Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/140103

Analyzing the South African Industrial Index Data using Numerous Standard Statistical Distributions

Sandile C. Shongwe, Zander Greyling, Thandiwe Zwane, and Sihle Gcanga*

Department of Mathematical Statistics and Actuarial Science, Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein, 9301, South Africa

Received: 1 Oct. 2024, Revised: 22 Nov. 2024, Accepted: 27 Nov. 2024.
Published online: 1 Jan. 2025.

Abstract: Model fitting and risk estimation are an important everyday aspect of a successful financial institution. Consequently, in this paper, we discuss risk quantification of the South African Industrial Index (also known as J520) using 22 standard light- and heavy-tailed statistical distributions. Given the importance of the J520 index (since it has the highest market capitalization in the Johannesburg Stock Exchange), investors may have a very keen interest in fully understanding the loss and gain returns characteristics and underlying statistical distribution's properties, including tail properties. Thus, an in-depth goodness-of-fit evaluation is conducted by assessing six different tests (i.e. Kolmogorov-Smirnov, Anderson-Darling, Cramer von Mises, negative log-likelihood, Akaike information criterion, Bayesian information criterion) as well as two risk measures (i.e., value-at-risk and tail value-at-risk) are computed and interpreted within the context of J520 index. It is observed that the best distribution to fit to loss returns are the inverse Burr or transformed beta while for the gain returns it is either transformed gamma, inverse Burr or generalized beta distributions. The latter distributions strike a better balance with respect to excellent goodness-of-fit and risk measures that are very close to the corresponding ones for the empirical distribution. Given our findings, it may not be advisable for investors to hold very long positions in the J520 index since loss returns have much higher leptokurtic and heavy-tailed as compared to the lighter-tailed gain returns. Therefore, the growth shares observed over the long term indicate that a more prudent strategy would be to consider shorting the index. By doing so, investors could better align their strategies with the highly likely potential for substantial drawdowns inherent in this market.

Keywords: Goodness-of-fit, Heavy-tailed, Industrial index, Tail value-at-risk (TVaR), Value-at-risk (VaR).

1 Introduction

Over the years, researchers and professionals have devoted significant efforts to exploring the complexities of financial markets. These intricate and dynamic systems have been subjected to thorough analysis, considering a range of factors including economic, political, and psychological influences. While much research has focused on forecasting and simulations, we recognize the need to delve deeper than predictive models to fully understand financial markets, see Chernobai et al. [1] and Sweeting [2]. A pivotal aspect of this endeavor involves scrutinizing the underlying probability distributions that govern the behavior of financial assets' returns. Traditionally, the normal distribution has served as the main model for characterizing returns on financial instruments, see for instance [3] and [4]. However, mounting empirical evidence challenges this age-old view. Security returns often exhibit deviations from the normal distribution, manifesting heavy tails, leptokurtic and skewness. This observation has profound implications for portfolio optimization, risk management, and the validation of market efficiency hypotheses.

1.1 Literature Review

Table 1 provides a summative overview for a better understanding of different publications that will be used as part of the literature review for this research work. The *twenty-two* distributions that are the focus of this study can be studied in detail in the Appendix of the book by Klugman et al. [5]. For a brief outline of these distribution's properties and an indication of which articles summarised in Table 1 discuss specific distributions that are provided on Table A1 in the Appendix. The objective of [5] is to provide a foundation in modelling data using well-known statistical distributions and thereafter compare them with the corresponding empirical distribution. In response to the observed discrepancies (between the fitted model and the actual data or empirical distribution), researchers have diligently investigated alternative distributional models that better align with empirical data.

Alternative complex model fitting, such as composite models which are defined as a combination of two or more standard distributions, at a calculated threshold value(s), see Cooray and Ananda [22]. Composite models are very common in actuarial applications due to their flexibility and generally better modelling of small, moderate and extreme claims simultaneously. Subsequent contributions, including those by [8, 9, 11, 14 to 21] introduced variations such as mixture

*Corresponding author e-mail: shongwesc@ufs.ac.za

models (convex combinations of two or more distributions on the same domain) and adjustments to standard distributions. In the next paragraphs, the single-distribution fitting articles are discussed in more detail; however, for more details on the composite and mixture distributions (with more emphasis on the Danish fire claims data) we refer the reader to [23].

Table 1: Summative summary of different publications that discuss single, composite and mixture distributions fitted on real and simulated financial data

Article Number	Criteria / Publication	Dataset(s)	Distributions Fitted	Model Type	Goodness-of-fit tests							Risk Metrics	
					KS	AD	CvM	NLL	BIC	AIC	Other(s)	VaR	TVaR / ES / CVaR
1*	Chikobvu and Ndlovu [6]	Bitcoin/USD ZAR/USD	3	Single		✓			✓	✓		✓	✓
2*	Marambakuyana and Shongwe [7]	SA taxi claims Danish fire claims	19	Single	✓	✓	✓	✓	✓	✓		✓	✓
3	Marambakuyana and Shongwe [8]	SA taxi claims Danish fire claims	512	Composite & Mixture				✓	✓	✓		✓	✓
4	van Dorp and Shittu [9]	Danish fire claims French interruption loss	6	Composite	✓		✓					✓	✓
5*	Shongwe et al. [10]	SA Financial Index (J580)	4	Single	✓	✓	✓	✓	✓	✓		✓	✓
6	Tomarchio et al. [11]	Swedish motorcycle losses French motor line losses	23	Mixture					✓	✓		✓	✓
7*	Chikobvu and Jakata [12]	J580	4	Single					✓	✓		✓	✓
8*	Maphalla et al. [13]	SA taxi claims	6	Single	✓	✓						✓	✓
9	Alkhairy et al. [14]	Unemployment insurance	5	Mixture	✓	✓	✓		✓	✓	✓	✓	✓
10	Zhao et al. [15]	Vehicle insurance loss	8	Mixture				✓	✓	✓	✓	✓	✓
11	Ahmad et al. [16]	Earthquake insurance loss	7	Mixture		✓			✓	✓	✓	✓	✓
12	Tung et al. [17]	Vehicle insurance loss Earthquake insurance loss	6	Mixture	✓	✓	✓		✓	✓			
13	Grün and Miljkovic [18]	Danish fire claims	256	Composite	✓	✓			✓	✓	✓	✓	✓
14	Miljkovic and Grün [19]	Danish fire claims	33	Mixture				✓	✓	✓		✓	✓
15	Abu Bakar et al. [20]	Danish fire claims	8	Composite				✓	✓	✓		✓	✓
16	Nadarajah and Abu Bakar [21]	Danish fire claims	17	Composite						✓			

Note: The '*' on the first column denotes the articles we are using as primary references.
Acronyms: KS – Kolmogorov Smirnov, AD – Anderson Darling, CvM – Cramer von Mises, NLL – Negative log likelihood, BIC – Bayesian Information Criterion, AIC – Akaike Information Criterion, VaR – Value at Risk, TVaR – Tail VaR, ES – Expected Shortfall, CVaR – Conditional VaR, SA – South Africa, ZAR – South African Rand, USD – United States Dollar.

Chikobvu and Ndlovu [6] fitted 3 light- and heavy-tailed distributions to 2 exchange rates' returns (i.e. the Bitcoin (BTC) and U.S. Dollar (USD) denoted as BTC/USD as well as South African Rand (ZAR) and USD denoted as ZAR/USD). Firstly, Chikobvu and Ndlovu (2024) split the data into loss and gains returns for each BTC/USD and ZAR/USD. Secondly, these loss and gain returns were separately observed to be leptokurtic and positively skewed. Thirdly, the Hill's plots as well as the generalised QQ and PP plots were analysed so that the tails of the returns can be classified as being light- or heavy-tailed. Fourthly, the ZAR/USD gains and losses returns were both classified to follow a Weibull distribution. However, the BTC/USD gains follow the Burr distribution, while the losses follow the exponential distribution (i.e. BTC/ZAR gains are therefore more volatile giving the potential for large gains when compared to their losses). Finally, the BTC/USD was classified as riskier (or volatile) than the ZAR/USD and that the upside risk (likelihood of potential gains) is greater than the downside risk (the prospects of potential losses) for both BTC/USD (more pronounced here) and ZAR/USD (less pronounced).

Maphalla et al. [13] fitted six common parametric loss distributions (i.e. Pareto, gamma, Burr, log-normal, Weibull and exponential) to the South African taxi claims dataset. They quantified the goodness-of-fit using the KS and AD tests, with the parameter estimates calculated using MLE. The lognormal distribution seemed to have a better goodness of fit than the others. Only the VaR was computed as a risk measure, this was compared with the corresponding VaR of the generalized extreme value distribution. Next, [7] extended on the work by [13] by fitting 13 additional distributions to the taxi claims data and used six different goodness-of-fit tests to find the best fit for the data. In addition, [7] analysed a dataset called

Danish fire loss data using the six goodness-of-fit tests and thereafter computed the risk measures of the 19 theoretical distributions and compared their values to the empirical risk estimates to assess the fit of the model. Overall, in this study, the transformed beta family offered the best fit for the taxi claims data; while for the Danish fire loss data, out of all those in the class of transformed beta family, the Burr distribution performs the best.

Chikobvu and Jakata [12] fitted 4 distributions (exponential, Weibull, gamma and Burr) into the J580 (South African Financial Index) and studied the downside / upside of investing in it. They showed that using the AIC and BIC, the loss and gain returns of the J580 index can be modelled using Burr and exponential distributions, respectively. They further calculated the risk measures for the loss returns (using the Burr distribution only) and for the gain returns (using the exponential distribution only). Given that the J580 dataset's expected losses have a higher likelihood of extreme events occurring as they are recommended to be modelled by the much heavier-tailed Burr distribution; however, the J580 dataset's expected gains have a light-tailed pattern as it is recommended to be modelled by the exponential distribution.

Next, Shongwe et al. [10] re-assessed the results in Chikobvu and Jakata [12] where several different deductions were observed, namely:

- While [12] computed the descriptive statistics for loss and gain returns combined, [10] argued that this was not a correct approach since the actual analysis is done separately for loss and gain returns, thus presenting a combined descriptive is incorrect. In addition, the combined returns presented in [12] were shown to be also incorrect and [10] computed the correct ones. Separately, loss and gain returns data were shown to be leptokurtic with losses having a heavier tail than the lighter-tailed gain returns.
- Results for the goodness-of-fit and parameter estimates were identical in both papers. In addition to AIC and BIC that were computed by [12], then [10] computed four additional goodness-of-fit tests (i.e. KS, AD, CvM and NLL). Given, the additional goodness-of-fit information, [10] deduced that the Burr distribution seems to be ideal for loss returns; however, the Weibull (instead of exponential) seemed ideal for gain returns.
- While [12] computed the VaR and ES for loss returns under the Burr distribution only and gain returns under the exponential distribution only, [10] computed the VaR and ES for loss and gains for all four distributions in each case. In [10], the Burr had the VaR and ES much closer to those of the empirical distribution under loss returns whereas the Weibull had the VaR and ES much closer to those of the empirical distribution under gains returns.

