Robust regression modelling for inflation factor in the Indonesian economy development

A’yunin Sofro¹*, Ika Aprilia Rizka Azzahro¹, and Khusnia Nurul Khikmah¹, and Orasa Nunkaw²

¹Mathematics Department, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Surabaya, East Java, Indonesia
²Mathematics and Statistics Department, Thaksin University, Phatthalung, Thailand

Abstract. Inflation is a general and continuous increase in the prices of goods and services over a while. Inflation is measured by the Consumer Price Index (CPI) indicator to calculate the average price change of a package of goods and services households consume over time. Inflation in Indonesia is not in a precarious condition. However, it also states that inflation in Indonesia could rise. This rise can affect the sustainable development goals (SDGs). Therefore, Indonesia needs to have inflation data scrutinized every year. This study aims to examine and show the factors that affect inflation. Robust regression is an important method to analyze data contaminated by outliers and provide more flexible results. So that it can produce robust regression models and determine what factors significantly affect inflation in Indonesia by the research objectives, this study's research type is experimental research with literature studies. The variable used is Inflation in Indonesia, which is the dependent variable of the study, and the other four independent variables are export value, interest rate, money supply, and exchange rate. After that, a robust regression model with maximum likelihood type (M) estimation, scale (S) estimation, and least median of squares (LMS) estimation will be obtained, and the best model will be selected. The results show that the least median of squares (LMS) estimation is the best with the acquisition of an Akaike information criteria value of -390.1363. Furthermore, the high and low inflation in 2018-2022 is influenced by interest rates, money supply and export value.

1 Introduction

An increase in other goods over time is the definition of inflation [1, 2]. The inflation calculation conducted by the Central Bureau of Statistics (BPS) in Indonesia states that the inflation rate is measured by the percentage change in the consumer price index (CPI). Inflation occurs due to pressure from the supply side, the demand side, and inflation expectations. Therefore, inflation control needs to be done optimally to avoid the negative impact [3]. Several factors influenced the high inflation in those years. The government's effort for sustainable economic growth for SDGs is vital in controlling inflation. The high

*Corresponding author: ayuninsofro@unesa.ac.id
and low level of inflation in Indonesia is undoubtedly influenced and related to other variables. An increase in inflation will cause the rupiah exchange rate to depreciate against the U.S. dollar [4]. Previous research found that one of the independent variables that affect the high and low inflation in Indonesia is the money supply, which has a positive relationship with inflation [5]. Therefore, researchers are interested in studying the problem of inflation in Indonesia.

Research on modelling is generally very often associated with outliers. Outliers can provide information that other data points cannot provide, one of which is that outliers occur due to unusual circumstances that may be very important and need to be investigated further [6]. Some studies applying outliers are only rejected if, after tracing, it turns out to be the result of errors [7], such as entering the wrong measure or analysis, inaccurate data recording, and damage to the measurement tool. If they do not result from such errors, careful investigation should be conducted [8]. Outlier data must be seen against the position and distribution of other data, so it needs to be evaluated whether the outlier data needs to be removed. According to previous research, there are several methods to determine outlier boundaries in an analysis, namely scatterplot, boxplot, and leverage values [9].

One of the regression methods to analyze data contaminated by outliers and provide more flexible results is robust regression [10]. Therefore, this study will apply one of the robust regression methods to the case of inflation in Indonesia. In addition, it can be known whether the variables in the model can explain the behavior of inflation control in Indonesia. Robust regression is a method used to analyze data affected by outliers to produce a model that is robust or resistant to outliers. A resistance estimate is relatively unaffected by significant changes in a small part of the data or minor changes in a large part. Robust regression procedures are designed to reduce the influence of outliers [11]. Therefore, robust regression procedures ignore residuals associated with significant outliers. Another advantage of this robust regression method is that it analyzes data affected by outliers to produce a model that is robust or resistant to outliers.

