Empirical Examination of Operational Loss Distributions

Svetlozar T. Rachev

Institut für Statistik und Mathematische Wirtschaftstheorie
Universität Karlsruhe

and

Department of Statistics and Applied Probability
UC Santa Barbara

Anna Chernobai

Department of Statistics and Applied Probability
UC Santa Barbara

Christian Menn

School of Operations Research and Industrial Engineering
Cornell University

1 Introduction

Until very recently, it has been believed that banks are exposed to two main types of risks: credit risk (the counterparty failure risk) and market risk (the risk of loss due to changes in market indicators, such as interest rates and exchange rates), in the order of importance. The remaining financial risks have been put in the category of *other* risks, operational risk being one of them. Recent developments in the financial industry have shown that the importance of operational risk has been largely under-estimated. Newly defined capital requirements set by the Basel Committee for Banking Supervision in 2004, require financial institutions to estimate the capital charge to cover their operational losses [6].

This paper is organized as follows. In Section 2 we give the definition of operational risk and describe the effect of the recent developments in the global financial industry on banks' exposures to operational risk. The following section, Section 3, will briefly outline the recent requirements set by the Basel Committee regarding the capital charge for operational risk. After that we proceed to Section 4 that presents several alternative models that can be used for operational risk modeling. In Section 5 the class of heavy-tailed αStable distributions and their extensions are defined and reviewed. Section 6 consists of the empirical analysis with real operational risk data. Finally, Section 7 summarizes the findings and discusses directions for future work.

2 Definition of Operational Risk in Finance

Operational risk has been recently defined as 'the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events' [4]. Examples include losses resulting from deliberate or accidental accounting errors, equipment failures, credit card fraud, tax noncompliance, unauthorized trading activities, business disruptions due to natural disasters and vandalism. Operational risk affects the soundness and efficiency of all banking activities.

Until recently, the importance of operational risk has been highly underestimated by the banking industry. The losses due to operational risk has been largely viewed as unsubstantial in magnitude, with a minor impact on the banking decision-making and capital allocation. However, increased investors' appetites have led to significant changes in the global financial industry during the last couple of decades - globalization and deregulation, accelerated technological innovation and revolutionary advances in the information network, and increase in the scope of financial services and products. These have caused significant changes in banks' risk profiles, making banks more vulnerable to various sources of risk. These changes have also brought the operational risk to the center of attention of financial regulators and practitioners.

A number of large-scale (exceeding $1 billion in value) operational losses, involving high-profile financial institutions, have shaken the global financial industry in the past two decades: BCCI (1991), Orange Country (1994), Barings Bank (1995), Daiwa Bank (1995), NatWest (1997), Allied Irish Banks (2002), the Enron scandal (2004), among others.

3 Capital Requirements for Operational Risk

The Basel Committee for Banking Supervision (BCBS) has brought into focus the importance of operational risk in 1998 [2], and since 2001 bank regulators have been working on developing capital-based counter-measures to protect the global banking industry against the risk of operational losses - that have demonstrated to possess a substantial, and at times vital, danger to banks. It has been agreed to include operational risk into the scope of financial risks for which the regulatory capital charge should be set [3]. Currently in progress is the process of developing models for the quantitative assessment of operational risk, to be used for measuring the capital charge.

The Basel Capital Accord (Basel II) has been finalized in June 2004 [6]. It explains the guidelines for financial institutions regarding the capital requirements for credit, market and operational risks (Pillar I), the framework for the supervisory capital assessment scheme (Pillar II), and the market discipline principles (Pillar III). Under the first Pillar, several approaches have been proposed for the estimation of the regulatory capital charge. Bank is allowed to adopt one of the approaches, dependent upon fulfillment of a number of quantitative and qualitative requirements. The Basic Indicator Approach (that takes the capital charge to be a fixed fraction of the bank's gross income) and the Standardized Approach (under which the capital charge is a sum of fixed proportions of the gross incomes across all business lines) are the 'top-down' approaches, since the capital charge is determined 'from above' by the regulators; the Advanced Measurement Approaches (that involve the exact form of the loss data distributions) are the 'bottom-up' approaches, since the capital charge is determined 'from below', being driven by individual bank's internal loss data history and practices.

3.1 Loss Distribution Approach

The Loss Distribution Approach (LDA) is one of the proposed Advanced Measurement Approaches. It makes use of the exact operational loss frequency and severity distributions. A necessary requirement for banks to adopt this approach is an extensive internal database.

In the LDA, all bank's activities are classified into a matrix of 'business lines/event type' combinations. Then, for each combination, using the internal loss data the bank estimates two distributions: (1) the loss frequency and (2) severity. Based on these two estimated distributions, the bank computes the probability distribution function of the cumulative operational loss. The operational capital charge is computed as the simple sum of the one-year Value-at-Risk (VaR) measure (with confidence level such as 99.9%) for each 'business line/ event type' pair. The capital charge for a general case (8 business lines and 7 event types) can be expressed as:[1]

$$K_{\text{LDA}} = \sum_{j=1}^{8} \sum_{k=1}^{7} \text{VaR}_{jk}. \tag{3.1}$$

where K_{LDA} is the one-year capital charge under the LDA, and VaR is the Value-at-Risk risk measure,[2] for a one-year holding period and high confidence level (such as 99.9%), based on the aggregated loss distribution, for each 'business line/event type' jk combination.

4 Aggregated Stochastic Models for Operational Risk

Following the guidelines of the Basel Committee, the aggregated operational loss process can be modeled by a *random sum* model, in which the summands are composed of random amounts, and the number of such summands is also a random process. The compound loss process is hence assumed to follow a stochastic process $\{S_t\}_{t \geq 0}$ expressed by the following equation:

$$S_t = \sum_{k=0}^{N_t} X_k, \quad X_k \overset{\text{iid}}{\sim} F_\gamma, \tag{4.1}$$

in which the loss magnitudes (severities) are described by the random independently and identically distributed (iid) sequence $\{X_k\}$ assumed to follow the distribution function (cdf) F_γ that belong to a parametric family of continuous probability distributions, and the density f_γ, and the counting process N_t is assumed to follow a discrete counting process. To avoid the possibility of negative losses, it is natural to restrict the support of the severity distribution to the positive half-line $\mathbb{R}_{>0}$. Representation (4.1) generally assumes (and we also use this assumption) independence between the frequency and severity distributions. The cdf of the compound

[1] Such representation *perfect correlation* between different 'business lines/ event type' combinations. Ignoring possible dependence structures within the banking business lines' and event type profiles may result in overestimation of the capital charge under the LDA approach. The latest Basel II proposal suggested to take into account possible dependencies in the model [6]. Relevant models would involve using techniques such as copulas (see for example numerous works by McNeil and Embrechts on the discussion of copulas), but this is outside the scope of this paper.