1.2 Research Problem

In recent years, researchers have diligently explored various modelling approaches to understand the intricate characteristics of financial data. Some studies (e.g. [10, 12]) have focused on fitting standard models to financial data, specifically the South African Financial Index (J580). Similarly, [7, 13] have investigated the fitting standard distributions to South African taxi claims and Danish fire claims datasets. The existing academic research landscape predominantly focuses on insurance data and publicly available financial datasets. While these datasets exhibit similar heavy-tailed features, there remains a critical gap: the lack of comprehensive risk assessment for other indices beyond insurance and finance. Considering the significant impact and importance of various indexes, it becomes imperative to extend risk calculations to these domains. By addressing this gap, researchers can enhance risk management strategies and provide valuable insights for decision-makers across diverse sectors.

In our study, we shift our focus to the South African Industrial Index (J520) data comprises of all Johannesburg Stock Exchange (JSE)-listed companies that do not fall into the financial (banking, investment, and security) or resource (oil, gas, and mining) indices. These companies are categorized into the following sub-category indices: Construction and Materials (J235), Aerospace and Defence (J271), General Industrials (J272), Electronic and Electrical Equipment (J273), Industrial Engineering (J275), Commercial Vehicles and Trucks Industrial Transportation (J277), and Support Services (J279); see Jakata and Chikobvu [24]. The dataset can also be accessed from SA Shares, Yahoo Finance websites or in [25]. The J520 data consist of 2 columns, i.e. sorted monthly dates (first column) and average monthly index (second column). This dataset has been used by other authors, i.e. [24, 26 to 29]; however, they used it in the context of extreme value theory, we intend to analyze it in the context of standard statistical distributions. By doing so, we contribute to the ongoing discourse on model selection and risk quantification, aiming to enhance decision-making processes for investors and practitioners.

1.3 Objective of the Paper

The main objective is to present the methodology (i.e., goodness-of-fit and risk estimation) and fit 22 different standard distributions that capture the intricate features of the leptokurtic, positively skewed and (to a certain degree) heavy-tailed returns of the J520 dataset so that we can deduce the best distribution for the losses and gains separately.

This information will provide historical information about the index on how it tends to perform so that investors can make rational decisions. Note that most investors are often driven by loss aversion, a psychological phenomenon that prioritizes the avoidance of losses over the pursuit of gains, see [30]. Given the importance of the J520 index (since it has the highest market capitalization in JSE), investors may have a very keen interest in fully understanding the loss and gain returns characteristics and underlying statistical distribution's properties, including tail properties in order to understand potential losses.

1.4 Research Questions

This research work intends to provide answers to the following questions:

- Which of the 22 standard distributions performs the best in terms of the goodness-of-fit and risk measures for the J520 loss and gain returns?
- What are the overall implications of the riskiness of investing in the J520 index for the considered period?

The rest of the paper is structured as follows: In Section 2, the theoretical methodology is discussed. In Section 3, the empirical analysis is conducted, and the corresponding concluding remarks are done in Section 4.

2 Methodology

2.1 Goodness-of-fit

The first three commonly used goodness-of-fit tests to assess the accuracy of a fitted model on a specific dataset or simulated data are the KS, CvM and AD. According to [1]:

- The KS statistic computes the maximum absolute vertical differences between the empirical cumulative distribution function (cdf) and the theoretical cdf.
- The CvM statistic considers the integral of the squared differences between the empirical cdf and the theoretical cdf rather than just considering differences between points.
- The AD statistic places emphasis on the tails of the distribution, i.e., where $F(x)$ or $1-F(x)$ are small.

Stated differently, KS statistic captures the differences between the middle of the data and the proposed model (recommended to assess middle part of the distribution) while the AD statistic prioritizes the tail component (recommended to assess tail part of the distribution). The latter test statistics are computed as follows [1]:

$$KS = \max_x |F_n(x) - F(x)| \tag{1}$$

$$CvM = n\int \left(F_n(x) - F(x)\right)^2 f(x)dx \tag{2}$$

$$AD = n\int \frac{\left(F_n(x) - F(x)\right)^2}{F(x)\left(1-F(x)\right)} f(x)dx \tag{3}$$

where n is the number of observations, $F_n(x)$ is the empirical cdf, $F(x)$ is the theoretical (fitted) cdf and $f(x)$ is the theoretical pdf. The theoretical distributions to be considered in this work are listed in Table A1 in the Appendix.

The second set of goodness-of-fit tests are usually called information criterions, and these are the negative log-likelihood (*NLL*), Akaike information criterion (*AIC*) and Bayesian information criterion (*BIC*). Assume that $\ell(\theta)$ denote the maximized log-likelihood function of a model, then the *NLL*, *AIC* and *BIC* are defined as [5]:

$$NLL = -\ell(\theta). \tag{4}$$

$$AIC = 2NLL + 2p \tag{5}$$

$$BIC = 2NLL + p \times log(n) \tag{6}$$

respectively, where p is the number of parameters or degrees of freedom and n is the number of observations. The metrics, *NLL*, *AIC*, and *BIC*, incorporate the model parameters in their evaluations, with the *AIC* and *BIC* imposing a greater penalty on models characterized by increased complexity (i.e., a greater number of parameters). A desirable model fit is indicated by lower test scores; therefore, given identical datasets, the distribution exhibiting the lowest value for separately Equations (1) to (6) is deemed the most appropriate fit.

2.2 Actuarial risk measures

Decision-making regarding risks is very complex and risk measures are essential for actuaries, investors, and financial institutions to make informed decisions about investments and risk management strategies. Two main risk measures are considered, i.e., VaR and TVaR. Let $F(\cdot)$ and $F^{-1}(\cdot)$ denote the cdf and inverse cdf of a continuous random variable X, respectively. Then, the VaR of X at a 100p% security level denoted by $\mathrm{VaR_p}(X)$, is the 100p% quantile of F such that

$$P\left(X < \mathrm{VaR_p}(X)\right) = p, \qquad F^{-1}(p) = \mathrm{VaR_p}(X) \tag{7}$$

which can be thought of as the lower bound for the capital required to avoid insolvency. Next, TVaR of X at a 100p% security level denoted by $\mathrm{TVaR_p}(X)$,

$$\mathrm{TVaR_p} = \frac{1}{1-p}\int_p^1 \mathrm{VaR_u}(X)du = \mathbb{E}\left[X|X > \mathrm{VaR_p}(X)\right] \tag{8}$$

which can be thought of as the expected value of total loss, given that it exceeds VaR. TVaR is also known as Expected Shortfall (ES), Average Value at Risk (AVaR), Conditional-Tail-Expectation (CTE), Tail Conditional Expectation (TCE), and Conditional Value at Risk (CVaR). The use of the term ES is popular in Europe, while CTE and TCE are more popular in North America (see Klugman et al. [5]).

In this research work, the fit of the theoretical model is also assessed by comparing the empirical risk estimates to the theoretical risk estimates. The cdf of the empirical distribution is computed by

$$F_n(x) = \frac{1}{n}\#\{i: x_i \leq x\}, \tag{9}$$

where # denotes the number of observations $\leq x$, and n is the total number of observations in the sample. This is done by computing the percentage deviation as follows:

$$\%\text{deviation } VaR = \frac{\left(\mathrm{VaR_{Theoretical}} - \mathrm{VaR_{Empirical}}\right)}{\mathrm{VaR_{Empirical}}} \tag{10}$$

$$\%\text{deviation } TVaR = \frac{\left(\mathrm{TVaR_{Theoretical}} - \mathrm{TVaR_{Empirical}}\right)}{\mathrm{TVaR_{Empirical}}} \tag{11}$$

where the theoretical implies any of the 22 standard distributions that are in Table A1. It is important to note that:

- *Underestimating* the risk measures may result in *under-reserving* - which may lead to insolvency, i.e., not enough capital to cover unexpected losses.
- *Overestimating* the risk measures may result in *over-reserving* - which may negatively affect the profitability due to fewer funds available for investment purposes.

3 Analysis

3.1 Descriptive statistics

Let Y_t be the monthly J520 index for the period June 1995 to January 2018; hence, the corresponding monthly returns (r_t) are computed by

$$r_t = \ln(Y_t/Y_{t-1}) = \begin{cases}\text{Gain if } r_t \geq 0\\ \text{Loss if } r_t < 0\end{cases} \tag{12}$$

since some of the returns are positive (i.e. denote as gains) and others are negative (i.e. denote as losses). The gains and losses are separately analysed in this study and for ease in analysis, each of the losses (negative returns) are redefined as positive by introducing absolute values thereof: $l_t = |r_t|$ and thus are positive.

In Table 2, the total of 271 returns consists of 160 (approximately 59%) gain returns and 111 (approximately 41%) loss returns. According to the boxplots in Figure 1, the 2nd quartile is closer to the 1st quartile, this suggests that the data is skewed to the right for both losses and gains, however, the difference in the values for the 1st quartile and 2nd quartile is much smaller for the losses than gains, further suggesting that the losses are more skewed to the right than the gains. The standard deviation (0.04605) and variance (0.00212) for losses show more variability around the mean whereas the standard deviation (0.03338) and variance (0.00111) for gains show less variability around the mean. The coefficient of variation is low for both losses and gains; thus, we do not have a greater level of dispersion around the mean. The distribution of both losses and gains is positively skewed, indicating a right-skewed distribution. However, the losses (3.05364) exhibit a higher degree of skewness compared to the gains (0.77519). The mean-median ratio for both losses and gains being greater than 1 and the median-mean being less than 1, indicating that the mean is larger than the median, providing more evidence of positive skewness.

Additionally, the histograms in Figure 1 visually demonstrate that the losses have heavier tails, extending up to 0.35, compared to the gains, which end at 0.15. Looking closely at the kurtosis, both values are positive, but it can be observed that the losses have a much higher kurtosis compared to gains which means that it has a higher peak than the gains, but both are leptokurtic, this is also visible in Figure 1, where we have the frequency of the histogram for losses being 80 and that of gains being 40. The higher the kurtosis the more extreme values tend to appear, the boxplot of losses in Figure 1 indicates this when compared to that of gains.

