BAREKENG: Journal of Mathematics and Its Applications
March 2023 Volume 17 Issue 1 Page 0225–0234
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss1pp0225-0234

ANALYSIS OF ROBUST CHAIN LADDER METHOD IN ESTIMATING AUSTRALIAN MOTOR INSURANCE RESERVES WITH OUTLYING DATASET

Jonathan Prasetyo Johan[1], Felivia Kusnadi[2*], Benny Yong[3]

[1,2,3] *Center for Mathematics and Society, Department of Mathematics, Faculty of Information Technology and Science, Parahyangan Catholic University*
94 Ciumbuleuit Street, West Java, 40141, Indonesia

Corresponding author e-mail: [2] felivia@unpar.ac.id*

Article History:

Received: 28th September 2022
Revised: 16th December 2022
Accepted: 26th January 2023

Keywords:

Mean Squared Error;
Motor Insurance;
Reserves, Chain Ladder;
Robust Chain Ladder.

ABSTRACT

Reserves are one of the most crucial components for an insurance company to ensure it has enough money to pay off all the incurred claims. The presence of outliers in the incurred claims data harbors risk on inaccurately predicting reserves to cover claim amounts, usually achieved by the standard chain ladder reserving method. To remedy the effect of the outliers, the robust chain ladder reserving method is used by setting the median value to predict the estimated reserve. In this research, we utilized both methods on various datasets. The purpose of this paper is to determine the best method that can be utilized by insurance company in various scenarios to obtain the most optimized reserved estimate that can minimize the risk of being unable to pay the insurance claim or even the risk of over-allocating reserves that could pose profitability issue. The primary data used are the Australian domestic motor insurance claims from 2012 to 2017, obtained from Australian Prudential Regulation Authority (APRA). The dataset is then manipulated to have outliers. After calculating the estimation, the result is compared to assess the strength of the methods using Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) calculation. In conclusion, we found that the robust chain ladder reserving method works better in an outlying dataset. We also identify cases in which robust chain ladders are not appropriately used.

How to cite this article:

J. P. Johan, F. Kusnadi and B. Yong., "ANALYSIS OF ROBUST CHAIN LADDER METHOD IN ESTIMATING AUSTRALIAN MOTOR INSURANCE RESERVES WITH OUTLYING DATASET," *BAREKENG: J. Math. & App.*, vol. 17, iss. 1, pp. 0225-0234, March 2023.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article • Open Access

1. INTRODUCTION

Claims that might occur in the insurance company have two main characterizations. The first one is that we cannot perfectly predict when a claim will occur. Secondly, the high variability in claim amounts despite having a strict underwriting process. These characterizations led to why claims may pose a significant financial risk for the companies. That is why it is important to have a reserve in the insurance companies to make sure that every claim is paid based on the initial agreement. One of the methods that is often used in general insurance companies is the chain ladder reserving method, which is not bounded by any distribution properties that may complicate the reserving process [1], [2].

Datasets in non-life insurance industry are often contaminated by outliers [3]. The standard chain ladder reserving method is sensitive towards outlier data, making the estimate of the future claim inaccurate, thus posing an additional risk on the solvency of the insurance company [4]. Another method discussed by Hubert et al. [5] uses the generalization of linear regression model with which the error terms are correlated. This results as an efficient estimator, but still prone to outlying data. Thus, another method is developed into a method that can reduce the risk of outliers and increase the accuracy of the claim estimate, called the robust chain ladder reserving method [6]–[8]. The robust chain ladder reserving method identifies and modifies the outlier data.

The purpose of this paper is to discuss the chain ladder and robust chain ladder reserving methods and their effectiveness on various datasets by calculating the claim estimate and the error that are produced. This is important to understand the best method that can be used on each dataset. Furthermore, this paper discusses the strength and weaknesses of the robust chain ladder reserving method to develop appropriate solutions. Many reserving methods for general insurance have been created in the past, such as the double chain ladder that calculates incurred but not reported (IBNR) and reported but not settled (RBNS) claims [9], multivariate chain ladder that considers the company's secondary portfolio in calculating the reserve estimate [10], Bornhuetter-Ferguson method that requires loss ratio assumption in order to do the estimation [5].

