Applying Robust Methods to Operational Risk Modeling

Anna Chernobai*
University of California
at Santa Barbara, USA

Svetlozar T. Rachev†
Universität Karlsruhe,
Germany and
University of California
at Santa Barbara, USA

March 7, 2006

Contents

*Email: annac@pstat.ucsb.edu. Parts of the research were done when A.Chernobai was visiting Institute of Econometrics, Statistics, and Mathematical Finance, School of Economics and Business Engineering, University of Karlsruhe, Germany, in 2005.

†Email: rachev@pstat.ucsb.edu. S.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).

1 Introduction

In 2001, the Basel Committee of the Bank of International Settlements (BIS) released new regulatory capital guidelines on operational risk (BIS, 2001a,b), finalized in 2004 (BIS, 2004). The nature of operational risk, fundamentally different from that of market risk and credit risk, is highly bank-specific and calls for the development of complex quantitative and qualitative solutions, new “know-how,” and setting additional standards for training bank personnel. Despite significant progress in the operational risk management, numerous challenges remain, and the current development in the area is criticized by many as being “more art than science.”

An actuarial type model dominates statistical models for operational risk under the Advanced Measurement Approach (see Section 2 for the discussion of the model; see also publications by BIS, 2001-2004). A quantitative model is required to have the capacity of accommodating peculiarities of the loss distribution: high kurtosis, severe right-skewness, and excessive heavy-tailedness. Model selection is complicated by scarcity of the available data along with the presence of “tail events”, the so-called “low frequency/ high severity” losses, that contribute to the heaviness of the upper tail of the loss distribution. Some critics of the Basel II framework argue that the standards required on calculation of regulatory capital are such that this figure might even exceed the economic capital (see, e.g., Currie (2005)), leaving decreased availability of funds required for financial needs and investments (FGG, 2005). This may be well due to misspecification in the model. In this paper, we propose an approach that can provide a solution to this dilemma.

In 2001, the Basel Committee made the following recommendation (BIS, 2001a, Annex 6, p. 26):

> “...data will need to be collected and *robust* estimation techniques (for event impact, frequency, and aggregate operational loss) will need to be developed.”

The notion of *robustness* can be given different interpretations. One such interpretation would be the distributional robustness, i.e., robustness of the assumed model to minor departures from the model assumptions. *Outlier-resistant* or *distributionally robust* (so-called robust) statistics methods[1] aim at constructing statistical procedures that are stable (robust) even when the underlying model is not perfectly satisfied by the available dataset. An example of departure from the assumed model is the presence of outliers – observations that are very different from the rest of the data. Outliers are “bad” data in the sense that they deviate from the pattern set by the majority of the data (Huber (1981), Hampel et al. (1986)). Hence, they tend to obscure its generic flow and may lack explanatory and predictive power regarding the generic portion of the data. Robust models focus on the statistical properties of the bulk of the data without being distracted by outliers, while in classical models all data equally participate in the analysis.

[1] We use terms *robust methods* and *robust statistics methods* interchangeably in this paper.