Table 2: Descriptives of the losses and gains returns for J520 data

Descriptive	(Absolute) Losses	Gains
No. of observations	111	160
Minimum	0.00006	0.00025
1st quartile	0.01051	0.01818
Median/2nd quartile	0.02979	0.03819
Mean/average	0.04123	0.04446
3rd quartile	0.05426	0.06609
Maximum	0.32847	0.14027
Standard deviation	0.04605	0.03338
Variance	0.00212	0.00111
Coefficient of variation	1.11707	0.75067
Skewness	3.05364	0.77519
Kurtosis	16.78625	3.00489
Mean-median ratio	1.38396	1.16426
Median-mean ratio	0.72257	0.85892

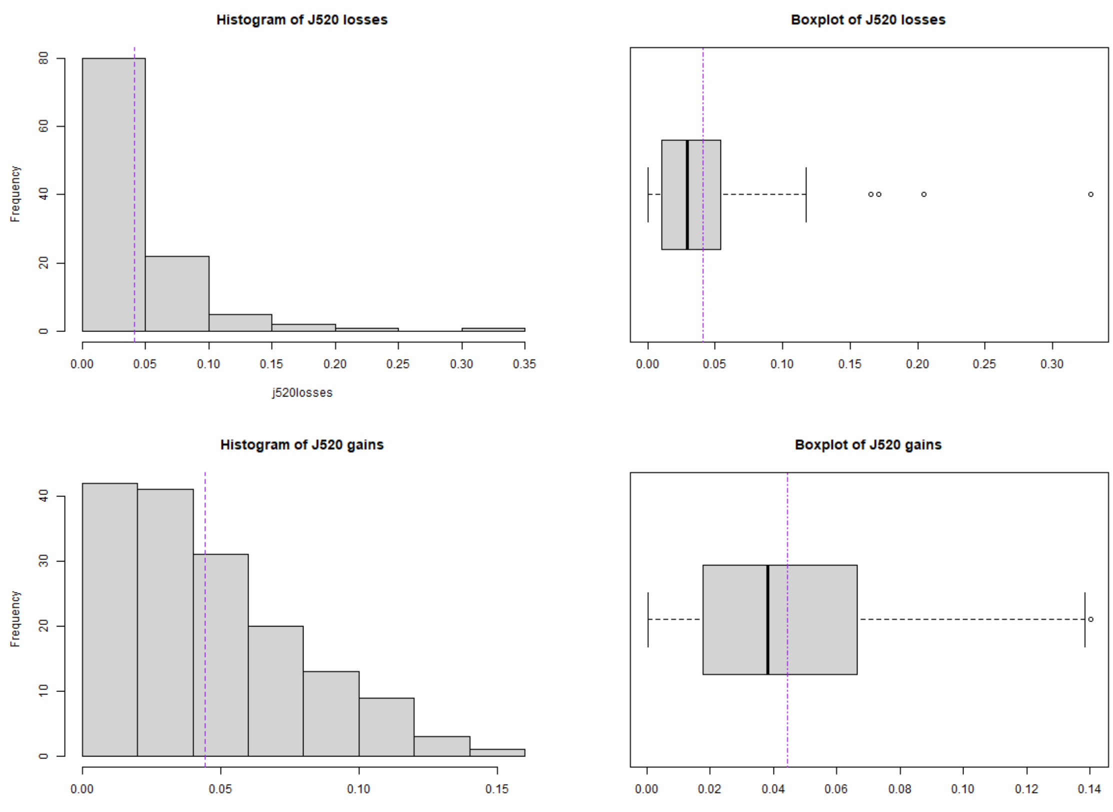
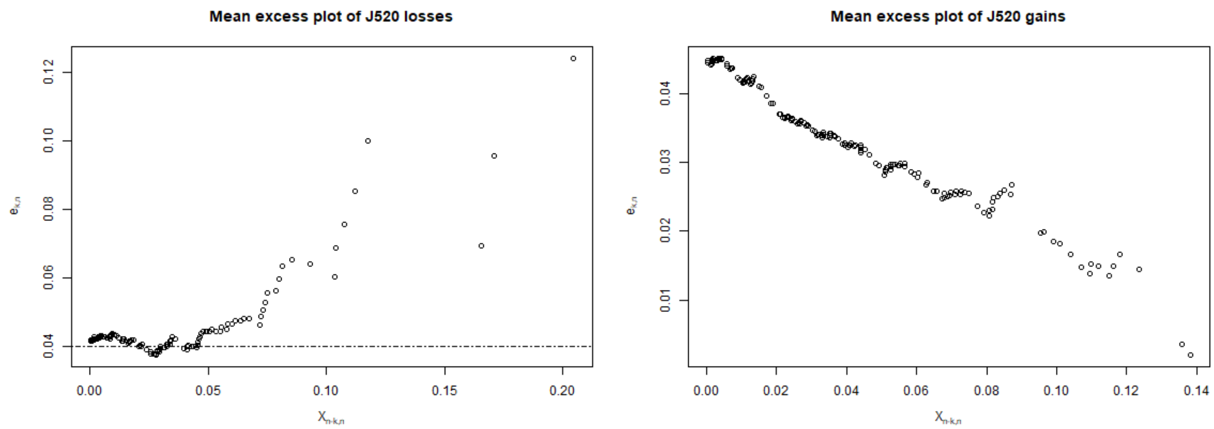


Fig. 1: Histogram and boxplot for the loss and gain returns of the J520 data

Beirlant et al. [31] and Nerantzaki and Papalexiou [32] stated that researchers often visually inspect the mean excess plot (also termed as mean residual life plot in extreme value analysis) to determine the heaviness of the tail of the dataset. The mean excess function, denoted by $e(k) = \mathbb{E}[X - k|X > k]$, plotted against various threshold levels yields the mean excess plot [1]. Note that $e(k)$ is the mean of all differences between the data values and the threshold value, given that data values exceed the threshold, where k denotes a threshold variable. An ultimately increasing (decreasing) mean excess plot suggests that the underlying distribution is heavy- (light-) tailed, respectively. The mean excess plot for losses in Figure 2(a) initially appears linear on the left tail, indicating a lighter-tailed distribution such as the exponential. However, the right tail section exhibits a heavier-tailed distribution, possibly corresponding to the lognormal, Weibull, or Pareto distributions. As for the gains in Figure 2(b), the mean excess plot depicts a decreasing trend suggesting that a distribution may be similar to, say, the gamma (with $\gamma > 1$), Weibull, or uniform distribution.



(a) Loss returns (b) Gain returns

Fig. 2: Mean excess plots for the losses and gains returns of the J520 data

3.2 Goodness-of-fit for losses and gains

3.2.1 Losses

Table 3 present the goodness-of-fit metrics for the KS, CvM, and AD statistics, which were computed for 22 standard probability distributions applied to the J520 loss returns. While all test values warrant consideration, the optimal distribution would ideally demonstrate a strong fit across the majority or all assessments. Notably, the KS, AD, and CvM tests (Equations (1) to (3)) typically converge upon similar distributions since they calculate differences in values. Additionally, the NLL, AIC, and BIC metrics (Equations (4) to (6)) account for the model parameters, thereby offering an alternative perspective that assists in avoiding overfitting through the principle of parsimony. Subsequently, it is important to evaluate each test value while interpreting their respective strengths and limitations.

Table 3: Goodness-of-fit metrics for the 22 standard distributions fitted on J520 *loss* returns

Distributions	KS	CvM	AD	NLL	AIC	BIC
Exponential	0.0574	0.0589	0.4948	-242.94	-483.89	-481.18
Gamma	0.0768	0.0853	0.4830	-243.53	-483.07	-477.65
Weibull	0.0789	0.0900	0.5094	-243.45	-482.9	-477.48
Pareto	0.0762	0.0835	0.5194	-243.40	-482.78	-477.36
Inverse Burr	0.0528	0.0275	0.1648	-245.64	-485.28	-477.15
Beta	0.0712	0.0784	0.4753	-243.13	-482.26	-476.84
Generalized Pareto	0.0815	0.0959	0.5199	-243.63	-481.25	-473.12
Transformed beta	0.0424	0.0183	0.1176	-245.94	-483.87	-473.04
Transformed gamma	0.0748	0.0809	0.4655	-243.54	-481.09	-472.96
Burr	0.0803	0.0925	0.5213	-243.49	-480.98	-472.85
Generalized beta	0.0748	0.0809	0.4656	-243.54	-479.08	-468.24
Paralogistic	0.0936	0.1875	1.5192	-236.90	-469.81	-464.39
Loglogistic	0.0992	0.2359	1.8423	-234.30	-464.61	-459.19
Inverse paralogistic	0.1056	0.3115	2.2339	-231.99	-459.99	-454.57
Inverse Pareto	0.1317	0.4760	3.099	-229.22	-454.43	-449.01

Lognormal	0.1459	0.5420	3.0207	-228.34	-452.67	-447.26
*ITG	0.1683	0.9257	5.0337	-217.41	-428.82	-420.69
Inverse Weibull	0.2069	1.4935	8.5801	-195.30	-386.60	-381.18
Inverse Gamma	0.2921	3.4723	17.4223	-161.96	-319.92	-314.50
Inverse Gaussian	0.4279	8.0978	39.4478	-155.58	-307.16	-301.74
Uniform	0.6535	19.8999	∞	-123.60	-243.19	-237.77
Inverse Exponential	0.5578	14.3839	87.9795	-80.49	-158.98	-156.27

*ITG: Inverse transformed gamma.
The grey highlighted cells denote the top 5 for that test.
The green highlighted cells denote the worst 5 for that test.

Firstly, from Table 3, the inverse Burr is the most consistent distribution among all the goodness-of-fit tests, having the second lowest (best) values in the KS (0.0528), CvM (0.0275), AD (0.1648) and NLL (-245.64) test statistics. Staying in the top 5 with AIC (-485.28) and BIC (-477.15) values, the inverse Burr performs the best across most metrics. Secondly, the transformed beta performs the best for the KS (0.0424), CvM (0.0183), AD (0.1176), NLL (-245.94) tests and second best for the AIC (-483.87); however, it is only seventh best by the BIC (-473.04) test statistic. Thirdly, the exponential is the best performing distribution in the AIC (-483.89) and BIC (-481.18) values (has more advantage on the information criterion as it has a single variable) but is the third best performing in KS (0.0574) and CvM (0.0589). Note that the exponential is not among the top 5 best performing distributions with respect to the AD (0.4948) and NLL (-242.94) test statistics, perhaps indicating deviation in the tail component. Finally, the transformed gamma and generalized beta both have similar KS, CvM, AD and NLL test values. However, the transformed gamma distribution has a stronger AIC and BIC values providing the better fit between the two.

Among the middle-performing distributions for loss returns, several models demonstrate adequate, though not exceptional, performance across the various goodness-of-fit metrics. These distributions (e.g. gamma and Pareto), while not as robust as the top performers like the inverse Burr and transformed beta, still manage to capture some of the underlying dynamics of the loss data. For instance, the beta and Weibull distributions show moderately low KS values (0.0712 and 0.0789, respectively) and perform reasonably well across NLL and AIC measures. Their performance is solid but not exemplary, as their ability to align with the empirical data falls short when compared to the best-fitting models. The generalized beta and transformed gamma distributions also fit within this middle category, exhibiting satisfactory results. They register decent KS values and similar CvM values, indicating that they can model the data without overfitting. However, these models do not provide the exceptional fit needed for accurate risk assessment or predictive modeling. Overall, these distributions may serve as viable alternatives in contexts where the top-performing models are not suitable, but these are not the most reliable choices for highly accurate modeling in practical situations, perhaps by businesses.

The inverse Weibull, inverse gamma, inverse Gaussian, uniform and inverse exponential distributions perform the worst with respect to most of the considered goodness-of-fit tests.

