FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 24, December, 2026, pp 287 – 296
DOI: https://doi.org/10.33003/fjs-2026-1024-6011

Modelling Loss and Allocated Loss Adjustment Expenses using Modified Bivariate Conditional Lomax Distribution with a Power-Link Fuction

***Audu Samuel Omeiza, Shobanke Dolapo Abidemi and Olajide Oluwamayowa Opeyimika**

Department of Statistics, Federal University Lokoja, Kogi State, Nigeria.

*Corresponding authors' email: audiosam360@gmail.com Phone: +2348161170359

Received: 20 Sep. 2026
Accepted: 10 Oct. 2026
Published: 16 Oct. 2026

ABSTRACT
This study developed a Modified Bivariate Conditional Lomax Distribution using Power-Link Function (MBCLD-PLF) in which the scale parameter is linked to the conditioning variable through a power function. The model was proposed to provide a simple but flexible way of describing dependence between two positive heavy-tailed variables. Instead of introducing dependence mainly through a copula, the proposed model allows the conditional scale parameter to change directly with the value of the first variable. A Monte Carlo simulation with 1,000 replications at each sample size of 50, 100, 150, 300, 500 and 1,000 was used to examine the finite-sample behaviour of the maximum likelihood estimators. The results showed that estimation generally improved as the sample size increased, with lower bias, mean squared error and root mean squared error, while the confidence-interval coverage probabilities remained close to the nominal 95% level. Numerical optimisation converged successfully across the simulation settings, although boundary estimates were more frequent at the smallest sample size. The model was also fitted to the Loss–ALAE insurance dataset containing 1,500 paired observations. Its performance was compared with the Classical Bivariate Lomax, Bivariate Power Lomax–FGM, Clayton Copula Bivariate Lomax and Bivariate Power Lomax–Sarmanov models. The proposed model recorded the highest maximised log-likelihood and the lowest AIC, BIC, CAIC and HQIC. However, the Clayton and Sarmanov models produced a smaller Kolmogorov–Smirnov, Anderson–Darling and Cramér–von Mises Rosenblatt-residual statistics. The findings therefore support MBCLD-PLF as a competitive conditional alternative with strong likelihood-based performance for heavy tailed data.

Keywords: Bivariate Lomax distribution, Power-link function, Conditional specification, Heavy-tailed data, Maximum likelihood estimation

INTRODUCTION

Extreme observations occur in many types of data, particularly in insurance, finance, reliability studies and environmental applications. In such situations, very large observations may occur more frequently than would be expected under light-tailed models. Distributions that allow greater probability in their tails are therefore important when the main interest includes rare but severe events (Embrechts et al., 1997; Beirlant et al., 2004). Baba et al. (2025) observed that many existing heavy-tailed distributions either lack sufficient flexibility to capture the full spectrum of real-world data characteristics or suffer from excessive complexity that limits their practical application. The Lomax distribution, also known as the Pareto Type II distribution, retains the heavy-tailed behaviour of the Pareto distribution but has support beginning from zero (Lomax, 1954). This property has made the distribution useful for modelling positive and highly skewed observations. It has been applied in reliability, survival, insurance and financial studies where unusually large observations may occur (Helu & Samawi, 2022; Qura et al., 2023).

The need for bivariate extensions arises when two heavy-tailed variables are observed together. Analysing such variables separately ignores information contained in their dependence. Bivariate distributions make it possible to study their individual behaviour and dependence within one probability model (Fayomi et al., 2024). Recent bivariate Lomax models have used dependence structures such as the FGM copula, Clayton copula and Sarmanov construction (Clayton, 1978; Qura et al., 2023; Fayomi et al., 2024; Abd Elgawad et al., 2025). Conditional specification provides another construction framework. Arnold and Strauss (1988), Arnold and Strauss (1991), and Arnold et al. (1999) showed how bivariate distributions can be constructed through specified conditional forms, while Arnold and Sarabia (2022) reviewed the broader conditional-specification framework. In the present model, the dependence is introduced through the conditional distributional parameter itself rather than through a separate copula.

Based on this conditional framework, the present study develops a bivariate conditional Lomax model in which the conditional scale parameter changes through a power-link function. Specifically, the scale of ALAE conditional on Loss is linked to Loss through $\lambda_2(x) = ax^{\mathrm{b}}$. This differs from a copula construction, where dependence is represented through a separate dependence function, and it also differs conceptually from a model with an externally supplied covariate because X is itself a random variable in the joint distribution. The Loss–ALAE dataset contains 1,500 paired observations and provides an application in which both variables are positive, highly skewed and positively associated.

The study has three main objectives: (i) to formulate the modified bivariate conditional Lomax distribution by linking the conditional scale parameter to the conditioning variable through a power function; (ii) to estimate the model

parameters by maximum likelihood and evaluate finite-sample performance using Monte Carlo simulation; and (iii) to apply the proposed distribution to the Loss–ALAE data and compare it with selected bivariate Lomax models. The contribution is the proposed power-link conditional scale construction, together with its likelihood-based estimation, simulation assessment and empirical application.

MATERIALS AND METHODS

Let

$$X \sim Lomax(\alpha_1, \lambda_1) \quad (1)$$

where $\alpha_1 > 0$ is the shape parameter and $\lambda_1 > 0$ is the scale parameter. The probability density function of X is

$$f_X(x) = \frac{\alpha_1}{\lambda_1}\left(1+\frac{x}{\lambda_1}\right)^{-(\alpha_1+1)}, x > 0. \quad (2)$$

$$f_X(x) = \alpha_1\lambda_1^{-1}\left(\frac{\lambda_1+x}{\lambda_1}\right)^{-(\alpha_1+1)}, x > 0. \quad (3)$$

$$f_X(x) = \alpha_1\lambda_1^{-1}\left(\lambda_1{}^{-1}\right)^{-(\alpha_1+1)}(\lambda_1+x)^{-(\alpha_1+1)}, x > 0. \quad (4)$$

Simplifies to

$$f_X(x) = \alpha_1\lambda_1^{\alpha_1}(\lambda_1+x)^{-(\alpha_1+1)}, x > 0. \quad (5)$$

Now, suppose that the conditional distribution of Y given $X = x$ is also Lomax. That is,

$$Y \mid X = x \sim Lomax(\alpha_2, \lambda_2) \quad (6)$$

where $\alpha_2 > 0$ is the conditional shape parameter and λ_2 is the conditional scale parameter.