[2] $\text{VaR}_{\triangle t,\ 1-\alpha}$ is the risk measure that determines the highest amount of loss that one can expect to lose over a pre-determined time interval (or holding period) $\triangle t$ at a pre-specified confidence level $1-\alpha$. Detailed analysis of VaR models can be found in [11].

process is given by:

$$P(S_t \le s) = \begin{cases} \sum_{n=1}^{\infty} P(N_t = n)\; F_\gamma^{n*}(s) & s > 0 \\ P(N_t = 0) & s = 0, \end{cases} \tag{4.2}$$

where F_γ^{n*} denotes the n-fold convolution with itself.

We summarize the basic properties of such compound process by the expressions for the mean and variance:[3]

$$\begin{aligned} &\text{Mean:} && \mathbb{E}S_t = \mathbb{E}N_t \cdot \mathbb{E}X, \\ &\text{Variance:} && \mathbb{V}S_t = \mathbb{E}N_t \cdot \mathbb{V}X + \mathbb{V}N_t \cdot (\mathbb{E}X)^2. \end{aligned} \tag{4.3}$$

The upper (right) tail behavior has a simple expression for the special case when X belongs to the class of sub-exponential distributions, $X \sim \mathcal{S}$, such as Lognormal, Pareto or the heavy-tailed Weibull. Then the upper tail of the compound process can asymptotically be approximated by (see e.g. [10]):

$$\text{Right tail:} \quad P(S_t > s) \propto \mathbb{E}N_t \cdot P(X > s), \quad s \to \infty. \tag{4.4}$$

4.1 Compound Homogeneous Poisson Process

A stochastic process of the form (4.1) and N_t following a Poisson process with a fixed intensity lambda (λ) is called a *compound homogeneous Poisson process*. It assumes a fixed intensity of the number of loss events in a unit of time. Incorporating the probability mass function of a Poisson distribution into the basic model of Equation (4.2), the cdf of the compound process becomes:

$$P(S_t \le s) = \begin{cases} \sum_{n=1}^{\infty} \frac{(\lambda t)^n e^{-\lambda t}}{n!}\; F_\gamma^{n*}(s) & s > 0 \\ e^{-\lambda t} & s = 0. \end{cases} \tag{4.5}$$

The basic properties of a compound Poisson process can be summarized using Equations (4.3) and (4.4) as follows:

$$\begin{aligned} &\text{Mean:} && \mathbb{E}S_t = \lambda t \cdot \mathbb{E}X, \\ &\text{Variance:} && \mathbb{V}S_t = \lambda t \cdot \mathbb{V}X + \lambda t \cdot (\mathbb{E}X)^2, \\ &\text{Right tail:} && P(S_t > s) \propto \lambda t \cdot P(X > s), \quad s \to \infty, \quad X \sim \mathcal{S}. \end{aligned} \tag{4.6}$$

[3] Of course, this requires the existence of the first and second raw moments of the loss severity distribution.

The mean of the Poisson distribution is the parameter lambda, which is estimated via Maximum Likelihood as the simple arithmetic mean number of events in a unit of time. A number of tests exist to test the Poisson assumption. A common formal test is the Chi-square test. If the model is rejected, then one should consider a more complex alternative model. Next, we briefly review some of them.

4.2 Compound Cox Process

The compound homogeneous Poisson process, discussed earlier, is based on the counting process that is characterized by a fixed intensity lambda. We now relax this assumption. In real life sometimes there are reasons to believe that flow of loss events is often chaotic in nature, and occurrence at each fixed time interval is inhomogeneous and not easily predictable. The compound Cox process, also known in literature as the *doubly stochastic compound Poisson process*, involves a non-constant (or non-homogeneous) form of the intensity factor of the Poisson component of the model. The associated Poisson process is said to be controlled by the random measure $\Lambda(t) = \int_0^t \lambda(s)ds$. A number of scenarios can be considered.

Example 1. The intensity is a random variable, that follows a distribution function, discrete or continuous: $\lambda(t) \sim \mathcal{L}$. For example, λ may take two values λ_1 and λ_2 with probabilities α and $1-\alpha$, respectively. Another example is a Poisson process with intensity λ that follows a two-parameter Gamma distribution. Such counting model is known as a Negative-Binomial model. Such counting models are often called *mixed Poisson models*. The basic properties of compound mixed Poisson processes are dependent on the distribution of the underlying intensity process. Let μ_λ and σ_λ^2 denote the expectation and variance of λ. Then

$$\begin{aligned}
&\text{Mean:} && \mathbb{E}S_t = \mu_\lambda t \cdot \mathbb{E}X, \\
&\text{Variance:} && \mathbb{V}S_t = \mu_\lambda t \cdot \mathbb{V}X + (\mu_\lambda t + t^2\sigma_\lambda^2) \cdot (\mathbb{E}X)^2, \qquad (4.7) \\
&\text{Right tail:} && P(S_t > s) \propto \mu_\lambda t \cdot P(X > s), \quad s \to \infty, \quad X \sim \mathcal{S}.
\end{aligned}$$

Example 2. The intensity is of form $\lambda(t)$ and is dependent on time. The associated cumulative intensity is of form $\Lambda(t)$, a positive non-decreasing process. Here, one example would be a deterministic process that fits the number of losses per unit of time, examined over a prolonged time interval. Another scenario would incorporate a random component into the model.

Here, Brownian Motion and other stochastic models can be of use. Given a particular value of the intensity, the conditional compound Cox process coincides with the compound homogeneous Poisson Process and preserves the properties.

