Modeling Operational Risk

ALEXANDER JÖHNEMARK

Master of Science Thesis
Stockholm, Sweden 2012

Modeling Operational Risk

ALEXANDER JÖHNEMARK

Degree Project in Mathematical Statistics (30 ECTS credits)
Degree Programme in Engineering Physics 270 credits
Royal Institute of Technology year 2012
Supervisor at KTH was Filip Lindskog
Examiner was Filip Lindskog

TRITA-MAT-E 2012:14
ISRN-KTH/MAT/E--12/14--SE

Royal Institute of Technology
School of Engineering Sciences

KTH SCI
SE-100 44 Stockholm, Sweden

URL: www.kth.se/sci

Abstract

The Basel II accord requires banks to put aside a capital buffer against unexpected operational losses, resulting from inadequate or failed internal processes, people and systems or from external events. Under the sophisticated Advanced Measurement Approach banks are given the opportunity to develop their own model to estimate operational risk. This report focus on a loss distribution approach based on a set of real data.

First a comprehensive data analysis was made which suggested that the observations belonged to a heavy tailed distribution. An evaluation of commonly used distributions was performed. The evaluation resulted in the choice of a compound Poisson distribution to model frequency and a piecewise defined distribution with an empirical body and a generalized Pareto tail to model severity. The frequency distribution and the severity distribution define the loss distribution from which Monte Carlo simulations were made in order to estimate the 99.9% quantile, also known as the the regulatory capital.

Conclusions made on the journey were that including all operational risks in a model is hard, but possible, and that extreme observations have a huge impact on the outcome.

Keywords: Operational risk, Advanced Measurement Approaches, Loss Distribution Approach

Acknowledgements

I would like to thank Carl Larsson, my supervisor at Swedbank, for encouragement, ideas and great feedback. My thanks also goes out to Ulf Holmqvist and Tanel Pärss for helping setting me up at the bank and providing me with highly appreciated data. Last but not least I would like to thank my supervisor at KTH, Filip Lindskog, for valuable tips and insightful comments about the writing and structure of the report.

Stockholm, November 2012

Alexander Jöhnemark

Contents

Chapter 1

Introduction

A financial institution is exposed to several major risks, such as market and credit risk, and are required to put aside a capital buffer against unexpected losses. With the implementation of the Basel II recommendations and regulations in Sweden in February 2007 a capital requirement for operational risk was set under regulation by Finansinspektionen[1].

Operational risk is a very broad concept and may include anything from bank robberies and unauthorized trading to terrorist attacks and natural disasters. In other words, it is everything that is not credit, systematic or financial risk, and which arises from the operation of a company's business functions.

Opposite to credit risk and market risk which can be exploited to generate profit, managing operational risk is not used to generate profit. However, it is still managed to keep losses within a company's risk appetite.

1.1 Background

Market risk and credit risk have for a long time been the subject of much debate and research, resulting in considerable progress in the identification, measurement and management of these risks. The increased globalization and the progress in financial technology has moved the financial world into a more complex realm, beyond normal expectations, were highly improbable and severe events reign. Standard models clearly fail to capture the extreme events, as we have seen in the 2008 global financial crisis. In other professional areas flawed models would be unacceptable and never be used. Would you ever cross a bridge that works most of the time? It leads us into more philosophical thoughts: Is it even possible to predict the highly improbable with accurate estimates of its probabilities? And say it does, would it not mean that the unexpected becomes expected and, by definition, is no longer improbable? This paper will not dig too deep into questions regarding the

[1] Finansinspektionen is the Swedish financial services authority.

use of probability distributions to predict processes over time, but rather use established methods and risk measures. Although, conventional tools such as the log-normal distribution and the value at risk measures are challenged. I can recommend reading The Black Swan by Nassim Nicholas Taleb for more in-depth discussions regarding these topics.

Many of these highly improbable events such as the terrorist attacks on September 11, unauthorized trading losses at Barings Bank, resulting in its collapse in 1995, and other rogue tradings resulting in large losses at Société Générale[2], AIB[3] and National Australia Bank[4], have contributed to a growing focus on identification and measurement of operational risk.

Historically, many companies have regarded operational risk as an unavoidable cost of doing business. With the decision by the Basel Committee on Banking Supervision to introduce a capital charge for operational risk as part of the Basel II framework, identification and measurement of operational risk is today a real and everyday function of a bank.

1.2 Definition

There are many different definitions of operational risk and many institutions have adopted their own definition which better reflects their area of business. However, the Basel Committee (2006) define operational risk as:

> *"The risk of loss resulting from inadequate or failed internal processes, people and systems or from external events"*

1.3 Problem Statement

Can operational risk be modeled in a sound and consistent manner and how should a model for operational risk be designed?

1.4 Purpose

The aim of this thesis is to study and evaluate different methods and approaches to modeling operational risk and then develop a sound model for the calculation of capital requirement that fulfills the AMA requirements (AMA will be explained in the next chapter).

[2] "Rogue trader blamed for 4.9 billion euro fraud at Société Générale". Agence France-Presse. 24 January 2008

[3] "Rogue trader 'Mr Middle America'". BBC News. 7 February 2002.

[4] "Former NAB foreign currency options traders sentenced". Australian Securities and Investments Commission. 4 July 2006.

1.5 Literature Overview

There is plenty of literature covering operational risk and various techniques and methods modeling it. Most papers are about theoretical approaches where a couple of operational risk categories with many data points are selected and used to estimate partial capital requirements. Therefore I would like to highlight the following articles:

- F. Aue and M. Kalkbrener (2006) present the capital model developed at Deutsche Bank. The model follows a so called Loss Distribution Approach (LDA), which will be explained in this paper, and describes the use of loss data and scenarios, frequency and severity modeling and the implementation of dependence and the capital calculation.

- Dutta & Perry (2007) perform a comprehensive evaluation of commonly used methods and introduce the g-and-h distribution in order to model the whole severity range with one distribution.

- A. Chapelle, Y. Crama, G. Hübner and J. P. Peters (2007) analyze the implications of AMA through a study on four categories of two business lines and two event types of a large financial institution. They use a mixed model by calibrating one distribution to describe normal losses and another for extreme losses. They also estimate the impact of operational risk management on bank profitability.

- K. Böcker and C. Klüppelberg (2005) propose a simple closed-form approximation for operational value at risk, when loss data are heavy tailed, and apply the approximation to the pareto severity distribution.

- E. W. Cope, G. Mignola, G. Antonini and R. Ugoccioni (2009) examine the data sufficiency in internal and external operational loss data and the difficulties in measuring operational risk.

- A. Colombo and S. Desando (2008) describe the implementation and development of a scenario based approach to measuring operational risk at the Italian bank Intesa Sanpaolo.

- E. Cope and G. Antonini (2008) study observed correlations among operational losses in the ORX database and the implications for diversification benefits when aggregating losses across different operational risk categories.

1.6 Terms & Abbreviations

AMA Advanced Measurement Approaches

BEICFs Business Environment and Internal Control Factors

BIA Basic Indicator Approach

BL Business Line

CDF Cumulative Distribution Function

ED External Data

EDA Exploratory Data Analysis

ES Expected Shortfall

ET Event Type

ILD Internal Loss Data

LDA Loss Distribution Approach

ORMF Operational Risk Management Framework

ORX Operational Riskdata eXchange Association, consortium of institution anonymously collecting operational loss data

PDF Probability Density Function

POT Peaks Over Threshold

QQ-plot Quantile-Quantile plot

SA Standard Approach

SBA Scenario Based Approach

Units of Measure Category for which operational losses that share the same risk profile are sorted in. Simply named *"cells"* in this paper.

VaR Value at Risk

Chapter 2

Theory

In this chapter I will go through the Basel II Framework, different risk measures, distributional assumptions, extreme value theory, popular AMA models, simulation methods and a closed form approximation of the loss distribution.

2.1 Basel II Framework

2.1.1 Categorization

The loss data are organized according to the seven official Basel II defined event types

1. Internal Fraud
2. External Fraud
3. Employment Practices & Workplace Safety
4. Clients, Products, & Business Practice
5. Damage to Physical Assets
6. Business Disruption & Systems Failures
7. Execution, Delivery, & Process Management

and eight defined business lines

1. Corporate Finance
2. Trading & Sales
3. Retail Banking
4. Commercial Banking

5. Payment & Settlement
6. Agency Services
7. Asset Management
8. Retail Brokerage

This categorization is based on the principle to organize losses that share the same basic risk profile and behaviour pattern. Sorting the data in these categories yields a matrix with 56 buckets, or cells, of data.

2.1.2 Methods

The Basel Committee on Banking Supervision have prescribed guidance for three types of methods for the calculation of capital requirement for operational risk. Those are the Basic Indicator Approach (BIA), Standard Approach (SA) and Advanced Measurement Approaches (AMA). The latter is the most sophisticated of the approaches and is what this thesis is about. Below follows a short presentation of the different approaches.

Basic Indicator Approach

Is the least complicated method and is based on the annual revenue of the financial institution. The bank must hold a capital for operational risk based on a fixed percentage, set to 15% by the Committee, of the past three years positive average annual gross income. Years where annual gross income is negative or zero should be excluded from the calculation of the average.

Standardized Approach

In the standardized approach, capital requirements are calculated based on the annual revenue of each business line. Similar to the BIA, capital charge is based on a fixed percentage (called beta factors) of the three year positive average annual gross income, but for each business line.

Advanced Measurement Approaches

The Advanced Measurement Approaches allows a bank to internally develop its own risk measurement model. The rules under advanced measurement approaches requires that a bank's operational loss data captures the operational risks to which the firm is exposed to. The model must include credible, transparent, systematic and verifiable approaches for weighting internal operational loss data, external operational loss data, scenario analysis and BEICFs (Basel Committee, 2011. p. 46) [1].

- Risk measure: Value at Risk

Business Lines	Beta Factor
Corporate Finance	18%
Trading & Sales	18%
Retail Banking	12%
Commercial Banking	15%
Payment & Settlement	18%
Agency Services	15%
Asset Management	12%
Retail Brokerage	12%

Table 2.1: List of beta factors for each business line (Basel Committee, 2006)

- Time horizon: 1 year
- Confidence level: 99.9%

In Sweden there are additional requirements, mostly qualitative, that also have to be fulfilled[1].

There are a variety of AMA models which differs in emphasis and ways of combining the four data elements. The most common methods used in modeling operational risk is the popular loss distribution approach (LDA) followed by scenario based approaches (SBA) (Basel Committee, 2011. p. 34) [1].

2.1.3 The Four Data Elements

The following section have some short descriptions of each of the four data elements as well as some longer quotes from the Supervisory Guidelines for the Advanced Measurement Approaches (Basel Committee, 2011) [1]. The initial number in every quote is simply the numbered subsection from which the quote is taken from in the Basel document.

Internal Data

The internal loss data are considered the most important input to the model and are thought to reflect the bank's risk profile most accurately. The data are exclusively used in calibrating the frequency parameters and are used in combination with external data to calibrate the severity distribution. Also frequency dependencies are analyzed using internal data.

> 247. While the Basel II Framework provides flexibility in the way a bank combines and uses the four data elements in its operational risk management framework (ORMF), supervisors expect

[1] Ansökan om internmätningsmetod, operativ risk. Finansinspektionen, 2007.

> that the inputs to the AMA model are based on data that represent or the bank's business risk profile and risk management practices. ILD is the only component of the AMA model that records a bank's actual loss experience. Supervisors expect ILD to be used in the operational risk measurement system (ORMS) to assist in the estimation of loss frequencies; to inform the severity distribution(s) to the extent possible; and to serve as an input into scenario analysis as it provides a foundation for the bank's scenarios within its own risk profile. The Committee has observed that many banks have limited high severity internal loss events to inform the tail of the distribution(s) for their capital charge modeling. It is therefore necessary to consider the impact of relevant ED and/or scenarios for producing meaningful estimates of capital requirements.

External

External data can be used to enrich the scarce internal data. They can also be used to modify parameters derived from internal data, or as input in scenarios and for benchmarking. Even though external data do not fully reflect the bank's risk profile, and are generally more heavy tailed than internal data, external data can still be more reliable for calibrating the tail distribution than internal.

> 248. ED provides information on large actual losses that have not been experienced by the bank, and is thus a natural complement to ILD in modelling loss severity. Supervisors expect ED to be used in the estimation of loss severity as ED contains valuable information to inform the tail of the loss distribution(s). ED is also an essential input into scenario analysis as it provides information on the size of losses experienced in the industry. Note that ED may have additional uses beyond providing information on large losses for modelling purposes. For example, ED may be useful in assessing the riskiness of new business lines, in benchmarking analysis on recovery performance, and in estimating competitors' loss experience.
>
> 249. While the ED can be a useful input into the capital model, external losses may not fit a particular bank's risk profile due to reporting bias. Reporting bias is inherent in publicly-sourced ED and therefore focuses on larger, more remarkable losses. A bank should address these biases in their methodology to incorporate ED into the capital model.
>
> 250. As ED may not necessarily fit a particular bank's risk profile, a bank should have a defined process to assess relevancy

and to scale the loss amounts as appropriate. A data filtering process involves the selection of relevant ED based on specific criteria and is necessary to ensure that the ED being used is relevant and consistent with the risk profile of the bank. To avoid bias in parameter estimates, the filtering process should result in consistent selection of data regardless of loss amount. If a bank permits exceptions to its selection process, the bank should have a policy providing criteria for exceptions and documentation supporting the rationale for any exceptions. A data scaling process involves the adjustment of loss amounts reported in external data to fit a bank's business activities and risk profile. Any scaling process should be systematic, statistically supported, and should provide output that is consistent with the bank's risk profile.