The major idea of this study is to allow the scale parameter of $Y \mid X = x$ to depend on x through a power-link function. Hence,

$$\lambda_2 = \lambda(x) = ax^b, a > 0, b > 0. \quad (7)$$

Here, a is a positive scale constant, while b is the power-link dependence parameter. The parameter b controls how strongly the value of X affects the conditional behaviour of Y. If $b = 1$, the scale changes linearly with x. If $b \neq 1$, the relationship becomes nonlinear. The restriction $b > 0$ is imposed in the main model because it makes the conditional scale $\lambda_2(x) = ax^{\mathrm{b}}$ increase with x and therefore focuses the model on positive scale dependence. The boundary case $b = 0$ is mathematically meaningful. It gives $\lambda_2(x) = a$, so the conditional distribution of Y no longer depends on X and the resulting joint distribution reduces to independence. This boundary case is excluded from the main estimation to retain the intended increasing conditional-scale interpretation, but it provides a useful nested reference case.

Since $Y \mid X = x$ follows a Lomax distribution, its conditional density is

$$f_{Y|X}(y \mid x) = \frac{\alpha_2}{\lambda(x)}\left(1+\frac{y}{\lambda(x)}\right)^{-(\alpha_2+1)}, y > 0. \quad (8)$$

Substituting $\lambda(x) = ax^b$ into the conditional density gives

$$f_{Y|X}(y \mid x) = \frac{\alpha_2}{ax^b}\left(1+\frac{y}{ax^b}\right)^{-(\alpha_2+1)}, y > 0, x > 0. \quad (9)$$

Simplifying

$$f_{Y|X}(y \mid x) = \alpha_2 a^{\alpha_2}x^{b\alpha_2}(ax^b+y)^{-(\alpha_2+1)}, x > 0, y > 0. \quad (10)$$

Model Formulation and Joint Density

The joint density of X and Y is obtained using

$$f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y \mid x) \quad (11)$$

Substituting the marginal density of X (Eqn 5) and the conditional density of $Y \mid X = x$ (Eqn 10), we get

$$f_{X,Y}(x,y) = \left[\alpha_1\lambda_1^{\alpha_1}(\lambda_1+x)^{-(\alpha_1+1)}\right] \left[\alpha_2 a^{\alpha_2}x^{b\alpha_2}(ax^b+y)^{-(\alpha_2+1)}\right] \quad (12)$$

Multiplying the constants and terms together gives

$$f_{X,Y}(x,y) = \alpha_1\alpha_2\lambda_1^{\alpha_1}a^{\alpha_2}x^{b\alpha_2}(\lambda_1+x)^{-(\alpha_1+1)}(ax^b+y)^{-(\alpha_2+1)} \quad (13)$$

Hence, the proposed MBCLD-PLF is

$$f_{X,Y}(x,y) = \alpha_1\alpha_2\lambda_1^{\alpha_1}a^{\alpha_2}x^{b\alpha_2}(\lambda_1+x)^{-(\alpha_1+1)}(ax^b+y)^{-(\alpha_2+1)} \quad (14)$$

$$f_{X,Y}(x,y) = \frac{\alpha_1\alpha_2\lambda_1^{\alpha_1}a^{\alpha_2}x^{b\alpha_2}}{(\lambda_1+x)^{(\alpha_1+1)}(ax^b+y)^{(\alpha_2+1)}}, x > 0, y > 0 \quad (15)$$

where

$x > 0, y > 0, \alpha_1 > 0, \alpha_2 > 0, \lambda_1 > 0, a > 0, b > 0.$

The full parameter vector of the proposed model is therefore $\Theta = (\alpha_1, \lambda_1, \alpha_2, a, b)$.

Validity of the Model

The total probability over the support must be equal to one, that is,

$$\int_0^\infty\int_0^\infty f_{X,Y}(x,y)\,dy\,dx = 1 \quad (16)$$

Substituting the proposed joint density gives

$$\int_0^\infty\int_0^\infty \frac{\alpha_1\alpha_2\lambda_1^{\alpha_1}a^{\alpha_2}x^{b\alpha_2}}{(\lambda_1+x)^{(\alpha_1+1)}(ax^b+y)^{(\alpha_2+1)}}dy\,dx \quad (17)$$

$$\int_0^\infty \frac{\alpha_1\lambda_1^{\alpha_1}}{(\lambda_1+x)^{(\alpha_1+1)}}\left[\int_0^\infty \frac{\alpha_2(ax^b)^{\alpha_2}}{(ax^b+y)^{(\alpha_2+1)}}dy\right]dx \quad (18)$$

Evaluating the inner integral,

$$\int_0^\infty \frac{\alpha_2(ax^b)^{\alpha_2}}{(ax^b+y)^{(\alpha_2+1)}}dy \quad (19)$$

Let $u = ax^{\mathrm{b}} + y$, so that $du = dy$. When $y = 0, u = ax^{\mathrm{b}}$, and when $y \to \infty, u \to \infty$. Therefore,

$$\int_0^\infty \frac{\alpha_2(ax^b)^{\alpha_2}}{(ax^{\mathrm{b}}+y)^{\alpha_2+1}}dy = \alpha_2(ax^b)^{\alpha_2}\int_{ax^{\mathrm{b}}}^\infty u^{-(\alpha_2+1)}du. \quad (20)$$

Since

$$\int u^{-(\alpha_2+1)}du = -\left(\frac{1}{\alpha_2}\right)u^{-\alpha_2}, \quad (21)$$

it implies that

$$\alpha_2(ax^b)^{\alpha_2}\left[-\left(\frac{1}{\alpha_2}\right)u^{-\alpha_2}\right]_{ax^b}^{\infty} \quad (22)$$

$$\alpha_2(ax^b)^{\alpha_2}\left[0 + \left(\frac{1}{\alpha_2}\right)(ax^b)^{-\alpha_2}\right] = 1. \quad (23)$$