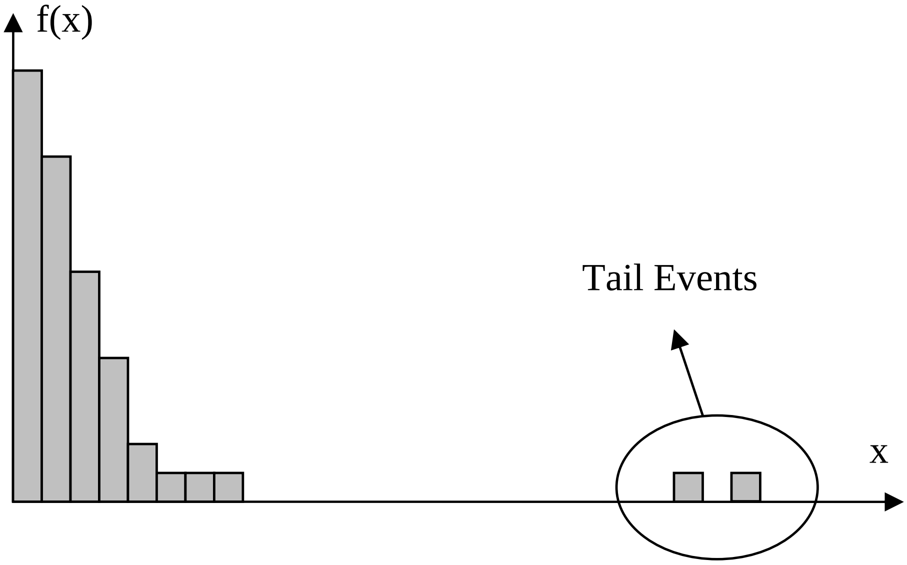

Figure 1: An examplary histogram of operational loss data.

Performing robust or classical analysis of the data is a trade-off between safety and efficiency: although some information may be lost while discarding or diminishing the contribution of the outlying events, one can significantly improve forecasts and produce more reliable estimates by applying a robust methodology.

The purpose of this paper is to lay ground for applying robust techniques to quantitative modeling of operational risk.[2] In Section 2 we give a brief overview of the actuarial type model for operational risk and discuss some distinctive characteristics of the operational loss data. Section 3 discusses robust models and their possible applications to operational risk. Empirical study with operational risk data is presented in Section 4, and Section 5 concludes the paper and states final remarks.

2 Actuarial Approach to Modeling Operational Risk

The Loss Distribution Approach (BIS, 2001a) suggests an actuarial type model for the aggregated operational losses. For a particular 'business line/ event type' combination the losses are assumed to follow a stochastic process $\{S_t\}_{t\geq 0}$ so that for a one-year time interval $\triangle t$

$$S_{\triangle t} = \sum_{k=0}^{N_{\triangle t}} X_k, \quad X_k \overset{\text{iid}}{\sim} F_\gamma, \tag{1}$$

in which the *iid* random sequence of loss magnitudes $\{X_k\}$ follows the distribution function (cdf) F_γ and the density f_γ, and the counting process $N_{\triangle t}$ is assumed to take a form of a homogeneous Poisson process (HPP) with intensity $\lambda > 0$ (or a non-homogeneous Poisson

[2]The material presented in this paper is based on the discussion of the chapter on robust methods for operational risk modeling in Chernobai et al. (2006b).

process (NHPP) with intensity $\lambda(t) > 0$), independent from $\{X_k\}$. Under the stated assumptions, the expected aggregate loss (EL) is calculated as the product of expected frequency and expected severity:

$$\text{EL} := \mathbb{E}S_{\triangle t} = \mathbb{E}X \times \mathbb{E}N_{\triangle t}. \tag{2}$$

The one-year operational risk regulatory capital charge can be estimated by the Value-at-Risk (VaR) measure. VaR is measured by the $(1-\alpha) \times 100^{\text{th}}$ quantile of the cumulative loss distribution over a one year period, with α taken to be, for example, $0.01 - 0.05$. VaR is obtained as the solution to:

$$P(S_{\triangle t} > \text{VaR}_{\triangle t, 1-\alpha}) = \alpha. \tag{3}$$

An alternative risk measure, Conditional VaR (CVaR),[3] is defined by:

$$\begin{aligned} \text{CVaR}_{\triangle t, 1-\alpha} : &= \mathbb{E}\left[S_{\triangle t} \mid S_{\triangle t} > \text{VaR}_{\triangle t, 1-\alpha}\right] \\ &= \frac{\mathbb{E}\left[S_{\triangle t}\ ;\ S_{\triangle t} > \text{VaR}_{\triangle t, 1-\alpha}\right]}{\alpha}. \end{aligned} \tag{4}$$

Existing empirical evidence suggests that the general pattern of operational loss severity data $\{X_k\}$ is characterized by high kurtosis, severe right-skewness, and a very heavy right tail created by several outlying events. Figure 1 portrays an illustrative example of the operational loss severity data. Studies supporting the claim include Cruz (2002), Medova (2002), Moscadelli (2004), Embrechts et al. (2004), De Fontnouvelle et al. (2005), Chernobai et al. (2005), and Neslehová et al. (2006).