4.3 Renewal Process

Another approach to aggregate losses occurring at random times would be to consider looking at the inter-arrival times, instead of the number of arrivals, in a fixed time interval. Such models are called the *renewal models*. A Poisson counting process implies that the inter-arrival times between the loss events are distributed as an Exponential random variable with mean $1/\lambda$. This assumption on the distribution of the inter-arrival times can be relaxed, and a wider class of distributions can be fitted to the loss inter-arrival times data.

An excellent reference on random sum models and applications to financial data is [1].

4.4 Aggregated Model for Left-Truncated Loss Data

In the operational loss modeling, one should be careful to possible rules that banks follow in recording their internal loss data into the databases. For example, a reality that is often neglected in practical models, is that banks record the losses beginning from a minimum collection threshold of \$5,000-\$10,000. In the external databases the threshold is even higher - \$1,000,000. Hence, the recorded (and observed) loss data is *left-truncated*, and the associated frequency is below the complete-data frequency. This has direct implication on the model specification. Correctly specifying the model for the truncated data, we arrive at the following expressions of the loss severity and frequency probability density/mass functions (pdf/pmf) (assuming, for simplicity, a simple homogeneous Poisson counting process):

$$\begin{aligned} &\text{Severity pdf:} \qquad f_\gamma(x \mid x > u) = \begin{cases} \frac{f_\gamma(x)}{\bar{F}_\gamma(u)} & \text{if } x > u \\ 0 & \text{if } x \le u, \end{cases} \\ &\text{Frequency pmf:} \quad P(N_t = n) = \frac{(\tilde{\lambda}t)^n e^{-\tilde{\lambda}t}}{n!}, \end{aligned} \tag{4.8}$$

where u is the threshold level, $\bar{F}_\gamma(u) = 1 - F_\gamma(u)$, λ is the complete-data frequency parameter, $\tilde{\lambda} = \lambda \cdot \bar{F}_\gamma(u)$ is the truncated data frequency parameter, and N_t is the counting process for the number of losses exceeding u.

In application to operational risk, the truncated compound Poisson model has been introduced and studied in [8]. Further studies include [7], [12].

5 Pareto αStable Distributions

A wide class of distributions that appear highly suitable for modeling operational losses is the class of αStable (or Pareto Stable) distributions. Although no closed-form density form (in the general case) poses difficulties in the estimations, αStable distributions possess a number of attractive features that make them relevant in applications to a variety of financial models. An excellent reference on αStable is [14]. A profound discussion of applications to the financial data can be found in [13]. We now review the definition and basic properties.

5.1 Definition of an αStable Random Variable

A random variable X is said to follow an αStable distribution – we use the notation $X \sim S_\alpha(\sigma, \beta, \mu)$ – if for any $n \geq 2$, there exist $C_n > 0$ and $D_n \in \mathbb{R}$ such that

$$\sum_{k=1}^{n} X_k \stackrel{d}{=} C_n X + D_n, \tag{5.1}$$

where $X_k,\ k = 1, 2, ..., n$ are iid copies of X. The stability property is governed by the constant $C_n = n^{1/\alpha},\ 0 < \alpha \leq 2$. The stability property is a useful and convenient property, and dictates that the distributional form of the variable is preserved under affine transformations. $\alpha = 2$ corresponds to the Gaussian case. $0 < \alpha < 2$ refers to the non-Gaussian case. When we refer to a $S_\alpha(\sigma, \beta, \mu)$ distribution, we mean the latter case. References on the $S_\alpha(\sigma, \beta, \mu)$ distributions and properties include [14], [13].

5.2 Key characteristics of an αStable Random Variable

For the $S_\alpha(\sigma, \beta, \mu)$ random variables, the closed form density exists only for Gaussian ($\alpha = 2$), Cauchy ($\alpha = 1, \beta = 0$) and Lévy ($\alpha = 1/2, \beta = \pm 1$) densities. For the general case, the distribution is expressed by its characteristic function that takes the form:

$$\mathbb{E}e^{iuX} = \begin{cases} \exp\left(-|\sigma u|^\alpha (1 - i\beta(\operatorname{sign} u) \tan\frac{\pi\alpha}{2}) + i\mu u\right), & \alpha \neq 1 \\ \exp\left(-\sigma|u|(1 + i\beta\frac{2}{\pi}(\operatorname{sign} u) \ln|u|) + i\mu u\right), & \alpha = 1 \end{cases} \tag{5.2}$$

The four parameters[4] defining the $S_\alpha(\sigma, \beta, \mu)$ distribution are:

α, the index of stability: $\alpha \in (0,\ 2)$;

β, the skewness parameter: $\beta \in [-1,\ 1]$;

σ, the scale parameter: $\sigma \in \mathbb{R}_+$;

μ, the location parameter: $\mu \in \mathbb{R}$.

Because of the four parameters, the distribution is highly flexible and suitable for fitting to the data which is non-symmetric (skewed) and possesses a high peak (kurtosis) and heavy tails. The heaviness of tails is driven by the *power tail decay* property (see next).

We briefly present the basic properties of the $S_\alpha(\sigma, \beta, \mu)$ distribution. Let $X \sim S_\alpha(\sigma, \beta, \mu)$ with $\alpha < 2$, then

$$\begin{aligned} &\text{Mean:} && \mathbb{E}X = \begin{cases} \mu & \text{if } \alpha > 1 \\ \infty & \text{otherwise,} \end{cases} \\ &\text{Variance:} && \mathbb{V}X = \infty \quad \text{(no second moment),} \\ &\text{Tail:} && P(|X| > x) \propto \text{const} \cdot x^{-\alpha},\ x \to \infty \text{ (power tail decay).} \end{aligned} \tag{5.3}$$

5.3 Useful Transformations of Pareto Stable Random Variables

For $\alpha > 1$ or $|\beta| < 1$, the support of $S_\alpha(\sigma, \beta, \mu)$ distribution equals the whole real line, and is useful for modeling data that can take negative and positive values. It would be unwise to directly apply this distribution to the operational loss data, because it takes only positive values. We suggest to use the following three transformations of the random variable to which the Stable law can be applied.