In the first section, we discuss the standard chain ladder and robust chain ladder methods. Next, the application of both methods to our data that has outliers and without outliers is analyzed. In the final section, we scrutinize scenarios in which robust chain ladders are not appropriately used.

2. RESEARCH METHODS

2.1 Chain Ladder Reserving Method

Usually, the outstanding claim liability that may occur in insurance companies for long-tail insurance is based on run-off triangle data. One of the methods that is used to take the run-off triangle advantage is the chain ladder reserving method [11]. Let $\boldsymbol{C_{i,j}}$ is a claim that occurs on accident year $\boldsymbol{i, 1 \le i \le I}$ that is settled in $\boldsymbol{j, 0 \le j \le I-1}$ year. All the observation of the actual $\boldsymbol{C_{i,j}}$ can be written down in the run-off triangle form with $\boldsymbol{i + j \le I}$. While $\boldsymbol{\widehat{C}_{i,j}}$ represents the estimate claim that occurs on year $\boldsymbol{i}$ and settled in $\boldsymbol{j}$ year which place on the run-off triangle table with $\boldsymbol{i + j > I}$. The table can be seen in Table 1.

Table 1. Run-off Triangle Table for Chain Ladder Reserving Method

Accident Year	Development Year					
	0	1	⋮	j	⋮	$I-1$
1	$C_{1,0}$	$C_{1,1}$	⋮	$C_{1,j}$	⋮	$C_{1,I-1}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	$C_{i,0}$	$C_{i,1}$	$C_{i,0}$	$C_{i,j}$	⋮	$\hat{C}_{i,,I-1}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$I-1$	$C_{I-1,0}$	$C_{I-1,1}$	⋮	$\hat{C}_{I,j}$	⋮	$\hat{C}_{I-1,I-1}$
I	$C_{I,0}$	$\hat{C}_{I,1}$	⋮	$\hat{C}_{I,j}$	⋮	$\hat{C}_{I,I-1}$

The next step is to find the cumulative claim that occurs on year i and settled in j year denoted with $Z_{i,j}$. The cumulative claim can be calculated using Equation (1).

$$Z_{i,j} = \sum_{k=0}^{j} C_{i,k} \tag{1}$$

The main feature of the chain ladder reserving method is the property of $E(Z_{i,j+1}|D) = Z_{i,j} \cdot f_{j+1}$ [8], with $D = \{Z_{i,j}|1 \leq i \leq I, 0 \leq j \leq I-1\}$ [12]. This property will be used to find the development factor. Let $f_{i,j}$ be the development factor for the claim that occur on year i and settled in j year that is estimated by $\hat{f}_{i,j}$. The development factor equation is defined by **Equation (2)**.

$$\hat{f}_{i,j} = \frac{Z_{i,j}}{Z_{i,j-1}} \quad (2)$$

To determine which development factor that will be used to calculate the estimate cumulative claim, the average method will be used on each development year to find the development factor that will be used for that development year, denoted by $\hat{\boldsymbol{f}}_{\boldsymbol{j}}$. From there, the estimated cumulative claim then can be calculated using **Equations (3)** and **(4)**.

$$\boldsymbol{\hat{Z}_{i,j} = Z_{i,j-1} \cdot \hat{f}_j \ for \ j = I - i - 1} \quad (3)$$

$$\boldsymbol{\hat{Z}_{i,j} = \hat{Z}_{i,j-1} \cdot \hat{f}_j \ for \ I + i - 2 \leq j \leq I - 1} \quad (4)$$

Finally, the estimated claim can be found by subtracting the estimated cumulative claim with the estimated cumulative claim from the previous development year with the same accident year, which can be seen on **Equations (5)**, **(6)**, and **(7)**.

