, repeat the steps (3–5) until convergence, where $\delta$ is a very small value.

## Simulation study

This section describes a simulation study prepared to examine the performance of the proposed algorithm WFTLE compared to EM and FTLE methods.

Consider the mixture linear regression model as follows:

$$Y = \begin{cases} \theta_0 + \theta_1 X_1 + \theta_2 X_2 + e, & if \ Z = 1 \\ \theta_0 - \theta_1 X_1 - \theta_2 X_2 + e & if \ Z = 2 \end{cases},$$

where $Z$ is the latent variable and when $P(Z=1)$ is construct $\pi_1 = 0.25$ of the components of $Y$ and $P(Z=2)$ is constructed $(\pi_2 = 1 - \pi_1)$. The $X_1$ and $X_2$ are sampled independently and are identically distributed from standard normal distribution individually with $n = \{25, 50, 75, 100, 200, 500\}$. The error terms $e$ has been considered in the context of the following cases:

I Let the initial values are $\theta_i = \begin{cases} \beta_1 = (1,2,2) & \sigma = 1 & \pi_1 = 0.25 \\ \beta_2 = (1,-2,-2) & \sigma = 1 & \pi_2 = 0.75 \end{cases}$,

generate $n$ data points from the standard normal distribution, $e \sim N(0,1)$ and the computing $Y$.

II Generate $n$ data points from a uniform distribution such that $e = \begin{cases} N(0,1) & if \ u \leq 0.90 \\ N(5,1) & if \ u > 0.90 \end{cases}$, and then finding $Y = \begin{cases} 1 + 2X_1 + 2X_2 + e & if \ u \leq \pi \\ 1 - 2X_1 - 2X_2 + e & if \ u > \pi \end{cases}$, where $\pi = 0.25$ of the components of $Y$.

III The same procedure as case II, with 10% of leverage points being in $X_1 = 100, X_2 = 100$.

IV Generate $n$ data points from a uniform distribution such that $e = \begin{cases} N(0,1) & if \ u \leq 0.90 \\ N(0, 5^2) & if \ u > 0.90 \end{cases}$ and calculate $Y$.

V The same procedure as case IV, with 10% of leverage points being in:

$X_1 = 100, X_2 = 100$.

The steps (I-V) are repeated 5,000 times, and the mean squares errors of the estimated component are computed as follows:

$$Mse\left(\hat{\theta}_i\right) = \frac{\sum_{j=1}^{5000}\left(\hat{\theta}_i - \theta_i\right)^2}{5000}, i = 1, 2,$$

$$Bias\left(\hat{\theta}_i\right) = \frac{\sum_{j=1}^{5000}\left|\hat{\theta}_i - \theta_i\right|}{5000}, i = 1, 2,$$

$$Mse\left(\hat{\sigma}_i\right) = \frac{\sum_{j=1}^{5000}\left(\hat{\sigma}_i - 1\right)^2}{5000}, i = 1, 2.$$

The best method is the one that has the lowest values of the above criteria.

It is interesting to note regarding the results of case I of the simulation study reported in Table 1 when the distribution of random errors is standard normal. The values of $Mse\left(\hat{\theta}_j\right)$, $Bias\left(\hat{\theta}_j\right)$, and (between two parentheses in the tables) $Mse\left(\hat{\sigma}_i\right)$ of the EM method, where $j = 1, 2, 3$ respectively, are lowest than their counterparts in FTLE and WFTLE methods when $n = (25, 50, 75)$. Notably, the values of $Mse\left(\hat{\theta}_j\right)$ become smaller little by little when the size $n$ increases from 25 to 500. Moreover, the values of $Bias\left(\hat{\theta}_i\right)$ are very small and coincide with $Mse\left(\hat{\theta}_i\right)$ values increase from 25 to 75, which is very close to each other where $n \to 500$. The results of $Mse\left(\hat{\sigma}_i\right)$ in Table 1 are reasonable, and the performance of EM is better than the others when dealing with a small sample size since the algorithm of EM does not include any trimming procedure.