One approach to calibrate operational losses is to fit a parametric family of distributions, such as Lognormal, Weibull, Gamma, Pareto, etc. These distributions may not be optimal in fitting well both the center and the tails (see, for example, discussion in Chernobai et al. (2005)). An alternative approach, pioneered by several works by P. Embrechts, uses the Extreme Value Theory (EVT) to fit a Generalized Pareto Distribution (GPD) to extreme losses exceeding a high pre-specified threshold; see Embrechts et al. (2003), Embrechts et al. (2004), Chavez-Demoulin and Embrechts (2004), and Neslehová et al. (2006); an excellent reference on EVT is Embrechts et al. (1997). The parameters of the GPD distribution obtained under the EVT approach, however, are highly sensitive to the choice of threshold and additional extreme observations. Furthermore, both approaches often produce an infinite mean and unreasonably high estimates for the capital charge (Neslehová et al. (2006), Chernobai et al. (2005)).

[3]CVaR is also called Expected Tail Loss (ETL) or Expected Shortfall (ES). Unlike VaR, CVaR satisfies the properties of a *coherent* risk measure (Artzner et al., 1999) and allows to capture better the tail events.

3 Why Robust Statistics?

3.1 Classical vs. Robust Methods

The presence of outlying events, the so-called "low frequency / high severity" events, in the operational loss data creates a following paradox.

On the one hand, the tail events correspond to the losses that, despite their low frequency of occurrence, are often the most destructive for the institution. In this sense, they cannot be ignored as they convey important information regarding the loss generation process and may signal important flaws in the system.

On the other hand, recent empirical findings suggest that classical methods will frequently fit neither the bulk of the operational loss data nor the outliers well, and the center and the tails of the data appear to conform to different laws. Classical estimators that assign equal importance to all available data are highly sensitive to outliers and in the presence of just a few extreme losses can produce arbitrarily large estimates of mean, variance, and other vital statistics. For example, a high mean and standard deviation values for operational loss data do not provide an indication as to whether this is due to generally large values of observations or just one outlier, and it may be difficult to give the right interpretation to such result. On the contrary, robust methods take into account the underlying structure of the data and "separate" the bulk of the data from outlying events, this way avoiding the upward bias in the vital statistics and forecasts.

A reader may argue that the extreme operational losses largely determine the shape of the upper tail of the loss distribution that in turn determines the operational risk capital charge (in the VaR approach); hence, they should not be blindly discarded. Robust methods *do not* aim at throwing away extreme observations. They focus on the behavior of the bulk of the data that can be easily distorted by outliers.

An important application of robust statistics is using them as a *diagnostic technique* for evaluating the sensitivity of the inference conducted under the classical model to the rare events and to reveal their possible economic role (Knez and Ready, 1997). We therefore emphasize that the classical model and the robust model are not competitors – we encourage the use of both models as important complements to each other, rather than advocating the use of the robust model instead of the classical. The results from both approaches are not expected to be the same, as they explain different phenomena dictated by the original data: the general tendency (the robust method) and the conservative view (the classical method).

Applications of robust analysis can be found in a variety of recent finance literature and are dominant in regression analysis. A classical example is the study of stock return anomalies by Fama and French (1992): they argue that there appears to be risk premia associated with the size of firm and book-to-market. Knez and Ready (1997), however, demonstrated that the results were driven by a small portion of firms, and use instead least trimmed squares (LTS) as a robust regression technique to trim the few outlying observations, and then perform OLS on the remainder of the data. Another recent study is due to Kim and White (2003) who apply robust estimators to examine the properties of the S&P500 index returns. They find evidence that the S&P500 index returns are composed of a mixture of two components, with a predominant component being nearly symmetric with mild kurtosis, and a relatively rare component generating extreme anomalies. Bassett et al. (2004) investigate the performance of portfolio return distribution using robust and quantile-based methods, and conclude that the resulting forecasts outperform those under a conventional classical analysis. Perret-Gentil and Victoria-Feser (2005) use robust estimates for mean and the covariance matrix in the mean-variance portfolio selection problem. They show that the robust portfolio outperforms the classical one, and the outlying observations that account for 12.5% of the dataset can have serious influence on portfolio selection under the classical approach.