$$\hat{C}_{i,j} = \hat{Z}_{i,j} - \hat{Z}_{i,j-1} \ for \ I - i + 1 \leq j \leq I - 1 \quad (5)$$

$$\hat{C}_{i,j} = \hat{Z}_{i,j} - Z_{i,j-1} \ for \ i = 1, \dots, 4 \ and \ j = I - i + 1 \quad (6)$$

$$C_{i,j} = Z_{i,j} - Z_{i,j-1} \ for \ i = I - 1 \ and \ 1 \leq j \leq I - i \quad (7)$$

2.2 Robust Chain Ladder Reserving Method

The main difference between the chain ladder and robust chain ladder reserving method is that the latter uses the factor development that is estimated using **Equation (8)**, as explained by Beaver et al. [13].

$$\hat{f}_j = median\left\{\frac{Z_{i,j}}{Z_{i,j-1}} | i = 1, \dots, I - j\right\}, 1 \leq j \leq I - 1 \quad (8)$$

The reason why the median is used is because median is less likely to be affected by outliers in comparison to mean, since most of the time outliers lie at the beginning or the end of a data set, whereas median takes the value of the data in the middle. After obtaining the development factor, it will be used to estimate the cumulative claim for every known claim in the upper triangle side of data. **Equations (9)** and **(10)** are used to calculate the estimation.

$$\hat{Z}_{i,j-1} = \frac{Z_{i,j}}{\hat{f}_j} \ for \ i = 1, \dots, I - 1 \ and \ j = I - i - 1 \quad (9)$$

$$\hat{Z}_{i,j-1} = \frac{\hat{Z}_{i,j}}{\hat{f}_j} \ for \ i = 1, \dots, I - 2 \ and \ j = I - i - 2 \quad (10)$$

Pearson residuals for each accident and development year that is represented by $r_{i,j}$ are formed to detect which data from the run-off triangle is an outlier, obtained using **Equation (11)**.

$$r_{i,j} = \frac{C_{i,j} - \hat{C}_{i,j}}{\sqrt{\hat{C}_{i,j}}} \quad (11)$$

To detect which Pearson residuals contains an outlier [14], it is tested using the boxplot method that is shown by **Equation (12)** to determine which residuals contains an outlier.

$$[Q_1 - 1{,}5IQR, Q_3 + 1{,}5IQR] \quad (12)$$

Should a residual value lie outside of the interval, then the claim forms the residual is an outlier and therefore should be modified to remove the outlier so that the estimation is not biased. Let $Z^*_{i,j}$ be a cumulative claim with the first adjustment that occurred on the year i and settled in j year. There are two options on how that adjustments can be made to the outlying data. Suppose that $r_{i,j}$ is the residual outlier.
If $r_{i,j+1}$ is an outlier, then

$$Z^*_{i,j} = Median\{Z_{i,j}|i = 1, \dots, I\} \tag{13}$$

If $r_{i,j+1}$ is not an outlier, then

$$Z^*_{i,j} = \frac{Z_{i,j+1}}{Median\left\{\frac{Z_{h,j+1}}{Z_{h,j}}|h=1,\dots,I-1\right\}} \tag{14}$$

If $r_{i,j}$ is not an outlier, then

$$Z^*_{i,j} = Z_{i,j} \tag{15}$$

The next step is to form another Pearson residual to ensure that every claim is not an outlier anymore, but the claims that are used to calculate the Pearson residuals are the claims based on the claims that were settled in 0 year. The reason for that is because from the previous step, the claim with the 0-development year on the first column is less likely being an outlier. Therefore, the claim with the 0-development year is used as the basis to test whether the other claim is an outlier. Suppose that $Z^0_{i,j}$ is a cumulative claim based on the claim that was settled in 0 year, we can calculate this cumulative claim using **Equations (16)** and **(17)**.

$$\hat{f}^0_j = Median\left\{\frac{Z_{i,j}}{Z_{i,0}}\right\} \tag{16}$$

$$Z^0_{i,j} = Z_{i,0} \cdot \hat{f}^0_j \tag{17}$$

After converting the cumulative claim into claim data, it is used to calculate the Pearson residuals for the claim based on the claim that was settled in year 0, which is represented by $r^0_{i,j}$, which then is tested using the boxplot method to identify the residual that is an outlier. If the residual is not an outlier, the claim data based on the claims that were settled in 0 year can be directly used. Should the residual be an outlier, the modification for the Pearson residuals is conducted using **Equation (18)**.