5.3.1 Symmetric αStable Random Variable

A random variable X is said to follow the symmetric αStable distribution, i.e. $X \sim S\alpha S$, if the $S_\alpha(\sigma, \beta, \mu)$ distribution is symmetric and centered around zero. Then there are only two parameters that need to be estimated, α and σ, and the remaining two are $\beta, \mu = 0$.

[4]The parametrization of αStable laws is not unique. The presented one has been propagated by Samorodnitsky and Taqqu [14]. An overview of the different approaches can be found in [15].

To apply $S\alpha S$ distribution to the operational loss severity data, one can do a simple transformation to the original data set: $Y = [-X;\ X]$. The resulting random variable Y is then symmetric around zero:

$$\begin{aligned} &f_Y(y) = g(y), \quad g \in S_\alpha(\sigma, 0, 0) \\ &x = |y|,\ \alpha \in (0, 2),\ \sigma > 0. \end{aligned} \tag{5.4}$$

5.3.2 log-αStable Random Variable

It is often convenient to work with the natural *log* transformation of the original data. A typical example is the Lognormal distribution: if X follows a Lognormal distribution, then $\log X$ follows a Normal distribution with the same location and scale parameters μ and σ. The same procedure can be applied here. A random variable X is said to follow a log-αStable distribution, i.e. $X \sim \log S_\alpha(\sigma, \beta, \mu)$, if

$$\begin{aligned} &f_X(x) = \tfrac{g(\ln x)}{x}, \quad g \in S_\alpha(\sigma, \beta, \mu) \\ &x > 0,\ \alpha \in (0, 2),\ \beta \in [-1, 1],\ \sigma, \mu > 0. \end{aligned} \tag{5.5}$$

Fitting $\log S_\alpha(\sigma, \beta, \mu)$ distribution to data is appropriate when there is reason to believe that the data is very heavy-tailed, and the regular $S_\alpha(\sigma, \beta, \mu)$ distribution may not capture the heaviness in the tails.

5.3.3 Truncated αStable Random Variable

Another scenario would involve a restriction on the density, rather than a transformation of the original data set. The support of the $S_\alpha(\sigma, \beta, \mu)$ distribution can be restricted to the positive half-line, and the estimation part would involve fitting a truncated Stable distribution of the form:

$$\begin{aligned} &f_X(x) = \tfrac{g(x)}{1 - G(0)} \times \mathbb{I}_{x>0} \\ &\text{where} \quad \mathbb{I}_{x>0} := \begin{cases} 1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x \leq 0, \end{cases} \end{aligned} \tag{5.6}$$

where $g(x) \in S_\alpha(\sigma, \beta, \mu)$, and $G(0)$ denotes the cdf of the $S_\alpha(\sigma, \beta, \mu)$ distribution at zero. Fitting the left-truncated Stable distribution to the data means fitting the right tail of the distribution.

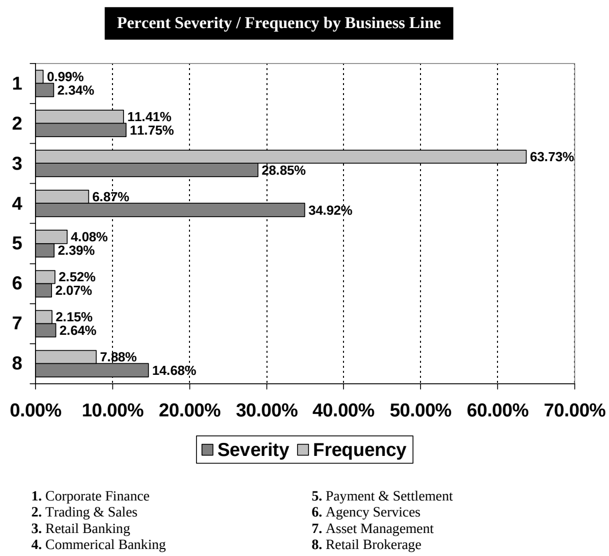
6 Empirical Examination of Operational Loss Data

In the previous sections we discussed the models that can be used to model operational losses. Choosing the right methodology is crucial for accurately estimating the operational risk regulatory capital charge. In addition, understanding the structure of the underlying model that drives the process of the occurrences and severity of the losses is vital for the sound risk management practices and control. In this section we examine the operational loss data and derive conclusions regarding the loss distributions that would be appropriate for modeling the data. We will in particular focus on the loss severity data.

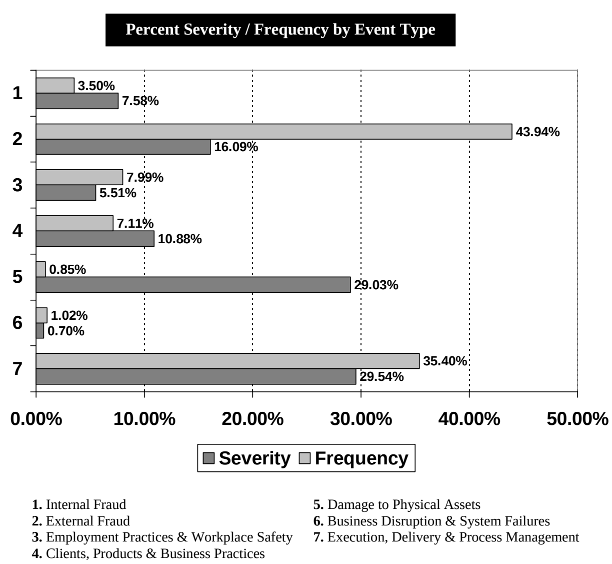
6.1 Results of Basel Committee Loss Data Collection Exercise

In 2002 the Risk Management Group of the BCBS carried out an Operational Loss Data Collection Excercise (OLDC) (also called the third Quantitative Impact Study (QIS3)) aimed at examining various aspects of banks' internal operational loss data [5]. Banks activities were broadly divided into eight business lines and seven loss event types (see Figure 1 for the full list of business lines and event types). Figure 1 demonstrates the severity of losses (i.e. loss amount) and the frequency of losses (i.e. number of losses) by business lines and event types, as a percentage of total. The results are based on the one year of loss data (collected in 2001) provided by 89 participant banks.