Therefore $\int_0^{\infty} f_Y|_x(y\,|x)dy = 1$. The double integral (Eqn 18) now becomes

$$\int_0^{\infty} \frac{\alpha_1\lambda_1^{\alpha_1}}{(\lambda_1+x)^{(\alpha_1+1)}}\,dx. \quad (24)$$

Let $v = \lambda_1 + x$, so that $dv = dx$. When $x = 0, v = \lambda_1$, and when $x \to \infty, v \to \infty$. Thus,

$$\int_0^{\infty} \frac{\alpha_1\lambda_1^{\alpha_1}}{(\lambda_1+x)^{(\alpha_1+1)}}\,dx = \alpha_1\lambda_1^{\alpha_1}\int_0^{\infty} v^{-(\alpha_1+1)}\,dv. \quad (25)$$

Using

$$\int v^{-(\alpha_1+1)}\,dv = -\left(\frac{1}{\alpha_1}\right)v_1^{-\alpha_1}, \quad (26)$$

we obtain

$$\alpha_1\lambda_1^{\alpha_1}\left[-\left(\frac{1}{\alpha_1}\right)v_1^{-\alpha_1}\right]_{\lambda_1}^{\infty} \quad (27)$$

$$\alpha_1\lambda_1^{\alpha_1}\left[0 + \left(\frac{1}{\alpha_1}\right)\lambda_1^{\alpha_1}\right] = 1. \quad (28)$$

Hence,

$$\int_0^{\infty}\int_0^{\infty} f_{X,Y}(x,y)\,dy\,dx = 1 \quad (29)$$

Thus, the proposed function is a valid bivariate probability density under the stated parameter restrictions.

The Theoretical Properties of the Model

The conditional cumulative distribution function of $Y \mid X = x$ is

$$F_{Y|X}(y \mid x) = 1 - \left(\frac{ax^b}{ax^b+y}\right)^{\alpha_2}. \quad (30)$$

Hence, the joint cumulative distribution function can be expressed as

$$F_{X,Y}(x,y) = \int_0^{x}\left[1 - \left(\frac{at^b}{at^b+y}\right)^{\alpha_2}\right]\frac{\alpha_1\lambda_1^{\alpha_1}}{(\lambda_1+t)^{\alpha_1+1}}\,dt. \quad (31)$$

For general values of b, this integral does not generally reduce to a simple closed-form expression and can therefore be evaluated numerically when required.

Similarly, the joint survival function is given by

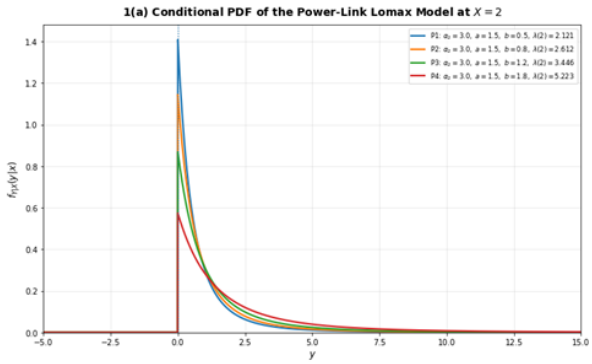
$$S_{X,Y}(x,y) = \int_x^{\infty}\left(\frac{at^b}{at^b+y}\right)^{\alpha_2}\frac{\alpha_1\lambda_1^{\alpha_1}}{(\lambda_1+t)^{\alpha_1+1}}\,dt. \quad (32)$$

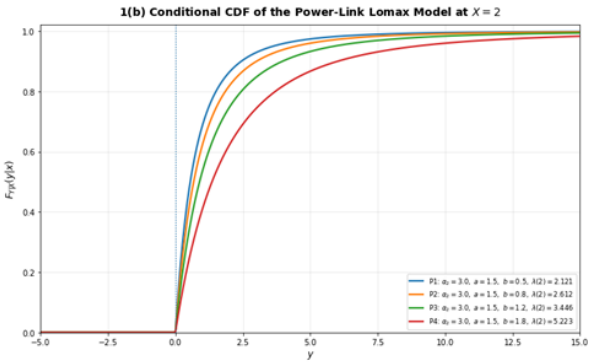
The corresponding joint hazard function is defined as

$$h_{X,Y}(x,y) = \frac{f_{X,Y}(x,y)}{S_{X,Y}(x,y)}. \quad (33)$$

Thus,

$$h_{X,Y}(x,y) = \frac{\frac{\alpha_1\alpha_2\lambda_1^{\alpha_1}a^{\alpha_2}x^{b\alpha_2}}{(\lambda_1+x)^{\alpha_1+1}(ax^b+y)^{\alpha_2+1}}}{\int_x^{\infty}\left(\frac{at^b}{at^b+y}\right)^{\alpha_2}\frac{\alpha_1\lambda_1^{\alpha_1}}{(\lambda_1+t)^{\alpha_1+1}}dt}. \quad (34)$$





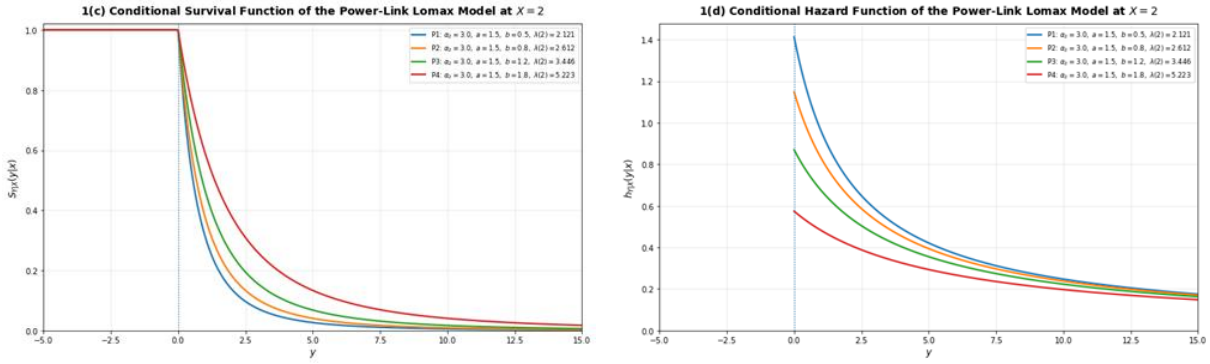


Figure 1: Plot of MBCLD-PLF (a) PDF, (b) CDF, (c) survival and (d) hazard functions at $X = 2$.