3.2.2 Gains

Table 4 present the goodness-of-fit metrics for the KS, CvM, and AD statistics, which were computed for 22 standard probability distributions applied to the J520 gain returns, see the footnote of Table 3 for the colour highlight description.

Table 4: Goodness-of-fit metrics for the 22 standard distributions fitted on J520 *gain* returns

Distribution	KS	CvM	AD	NLL	AIC	BIC
Transformed gamma	0.0391	0.0285	0.2018	-347.70	-689.40	-680.17
Inverse Burr	0.0347	0.0289	0.2020	-346.49	-686.99	-677.76
Generalized beta	0.0391	0.0285	0.2018	-347.70	-687.40	-675.10
Transformed beta	0.0390	0.0285	0.2017	-347.70	-687.40	-675.10
Weibull	0.0610	0.1581	1.1543	-341.90	-679.80	-673.65
Beta	0.0737	0.2430	1.4079	-341.07	-678.14	-671.99
Exponential	0.1134	0.5393	2.6649	-338.09	-674.18	-671.10
Gamma	0.0776	0.2777	1.6049	-339.91	-675.81	-669.66
Burr	0.0610	0.1582	1.1548	-341.90	-677.79	-668.57
Pareto	0.1134	0.5391	2.6638	-338.09	-672.18	-666.03
Paralogistic	0.0703	0.2664	2.8234	-327.69	-651.38	-645.22
Loglogistic	0.0957	0.3868	3.6396	-320.76	-637.51	-631.36
Uniform	0.2977	6.6343	∞	-314.56	-625.11	-618.96
Inverse Paralogistic	0.1260	0.6171	4.7448	-314.17	-624.34	-618.19

Lognormal	0.1461	1.0081	5.8696	-312.52	-621.05	-614.90
Inverse Pareto	0.1754	1.1335	7.2016	-304.62	-605.24	-599.10
Inverse Weibull	0.2138	2.3205	12.9403	-267.54	-531.08	-524.93
Generalized Pareto	0.2556	3.2321	16.2132	-260.68	-515.37	-506.14
Inverse Gaussian	0.3491	6.6237	30.8672	-255.51	-505.02	-495.80
Inverse Gamma	0.2785	4.0194	19.9375	-240.67	-477.33	-471.18
Inverse exponential	0.4325	10.8015	54.1290	-202.93	-403.87	-400.80
ITG	1.00000	53.33333	∞	35198.29	70402.58	70411.8

From Table 4, the transformed gamma seems the most suitable distribution with the lowest CvM (0.0285), AD (0.2018), KS (0.0391), NLL (-347.7) test statistics indicating it is approximately tied (with generalized beta and transformed beta) for the best performing distribution. The AIC (-689.4) and BIC (-680.17) are the lowest for the transformed gamma distribution. The inverse Burr remains among the most competitive distributions with the second strongest values in the AIC (-686.99) and BIC (-677.76), due to the fewer parameters compared to the other top 5 distributions. The KS (0.0391), CvM (0.0289), AD (0.202) and NLL (-346.49) test statistics are slightly lower than the top 3 that have very similar values. The generalized beta and transformed beta have extremely similar values in all test statistics; with the CvM (0.0285), NLL (-347.70), AIC (-687.40) and BIC (-675.10) being the same when rounding, however differing slightly at the decimals. The generalized beta has slightly different worse performing KS (0.0391) and AD (0.2018) as compared to the transformed beta's KS (0.0390) and AD (0.2017), ultimately being negligible with such little difference at the fourth decimal. Weibull shows the worst noticeable performance among the top 5 best distributions. The KS (0.061), CvM (0.1581), AD (1.1543) and NLL (-341.9) test statistics are fifth best performing, tied with worse performing Burr distribution. The Weibull outperforming Burr in the AIC (-679.8) and BIC (-673.65) test statistics, with the Burr being relatively not as good.

In the context of gain returns, the middle-performing distributions, such as the paralogistic, Burr, beta and gamma distributions, demonstrate reasonable but not outstanding goodness-of-fit. With relatively low KS values, these distributions perform adequately enough in capturing the general shape of the gain return data. However, their higher CvM and AD values indicate a degree of misalignment with the empirical data. These models may be useful for preliminary analysis but are not the best choices for highly precise risk estimation or predictive modeling; thus, making them less ideal compared to top performers. Finally, the generalized Pareto, inverse Gaussian, inverse gamma, inverse exponential and inverse transformed gamma distributions are among the top 5 worst performing distributions with respect to most of the test statistics.

3.3 Risk measures for Losses and gains

3.3.1 Losses

Table 5 present the VaR and TVaR metrics, which were estimated at confidence intervals of 95% (α = 0.05), 99% (α = 0.01), and 99.5% (α = 0.005) for the J520 loss returns across 22 standard distributions. The distributions are arranged in the same sequence as those displayed in the corresponding goodness-of-fit in Table 3. Here, the empirical distribution is used as a reference 'model', while the deviations (shown in brackets) from the theoretical VaR and TVaR values are computed using Equations (10) and (11) for comparison purposes. In the context of loss returns, certain distributions stand out for their close approximation to empirical values, while others exhibit significant deviations.

Table 5: Risk measures for the 22 standard distributions fitted on the J520 *loss* returns and the % deviation from the empirical distribution's risk measure

Distribution	$\mathbf{VaR_{0.95}}$	$\mathbf{VaR_{0.99}}$	$\mathbf{VaR_{0.995}}$	$\mathbf{TVaR_{0.95}}$	$\mathbf{TVaR_{0.99}}$	$\mathbf{TVaR_{0.995}}$
Empirical	0.1099	0.2011	0.2603	0.1831	0.2665	0.3285
Exponential	0.1235 (12.4%)	0.1899 (-5.6%)	0.2184 (-16.1%)	0.1647 (-10.1%)	0.2311 (-13.3%)	0.2596 (-21.0%)
Gamma	0.1291 (17.5%)	0.2023 (0.6%)	0.2340 (-10.1%)	0.1746 (-4.6%)	0.2481 (-6.9%)	0.2798 (-14.8%)
Weibull	0.1298 (18.1%)	0.2061 (2.5%)	0.2397 (-7.9%)	0.1774 (-3.1%)	0.2549 (-4.4%)	0.2889 (-12.1%)
Pareto	0.1281 (16.6%)	0.2102 (4.5%)	0.2489 (-4.4%)	0.1800 (-1.7%)	0.2690 (0.9%)	0.3109 (-5.4%)
Inverse Burr	0.1164 (5.9%)	0.2058 (2.3%)	0.2602 (-0.04%)	0.1790 (-2.2%)	0.3098 (16.2%)	0.3907 (18.9%)
Beta	0.1287 (17.1%)	0.1951 (-3%)	0.2223 (-14.6%)	0.1696 (-7.4%)	0.2333 (-12.5%)	0.2594 (-21%)
Generalized Pareto	0.1301 (18.4%)	0.2100 (4.4%)	0.2462 (-5.4%)	0.1802 (-1.6%)	0.2638 (-1%)	0.3017 (-8.2%)
Transformed beta	0.1163 (5.8%)	0.2322 (15.5%)	0.3125 (20.1%)	0.2038 (11.3%)	0.4064 (52.5%)	0.5470 (66.5%)
Transformed gamma	0.1285 (16.9%)	0.2000 (-0.5%)	0.2308 (-11.3%)	0.1729 (-5.6%)	0.2444 (-8.3%)	0.2750 (-16.3%)
Burr	0.1299 (18.2%)	0.2102 (4.5%)	0.2467 (-5.4%)	0.1802 (-1.6%)	0.2645 (-0.8%)	0.3027 (-7.9%)
Generalized beta	0.1285 (16.9%)	0.2000 (-0.5%)	0.2307 (-11.4%)	0.1729 (-5.6%)	0.2442 (-8.4%)	0.2749 (-16.3%)
Paralogistic	0.1792 (63.1%)	0.4857 (141.5%)	0.7346 (182.2%)	0.4536 (147.7%)	1.1826 (343.8%)	1.7775 (441.1%)
Loglogistic	0.2356 (114.4%)	0.8236 (309.6%)	1.3985 (437.3%)	1.0058 (449.3%)	3.428 (1186.6%)	5.804 (1666.8%)
Inverse Paralogistic	0.3042 (176.8%)	1.2695 (531.3%)	2.3248 (793.1%)	2.371 (1195.3%)	9.609 (3505.6%)	17.53 (5237.3%)
Inverse Pareto	0.4463 (306.1%)	2.3190 (1053.6%)	4.661 (1690.9%)	∞	∞	∞
Lognormal	0.2333 (112.3%)	0.6297 (213.1%)	0.9058 (248%)	0.5224 (185.3%)	1.1808 (343.1%)	1.6161 (392%)
*ITG	0.3888 (253%)	1.6768 (733.8%)	2.965 (1039.1%)	1.8733 (923.1%)	6.399 (2301.4%)	10.60 (3129.1%)
Inverse Weibull	3.315 (2916%)	82.881 (41114%)	327.43(125692%)	Divergent	Divergent	Divergent

Inverse Gamma	13.11 (11838%)	1989.3 (989094%)	17294 (7×10⁶%)	Divergent	Divergent	Divergent
Inverse Gaussian	0.1749 (59.1%)	0.6960 (246.1%)	1.0501 (303.4%)	0.5237(186%)	1.2835 (381.6%)	1.7229 (424.5%)
Uniform	0.3121 (184%)	0.3252 (61.7%)	0.3268 (25.5%)	0.3203 (74.9%)	0.3268 (22.6%)	0.3277 (-0.2%)
Inverse exponential	0.0499 (-54.6%)	0.2548 (26.7%)	0.5109 (-42%)	∞	∞	∞

*ITG: Inverse transformed gamma
The grey highlighted cells denote the top best for that risk measure.
The green highlighted cells denote the worst ones for that risk measure (due to at least one of the measures at 1-α level being greater or equal to 1).

To illustrate how to explain VaR and TVaR values simultaneously at a specific (1-α)×100% level, then from Table 5, consider the Pareto distribution at a 95% confidence level with VaR = 0.1281 and TVaR = 0.1800: This means that the loss returns are not expected to go beyond 12.81% at a 95% confidence level which is 16.6% higher than the empirical distribution one. However, if it goes beyond 12.81%, it will average 18.00% at a 95% confidence level which is 1.7% lower than the empirical distribution. The Pareto distribution performs well across all quantiles, but it contains slight underestimations in the extreme tail.

Similarly, the inverse Burr tends to have a good performance for the risk measures, with the distribution fitted well for VaR, it usually tends to overestimate tail risks in TVaR. Next, the Weibull distribution seems to be adequate in fitting VaR but less optimal at the extreme tail in the estimation of TVaR. While the Burr distribution fits well in VaR, with a slight underestimation of extreme tail risks in TVaR, the generalized Pareto distribution fits well for both VaR and TVaR, but slightly underestimates the tail risks at the extreme tail positions. These latter distributions, including the paralogistic and exponential distributions show similar trends, with moderate deviations in VaR and TVaR values, indicating good overall performance (capturing middle quantiles) but underestimates extreme tails risks.