3.2 Some Examples

In this section we illustrate some dangers of using only the classical approach in modeling operational risk.

Suppose, a risk expert constructs a one quarter ahead forecast of the total operational loss, based on the historic data of his institution that includes the events of the order of magnitude of “9/11” or the Hurricane Andrew (1992) and the Hurricane Katrina (2005) in the model. Would his forecast be robust? Most likely, his forecasts would indicate that his bank will have little reserves left if it decides to cover the potential loss.

As another example, suppose a risk analyst fits a heavy-tailed loss distribution to full data which include several “low frequency/ high severity” data points. The estimate of the aggregate expected loss (EL) is likely to be very high. In particular, if the fitted distribution is very heavy-tailed, such as some cases of Pareto or α-Stable, he may get an infinite mean and infinite second and higher moments’ estimates. Occasionally, EL may even exceed VaR. Ongoing discussions by the Risk Management Group of the BIS suggest excluding the EL amount from the total estimated capital charge (e.g., VaR or CVaR) and set the charge on the basis of the marginal unexpected loss[4] (UL), provided that the bank can demonstrate its ability to effectively monitor expected operational losses. The danger of treating outliers equally with the rest of the data is that the resulting UL-based capital charge may appear insufficient to cover the true exposure to the risk.

[4]BIS defines UL as VaR − EL.

3.3 Overview of Literature on Robust Statistics

We here give an overview of the literature on robust statistics.A pioneering work on robust statistics is due to Huber (1964) and Hampel (1968). Robust statistics is the generalization of the classical theory: it takes into account the possibility of model misspecification, and the inferences remain valid not only at the parametric model but also in the neighborhood. The aims of robust statistics are (Hampel et al., 1986):

1. To describe the structure best fitting the bulk of the data;
2. To identify outliers (for possible further treatment);
3. To give a warning about highly influential data points ("leverage points");
4. To deal with deviations from the assumed correlation structures.

5-10% of wrong values in the data appears to be the rule rather than the exception (Hampel, 1973). Outliers may appear in data due to (a) gross errors, (b) wrong classification of the data (outlying observations may not belong to the model followed by the bulk of the data), (c) grouping, and (d) correlation in the data (Hampel et al., 1986).

Let $(1 - \varepsilon)$ be the probability of well-behaved data, and ε be the probability of data being contaminated by "bad" observations. If $H(x)$ is an arbitrary distribution defining a neighborhood of the parametric model F_γ, then G is the two-point mixture of the parametric model and the contamination distribution:

$$G(x) = (1 - \varepsilon)F_\gamma(x) + \varepsilon H(x). \tag{5}$$

Under traditional robust models, outliers are exogenously detected and excluded from the dataset, and the classical analysis is performed on the "cleaned" data. Data editing, screening, truncation, censoring, Winsorizing, and trimming are various methods for data cleaning. Outlier rejection approach is the simplest robust approach. Such procedures for outlier detection are referred in literature as "forwards-stepping rejection", or "outside-in rejection" of outliers (see, e.g., Simonoff (1987a,b)).

There are two kinds of outlier detection methods in the forwards-stepping rejection procedure: informal and formal. The former approach is rather subjective: a visual inspection of the database may be performed by a risk expert, and data points that clearly do not follow "the rule of the majority" are excluded. A risk expert may further conduct a background analysis of extreme losses, analyze whether they follow a pattern, and decide whether they are likely to repeat in future. Which losses and how many to exclude is left up to his subjective judgment. For example, Moscadelli (2004) examines the operational loss data[5] and excludes one outlier from the Retail Brokerage loss data (that consists of a total of 3,267

[5]The data are taken from the second Loss Data Collection Exercise (Quantitative Impact Study 3), see also BIS (2003).

observations) and five outliers from the Commercial Banking loss database (that consists of a total of 3,414 observations).