2 Literature Review

2.1 Robust Regression with Maximum Likelihood Type (M) Estimation

According to Montgomery [12], M estimation minimizes the objective function of the residuals. The stages of robust regression with M estimation are as follows [13,14]. Calculate the parameter estimation $\hat{\beta}$ to obtain $\hat{Y}$, then, calculate the residual value with $\epsilon_i = Y_i - \hat{Y}_i$. Third, calculate the $S_M$ value. $S_M = \frac{\text{med}[\epsilon_i - \text{med}(\epsilon_i)]}{0.6745}$, $i = 1, 2, ..., n$. Calculating the value of $u_i = \frac{\epsilon_i}{S_i}$. Fifth, calculating the value of the weighting function $w_i = w(u_i)$ with the constant Huber weighting function used is $c = 1.345$, $w(u) = \begin{cases} 1, & |u_i| \leq c \\ \frac{c}{|u_i|}, & |u_i| > c \end{cases}$. Estimate the value of $\hat{\beta}_M$ using the weight $w_i$ to obtain a one-stage M estimation. At each $t$-th iteration, calculate the residual $\epsilon_i^{(t-1)}$ and using the weight $w_i^{(t-1)}$ from the previous iteration so that, a new $\hat{\beta}_M$ parameter estimate is obtained. Last, perform steps 2 to 6 until a converged $\hat{\beta}_M$ parameter estimate is obtained,
with $\beta$ as the parameter value. $\epsilon_i$ is the residual value. $Y_i$ is the dependent variable of the $i$-th observation. $\hat{Y}_i$ is the observation result of the dependent variable of the $i$-th observation. $S_M$ is the robust regression value with maximum likelihood type (M) estimation (the general equation for robust regression with M estimation is often used). $u_i$ is the contribution value to each objective function residual. $S_j$ is $S_M$ value of the $i$-th observation. $w_i$ is the weighting function value (if $\alpha = 5\%$ is taken, then the estimation of M will effectively use the value of $c = 1.345$). The last, $t$ is the number of iterations performed until a convergent parameter estimate is obtained.

### 2.2 Robust Regression with Scale (S) Estimation

When the data is contaminated with outliers in variable $X$, M estimation does not work well. M estimation cannot identify wrong observations, which means it cannot distinguish good and lousy leverage points; high breakdown estimation is needed. One estimate that has a high breakdown value is the S estimate [15, 16]. The S estimate can reach a breakdown point of up to 50% and providing a good influence for other observations. The stages of robust regression with S estimation are as follows [16]. First, calculate the parameter estimation $\hat{\beta}$ to obtain $\hat{Y}_i$. Then, calculate the residual value with $\epsilon_i = Y_i - \hat{Y}_i$. Third, calculate the robust scale estimation with $K = 0.199$, then find the value of $S_S$. $S_S = \left\{ \begin{array}{ll} \frac{\text{med} | e_i - \text{med}(\epsilon_i) |}{0.6745} & \text{iteration } i = 1 \\ \frac{1}{n_k} \sum_{i=1}^n w_i e_i^2 & \text{iteration } i > 1 \end{array} \right.$. Calculate the value of $u_i = \frac{\epsilon_i}{S_j}$. Fifth, calculating the value of the weighting function $w_i = w(u_i)$ with the weighting function using the weights $w_i$ to obtain a one-stage estimate of S one stage. At each iteration $t$, the residual $\epsilon_i^{(t-1)}$ is calculated and uses weight $w_i^{(t-1)} = w(u_i^{(t-1)})$ from the previous iteration to obtain the new $\hat{\beta}_S$ parameter estimate. Last, perform steps 2 to 6 until a converged $\hat{\beta}_S$ parameter estimate is obtained, with $\beta$ as the parameter value. $Y_i$ is the residual value. $Y_i$ is the dependent variable of the $i$-th observation. $\hat{Y}_i$ is the observation result of the dependent variable of the $i$-th observation. $S_S$ is the robust regression value with scale type (S) estimation (the general equation for robust regression with S estimation is often used). $u_i$ is the contribution value to each objective function residual. $S_j$ is $S_S$ value of the $i$-th observation. $w_i$ is the weighting function value (if $\alpha = 5\%$ is taken, then the estimation of S will effectively use the value of $c = 1.345$). The last, $t$ is the number of iterations performed until a convergent parameter estimate is obtained.

