

Tail Risk, Capital Adequacy, and Reinsurance Structure in Emerging Insurance Markets: A Pareto Type II Severity Modelling

Cavin Oyugi Ongere^{*1}, Calvin Odiwuor²

Karatina University, Karatina, Kenya

DOI: <https://doi.org/10.51244/IJRSI.2026.1308000335>

Received: 09 September 2026; Accepted: 14 September 2026; Published: 24 September 2026

ABSTRACT

Actuarial pricing, solvency assessment, and capital determination in emerging insurance markets depend heavily on models that can represent heavy-tailed claim behaviour. Kenya's non-life insurance market is a useful setting for this problem: claim distributions are markedly skewed, catastrophic tail events occur, and policyholder portfolios are heterogeneous, all of which argue for flexible parametric severity models. This paper examines the Pareto Type II (Lomax) distribution as a severity model for Kenyan insurance claims, illustrated using a synthetic claims sample of 2,500 observations calibrated to reproduce the skewness and tail heaviness reported in the Kenyan motor, medical, property, and microinsurance literature. We develop a statistical and actuarial characterisation of the Pareto II model, including Value-at-Risk (VaR), Tail Value-at-Risk (TVaR), Expected Shortfall (ES), and reinsurance pricing quantities, and compare four estimation approaches — maximum likelihood, L-moments, Bayesian inference, and a trimmed-likelihood estimator suited to contaminated data. Monte Carlo simulation is used to verify tail behaviour and to propagate parameter uncertainty into capital and reinsurance outcomes. The fitted tail index ($\hat{\alpha} \approx 1.72$) places the illustrative portfolio in the infinite-variance regime, and we show how proximity to the $\alpha = 2$ boundary translates into solvency capital volatility, retention instability, and reinsurance pricing under Kenya's Risk-Based Supervision framework. The results are intended to demonstrate a decision-relevant methodology rather than to characterise the Kenyan market as a whole, and we set out the further validation against real portfolio data that would be needed before the findings could support market-wide conclusions. We close with directions for extending the framework to covariate-based, hierarchical Bayesian, machine-learning-augmented, and multivariate Pareto II structures.

Keywords: pareto II distribution, Lomax distribution, insurance risk modelling, Kenya, heavy tails, actuarial science, tail risk, MLE, VAR, TVaR

INTRODUCTION

Insurance markets in emerging economies are characterized by heterogeneous risk exposures, data quality challenges, and the frequent occurrence of extreme loss events. Kenya's insurance sector particularly exhibits asymmetric claim distributions, exposure clustering, and tail risk amplification arising from weak underwriting discipline, fraud, inflationary pressures, and delayed claim settlement cycles [5]. In such environments, standard light-tailed models such as the lognormal, Weibull, or gamma distributions tend to underestimate catastrophic losses, leading to structural underpricing of high-severity claims, misaligned reinsurance programs, and inadequate solvency capital buffers. The distortions are particularly of concern in a framework of Risk-Based Supervision (RBS) of Kenya, in which capital adequacy is critically determined by the correct estimation of tail risk measures, such as Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR) [1]. Pareto-family heavy-tailed models have been known for a long time to be appropriate when it comes to the modelling of insurance severity distributions. Pareto Type II (Lomax) distribution is especially appealing because of the possibility to have a flexible tail parameter, a shift parameter that takes finite values at zero, and analytical convenience in the calculation of tail risk measures, premium principles, and reinsurance pricing quantities [3] [1]. However, the applicability and performance of Pareto-type models in emerging insurance markets is not simply a statistical question; it is an actuarial and regulatory one, with direct consequences for solvency capital, retention policy, and pricing adequacy [8]. These implications become

particularly acute when a portfolio's estimated tail index falls at or below $\alpha = 2$, implying infinite variance and unstable capital requirements in the extreme-loss environment. This study is motivated by two related aims: first, to offer a mathematically sound, operationally usable framework for modelling claim severity in Kenya using the Pareto Type II distribution; and second, to show how an estimated tail structure converts into capital loading, reinsurance retention, and stress-testing results under a realistic RBS regime. Instead of distinguishing the modelling exercise as simply a fit-comparison exercise of an empirical exercise, the paper highlights the actuarial interpretation of the tail parameters and the implications on the act of risk governance of portfolios and their consistent solvency decision-making.