$$r^0_{i,j} = Median\{r^0_{i,a}, 0 \leq a \leq j\} \tag{18}$$

After that, the residual that has been modified can be backtracked into the claim using **Equation (19)**.

$$C^0_{i,j} = r^0_{i,j} \cdot \sqrt{C^0_{i,j}} + C^0_{i,j} \tag{19}$$

The last step that needs to be done is to apply the chain ladder reserving method to the modified claim data to get the estimated claim in the future for reserving purposes.

2.3 Mean Squared Error

Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) are utilized to observe the variance and standard deviation of the estimated claim **[13]**. The higher the MSE or RMSE value means the higher the possibility for the estimated claim to be an outlier. Both formulas are in **Equation (20)**.

$$MSE = \frac{\Sigma(C_{i,j}-\hat{C}_{i,j})^2}{n}, \quad RMSE = \sqrt{\frac{\Sigma(C_{i,j}-\hat{C}_{i,j})^2}{n}} \tag{20}$$

3. RESULTS AND DISCUSSION

The data used for the reserves estimate is the Australian domestic motor insurance claim record from 2012 to 2017 **[15]**. There will be two sets of data, one is the actual data that represent a dataset without outliers, and the other is the actual data but with the claim that occurred in 2015 and resolved in one-year times 10 to make the data outlying, which represents a dataset with outliers.

Table 2. Domestic Motor Insurance Claim without Outlier Data

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	4.689
2013	3.783	4.574	4.601	4.602	4.593	
2014	3.891	4.751	4.779	4.778		
2015	4.503	5.507	5.569			

Accident Year	Development Year					
	0	1	2	3	4	5
2016	4.451	5.575				
2017	4.812					

Table 3. Domestic Motor Insurance Claim with Outlier Data

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	4.689
2013	3.783	4.574	4.601	4.602	4.593	
2014	3.891	4.751	4.779	4.778		
2015	4.503	55.070	5.569			
2016	4.451	5.575				
2017	4.812					

3.1. Reserve Estimates on Dataset without Outliers

Firstly, the data from Table 2 are used to estimate the reserve of the future claim using the chain ladder reserving method and robust chain ladder reserving method. The result of the estimate can be seen in Table 4 for the chain ladder reserves and Table 5 for the robust chain ladder reserves.

Table 4. Reserve Estimate of Domestic Motor Insurance Claim without Outlier Data Using Chain Ladder Method

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	4.689
2013	3.783	4.574	4.601	4.602	4.593	4.618
2014	3.891	4.751	4.779	4.778	4.784	4.791
2015	4.503	5.570	5.569	5.554	5.555	5.564
2016	4.451	5.575	5.559	5.556	5.557	5.566
2017	4.812	5.905	5.942	5.940	5.940	5.949

Table 5. Reserve Estimate of Domestic Motor Insurance Claim without Outlier Data Using Robust Chain Ladder Method

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	4.689
2013	3.783	4.574	4.601	4.602	4.593	4.805
2014	3.891	4.751	4.779	4.778	4.873	4.982
2015	4.503	5.570	5.569	5.508	5.642	5.768
2016	4.451	5.575	5.482	5.443	5.575	5.700
2017	4.812	5.885	5.926	5.884	6.026	6.161

As shown in Table 4 and Table 5, the results of the estimation between the chain ladder and robust chain ladder on a dataset without an outlier are quite similar. Therefore, to see in a more detailed manner, the comparison with the actual data and the calculated error for each estimated claim can be seen inTable 6. Let CL be the estimated claim using the chain ladder method and RCL be the estimated claim using the robust chain ladder method.