Table 2 shows the simulation results of case II when the distribution of random errors is a mixture of normal; one is standard normal, and the other is shifted only by the mean. The results shown in this table suggest that the performance of FTLE and WFTLE methods is better than the EM method by having lower $Mse\left(\hat{\theta}_j\right)$. Similarly, the WFTLE method is more efficient than the FTLE method when $n = \{25, 50\}$. Notably, both methods are very close when $n \geq 75$. Regarding $Bias\left(\hat{\theta}_i\right)$, FTLE and WFTLE methods produce the smallest bias than the EM method, but both

are unstable with small samples of $n = \{25, 50\}$ and have approximately the same results when $n \geq 75$. It is clear that $Mse(\hat{\sigma}_i)$ of FTLE and WFTLE are better than the EM method with different sample sizes. However, FTLE is the best only when $n = 25$. This outperforms FTLE with this sample size not appearing with $Bias(\hat{\theta}_1)$ and $Bias(\hat{\theta}_2)$, except $Bias(0)$. It is notable that the $Bias(\hat{\theta}_1)$ and $Bias(\hat{\theta}_2)$ are the same when $n > 25$.

Table 3 presents the results of three methods in the presence of leverage points for the data generated by case III of the simulation. Indeed, the performance of the WFTLE method is more efficient than others even when the sample size is $n = \{25, 50, 75\}$, as summarized in Table 3. When the sample size is $n \geq 100$, both methods have approximately the same performance.

## The market value of Iraq's trade banks

To better demonstrate the robust performance of the WFTLE method compared to the FTLE and EM methods, we selected the market value data of Iraqi trade banks presented by Uraibi and Haraj (2022). From the original dataset, spanning 2011–2015, the Trading Rate and Earning Per Share (EPS) variables were selected out of nine financial ratios affecting market value. First, the regression model was fitted using the EM method, and the plotted residuals are shown in Figure 1.

It is clear that there is more than one pattern in the residuals, each potentially representing a random distribution. This indicates a heterogeneity problem, leading to a mixture in the

random distribution of residuals. Figure 2 further illustrates that the normalized EPS and trading rate ratios contain outliers (leverage points), as some data points lie far from the main cluster, suggesting heterogeneity. The performance of the WFTLE method proved superior to both the EM and FTLE methods, as evidenced by the market value data from the Iraq stock market presented in Table 4. The mean squared errors (MSE) and mean absolute errors (MAE) for WFTLE were 0.053 and 0.865, respectively, both of which are lower than the MSE and MAE of the other methods.

## Conclusion

The primary focus of this study is to improve the performance of the FTLE method in the presence of leverage points in mixture regression data. From the results presented in Tables 2–4, it can be concluded that the EM method is the most suitable choice where there are no outliers or leverage points. However, the EM method is highly susceptible to outliers and leverage points, as demonstrated by the variation in its results between Tables 2, 3. In contrast, the WFTLE method, especially for small sample sizes without trimming, demonstrates superior performance, even when outliers are present, as shown in Table 2. It also outperforms other methods when applied to real data. Trimming leverage points in small samples is not advisable, as it reduces the degrees of freedom, thereby increasing the value of MSE. Instead of discarding rows with leverage points by assigning zero weight, the WFTLE method assigns lower weights to those rows, preserving the degrees of freedom. This approach explains why

TABLE 1 $Mse(\hat{\theta}_i)$, $Bias(\hat{\theta}_i)$, and $Mse(\hat{\sigma}_i)$ for case I of the simulation study.