Figure 1(a) shows the conditional probability density, Figure 1(b) the conditional cumulative distribution function, Figure 1(c) the conditional survival function, and Figure 1(d) the conditional hazard function. The four plots illustrate how changes in the power-link parameter alter the conditional behaviour of Y for a fixed value of X.

The marginal distribution of Y is induced by the joint model and is obtained by integrating out X, $fY(y) = \int_0^\infty f_{X,Y}(x,y)dx$. In general, this integral does not have a simple closed form, so the marginal distribution of Y is evaluated numerically for chosen parameter values and is not assumed to remain Lomax in closed form. The conditional quantile function is available directly from the conditional Lomax distribution and shows how the conditional level of Y changes with x through the power-link scale.

Dependence can also be examined through tail behaviour. The upper and lower tail-dependence coefficients are defined in the usual way from the joint survival function and joint distribution function. For the proposed model, substitution of the derived conditional and marginal quantities leads to semi-closed-form expressions that require numerical evaluation in general. These quantities provide an additional way of studying whether extreme values of the two variables occur together.

The parameter roles are distinct: α_1 and λ_1 determine the marginal distribution of X; α_2 controls the conditional shape of Y given X; a is the positive scale constant; and b controls how the conditional scale changes with X. These distinct roles support practical parameter interpretation.

Maximum Likelihood Estimation

The parameters of the proposed MBCLD-PLF are estimated using the maximum likelihood estimation (MLE) method. Let the observed bivariate sample be

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), \quad (35)$$

where $x_i > 0$ and $y_i > 0$ for $i = 1,2, \dots, n$. The parameter vector of the proposed model is

$$\Theta = (\alpha_1, \lambda_1, \alpha_2, a, b), \quad (36)$$

where

$$\alpha_1 > 0, \lambda_1 > 0, \alpha_2 > 0, a > 0, b > 0.$$

Using the conditional specification approach, the joint density is written as

$$f_{X,Y}(x_i, y_i) = f_X(x_i) f_{Y|X}(y_i \mid x_i). \quad (37)$$

Therefore, the likelihood function is

$$L(\Theta) = \prod_{i=1}^n f_X(x_i) \prod_{i=1}^n f_{Y|X}(y_i \mid x_i). \quad (38)$$

Taking logarithm gives the log-likelihood function:

$$\ell(\Theta) = \sum_{i=1}^n \log f_X(x_i) + \sum_{i=1}^n \log f_{Y|X}(y_i \mid x_i). \quad (39)$$

For the marginal distribution of X,

$$f_X(x_i) = \alpha_1 \lambda_1^{\alpha_1} (\lambda_1 + x_i)^{-(\alpha_1+1)} \quad (40)$$

Hence,

$$\log f_X(x_i) = \log \alpha_1 + \alpha_1 \log \lambda_1 - (\alpha_1 + 1) \log(\lambda_1 + x_i). \quad (41)$$

So,

$$\sum_{i=1}^n \log f_X(x_i) = n \log \alpha_1 + n\alpha_1 \log \lambda_1 - (\alpha_1 + 1) \sum_{i=1}^n \log(\lambda_1 + x_i) \quad (42)$$

For the conditional distribution,

$$f_{Y|X}(y_i \mid x_i) = \alpha_2 a^{\alpha_2} x_i^{b\alpha_2} (a x_i^b + y_i)^{-(\alpha_2+1)}. \quad (43)$$

Therefore,

$$\log f_{Y|X}(y_i \mid x_i) = \log \alpha_2 + \alpha_2 \log a + b\alpha_2 \log x_i - (\alpha_2 + 1) \log\left(a x_i^b + y_i\right) \quad (44)$$

Summing over all observations gives

$$\sum_{i=1}^n \log f_{Y|X}(y_i \mid x_i) = n \log \alpha_2 + n\alpha_2 \log a + b\alpha_2 \sum_{i=1}^n \log x_i - (\alpha_2 + 1) \sum_{i=1}^n \log(a x_i^b + y_i). \quad (45)$$

Thus, the full log-likelihood function is

$$\begin{aligned} \ell(\Theta) = \; & n \log \alpha_1 + n\alpha_1 \log \lambda_1 - (\alpha_1 + 1) \sum_{i=1}^n \log(\lambda_1 + x_i) \\ & + n \log \alpha_2 + n\alpha_2 \log a + b\alpha_2 \sum_{i=1}^n \log x_i \\ & - (\alpha_2 + 1) \sum_{i=1}^n \log(a x_i^b + y_i). \end{aligned} \quad (46)$$

The maximum likelihood estimators are obtained by finding the parameter values that maximise $\ell(\Theta)$. That is, $\widehat{\Theta} = \left(\hat{\alpha}_1, \hat{\lambda}_1, \hat{\alpha}_2, \hat{a}, \hat{b}\right) = \arg\max_{\Theta} \ \ell(\Theta)$.

Because the likelihood equations do not have simple closed-form solutions, particularly because of the nonlinear power-link term, the parameters were estimated by constrained numerical optimisation. The optimisation was performed on the logarithms of the five parameters to preserve positivity. The primary optimiser was L-BFGS-B with parameter bounds of $log(0.20) - log(20.0)$ for α_1, λ_1 and α_2, $log(0.10) - log(20.0)$ for a, and $log(0.10) - log(5.0)$ for b. Eight starting values were used for each fit: one deterministic starting value and seven random perturbations. The L-BFGS-B settings included a maximum of 2,000 iterations, $ftol = 10^{-11}$, $gtol = 10^{-8}$ and $maxls = 50$; Powell optimisation was used as a fallback when necessary. Standard errors and confidence intervals were obtained from the inverse observed Hessian only when the numerical Hessian had positive eigenvalues and a condition number below 10^{12}. The simulation code was implemented in Spyder (Python 3.8.5) using NumPy, pandas and SciPy packages.