Table 6. Comparison of Reserve Estimate of Domestic Motor Insurance Claim without Outlier Data

$C_{i,j}$	CL	RCL	Actual Claim	% Error CL	% Error RCL
$C_{2013,5}$	4.618	4.805	4.583	0,764%	4,844%
$C_{2014,4}$	4.784	4.873	4.771	0,272%	2,138%
$C_{2014,5}$	4.791	4.962	4.775	0,335%	3,916%
$C_{2015,3}$	5.554	5.508	5.568	0,251%	1,078%
$C_{2015,4}$	5.555	5.642	5.574	0,341%	1,220%

$C_{i,j}$	CL	RCL	Actual Claim	% Error CL	% Error RCL
$C_{2015,5}$	5.564	5.768	5.584	0,358%	3,295%
$C_{2016,2}$	5.559	5.482	5.628	1,226%	2,594%
$C_{2016,3}$	5.556	5.443	5.646	1,594%	3,595%
$C_{2016,4}$	5.557	5.575	5.657	1,768%	1,450%
$C_{2017,1}$	5.905	5.885	6.035	2,154%	2,486%
$C_{2017,2}$	5.942	5.926	6.103	2,638%	2,900%
$C_{2017,3}$	5.940	5.884	6.128	3,068%	3,982%
Average % Error				1,231%	2,791%

From the comparison above the estimated reserved claim with chain ladder method has slightly better performance in comparison to the robust chain ladder, with a 1,5% error decrease in the average percentage error. Furthermore, through the calculation of the MSE and RMSE we got the result as **Table 7**.

Table 7. MSE and RMSE of the Estimated Reserved Claim by Chain Ladder and Robust Chain Ladder Reserving Method on Dataset Without Outliers

Method	MSE	RMSE
CL	8.636	92,931
RCL	26.136	163,134

This result also shows that chain ladder performs better than robust chain ladder on dataset without outliers, especially on the MSE metric; there is a 202,64% difference that shows higher variability in robust chain ladder estimates which makes it inconsistent. Furthermore, the RMSE also shows that the deviation created by the chain ladder reserving method is lower by 75,54%, which is beneficial for an insurance company that prefers the most accurate reserve estimation to ensure the maximum profit while still being able to cover all the claims.

3.2. Reserve Estimates on Dataset with Outliers

In this section, the data from **Table 3** will be used to find the reserves estimate using both the chain ladder method and robust chain ladder method on a dataset with outlier. The result of the estimate can be seen in **Table 8**. for chain ladder reserves and **Table 9** for robust chain ladder reserves.

Table 8. Reserve Estimate of Domestic Motor Insurance Claim with Outlier Data Using Chain Ladder Method

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	4.689
2013	3.783	4.574	4.601	4.602	4.593	4.618
2014	3.891	4.751	4.779	4.778	4.784	4.791
2015	4.503	55.070	5.569	23.225	23.228	23.264
2016	4.451	5.575	4.399	5.143	5.143	5.151
2017	4.812	16.498	9.349	10.931	10.932	10.949

Table 9. Reserve Estimate of Domestic Motor Insurance Claim with Outlier Data Using Robust Chain Ladder Method

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	4.689
2013	3.783	4.574	4.601	4.602	4.593	4.783
2014	3.891	4.751	4.779	4.778	4.872	4.944
2015	4.503	55.070	5.569	5.491	5.637	5.721
2016	4.451	5.575	5.947	5.426	5.571	5.654
2017	4.812	5.913	5.939	5.862	6.018	6.108

As we can see, the results of the estimated claim from chain ladder and robust chain ladder method are significantly different, especially on the claim that occurred in 2015 and 2017. Both **Tables 8** and **9** are used to present the comparison with the actual data, and the calculated error for each estimated claim can be seen in **Table 10**. In contrast, when dealing with dataset with outliers, the chain ladder reserving method performs poorly by producing 108,05% error on the estimates, especially in the accident year 2015 and 2017. This is because the outlying data happen in the accident year 2015, therefore making the estimation for the rest of the year also become an outlier. In 2017 because we estimate the claim starting from the development year 1, which is the development year of the outlier, we produce a development factor that is calculated based on the outlier therefore affecting every calculation for every estimate that starts from the development year 1. However, the robust chain ladder performs well in estimating using the dataset with outliers by only providing 2,648% average error. To crystalize it on an absolute number, **Table 11** show the MSE and RMSE for both method estimation on a dataset with outliers.