251. To the extent that little or no relevant ED exists for a bank, supervisors would expect the model to rely more heavily on the other data elements. Limitations in relevant ED most frequently arise for banks operating in distinct geographic regions or in specialised business lines.

Scenario Analysis

Scenarios are used as a complement to historical loss data and are used where data are scarce. Scenario data are forward looking, unlike external and internal data, and include events that have not yet occurred. Scenario analysis is inherently biased, such as anchoring, availability and motivational biases.

252. A robust scenario analysis framework is an important element of the ORMF. This scenario process will necessarily be informed by relevant ILD, ED and suitable measures of BEICFs. While there are a variety of integrated scenario approaches, the level of influence of scenario data within these models differs significantly across banks.

253. The scenario process is qualitative by nature and therefore the outputs from a scenario process necessarily contain significant uncertainties. This uncertainty, together with the uncertainty from the other elements, should be reflected in the output of the model producing a range for the capital requirements estimate. Thus, scenario uncertainties provide a mechanism for estimating an appropriate level of conservatism in the choice of the final regulatory capital charge. Because quantifying the uncertainty arising from scenario biases continuous to pose significant challenges, a bank should closely observe the integrity of the modelling process and engage closely with the relevant supervisor.

> 254. Scenario data provides a forward-looking view of potential operational risk exposures. A robust governance framework surrounding the scenario process is essential to ensure the integrity and consistency of the estimates produced. Supervisors will generally observe the following elements in an established scenario framework: (a) A clearly defined and repeatable process; (b) Good quality background preparation of the participants in the scenario generation process; (c) Qualified and experienced facilitators with consistency in the facilitation process; (d) The appropriate representatives of the business, subject matter experts and the corporate operational risk management function as participants involved in the process; (e) A structured process for the selection of data used in developing scenario estimates; (f) High quality documentation which provides clear reasoning and evidence supporting the scenario output; (g) A robust independent challenge process and oversight by the corporate operational risk management function to ensure the appropriateness of scenario estimates; (h) A process that is responsive to changes in both the internal and external environment; and (j) Mechanisms for mitigating biases inherent in scenario processes. Such biases include anchoring, availability and motivational biases.

Business Environment & Internal Control Factors

The fourth element and input in an AMA model are the Business Environment and Internal Control Factors (BEICFs). LDA models mainly depend on historical loss data which are backward looking. Therefore it is necessary with the ability to make continuous qualitative adjustments to the model that reflects ongoing changes in business environment and risk exposure. There are several ways these adjustment can be made and they can vary a lot depending on institution and relevant information collecting process.

> 255. BEICFs are operational risk management indicators that provide forward-looking assessments of business risk factors as well as a bank's internal control environment. However, incorporating BEICFs directly into the capital model poses challenges given the subjectivity and structure of BEICF tools. Banks continue to investigate and refine measures of BEICFs and explore methods for incorporating them into the capital model.
>
> 256. BEICFs are commonly used as an indirect input into the quantification framework and as an ex-post adjustment to model output. Ex-post adjustments serve as an important link between the risk management and risk measurement processes and may result in an increase or decrease in the AMA capital charge at the

group-wide or business-line level. Given the subjective nature of BEICF adjustments, a bank should have clear policy guidelines that limit the magnitude of either positive or negative adjustments. It should also have a policy to handle situations where the adjustments actually exceed these limits based on the current BEICFs. BEICF adjustments should be well-supported and the level of supervisory scrutiny will increase with the size of the adjustment. Over time, the direction and magnitude of adjustments should be compared to ILD, conditions in the business environment and changes in the effectiveness of controls to ensure appropriateness. BEICFs should, at a minimum, be used as an input in the scenario analysis process.

2.2 Risk Measures

Paragraph 667 of the Basel II Framework states that

> *"Given the continuing evolution of analytical approaches for operational risk, the Committee is not specifying the approach or distributional assumptions used to generate the operational risk measure for regulatory capital purposes. However, a bank must be able to demonstrate that its approach captures potentially severe 'tail' loss events. Whatever approach is used, a bank must demonstrate that its operational risk measure meets a soundness standard comparable to that of the internal ratings-based approach for credit risk (i.e. comparable to a one year holding period and a 99.9th percentile confidence interval)."*

2.2.1 Value at Risk

The mathematical definition of Value at Risk at a confidence level $\alpha \in (0,1)$ is given by the smallest number l such that the probability that the loss L exceeds l is at most $(1-\alpha)$. VaR_α equals the α-quantile of the loss distribution[2].

$$\begin{aligned} VaR_\alpha(L) &= \inf\{l \in \mathbb{R} : P(L > l) \leq 1-\alpha\} \\ &= \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\} = F_L^{-1}(\alpha) \end{aligned}$$

Where F_L is a continuous and strictly increasing loss distribution function. Consider an ordered sample independent and identically distributed variables $X_{1,n} \geq \cdots \geq X_{n,n}$. The empirical quantile function is given by

$$F_{n,X}^{-1}(\alpha) = X_{\lfloor n\alpha \rfloor + 1, n}$$

[2] A similar definition can be found in Hult et. al. (2012) with confidence level p defined as $p = 1 - \alpha$

Which is also the empirical Value at Risk. Value at Risk has a big weakness that it ignores the tail beyond α. A risk measure that takes the tail into consideration is the Expected Shortfall. Some also criticize VaR for trying to estimate something that is not scientifically possible to estimate, namely the risks of rare events, and that you are better off with no information at all, than relying on misleading information[3]. For other shortcomings with the risk measure there is a paper published in 2012 by Embrecths et al. called Model uncertainty and VaR aggregation.

2.2.2 Expected Shortfall

Expected shortfall at confidence level α can be seen as the average loss, given that the loss exceeds Value at Risk at confidence level α.

$$ES_\alpha(L) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_t(L)dt \tag{2.1}$$

For a loss L with a continuous loss distribution (2.1) can be rewritten as

$$ES_\alpha(L) = E(L|L \geq VaR_\alpha(L))$$

Expected Shortfall measures what the expected loss is if things get bad, rather than measuring how bad things can get. Other modifications of Expected Shortfall is called Avarage Value at Risk (AVaR), Conditional Value at Risk (CVaR), Tail Value at Risk (TVaR) and Tail Conditional Expectation (TCE) (H. Hult et al., 2012. p. 178).

2.3 Probability Distributions

The continuous probability distribution of a random variable, X, is defined as

$$P(X \leq x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = F(x)$$

where $f(t)$ is the probability density function and $F(x)$ the cumulative distribution function. If the distribution of X is discrete, $f(t)$ is called the probability mass function and $F(x)$ is defined as

$$F(x) = \sum_{t \leq x} f(t)$$

Here follows a couple of distributions that will be used in the thesis. Definitions were taken from M. P. McLaughlin (2001), G. Blom (2005) and the Matlab statistic toolbox documentation[4].

[3] Interview with Nassim Taleb, Derivative Strategy, http://www.derivativesstrategy.com/magazine/archive/1997/1296qa.asp

[4] http://www.mathworks.se/help/stats/index.html

2.3.1 Empirical Distribution

The empirical distribution function is a step function that jumps $1/n$ at each step in a set of n data points. The standard estimator for empirical distribution is defined as

$$\widehat{F_n}(t) = \frac{\#observations \leq t}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{x_i \leq t\} \tag{2.2}$$

where the x_is are the data points in the sample, and by the strong law of large numbers, $\widehat{F_n}(t) \rightarrow F_n$ almost surely as $n \rightarrow \infty$.

Some advantages are that no underlying distribution assumption is needed, thus no parameters needs to be estimated and it is very flexible. Drawbacks are that it is entirely based historical data and cannot generate numbers outside the historical data set. However, I will use a smoothed empirical CDF where the last element of the output quantile is linearly extrapolated to make sure it covers the closed interval $[0, 1]$.

2.3.2 Exponential Distribution

The standard exponential distribution is a one parametric continuous distribution with PDF

$$f(x) = \frac{1}{\mu} exp^{-x/\mu}$$

where μ is the scale parameter and also the expected value.

2.3.3 Truncated Log-normal Distribution

If $Z \sim N(0, 1)$ then $X = e^{\mu+\sigma Z}$ is log-normal distributed. Operational loss data are often truncated at a certain reporting threshold and it is therefore necessary to adjust the log-normal distribution so it takes the truncation threshold into account. The adjustment is done by re-normalizing the PDF so that it sums to one, which yields the following PDF[5]

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left[-\left(\frac{\ln x - \mu}{\sigma\sqrt{2}}\right)^2\right] \frac{1}{1 - F(a)} \mathbb{1}\{x \in (a, \infty)\}$$

where a is the lower truncation point of the data and $F(a)$ is the CDF of X at the truncation point a. To simulate from the truncated log-normal distribution the inverse CDF is used so that a random variate x can be defined as

$$x = F^{-1}(F(a) + U(1 - F(a)))$$

where F^{-1} is the inverse cumulative distribution function and U a uniform random variable on [0,1].

[5] From Dutta & Perry (2007). p. 14.

2.3.4 Weibull Distribution

The Weibull distribution is a two parametric continuous distribution with PDF

$$f(x) = \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} e^{-(x/a)^b}$$

Where a is the scale parameter and b the location parameter.

2.3.5 Log-logistic Distribution

Also known as the Fisk distribution with location parameter μ and scale parameter σ, has the PDF

$$f(x) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{[1+(x/\alpha)^{\beta}]^2}$$

The log-logistic CDF can be expressed on closed form

$$F(x) = \frac{1}{1+(x/\alpha)^{-\beta}}$$

as well as the inverse CDF

$$F^{-1}(p) = \alpha\left(\frac{p}{1-p}\right)^{1/\beta}$$

The log-logistic distribution is similar to the log-normal distribution, but it has heavier tails.

2.3.6 Gamma Distribution

The gamma PDF is

$$f(x) = \frac{1}{b^a\Gamma(a)}x^{a-1}e^{-x/b}$$

where $\Gamma(\cdot)$ is the Gamma function, a the shape parameter and b the scale parameter.

2.3.7 Location-scale Student's t-distribution

The location-scale Student's t-distribution is a symmetric continuous distribution similar in shape to the normal distribution, with the exception that it can produce values that are further from the mean, i.e. it has heavier tails. The PDF is

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu\sigma^2}}\left(1+\frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$$

2.3.8 Poisson Distribution

Poisson distribution is a discrete probability distribution that describes the probability of a given number of events occurring in a fixed time window, i.e. the number of traffic accidents in a year. The probability mass function is defined as

$$f(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

The parameter λ is estimated based on desired frequency. If the Poisson distribution is supposed to generate a yearly frequency, λ is simply the yearly average of the sample.

$$\widehat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

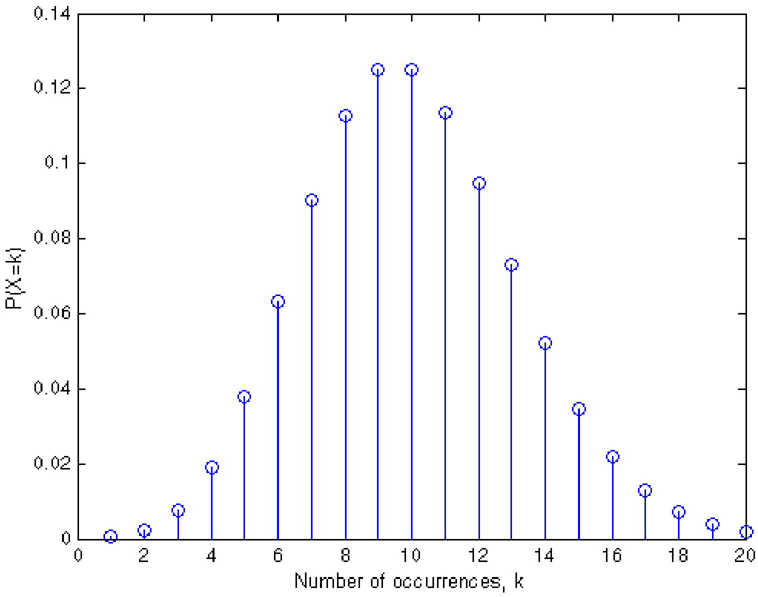


Figure 2.1: A Poisson distribution with $\lambda = 10$.