Monte Carlo Simulation Design

A Monte Carlo simulation study was carried out to evaluate the finite-sample behaviour of the maximum likelihood estimators of the proposed distribution. The parameter vector and admissible parameter space were defined. The true shape parameters were selected above two so that the corresponding Lomax moments required for stable summary assessment exist. The true parameter values were taken as $\alpha_1 = 2.50$, $\lambda_1 = 2.00$, $\alpha_2 = 3.00$, $a = 1.50$, $b = 1.20$ as representative positive values within the admissible parameter space of the proposed model (Bourguignon et al., 2022; Ríos-Ocampo et al., 2022). A nonlinear value of the power parameter was used to examine the proposed link away from the linear case. Samples of sizes 50, 100, 150, 300, 500 and 1000 were generated, with 1,000 independent replications at each sample size. Estimator performance was assessed using bias, mean squared error (MSE), root mean squared error (RMSE) and empirical coverage of nominal 95% confidence intervals. Confidence intervals were based on standard errors obtained from the inverse observed Hessian. The same likelihood specification and parameter constraints were applied in every replication.

Data and Empirical Analysis

For empirical assessment, the model was applied to the Comprehensive R Archive Network (CRAN) Loss–ALAE dataset containing 1,500 general-liability claims. Each record contains an indemnity payment (Loss) and its allocated loss adjustment expense (ALAE), such as legal and claim-investigation expenses. Descriptive statistics and Pearson, Spearman and Kendall dependence measures were computed before model fitting. The proposed model was compared with the Classical Bivariate Lomax, Bivariate Power Lomax–FGM, Bivariate Power Lomax–Sarmanov and Clayton Copula Bivariate Lomax models. Model selection used the maximised log-likelihood, AIC, BIC, CAIC and HQIC. Rosenblatt-transformed residuals were additionally assessed using Kolmogorov–Smirnov, Anderson–Darling and Cramér–von Mises statistics. Smaller information criteria and residual statistics indicate better performance under their respective criteria.

RESULTS AND DISCUSSION

Monte Carlo Results

Across the simulated sample sizes, estimation error decreased as the amount of information increased. The largest RMSE, MSE, and bias values were observed at $n = 50$, while considerably smaller values were obtained at $n = 500$ and $n = 1000$. The power-link parameter b showed relatively small estimation error across all sample sizes. Coverage probabilities were generally close to the nominal 0.95 level. Tables 1–4 summarise the principal finite-sample results.

Table 1: Bias of Parameter Estimates

Sample Size	α_1	λ_1	α_2	a	b
50	2.1852	2.2193	3.8434	2.4661	0.0003
100	0.8723	0.9643	1.6876	1.0904	0.0008
150	0.4807	0.5220	0.8143	0.5153	-0.0018
300	0.1653	0.1775	0.3598	0.2337	0.0008
500	0.1168	0.1241	0.1588	0.0982	-0.0005
1000	0.0330	0.0361	0.0904	0.0584	0.0012

Table 1 shows that the bias of the estimators generally decreased as the sample size increased. Relatively large positive bias was observed for the four distributional parameters at $n = 50$, particularly α_2, which had a bias of 3.8434. The bias of α_2 decreased to 0.0904 at $n = 1000$, while those of α_1, λ_1, and a decreased to 0.0330, 0.0361, and 0.0584, respectively. In contrast, the bias of the power-link parameter b remained close to zero throughout the simulation. The general decrease in bias indicates that the average estimates move closer to their true parameter values as the sample size increases.

Table 2: Mean Squared Error (MSE) of Parameter Estimates

Sample Size	α_1	λ_1	α_2	a	b
50	28.0629	29.1025	55.9277	24.1363	0.0167
100	7.6410	8.7052	18.8114	7.8748	0.0083
150	2.8574	3.4292	5.8717	2.3143	0.0055
300	0.4480	0.5217	1.5056	0.6018	0.0029
500	0.2069	0.2354	0.4366	0.1849	0.0016
1000	0.0780	0.0911	0.1766	0.0751	0.0008

The MSE results in Table 2 show a steady decline as sample size increases for each parameter, indicating lower combined squared bias and sampling variability. The largest MSE was recorded for α_2 at $n = 50$, with a value of 55.9277, but this decreased to 0.1766 at $n = 1000$. Similarly, the MSE of α_1 decreased from 28.0629 to 0.0780, while that of a declined from 24.1363 to 0.0751. The MSE of b was much smaller throughout the simulation, declining from 0.0167 at $n = 50$ to 0.0008 at $n = 1000$. Thus, the MSE results directly show improved finite-sample estimation as sample size increases.

Table 3: Root Mean Squared Error (RMSE) of Parameter Estimates

Sample Size	α_1	λ_1	α_2	a	b
50	5.2974	5.3947	7.4785	4.9129	0.1292
100	2.7642	2.9505	4.3372	2.8062	0.0908
150	1.6904	1.8518	2.4232	1.5213	0.0740
300	0.6693	0.7223	1.2270	0.7758	0.0538
500	0.4549	0.4851	0.6608	0.4300	0.0400
1000	0.2792	0.3019	0.4202	0.2741	0.0288

Table 3 shows a consistent reduction in RMSE for all five parameters as the sample size increased. The reduction was more pronounced for α_1, λ_1, α_2 and a, while the power-link parameter b maintained comparatively small RMSE values throughout. For example, the RMSE of α_2 decreased from 7.4785 at $n = 50$ to 0.4202 at $n = 1000$, while that of a decreased from 4.9129 to 0.2741. The RMSE of b decreased from 0.1292 to 0.0288. The comparatively small error for b indicates that it was estimated more stably than the other parameters under the selected simulation settings, although this does not establish the same behaviour under other parameter configurations. The relatively large RMSE values at n=50 also show that some parameters can be unstable in small samples.