| $n$ | Method | $Mse(\hat{\theta}_0)$ | $Mse(\hat{\theta}_1)$ | $Mse(\hat{\theta}_2)$ | $Mse(\hat{\sigma}_i)$ |
|---|---|---|---|---|---|
| 25 | EM | 0.64 (0.01) | 0.71 (0.04) | 0.5 (0.00) | 0.07 |
| | FTLE | 0.65 (0.09) | 1.68 (0.42) | 1.34 (0.33) | 0.18 |
| | WFTLE | 0.75 (0.05) | 2.35 (0.59) | 1.91 (0.43) | 0.21 |
| 50 | EM | 0.24 (0.02) | 0.18 (0.02) | 0.23 (0.05) | 0.02 |
| | FTLE | 0.26 (0.02) | 0.59 (0.1) | 0.83 (0.27) | 0.1 |
| | WFTLE | 0.32 (0.09) | 0.66 (0.11) | 0.91 (0.31) | 0.1 |
| 75 | EM | 0.09 (0.02) | 0.13 (0.01) | 0.13 (0.04) | 0.01 |
| | FTLE | 0.15 (0.03) | 0.69 (0.17) | 0.76 (0.14) | 0.08 |
| | WFTLE | 0.2 (0.09) | 0.6 (0.13) | 0.78 (0.17) | 0.07 |
| 100 | EM | 0.06 (0.02) | 0.08 (0.01) | 0.05 (0.01) | 0.01 |
| | FTLE | 0.11 (0.02) | 0.08 (0.01) | 0.08 (0.03) | 0.07 |
| | WFTLE | 0.11 (0.06) | 0.08 (0.02) | 0.11 (0.05) | 0.07 |
| 200 | EM | 0.03 (0.02) | 0.04 (0.01) | 0.03 (0.01) | 0.00 |
| | FTLE | 0.05 (0.04) | 0.06 (0.04) | 0.04 (0.00) | 0.06 |
| | WFTLE | 0.06 (0.12) | 0.06 (0.05) | 0.04 (0.00) | 0.05 |
| 500 | EM | 0.01 (0.00) | 0.01 (0.01) | 0.01 (0.01) | 0.00 |
| | FTLE | 0.02 (0.01) | 0.03 (0.04) | 0.03 (0.07) | 0.05 |
| | WFTLE | 0.03 (0.07) | 0.02 (0.03) | 0.03 (0.07) | 0.05 |

TABLE 2 $Mse(\hat{\theta}_i)$, $Bias(\hat{\theta}_i)$, and $Mse(\hat{\sigma}_i)$ for case II of the simulation study.

| n | Method | $Mse(\hat{\theta}_0)$ | $Mse(\hat{\theta}_1)$ | $Mse(\hat{\theta}_2)$ | $Mse(\hat{\sigma}_i)$ |
|---|--------|------|------|------|------|
| 25 | EM | 3.86 (0.88) | 2.56 (0.36) | 3.15 (0.30) | 0.26 |
| | FTLE | 1.03 (0.07) | 2.32 (0.25) | 2.03 (0.19) | 0.1 |
| | WFTLE | 0.6 (0.14) | 1.86 (0.16) | 1.29 (0.16) | 0.12 |
| 50 | EM | 1.49 (0.68) | 1.31 (0.16) | 1.24 (0.05) | 0.31 |
| | FTLE | 0.46 (0.07) | 0.37 (0.05) | 0.39 (0.00) | 0.04 |
| | WFTLE | 0.37 (0.03) | 0.37 (0.05) | 0.38 (0.00) | 0.04 |
| 75 | EM | 1.06 (0.69) | 0.57 (0.03) | 0.48 (0.02) | 0.41 |
| | FTLE | 0.21 (0.11) | 0.23 (0.04) | 0.17 (0.00) | 0.03 |
| | WFTLE | 0.2 (0.06) | 0.23 (0.04) | 0.17 (0.00) | 0.03 |
| 100 | EM | 0.87 (0.71) | 0.64 (0.13) | 0.4 (0.09) | 0.46 |
| | FTLE | 0.13 (0.08) | 0.09 (0.00) | 0.11 (0.04) | 0.02 |
| | WFTLE | 0.12 (0.02) | 0.08 (0.00) | 0.10 (0.04) | 0.02 |
| 200 | EM | 0.51 (0.59) | 0.15 (0.00) | 0.13 (0.01) | 0.49 |
| | FTLE | 0.05 (0.04) | 0.05 (0.01) | 0.05 (0.02) | 0.01 |
| | WFTLE | 0.05 (0.02) | 0.05 (0.01) | 0.05 (0.02) | 0.01 |
| 500 | EM | 0.37 (0.56) | 0.04 (0.02) | 0.04 (0.01) | 0.52 |
| | FTLE | 0.02 (0.02) | 0.02 (0.02) | 0.02 (0.02) | 0.01 |
| | WFTLE | 0.02 (0.02) | 0.02 (0.02) | 0.02 (0.02) | 0.01 |