2.3.9 Negative Binomial Distribution

Is a two parametric discrete probability distribution of the number of successes in a sequence of Bernoulli trials before a specified number r of failures. The sucess probability is denoted p and the PDF is

$$f(k) = \binom{r+k-1}{k} p^r q^k$$

where $q = 1 - p$ and if r is not an integer the binomial coefficient in the expression is replaced by[6]

$$\frac{\Gamma(r+k)}{\Gamma(r)\Gamma(k+1)}$$

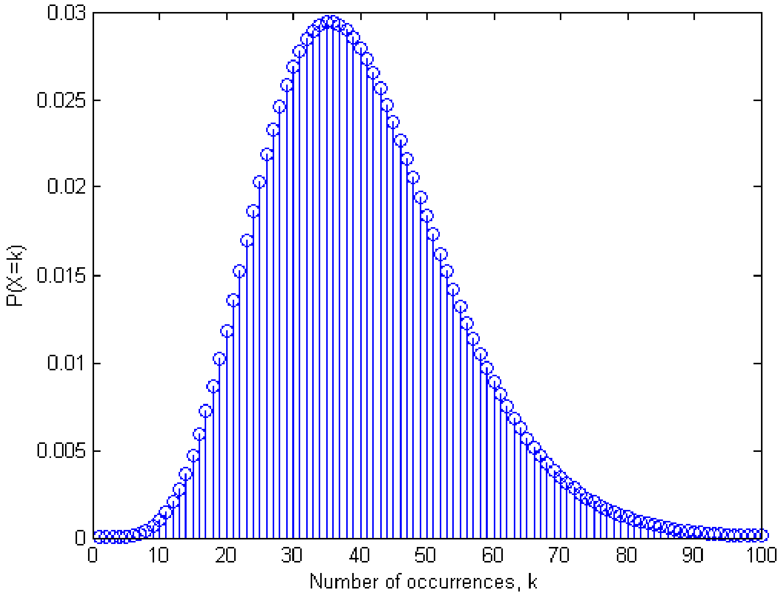


Figure 2.2: A negative binomial distribution with $r = 10$ and $p = 0.2$.

2.4 Peaks Over Threshold

The peaks over threshold (POT) method is part of extreme value theory and is a way to extrapolate the empirical tail outside the range of the sample. It turns out that the distribution of excesses $X_k - u$ over a high threshold u of a sample independent and identically distributed random variables is well approximated by the generalized Pareto distribution, with CDF

$$G_{\gamma,\beta}(x) = 1 - (1 + \gamma x/\beta)^{-1/\gamma}.$$

Here follows a derivation of the estimator of the quantile function (H. Hult et al., 2012). Given a sample $X_1, ..., X_n$ iid with regularly varying right tail

$$\lim_{t\to\infty} \frac{\overline{F}(t\lambda)}{\overline{F}(t)} = \lambda^{\rho}$$

[6] http://www.mathworks.se/help/stats/negative-binomial-distribution.html

The excess distribution function of X over the threshold u is given by

$$F_u(x) = P(X - u \leq x | X > u), \quad x \geq 0.$$

We have

$$\overline{F_u}(x) = \frac{\overline{F}(u+x)}{\overline{F}(u)} = \frac{\overline{F}(u(1+x/u))}{\overline{F}(u)}$$

Since $\overline{F}$ is regularly varying with index $-\rho < 0$ it holds that

$$\frac{\overline{F}(t\lambda)}{\overline{F}(t)} \to \lambda^{\rho}$$

uniformly in $\lambda \geq 1$ as $u \to \infty$

$$\lim_{t \to \infty} \sup_{x > 0} |\overline{F_u}(x) - G_{\gamma,\beta(u)}(x)| = 0$$

where $\gamma = 1/\rho$ and $\beta(u) \sim u/\rho$ as $u \to \infty$. Now let N_u denote the number of exceedences of u by $X_1, ..., X_n$. Recall that

$$\overline{F}(u+x) = \overline{F}(u)\overline{F_u}(x)$$

If u is not too far out in the tail then the empirical approximation $\overline{F}(u) \approx \overline{F_n}(u) = N_n/n$ holds. Moreover the approximation

$$\overline{F_u}(x) \approx \overline{G}_{\gamma,\beta(u)}(x) \approx \overline{G}_{\widehat{\gamma},\widehat{\beta}}(x) = \left(1 + \widehat{\gamma}\frac{x}{\widehat{\beta}}\right)^{-1/\widehat{\gamma}},$$

where $\widehat{\gamma}$ and $\widehat{\beta}$ are the estimated parameters makes sense. We then get the tail estimator

$$\widehat{\overline{F}(u+x)} = \frac{N_u}{n}\left(1 + \widehat{\gamma}\frac{x}{\widehat{\beta}}\right)^{-1/\widehat{\gamma}},$$

and the quantile estimator

$$\widehat{F^{-1}}(p) = u + \frac{\widehat{\beta}}{\widehat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\widehat{\gamma}} - 1\right).$$

One of the key elements in the POT method is to find a suitable high threshold u where the GPD approximation is valid, yet not too high were not enough data are available. One approach is to start with a rather low threshold and examining the GPD goodness-of-fit with the excess data by comparing respective quantiles in a quantile-quantile-plot. Then repeat the procedure with a slightly higher threshold and gradually increase it until the best fit is found.

2.5 Dependence

There are several ways to analyze dependency in the data. There might be dependencies between the number of occurrences or severity impact within a cell and between cells. If the number of losses in a cell are not independent of each other they are not Poisson distributed. In that case, it might be more appropriate to model the frequency with a negative binomial distribution.

Under the standard LDA model, frequency and severity distributions within a cell are assumed independent and the severity samples are independent and identically distributed. The high quantile of the total annual loss distribution is computed by simply adding together the high quantiles of the loss distribution for each BL/ET cell. Summing the quantiles implies a perfect correlation among the cells. To avoid this assumption, typically three methods are considered to implement dependence in the model (E. Cope and G. Antonini, 2008)

1. The frequency distribution between cells are dependent

2. The severities between cells are dependent

3. The aggregated loss between cells are dependent

It is often difficult to define dependencies between severities. One approach is to introduce common "shock models" affecting several cells simultaneously (Lindskog and McNeil, 2003). Dependencies of the aggregated loss distributions between different cells in the ORX database have been analyzed by E. Cope and G. Antonin (2008) with the following conclusion:

> *" (...) most of the correlations among quarterly aggregate losses are low, generally less than 0.2, and rarely exceeding 0.4. Moreover, the correlation structures of individual banks appear to be largely homogeneous. A formal statistical test for the equality of correlation matrices indicated that the majority of individual banks' correlation matrices were found to be statistically equal to the average correlation matrix. Therefore, the average correlation matrix is representative of the correlation of most ORX members. "*

Model dependency between random frequency variables can be done by sorting the number of loss occurrences in buckets of periods and estimate correlations such as Pearson's correlation coefficient and Kendall's rank correlation. The estimated correlation can be used to specify the correlation matrix for different copulas which in turn can describe the dependence structure between the random variables. The most commonly used copula in AMA is the Gaussian copula. (Basel Committee, 2009. p. 50) [3].

2.5.1 Copula

A copula is a sort of distribution function that is used to describe the dependence structure between random variables. The d-dimensional Gaussian copula G_R^{Ga} is defined as (H. Hult et al., 2012. p. 303.)

$$G_R^{Ga}(u) = P(\Phi(X_1) \leq u_1, ..., \Phi(X_d) \leq u_d) = \Phi_R^d(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)),$$

where Φ_R^d is the distribution function of X, R is the linear correlation matrix and Φ the standard normal distribution function. Let $F_1, ..., F_n$ be frequency distributions and C a copula. Then $C(F_1(x_1), ..., F_n(x_n))$ specifies the n-variate frequency distribution.

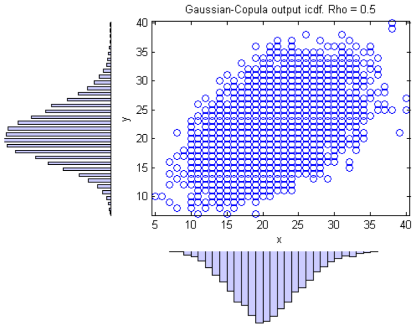


Figure 2.3: Example of Gaussian copula with linear correlation 0.5 and marginal distributions transformed to Poisson.

The d-dimensional Student's t copula $C_{\nu,R}^t$ is defined (H. Hult et al., 2012. p. 303.)

$$C_{\nu,R}^t(u) = P(t_\nu(X_1) \leq u_1, ..., t_\nu(X_d) \leq u_d) = t_{\nu,R}^d(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d))$$

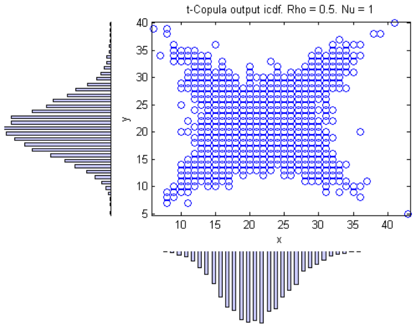


Figure 2.4: Example of Student's t copula with linear correlation 0.5 and marginal distributions transformed to Poisson.

2.6 Scenario Based Approach

Unlike traditional techniques, scenario analysis uses expert opinion as input rather than historical data. A common approach is to calibrate an underlying distribution, often a log-normal distribution for severity and Poisson for frequency, based on expert judgment estimates on frequency, most frequent loss and worst case loss of an operational risk.

There are many ways to interpret the worst case scenario, including a fixed high quantile, worst single loss in a period and a quantile of severity distribution with a probability level depending on average frequency (A. Colombo and S. Desando, 2008). Scenario analysis typically includes some degree of bias and subjectivity. Biases in scenario analysis development processes can include overconfidence, motivational bias, availability bias, partition dependence, and anchoring (Federal Reserve System, 2011)[7].

Another issue is that the implied severity distribution is often assumed log-normal distributed without further ado. If the implied severity distribution is assumed log-logistic distributed, the worst case (WC) loss is regarded as the p-quantile[8] and typical loss (M) as the median (50% quantile) the relationship between these inputs and the distribution parameters can easily be obtained using the quantile function. The scale parameter α can be expressed as:

$$\begin{aligned} M &= F^{-1}(0.5) = \alpha\left(\frac{0.5}{1-0.5}\right)^{1/\beta} \\ \Rightarrow & \quad \alpha = M \end{aligned}$$

And the shape parameter β as follows

$$\begin{aligned} WC &= F^{-1}(p) = \alpha\left(\frac{p}{1-p}\right)^{1/\beta} \\ \Rightarrow & \quad (WC)^{\beta} = M^{\beta}\frac{p}{1-p} \\ \Rightarrow & \quad \frac{p}{1-p} = \left(\frac{WC}{M}\right)^{\beta} \\ \Rightarrow & \quad \ln\frac{p}{1-p} = \beta\ln\frac{WC}{M} \\ \Rightarrow & \quad \beta = \frac{\ln\frac{p}{1-p}}{\ln\frac{WC}{M}} \end{aligned}$$

And if the underlying distribution is assumed log-normal, most occuring loss

[7] http://www.occ.gov/news-issuances/bulletins/2011/bulletin-2011-21a.pdf

[8] Typically a very high quantile, such as $p = 0.99$

as the mode and worst case the p-quantile we get the following relationship:

$$\begin{aligned} M &= mode(X) = e^{\mu-\sigma^2}, \quad X \sim LN(\mu,\sigma^2) \\ \Rightarrow & \quad \mu = \ln M + \sigma^2 \end{aligned}$$

And the standard deviation σ

$$\begin{aligned} WC &= e^{\mu+\sigma\Phi^{-1}(p)} \\ \Rightarrow & \quad \ln(WC) = \ln(M) + \sigma^2 + \sigma\Phi^{-1}(p) \\ \Rightarrow & \quad (\sigma + \frac{1}{2}\Phi^{-1}(p))^2 - \frac{1}{4}(\Phi^{-1}(p))^2 = \ln WC - \ln M \\ \Rightarrow & \quad \sigma + \frac{1}{2}\Phi^{-1}(p) = \sqrt{\ln WC - \ln M + \frac{1}{4}(\Phi^{-1}(p))^2} \\ \Rightarrow & \quad \sigma = -\frac{1}{2}\Phi^{-1}(p) + \sqrt{\ln WC - \ln M + \frac{1}{4}(\Phi^{-1}(p))^2} \end{aligned}$$

where $\Phi^{-1}(p)$ is the inverse standard normal CDF (or quantile function).