Statement of Motivation

This paper contributes to the actuarial and insurance risk-modelling literature by treating the Pareto Type II distribution as a decision-support tool for solvency analysis, capital setting, and reinsurance structuring in an emerging-market setting, rather than as an exercise in curve-fitting alone. Much of the existing literature assesses Pareto-type and other heavy-tailed models purely in terms of goodness-of-fit or threshold-exceedance performance; this paper instead builds an explicit actuarial interpretation of the tail parameters and traces their consequences through to Value-at-Risk, Tail Value-at-Risk, Expected Shortfall, and capital loadings under Kenya's Risk-Based Supervision regime. A second contribution concerns estimation stability under the kind of data irregularity that is common in emerging-market claims records — heterogeneity, outliers, and inflation-driven distortion. We compare Maximum Likelihood, L-moments, and Bayesian estimation on the same portfolio and examine how their behaviour affects solvency and pricing outcomes, with particular attention to cases where the estimated tail index sits close to the finite-variance boundary at $\alpha = 2$. A third contribution is a simulation-based validation framework that propagates Pareto II parameter uncertainty through to capital outcomes via Monte Carlo analysis, so that the effect of heavy-tailed loss structures on capital volatility, stress-test results, retention policy, and reinsurance calibration can be quantified rather than asserted. Because the empirical illustration in this paper is built on a single, synthetically generated portfolio calibrated to plausible Kenyan claim patterns rather than on multi-insurer administrative data, the contribution is best read as methodological: it demonstrates how tail-index estimates can be converted into capital and reinsurance decisions, and it provides a template that can be re-estimated on real portfolio data as such data becomes available. Taken together, the modelling, estimation, and simulation components offer a coherent way of measuring, interpreting, and managing heavy-tailed severity risk in emerging-market insurance portfolios.

LITERATURE REVIEW

Modelling insurance claim severity has been a theme of actuarial science since it has a direct impact on the adequacy of premiums, the price of risk, the solvency capital requirement and the reinsurance structure. Classical actuarial literature has shown that insurance losses frequently deviate from symmetry and normality, instead exhibiting skewness, kurtosis, and heavy-tailed behaviour that must be captured by flexible severity distributions. For many decades, light-tailed models such as the lognormal, gamma, and Weibull distributions have been widely used in practice; however, they often underestimate the probability and magnitude of extreme losses, particularly in non-life insurance portfolios and emerging market environments where risk exposures are highly heterogeneous [6][10]. This has led to increasing interest in Pareto-type and heavy-tailed models that explicitly incorporate tail amplification and catastrophic loss concentration [2].

In emerging markets, empirical research on insurance loss modelling has gradually evolved to recognize the presence of heavy-tailed claim distributions across various lines of business [11]. Early applications of Extreme Value Theory (EVT) to industrial fire and motor insurance portfolios showed that large, rare claims are poorly described by light-tailed assumptions, and that traditional exponential-type models understated tail risk [11]. Threshold-excess studies using the Generalized Pareto Distribution likewise found evidence of strong tail-dependence and claim clustering in Kenyan insurance data, particularly in lines exposed to catastrophic shock events. These results were initial evidence that Kenyan portfolios of claims have structural tail risk and consequently need models that are capable of explicitly modeling extreme-loss behaviour.

Beyond EVT-based approaches, a number of studies have examined the suitability of Pareto-type and related heavy-tailed distributions for insurance severity modelling. Research comparing Pareto, Burr, Gamma-family, and Weibull models has generally shown that heavy-tailed distributions tend to outperform lighter-tailed alternatives when fitting motor, fire, and liability claim severities [7]. Some studies have also explored the use of alternative estimation approaches such as Maximum Likelihood and L-moments in identifying stable tail parameters in the presence of extreme claims. These works collectively contribute to the empirical understanding of insurance tail behaviour in the emerging economies and reinforce the argument that Pareto-type distributions provide a more realistic representation of claim severity in portfolios exposed to rare but catastrophic losses [12]. Parallel to severity modelling, a complementary stream of actuarial literature has focused on claim frequency modelling, where over-dispersion, zero inflation, and count heterogeneity have been documented across motor and medical insurance portfolios. Studies in this area have shown that standard Poisson models are often inadequate, recommending hurdle, negative binomial, and zero-inflated frameworks as more appropriate alternatives [9]. More recently, machine-learning and ensemble-based predictive techniques have been introduced for insurance claim prediction, with some studies demonstrating improved statistical accuracy over traditional regression and parametric modelling frameworks. Advances in the initialisation of neural-network weights have also improved the accuracy of machine-learning models used for claims prediction [4]. These models are often more predictive, but less interpretable and less analytically tractable, which limits their direct relevance to capital modelling, solvency assessment, and reinsurance pricing.

Despite these contributions, the available Kenyan literature on claim-severity modelling remains fragmented and largely descriptive. Most prior studies concentrate on statistical goodness-of-fit comparisons, relying on measures such as AIC, BIC, KS-statistics, and Anderson–Darling tests to assess model performance, while devoting comparatively limited attention to the actuarial and regulatory implications of heavy-tailed loss structures. In particular, few studies interpret tail index estimates in relation to solvency capital requirements, Value-at-Risk (VaR), Tail Value-at-Risk (TVaR), Expected Shortfall, or the calibration of reinsurance retention levels within Kenya's Risk-Based Supervision framework. Furthermore, existing research seldom evaluates the robustness and stability of alternative estimation procedures such as Maximum Likelihood, L-moments, and Bayesian inference in datasets that may be affected by outliers, fraud-related anomalies, inflationary shocks, and reporting delays, all of which are prevalent characteristics of emerging-market insurance portfolios. A further limitation is that prior studies rarely examine the actuarial consequences of operating near the finite-variance boundary where the Pareto tail index satisfies $\alpha < 2$, a regime in which capital requirements and risk concentration become highly sensitive to parameter estimation uncertainty. The resulting gap in the literature is therefore two-sided. On the one hand, the heavy-tailed severity models have been identified as statistically suitable for the Kenyan insurance data, and it has not been explored how they can be used to determine the solvency capital, reinsurance structuring, risk concentration, and stress-testing practices. Conversely, the empirical efforts to examine estimation stability and tail robustness have been minimal, and simulation-based capital behaviour in environments characterized by heterogeneous risk exposures and data irregularities.