Formal approaches to discriminate outliers include trimming and Winsorizing data. For example, $(\delta, 1-\gamma)$-trimmed data has the lowest δ and the highest γ fractions of the original data removed. For symmetrically contaminated data, $\delta = \gamma$. In the context of operational risk, contamination is asymmetric (on the right) and $\delta = 0$. For an original sample $x_j, j = 1, \ldots, n$ of size n, define $L_n = \lfloor n\delta \rfloor$ and $U_n = \lfloor n\gamma \rfloor$, where $\lfloor a \rfloor$ denotes the floor of a, and let $x_{(k)}$ denote the kth order statistic such that $x_{(1)} \leq \ldots \leq x_{(n)}$. Winsorizing data is more efficient than trimming: the lowest $\lfloor n\delta \rfloor$ observations in the original dataset are set equal to the lowest observation in the "cleaned" data, and the highest $\lfloor n\gamma \rfloor$ observations are set equal to the greatest observation in the "cleaned" data. The Winsorized sample $y_j, j = 1, \ldots, n$ is thus obtained by transforming $x_j, j = 1, \ldots, n$ in the following way:

$$y_j = \begin{cases} x_{(L_n+1)} & j \leq L_n \\ x_{(j)} & L_n + 1 \leq j \leq U_n, \quad j = 1, \ldots, n \\ x_{(U_n)} & j \geq U_n + 1. \end{cases} \tag{6}$$

Other outlier rejection principles include rejection rules based on kurtosis, largest Studentized residual, Studentized range, Shapiro-Wilk statistic, and Dixon's rule. A variety of outlier rejection methods have been discussed by Hampel (1973, 1985), Hampel et al. (1986), Simonoff (1987a,b), Stigler (1973), (David, 1981, Ch.8), to name a few. The main criticism of the outlier rejection approach is that information is lost due to discarding several data points. One possibility is to choose to allow a fixed efficiency loss of, say, 5% or 10% (Hampel et al., 1986, p.44). Hampel (1985) also showed that outside-in outlier rejection procedures possess low *breakdown points*[6] and estimators can be severily affected by a relatively small number of extreme observations, which means that estimators are not robust to heavy contamination. Nevertheless, despite the criticism, "any way of treating outliers which is not totally inappropriate, prevents the worst" (Hampel, 1973).

Examples of non-robust estimators include the arithmetic mean, standard deviation, mean deviation and range, covariance and correlation, ordinary least squares (OLS). Robust measures of center include median, trimmed mean, and Winsorized mean. Robust measures of spread include inter-quartile range (IQR), median absolute deviation (MAD), mean absolute deviation, and Winsorized standard deviation; more estimators of scale were proposed by Rousseeuw and Croux (1993). Robust estimators of skewness were studied by Kendall and Stuart (1977), Bowley (1920), Hinkley (1975), and Groeneveld and Meeden (1984). Robust estimators of kurtosis for heavy-tailed distributions were proposed by Hogg (1972, 1974); others are due to Moors (1988), Hogg (1972, 1974), and Groeneveld and Meeden (1984).

Under more modern robust models, outliers are given a further treatment rather than being simply discarded. Outliers can be detected and rejected using a "backwards-stepping"

[6]The breakdown of an estimator is the maximum fraction of outliers that an estimator can tolerate (Huber, 1981), Simonoff (1987a).

or "inside-out" rejection procedure. One approach is based on the *influence functions* (IF), proposed by Hampel (1968, 1974); see also Hampel et al. (1986). IF measures the differential effect of an infinitesimal amount of contamination (e.g., one additional observation) in an uncontaminated sample on the value of the estimator T at a point x, standardized by the amount of contamination:

$$\mathrm{IF}(x; T, F_\gamma) = \lim_{\varepsilon \searrow 0} \frac{T(G) - T(F_\gamma)}{\varepsilon}. \tag{7}$$

IF can be used to measure the *gross-error sensitivity* (GES) – the worst (approximate) influence which a small amount of contamination of fixed size can have on the value of the estimator T (Hampel et al., 1986):

$$\mathrm{GES}(T, F_\gamma) = \sup_x |\mathrm{IF}(x; T, F_\gamma)|. \tag{8}$$

GES can be used as a tool to detect the observations having a large influence on the value of the estimator. "Inside-out" outlier rejection rules have high breakdown points and the estimators can tolerate up to 50% of contamination (Simonoff, 1987a). Further discussion on IF and "inside-out" outlier treatment procedures can be found in Huber (1981) and Simonoff (1987a,b). Other references on robust statistics include Rousseeuw and Leroy (1987), Martin and Simin (2003), Kim and White (2003), Aucremanne et al. (2004), Hubert et al. (2004), and Olive (2005).