2.7 Loss Distribution Approach

A general model for measuring operational risks is the loss distribution approach. It arises if the number N of internal operational losses and the severity X of a single loss are assumed to be independent random variables. One can interpret it as two different dice which are thrown one after another. The first die represents the number of operational losses and the second die the single loss severity. If the first die is thrown and shows the number n the second die has to be thrown n times. If x_i is the result of throw i, the respective loss is given by

$$l = x_1 + ... + x_n = \sum_{i=1}^{n} x_i$$

The random total loss L before throwing any dice is given by

$$L = X_1 + ... + X_N = \sum_{i=1}^{N} X_i$$

and the aggregated loss distribution function $G(x)$ is then given by

$$G(x) = P(L \leq x).$$

This model is called the Standard Loss Distribution Approach (LDA) and requires the determination of the probability distribution of the the frequency of operational loss events and the conditional probability distribution of the severity of operational losses given an operational loss event.

2.8 Monte Carlo Method

Since the loss distribution can not be represented on analytical form, one of the most common approaches to estimate it is by using Monte Carlo simulations. A number of losses are simulated and aggregated from the underlying frequency and severity distributions to form a potential one year loss. This step is then repeated a large number of times, say n times, resulting in an empirical loss distribution, which is a reasonably good approximation of the "true" loss distribution. The Value at Risk with confidence level α, also known as the α-quantile, can then simply be read from the distribution as the $n - \lfloor n \cdot \alpha \rfloor$ largest value.

2.8.1 Remarks

One problem with the Monte Carlo method is that the underlying distributions have to be known over the whole possible region. Since the severity and frequency distributions are fitted to historical data, which often contains few data points and data from different sources (internal and external etc.), Monte Carlo simulations may yield unstable in estimates of the high quantiles of heavy tailed distributions.

2.9 Analytical Approximation

When loss data are heavy-tailed, or more precisely subexponential, there exists a simple closed-form approximation of the high quantile shown by Klüppelberg and Böcker (2006). A non-negative random variable X belongs to the subexponential distribution family if

$$\lim_{x\to\infty} \frac{P(X_1 + ... + X_n > x)}{P(X > x)} = n, \quad n \geq 2$$

The subexponentiality implies that the sum of operational losses are most likely large due to a single large loss rather than several smaller losses. The aggregated loss distribution $G(x)$ can be written as

$$\begin{aligned} G(x) &= P(L \leq x) \\ &= \sum_{n=0}^{\infty} P(N = n) P(L \leq x | N = n) \\ &= \sum_{n=0}^{\infty} P(N = n) P(X_1 + ... X_n \leq x) \\ &= \sum_{n=0}^{\infty} P(N = n) F_n(x) \end{aligned}$$

where F is the distribution function of X and F_n is the n-fold convolution of F. N is the random frequency variable. Assuming that the severities $X_1, \ldots, X_n$ are subexponential with distribution function F, and assume that for some $\epsilon > 0$

$$\sum_{n=0}^{\infty} (1+\epsilon)^n p(n) < \infty$$

Then, the aggregated loss distribution $G(x)$ is subexponential with tail behavior given by

$$\overline{G}(x) \sim E(N)\overline{F}(x), \quad x \to \infty \tag{2.3}$$

Where $E(N)$ is the expected frequency and $\overline{G}(x) = 1 - G(x)$ and $\overline{F}(x) = 1 - F(x)$ (Klüppelberg and Böcker, 2006. Theorem 1.3.5 in appendix). Hence, according to (2.3) we get

$$G(x) \approx 1 - E(N)(1 - F(x))$$

The value at risk at confidence α of the aggregated loss distribution $G(x)$ is defined the α-quantile of the aggregated loss distribution

$$VaR_\alpha(X) = G^{-1}(\alpha)$$

and consequently

$$\begin{aligned} \alpha &= G(VaR_\alpha(X)) \approx 1 - E(N)(1 - F(VaR_\alpha(X)) \\ \Rightarrow & \quad F(VaR_\alpha(X)) \approx 1 - \frac{1-\alpha}{E(N)}, \quad \alpha \text{ close to } 1 \end{aligned}$$

Which leads to

$$VaR_\alpha(X) \approx F^{-1}\left(1 - \frac{1-\alpha}{E[N]}\right), \quad \alpha \text{ close to } 1 \tag{2.4}$$

What is interesting about the result is that according to this approximation, operational value at risk only depends on the tail of the severity distribution, and ignores the body. Furthermore, since the expected frequency is sufficient, calibrating some counting process, such as Poisson or negative binomial, is not necessary. Just the sample mean is needed.

Recall the estimator of the generalized Pareto quantile explained in the peaks over threshold section.

$$\widehat{F^{-1}}(p) = u + \frac{\widehat{\beta}}{\widehat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\widehat{\gamma}} - 1\right)$$

Since the generalized Pareto distribution is a subexponential distribution for $\gamma > 0$ we may use (2.4) to obtain the analytical Value at Risk for the generalized Pareto model. We get

$$\begin{aligned} VaR_\alpha(X) &\approx \widehat{F^{-1}}\left(1-\frac{1-\alpha}{E[N]}\right) \\ &= u+\frac{\widehat{\beta}}{\widehat{\gamma}}\left(\left(\frac{n}{N_u}\left(\frac{1-\alpha}{E[N]}\right)\right)^{-\widehat{\gamma}}-1\right) \end{aligned}$$

A comparison of the analytical VaR and monte carlo simulated VaR is shown in Figure 2.5.

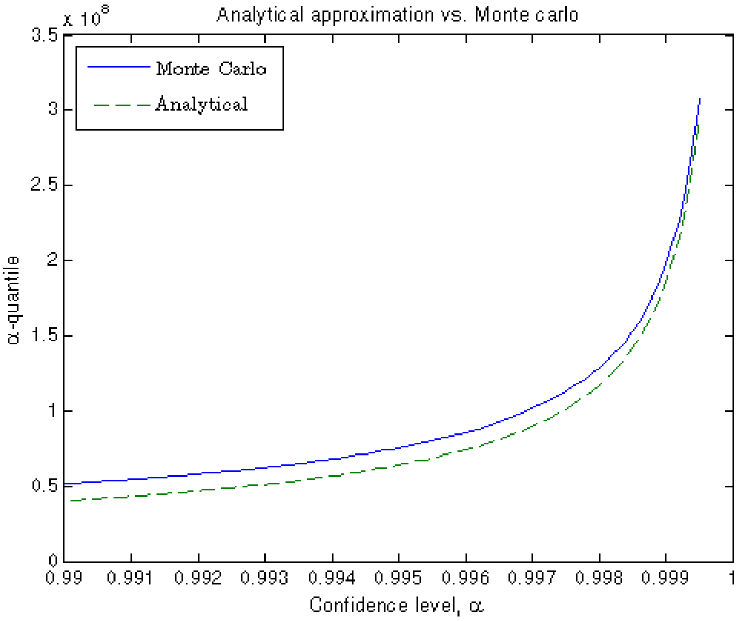


Figure 2.5: Comparison of analytical VaR with monte carlo simulated VaR based on a Poisson process and a generalized Pareto severity distribution.

Chapter 3

Data

The operational loss data used in this thesis consist of internally collected data by the bank and external data privately released by a consortium of institutions. This section will present some characteristics of the data, an exploratory data analysis and also some challenges in measuring operational risk from loss data. Note that all losses have been multiplied by a constant to anonymize the data.

3.1 Exploratory Data Analysis

EDA is an approach promoted by John Tukey to analyze data sets, often visually, in order to assess assumptions on which hypotheses and appropriate statistical models can be selected.[1] The following EDA was performed on all internal and external data available.

Min	25%	50%	75%	Max
0	590.87	2469.46	9073.85	41544970.92

Table 3.1: Min/max and quartilels of modified internal data

These five number statistics are functions of the empirical distribution and are defined for all distributions unlike moments such as the mean and standard deviation. Nevertheless, the moments are also computed and displayed in Table 3.2 below.

Mean	Std. deviation	Skewness	Kurtosis
53883.26	868820.96	42.68	1995.02

Table 3.2: Min/max and quartiles of modified internal data

[1] In 1977 John Tukey published the book Exploratory Data Analysis.

The sample shows a significant difference in the mean and median which may be caused by a heavy right tail. The high kurtosis indicates that the high standard deviation is a result of extreme observations, far from the sample's mean.

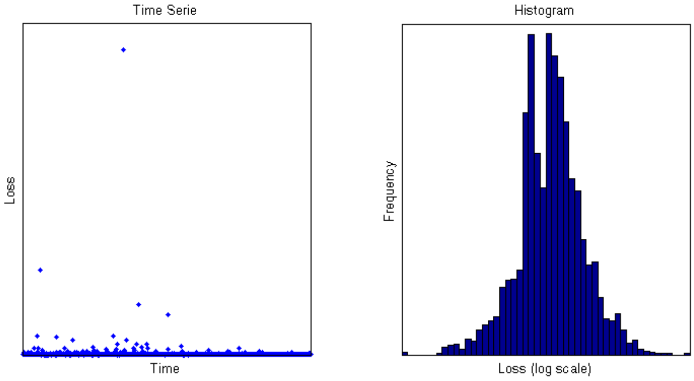


Figure 3.1: Time series and histogram of internal loss data

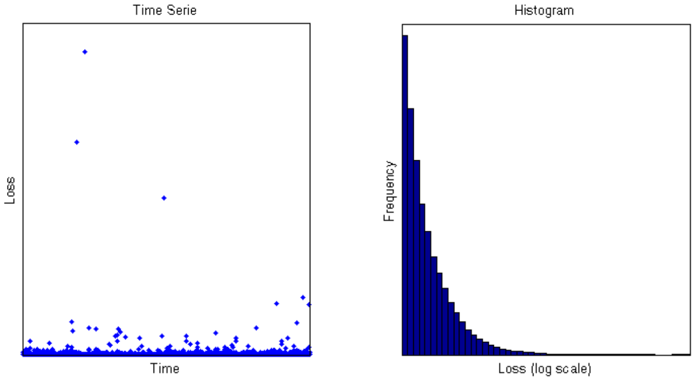


Figure 3.2: Time series and histogram of external loss data. Note that the external data are truncated at a certain threshold.

The data show that there are two types of losses, the high frequency low severity losses and the low frequency high severity losses. It is possible that the tail and the body do not necessarily belong to the same distribution. In Figure 3.3 the empirical quantiles are compared to the quantiles of some reference distributions in a QQ-plot to determine whether the data have heavier or lighter tails than the reference distributions.

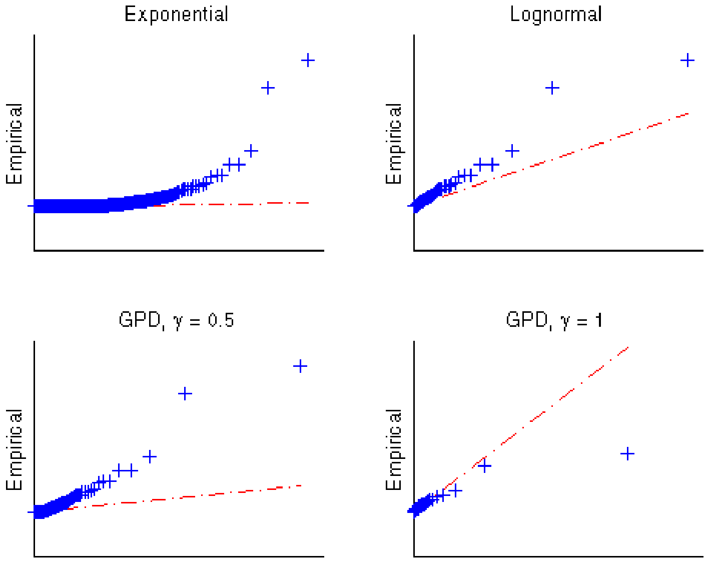


Figure 3.3: QQ-plots of loss data against selected reference distributions. In the top left plot we see that the curve deviates up from the dashed line, indicating that the data are heavier than exponential. Top right indicates that the data are heavier than log-normal, though it is a better fit than exponential. Lower plots suggest that if the data are fitted to a generalized Pareto distribution, the shape parameter γ should lie between 0.5 and 1. We have indeed heavy tailed data.

3.2 Characteristics of Data

Heavy tailed data can be observed in nearly every category of the external operational losses, regardless of event type or business line. There are some implications working with heavy tailed data which E. W. Cope et. al. (2009) sum up in three pieces, together with some examples from H. Hult. et. al. (2012).

1. Instability of estimates - A single observation can have a drastic impact on the estimated variables, even when there is a large underlying data set.

2. Dominance of sums - Annual total operational losses for a cell will typically be driven by the most extreme losses. From H. Hult. et. al. (2012) p. 257 we get that for subexponential variables $X_1, \ldots, X_n$

$$\lim_{x\to\infty} \frac{P(\max(X_1, \ldots, X_n) > x)}{P(X_1 + \cdots + X_n > x)} = 1.$$

The interpretation is that the sum takes a very large value due to precisely one of the terms taking a very large value and the sum of the remaining terms being small.