Table 10. Comparison of Reserve Estimate of Domestic Motor Insurance Claim with Outlier Data

$C_{i,j}$	CL	RCL	Actual Claim	% Error CL	% Error RCL
$C_{2013,5}$	4.618	4.783	4.583	0,764%	4,364%
$C_{2014,4}$	4.784	4.872	4.771	0,272%	2,117%
$C_{2014,5}$	4.791	4.944	4.775	0,335%	3,539%
$C_{2015,3}$	23.225	5.491	5.568	317,116%	1,383%
$C_{2015,4}$	23.228	5.637	5.574	316,720%	1,130%
$C_{2015,5}$	23.264	5.721	5.584	316,619%	2,453%
$C_{2016,2}$	4.399	5.497	5.628	21,837%	2,328%
$C_{2016,3}$	5.143	5.426	5.646	8,909%	3,897%
$C_{2016,4}$	5.143	5.571	5.657	9,086%	1,520%
$C_{2017,1}$	16.498	5.913	6.035	173,372%	2,022%
$C_{2017,2}$	9.349	5.939	6.103	53,187%	2,687%
$C_{2017,3}$	10.931	5.862	6.128	78,378%	4,341%
Average % Error				108,050%	2,648%

Table 11. MSE and RMSE of the Estimated Reserved Claim by Chain Ladder and Robust Chain Ladder Reserving Method on Dataset with Outliers

Method	MSE	RMSE
CL	90.093.730	9.491,772
RCL	24.410	156,238

3.3. Analysis on Different Scenarios

Even though from the results above robust chain ladder performs exceptionally well in estimating claims with outliers, note that there are several cases when robust chain ladder is unable to give precise estimates and instead failed to remove the outlier factors. This section will discuss and analyze those scenarios. From the observation **Table 11**, clearly, robust chain ladder outperforms chain ladder in terms of accuracy of the estimation by providing significantly less variability and deviation on the estimates. Therefore, a robust chain ladder should be the preferred method for insurance companies to use when dealing with estimation that involves a dataset with outliers.

3.3.1. Scenario when outliers were placed on the claims with 0-development year

Assume that the data that will be used is the same data in **Table 2**, but the claim that occured in 2015 and settled in year 0 will be an outlier by multiplying it with 10, like **Table 12**. Firstly, both chain ladder and robust chain ladder reserving methods will be used to calculate the reserved claims. Then by calculating the percentage error of each method, the comparison for both reserved claims can be analyzed. The result of the calculation can be seen in **Table 13**.

Overall, **Table 13** shows that both methods provide a poor estimation on this scenario, especially on the claim that occured in 2015. The reason why robust chain ladder is unable to estimate accurately is because one of the steps in robust chain ladder is calculating the claim using the development factor of the claims that are based on the claim that is settled in year 0. Therefore, if the claim with the zero-development year is an outlier, the result of the calculation that is using that claim will also be an outlier, resulting in every claim in 2015 becoming an outlier. One of the methods to mitigate this scenario is taking a wider range of time frames so the outlier would not be on the development year 0. For example, in this case, the data time frame is from

2012 to 2017. To ensure that the outlier is not placed on the development year 0, the time frame can be enlarged from 2011 or prior to 2017. In the case where this is the only data that can be used, from average percentage error, the chain ladder method provides a significantly lower error. However, the reason why robust chain ladder average percentage error is bigger than chain ladder is because of the error spike on claims that occurred in 2015, but the rest of the error is far more stable in comparison to the chain ladder method by having an error that is less than 5%. Therefore, the conclusion is that if what insurance companies are looking for is stability, robust chain ladder is the better method because now the company can focus on mitigating the error in a narrower time frame in comparison to chain ladder method, where the big error is spreading in all the estimates.