The results of the data collection exercise demonstrate a rough picture for the non-uniform nature of the distribution of loss amounts and frequency across various business lines and event types. The results also suggested that the losses are highly right-skewed and have a heavy right tail (i.e. the total loss amounts are highly driven by the high-magnitude 'tail events') [5]. For example, the Commercial Banking business line includes losses of a relatively low frequency (roughly 7% of total) but the second highest severity (roughly 35% of total). As for the losses classified by event type, the losses in the Damage to Physical Assets category (such as losses due to natural disasters) account for less than 1% of the total number of losses, but almost 30% of the aggregate amount. In particular, the 'retail banking/ external fraud' and 'commercial banking/ damage to physical assets' combinations account for over 10% of the total loss amount each, with the first pair accounting for 38% and the second for merely 0.1% of the total number of loss events [5].



(a) Business Line



(b) Event Types

Figure 1: Percentage Frequency and Severity of Operational Losses across Business Lines and Event Types

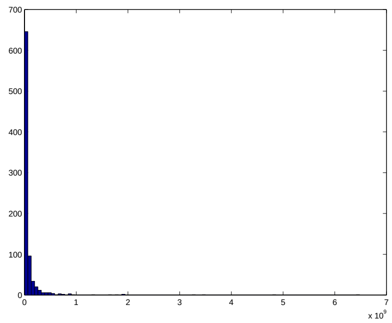
6.2 Analysis of 1980-2002 External Operational Loss Data

In this section we fit various distributions to the operational risk data, obtained from a major European operational loss data provider. The external database is comprised of operational loss events throughout the world. The original loss data cover losses in the period 1950-2002. As discussed earlier, the data in external databases are subject to minimum recording thresholds of $1 million. A few recorded data points were below $1 million in nominal value, so we excluded them from the dataset. Furthermore, we excluded the observations before 1980 because of relatively few data points available (which is most likely due to poor data recording practices). The final dataset for the analysis covered losses in US dollars for the time period between 1980 and 2002. It consists of five types of losses: "Relationship" (such as events related to legal issues, negligence and sales-related fraud), "Human" (such as events related to employee errors, physical injury and internal fraud), "Processes" (such as events related to business errors, supervision, security and transactions), "Technology" (such as events related to technology and computer failure and telecommunications) and "External" (such as events related to natural and man-made disasters and external fraud). The loss amounts have been adjusted for inflation using the Consumer Price Index from the U.S. Department of Labor. The numbers of data points of each of the "Relationship", "Human", "Processes", "Technology", and "External" types are $n = 849$, 813, 325, 67, and 233, respectively. Figure 2 presents the histograms for the five loss types of data. The histograms (the horizontal axis covers the entire range of the data) indicate the leptokurtic nature of the data: a very high peak is observed close to zero, and an extended right tail indicates the right-skewness and high dispersion of the data values.

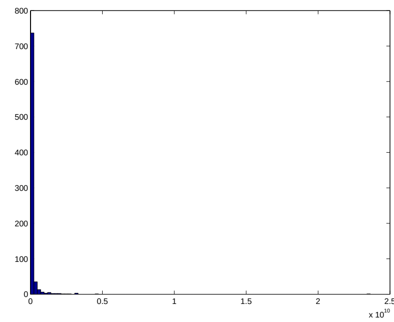
6.2.1 Operational Loss Frequency Process

Figure 3 portrays the annually aggregated number of losses for the "External" type losses, shown by the dotted-line. It suggests that the accumulation is somewhat similar to a continuous cdf-like process, supporting the use of a non-homogeneous Poisson process. We consider two following functions,[5] each with four parameters:

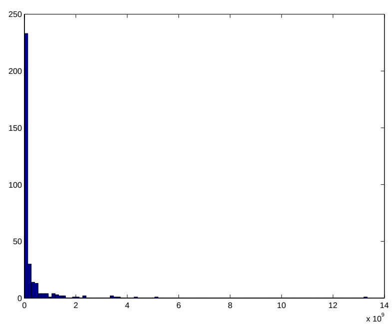
[5] Of course, asymptotically (as time increases) such functions would produce a constant cumulative intensity. However, for this particular sample and this particular time frame, this form of the intensity function appears to provide a good fit.



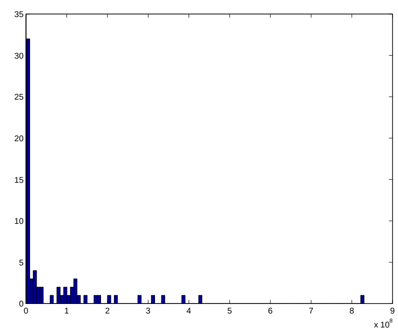
(a) “Relationship”



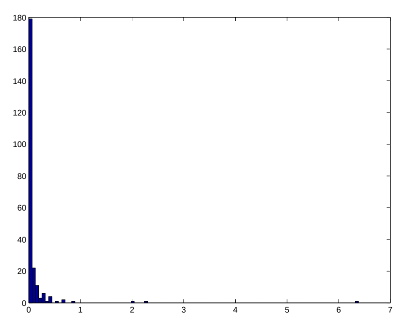
(b) “Human”



(c) “Processes”



(d) “Technology”



(e) “External”

Figure 2: Histograms for operational loss data from external sources.

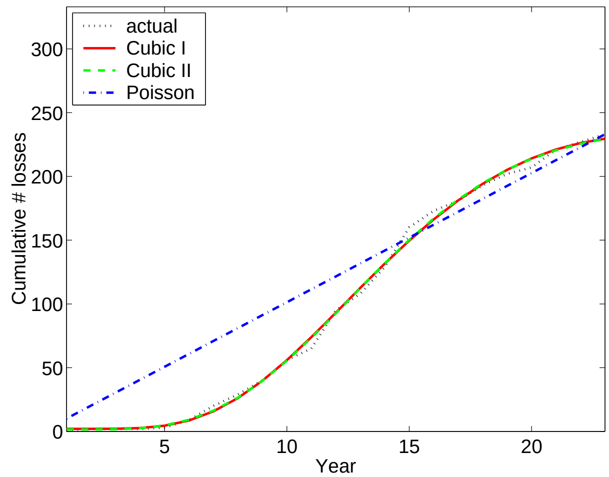


Figure 3: Annual accumulated number of "External" operational losses, with fitted non-homogeneous and simple Poisson models.