3. Dominance of mixtures - If losses that are generated in two different cells are pooled together, the tail distribution of the total losses will generally follow the distribution of the cell with the heavier tailed distribution of the two of them. Let X and Y be losses from two different business lines. Suppose that X has a distribution function with a regularly varying right tail and $|Y|$ has a finite order of moments. This gives the expression

$$\lim_{x\to\infty} \frac{P(X+Y>x)}{P(X>x)} = 1,$$

which shows that only the loss variable with the heaviest right tail matters for probabilities of very large losses. (H. Hult. et. al., 2012. p. 261)

3.3 Data Sufficiency

Recall the standard empirical estimator (2.2) from chapter 2

$$\widehat{F_n}(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{x_i \leq t\}$$

and notice that the sum is Bin$(n, F(t))$-distributed. The relative error is the standard deviation of the estimator divided by the estimated quantity, i.e.

$$\frac{Var(\widehat{F_n}(t))^{1/2}}{F(t)} = n^{-1/2}\left(\frac{1}{F(t)} - 1\right)^{1/2}$$

A natural requirement is that the standard deviation of the estimator must not be greater than the probability to be estimated. Under this requirement $n \approx 1/F(t)$, which is very large due to $F(t)$ being very small (H. Hult, 2012. p. 203).

3.4 Scaling & Weighting

Integrating external data in the internal loss database requires an appropriate scaling technique. Most research propose some non-linear relationship between a loss and firm size, i.e. gross income or equity. The only thing that can be associated with a loss in the consortium database is the region where the loss occurred and the size of the loss. Therefore any scaling method where firm size is involved can not be carried through. More on scaling can be found in E. Cope and A. Labbi (2008).

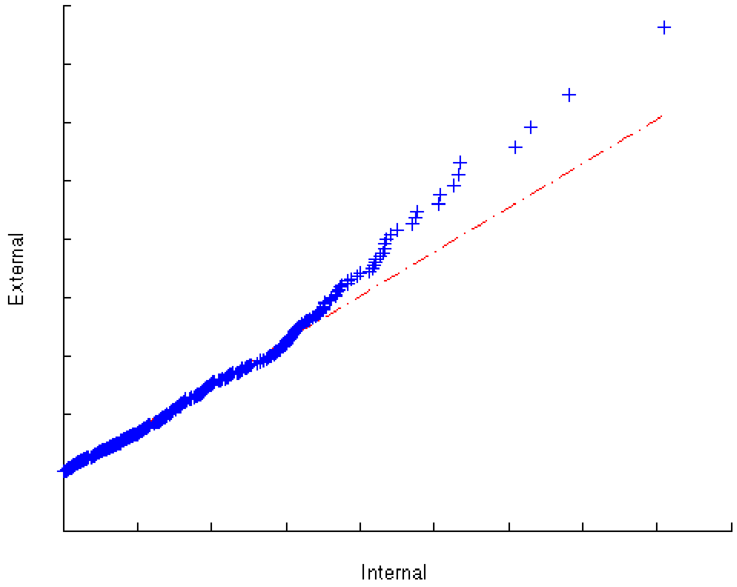


Figure 3.4: QQ-plot of external vs. internal loss data

We see that the external data have a heavier right tail than the internal data, so simply mixing the two data sets will most likely lead to an overestimation of capital. However, mixing the data may yield more realistic results by applying a weighted sum to the internal and external distribution functions:

$$F(x) = w_{int}F_{int}(x) + w_{ext}F_{ext}(x), \quad w_{int} + w_{ext} = 1.$$

Chapter 4

Method

In this chapter different methods and approaches based on the theory will be evaluated and culminate in a model, which will be described in more detail. Assumptions and methods are judged according to different performance measures proposed by Dutta and Perry, 2007.

- Good Fit - Statistically, how well does the method fit the data?
- Realistic - If a method fits well in a statistical sense, does it generate a loss distribution with a realistic capital estimate?
- Well-Specified - Are the characteristics of the fitted data similar to the loss data and logically consistent?
- Flexible - How well is the method able to reasonably accommodate a wide variety of empirical loss data?
- Simple - Is the method easy to apply in practice, and is it easy to generate random numbers for the purposes of loss simulation?

4.1 Modeling Frequency

The frequency plays an important role in the estimation of capital. The two most common discrete parametric distributions for modeling frequency of operational losses are the Poisson distribution and the negative binomial distribution[1].

To determine which of the two frequency distribution proposed meet the different performance measures best, a quick analysis is carried through. If the mean and the variance is of the same magnitude, the Poisson distribution is an appropriate choice. A variance that is higher than the sample mean, suggests a negative binomial distribution. Only internal data from the past

[1] Observed range of practice in key elements of Advanced Measurement Approaches. Basel Committee, 2009

five years are considered relevant and are used in the calibration. No external data are being used here since internal data reflects the banks loss profile more accurately and calibrating frequency parameters requires fewer data points.

Generally, we notice a higher number of occurrences in recent years compared to past years, which may interfere in the parameter estimation. This difference is most likely due to an increase in reporting frequency and not in the actual loss frequency. An example of parameter estimates and goodness-of-fit tests are presented in the table below.

Distribution	Poisson	N. binomial
Parameter 1	5.6	5.9
Parameter 2	-	0.51
χ^2-test	0.249	0.415
K-S test	0.551	0.640
Empirical variance	11.3	11.3

Table 4.1: Estimated parameters of frequency distributions and the p-values of Kolmogorov-Smirnov and χ^2 goodness-of-fit tests. Parameter 1 corresponds to $\widehat{\lambda}$ for Poisson and $\widehat{r}$ for negative binomial. Paramater 2 corresponds only to the $\widehat{p}$ parameter for negative binomial, since Poisson only have one parameter.

Based on the Kolmogorov-Smirnov and χ^2 goodness-of-fit tests none of the null hypotheses at significance level 5% are rejected. Even though the variance is higher than the sample mean the primary choice of frequency distribution is, due to its flexibility and simplicity, the Poisson distribution. Also, the analytical approximation of the high-end quantile of the loss distribution revealed that the actual choice of frequency distribution is less important.

4.2 Modeling Severity

Severity is much harder to model than frequency and requires a lot more data points. The choice of how to model severity has usually a much higher impact on the capital than the choice of how to model the frequency. All severities are assumed independent and identically distributed. Three different modeling techniques are performed.

4.2.1 Method 1

First, commonly used parametric distributions such as the log-normal, weibull and log-logistic distributions are fitted to the whole data set. Also a location-scale student's t distribution is fit to log losses, using both maximum likelihood (MLE) and least squares (LSE) to estimate the parameters. LSE is used due to the log-likelhood function typically beging rather flat and small changes in the data can result in big changes in the parameter estimates, as explained by H. Hult et al., (2012). Figure 4.1 contains the resulting QQ-plots and Figure 4.2 contains the corresponding histograms with the imposed density function. Plots are in log scale. Parameters are shown in Table 4.2

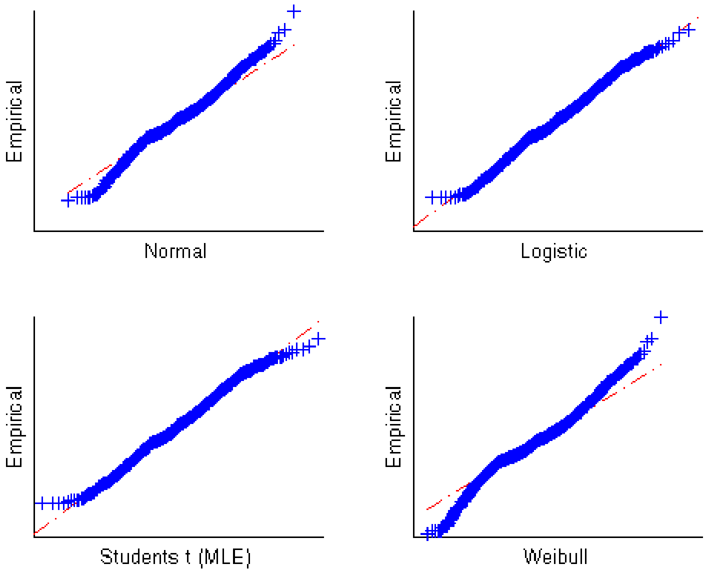


Figure 4.1: Empirical quantiles compared with parametric distributions fitted to data. Upper plots: Normal left, logistic right. Lower plot: Student's t left, Weibull right.

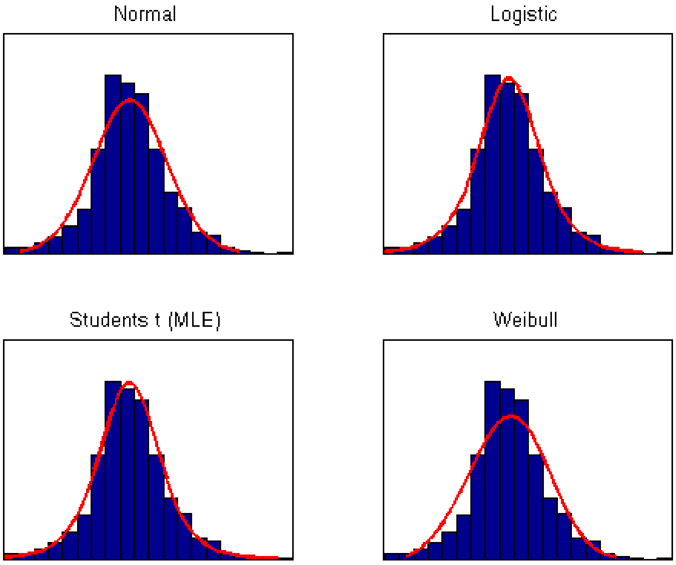


Figure 4.2: Histogram of log losses with imposed density function. Upper left: Normal distribution fitted to log losses. Upper right: Logistic distribution fitted to log losses. Lower left: Student's t distribution fitted to log losses. Lower right: Log transformed Weibull fitted to losses.

Distribution	Log-normal	Weibull	Log-logistic	Student's t (MLE)	Student's t (LSE)
Parameter 1	7.68737	6704.08	2144.73	7.67062	7.68737
Parameter 2	2.24207	0.421686	0.81873	1.83439	1.91251
Parameter 3	-	-	-	5.74174	7.11015
$\widehat{F^{-1}}(0.999)$	$2.23 \cdot 10^6$	$6.56 \cdot 10^5$	$9.88 \cdot 10^6$	$3.93 \cdot 10^7$	$1.92 \cdot 10^7$

Table 4.2: Estimated parameters of the distributions fitted to the whole data set

At first glance, they all seem to fit the data pretty well. Looking closer at the high quantile, which is the main driver of capital, we see that Weibull, as expected, has the lowest severity and Student's t the highest. Monte Carlo simulations on cell level[2] yields that the Student's t distribution gives unrealistically high capital estimates and is therefore discarded. Likewise is the Weibull distribution discarded due to very low capital estimates. What is left is the log-normal distribution which gives a capital estimate lower than the current capital, and log-logistic distribution which has the best goodness-of-fit.

To verify how well the log-normal distribution fits operational loss data, a truncated log-normal distribution is fitted to the external data, which contains more extreme observations than the internal loss data. The probability

[2] Parameter estimates on cell level are presented in Results.

distribution is defined as the usual PDF renormalized with the CDF evaluated at the truncation x_{Trunc}.

$$\frac{1}{1 - F(x_{Trunc}; \mu, \sigma)} f(x; \mu, \sigma)$$

and then the parameters $\widehat{\mu} = 1.682$ and $\widehat{\sigma} = 0.323$ are estimated using maximum likelihood.

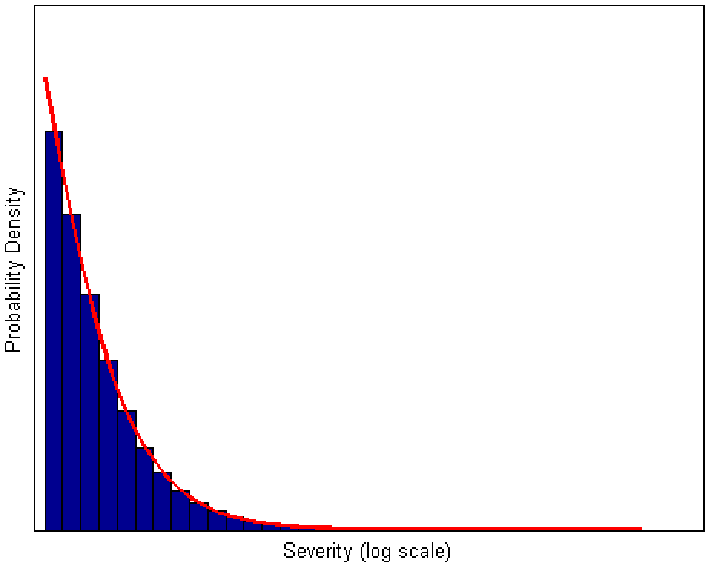


Figure 4.3: Histogram of external log data with imposed truncated normal distribution.