### 2.3 Robust Regression with Least Median of Squares (LMS) Estimation

LMS is one of the robust regression estimates with a high breakdown point [17]. The LMS algorithm minimizes the squared ordered residuals' median. The stages of robust regression with LMS estimation are as follows [18]. First, calculate the parameter estimation $\hat{\beta}$ to obtain $\hat{Y}_i$. Then, calculate the residual value with $\epsilon_i = Y_i - \hat{Y}_i$. Third, calculating the median squared residual value, where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, S$. 

https://doi.org/10.1051/e3sconf/202451301006

E3S Web of Conferences 513, 01006 (2024)
Calculating the value of $u_i = \frac{\epsilon_i}{S_i}$. Fifth, calculating the value of the weighting function $w_i = w(u_i)$ with the Huber weighting function, the constant used is $c = 1.345$, $w(u_i) = \begin{cases} 1, & |u_i| \leq c \\ 0, & |u_i| > c \end{cases}$. Estimate the value of $\hat{\beta}_{LMS}$ using the weight $w_i$ to obtain a one-stage LMS iteration. At each $t$-th iteration, calculate the residual $\epsilon_i^{(t-1)}$ and use the weight $w_i^{(t-1)} = w(u_i^{(t-1)})$ from the previous iteration to obtain the new $\hat{\beta}_{LMS}$ parameter estimate. Last, perform steps 2 to 6 until a converged $\hat{\beta}_{LMS}$ parameter estimate is obtained, with $\beta$ as the parameter value. $\epsilon_i$ is residual value. $Y_i$ is the dependent variable of the $i$-th observation. $\hat{Y}_i$ is the observation result of the dependent variable of the $i$-th observation. $S_{LMS}$ is the robust regression value with the least median of squares (LMS) estimation. $u_i$ is the contribution value to each objective function residual. $S_i$ is $S_{LMS}$ value of the $i$-th observation. $w_i$ is the weighting function value (if $\alpha = 5\%$ is taken, then the estimation of LMS will effectively use the value of $c = 1.345$). The last, $t$ is the number of iterations performed until a convergent parameter estimate is obtained.

2.4 Best Model Selection with Akaike Information Criteria (AIC)

The method for determining the best model is the Akaike Information Criteria (AIC) [19]. The smaller the residual value, the smaller the error rate the model produces. Therefore, the best model is the model with the smallest AIC value. The AIC value, where $L(\theta)$ is the likelihood value, and $p$ is the number of parameters is defined as follows:

$$AIC = -2 \ln L(\theta) + 2p$$

3 Methodology

Inflation in Indonesia is the dependent variable, and the four independent variables are export value, interest rate, money supply, and exchange rate. The underlying reason for the variables is the research conducted (5). The data used is secondary data obtained from two different websites, first the official website of Bank Indonesia publications (https://www.bi.go.id/) and the second, the official website of the BPS (https://www.bps.go.id/). The procedure of the robust regression method is collecting data, performing descriptive analysis, and analyzing outlier identification. Regardless of whether the data contains outliers, an evaluation will be conducted. This study will be used in more detail by performing statistical calculations of the scatterplot method. Third, analyzing using the robust regression method. Robust regression procedures tend to ignore residuals associated with significant outliers. This research will use the robust regression method with three estimates. There are robust regression models with M estimation, S estimation, and LMS estimation for inflation in Indonesia and finding factors that significantly affect inflation in Indonesia. Fourth, analyze validity using a parameter significance test. Based on the significant parameter value, several steps are carried out. There is a simultaneous test using the F test and a partial test using the t-test. Fifth, finding the best model with AIC. The last is making conclusions.