- **Type I**: Lognormal cdf-like process of form:

$$\Lambda(t) = a + \frac{b}{\sqrt{2\pi}c} \exp\left\{-\frac{(\log t - d)^2}{2c^2}\right\};$$

- **Type II**: Logweibull cdf-like process of form

$$\Lambda(t) = a - b \exp\left\{-c \log^d t\right\}.$$

We obtain the four parameters a, b, c, d by minimizing the Mean Square Error. Table 1 demonstrates the estimated parameters and the Mean Square Error (MSE) and the Mean Absolute Error (MAE) for the cumulative intensities and a simple homogeneous Poisson process with a constant intensity factor. Figure 3 shows the three fits plotted together with the actual aggregated number of events. The two non-linear fits appear to be superior to the standard Poisson, for all 5 loss datasets. We thus reject the conjecture that the counting process is simple Poisson.

Table 1: Fitted frequency functions to the "External" type losses.

process					MSE	MAE
Type I	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{c}$	$\boldsymbol{d}$		
	2.02	305.91	0.53	3.21	16.02	2.708
Type II	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{c}$	$\boldsymbol{d}$		
	237.88	236.30	0.00026	8.27	14.56	2.713
Poisson				λ		
				10.13	947.32	24.67

6.2.2 Operational Loss Distributions

The following loss distributions were considered for this study.

- Exponential $\mathcal{E}xp(\beta)$ $f_X(x) = \beta e^{-\beta x}$
  $x \geq 0,\ \beta > 0$
- Lognormal $\mathcal{LN}(\mu, \sigma)$ $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\}$
  $x \geq 0,\ \mu, \sigma > 0$
- Weibull $\mathcal{W}eib(\beta, \tau)$ $f_X(x) = \tau\beta x^{\tau-1} \exp\{-\beta x^\tau\}$
  $x \geq 0,\ \beta, \tau > 0$
- Logweibull $\log \mathcal{W}eib(\beta, \tau)$ $f_X(x) = \frac{1}{x}\tau\beta(\log x)^{\tau-1} \exp\{-\beta(\log x)^\tau\}$
  $x \geq 0,\ \beta, \tau > 0$
- Log-αStable $\log S_\alpha(\sigma, \beta, \mu)$ $f_X(x) = \frac{g(\ln x)}{x}, g \in S_\alpha(\sigma, \beta, \mu)$
  no closed-form density
  $x > 0,\ \alpha \in (0, 2),\ \beta \in [-1, 1],\ \sigma, \mu > 0$
- Symmetric αStable $S\alpha S(\sigma)$ $f_Y(y) = g(y), g \in S_\alpha(\sigma, 0, 0),$
  no closed-form density
  $x = |y|,\ \alpha \in (0, 2),\ \sigma > 0$

All except the $S\alpha S$ distributions are defined on $\mathbb{R}_+$, making them applicable for the operational loss data. For the $S\alpha S$ distribution, we symmetrized the data by multiplying the losses by -1 and then adding them to the original dataset.

We fitted conditional loss distribution to the data (see Equation (4.8)) with the minimum threshold of $u = 1,000,000$, using the method of Maximum Likelihood. Parameter estimates are presented in Table 2.

Table 2: Estimated parameters for the loss data separated by event type.

$\widehat{\gamma}_{\text{MLE}}$	"Rel.-ship"	"Human"	"Proc."	"Tech."	"Ext."
$\mathcal{E}xp(\beta)$					
β	$11.25{\cdot}10^{-9}$	$7.27{\cdot}10^{-9}$	$3.51{\cdot}10^{-9}$	$13.08{\cdot}10^{-9}$	$9.77{\cdot}10^{-9}$
$\mathcal{LN}(\mu,\sigma)$					
μ	16.1911	15.4627	17.1600	15.1880	15.7125
σ	2.0654	2.5642	2.3249	2.7867	2.3639
$\mathcal{W}eib(\beta,\tau)$					
β	0.0032	0.0240	0.0021	0.0103	0.0108
τ	0.3538	0.2526	0.3515	0.2938	0.2933
$\log\mathcal{W}eib(\beta,\tau)$					
β	$0.27{\cdot}10^{-8}$	$30.73{\cdot}10^{-8}$	$0.11{\cdot}10^{-8}$	$11.06{\cdot}10^{-8}$	$2.82{\cdot}10^{-8}$
τ	7.0197	7.0197	7.1614	5.7555	6.2307
$\log S_\alpha(\sigma,\beta,\mu)$					
α	1.9340	1.4042	2.0000	2.0000	1.3313
β	-1	-1	0.8195	0.8040	-1
σ	1.5198	2.8957	1.6476	1.9894	2.7031
μ	15.9616	10.5108	17.1535	15.1351	10.1928
$S\alpha S(\sigma)$					
α	0.6592	0.6061	0.5748	0.1827	0.5905
σ	$1.0{\cdot}10^{7}$	$0.71{\cdot}10^{7}$	$1.99{\cdot}10^{7}$	$0.17{\cdot}10^{7}$	$0.71{\cdot}10^{7}$

6.2.3 Validation Tests for Loss Models

A variety of tests can be considered to examine the goodness of fit of the distributions to the data. Typical tests include performing the Likelihood-Ratio test, examination of Quantile-Quantile plots, forecasting, and various tests based on the comparison of the fitted distribution and the empirical distribution (the so-called EDF-based tests). In this paper, we focus on the last testing procedure, because it allows us to compare separately the fits of the distributions around the center and around the tails of the data.