Table 12. Domestic Motor Insurance Claim with Outlier Data on the Development Year 0

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	4.689
2013	3.783	4.574	4.601	4.602	4.593	
2014	3.891	4.751	4.779	4.778		
2015	45.030	5.507	5.569			
2016	4.451	5.575				
2017	4.812					

Table 13. Comparison of Reserve Estimate of Domestic Motor Insurance Claim with Outlier Data on the Development Year 0

$C_{i,j}$	CL	RCL	Actual Claim	% Error CL	% Error RCL
$C_{2013,5}$	4.618	4.846	4.583	0,764%	5,739%
$C_{2014,4}$	4.784	4.863	4.771	0,272%	1,928%
$C_{2014,5}$	4.791	5.011	4.775	0,335%	4,942%
$C_{2015,3}$	20.004	55.417	5.568	259,267%	895,277%
$C_{2015,4}$	20.006	56.275	5.574	258,916%	909,598%
$C_{2015,5}$	20.037	57.993	5.584	258,829%	938,557%
$C_{2016,2}$	4.440	5.455	5.628	21,109%	3,074%
$C_{2016,3}$	5.158	5.478	5.646	8,643%	2,976%
$C_{2016,4}$	5.158	5.563	5.657	8,821%	1,662%
$C_{2017,1}$	4.846	5.876	6.035	19,702%	2,635%
$C_{2017,2}$	4.277	5.897	6.103	29,920%	3,375%
$C_{2017,3}$	4.968	5.922	6.128	18,930%	3,362%
Average % Error				73,792%	231,094%

3.3.2. Scenario when outliers were placed on the claims with 0-development year

For this subsubsection, the data are based on the data in **Table 2**. The difference from the data is that the claim that occurred in 2012 and settled in 5 years will be an outlier by multiplying the value by 10. The data can be seen in **Table 14**. Again, both chain ladder and robust chain ladder reserving methods are used with the **Table 14** data. After estimating the claims, all calculations are compared using their percentage error to gauge the effectiveness of each method, shown in **Table 15**.

Table 14. Domestic Motor Insurance Claim with Outlier Data on the Claims that Occurred in the Past but Have Just Been Recently Settled

Accident Year	Development Year					
	0	1	2	3	4	5
2012	3.770	4.638	4.690	4.695	4.700	46.890
2013	3.783	4.574	4.601	4.602	4.593	
2014	3.891	4.751	4.779	4.778		
2015	4.503	5.507	5.569			
2016	4.451	5.575				
2017	4.812					

According to the result, both chain ladder and robust chain ladder reserving methods produced a similar inaccurate result, mainly on every claim estimate with development year 5. This is because the outlier was placed in the development year 5, in which there is only 1 claim known. The main feature of robust chain ladder is the usage of median to eliminate outlying development factors. However, due to the fact, there is only 1 data existing, the median only takes that data. If it is an outlier, the development factor is an outlier as well, making every estimation using that development factor an outlier as well. Should an insurance company be faced with this situation, the best method that can be taken is choosing the standard chain ladder. This is because the outlier errors are placed in the same claim estimates, so there is no significant difference, but furthermore, the overall error produced by the chain ladder method is lower in comparison to the robust chain ladder method, providing more accurate results in the other estimates.

Table 15. Comparison of Reserve Estimate of Domestic Motor Insurance Claim with Outlier Data on the Claims that Occurred in the Past but Have Just Been Recently Settled

$C_{i,j}$	CL	RCL	Actual Claim	% Error CL	% Error RCL
$C_{2013,5}$	46.181	46.846	4.583	907,659%	922,169%
$C_{2014,4}$	4.784	4.860	4.771	0,272%	1,865%
$C_{2014,5}$	47.911	48.512	4.775	903,372%	915,958%
$C_{2015,3}$	5.554	5.506	5.568	0,251%	1,114%
$C_{2015,4}$	5.555	5.625	5.574	0,341%	0,915%
$C_{2015,5}$	55.636	56.142	5.584	896,347%	905,408%
$C_{2016,2}$	5.559	5.481	5.628	1,226%	2,612%
$C_{2016,3}$	5.556	5.442	5.646	1,594%	3,613%
$C_{2016,4}$	5.557	5.560	5.657	1,768%	1,715%
$C_{2017,1}$	5.905	5.585	6.035	2,154%	7,457%
$C_{2017,2}$	5.942	5.926	6.103	2,638%	2,900%
$C_{2017,3}$	5.940	5.884	6.128	3,608%	3,982%
Average % Error				226,724%	230,809%