4 Results
Descriptive analysis was conducted to describe the data of each research variable used. Table 1 shows a description of the data on each variable used.

Table 1. Descriptive Analysis Results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min.</th>
<th>Q₁</th>
<th>Q₂</th>
<th>Mean</th>
<th>Q₃</th>
<th>Max.</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>x₁</td>
<td>13413.00</td>
<td>14138.00</td>
<td>14360.00</td>
<td>14457.00</td>
<td>14662.00</td>
<td>16367.00</td>
<td>5306834.00</td>
</tr>
<tr>
<td>x₂</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>x₃</td>
<td>5351650.00</td>
<td>5756846.00</td>
<td>6517959.00</td>
<td>6617291.00</td>
<td>7348617.00</td>
<td>8528022.00</td>
<td>917515.8</td>
</tr>
<tr>
<td>x₄</td>
<td>10453.00</td>
<td>14071.00</td>
<td>15257.00</td>
<td>17241.00</td>
<td>20509.00</td>
<td>27862.00</td>
<td>4524464.00</td>
</tr>
</tbody>
</table>

Table 1 shows that the minimum inflation value in Indonesia every month in the last five years is 0.01, and the maximum is 0.06. The difference in the inflation value in each month can be interpreted as several factors influencing inflation. Furthermore, the data used for outlier identification is data or inflation data with each independent variable. Based on the scatterplot in Figure 1, check for outliers in the data. The scatterplot was created by plotting the data against the i-th observation (i = 1, 2, ..., n).

Fig. 1. Scatterplot (a). Y and x₁. (b) Y and x₂. (c) Y and x₃. (d) Y and x₄.

Each independent variable x has some outlier values. Outliers are sets significantly different from the overall pattern of the data set. Leverage value measures the effect of an observation on the estimated parameter size, which can be seen from all distances of the observation’s x-value. If critical region $h_{ii} > cutoff$, then outliers are detected. $Cutoff = \frac{2p}{n} = \frac{2 \times 4}{60} = \frac{8}{60} = 0.1333$, where p is the number of independent variables x and n is the number of observations. The results of identifying outliers using the leverage value more significant than the cutoff are the 10-th is 0.1454, 27-th is 0.4046, 59-th is 0.1857, and 60-th data is 0.2515. Among the outlier data identification used, one of them
must show the presence of outlier data. The results obtained from all methods used indicate the presence of outlier data.

4.1 Robust Regression with Maximum Likelihood Type (M) Estimation

Maximum likelihood (M) estimation is a direct estimation method developed by Huber theoretically and mathematically. Most of the outliers found in the data are assumed to be the dependent variable. The result of parameter estimation is:

Table 2. Maximum Likelihood Type (M) Estimation Results.

<table>
<thead>
<tr>
<th>Data</th>
<th>Estimation Value</th>
<th>Error</th>
<th>$t - value$</th>
<th>$p - value$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>$1.0253 \times 10^{-1}$</td>
<td>0.0313</td>
<td>3.2762</td>
<td>0.1025</td>
</tr>
<tr>
<td>$x_1$</td>
<td>$3.0602 \times 10^{-6}$</td>
<td>0.0000</td>
<td>1.1350</td>
<td>0.0000</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$-9.7981 \times 10^{-1}$</td>
<td>0.1716</td>
<td>$-5.7082$</td>
<td>$-0.9798$</td>
</tr>
<tr>
<td>$x_3$</td>
<td>$-1.7746 \times 10^{-8}$</td>
<td>0.0000</td>
<td>$-6.2366$</td>
<td>0.0000</td>
</tr>
<tr>
<td>$x_4$</td>
<td>$2.43379 \times 10^{-6}$</td>
<td>0.0000</td>
<td>5.4414</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Based on Table 2, it is concluded that the M-estimation robust regression equation model is obtained as follows:

\[
Y = 1.0253 \times 10^{-1} + 3.0602 \times 10^{-6}x_1 - 9.7981 \times 10^{-1}x_2 - 1.7746 \times 10^{-8}x_3 \\
+ 2.43379 \times 10^{-6}x_4
\]