First, we compare the magnitudes of several EDF test statistics between different models. A lower test statistic value indicates a better fit (in the sense that the value of the norm, which is based on the distance between the fitted and empirical cdf's, is smaller). Second, we compare the p-values based on the EDF tests. p-values indicate the proportion of times in which the samples drawn from the same fitted distributions have a higher statistic value. In other words, a higher p-value suggests a better fit. The following test statistics are considered: Kolmogorov-Smirnov (D), Kuiper (V), quadratic Anderson-Darling (A^2), quadratic *"upper tail"* Anderson-Darling (A^2_{up}) and Cramér-von Mises (W^2), computed as

$$\begin{aligned}
D &= \max\left\{D^+, D^-\right\}, \\
V &= D^+ + D^-, \\
A^2 &= n\int_{-\infty}^{\infty} \frac{(F_n(x) - \widehat{F}(x))^2}{\widehat{F}(x)(1-\widehat{F}(x))} d\widehat{F}(x), \\
A^2_{up} &= n\int_{-\infty}^{\infty} \frac{(F_n(x) - \widehat{F}(x))^2}{(1-\widehat{F}(x))^2} d\widehat{F}(x), \\
W^2 &= n\int_{-\infty}^{\infty} (F_n(x) - \widehat{F}(x))^2 d\widehat{F}(x),
\end{aligned}$$

where $D^+ = \sqrt{n}\sup_x\{F_n(x) - \widehat{F}(x)\}$ and $D^- = \sqrt{n}\sup_x\{\widehat{F}(x) - F_n(x)\}$. The A^2_{up} statistic was introduced and studied in [9], and designed to put most of the weight on the upper tail. $F_n(x)$ is the empirical cdf, and $\widehat{F}(x)$ is defined as

$$\widehat{F}(x) = \begin{cases} \frac{\widehat{F}_\gamma(x) - \widehat{F}_\gamma(u)}{1-\widehat{F}_\gamma(u)} & x > u \\ 0 & x \le u. \end{cases}$$

Table 3[6] demonstrates that on the basis of the statistic values we would tend to conclude that Logweibull, Weibull or Lognormal densities describe best the dispersion of the operational loss data: the statistics are the lowest for these models in most cases. However, if we wish to test the null that a given dataset belongs to a family of distributions (such as Lognormal or

[6] The fit of the exponential distribution is totally unsatisfactory and the results have been omitted for saving space.

Table 3: Goodness-of-fit test statistics for the loss data.

	"Rel.-ship"	"Human"	"Proc."	"Tech."	"Ext."
			D		
$\mathcal{LN}$	0.8056	0.8758	0.6854	1.1453	0.6504
$\mathcal{W}eib$	0.5553	0.8065	0.6110	1.0922	0.4752
$\log\mathcal{W}eib$	0.5284	0.9030	0.5398	1.1099	0.6893
$\log S_\alpha(\sigma,\beta,\mu)$	1.5929	9.5186	0.6931	1.1540	7.3275
$S\alpha S$	1.1634	1.1628	1.3949	2.0672	0.7222
			V		
$\mathcal{LN}$	1.3341	1.5265	1.1262	1.7896	1.2144
$\mathcal{W}eib$	1.0821	1.5439	1.0620	1.9004	0.9498
$\log\mathcal{W}eib$	1.0061	1.5771	0.9966	1.9244	1.1020
$\log S_\alpha(\sigma,\beta,\mu)$	1.6930	9.5619	1.1490	1.7793	7.4089
$S\alpha S$	2.0695	2.1537	1.9537	2.8003	1.4305
			A^2		
$\mathcal{LN}$	0.7554	0.7505	0.4624	1.3778	0.5816
$\mathcal{W}eib$	0.7073	0.7908	0.2069	1.4536	0.3470
$\log\mathcal{W}eib$	0.4682	0.7560	0.1721	1.5355	0.4711
$\log S_\alpha(\sigma,\beta,\mu)$	3.8067	304.61	0.4759	1.3646	194.74
$S\alpha S$	4.4723	11.9320	6.5235	19.6225	1.7804
			A^2_{up}		
$\mathcal{LN}$	4.6122	4.5160	4.0556	6.4213	2.5993
$\mathcal{W}eib$	13.8191	8.6610	2.2340	4.8723	5.3662
$\log\mathcal{W}eib$	5.2316	4.5125	1.4221	5.2992	4.1429
$\log S_\alpha(\sigma,\beta,\mu)$	10.1990	4198.9	4.0910	6.4919	3132.6
$S\alpha S$	$2.6\cdot 10^{14}$	$3.3\cdot 10^{11}$	$6.8\cdot 10^{14}$	$7.2\cdot 10^{10}$	$1.2\cdot 10^{10}$
			W^2		
$\mathcal{LN}$	0.1012	0.0804	0.0603	0.2087	0.0745
$\mathcal{W}eib$	0.0716	0.0823	0.0338	0.2281	0.0337
$\log\mathcal{W}eib$	0.0479	0.0915	0.0241	0.2379	0.0563
$\log S_\alpha(\sigma,\beta,\mu)$	0.7076	44.5156	0.0660	0.2072	24.3662
$S\alpha S$	0.3630	0.2535	0.3748	1.4411	0.1348

Table 4: p-values associated with the goodness-of-fit test statistics for the loss data. p-values were obtained from 1,000 simulated samples. Figures in bold show p-values whenever their values were the first or second highest.

	"Rel.-ship"	"Human"	"Proc."	"Tech."	"Ext."
			D		
$\mathcal{LN}$	0.082	0.032	0.297	<0.005	0.326
$\mathcal{W}eib$	0.625	0.103	0.455	<0.005	0.852
$\log \mathcal{W}eib$	0.699	0.074	0.656	<0.005	0.296
$\log S_\alpha(\sigma,\beta,\mu)$	0.295	**0.319**	0.244	<0.005	0.396
$S\alpha S$	0.034	**0.352**	0.085	**0.085**	**0.586**
			V		
$\mathcal{LN}$	0.138	0.039	0.345	0.005	0.266
$\mathcal{W}eib$	0.514	0.051	0.532	<0.005	0.726
$\log \mathcal{W}eib$	0.628	0.050	0.637	<0.005	0.476
$\log S_\alpha(\sigma,\beta,\mu)$	0.295	**0.324**	0.342	**0.007**	0.458
$S\alpha S$	<0.005	0.026	0.067	**0.067**	0.339
			A^2		
$\mathcal{LN}$	0.043	0.408	0.223	<0.005	0.120
$\mathcal{W}eib$	0.072	0.112	0.875	<0.005	0.519
$\log \mathcal{W}eib$	0.289	0.392	0.945	<0.005	0.338
$\log S_\alpha(\sigma,\beta,\mu)$	**0.290**	0.215	0.202	<0.005	0.284
$S\alpha S$	**0.992**	**0.436**	**0.964**	>**0.995**	**0.841**
			A^2_{up}		
$\mathcal{LN}$	0.401	0.408	0.367	0.067	0.589
$\mathcal{W}eib$	0.081	0.112	0.758	0.087	0.164
$\log \mathcal{W}eib$	0.282	0.392	0.977	0.114	0.283
$\log S_\alpha(\sigma,\beta,\mu)$	**0.288**	0.215	0.361	0.060	0.128
$S\alpha S$	<0.005	**0.436**	0.193	>**0.995**	**0.841**
			W^2		
$\mathcal{LN}$	0.086	0.166	0.294	<0.005	0.210
$\mathcal{W}eib$	0.249	0.188	0.755	<0.005	0.781
$\log \mathcal{W}eib$	0.514	0.217	0.918	<0.005	0.458
$\log S_\alpha(\sigma,\beta,\mu)$	**0.292**	**0.315**	0.258	<0.005	0.366
$S\alpha S$	<0.005	0.027	0.102	**0.964**	0.265