Now, generating numbers from the truncated distribution can be done using the estimated inverse CDF

$$x = \widehat{F^{-1}}(\widehat{F}(x_{Trunc}) + U \cdot (1 - \widehat{F}(x_{Trunc})))$$

The problem here is that $\widehat{F}(x_{Trunc}) \approx 1$, which means that the truncated normal fit with the external data are so far in the right tail that it practically makes it impossible to simulate from the distribution. Compared to the relative good log-normal fit with internal data signals that the tail indeed has different characteristics than the body and as a consequence, finding one distribution that fits well over the entire range is hard.

4.2.2 Method 2

Difficulties in fitting one distribution to the entire range leads us in to the second severity modeling approach. Namely, using a mixed model with different body and tail distributions. First, a couple of commonly used distributions are evaluated, as seen in Figure 4.4

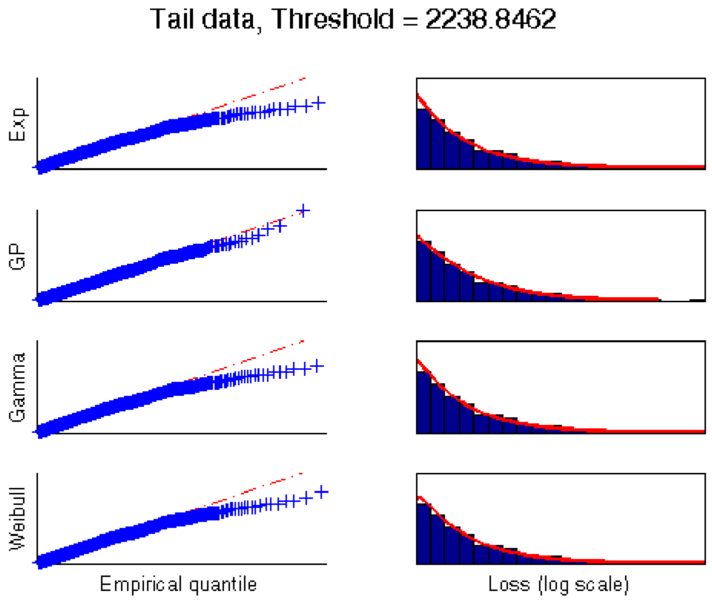


Figure 4.4: Left: QQ-plots of empirical quantiles versus quantiles of parametric distributions fitted to data above the threshold. Right: Corresponding histograms with imposed probability density function in log scale. Distributions are from top to bottom: Exponential, Generalizde pareto, Gamma and Weibull.

The generalized Pareto distribution has the best fit, thus the peaks over threshold method is applied to extrapolate losses beyond the historical data. However, fitting a different generalized Pareto distribution for each cell can not be done with internal data only, and using unscaled external data yields too high estimates of the tail. However, with the assumption that the extreme tail of all severity distributions across all cells have something in common, all internal losses are combined to provide better information on extreme tail of the severity distribution. Starting at a low threshold, which is gradually increased until a good fit is achieved at about the top 6% of all losses, giving the estimates

$$\left(\widehat{\gamma}, \widehat{\beta}\right) \approx (0.79568, 136866)$$

Theory also predicts the shape parameter to stabilize as the threshold becomes large. However, as seen in Figure 4.5 no stabilization can be noted, but the QQ-plot in Figure 4.6 indicates a rather good fit.

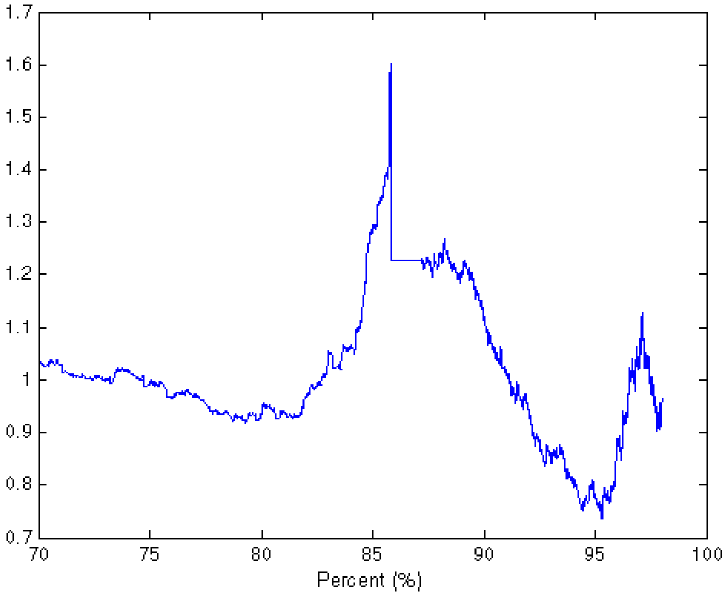


Figure 4.5: Estimated shape parameter γ as a function of threshold in per cent, where a generalized Pareto distribution has been fit to the data above various thresholds.

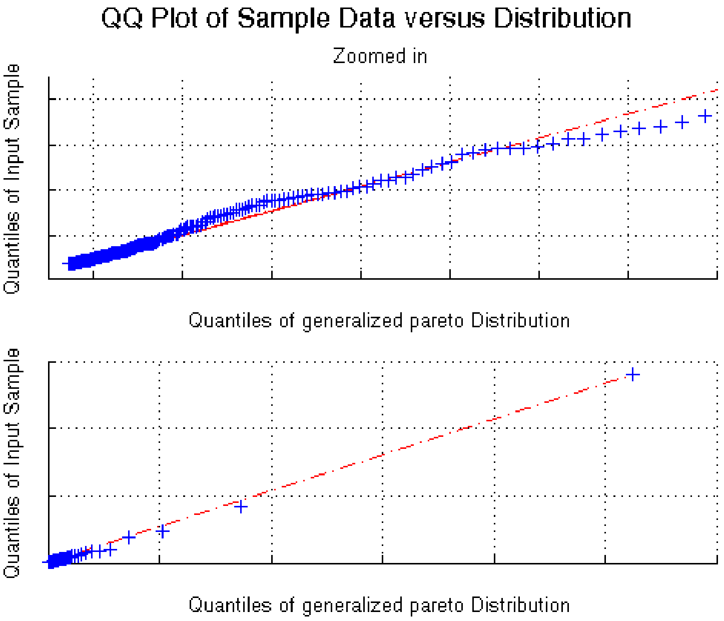


Figure 4.6: QQ-plot of loss data above a threshold versus fitted generalized Pareto distribution.

Below the POT threshold we model the body of the distribution using the weighted sum of the smoothed empirical distribution function of internal and external data. Losses below the external reporting threshold does not affect capital estimate and are not in the model. The frequency is adjusted by only counting the number of internal losses above the external reporting threshold. The resulting mixed severity distribution is described below and can be seen in Figure 4.7.

$$F(x) = \begin{cases} F_1(x) & x < u \\ F_1(u) + (1 - F_1(u))F_2(x) & u \leq x \end{cases}$$

where F_1 is the weighted empirical distribution of the body and F_2 is the generalized Pareto distribution of the tail.

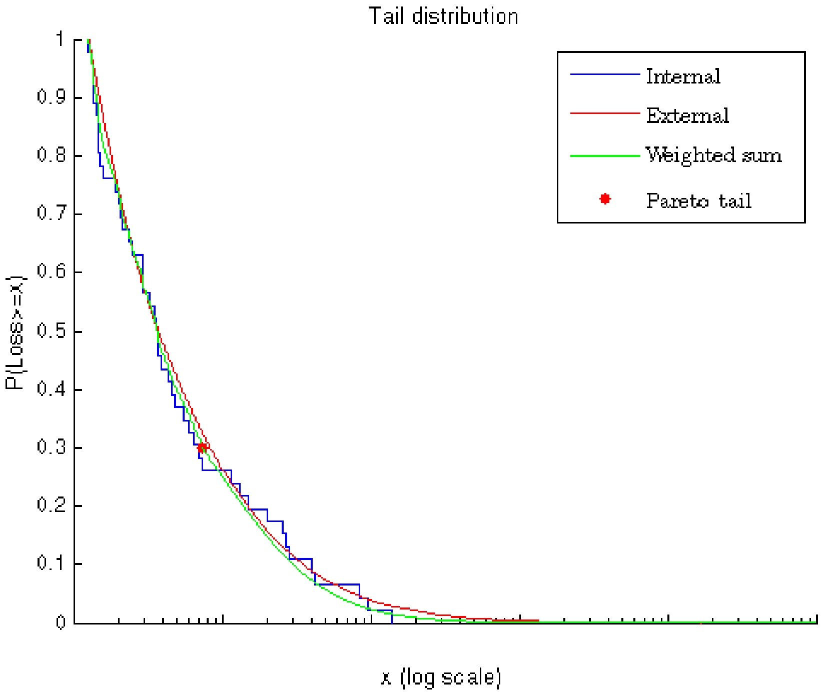


Figure 4.7: The resulting mixed severity distribution is the green line consisting of the smoothed empirical weighted sum body and generalized Pareto tail. The higher quantile of the piecewise distribution function is $\widehat{F^{-1}}(0.999) = 4.37 \cdot 10^6$.

4.2.3 Method 3

The third method is to not make any distributional assumptions at all. The severity distribution is defined as the weighted sum applied on the interpolated empirical cumulative distribution functions for the internal and external data over the whole region, including the tail. Since the external data have more observations in the right tail, beyond the range of the internal data, the external empirical distribution works rather good as a substitute to extrapolating from a parametric distribution.

$$F(x) = w_{int}F_{int}(x) + w_{ext}F_{ext}(x)$$

where $w_{int} + w_{ext} = 1$. A sample of the distribution in one cell can be seen in Figure 4.8

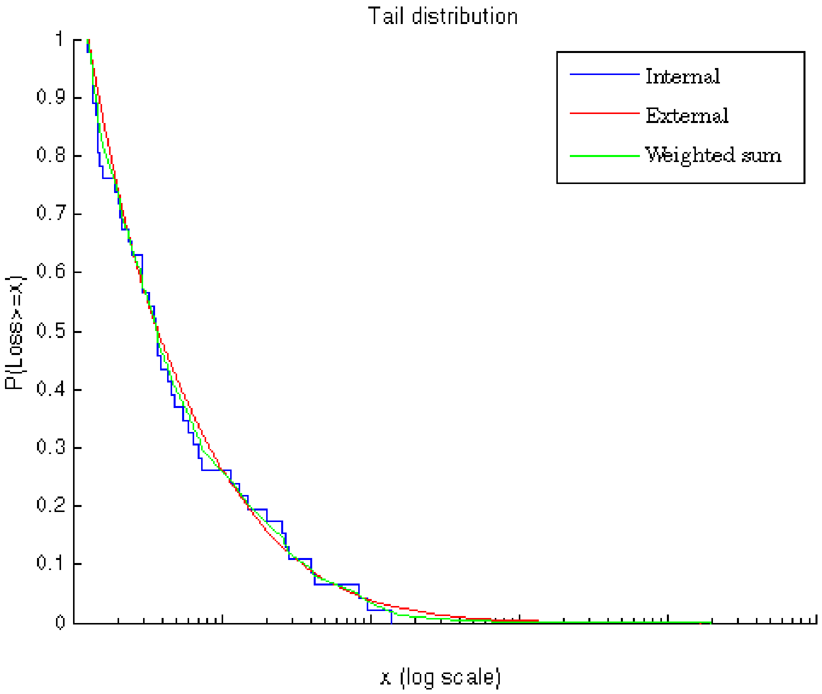


Figure 4.8: Empirical tail distributions of internal and external data, and the resulting weighted sum of in green. The higher quantile of the resulting weighted empirical distribution function is $\widehat{F^{-1}}(0.999) = 5.87 \cdot 10^6$.

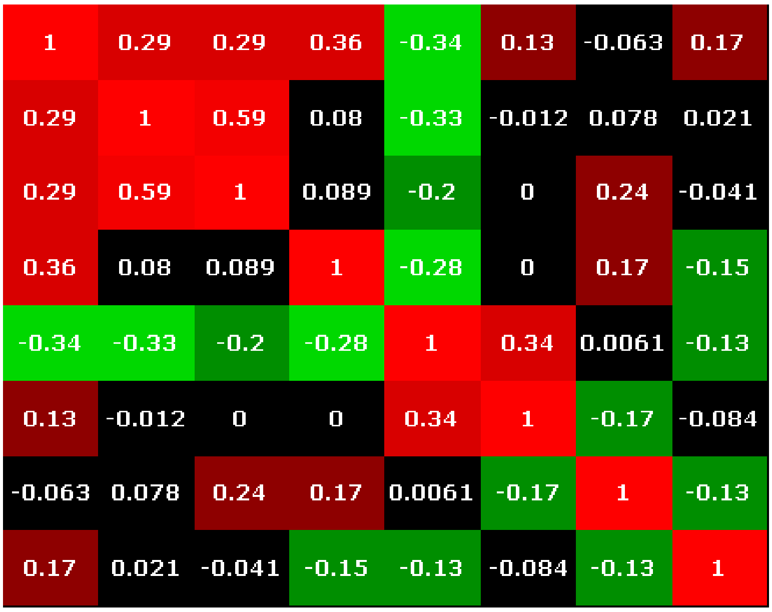
4.3 Scenario Analysis

Several of the cells have no data, or very few data points. Cells with no recorded losses in the entire internal loss record are considered to not be exposed to any risk and is therefore left out of the model. Cells with too few losses to make any statistical analysis are being replaced by scenario analysis. Frequency is still estimated based on the few internal losses we have. Even if there is only one recorded loss in the five year data span, the frequency will simply be estimated as once every five year. I will propose a way to estimate worst case (WC) scenario and typical loss (M) from external loss data by using a pivot cell. The assumption that the ratio between different cells quantiles and the pivot cell in the external data are the same as the ratio between the internal pivot cell and the other internal data cell quantiles makes it possible to estimate WC and M for cells with little or no internal data.