Table 4: Coverage Probability of the 95% Confidence Intervals

Sample Size	α_1	λ_1	α_2	a	b
50	0.978	0.980	0.976	0.977	0.950
100	0.975	0.978	0.975	0.978	0.943
150	0.967	0.972	0.964	0.976	0.949
300	0.968	0.960	0.974	0.969	0.940
500	0.961	0.959	0.955	0.959	0.948
1000	0.969	0.965	0.954	0.953	0.943

Table 4 reports the empirical coverage probabilities of the nominal 95% confidence intervals. Coverage for the four distributional parameters remained generally close to the nominal level across the sample sizes considered. At $n = 1000$, the coverage probabilities were 0.969, 0.965, 0.954, and 0.953 for α_1, λ_1, α_2, and a, respectively. Coverage for the power-link parameter b ranged from 0.940 to 0.950 and also remained reasonably close to 0.95. With the decreasing RMSE, MSE, and bias values, these results indicate satisfactory finite-sample performance, particularly at the moderate and larger sample sizes considered.

Table 5: Optimisation Diagnostics

Sample Size	Success Rate	Hessian OK	Any Boundary	Powell Rate
50	1.000	1.000	0.231	0.000
100	1.000	1.000	0.052	0.000
150	1.000	1.000	0.012	0.000
300	1.000	1.000	0.001	0.000
500	1.000	1.000	0.000	0.000
1000	1.000	1.000	0.000	0.000

The optimisation diagnostics show a 100% success rate and 100% Hessian-validity rate at all sample sizes. The Powell fallback was not required in the simulation. However, the proportion of replications containing at least one estimate close to an optimisation boundary was 0.231 at $n = 50$ and decreased to 0.052, 0.012, 0.001 and 0.000 at $n = 100, 150, 300, 500\ and\ 1000$, respectively. This indicates that boundary behaviour was mainly a small-sample phenomenon rather than a failure of numerical convergence.

Loss–ALAE Data Characteristics

The Loss–ALAE dataset contained 1,500 complete positive claim pairs. Both variables were strongly right-skewed and displayed extreme observations, making the dataset appropriate for evaluating a heavy-tailed bivariate model. Table 6 presents the descriptive statistics.

Table 6: Descriptive Statistics of the Loss–ALAE Dataset

Variable	Loss	ALAE
N	1500	1500
Mean	41,208.42	12,588.16
Median	12,000	5,471
Standard Deviation	102,747.72	28,145.64
Minimum	10	15
Q1	4,000	2,333
Q3	35,000	12,571.75
Maximum	2,173,595	501,863
Skewness	9.173	9.259
Excess Kurtosis	142.651	125.136

The descriptive statistics clearly indicate that both variables are highly positively skewed. The mean values of both Loss and ALAE are considerably larger than their corresponding median values, indicating the presence of a relatively small number of very large insurance claims. For instance, the average Loss claim was 41,208.42, whereas the median claim was only 12,000. Likewise, the average ALAE was 12,588.16 compared with a median of 5,471. This difference between the mean and median confirms that both variables possess long right tails, which is a common characteristic of insurance loss data.

The estimated skewness coefficients of 9.173 for Loss and 9.259 for ALAE further confirm the presence of strong positive skewness. Similarly, the excess kurtosis values of 142.651 and 125.136 indicate extremely heavy tails relative to the normal distribution. These results justify the use of heavy-tailed probability distributions such as the Lomax family for modelling the data. Conventional light-tailed distributions would likely underestimate the probability of extremely large insurance claims and consequently produce poor actuarial risk estimates. To further examine the dependence between Loss and ALAE, three commonly used correlation measures were computed. The results are presented in Table 7

Table 7: Dependence Measures for the Loss–ALAE Data

Measure	Estimate	p-value
Pearson Correlation	0.4022	<0.001
Spearman Correlation	0.4519	<0.001
Kendall's Tau	0.3154	<0.001

The three dependence measures consistently indicate a statistically significant positive association between Loss and ALAE. The Pearson correlation coefficient of 0.4022 suggests a moderate positive linear relationship between the two variables. Since insurance loss data are highly skewed and contain extreme observations, rank-based measures provide additional information regarding dependence. The Spearman rank correlation coefficient of 0.4519 and Kendall's tau of 0.3154 are both statistically significant with p-values less than 0.001, confirming that larger insurance losses are generally associated with larger adjustment expenses. These findings support the need for a flexible bivariate model capable of simultaneously accommodating heavy tails and positive dependence.

Model Comparison

Table 8: Likelihood and Information-Criterion Comparison of Competing Models

Model	Log-Likelihood	AIC	BIC	CAIC	HQIC	Number of Parameters
Proposed MBCLD-PLF	−32163.163	64336.325	64362.891	64367.891	64346.222	5
Classical Bivariate Lomax	−32177.084	64360.169	64376.109	64379.109	64366.107	3
Bivariate Power Lomax–FGM	−32199.970	64413.940	64451.132	64458.132	64427.795	7
Bivariate Power Lomax–Sarmanov	−32212.589	64439.177	64476.370	64483.370	64453.033	7
Clayton Copula Bivariate Lomax	−32247.617	64505.234	64531.800	64536.800	64515.131	5

Table 8 compares the proposed model with four competing bivariate Lomax specifications. The comparison shows that the proposed MBCLD-PLF produced the largest maximised log-likelihood and the smallest AIC, BIC, CAIC and HQIC among the five fitted models. These information criteria indicate strong likelihood-based performance after accounting for model complexity. The result is particularly notable because the proposed model has five parameters, compared with three for the classical bivariate Lomax and seven for each of the Power Lomax–FGM and Power Lomax–Sarmanov models. Relative to the Classical Bivariate Lomax model, the proposed model reduced the AIC from 64,360.169 to 64,336.325, a difference of approximately 23.84 units. The lower information criteria indicate that the likelihood improvement associated with the conditional power-link construction was sufficient to offset its additional parameters under these criteria. Similar likelihood-based differences were observed relative to the other competing models.