Furthermore, simultaneous significance testing was carried out to determine the effect of independent variables simultaneously on the dependent variable. If the $p - value < \alpha = 0.05$, then the conclusion is drawn reject $H_0$. The test hypothesis used is $H_0: \beta_1 = \beta_2 = \cdots = \beta(p - 1) = 0$ and $H_1$: not all $\beta_k$ are equal to zero ($k = 1, \ldots, p - 1$). Based on the test results, it is found that the $p - value$ of $x_1$, $x_3$, and $x_4$ is $0.0000 < \alpha$ and $x_2$ is $-0.9798 < \alpha$, which means that the model is feasible to use because there are independent variables that affect the dependent variable. The partial significance test will also be used to see how the independent variables affect the dependent variable simultaneously. The partial significance test aims to test the effect of each independent variable separately. If the $|t - value| > t - table$, then the conclusion is drawn reject $H_0$. The test hypothesis used is $H_0: \beta_k = 0$ and $H_1: \beta_k \neq 0$. If the partial significance test shows the result of $H_0$, then the $k$-th independent variable does not significantly affect the dependent variable. If the result shows $H_1$, it means that the $k$-th independent variable has a significant influence on the dependent variable ($k = 1, 2, \ldots, p - 1$). Based on the test results, it is found that the $t - value$ of $-5.7082$, $x_3$ of $-6.2366$, and $x_4$ of 5.4414 affect the dependent variable because absolutely the three independent variables are more significant than the $t - table$ value of 2.132. The $x_3$ value has a $t - value$ of 1.1350, which means it is smaller than the $t - table$ value. So, $x_1$ has a $t - value$ of 1.1350, then $x_1$ does not partially affect the dependent variable.

4.2 Robust Regression with Scale (S) Estimation

Scale (S) estimation is a method with a high breakdown point. The following is the result of parameter estimation is:

Table 3. Scale (S) Estimation Results.
Based on Table 3, it is concluded the robust regression equation model of S estimation is obtained as follows:

\[ Y = -8.643 \times 10^{-2} + 9.802 \times 10^{-6} x_1 - 5.643 \times 10^{-10} x_2 + 1.965 \times 10^{-10} x_3 - 2.885 \times 10^{-7} x_4 \]

Furthermore, simultaneous significance testing will be carried out to see the effect of the independent variables simultaneously on the dependent variable. If the \( p-value < \alpha \) with \( \alpha = 0.05 \), it is concluded to reject \( H_0 \). The test hypothesis used is \( H_0 : \beta_1 = \beta_2 = \cdots = \beta(p-1) = 0 \) and \( H_1 : \) not all \( \beta_k \) are equal to zero \( (k = 1, \ldots, p-1) \). Based on the test results, it is found that the \( p-value \) of \( x_1 \) is \( 8.92 \times 10^{-7} \) and \( x_2 \) is \( 2.87 \times 10^{-8} \), meaning that \( x_1, x_2 < \alpha \), which means that the model is feasible to use, because there are independent variables that affect the dependent variable. The value of \( x_3 \) is 0.907, and \( x_4 \) is 0.256, meaning that \( x_3, x_4 > \alpha \) means that \( x_3 \) and \( x_4 \) have no simultaneous effect on the dependent variable. The partial significance test will also be used to see how the independent variables affect the dependent variable simultaneously. The partial significance test aims to test the effect of each independent variable separately. If the \( |t-value| > t-table \), the conclusion is drawn as reject \( H_0 \). The test hypothesis used is \( H_0 : \beta_k = 0 \) and \( H_1 : \beta_k \neq 0 \). If the partial significance test shows the result \( H_0 \), then the \( k \)-th independent variable does not significantly affect the dependent variable. If the result shows \( H_1 \), it means that the \( k \)-th independent variable has a significant influence on the dependent variable \( (k = 1, 2, \ldots, p-1) \). Based on the test results, it is found that the \( t-value \) of \( x_1 \) is 5.536 and \( x_2 \) is 5.536, affecting the dependent variable because both independent variables are more significant than the \( t-table \) value of 2.132.