3.4 Outlier Rejection Approach and Stress Tests: A Parallel

We note a parallel of data trimming with Stress tests that are widely applied in the operational risk modeling. In Stress tests, by adding a few high observations to the dataset one aims at examining the incremental effect of potentially hazardous events on VaR and other measures. With the robust methodology, outlying observations are excluded, rather than included, from the dataset with the purpose of examining the fundamental properties of the main subset of the data in the absence of these potentially impossible events as well as investigating their incremental effects on risk measures.

Decisions on whether to include (Stress tests) or exclude (robust method) high-magnitude events, or whether to perform both tests, as well as how many points and of what magnitudes to include or exclude, can be left up to the subjective judgement of the risk expert or can be performed using a formal (objective) procedure discussed earlier.

		Classical	Robust
n		233	221
min	($ '000,000)	1.1	1.1
max	($ '000,000)	6,384	364.80
mean	($ '000,000)	103.35	39.7515
median	($ '000,000)	12.89	11.40
st.dev.	($ '000,000)	470.24	63.84
MAD	($ '000,000)	11.17	9.64
skewness		11.0320	2.5635
kurtosis		140.8799	10.0539

Table 1: Descriptive sample statistics of full and top-5%-trimmed operational loss data.

4 Application to Operational Loss Data

The empirical section of this paper applies the simple data trimming technique to historic operational loss data.[7] The purpose of the study is to investigate the impact of outlying tail events on the performance of EL, VaR, and CVaR.

The dataset used in the study was obtained from a major European operational public loss data provider. The database is comprised of operational loss events throughout the world. The dataset used for the analysis covers losses in US$ exceeding 1,000,000 for the time period between 1980 and 2002. Our analysis is restricted to the data of loss type *"External"* that includes events related to natural and man-made disasters and external fraud.

Contamination of the data is located in the far right tail of the loss distribution. In this sense, the contamination is of a non-symmetric nature. We trim the original data by cutting off the *highest* 5% of losses. These correspond to twelve observations. Table 1 summarizes the descriptive statistics of the full and cleaned data. A dramatic change in the statistics is evident when the robust methodology is applied: the mean and the standard deviation have decreased to roughly 1/3 and 1/7 of the initial values, respectively; the skewness coefficient has dropped 4 times, and the kurtosis coefficient has decreased roughly 14 times. Note that the robust measures of center and spread – median and median absolute deviation (MAD), respectively – remain practically unchanged.

In the next step, we fit loss distributions to both complete and trimmed datasets.[8] Ta-

[7]Another application can be found in Chernobai et al. (2006a), in which the robust methodology was applied to the US natural catastrophe claims data. Results presented here are partially reproduced from Chernobai et al. (2005).

[8]To account for the reporting bias, left-truncated loss distributions were fitted using the method of maximizing the restricted likelihood function. See Chernobai et al. (2005a) and Chernobai et al. (2005) for a detailed description of the methodology, its theoretical implication on the estimates of the loss and frequency distribution parameters and the capital charge, and the empirical application to operational loss data.

	Lognormal			Weibull	
	Classical	Robust		Classical	Robust
μ	15.7125	15.8095	β	0.0108	0.0012
σ	2.3639	1.9705	τ	0.2933	0.4178
mean	$1.09{\cdot}10^8$	$0.51{\cdot}10^8$	mean	$5.21{\cdot}10^7$	$2.90{\cdot}10^7$
st.dev.	$1.78{\cdot}10^9$	$0.35{\cdot}10^9$	st.dev.	$2.96{\cdot}10^8$	$0.85{\cdot}10^8$

Table 2: Estimated parameters, mean, and standard deviation for loss distributions fitted to the full and top-5%-trimmed operational loss data.

ble 2 exhibits parameter estimates, mean, and standard deviation[9] for the Lognormal and Weibull distributions.[10] Outlier rejection has resulted in significantly decreased mean and the standard deviation estimates.