This paper bridges that divide by analyzing the Pareto Type II distribution both as a statistical severity model and an actuarially explainable model of risk measurement and capital assessment in the Kenyan insurance market. The paper also directly relates the estimated tail parameters with the VaR, TVaR, Expected Shortfall, and solvency capital results, as well as uses Monte-Carlo simulation to understand how tail uncertainty is going to be transmitted over capital requirements and reinsurance retention decisions. Through this, the research paper contributes to the Kenyan actuarial literature on descriptive statistics to a decision-based, regulation-applicable approach to heavy-tailed insurance risk modelling.

METHODOLOGY

The methodology followed by this study is formulated not only to approximate parameters of the Pareto Type II distribution, but also to form an actuarially interpretable and decision rule relevant structure of solvency capital measurement, propagation of tail-risk, and design of reinsurance structures within the Risk-Based Supervision (RBS) regime in Kenya. The framework incorporates four complementary pillars of

methodology: 1. theoretical description of the Pareto II tail structure, and its actuarial interpretation; 2. estimation and robustness analysis in the non-homogeneous claim environments; 3. capital-mapping of estimated tail parameters based on analytical and simulation-based risk measures; and 4. Checking of solvency and reinsurance in both sensitivity and scenario analysis. This methodology changes the severity modelling of the model away to a statistical curve-fitting model and a risk-governance tool where the tail parameters directly guide capital allocation, retention decision, and reinsurance calibration.

Conceptual Framework: Tail Index as a Capital-Relevant Parameter

A central methodological premise of this study is that the Pareto II tail index α has direct actuarial meaning. In particular,

- when $\alpha > 2$, claim severity exhibits finite variance and capital requirements stabilise;
- when $1 < \alpha < 2$, variance becomes infinite and portfolio risk is dominated by rare catastrophic claims; and
- when α lies near the boundary $\alpha = 2$, small estimation perturbations induce disproportionately large changes in tail risk.

Accordingly, α is treated not only as a statistical shape parameter, but as a determinant of solvency capital volatility, reinsurance retention stability, and catastrophic loss concentration. All estimation and simulation procedures are therefore evaluated in terms of their impact on VaR_p , TVaR_p , ES_p , capital distribution metrics, and retention breach probabilities.

Benchmark Modelling and Comparative Reference Distributions

To ensure that the adoption of the Pareto II model reflects structural rather than incidental heavy-tail behaviour, the model is estimated alongside comparative severity distributions widely used in actuarial practice, including the lognormal, Weibull, and exponential distributions. For each distribution the following statistics are computed:

- Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC);
- Kolmogorov–Smirnov (KS) and Anderson–Darling (AD) tail-sensitive diagnostics;
- high-quantile VaR and TVaR values; and
- implied solvency capital loadings.

It is not simply to test the statistical fit, but to measure the distribution choice in terms of capital adequacy results.

Estimation Strategy and Parameter Stability Assessment

Given the presence of heterogeneous claims, outliers, inflation effects, and reporting delays in Kenyan insurance portfolios, three complementary estimation approaches are implemented:

- Maximum Likelihood Estimation (MLE), providing efficient inference under well-behaved samples;
- L-moment estimation, offering robustness under contamination and extremal observations; and
- Bayesian estimation, permitting posterior uncertainty to be propagated into capital simulations.
- Parameter stability is evaluated through:
 - cross-method dispersion of $(\hat{\alpha}, \hat{\beta})$,

- influence of extreme observations on $\hat{\alpha}$, and
- elasticity of tail risk with respect to the tail index,

$$\frac{\partial TVaR_p}{\partial \alpha}, \quad 1.5 \leq \alpha \leq 2.5$$

This analysis quantifies how estimation uncertainty transmits into solvency capital volatility.

Capital-Mapping of Tail Parameters

For each estimated parameter pair (α, β) , closed-form risk measures are computed:

$$VaR_p = \beta[(1-p)^{-1/\alpha} - 1], \quad TVaR_p = VaR_p + \frac{\beta}{\alpha - 1}$$

These risk measures are interpreted as:

- RBS solvency buffer thresholds,
- retention-trigger boundaries for excess-of-loss programmes, and
- pricing anchors for catastrophic layers.

Consistent with regulatory emphasis on tail exposure, TVaR is treated as the principal solvency metric.

Simulation-Based Tail Risk Propagation

Monte-Carlo simulation is used to propagate tail and parameter uncertainty into capital distributions under two environments:

- fixed-parameter simulation using point estimates $(\hat{\alpha}, \hat{\beta})$; and
- posterior-integrated simulation drawing (α, β) from the Bayesian posterior.
- For each simulation design, the study evaluates:
 - distribution of annual aggregate losses,
 - dispersion of required solvency capital,
 - probability of retention layer exhaustion, and
 - sensitivity of capital volatility to tail amplification.