In particular, the grey shaded risk measures denote those whose values are not too far from the empirical distribution and the unshaded ones being comparatively acceptable too. However, the green shaded ones denote those that have irrationally high value(s) of VaR and TVaR for at least one of the risks at different significance levels. This is because at least one of the VaR or TVaR is greater or equal to 1; thus, must be ignored as it is both impractical and impossible to keep more than a 100% of cash on reserves. Stated differently, maintaining reserves exceeding 100% of cash is both impractical and impossible; therefore, any VaR or TVaR greater than 1 (or approaching infinity / divergent) must be disregarded.

3.3.2 Gains

Table 6 presents the VaR and TVaR metrics for the J520 loss returns across 22 standard distributions which are arranged in the same sequence as those displayed in the corresponding goodness-of-fit in Table 4. Using the same analogue as in Table 5, it follows that the grey shaded risk measures (i.e., transformed gamma, inverse Burr, generalized beta, transformed beta, as well as in part the Weibull and Burr) denote those whose values are not too far from the empirical distribution and the unshaded ones (i.e., beta, exponential, gamma and Pareto) being comparatively acceptable too. However, the green shaded ones denote those that have irrationally high value(s) of VaR and TVaR for at least one of the risks at different significance levels.

Table 6: Risk measures for the 22 standard distributions fitted on the J520 *gains* returns and the % deviation from the empirical distribution's risk measure

Distribution	$VaR_{0.95}$	$VaR_{0.99}$	$VaR_{0.995}$	$TVaR_{0.95}$	$TVaR_{0.99}$	$TVaR_{0.995}$
Empirical	0.1098	0.1368	0.1387	0.1249	0.1393	0.1403
Transformed gamma	0.1074 (-2.2%)	0.1348 (-1.5%)	0.1445 (4.2%)	0.1242 (-0.6%)	0.1477 (6%)	0.1562 (11.3%)
Inverse Burr	0.1067 (-2.8%)	0.1461 (6.8%)	0.1657 (19.5%)	0.1323 (5.9%)	0.1781 (27.9%)	0.2017 (43.8%)
Generalized beta	0.1074 (-2.2%)	0.1348 (-1.5%)	0.1444 (4.1%)	0.1241 (-0.6%)	0.1477 (6%)	0.1562 (11.3%)
Transformed beta	0.1074 (-2.2%)	0.1348 (-1.5%)	0.1445 (4.2%)	0.1242 (-0.6%)	0.1477 (6%)	0.1563 (11.4%)
Weibull	0.1168 (6.4%)	0.1669 (22%)	0.1875 (35.2%)	0.1478 (18.3%)	0.1960 (40.7%)	0.2160 (54%)
Beta	0.1220 (11.1%)	0.1772 (29.5%)	0.1997 (44%)	0.1560 (24.9%)	0.2088 (49.9%)	0.2304 (64.2%)
Exponential	0.1332 (21.3%)	0.2048 (49.7%)	0.2356 (69.9%)	0.1777 (42.3%)	0.2492 (78.9%)	0.2800 (99.6%)
Gamma	0.1245 (13.4%)	0.1859 (35.9%)	0.2122 (53%)	0.1626 (30.2%)	0.2237 (60.6%)	0.2498 (78%)
Burr	0.1168 (6.4%)	0.1669 (22.0%)	0.1875 (35.2%)	0.1478 (18.3%)	0.1960 (40.7%)	0.2160 (54%)
Pareto	0.1332 (21.3%)	0.2048 (49.7%)	0.2356 (69.9%)	0.1777 (42.3%)	0.2492 (78.9%)	0.2801 (99.6%)
Paralogistic	0.1571 (43.1%)	0.3311 (142%)	0.4491 (223.8%)	0.2870 (129.8%)	0.5830 (318.5%)	0.7855(459.9%)
Loglogistic	0.2170 (97.6%)	0.6207 (353.7%)	0.9682 (598.1%)	0.6119 (389.9%)	1.7168 (1132.4%)	2.6715(1804%)
Uniform	0.1333 (21.4%)	0.1389 (1.5%)	0.1396 (0.6%)	0.1368 (9.5%)	0.1396 (0.2%)	0.1399 (-0.3%)
Inverse Paralogistic	0.3018 (174.9%)	1.0905 (697.1%)	1.8796 (1255.2%)	1.4152 (1033.1%)	4.9943 (3485%)	8.5835 (6018%)
Lognormal	0.2106 (91.8%)	0.4861 (255.3%)	0.6603 (376.1%)	0.4018 (221.7%)	0.8074 (479.6%)	1.0548 (651.8%)
Inverse Pareto	0.5461 (397.4%)	2.8274 (1966.8%)	5.6790 (3994.4%)	∞	∞	∞
Inverse Weibull	1.702 (1450.2%)	23.6829 (17212%)	72.8547 (52426%)	Divergent	Divergent	Divergent
Generalized Pareto	1.0363 (843.8%)	14.6131 (10582%)	45.6117 (32785%)	Divergent	Divergent	Divergent
Inverse Gaussian	0.1893 (72.4%)	0.4828 (252.9%)	0.6451 (365.1%)	0.3775 (202.2%)	0.7363 (428.6%)	0.9194 (555.3%)
Inverse gamma	2.197 (1901.1%)	59.7619 (43585%)	247.777 (178542%)	Divergent	Divergent	Divergent
Inverse exponential	0.1472 (34.1%)	0.7512 (449.1%)	1.5061 (985.9%)	∞	∞	∞
ITG	68.5003 (62286%)	124.988 (91285%)	157.8033 (113673%)	106.7109 (85337%)	183.2893(131479%)	227.5574 (162093%)

3.4 Discussion

To mathematically evaluate the performances, we gave each distribution a rank, ordered from first (for best performance) to twenty second (for worst performance). The average ranking was calculated by averaging each distribution ranking across all tests; thus, it provided a general sense of how well the model performed across different tests. For the loss returns the inverse Burr and transformed beta ranked best for goodness-of-fit metrics, with the inverse Burr ranking comfortably better when risk metrics were considered. The Pareto and gamma distributions were ranked among the top 5 distributions, due to top performance in the risk metrics. The gain returns have three top performing distributions, namely transformed gamma, generalized beta and transformed beta, with very little difference in ranking total. For the less extreme gain returns, transformed gamma, generalized beta and transformed beta provided estimations very close to the empirical data, with the fewer parameters, transformed gamma, being ultimately preferred, due to the principle of parsimony.

It is important that we analyze the descriptive statistics for losses and gains separately. These statistics indicate whether gains or losses have been more prevalent during the index's lifetime, specifically examining the outliers and extreme values identified by the leptokurtic and skewness measures. Our findings indicate that losses tend to exhibit more extreme values and outliers, suggesting that the index is more likely to experience decreasing values on a month-to-month basis.

Both distributions, the inverse Burr and transformed gamma which proved to be the best fits for loss and gain returns, respectively, have remarkable tail properties, thus carrying important information about the risk profile of the J520 index. The heavy-tailed nature of inverse Burr distribution of loss returns signals a high probability of extreme losses, hence a risk profile characterized by substantial downside potential. This means that an investment in the J520 index is more exposed to sudden market downturns, which can be further adversely impacted by economic shocks or liquidity constraints. Additionally, this tendency towards extreme decreases in index prices implies that investors face a higher probability of incurring losses, including substantial losses, rather than making profits.

Heavy-tailed distributions may be considered as rooted in some fundamental market risk drivers, such as volatility in macroeconomic conditions, fluctuations in demand within various industries, and sector-specific risks related to the wide varieties of industries represented in the J520 index (see the sub-category indices in Section 1.1). We present results on the tail properties that remarkably capture these events of extreme loss and underline the necessity for investors, especially long position holders, to exercise increased caution. The tail properties here reflect a market environment in which systematic risks and large-scale drawdowns are more likely to occur; portfolio managers may benefit from strategies that mitigate extreme downside risk.

Heavy-tailed distributions indicate a higher probability of extreme losses and investors can use the best-fitting distribution to model downside risk more accurately. VaR can be calculated based on these distributions and can provide investors with insights into the maximum potential losses at specified confidence levels. This can guide the sizing of short positions to ensure they stay within risk tolerance. The insights on the tail properties of the inverse Burr distribution can be used to construct a risk management strategy for investors through tighter stop-loss thresholds and more frequent portfolio rebalancing to hedge against such extreme downside risks. For instance, using the heavier-tailed inverse Burr distribution when informing hedge ratios allows for a better accounting of potential losses and therefore improvement in risk coverage. The same properties also suggest caution in holding long positions in the J520 index, advocating instead for strategic shorting or hedging. By using known distributions and seeking the closest fitting distribution, traders could try and predict future prices by continuing the modeled distribution. Deviations from the model could be early warning indications when to exit or reverse the current position being held.

Heavy-tailed distributions are sensitive to outliers and possible market anomalies, and while they are useful for capturing extreme losses, these unusual data points can sometimes lead to overestimations of risk. Acknowledging this limitation would provide a balanced view and could prompt future researchers to explore techniques for outlier mitigation, such as robust statistical methods or volatility filtering.

New research might be devoted to composite and mixture models (models that combine two or more distributions). This might allow a much more subtle investigation of J520, possibly yielding enhanced risk management strategies directed at the various return distribution segments of the index. These models can capture both the usual and extreme markets' stochastic periodicity better than single distributions. The losses, as evaluated through the mean excess plot (see Figure 2), exhibit characteristics indicative of multiple distributions, thus the composite and mixture models may be better suited for this data pattern. From Figure 2, it follows that the smaller losses demonstrate less complexity and greater linearity, thereby allowing for easier predictability in their occurrence within the index (investors can anticipate these small losses with relative ease). However, the larger losses display more complex characteristics, making them more challenging for investors to predict. Consequently, more sophisticated models, such as composite or mixture models, should be explored to provide more insights into the index characteristics.

4 Conclusion

Our study adds to the growing field of financial index and extreme value analysis, providing insight into an index with the highest market capitalization on the JSE for South African listed companies. While [6, 7, 10, 12, and 13] fitted 3, 19, 4, 4, and 6 standard distributions to different datasets, respectively, in this study, a model fitting analysis is conducted to a larger class of 22 standard distributions that were separately fitted to loss and gain returns. Some of these distributions (especially the flexible 3 or 4 parameters; not the less flexible 1 or 2 parameters) were observed to be able to capture the tail / extreme occurrences of the index over the long-term, to provide valuable insight investors could use. We considered the large number of extreme losses occurrences and concluded too short stock over the long-term would be the most profitable course of action. The large class of distributions provided more insight into the flexible nature of the empirical and financial data since we noticed that 3- to 4-parameter distributions tend to model the data closely, given the flexible nature. Lesser parameter distributions tend to underestimate the complex nature of the heavy-tailed losses. Overall, for loss returns, the inverse Burr distribution provided the best estimation of expected losses (goodness-of-fit), reserve cash required (risk measures), and managed to account for the extremes of the J520 loss returns. However, for gain returns, the transformed gamma distribution, with 3 parameters, outperforms the competitors (as it has fewer parameters than the other two similarly performing competitors which have 4 parameters), thus preferred due to principle of parsimony.