Next, we examine the aggregated 1-year EL, $\text{VaR}_{0.95}$, $\text{VaR}_{0.99}$, $\text{CVaR}_{0.95}$, and $\text{CVaR}_{0.99}$. The estimates are based on out-of-sample 1-year ahead forecast. For the frequency distribution, a Cox process with a non-homogeneous intensity rate function was used. We omit the estimated parameter values of the frequency distribution from this paper – see Chernobai et al. (2005) for the details. We note, however, that robust methods have a negligible effect on the parameters of the frequency distribution. Table 3 reports the findings. The estimates of the risk measures are considerably lower under the robust method in all cases. Hence, robust methods can prevent over-estimation of the capital charge.

Finally, we examine the incremental effect inflicted on these measures by the top 5% observations. The marginal impact was computed by

$$\triangle = \frac{T_{class.} - T_{robust}}{T_{class.}} \times 100\%, \tag{9}$$

with T being the appropriate measure – one of EL, VaR, and CVaR. Table 3 demonstrates that the twelve extreme data points account for up to 58% of the total EL, and up to 76% of the total operational risk capital charge (VaR or CVaR) for the data sample under consideration. The magnitude of the impact can serve as an important guideline for a bank to decide whether and at what price it should use insurance against extreme losses.

5 Conclusions

Classical estimation procedures, while being mathematically elegant, may appear inadequate under minor departures in the data from the model assumptions. In particular, in the pres-

[9]Note that in Table 1 the estimates of location and spread are based on the observed data that exceeds $1,000,000. In Table 2 the population estimates of location and spread are extrapolated to correspond to complete data.

[10]A large variety of distributions were applied in an earlier paper by the authors, see Chernobai et al. (2005).

	Lognormal			Weibull		
	Classical	Robust	△	Classical	Robust	△
EL	0.0327	0.0154	53%	0.0208	0.0088	58%
$\mathrm{VaR}_{0.95}$	0.1126	0.0580	48%	0.0885	0.0354	60%
$\mathrm{VaR}_{0.99}$	0.4257	0.1642	61%	0.2494	0.0715	71%
$\mathrm{CVaR}_{0.95}$	0.3962	0.1397	65%	0.2025	0.0599	70%
$\mathrm{CVaR}_{0.99}$	1.1617	0.3334	71%	0.4509	0.1066	76%

Table 3: Estimated 1-year EL, VaR, and CVaR values ($\times 10^{10}$) for the full and top-5%-trimmed operational loss data, and incremental effect (△) of the highest 5% losses.

ence of outliers, classical procedures may produce biased estimates of the model parameters and vital statistics. Robust statistics methodology provides a solution to such problem.

This paper discusses robust estimation techniques and their application to modeling heavy-tailed operational losses. "Low frequency/ high severity" events are an important characteristic of the operational loss data but possess the properties of outliers. A non-discriminant treatment of the bulk of the data together with these extreme events would drive the estimates of the mean and scale of the loss distribution upward, and the resulting estimates of the Value-at-Risk and Conditional Value-at-Risk measures would be unreasonably over-stated. Extreme events may occur once in 20 or 50 years, and hence involving them into the routine estimation procedures and forecasting may not always be reasonable.

This paper suggest complementing classical estimation procedures with robust statistics methods. Robust methods use a subjective or a formal routine to reject outliers. Robust statistics approach focuses on the key characteristics of the generic stream of the data.

Applying robust methods also enables a risk analyst to investigate the marginal contribution of extreme low-frequency events to various risk measures. Empirical adaptation of a robust methodology applied to operational loss data has demonstrated that the incremental contribution of extreme events that account for the highest 5% of data, stands at roughly 58-76% of the annual aggregate expected loss and the operational risk regulatory capital charge. The magnitude of the impact can serve as an important guideline for a bank to decide whether and at what price it should use insurance against the extreme losses.

Employing robust methods may signal important flaws in the models, such as outliers, poor classification of loss events, or failure of the *iid* assumption. We hope that this paper will serve as a starting point for further constructive research on robust statistics and their applications to the operational risk modeling. Finally, we would like to quote Hampel (1973): "Robust methods, in one form or another (and be it a glance at the data), are necessary; those who still don't use them are either careless or ignorant".