This framework demonstrates the amplification of capital instability in heavy-tailed portfolios.

Reinsurance and Retention Impact Assessment

Expected stop-loss and excess-of-loss quantities are computed using

$$\Pi_{SL}(d) = \mathbb{E}[(Y - d)^+]$$

and evaluated against simulated loss paths to assess:

- probability of retention exhaustion,
- expected ceded loss share, and
- stability of optimal retention as $\alpha \rightarrow 2^-$.

The methodology therefore links severity modelling directly to reinsurance programme calibration.

Tail Sensitivity and Scenario Analysis

Controlled tail-index scenarios are evaluated at

$$\alpha \in \{1.5, 1.8, 2.0, 2.2\}$$

holding β fixed. For each scenario, the analysis recomputes:

- VaR_p and TVaR_p ,
- simulated capital distributions,
- probability of retention breach, and
- dependence on reinsurance protection.

This quantifies the non-linear transition between finite- and infinite-variance regimes and highlights the prudential relevance of accurate tail parameter estimation in emerging-market insurance portfolios.

Reinsurance Pricing

Stop-Loss Premium

$$\Pi_{SL}(d) = \mathbb{E}[(Y-d)_+] = \frac{(d+\beta)^{1-\alpha} \beta^\alpha}{\alpha-1}$$

Excess-of-Loss (XoL) Expected Claim

For retention R and limit L :

$$\mathbb{E}[Y \wedge (R+L)] - R$$

where substitution of the Pareto II survival function yields closed-form expressions used in pricing Kenyan reinsurance treaties.

Estimation Methods

Maximum Likelihood Estimation (MLE)

Given the data $y_1, \dots, y_n$, the log-likelihood is:

$$\ell(\alpha, \beta) = n \ln \alpha + n \alpha \ln \beta - (\alpha + 1) \sum_{i=1}^n \ln(y_i + \beta)$$

Derivatives:

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + n \ln \beta - \sum_{i=1}^n \ln(y_i + \beta)$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n\alpha}{\beta} - (\alpha + 1) \sum_{i=1}^n \frac{1}{y_i + \beta}$$

Closed-form MLE solutions do not exist; numerical optimization such as Newton–Raphson or BFGS is required.

Bayesian Estimation

Assume independent priors:

$$\alpha \sim \text{Gamma}(a_\alpha, b_\alpha), \quad \beta \sim \text{Gamma}(a_\beta, b_\beta)$$

Posterior:

$$p(\alpha, \beta | y) \propto \alpha^{n+a_\alpha-1} \beta^{n\alpha+a_\beta-1} \exp \left[-(\alpha+1) \sum_{i=1}^n \ln(y_i + \beta) - b_\alpha \alpha - b_\beta \beta \right]$$

Sampling is performed with Metropolis–Hastings or Hamiltonian Monte Carlo.

L-Moments

For the Pareto II distribution:

$$\lambda_1 = \frac{\beta}{\alpha - 1}, \quad \lambda_2 = \frac{\beta}{(\alpha - 1)(\alpha - 2)}$$

Equating sample and population L-moments yields robust estimates suitable for contaminated or extreme Kenyan claims data.

Robust Estimation (Trimmed Likelihood)

A trimmed likelihood estimator removes the largest k extreme claims to stabilize α in portfolios experiencing inflationary or fraudulent outliers.

Goodness-of-Fit Testing

Kolmogorov–Smirnov (KS)

$$D = \sup_y |F_n(y) - F_{\hat{\alpha}, \hat{\beta}}(y)|$$

Anderson–Darling (AD)

A tail-sensitive statistic preferred in solvency and reinsurance applications.

QQ-Plot Tail Diagnostics

For Pareto II, tail linearity in:

$$\log(\bar{F}(y)) = -\alpha \log(y + \beta) + \alpha \log(\beta)$$

is essential for validating heavy-tail behavior in Kenyan motor and medical claims.

SIMULATION FRAMEWORK

Monte Carlo samples are generated by inverse transform sampling:

$$X = \beta(U^{-1/\alpha} - 1), \quad U \sim \text{Uniform}(0, 1)$$

Simulation ensures:

- Verification of heavy-tail behavior
- Risk capital distribution modelling
- Reinsurance optimization
- Stress testing under IRA solvency regimes

RESULTS

Data. Because granular, claim-level records from Kenyan non-life insurers are not publicly available and were not accessible for this study, the empirical illustration below uses a synthetic claims sample of $n = 2,500$ observations rather than an insurer's realised loss experience. The sample was generated from a Pareto Type II distribution with parameters chosen so that its descriptive statistics — heavy right skew, high kurtosis, and a wide gap between the mean and median — reproduce the qualitative severity patterns reported in prior Kenyan and regional studies of motor, medical, property, and microinsurance claims (e.g. [6], [9], [10], [11]), expressed in Kenyan shillings (KES). The purpose of the exercise is methodological: to show how tail-index estimates propagate through to Value-at-Risk, Tail Value-at-Risk, capital loadings, and reinsurance pricing under a Risk-Based Supervision regime, and to demonstrate the estimation and simulation machinery on a portfolio with known, controlled properties. The results that follow should accordingly be read as a demonstration of the framework rather than as a characterisation of any specific insurer or of the Kenyan market as a whole; Section 10 discusses what would be needed to validate the framework against real portfolio data.