Table 9: Parameter Estimates and Standard Errors of the Fitted Models

Model	Parameter	Estimate	Standard Error
Classical Bivariate Lomax	A	1.4862	0.0766
	λ_x	21,509.91	1,700.04
	λ_y	8,385.07	626.30
Bivariate Power Lomax-FGM	λ_1	1.2039	0.1396
	β_1	16,143.07	2,857.43
	γ_1	1.0234	0.0424
	λ_2	1.7453	0.2031
	β_2	10,931.98	1,714.10
	γ_2	1.1080	0.0387
	Θ	0.9993	0.0040
Clayton Copula Bivariate Lomax	α_1	1.2486	0.0745
	$Scale_1$	16,794.58	1,608.91
	α_2	2.3218	0.1888
	$Scale_2$	16,622.94	1,827.28
	Δ	0.5661	0.0443
Bivariate Power Lomax-Sarmanov	λ_1	1.2160	0.1440
	β_1	16,254.76	2,963.86
	γ_1	1.0112	0.0425
	λ_2	1.7062	0.2014
	β_2	10,428.71	1,671.54
	γ_2	1.1004	0.0396
	Ω	3.7149	0.1614
Proposed MBCLD-PLF	α_1	1.2377	0.0736
	λ_1	16,228.27	1,571.61
	α_2	4.0550	0.4554
	A	847.71	190.69
	B	0.3802	0.0185

The fitted models were further examined using parameter estimates and their standard errors, parameter counts, convergence information, and goodness-of-fit measures. The standard errors were obtained from the inverse of the observed numerical Hessian of the negative log-likelihood and transformed to the reported parameter scale using the delta method. The results for the estimated parameters are presented in Table 4.9. The results show that all estimated parameters of the five fitted models have finite standard errors, indicating that parameter uncertainty could be quantified for each fitted model. For the proposed MBCLD-PLF model, the estimates of $\alpha_1, \lambda_1, \alpha_2, a \text{ and } b$ were 1.2377, 16,228.27, 4.0550, 847.71 and 0.3802, respectively, with corresponding standard errors of 0.0736, 1,571.61, 0.4554, 190.69 and 0.0185. The relatively small standard errors compared with the corresponding estimates indicate reasonably precise parameter estimation. In particular, the standard error of the power parameter b is only 0.0185 relative to an estimate of 0.3802, showing that the estimated power relationship is reasonably well determined by the data. The Wald ratios, calculated as the parameter estimate divided by its standard error, were also large for the proposed model: 16.82 for α_1, 10.33 for λ_1, 8.90 for α_2, 4.45 for a and 20.58 for b. This indicates that the estimated parameters are clearly separated from zero under the usual Wald approximation.

The proposed model contains five estimated parameters, compared with three for the classical Bivariate Lomax and seven each for the Power Lomax-FGM and Power Lomax-Sarmanov models. Thus, the proposed model achieves its fit with substantially fewer parameters than the two seven-parameter alternatives. This difference in parameter complexity is important because the information criteria explicitly penalise models for additional parameters.

Table 4.9 also shows that all five models converged successfully using the L-BFGS-B optimisation algorithm. The proposed MBCLD-PLF model converged after 20 iterations and 150 function evaluations, with a Hessian condition number of 76.52. Its convergence message indicated successful convergence based on the relative reduction in the objective function. The standard errors were also available for all five estimated parameters. The condition number of the proposed model was relatively moderate compared with the Power Lomax-FGM model (28,728.42) and the Power Lomax-Sarmanov model (809.66), suggesting that the numerical Hessian for the proposed model was considerably better conditioned than those of the two seven-parameter alternatives. This provides additional support for the numerical stability of the proposed parameter estimates.

Table 10: Model Selection, Convergence and Rosenblatt Goodness-of-Fit Results

Model	KS	AD	CvM	KS Pooled	KS p-value
Proposed MBCLD-PLF	0.02056	1.77331	0.19444	0.0206	**0.1560**
Classical Bivariate Lomax	0.02301	1.55053	0.17748	0.0230	0.0821
Bivariate Power Lomax-FGM	0.02101	0.82854	0.12629	0.0210	0.1392
Bivariate Power Lomax-Sarmanov	0.01786	0.79793	0.13413	0.0179	0.2908
Clayton Copula Bivariate Lomax	0.01432	0.70536	0.09312	0.0143	0.5648

The pooled Rosenblatt goodness-of-fit test provides additional evidence concerning the adequacy of the fitted models. For the proposed model, the pooled Rosenblatt KS statistic was 0.0206 with an approximate p-value of 0.1560.

At the 5% significance level, this result does not provide evidence against the null hypothesis that the transformed observations follow the required uniform distribution. The proposed model therefore passes the pooled Rosenblatt KS diagnostic at the conventional 5% level. Its pooled KS p-value was also larger than that of the classical Bivariate Lomax (0.0821) and the Power Lomax-FGM model (0.1392), although it was lower than the corresponding values for the Power Lomax-Sarmanov (0.2908) and Clayton Copula (0.5648) models.

The Anderson-Darling and Cramér-von Mises statistics provide a more mixed picture. The proposed model produced an Anderson-Darling statistic of 1.7733 and a Cramér-von Mises statistic of 0.1944. These were larger than the corresponding values for the other fitted models. Thus, although the proposed model performs best according to the likelihood-based information criteria and passes the pooled KS test, it does not provide the smallest Rosenblatt Anderson-Darling or Cramér-von Mises statistics. This difference is important and suggests that the evidence of superiority is not uniform across every goodness-of-fit diagnostic.

Further examination of the separate Rosenblatt transformations gives an additional qualification. For the proposed model, the KS test for the first transformation U produced a statistic of 0.0401 with an approximate p-value of 0.0156, while the conditional transformation $V \mid U$ produced a statistic of 0.0427 with an approximate p-value of 0.0081. Both p-values are below 0.05. Therefore, while the pooled Rosenblatt KS test does not reject the fitted model, the separate transformation diagnostics indicate statistically detectable departures from the ideal uniform distribution in both components.