4. CONCLUSIONS

Based on the results above, there are a few conclusions that can be made. Firstly, the robust chain ladder reserving method provides better effectiveness than the chain ladder reserving method when dealing with a dataset that has outliers. While on the other hand, the chain ladder reserving method provides better effectiveness in comparison to the robust chain ladder reserving method when dealing with dataset without outliers. However, when it comes to reserving with robust chain ladder reserving method on a dataset that has outliers, it should be noted that this method comes with weakness that it cannot estimate accurately if the outlier was placed on the claims with development year 0, or if it was placed on a claim that occur in the past but have just been recently settled. On these cases, the chain ladder reserving method also thrive over the robust chain ladder reserving method, even though both results are not preferable.

REFERENCES

[1] T. Mack, "Distribution-free calculation of the standard error of chain ladder reserve estimates," *ASTIN Bull.*, vol. 23, no. 2, pp. 213–225, 1993.

[2] S. A. et al Klugman, *Loss Models: From Data to Decisions*, 5th ed. John Wiley & Sons, 2019.

[3] M. Barlak, J.; Bakon, M.; Rovnak, M.; Mokrisova, "Heat Equation as a Tool for Outliers Mitigation in Run-Off Triangles for Valuing the Technical Provisions in Non-Life Insurance Business," *Risks*, vol. 10, no. 9, p. 171, 2022.

[4] B. Avanzi, B.; Lavender, M.; Taylor, G.; Wong, "On the impact detection and treatment of outliers in robust loss reserving," *Proc. Actuar. Institute, Gen. Insur. Semin.*, pp. 13–15, 2016.

[5] Ö. Hubert, M.; Verdonck, T.; Yorulmaz, "Fast robust SUR with economical and actuarial applications," *Stat. Anal. Data Min. ASA Data Sci. J.*, vol. 10, no. 2, pp. 77–88, 2017.

[6] J. Verdonck, T.; Wouwe, M. V.; Dhaene, "A robustification of chain ladder method," *North Am. Actuar. J.*, vol. 13, no. 2, pp. 280–298, 2009.

[7] B. Pitselis, G.; Vasiliki, G.; Ioannis, "Robust loss reserving in a log-linear model," *Insur. Math. Econ.*, vol. 64, pp. 14–27, 2015.

[8] Y. Suwardi; D.A.; Purwono, "The Analysis of Motor Vehicle Insurance Claim Reserve Using Robust Chain Ladder," *5th Glob. Conf. Business, Manag. Entrep. (GCBME 2020)*, 2021.

[9] M. Schoepf, "On The Double Chain Ladder for Reserve Estimation with Bootstrap Applications," Missouri University of

Science and Technology, 2016.
[10] T. Peremans, K.; Aelst, S. V.; Verdonck, “A Robust General Multivariate Chain Ladder Method,” *Risks*, vol. 6, no. 4, p. 108, 2018.
[11] G. Taylor, *Loss Reserving An Actuarial Perspective*, 1st ed. New York: Springer, 2000.
[12] L. R. Brown, R. L.; Gottlieb, *Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance*. Winsted: ACTEX Publications, Inc, 2007.
[13] R. J. Beaver, B. M.; William Mendehall, I.; Beaver, *Introduction to Probability and Statistics*, 14th ed. Boston: Brooks/Cole, 2013.
[14] Y. A. Warton, D. I.; Thibaut, L.; Wang, “The PIT-trap—A ‘model-free’ bootstrap procedure for inference about regression models with discrete, multivariate responses,” *PLoS One*, vol. 12, no. 7, p. e0181790, 2017.
[15] APRA, “General insurance claims development statistics,” *Australian Prudential Regulation Authority*, 2021. .