For future research, mixed and composite models could fit the J520 by fitting different head and tail components as the mean-excess plot for losses seem to suggest two separate distributions may be ideal rather than a single one in the whole interval. Also, given that [6] only fitted 3 standard distributions to the BTC/USD and ZAR/USD returns, this work can be extended to fit the 22 distributions discussed here for a thorough characterization of the returns data. The flexibility of standard distributions could also be adjusted by way of Z-family [16], Arcsine-family [17] or Arctan-family [14] of distributions to name a few. This could provide a better fit to the complex nature of loss/gain returns.

Appendix

In this section, we provide additional information that was omitted in the main portions of the paper. Firstly, the basic properties of the 22 considered distributions and indications of which article discuss them are provided in Table A1 (given on the last page before References). It is observed that most of the articles seem to prefer to fit the Burr and Weibull distributions; however, the uniform and generalized beta were not fitted by any of the studies in the literature review.

Secondly, in Tables A2 and A3, the parameter estimates using the maximum likelihood estimation are presented for the 22 standard distributions for the loss and gain returns. The parameter estimates that are highlighted are not significant since the corresponding standard error(s) are greater than the parameters themselves. Certain values are denoted as NA (not any), these occur when the variance-covariance matrix is divergent or tends to occur when using the Maximum Likelihood Estimation method. To understand whether the parameter is for the location, shape or scale, see Table A1 for that information.

Finally, to illustrate the importance of understanding the key difference between light- to heavy-tailed distributions as well as which of these is more flexible, consider the sensitivity analysis presented in Figures A1 to A4. The exponential distribution in Figure A1 is more rigid and less flexible since it has a single parameter. That is, the larger the scale parameter (θ), the relatively heavier-tailed the distribution becomes.

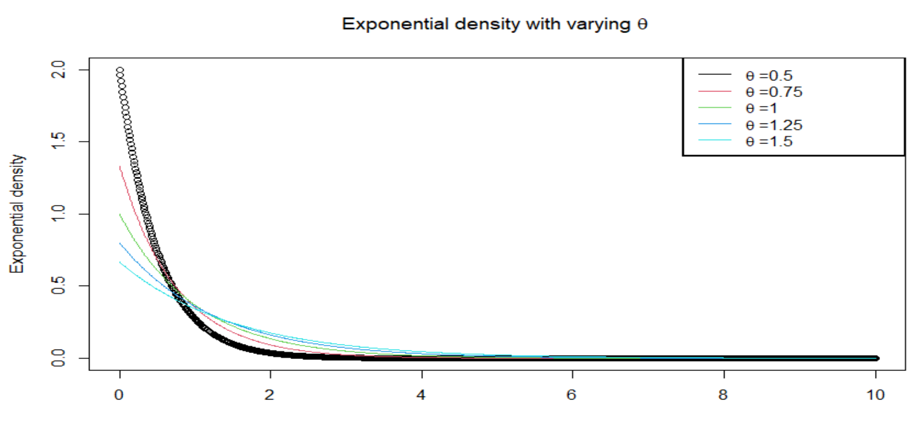


Fig A1: Sensitivity analysis of the one-parameter exponential distribution

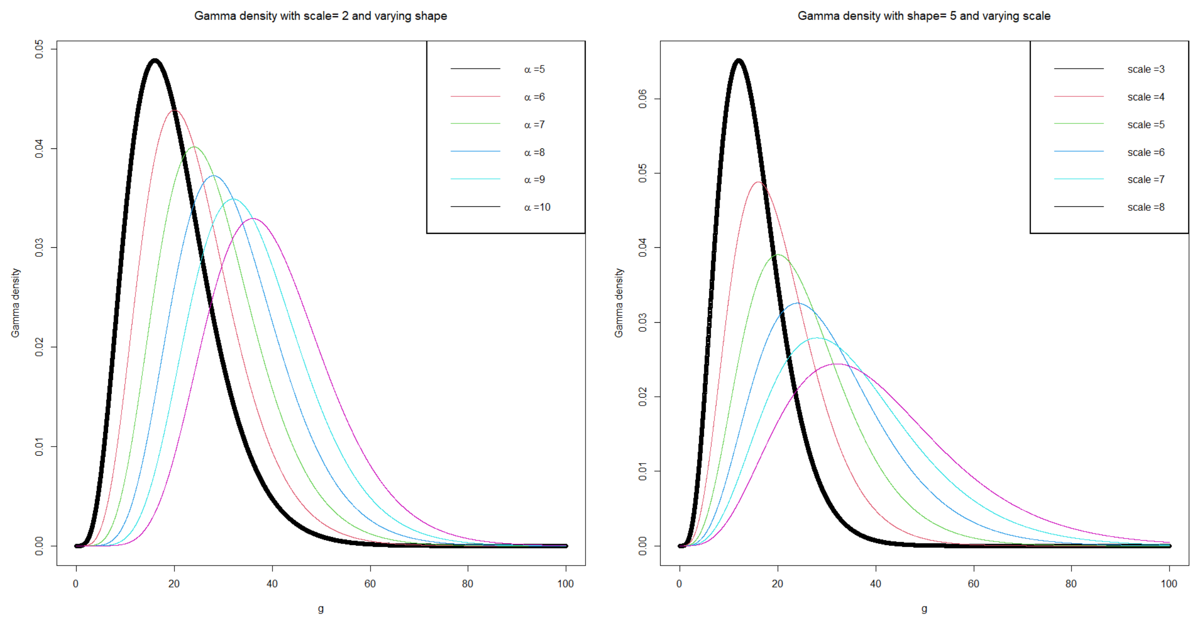


Fig A2: Sensitivity analysis of the two-parameter gamma distribution

Considering the gamma distribution which has an additional parameter to the exponential distribution, this means flexibility is increased, see Figure A2. With a fixed scale parameter (θ), the larger the shape parameter (α), the heavier the tail becomes; however, if the shape parameter is fixed, then the heavier the tail is when the scale parameter is increased.

The three-parameter inverse Burr provides more flexibility as it gains an increase in strength in modelling different kinds of data patterns. For a fixed shape parameter (γ) and scale parameter (θ), the larger the shape parameter (τ), the heavier the tail. Secondly, for a fixed τ and scale parameter, the smaller the γ, the heavier the tail. Thirdly, for a fixed τ and γ, the larger the scale parameter, the heavier the tail.