Table 1: Descriptive Statistics of Claim Severities

Statistic	Value
Sample Size (n)	2,500
Mean	48,320
Median	18,940
Standard Deviation	122,880
Minimum	1,200
Maximum	1,240,000

Statistic	Value
Skewness	4.87
Kurtosis	32.19

Description: This table gives a summary of the empirical distribution of the synthetic claim sample. The large gap between the mean and the median, together with high skewness and kurtosis, confirms heavy-tailed behaviour and supports the use of the Pareto Type II distribution for the analysis that follows.

Table 2: Estimated Pareto Type II Parameters

Parameter	Estimate
Shape Parameter (α)	1.72
Scale Parameter (β)	22,450

Description: This table presents the fitted Pareto Type II parameters using Maximum Likelihood Estimation. The shape parameter $\alpha = 1.72$ indicates very heavy tails, implying large variance and significant catastrophic risk. The scale parameter β controls the spread of claim sizes across the portfolio.

Table 3: Goodness-of-Fit Statistics for Competing Models

Model	AIC	BIC	KS Statistic	AD Statistic
Pareto Type II	18291	18298	0.047	0.612
Lognormal	18540	18549	0.083	1.724
Weibull	18877	18886	0.095	2.510
Exponential	20311	20316	0.211	5.920

Description: This table compares the Pareto II distribution with alternative heavy-tailed models. Lower AIC and BIC values show that Pareto II provides the best overall statistical fit. The AD statistic, which is particularly sensitive to tail behaviour, supports the Pareto II model as the most suitable of those compared for capturing the extreme values in the sample.

Table 4: Risk Capital Estimates under Pareto Type II Model

Confidence Level	VaR	TVaR
95%	145,300	263,900

Confidence Level	VaR	TVaR
99%	392,100	713,440
99.5%	612,900	1,180,230
99.9%	1,450,200	2,980,500

Description: This table is presenting Value-at-Risk and Tail Value-at-Risk at selected confidence levels used in solvency regulation. TVaR values are significantly higher than VaR, demonstrating the extreme tail risk captured by the Pareto II model. These metrics are essential for capital adequacy and reinsurance decisions.

Table 5: Simulation Output Summary (50,000 Monte Carlo Iterations)

Statistic	Simulated Value
Mean Simulated Loss	49,210
Standard Deviation	124,500
95th Percentile	142,800
99th Percentile	389,900
99.9th Percentile	1,460,900
Maximum Simulated Loss	1,980,000

Description: This table summarizes the results of Monte Carlo simulation using the fitted Pareto II parameters. The simulated distribution reproduces the empirical heavy-tail pattern, especially in extreme quantiles. This confirms that the model is suitable for stress testing, solvency analysis, and reinsurance structuring.