### 4.3 Robust Regression with Least Median of Squares (LMS) Estimation

The least median of squares (LMS) estimation has a high breakdown point and is one of the robust regression estimates. The LMS calculation limits the middle of the squared residuals of the demand. The following is the result of parameter estimation:

<table>
<thead>
<tr>
<th>Data</th>
<th>Estimation Value</th>
<th>Error</th>
<th>( t-value )</th>
<th>( p-value )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>(-8.643 \times 10^{-2})</td>
<td>(1.870 \times 10^{-2})</td>
<td>4.622</td>
<td>(2.34 \times 10^{-5})</td>
</tr>
<tr>
<td>(x_1)</td>
<td>(9.802 \times 10^{-6})</td>
<td>(1.771 \times 10^{-6})</td>
<td>5.536</td>
<td>(8.92 \times 10^{-7})</td>
</tr>
<tr>
<td>(x_2)</td>
<td>(-5.643 \times 10^{-1})</td>
<td>(8.734 \times 10^{-2})</td>
<td>-6.461</td>
<td>(2.87 \times 10^{-8})</td>
</tr>
<tr>
<td>(x_3)</td>
<td>(1.965 \times 10^{-10})</td>
<td>(1.676 \times 10^{-9})</td>
<td>0.117</td>
<td>0.907</td>
</tr>
<tr>
<td>(x_4)</td>
<td>(-2.885 \times 10^{-7})</td>
<td>(2.512 \times 10^{-7})</td>
<td>-1.148</td>
<td>0.256</td>
</tr>
</tbody>
</table>

Based on Table 4, it is concluded that the robust regression equation model of LMS estimation is obtained as follows:
\[ Y = 8.009 \times 10^{-2} + 3.358 \times 10^{-6} x_1 - 8.132 \times 10^{-1} x_2 - 1.603 \times 10^{-8} x_3 + 2.442 \times 10^{-6} x_4 \]

Furthermore, simultaneous significance testing will be carried out to see the effect of independent variables simultaneously on the dependent variable. If the \( p \)-value < \( \alpha \) with \( \alpha = 0.05 \), it is concluded to reject \( H_0 \). The test hypothesis used is \( H_0: \beta_1 = \beta_2 = \cdots = \beta (p - 1) = 0 \) and \( H_1: \) not all \( \beta_k \) are equal to zero \((k = 1, \ldots , p - 1)\). Based on the test results, it is found that the \( p \)-value of \( x_2 \) is \( 1.33 \times 10^{-5} \), \( x_3 \) is \( 5.09 \times 10^{-7} \), and \( x_4 \) is \( 1.00 \times 10^{-6} \). Then \( x_2, x_3, x_4 < \alpha \), which means that the model is feasible to use because independent variables affect the dependent variable. In contrast, \( x_1 \) is \( 0.2139 > \alpha \), and \( x_1 \) is an independent variable that does not affect the dependent variable simultaneously, which does not simultaneously affect the dependent variable. The partial significance test will also be used to see how the independent variables affect the dependent variable simultaneously. The partial significance test aims to test the effect of each independent variable separately. If the \(|t - value| > t - table\), then the conclusion is drawn reject \( H_0 \). The test hypothesis used is \( H_0: \beta_k = 0 \) and \( H_1: \beta_k \neq 0 \). Based on the test results, it is found that the \( t - value \) \( x_2 \) is \( -4.580 \), \( x_3 \) is \( -5.447 \), and \( x_4 \) is \( 5.269 \), affecting the dependent variable because absolutely the three independent variables are more significant than the \( t - table \) value of 2.132.