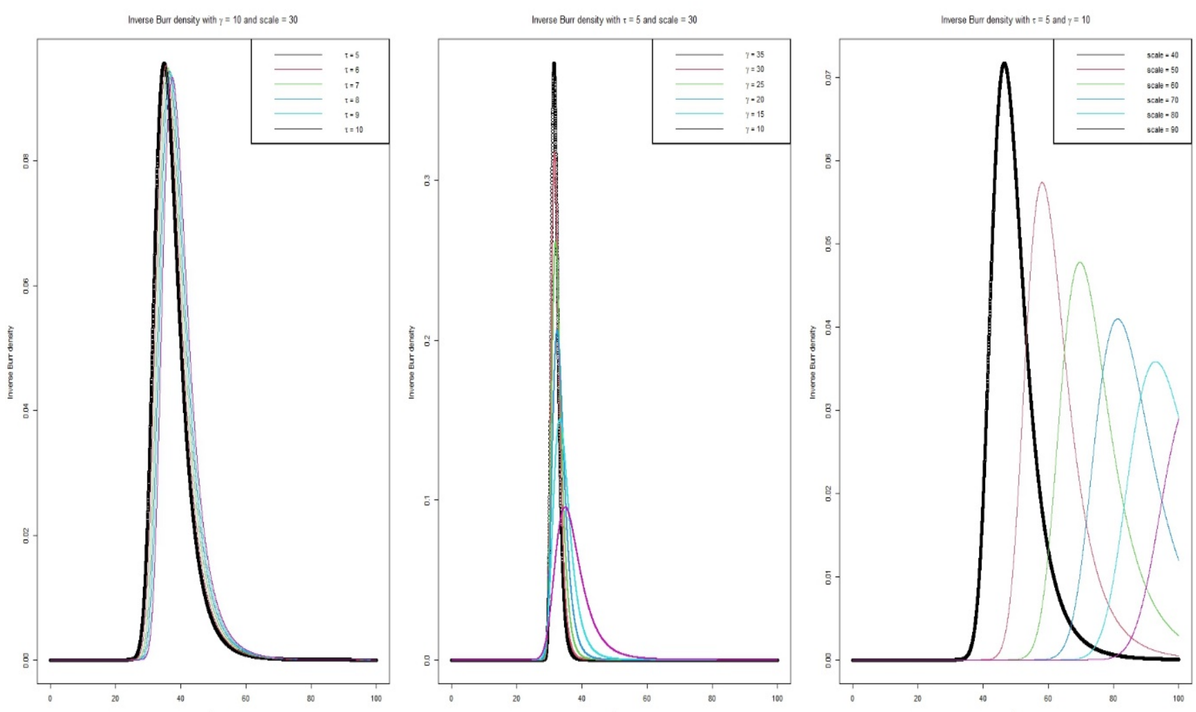


Fig A3: Sensitivity analysis of the three-parameter Inverse Burr distribution

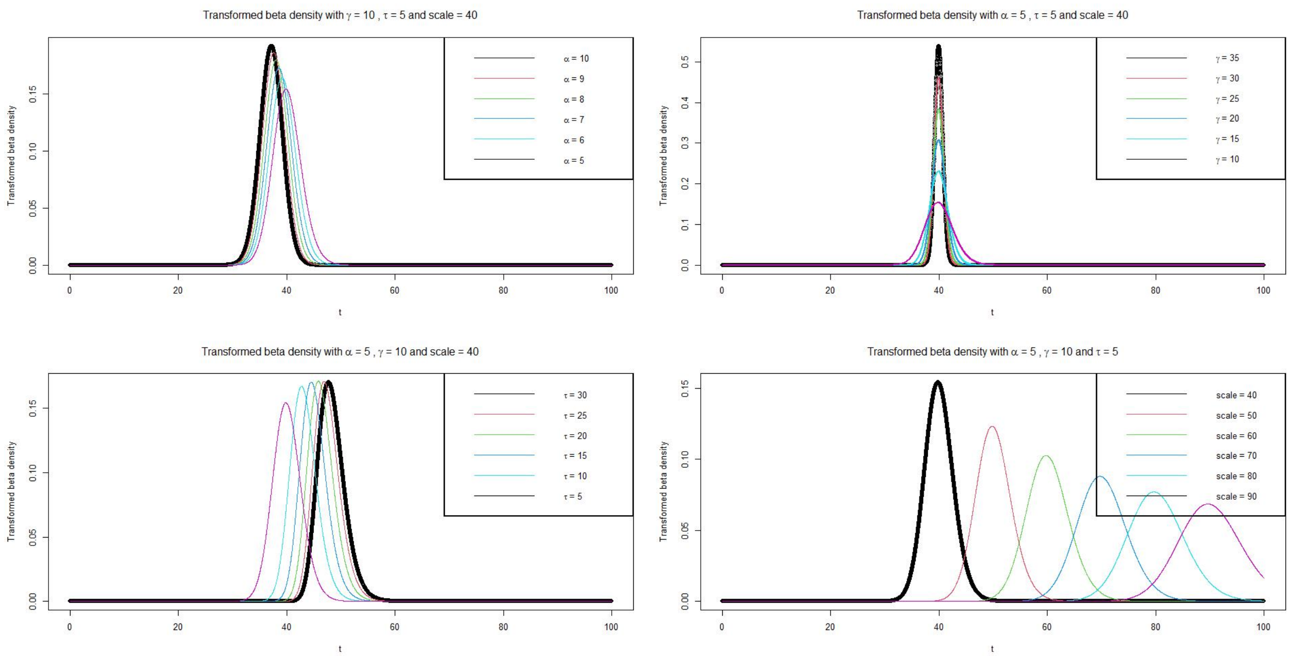


Fig A4: Sensitivity analysis of the four-parameter transformed beta distribution

In Figure A4, the four-parameter transformed beta has the largest flexibility enabling it to model complex structures in data. For a fixed γ, τ and scale (θ), the smaller the α, the heavier the tail. Secondly, for a fixed α, τ and scale, the smaller γ, the heavier the tail is. Thirdly, for a fixed α, γ and scale, the larger the τ, the heavier the tail. Lastly, for a fixed α, γ and τ, the larger the scale, the heavier the tail.

Table A1: Outline of the *twenty-two* distributions' properties and the summary of some of the articles that have fitted these distributions

			Article number from Table 1 using the first column numbering															
Distribution	**PDF**	**Parameters**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Burr	$f(x)=\frac{\alpha\gamma(x/\theta)^{\gamma}}{x\left(1+\left(\frac{x}{\theta}\right)^{\gamma}\right)^{\alpha+1}}, \quad x>0$	Shape = α & γ, scale = θ	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Exponential	$f(x)=\frac{e^{-x/\theta}}{\theta}, \quad x>0$	Scale = θ	✓	✓	✓		✓		✓	✓					✓			✓
Gamma	$f(x)=\frac{(x/\theta)^{\alpha}e^{-(x/\theta)}}{x\Gamma(\alpha)}, \quad x>0$	Shape = α, scale = θ		✓	✓		✓	✓	✓	✓					✓	✓		✓
Generalized Pareto	$f(x)=\frac{\Gamma(\alpha+\tau)\theta^{\alpha}x^{\tau-1}}{\Gamma(\alpha)\Gamma(\tau)(x+\theta)^{\alpha+\tau}}, \quad x>0$	Shape = α & τ, scale = θ		✓	✓										✓		✓	✓
Inverse Burr	$f(x)=\frac{\tau\gamma(x/\theta)^{\tau\gamma}}{x\left(1+\left(\frac{x}{\theta}\right)^{\gamma}\right)^{\tau+1}}, \quad x>0$	Shape = γ & τ, scale = θ		✓	✓										✓	✓	✓	✓
Inverse Exponential	$f(x)=\frac{\theta e^{-\theta/x}}{x^{2}}, \quad x>0$	Scale = θ		✓	✓										✓			✓
Inverse Gamma	$f(x)=\frac{(\theta/x)^{\alpha}e^{-(\theta/x)}}{x\Gamma(\alpha)}, \quad x>0$	Shape = α, scale = θ		✓	✓										✓			✓
Inverse Gaussian	$f(x)=\sqrt{\frac{\theta}{2\pi\sigma x^{3}}}\,e^{-\frac{\theta(x-\mu)^{2}}{2\mu^{2}x}}, \quad x>0$	Location = μ, scale = θ		✓				✓							✓	✓		
Inverse Paralogistic	$f(x)=\frac{\tau^{2}(x/\theta)^{\tau^{2}}}{x[1+(x/\theta)^{\tau}]^{\tau+1}}, \quad x>0$	Shape = τ, scale = θ		✓	✓										✓		✓	✓
Inverse Pareto	$f(x)=\frac{\tau\theta x^{\tau-1}}{(x+\theta)^{\tau+1}}, \quad x>0$	Shape = τ, scale = θ		✓	✓										✓		✓	✓
Inverse Transformed Gamma	$f(x)=\frac{\tau(\theta/x)^{\alpha\tau}e^{-(\theta/x)^{\tau}}}{x\Gamma(\alpha)}, \quad x>0$	Shape = α & τ, scale = θ		✓	✓													✓

Inverse Weibull	$f(x)=\frac{\tau(\theta/x)^{\tau}e^{-(\theta/x)^{\tau}}}{x}, \quad x>0$	Shape = τ, scale = θ		✓	✓										✓			
Loglogistic	$f(x)=\frac{\gamma(x/\theta)^{\gamma}}{x[1+(x/\theta)^{\gamma}]^2}, \quad x>0$	Shape = γ, scale = θ		✓	✓	✓									✓		✓	✓
Lognormal	$f(x)=\frac{1}{\sigma x\sqrt{2\pi}}e^{\frac{-(\ln x-\mu)^2}{2\sigma^2}}, \quad x>0$	Location = μ, scale = σ		✓				✓		✓					✓	✓		✓
Paralogistic	$f(x)=\frac{\alpha^2(x/\theta)^{\alpha}}{x[1+(x/\theta)^{\alpha}]^{\alpha+1}}, \quad x>0$	Shape = α, scale = θ		✓	✓	✓									✓		✓	✓
Pareto	$f(x)=\frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}}, \quad x>0$	Shape = α, scale = θ		✓	✓					✓					✓		✓	
Transformed Beta	$f(x)=\frac{\Gamma(\alpha+\tau)\gamma(x/\theta)^{\tau\gamma}}{\Gamma(\alpha)\Gamma(\tau)x(1+[x/\theta])^{\alpha+\tau}}, \quad x>0$	Shape = α,γ & τ, scale = θ		✓	✓													✓
Transformed Gamma	$f(x)=\frac{\tau(x/\theta)^{\alpha\tau}e^{-(x/\theta)^{\tau}}}{x\Gamma(\alpha)}, \quad x>0$	Shape = α & τ, scale = θ		✓	✓													✓
Weibull	$f(x)=\frac{\tau(x/\theta)^{\tau}e^{-(x/\theta)^{\tau}}}{x}, \quad x>0$	Shape = τ, scale = θ	✓	✓	✓		✓		✓		✓	✓	✓	✓	✓	✓	✓	✓
Generalized Beta	$f(x)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left(\frac{x}{\theta}\right)^{\tau a}\left(1-\left[\frac{x}{\theta}\right]^{\tau}\right)^{b-1}\frac{\tau}{x}, 0<x<\theta$	Shape = a, b & τ, scale = θ																
Beta	$f(x)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left(\frac{x}{\theta}\right)^{a}(1-x/\theta)^{b-1}\frac{1}{x}, 0<x<\theta$	Shape = a & b, scale = θ				✓					✓		✓	✓				
Uniform	$f(x)=\frac{1}{b-a}, a\le x\le b$	Location = a, scale = b																
Other	-	-						✓			✓	✓	✓	✓		✓		✓

Table A2: The parameter estimates using MLE for each of the 22 fitted distributions for the J520 loss returns

		Parameter 1		Parameter 2		Parameter 3		Parameter 4	
Distribution	**No. of parameters**	**Estimate**	**Std error**	**Estimate**	**Std error**	**Estimate**	**Std error**	**Estimate**	**Std error**
Burr	3	$\frac{1}{\theta}=0.7592$	3.3145	α=28.6937	108.20	γ=0.9525	0.1085	-	-
Exponential	1	$\frac{1}{\theta}=$24.2566	2.3023	-	-	-	-	-	-
Gamma	2	$\theta=0.0467$	0.0072	α=0.8827	0.1030	-	-	-	-
Generalized Pareto	3	$\frac{1}{\theta}=0.9765$	2.5175	α=23.6119	56.85	τ=0.9099	0.1252	-	-
Inverse Burr	3	$\frac{1}{\theta}=$14.0646	2.1026	γ=0.2505	0.0662	τ=3.0070	0.5838	-	-
Inverse exponential	1	$\theta=0.0026$	0.0002	-	-	-	-	-	-
Inverse gamma	2	$\frac{1}{\theta}=$1236.48	246.31	α=0.3205	0.0342	-	-	-	-
Inverse Gaussian	2	$\theta=0.0027$	0.0003	μ=0.0412	0.0152	-	-	-	-
Inverse paralogistic	2	$\frac{1}{\theta}=$47.9124	6.9027	τ=1.1535	0.0643	-	-	-	-
Inverse Pareto	2	$\theta=0.0208$	0.0056	τ=1.1286	0.1947	-	-	-	-
ITG	3	$\theta=128.57$	93.683	α=7.5440	0.8401	τ=0.2237	0.0110	-	-
Inverse Weibull	2	$\theta=0.0094$	0.0018	τ=0.5063	0.0301	-	-	-	-
Loglogistic	2	$\frac{1}{\theta}=$39.5878	4.8959	γ=1.3187	0.1061	-	-	-	-
Lognormal	2	$\sigma=1.4573$	0.0978	μ=-3.8526	0.1383	-	-	-	-
Paralogistic	2	$\frac{1}{\theta}=$29.4934	3.9062	α=1.3096	0.0842	-	-	-	-
Pareto	2	$\theta=0.4915$	0.6241	α=12.9314	15.29	-	-	-	-
Transformed beta	4	$\theta=0.0581$	0.0126	α=0.4141	0.4805	γ=5.6328	5.3869	τ=0.1290	0.1294

Transformed gamma	3	$\frac{1}{\theta}=20.2058$	9.18586	α=0.8410	0.3231	τ=1.0327	0.2516	-	-
Weibull	2	$\theta=0.0399$	0.00427	τ=0.9299	0.0686	-	-	-	-
Generalized beta	4	$\theta=57.8142$	NA	α=0.8423	N/A	b=1460.10	N/A	τ=1.0314	NA
Beta	3	-	-	α=0.8493	0.0989	b=19.6541	3.0081	-	-
Uniform	2	$b=0.0001$	NA	α=0.3285	N/A	-	-	-	-

Table A3: The parameter estimates using MLE for each of the 22 fitted distributions for the J520 gain returns