Overall, the results provide substantial but not completely uniform evidence in support of the proposed Power-Link Conditional Lomax model. The model has the highest maximised log-likelihood, the lowest AIC, BIC, CAIC and HQIC values, successful numerical convergence, finite and reasonably precise standard errors, and a well-conditioned numerical Hessian relative to the more highly parameterised alternatives. These features indicate that the proposed model provides a strong balance between fit, parsimony and parameter stability. However, its relatively larger Anderson-Darling and Cramér-von Mises statistics and the significant separate KS tests for U and $V \mid U$indicate that the model should not be described as providing a uniformly superior fit under every diagnostic criterion.

Discussion

The simulation results indicate that the maximum likelihood procedure becomes more accurate as sample size increases. This pattern is consistent with simulation findings reported for related bivariate Lomax extensions (Qura et al., 2023; Fayomi et al., 2024; Abd Elgawad et al., 2025). At the same time, the large RMSE and MSE values for some parameters at $n = 50$ show that the proposed model is not free from small-sample instability. For the Loss–ALAE data, the proposed model achieved the highest maximised log-likelihood and the lowest AIC, BIC, CAIC and HQIC among the five fitted models. The AIC difference relative to the Classical Bivariate Lomax model was about 23.84 units. This indicates that allowing the conditional scale of ALAE to change with Loss captures useful structure under the likelihood-based criteria. The mechanism differs from the Clayton copula and Sarmanov constructions (Fayomi et al., 2024; Abd Elgawad et al., 2025) because dependence is introduced directly through a conditional distributional parameter rather than through a separate copula.

The residual diagnostics provide an important qualification. The Clayton Copula Bivariate Lomax model had the smallest KS, Anderson–Darling and Cramér–von Mises statistics, and the Sarmanov model also had smaller values than the proposed model. Thus, the evidence differs by diagnostic: the proposed model performed most strongly on the likelihood and information criteria, while Clayton and Sarmanov performed better on the reported Rosenblatt-residual statistics. The proposed distribution is therefore best interpreted as a competitive conditional alternative with strong likelihood-based performance for heavy tailed data.

Conclusion

This study developed a MBCLD-PLF by allowing the scale parameter of the conditional Lomax component to vary with the conditioning variable through a power function. Maximum likelihood estimation was implemented through constrained numerical optimisation, and the Monte Carlo results showed improved estimation as sample size increased. In the Loss–ALAE application, the proposed model achieved the highest maximised log-likelihood and the lowest AIC, BIC, CAIC and HQIC, whereas the Clayton and Sarmanov models produced smaller Rosenblatt-residual goodness-of-fit statistics. The proposed distribution therefore provides a competitive conditional alternative for modelling dependent heavy-tailed data, with strong likelihood-based performance in the application considered. Its broader performance should be established through further theoretical study and applications to other datasets.

Acknowledgements

The authors acknowledge the custodians of the publicly available CRAN Loss–ALAE dataset used for the empirical illustration.

Data Availability

The Loss–ALAE data are publicly available through the R package evd and related actuarial-data repositories.

REFERENCES

Abd Elgawad, M. A., Alawady, M. A., Barakat, H. M., Mansour, G. M., Husseiny, I. A., Alyami, S. A., Hashem, A. F., & Mohamed, M. O. (2025). Bivariate power Lomax Sarmanov distribution: Statistical properties, reliability measures, and parameter estimation. *Alexandria Engineering Journal, 113*, 593–610. https://doi.org/10.1016/j.aej.2024.10.074

Arnold, B. C., & Strauss, D. J. (1991). Bivariate distributions with conditionals in prescribed exponential families. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *53*(2), 365-375.

Arnold, B. C., & Sarabia, J. M. (2022). Conditional specification of statistical models: classical models, new developments and challenges. *Journal of Multivariate Analysis*, *188*, 104801.

Arnold, B. C., Castillo, E., & Sarabia, J. M. (1999). Conditional specification. In *Conditionally Specified Distributions* (pp. 1-6). New York, NY: Springer New York.

Arnold, B. C., & Strauss, D. (1988). Bivariate distributions with conditionals in prescribed exponential families. Journal of the Royal Statistical Society: Series B, 50(3), 540–544.

Baba, L. Y., Bawa, M. U., & Rashid, B. (2025). Derivation of the Sine Exponentiated Lomax Distribution for Modelling Right Skewed and Heavy Tailed Data. *FUDMA Journal of Sciences*, *9*(12), 859-874. https://doi.org/10.33003/fjs-2025-0912-4495

Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. L. (2006). *Statistics of extremes: theory and applications*. John Wiley & Sons.

Bourguignon, M., Gallardo, D. I., & Gómez, H. J. (2022). A Note on Pareto-type distributions parameterized by its mean and precision parameters. *Mathematics*, *10*(3), 528.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65(1), 141–151. https://doi.org/10.1093/biomet/65.1.141

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer.

Fayomi, A., Refaey, R. M., AL-Dayian, G. R., EL-Helbawy, A. A., Mahmoud, A. A. M., & Abd Elaal, M. (2024). Bivariate Lomax distribution based on Clayton copula: Estimation and prediction. Journal of Statistics Applications & Probability, 13(4), 1321–1337. https://doi.org/10.18576/jsap/130415

Helu, A., & Samawi, H. (2022). Inference on P(X < Y) in bivariate Lomax model based on progressive type II censoring. *PLOS ONE, 17*(5), e0267981.

Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association, 49*(268), 847–852. https://doi.org/10.1080/01621459.1954.10501239

Qura, M. E., Fayomi, A., Kilai, M., & Almetwally, E. M. (2023). Bivariate power Lomax distribution with medical applications. *PLOS ONE, 18*(3), e0282581. https://doi.org/10.1371/journal.pone.0282581

Ríos-Ocampo, J., & Gary, M. S. (2022). Using analytical equations to represent nonlinear relationships. *System Dynamics Review*, *38*(4), 354-370.

©2026 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via https://creativecommons.org/licenses/by/4.0/ which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.