TABLE 3 $Mse(\hat{\theta}_i)$, $Bias(\hat{\theta}_i)$, and $Mse(\hat{\sigma}_i)$ for case III of the simulation study.

| n | Method | $Mse(\hat{\theta}_0)$ | $Mse(\hat{\theta}_1)$ | $Mse(\hat{\theta}_2)$ | $Mse(\hat{\sigma}_i)$ |
|---|--------|------|------|------|------|
| 25 | EM | 11.1 (1.00) | 8.28 (2.12) | 7.35 (1.89) | 1.29 |
| | FTLE | 1.37 (0.18) | 2.24 (0.23) | 2.57 (0.26) | 0.11 |
| | WFTLE | 0.76 (0.21) | 2.06 (0.13) | 2.27 (0.20) | 0.10 |
| 50 | EM | 11.47 (1.44) | 7.04 (1.99) | 7.13 (2.02) | 2.23 |
| | FTLE | 0.35 (0.09) | 0.44 (0.08) | 0.47 (0.11) | 0.05 |
| | WFTLE | 0.32 (0.02) | 0.40 (0.06) | 0.42 (0.11) | 0.05 |
| 75 | EM | 12.13 (1.82) | 6.07 ((2.06) | 5.69 (1.97) | 2.72 |
| | FTLE | 0.18 (0.07) | 0.22 (0.07) | 0.17 (0.04) | 0.03 |
| | WFTLE | 0.15 (0.01) | 0.20 (0.07) | 0.16 (0.05) | 0.03 |
| 100 | EM | 10.81 (1.75) | 6.1 (2.06) | 5.78 (1.96) | 3.19 |
| | FTLE | 0.11 (0.03) | 0.1 (0.02) | 0.1 (0.02) | 0.02 |
| | WFTLE | 0.11 (0.03) | 0.1 (0.02) | 0.1 (0.02) | 0.02 |
| 200 | EM | 9.78 (1.93) | 5.76 (2.15) | 5.85 (2.15) | 3.48 |
| | FTLE | 0.05 (0.04) | 0.05 (0.02) | 0.05 (0.02) | 0.01 |
| | WFTLE | 0.04 (0.02) | 0.05 (0.02) | 0.05 (0.02) | 0.01 |
| 500 | EM | 0.74 (0.64) | 9.88 (2.69) | 9.91 (2.97) | 1.98 |
| | FTLE | 0.02 (0.02) | 0.02 (0.03) | 0.02 (0.02) | 0.01 |
| | WFTLE | 0.02 (0.02) | 0.02 (0.03) | 0.02 (0.02) | 0.01 |

the WFTLE method outperforms others in small sample sizes. Additionally, it is resistant to both outliers and leverage points, with both methods performing similarly in large sample sizes.