Stable), then the test is not parameter-free, and we need to estimate the p-values for each hypothetical scenario. These results are demonstrated in Table 4. Now the situation is quite different from the one in Table 3. The numbers in bold indicate the cases in which $\log S_\alpha(\sigma, \beta, \mu)$ or $S\alpha S$ fit resulted in first or second highest p-values across the same group (i.e. for the same type of EDF test for a range of distributions, with a particular dataset). As is clear from the table, in the majority of cases (17 out of 25) either $\log S_\alpha(\sigma, \beta, \mu)$ or $S\alpha S$, or even both, resulted in the highest p-values. This supports the conjecture that the overall distribution of operational losses[7] are heavy-tailed. Fitting $\log S_\alpha(\sigma, \beta, \mu)$ or $S\alpha S$ distributions to the data appears a valid solution.

7 Summary

The objective of this paper was to examine the models underlying in the operational risk process. The conjecture that operational losses follow a compound Cox process was investigated for the external operational loss data of five loss types covering a 23 year period. The results of the empirical analysis provide evidence of heavy tailedness of the data in the right tail. Moreover, fitting $\log S_\alpha(\sigma, \beta, \mu)$ distribution to the loss severity data or symmetric $S_\alpha(\sigma, \beta, \mu)$ distribution to the symmetrized data resulted in high p-values in a number of goodness of fit tests, suggesting a good fit. In particular, the two distributions are shown to fit the data very well in the upper tail, which remains the central concern in the framework of operational risk modeling and regulation.

Furthermore, the paper suggested a number of models for the frequency of losses. A simple Poisson process with a fixed intensity factor appears too restrictive and unrealistic. A non-homogeneous Poisson process with a time-varying intensity function was fitted to the loss data and showed a superior fit to the homogeneous Poisson process.

Directions for future research include developing robust models for the operational risk modeling. For example, with $S_\alpha(\sigma, \beta, \mu)$ distributions, for the case when the shape parameter α is below or equal to unity, the first moment (and hence the mean) and the second moment (hence the variance) do not exist, making such distribution difficult to use for practical purposes. Possible solutions would include working with trimmed data, truncated data, or 'Winsorized' data, or splitting the dataset into two parts - the low-

[7] Another approach would be to split each data set into two parts: the main body of the data and the right tail. Some empirical evidence suggests that the two parts of the data follow different laws. Extreme Value Theory is an approach that can be used for such analysis.

and medium-size losses and the tail losses - and analyzing the properties of each separately.

8 Acknowledgements

We are thankful to S. Stoyanov of the FinAnalytica Inc. for computational assistance. We are also grateful to S. Trück for providing us with the data on which the empirical study was based.

C. Menn is grateful for research support provided by the German Academic Exchange Service (DAAD).

S.T. Rachev gratefully acknowledges research support by grants from Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara, the German Research Foundation (DFG) and the German Academic Exchange Service (DAAD).

References

[1] V. E. Bening and V. Y. Korolev. *Generalized Poisson Models and their Applications in Insurance and Finance*. VSP International Science Publishers, Utrecht, Boston, 2002.

[2] BIS. Operational risk management. http://www.bis.org, 1998.

[3] BIS. Consultative document: operational risk. http://www.bis.org, 2001.

[4] BIS. Working paper on the regulatory treatment of operational risk. http://www.bis.org, 2001.

[5] BIS. The 2002 loss data collection exercise for operational risk: summary of the data collected. http://www.bis.org, 2003.

[6] BIS. International convergence of capital measurement and capital standards. http://www.bis.org, 2004.

[7] A. Chernobai, C. Menn, S. T. Rachev, and S. Trück. Estimation of operational value-at-risk in the presence of minimum collection thresholds. Technical report, Department of Statistics and Applied Probability, University of California Santa Barbara, 2005.

[8] A. Chernobai, C. Menn, S. Trück, and S. Rachev. A note on the estimation of the frequency and severity distribution of operational losses. *Mathematical Scientist*, 30(2), 2005.

[9] A. Chernobai, S. Rachev, and F. Fabozzi. Composite goodness-of-fit tests for left-truncated loss samples. Technical report, Department of Statistics and Applied Probability, University of California Santa Barbara, 2005.

[10] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modeling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin, 1997.

[11] P. Jorion. *Value-at-Risk: the New Benchmark for Managing Financial Risk*. McGraw-Hill, New York, second edition, 2000.

[12] M. Moscadelli, A. Chernobai, and S. T. Rachev. Treatment of missing data in the field of operational risk: Effects on parameter estimates, EL, UL and CVaR measures. *Operational Risk*, June 2005.

[13] S. T. Rachev and S. Mittnik. *Stable Paretian Models in Finance*. John Wiley & Sons, New York, 2000.

[14] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Variance*. Chapman & Hall, London, 1994.

[15] V. M. Zolotarev. *One-dimensional Stable Distributions*. Translations of Mathematical Monographs, vol. 65. American Mathematical Society, Providence, 1986.