		Parameter 1		**Parameter 2**		**Parameter 3**		**Parameter 4**	
Distribution	**No. of parameters**	**Estimate**	**Std error**	**Estimate**	**Std error**	**Estimate**	**Std error**	**Estimate**	**Std error**
Burr	3	$\theta=96.1784$	52.1537	α=9756.96	3234.38	γ=1.2049	0.0771	-	-
Exponential	1	$\frac{1}{\theta}=22.4897$	1.7780	-	-	-	-	-	-
Gamma	2	$\frac{1}{\theta}=27.3520$	3.3679	α=1.2160	0.1217	-	-	-	-
Generalized Pareto	3	$\theta=$ 245212.8	9021.96	α=5977315	2529.29	τ=1.1284	0.0723	-	-
Inverse Burr	3	$\frac{1}{\theta}=10.8994$	0.9924	γ=0.1443	0.0421	τ=5.6538	1.3393	-	-
Inverse Exponential	1	$\theta=0.0076$	0.0006	-	-	-	-	-	-
Inverse Gamma	2	$\frac{1}{\theta}=272.221$	39.8334	α=0.4874	0.0449	-	-	-	-
Inverse Gaussian	2	$\theta=109.991$	12.2980	μ=0.0445	0.0078	-	-	-	-
Inverse Paralogistic	2	$\frac{1}{\theta}=40.1947$	4.1513	τ=1.2814	0.0597	-	-	-	-
Inverse Pareto	2	$\theta=0.0198$	0.0046	τ=1.4418	0.2271	-	-	-	-
ITG	3	$\theta=1726.03$	NA	α=9.0696	0.2347	τ=0.4826	0.0003	-	-
Inverse Weibull	2	$\frac{1}{\theta}=71.2308$	9.7030	τ=0.6191	0.0316	-	-	-	-
Loglogistic	2	$\frac{1}{\theta}=30.0447$	2.5918	γ=1.5705	0.1059	-	-	-	-
Lognormal	2	$\sigma=1.2277$	0.0686	μ=-3.5774	0.0971	-	-	-	-
Paralogistic	2	$\frac{1}{\theta}=20.3749$	1.7912	α=1.5397	0.0853	-	-	-	-
Pareto	2	$\theta=17080409$	NA	α=384106952	2652.71	-	-	-	-
Transformed beta	4	$\theta=1.0724$	NA	α=1088.42	NA	γ=2.9101	NA	τ=0.2906	NA
Transformed gamma	3	$\frac{1}{\theta}=10.3083$	1.1654	α=0.2908	0.0988	τ=2.9073	0.7692	-	-
Weibull	2	$\theta=0.0470$	0.0032	τ=1.2059	0.0785	-	-	-	-
Generalized beta	4	$\theta=2.2208$	1.1857	α=0.2907	0.0974	b=9000.94	NA	τ=2.9083	0.7578
Beta	3	-	-	α=1.1862	0.1185	b=25.6398	3.1306	-	-
Uniform	2	$b=0.0003$	NA	α=0.1403	NA	-	-	-	-

Data Availability Statement: All the data are publicly available on the following links:

https://figshare.com/articles/dataset/Modelling_Extremes_of_the_Johannesburg_Stock_Exchange_Industrial_Index_J520_using_the_Generalised_Pareto_Distribution_/6935810/1

Conflicts of Interest Statement

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

[1] Chernobai A.S., Rachev S.T. & Fabozzi F.J. (2007). *Operational Risk: A Guide to Basel II Capital Requirements,*

Models, and Analysis. John Wiley & Sons: New Jersey. https://doi.org/10.1002/9781119201922.

[2] Sweeting P. (2017). *Financial Enterprise Risk Management* (2nd ed). Cambridge University Press: Cambridge. ISBN: 9781107184619. www.cambridge.org/9780521111645.

[3] Harckbart G. (2019). Laplace versus the normal distribution for daily stock market returns. *SSRN Electronic Journal*, 1-7. http://dx.doi.org/10.2139/ssrn.3479681.

[4] Eling M. (2014). Fitting asset returns to skewed distributions: Are the skew-normal and skew-student good models? *Insurance: Mathematics and Economics*, 59, 45-56. https://doi.org/10.1016/j.insmatheco.2014.08.004.

[5] Klugman S.A., Panjer H.H. and Willmot G.E. (2019). *Loss Models: From Data to Decisions* (5th ed.). Wiley: New York. ISBN: 978-1-119-52378-9.

[6] Chikobvu D. and Ndlovu T. (2024). Statistical Analysis of the bitcoin and South African Rand exchange rates risks when the tails are somewhat heavy. *Journal of Statistics Application & Probability*, 13(5), 1411-1429. https://doi.org/10.18576/jsap/130501.

[7] Marambakuyana W.A. and Shongwe S.C. (2024). Quantifying risk of insurance claims data using various loss distributions. *Journal of Statistical Applications & Probability*, 13(3), 1031-1044. https://doi.org/ 10.18576/jsap/130315.

[8] Marambakuyana W.A. and Shongwe S.C. (2024). Composite and mixture distributions for heavy-tailed data – An application to insurance claims. *Mathematics*, 12(2), 335. https://doi.org/10.3390/math12020335.

[9] Rene van Dorp J. and Shittu E. (2024). Two-sided distributions with applications in insurance loss modelling. *Statistical Methods & Applications*. https://doi.org/10.1007/s10260-024-00749-x.

[10] Shongwe S.C., Paulse N.A. and Malumane L.E. (2024). Risk quantification using skewed distributions: An application to the South African Financial Index (J580). *ITM Web of Conferences*, 67, 01009. https://doi.org/10.1051/itmconf/20246701009.

[11] Tomarchio S.D., Punzo A., Ferreira J.T. and Bekker A. (2024). Mode mixture of unimodal distributions for insurance loss data. *Annals of Operations Research*. https://doi.org/10.1007/s10479-024-06063-9.

[12] Chikobvu D. and Jakata O. (2023). Fitting statistical parent distributions to quantify financial risk in the South African Financial Index (J580). *Journal of Statistics Applications & Probability*, 12(3), 1231-1245. http://doi.org/10.18576/jsap/120329.

[13] Maphalla R., Mokhoabane M., Ndou M. and Shongwe S.C. (2023). Quantifying risk using loss distributions. In A. A. Jaoude (Ed.), *Applied Probability Theory - New Perspectives, Recent Advances and Trends*, pp.139-161. London, United Kingdom: IntechOpen. http://doi.org/10.5772/intechopen.108856.

[14] Alkhairy I., Nagy M., Muse A.H. and Hussam E. (2021). The Arctan–X family of distributions: properties, simulation, and applications to actuarial sciences. *Complexity*, 2021, Article ID: 4689010. https://doi.org/10.1155/2021/4689010.

[15] Zhao J., Ahmad Z., Mahmoudi E., Hafez E.H. and Mohie El-Din M.M. (2021). A new class of heavy-tailed distributions: Modeling and simulating actuarial measures. *Complexity*, 2021, Article ID: 5580228. https://doi.org/10.1155/2021/5580228.

[16] Ahmad Z., Mahmoudi E. and Kharazmi O. (2020). On modeling the earthquake insurance data via a new member of the T-X family. *Computational Intelligence and Neuroscience*, 2020, Article ID: 7631495. https://doi.org/10.1155/2020/7631495.

[17] Tung Y.L., Ahmad Z. and Mahmoudi E. (2020). The Arcsine-X family of distributions with applications to financial sciences. *Computer Systems Science and Engineering*, 39(3), 351-363. https://doi.org/10.32604/csse.2021.014270.

[18] Grün B. and Miljkovic T. (2019). Extending composite loss models using a general framework of advanced computational tools. *Scandinavian Actuarial Journal*, 2019(8), 642-660. https://doi.org/10.1080/03461238.2019.1596151.

[19] Miljkovic T. and Grün B. (2016). Modeling loss data using mixtures of distributions. *Insurance: Mathematics and Economics*, 70, 387-396. https://doi.org/10.1016/j.insmatheco.2016.06.019.

[20] Abu Bakar S.A., Hamzah N.A., Maghsoudi M. and Nadarajah, S. (2015). Modeling loss data using composite models. *Insurance: Mathematics and Economics*, 61, 146-154. https://doi.org/10.1016/j.insmatheco.2014.08.008.

[21] Nadarajah S. and Abu Bakar S.A. (2014). New composite models for the Danish fire insurance data. *Scandinavian Actuarial Journal*, 2014(2), 180-187. https://doi.org/10.1080/03461238.2012.695748.

[22] Cooray K. and Ananda M.M. (2005). Modeling actuarial data with a composite lognormal-Pareto model. *Scandinavian Actuarial Journal*, 2005(5), 321-334. https://doi.org/10.1080/03461230510009763.

[23] Shongwe S.C. and Marambakuyana W.A. (2024). Danish fire insurance data: A review and additional analysis. *ITM Web of Conferences*, 67, 01011. https://doi.org/10.1051/itmconf/20246701011.

[24] Jakata O. and Chikobvu D. (2022). Extreme value modelling of the monthly South African Industrial Index (J520) returns. *Statistics, Optimization & Information Computing*, 10(2), 638-655. https://doi.org/10.19139/soic-2310-5070-906.

[25] FigShare (2024). The South African Industrial Index (J520) data: 06/1995 to 01/2018. https://figshare.com/articles/dataset/Modelling_Extremes_of_the_Johannesburg_Stock_Exchange_Industrial_Index_J520_using_the_Generalised_Pareto_Distribution_/6935810/1. Date Accessed: 24 May 2024.

[26] Chikobvu D. and Jakata O. (2023). Quantifying diversification effects of a portfolio using the generalised extreme value distribution-Archimedean Gumbel Copula model. *Applied Mathematics & Information Sciences*, 17(6), 967-981. https://doi.org/10.18576/amis/170603.

[27] Jakata O. and Chikobvu D. (2022). Extreme value modelling of the South African Industrial Index (J520) returns using the generalised extreme value distribution. *International Journal of Applied Management Science*, 14(4), 299-315. https://doi.org/10.1504/IJAMS.2022.127009.

[28] Jakata O. and Chikobvu D. (2023). Estimation of diversification effects/benefits using the generalised Pareto distribution – Extreme value Gumbel Copula model. *Journal of Statistics Application & Probability*, 12(S1), 1515-1534. https://doi.org/10.18576/jsap/12S111.

[29] Jakata O. and Chikobvu D. (2023). Modelling portfolio risk and diversification effects of a portfolio using the exponential distribution – Bivariate Archimedean Gumbel Copula model. *The Philippine Statistician*, 72(1), 71-89.

[30] Moji N., Matlou B., Mncela S., Koning F. and Shongwe S.C. (2024). A statistical analysis to evaluate whether gold is still the safe haven of old. *Journal of Statistics Applications & Probability*, 13(5), 1447-1463. http://doi.org/10.18576/jsap/130503.

[31] Beirlant J., Goegebeur Y., Teugels J. and Segers J. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley & Sons: UK. https://doi.org/10.1002/0470012382.

[32] Nerantzaki S.D. and Papalexiou S.M. (2019). Tails of extremes: Advancing a graphical method and harnessing big data to assess precipitation extremes. *Advances in Water Resources*, 134, 103448. https://doi.org/10.1016/j.advwatres.2019.103448.