The WFTLE method excels in small samples, though its performance may converge with FTLE when the number of observations reaches 100.
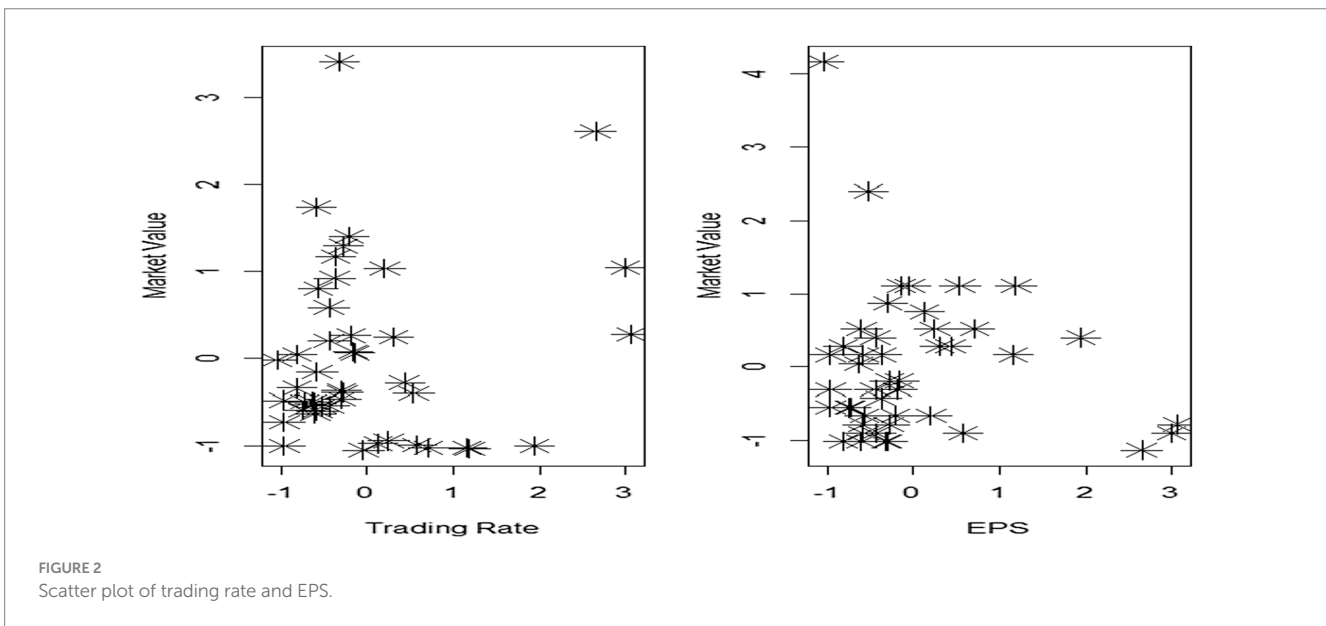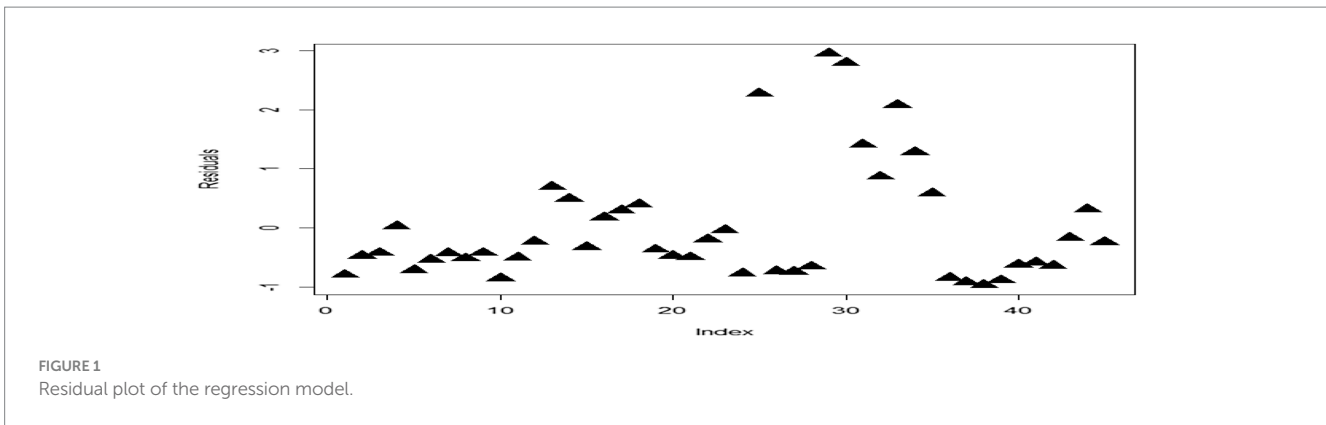
FIGURE 1
Residual plot of the regression model.



FIGURE 2
Scatter plot of trading rate and EPS.