The internal cell with most data points is selected as the pivot. The same external cell is selected and the 50% and 99% quantiles are computed representing the typical loss and the worst case loss. Now the same quantiles for every external cell is computed and divided by the 50% and 99% quantiles in the external pivot cell to get the external quantile ratios. Now the implied internal quantiles of an internal cell with few data points can be computed by taking the internal pivot cell quantiles and multiply it with the corresponding external cells quantile ratio. These quantiles are then used to estimate either log-normal parameters or log-logistic parameters as described in Chapter 2. With the selected severity distribution and frequency distribution the aggregated loss distribution can be generated and contribute to the estimated capital. Since scenarios are only used where there are very few internal losses, it will naturally have a relatively small impact on the total capital estimate, so to simplify, the whole severity is modeled by one distribution. WC, M and the frequency can of course be changed based on expert opinions.

4.4 Dependency & Correlations

The number of internal losses for each business line and event type are sorted into yearly and quarterly buckets. Estimating correlations are hard. Especially with the scarce data that operational losses consist of. Since we use five years of data, we only have five buckets of annual data for each business line and event type. Even though it is the yearly correlation we are after, sorting them on a quarterly basis seems more reliable. However even though there are more buckets of data, each bucket will contain less number of occurrences and it will be more sensitive on errors in the loss reporting date. The quarterly estimated Pearson's correlation coefficients and Kendall tau rank correlation coefficient between business lines are estimated, of which the latter is presented in Figure 4.9



1	0.29	0.29	0.36	-0.34	0.13	-0.063	0.17
0.29	1	0.59	0.08	-0.33	-0.012	0.078	0.021
0.29	0.59	1	0.089	-0.2	0	0.24	-0.041
0.36	0.08	0.089	1	-0.28	0	0.17	-0.15
-0.34	-0.33	-0.2	-0.28	1	0.34	0.0061	-0.13
0.13	-0.012	0	0	0.34	1	-0.17	-0.084
-0.063	0.078	0.24	0.17	0.0061	-0.17	1	-0.13
0.17	0.021	-0.041	-0.15	-0.13	-0.084	-0.13	1

Figure 4.9: Kendall τ rank correlation coefficient of quarterly frequency between the eight business lines.

The observed correlations are rather low as expected, rarely exceeding 0.3. The correlation matrix could be used to specify a copula in order to model frequency dependency between business lines.

4.5 Aggregated Loss Distribution

Samples from the loss distributions are generated using Monte Carlo simulations. For each cell a number N is simulated from its calibrated frequency distribution. Then N amount of losses are simulated from the corresponding severity distribution and are summed up. This sum corresponds to a possible annual loss in that cell. This step is repeated 10 000 times to get a large sample of the cell specific loss distribution. Now, assuming a perfect correlation between the 99.9% quantiles of each cell the $VaR_{99.9\%}$ of each cell can be added together to get the total capital requirement, as suggested by the Basel Committee. However, by assuming independence between the high quantiles, diversification benefits can be utilized and all the samples in all cells can be summed up to create a sample from the total aggregated loss distribution, from which the 99.9% quantile can be read off. Implementing dependency in the model yields on average an 8.3% higher capital compared to the full independent assumption (Basel Comittee, 2009. p. 5) [4]. Since correlations between cells are low, a full independence assumption is made over the more conservative perfect correlation assumption.

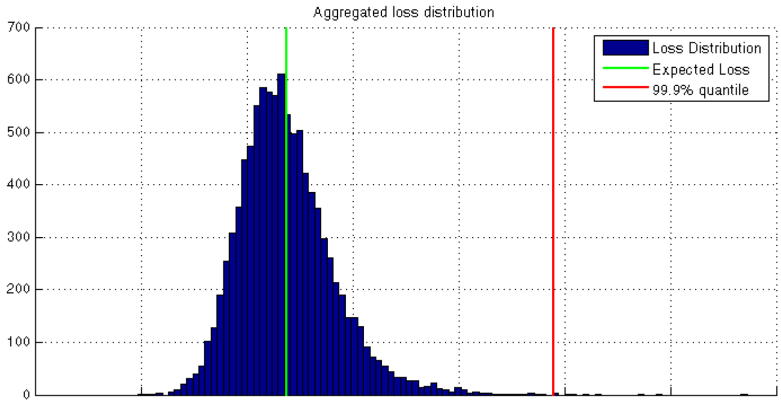


Figure 4.10: Total aggregated loss distribution with the 99.9% quantile and its mean marked.

There are some rare cases when dealing with really heavy tailed distributions where the full dependence assumption actually yields a lower capital estimate than the independent assumption. This deviation has to do with value at risk not being subadditive, i.e.

$$VaR_\alpha(X+Y) \nleq VaR_\alpha(X) + VaR_\alpha(Y)$$

This can be shown with a simple example from H. Hult et. al. (2012) p. 260. Consider two independent nonnegative random variables X_1 and X_2 with common distribution function F with a regularly varying tail $\overline{F} = 1 - F$ with index $\gamma^{-1} \in (0,1)$. For sufficiently large x it can be shown that $\mathrm{P}(X_1 + X_2 > x) > \mathrm{P}(2X_1 > x)$. For $p \in (0,1)$ sufficiently large

$$\begin{aligned} F_{X_1}^{-1}(p) + F_{X_2}^{-1}(p) &= 2F_{X_1}^{-1}(p) = F_{2X_1}^{-1}(p) \\ &= \min\{x : \mathrm{P}(2X_1 > x) \leq 1-p\} \\ &< \min\{x : \mathrm{P}(X_1 + X_2 > x) \leq 1-p\} \\ &= F_{X_1+X_2}^{-1}(p). \end{aligned}$$

This holds for γ^{-1} close to 1. The conclusion is that the sum of the quantiles for two independent and identically distributed random variables is not necessarily greater than the quantile of the sum.

4.6 Allocation

Allocating the capital back to business lines are done by using expected shortfall. We choose a confidence level α such as the expected shortfall at confidence α corresponds to the 99.9% quantile, i.e.

$$E(L | L \geq VaR_\alpha(L)) = VaR_{0.999}(L)$$

Then the cell specific expected shortfall at confidence level α is computed on every cell and summed up. The cell specific expected shortfall divided by the sum of expected shortfall times the total economic capital is the allocated economic capital to that cell. Value at Risk can also be used but the reason expected shortfall is preferred is due to it being sub-additive and it takes the tail beyond the 99.9% quantile in to consideration. Allocated capital within the same business line is summed up to obtain the individual business line capital, which can be compared to the standardized approach.

4.7 Qualitative Adjustments

Qualitative adjustments can be done to

- the frequency parameters,
- the severity parameters
- or directly to the estimated capital for different business lines.

The idea is to implement some scoring mechanism based on BEICFs such key risk indicators (KRIs) and risk self assessments (RSA) from which an adjustment up or down to the capital can be made. To explore which BEICFs are the most suitable has been out of scoop of this project. Below is an example of capital can be adjusted based on a score.

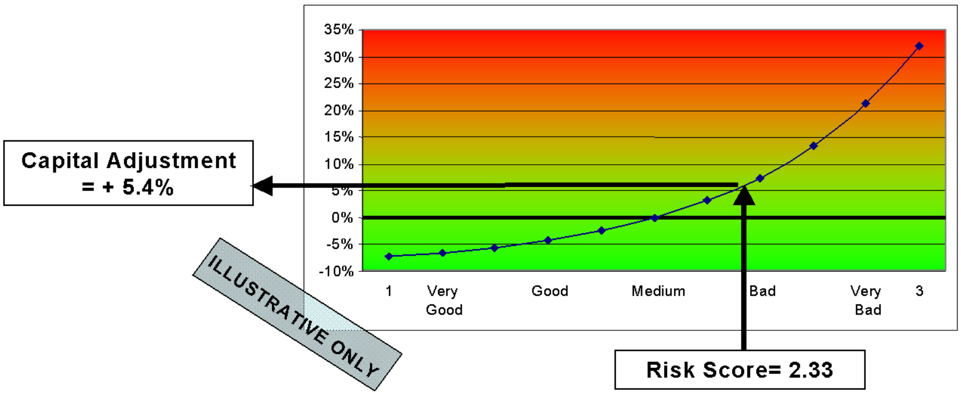


Figure 4.11: Example of how a qualitiative adjustment based on some scoring mechanism can look like

4.8 Model Description

Here follows a more straight walk through of the model

4.8.1 Data Selection, Analysis & Mapping Process

The corresponding loss accounting date, business line, event type, region and gross loss amount for every single data point in the internal and external loss database is loaded into Matlab. Dates, event types, business lines and regions are converted into numbers according to the tables below.

#	Business Lines	Event Type
1	Corporate Finance	Internal Fraud
2	Trading & Sales	External Fraud
3	Retail Banking	Employment Practices & Workplace Safety
4	Commercial Banking	Clients, Products & Business Practice
5	Payment & Settlement	Damage to Physical Assets
6	Agency Services	Business Disruption & System Failures
7	Asset Management	Execution, Delivery & Process Management
8	Retail Brokerage	
9	Insurance (excluded)	
#	External Region	Internal Region
1	Africa	Sweden
2	Asia/Pacific	Baltic
3	Eastern Europe	International
4	Latin America & Caribbean	
5	North America	
6	Western Europe & System Failures	

Table 4.3: Numbers assigned to business lines, event types and regions

The first internal loss accounting date is identified as well as the last date. A limit for which data are relevant is set to five years before the last loss accounting date. Thereafter a 7x8 sized matrix is constructed, where each cell corresponds to a business line and an event type.

Now the model scans through all the internal loss data and maps every loss to its correct BL/ET cell, but only if the loss occurred after the relevant date, i.e. loss accounting date is no older than five years. Older losses are stored in a separate matrix, since we still might want to use these when modeling the severity. Each loss is also adjusted for inflation. After determining how many years ago the loss occurred the models goes through the past years annual inflation rates and adjust the loss accordingly.

The mapping process is now repeated for three new matrices which correspond to the three regions Sweden, Baltics and International. E.g. in the Sweden matrix only losses that occurred in Sweden no longer than five years ago are selected and mapped to the corresponding BL/ET cell. In a similar way, all external data are converted into SEK and mapped to a new matrix.

The data in every cell in all matrices are now analyzed and inspected with histograms to look for any anomalies. Certain internal losses are discarded due to errors made in the reporting process and a few losses are adjusted down due to a fixed amount had been added to them which caused irregularities in the data set.

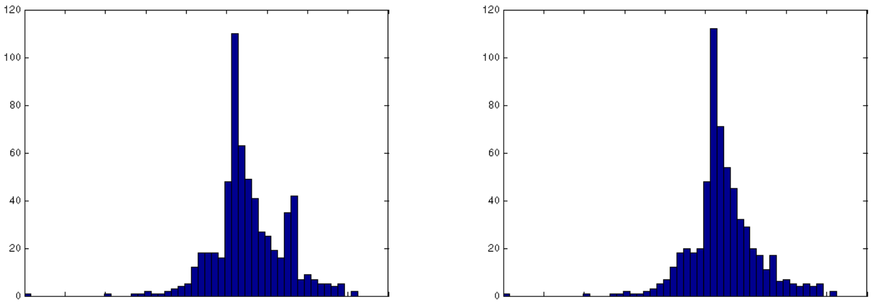


Figure 4.12: Example of data irregularities before and after adjusted.

The mean, median, min/max-values, standard deviation, skewness, kurtosis and number of occurrences is computed and stored. The number of occurrences for all internal losses over the past five year can be seen in Table 4.4 (not the actual figures).

BL/ET	1	2	3	4	5	6	7	8
1	109	NaN	NaN	21	NaN	54	NaN	438
2	141	NaN	NaN	260	117	126	NaN	205
3	NaN	2	NaN	108	43	NaN	86	81
4	NaN	NaN	60	NaN	NaN	336	NaN	NaN
5	NaN	NaN	NaN	153	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Table 4.4: Numbers of occurences (Illustrative only)

4.8.2 The Calibration & Simulation Process

Now the model goes through every cell, starting with commercial banking at position (1,1) and goes down the column, cell by cell, calibrating frequency, severity and estimating aggregated loss distribution using monte carlo simulations, until it reaches the last cell (7,8) in Retail Brokerage. First it checks that the cell is not empty. If it is empty, the model check for data older than 5 years. If it is still empty, the cell is left out of the model and it starts with a new cell.