### 4.4 Akaike Information Criteria (AIC)

The best model selection is done by comparing the best model among all models obtained. In this study, modelling was carried out using three estimates, the M, S, and LMS. The following are the results of the best model seen from the AIC value. The AIC values of Maximum Likelihood Type (M) estimation, Scale (S) estimation, and Least Median of Squares (LMS) estimation are \(-387.913\), \(6.027\), and \(-390.136\). The smallest value obtained by the LMS estimate is \(-390.136\), so the best model obtained from the AIC value is to use the LMS estimate.

### 4.5 Discussion

The interpretation of the results obtained is the factors that significantly affect inflation in Indonesia will be shown according to the best model. The robust regression model with the LMS estimation shows that the coefficient value for \( x_1 \) is \( 0.000003358 \). It can be explained that \( x_2 \) affects \( 0.0003358\% \) of inflation. The coefficient is positive, meaning a positive relationship exists between \( x_1 \) and \( Y \). The lower the exchange rate, the lower the inflation rate. The coefficient value for \( x_2 \) is \( -0.8132 \). It can be explained that \( x_2 \) affects \( 81.32\% \) of inflation. The negative coefficient means that there is a negative relationship between \( x_2 \) and \( Y \). The lower the exchange rate, the lower the inflation rate. The coefficient value for \( x_3 \) is \( -0.0000001603 \). It can be explained that \( x_3 \) affects \( 0.00001603\% \) of inflation. The negative coefficient means that there is a negative relationship between \( x_3 \) and \( Y \). The lower the value of the money supply, the higher the inflation rate. Finally, the coefficient value for \( x_4 \) is \( 0.00002442 \). It can be explained that \( x_4 \) affects \( 0.0002442\% \) of inflation. The coefficient is positive, meaning a positive relationship exists between \( x_4 \) and \( Y \). The lower the export value, the lower the inflation value.
This study provides the results of inflation modelling in Indonesia based on the method used to determine the factors that affect the inflation rate in Indonesia [1, 3]. The study found that inflation in Indonesia is significantly influenced by interest rate, money supply and export value. The finding of this study is that, in the presence of outlier data, the proposed methods show competitive results. The study further found that several approaches used to model the inflation rate in Indonesia based on the data provided different analytical results. In addition, we also found that the analysis used in the data analysis based on this statistical model is helpful in decision-making. Inflation rate results in various impacts in various fields, which requires policymakers, especially the government, to utilize statistics through statistical modelling in making decisions to obtain information [6, 9]. This research shows that various approaches can compare methods/models that produce the best modelling to provide projections or alternative problem-solving. Researchers also believe that selecting methods and models and handling outlier data in this study are essential in providing the correct information for policymakers to make the right decisions.

5 Conclusions

The best inflation model with a robust regression method is done with the least median of squares (LMS) estimation. Based on the results of simultaneous and partial significance testing, shows that the independent variables that affect the dependent variable are interest rate, money supply and export value. The best model selection is done by comparing the best model among all the models obtained from the AIC value. The smallest AIC value obtained by LMS estimation is \(-390.1363\). The LMS estimation method outperformed two of the three other robust estimation techniques examined in this study. Future research can use the difference to compare between S-estimation and LMS estimation, both of which have high breakdown point values.

Acknowledgment

Thanks to DAPT-EQUITY Program, Lembaga Pengelola Dana Pendidikan (LPDP), Ministry of Finance, Indonesia for supporting this publication.

References

3. M. McLeay, S. Tenreyro, NBER Macroecon Annu. 34, 1 (2020)
15. L. Mili, V. Phaniraj, P. J. Rousseeuw, Control and Dynamic Systems 37, 2 (1990)