TABLE 4  The mean squared of errors (MSE) and mean absolute errors (MAE) of market value data.

| Method | EM | FTLE | WFTLE |
|--------|--------|---------|-------|
| MSE | 169.38 | 132.086 | 0.053 |
| MAE | 2.006 | 1.976 | 0.865 |

## Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: no any restrictions that apply to the dataset. Requests to access these datasets should be directed to hssn.sami1@gmail.com.

## Author contributions

HU: Project administration, Software, Supervision, Writing – review & editing. MW: Conceptualization, Data curation, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft.

## Funding

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

# References

1. Quandt RE. A New Approach to Estimating Switching Regressions. *J Am Stat Assoc.* (1972) 67:306–10. doi: 10.1080/01621459.1972.10482378

2. Quandt RE, Ramsey JB. Estimating mixtures of normal distributions and switching regressions. *J Am Stat Assoc.* (1978) 73:730–8. doi: 10.1080/01621459.1978.10480085

3. Huber PJ. Robust statistics. New York: Wiley (1981).

4. Rousseeuw PJ, Leroy AM. Robust regression and outlier detection. New York: Wiley (1987).

5. Neykov N, Filzmoser P, Dimova R, Neytchev P. Robust fitting of mixtures using the trimmed likelihood estimator. *Comput Stat Data Anal.* (2007) 52:299–308. doi: 10.1016/j.csda.2006.12.024

6. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via theEMAlgorithm. *J R Stat Soc Series B Stat Methodol.* (1977) 39:1–22. doi: 10.1111/j.2517-6161.1977.tb01600.x

7. Peel D, McLachlan GJ. Robust mixture modelling using the t distribution. *Stat Comput.* (2000) 10:339–48. doi: 10.1023/A:1008981510081

8. Hennig C. Clusters, outliers, and regression: fixed point clusters. *J Multivar Anal.* (2003) 86:183–212. doi: 10.1016/S0047-259X(02)00020-9

9. Markatou M. Mixture models, robustness, and the weighted likelihood methodology. *Biometrics.* (2000) 56:483–6. doi: 10.1111/j.0006-341X.2000.00483.x

10. Shen R, Ghosh D, Chinnaiyan AM. Prognostic meta-signature of breast cancer developed by two-stage mixture modeling of microarray data. *BMC Genomics.* (2004) 5:1–16. doi: 10.1186/1471-2164-5-94

11. Hadi AS, Luceño A. Maximum trimmed likelihood estimators: a unified approach, examples, and algorithms. *Comput Stat Data Anal.* (1997) 25:251–72. doi: 10.1016/S0167-9473(97)00011-X

12. Vandev DL, Neykov NM. About regression estimators with high breakdown point. *Stat J Theor Appl Stat.* (1998) 32:111–29. doi: 10.1080/02331889808802657

13. Müller CH, Neykov N. Breakdown points of trimmed likelihood estimators and related estimators in generalized linear models. *J Stat Plan Inference.* (2003) 116:503–19. doi: 10.1016/S0378-3758(02)00265-3

14. Rousseeuw PJ, Driessen KV. A fast algorithm for the minimum covariance determinant estimator. *Technometrics.* (1999) 41:212–23. doi: 10.1080/00401706.1999.10485670

15. García-Escudero LA, Gordaliza A, Matrán C, Mayo-Iscar A. Avoiding spurious local maximizers in mixture modeling. *Stat Comput.* (2015) 25:619–33. doi: 10.1007/s11222-014-9455-3

16. Uraibi HS. Weighted Lasso Subsampling for HighDimensional Regression. *Electron J Appl Stat Anal.* (2019) 12:69–84. doi: 10.1285/i20705948v12n1p69

17. Olive D. J., Hawkins D. M. (2010), Robust multivariate location and dispersion. Preprint. Available at: https://lagrange.math.siu.edu/Olive/pphbmld.pdf

18. Uraibi HS, Haraj SAA. (2022). Group diagnostic measures of different types of outliers in multiple linear regression model. *Mal. J. Fund. Appl. Sci.* 41:23–33.