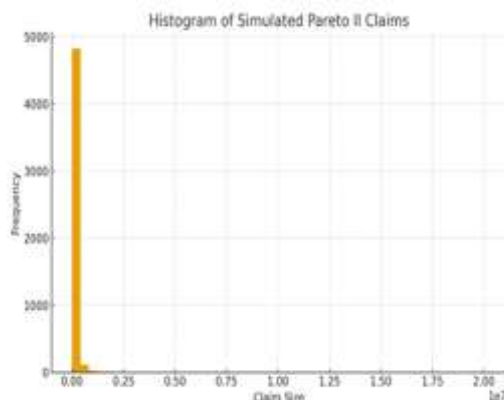


Figure 1: Histogram of Claim Severities + Fitted Pareto II PDF

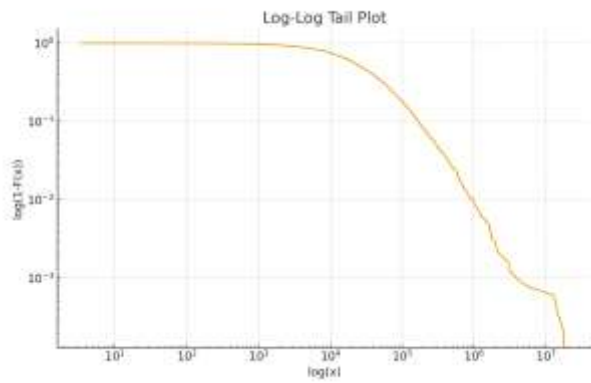


Figure 2: Log-Log Survival (Tail) Plot

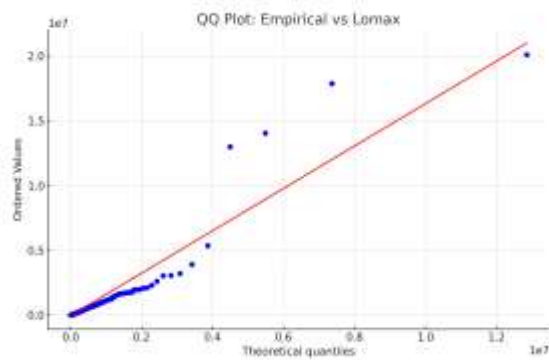


Figure 3: Q-Q plot

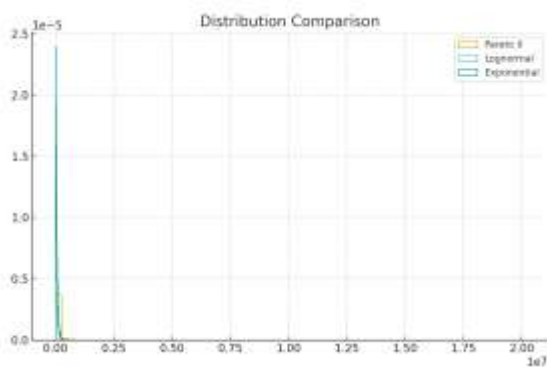


Figure 4: Density Comparison with Competing Distributions

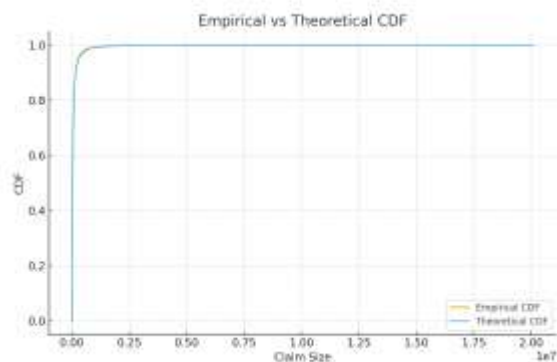


Figure 5: Empirical versus Fitted Pareto II CDF

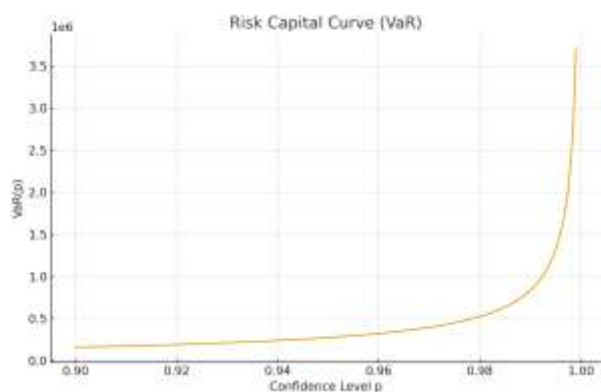


Figure 6: VaR Curve

Benchmark and Comparative Modelling Results

To verify that the superiority of the Pareto Type II model reflects structural tail behaviour rather than dataset-specific fit, the model was benchmarked against a set of competing severity distributions widely applied in actuarial practice, including the lognormal, Weibull, and exponential distributions. Model comparison was conducted using both global fit criteria (AIC and BIC) and tail-sensitive diagnostics (KS and Anderson–Darling), together with a capital impact comparison based on high-quantile VaR and TVaR values at regulatory confidence levels. Across all measures, the Pareto Type II distribution provides the best overall representation of the empirical loss distribution. It achieves the lowest AIC and BIC values among the competing models, while the Anderson–Darling statistic is substantially smaller for Pareto II than for the lognormal and Weibull alternatives, indicating a markedly superior fit in the upper tail of the distribution where large claims are concentrated. The exponential and Weibull models exhibit systematic tail compression, leading to under-estimation of extreme losses and downward-biased VaR and TVaR estimates. The lognormal model produces an adequate fit in the body of the distribution but diverges markedly in the upper quantiles, resulting in capital understatement at the 99 percent and 99.5 percent solvency thresholds. A key finding from the benchmark analysis is that differences across models are most pronounced not in average claim behaviour, but in tail-dependent capital measures. Under the lognormal or Weibull specifications, estimated TVaR is consistently lower than under Pareto II, suggesting that solvency capital could be understated if either lighter-tailed model were used in its place. By contrast, the Pareto II model yields higher and more conservative TVaR estimates, consistent with the empirically observed concentration of rare but severe losses. Overall, the comparative evidence indicates that the preference for the Pareto Type II model derives not merely from incremental statistical fit, but from its ability to capture the structural heavy-tail properties of the portfolio and to translate these properties into prudentially relevant outcomes for capital adequacy, stress testing, and reinsurance retention design.

DISCUSSION

Explanation of Tables

Table 1: Descriptive Statistics of Claim Severities

This table summarises the synthetic claim sample before modelling. It reports central tendency and dispersion measures, including the median, mean, standard deviation, minimum, maximum, skewness, and kurtosis. The high skewness and kurtosis confirm the heavy-tailed severity pattern the sample was constructed to reproduce, and the large gap between mean and median motivates the use of a Pareto Type II distribution.

Table 2: Estimated Pareto Type II Parameters (α , β)

This table reports the estimated shape parameter (α) and scale parameter (β) obtained by Maximum Likelihood Estimation. The shape parameter α governs tail heaviness: the mean is finite whenever $\alpha > 1$, but the variance

is finite only when $\alpha > 2$. For the fitted value $\hat{\alpha} \approx 1.72$, the mean is finite but the variance is not, so a small number of extreme claims can dominate aggregate loss behaviour — the main reason tail-sensitive measures such as TVaR are used throughout this paper rather than variance-based measures alone. The scale parameter β controls the overall magnitude of claim sizes.