4.8.3 Calibrating Frequency

For every cell that is not empty a Poisson distribution is fit by calibrating its parameter, taking the number of losses over the past five years in that cell divided by the number of years the relevant data set spans. In our case five years. The Poisson parameter is simply the yearly mean of the sample's occurrence. All frequencies between cells are currently assumed to be independent of each other.

Remarks

Dependencies between cells may be modeled with a copula, specified by a correlation matrix, and with marginal distributions transformed to Poisson. Dependence is not currently modeled due to too much uncertainty in the estimation of correlation and choice of dependence structure.

4.8.4 Calibrating Severity

First of all the generalized Pareto tail is calibrated numerically on all internal losses according to the peaks over threshold method. Then, for every cell with above 10 data points, the severity distribution is calibrated with a mixed model. The empirical CDF for both internal and external data are computed. The empirical CDF is a step function. To smooth it we take the midpoint at each jump. To ensure that the smoothed CDF covers the closed interval $[0, 1]$ the first and last elements of are linearly extrapolated using the adjacent slope of the piecewise linear CDF. Now to apply the weighted sum on the internal and external interpolation nodes they have to be of equal length. Therefore all nodes are cumulated and non unique values are removed. The cumulated nodes are evaluated in both the CDFs resulting in two empirical estimators, $\widehat{F}_{ext}$ and $\widehat{F}_{int}$, of equal length which we can apply a weighted sum to. The generalized Pareto distribution can then be "glued" on to the empirical weighted sum. The weighted sum estimator $\widehat{F}$ is evaluated at the threshold parameter u, which yields the upper probability where the generalized Pareto distribution begins.

4.8.5 Scenario Analysis

Cells few data points have the frequency calibrated exactly as before, but the severity distribution is calibrated with a log-normal (can be changed to log-logistic) distribution. The parameters are calibrated using typical loss (M) and worst case (WC) scenario estimated by multiplying the internal pivot cell to the current external loss cell ratio.

4.8.6 Capital Calculation

The aggregated loss distribution is estimated for each cell by Monte Carlo simulation. A random number n, is drawn from the corresponding Poisson distribution. Then n losses from the severity distribution is drawn and summed up. If the severity distribution is a mixed model with empirical body and pareto tail, n uniform variables on the interval [0,1] are instead generated. Then for every n less than the upper probability where the generalized Pareto tail begins a loss from the weighted empirical distribution is interpolated. And for every n larger or equal to the upper probability level a loss from the generalized Pareto distribution is generated. These are all summed up and this step is repeated 10000 times to get the cell specific loss distribution. The 99.9% quantile of the cell specific aggregated loss distribution is computed and stored in the matrix. When all cells have been evaluated, all aggregated loss distributions are summed to get the total aggregated loss distribution, from which the 99.9% quantile is computed and compared to the sum of all cell specific 99.9% quantiles.

Chapter 5

Results

	Log-normal		Weibull		Log-logistic	
Cell	μ	σ	a	b	α	β
1	7.98	2.58	9854.87	0.47	3640.95	0.68
2	7.32	2.10	4469.47	0.44	1422.26	0.84
3	8.02	2.66	10332.46	0.49	3827.63	0.66
4	8.32	1.68	9642.73	0.55	3904.95	1.15
5	7.37	1.94	4179.93	0.50	1556.20	0.93
6	6.85	2.54	3811.34	0.32	699.24	0.72
7	7.41	1.79	4227.73	0.50	1394.09	1.08
8	6.90	1.59	2247.93	0.55	934.49	1.20
9	10.27	2.16	86217.69	0.46	27446.67	0.79
10	10.02	2.90	81688.18	0.49	32859.63	0.61
11	9.32	2.13	29514.11	0.59	13095.19	0.87
12	9.05	2.80	35033.01	0.37	7707.89	0.61
13	2.57	3.09	73.76	0.26	7.54	0.64
14	5.86	2.64	1234.49	0.47	424.11	0.62
15	6.34	2.31	1790.16	0.46	578.25	0.72
16	8.70	1.48	13171.11	0.61	5377.61	1.23
17	9.85	2.22	66156.68	0.36	12964.89	0.87
18	9.81	1.63	45169.38	0.43	16317.61	1.18
19	9.70	2.13	45316.93	0.57	18214.99	0.78

Table 5.1: Estimated parameters of log-normal, weibull and log-logistic distribution on cell level.

	MLE			LSE		
Cell	μ	σ	ν	μ	σ	ν
1	7.98	2.58	3999005.41	7.99	2.88	64.59
2	7.29	1.97	16.86	7.32	1.93	10.57
3	8.02	2.66	1938524.16	8.02	2.88	31.33
4	8.22	1.09	2.90	8.32	1.21	3.78
5	7.37	1.68	7.61	7.37	1.65	5.35
6	6.49	2.03	5.28	6.85	2.18	5.04
7	6.99	1.00	2.11	7.41	1.21	2.75
8	6.76	0.98	2.54	6.90	1.19	4.22
9	10.27	2.16	6939667.79	10.27	2.31	113.35
10	10.09	2.83	38.35	10.02	2.98	14.90
11	9.51	1.71	5.52	9.32	1.66	3.16
12	9.05	2.80	5559061.55	9.05	2.98	∞
13	1.22	0.87	1.08	2.57	1.89	1.97
14	5.86	2.64	5169736.78	5.86	3.04	∞
15	6.34	2.31	5284308.24	6.34	2.39	∞
16	8.51	1.02	2.96	8.70	1.10	2.14
17	8.76	0.81	1.23	9.85	1.44	1.43
18	9.66	1.18	4.13	9.81	1.08	2.60
19	9.70	2.13	3701628.06	9.70	2.58	∞

Table 5.2: Estimated parameters of Student's t distribution on cell level.

	$VaR_{99.9\%}$	
Cell	Monte Carlo	Analytical
1*	7	-
2*	28	-
3*	30	-
4*	2	-
5	39	38
6	16	12
7	57	58
8	13	16
9*	1	-
10	12	15
11	17	24
12	44	36
13	27	28
14*	1	-
15*	2	-
16	51	24
17*	2	-
18	22	20
19*	1	-
20*	13	-
21*	1	-
22	7	7
23*	33	-
24*	1	-
25*	5	-
26*	24	-
27*	1	-
28	25	37
29*	1	-
(*) Based on scenario analysis		
Perfect correlated $VaR_{99.9\%}$		480
Independent $VaR_{99.9\%}$		161

Table 5.3: The 99.9% quantiles for each cell based on 10 000 monte carlo simulations from a mixed empirical+pareto distribution and the 99.9% quantiles based on the analytical approximation. The perfect correlated $VaR_{99.9\%}$ is the sum of all 99.9% quantiles on cell level, assuming a perfect correlation. The independet $VaR_{99.9\%}$ is the 99.9% quantile of the total aggregated loss distribution, i.e. the sum of independent samples from the loss distributions on cell level.

Chapter 6

Concluding Chapter

In this last chapter we will summarize the conclusions that can be drawn from this thesis together with suggestions of further research.

6.1 Conclusions & Thoughts

Modeling operational risk seems very easy at first. Take a sample of data, fit a frequency distribution and a severity distribution and start simulating losses. And perhaps for certain cells with many data points it is relatively easy. But operational risk is such a wide concept and to develop a model that must include all types makes it really hard. Especially when you have less data in a cell than you have fingers on your hands. And even if you have a lot of data, one single observation can have a huge impact on the outcome. Nevertheless, there are some techniques and methods that have proven to work well.

We saw that fitting one distribution over the whole range of data worked for some internal cells, where log-normal and log-logistic had the best goodness of fit, but it did not work well with heavier tailed cells and external data. Instead the empirical-pareto mixed model had the best goodness of fit and was able to generate a loss distribution with realistic capital estimates. A shortcoming is that the tail is calibrated entirely on sparse internal tail data, which makes it very sensitive to extreme observations. With scaled external data we could get a denser tail which would make the estimates more stable. The analytical approximation seemed to coincide well with the monte carlo simulations at higher quantiles. But it is also very sensitive to its parameter estimates. Finally, using the external empirical distribution instead of extrapolating losses beyond internal loss range from a parametric distribution worked well, but yielded a slightly higher estimated capital than in the case with the mixed model.

6.2 Dependence

Generally, there are very little evidence of strong correlations and dependence structures among operational losses. This observation can obviously in no reasonable way be interpreted as 'evidence of no correlation'. But for now, no dependence is currently in the model, instead full independence is assumed. If the only option to not include any dependence in the model is to assume perfect correlation between the aggregated loss quantiles, then there is much reason to incorporate some sort of dependency in the model, and do further research on the estimation of correlations and the choice of dependence structure.

6.3 Scaling

To make estimates of the tail more stable, external data must be added. Finding a good way to scale or filter external loss data would make the modeling a lot easier.

6.4 Final Words...

Last but not least actual BEICFs needs to be implemented in the model, as well as expert opinions on scenario input and perhaps even correlations in the future.

Bibliography

[1] Basel Committee on Banking Supervision (2011), *Operational Risk - Supervisory Guidelines for the Advanced Measurement Approaches*, Available at http://www.bis.org/

[2] Basel Committee on Banking Supervision (2011), *Principles for the Sound Management of Operational Risk*, Available at http://www.bis.org/

[3] Basel Committee on Banking Supervision (2009), *Observed range of practice in key elements of Advanced Measurement Approaches (AMA)*, Available at http://www.bis.org/

[4] Basel Committee on Banking Supervision (2009), *Results from the 2008 Loss Data Collection Exercise for Operational Risk*, Available at http://www.bis.org/

[5] Basel Committee on Banking Supervision (2006), *International Convergence of Capital Measurement and Capital Standards*, Available at http://www.bis.org/

[6] F. Aue, and M. Kalbrenner (2006), LDA at work, *Journal of Operational Risk*, 1(4), 49-93.

[7] K. Dutta, and J. Perry (2006), A Tail of Tails: An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital, Federal Reserve Bank of Boston, Working Paper No 06-13. Available at http://www.bos.frb.org/.

[8] A. Chapelle, Y. Crama, G. Hübner and J. P. Peters (2007), Practical methods for measuring and managing operational risk in the financial sector: A clinical study, *Journal of Banking & Finance*, 32, 1049-1061.

[9] K. Böcker and C. Klüppelberg (2005), Operational VAR: a closed-form approximation, *RISK*, 90-93.

[10] A. Colombo and S. Desando (2008), Developing and Implementing Scenario Analysis Models to Measure Operational Risk at In-

tesa Sanpaolo, *TheMathWorks News&Notes*, 91606v00. Available at: http://www.mathworks.com/.

[11] E. Cope and G. Antonini (2008), Observed correlations and dependencies among operational losses in the ORX consortium database, *Journal of Operational Risk*, 3(4), 47-76.

[12] E. Cope and A. Labbi (2008), Operational loss scaling by exposure indicators: evidence from the ORX database, *Journal of Operational Risk*, 3(4), 25-46.

[13] E. W. Cope, G. Mignola, G. Antonini and R. Ugoccioni (2009), Challenges in Measuring Operational Risk from Loss Data, *Journal of Operational Risk*, 4(4), 3-27.

[14] H. Hult, F. Lindskog, O. Hammarlid, and C.J. Rehn, *Risk and Portfolio Analysis*, Springer Series in Operations Research and Financial Engineering, Springer, 2012.

[15] N.N. Taleb, *The Black Swan: The Impact of the Highly Improbable*, second edition, Penguin Books, 2010.

[16] G. Blom, *Sannolikhetsteori och statistikteori med tillämpningar*, 5. uppl., Studentlitteratur, Lund, 2005

[17] Finansinspektionen (2007), *Ansökan om internmätningsmetod, operativ risk*, Available at: http://www.fi.se/.

[18] F. Lindskog, A. McNeil (2003), Common Poisson shock models: applications to insurance and credit risk modelling, *ASTIN Bulletin*, 33(2), 209-238.

[19] P. Embrechts, G. Puccetti, L. Rueschendorf (2012), Model uncertainty and VaR aggregation, Available at: http://www.math.ethz.ch/~embrecht/papers.html.

[20] M.P. McLaughlin (2001), *A Compendium of Common Probability Distributions*, Available at http://www.causascientia.org/math_stat/Dists/Compendium.pdf.

[21] Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, Office of the Comptroller of the Currency, Office of Thrift Supervision (2011), Interagency Guidance on the Advanced Measurement Approaches for Operational Risk, *Federal Reserve System: Supervision and Regulation Letters*, SR 11-8 (Attachment).

TRITA-MAT-E 2012:14
ISRN-KTH/MAT/E--12/1-SE

www.kth.se