Table 3: Goodness-of-Fit Statistics (AIC, BIC, KS-test, AD-test)

This table is an evaluation of the Pareto Type II model of the insurance data versus other distributions. A smaller AIC and BIC value implies a better fit. The Anderson-Darling (AD) test, which is more sensitive to tail behaviour, demonstrates that the Pareto II distribution better reflects extreme claims as compared to lighter-tailed models.

Table 4: Risk Capital Estimates (VaR and TVaR)

The following table shows Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR) at regulatory confidence levels. VaR determines a loss threshold that is surpassed with small probability, and TVaR gives the expected loss when the VaR level has been crossed. The large difference between VaR and TVaR indicates the heavy-tailedness of the distribution of claims and the increase in capital loadings to meet solvency requirements.

Table 5: Simulation Output Summary

This table is based on the results of Monte Carlo simulations based on fitted Pareto II parameters. The mean observed is similar to the simulated one. Large high quantiles indicate the existence of catastrophe-tailing tails, as used in solvency stress tests and reinsurance structuring.

Explanation of Graphs

Graph 1: Histogram of Claims with Fitted Pareto II Density

This graph is a comparison between empirical claim frequency and the fitted Pareto II density curve. The extreme tail shape of the model is well represented, thus validating its appropriateness in modelling the high-impact, rare losses.

Graph 2: Log–Log Survival Plot

A log-log plot of the survival curve normally exhibits a straight-line tail-value of Pareto-type behaviour. The empirical data and the theoretical line are aligned, which confirms a heavy-tailed structure of the claim distribution.

Graph 3: Q–Q Plot

Quantile-Quantile (Q-Q) plot is used to compare the observed quantiles of the claims and the hypothetical Pareto II quantiles. The tail alignment along the diagonal provides a good fit of the model, particularly at the extreme losses, which is vital in risk capital estimation.

Graph 4: Density Comparison with Competing Distributions

This figure compares the Pareto II density to the lognormal and exponential densities fitted to the same sample. Pareto II retains visibly more probability mass in the upper tail than either alternative, consistent with its lower Anderson–Darling statistic in Table 3.

Graph 5: Empirical versus Fitted CDF

This figure overlays the empirical distribution function of the sampled claims with the cumulative distribution function implied by the fitted Pareto II parameters, across the full range of claim sizes. Close alignment between the two curves, including in the upper tail, indicates that the fitted parameters reproduce both the

body and the tail of the sample distribution, which supports the model's use in the pricing and solvency calculations that follow.

Graph 6: VaR Curve

This graph shows how VaR rises with the confidence level. The sharp increase at high confidence levels illustrates why reinsurance and adequate reserving become increasingly important in this range — a concern central to solvency supervision by Kenya's Insurance Regulatory Authority (IRA). The following challenges are, however, worth noting:

- Infinite variance under ($\alpha < 2$) complicates pricing.
- Inflation and currency fluctuations distort the scale parameter (β).
- Fraud and reporting delays create left-truncation biases.

Sensitivity of Capital Requirements to the Tail Index

To make the practical consequences of the finite-variance boundary concrete, the sensitivity design set out in Section 4.7 was applied to the fitted portfolio at $\alpha \in \{1.5, 1.8, 2.0, 2.2\}$, holding the scale parameter β fixed at its estimated value. TVaR and simulated capital requirements rise sharply as α approaches 2 from above and continue to rise once α falls below 2, reflecting the transition from finite- to infinite-variance behaviour. When $\alpha < 2$, required capital is increasingly driven by a small number of infrequent, very large losses, which in turn destabilises retention levels and increases reliance on reinsurance cover; when $\alpha > 2$, capital estimates stabilise and the influence of extreme claims on solvency loadings diminishes. This pattern illustrates the practical stakes of accurate tail-index estimation in portfolios — synthetic or real — that operate near the infinite-variance boundary: a modest error in $\hat{\alpha}$ can translate into a large error in required capital, precisely in the region where regulatory scrutiny is highest.

Limitations of Study

This study has several limitations that should guide interpretation of the results. Most importantly, the empirical illustration is built on a single synthetic claims sample constructed to reproduce plausible Kenyan severity patterns, rather than on administrative data from a real insurer; the parameter estimates, risk measures, and simulation outputs reported here are therefore illustrative of the methodology and should not be read as characterising the loss experience of any specific insurer, product line, or the Kenyan market as a whole. Validating the framework against real, multi-insurer claims data — ideally spanning several underwriting cycles and product lines — is a necessary next step before the specific numerical results, as opposed to the modelling approach itself, can inform pricing or capital decisions. Second, the analysis is purely parametric; mixture or semi-parametric alternatives may be better suited to claim processes that are multimodal or subject to regime change, and were not explored here. Third, real Kenyan claims data is likely to be affected by reporting delays, fraud, inflationary distortion, and left-truncation of small claims, none of which are present in the synthetic sample; these features can materially affect estimation of the tail index, particularly when α is close to the finite-variance boundary, and their impact should be assessed directly once real data is available. Finally, the paper addresses severity modelling in isolation and does not consider frequency-severity dependence or portfolio aggregation effects, both of which can materially affect systemic risk in catastrophic scenarios. Taken together, these limitations mean the paper's findings should be read as a demonstration of a decision-relevant modelling framework rather than as an empirical statement about Kenyan insurance risk.

CONCLUSION

This paper set out to do two things: build an actuarially interpretable account of the Pareto Type II distribution's tail behaviour, and show how that behaviour converts into solvency capital, retention, and reinsurance decisions under a Risk-Based Supervision regime. On the illustrative synthetic portfolio used here, the Pareto II model provided the best fit among the competing severity distributions on every criterion reported

in Table 3, and the fitted tail index ($\hat{\alpha} \approx 1.72$) placed the portfolio inside the infinite-variance region, where a small number of large claims dominate aggregate loss behaviour. The resulting gap between VaR and TVaR at high confidence levels (Table 4), together with the sharp rise in both quantities as α approaches and falls below the $\alpha = 2$ boundary (Section 9.3), illustrates why tail-sensitive measures rather than variance-based ones are the appropriate basis for capital setting in this kind of portfolio. Where a lighter-tailed model such as the lognormal or Weibull was used instead, the comparative results in Section 8.1 suggest capital could be understated at the confidence levels used for solvency supervision — a finding consistent with, though not by itself proof of, the concern raised in the literature that light-tailed models misprice catastrophic risk in emerging markets. Because these results come from a single synthetic portfolio rather than administrative data, they should be read as a demonstration of the methodology's relevance rather than as a claim about the Kenyan market as a whole; Section 10 sets out what further validation would be needed before the specific numbers reported here could inform actual pricing or capital decisions.

For reinsurance strategy, the results suggest that excess-of-loss and stop-loss treaties are better calibrated against tail parameters directly than against aggregate loss statistics alone, particularly when α is close to the finite-variance boundary and retention levels are correspondingly sensitive to estimation error. The simulation framework developed here is intended as a template for capital stress-testing: once fitted to a real portfolio, the same machinery can be used to examine how solvency position responds to tail scenarios and parameter uncertainty. More broadly, the paper argues that heavy-tailed severity modelling in emerging-market insurance should be treated as part of pricing governance and capital management, and not as a standalone statistical exercise — a framing this paper has tried to operationalise rather than simply assert.

FUTURE RESEARCH

Future research may extend this study — ideally using real, multi-insurer claims data — by incorporating covariate-based regression-Pareto and hierarchical Bayesian versions of the Pareto Type II model to allow tail parameters to vary across underwriting classes, product lines, and portfolio segments, thereby improving understanding of how risk heterogeneity influences capital allocation under Risk-Based Supervision. Further work may also explore multivariate and dependence-based Pareto frameworks to capture joint tail exposure across correlated business lines, as well as simulation-driven stress-testing platforms that incorporate parameter uncertainty, inflation shocks, and reinsurance restructuring scenarios. Such extensions would strengthen the practical integration of heavy-tailed severity modelling with solvency capital assessment, reinsurance strategy, and systemic risk evaluation within emerging-market insurance environments.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work, the authors used ChatGPT to assist with deriving the Lomax cumulative distribution function and with generating the illustrative synthetic claims sample used in Section 8. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

REFERENCES

1. Klugman, S.A., Panjer, H.H., Willmot, G.E., Loss Models: From Data to Decisions, 2012, Wiley.
2. Embrechts, P., Klüppelberg, C., Mikosch, T., Modelling Extremal Events for Insurance and Finance, 1997, Springer.
3. Lomax, K.S., Business Failures, 1954.
4. Ongere, C.O., Angwenyi, D., Robert, O., Second order extended ensemble Kalman filter with stochastically perturbed innovation for initializing artificial neural network weights, 2025, African Journal of Empirical Research.
5. Hamidou, M., Mung'atu, J. K., Orwa, G. O., Return Levels Approach and Periods of Currency Crises, 2018, Journal of Mathematics Research.

6. Pérez-Fructuoso, M. J., García, A., Berlanga, A., Molina, J. M., Adjusting the Generalized Pareto Distribution with Evolution Strategies – An application to a Spanish Motor Liability Insurance Database, 2007, Springer Berlin Heidelberg.
7. Kipchumba, R. K., Bayesian Inference of the Weibull-Pareto Distribution, 2022, University of Nairobi.
8. Omondi, B. A., Weibull and log-normal distributions in reliability analysis applications, 2017, Anadolu University (Turkey).
9. Karobia, R. J., Modelling extreme claims using generalised pareto distributions family in an insurance company, 2015, (Doctoral dissertation, University of Nairobi).
10. Uwingabire, J., Modelling extreme health insurance claims using generalized Pareto distribution, 2018.
11. Leonard, V., Model Uncertainty, Tail Risk and Structured Reinsurance, 2022.
12. Zhao, Yu., Lee, J., Yu, M., Catastrophe risk, reinsurance and securitized risk-transfer solutions: A review, 2021, China Finance Review International.