6 Acknowledgements

The authors greatly acknowledge helpful comments by L. Rüschendorf and D. Hinkley, and the critical suggestions provided by anonymous referees.

References

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk. Mathematical Finance 9, 203–228.

Aucremanne, L., Brys, G., Hubert, M., Rousseeuw, P. J., Struyf, A., 2004. A stufy of Belgian inflation, relative prices and nominal rigidities using new robust measures of skewness and tail weight. In: Hubert et al. (2004), pp. 13–25.

Bassett, G., Gerber, G., Rocco, P., 2004. Robust strategies for quantitative investment management. In: Hubert et al. (2004), pp. 27–37.

BIS, 2001a. Consultative document: operational risk. `www.bis.org`.

BIS, 2001b. Working paper on the regulatory treatment of operational risk. `www.bis.org`.

BIS, 2003. The 2002 loss data collection exercise for operational risk: summary of the data collected. `www.bis.org`.

BIS, 2004. International convergence of capital measurement and capital standards. `www.bis.org`.

Bowley, A. L., 1920. Elements of Statistics. Charles Scribner's Sons, New York.

Chavez-Demoulin, V., Embrechts, P., 2004. Advanced extremal models for operational risk. Tech. rep., ETH Zurich.

Chernobai, A., Burnecki, K., Rachev, S., Trück, S., Weron, R., 2006a. Modelling catastrophe claims with left-truncated severity distributions. Journal of Computational Statistics 3-2006.

Chernobai, A., Menn, C., Rachev, S., Trück, S., 2005. Estimation of operational value-at-risk in the presence of minimum collection thresholds. Tech. rep., University of California Santa Barbara.

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

Chernobai, A., Rachev, S., Fabozzi, F., 2006b. Operational Risk: A Guide to Basel II Capital Requirements, Models and Analysis. John Wiley & Sons, preprint.

Cruz, M. G., 2002. Modeling, Measuring and Hedging Operational Risk. John Wiley & Sons, New York, Chichester.

Currie, C., 2005. A Test of the Strategic Effect of Basel II Operational Risk Requirements for Banks. Tech. Rep. 141, University of Technology, Sydney.

David, H. A., 1981. Order Statistics, 2nd Edition. John Wiley & Sons, New York.

De Fontnouvelle, P., Rosengren, E., Jordan, J., 2005. Implications of alternative operational risk modelling techniques. Tech. rep., Federal Reserve Bank of Boston and FitchRisk.

Embrechts, P., Furrer, H., Kaufmann, R., 2003. Quantifying regulatory capital for operational risk. Derivatives Use, Trading, and Regulation 9 (3), 217–233.

Embrechts, P., Kaufmann, P., Samorodnitsky, G., 2004. Ruin theory revisited: Stochastic models for operational risk. in Risk Management for Central Bank Foreign Reserves (Eds. C. Bernadell et al), European Central Bank, Frankfurt, 243–261.

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

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–466.

FGG, 2005. The Risk of Operational Risk-Based Capital: Why Cost and Competitive Implications Make Basel II's Requirement Ill-Advised in the United States. Tech. rep., Financial Guardian Group.

Groeneveld, R. A., Meeden, G., 1984. Measuring skewness and kurtosis. The Statistician 33, 391–399.

Hampel, F. R., 1968. Contribution to the theory of robust estimation. Tech. rep., PhD Thesis, University of California Berkeley.

Hampel, F. R., 1973. Robust estimation: A condensed partial survery. Z. Wahrscheinlichkeitstheorie verw. Gebiete 27, 87–104.

Hampel, F. R., 1974. The influence curve and its role in robust estimation. Journal of the American Statistical Association 69, 383–393.

Hampel, F. R., 1985. The breakdown points of the mean combined with some rejection rules. Technometrics 27, 95–107.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, R. J., Stahel, W. A., 1986. Robust Statistics: The Approach Based on Influence Functions. John Wiley & Sons.

Hinkley, D. V., 1975. On power transformations to symmetry. Biometrika 62, 101–111.

Hogg, R. V., 1972. More light on the kurtosis and related statistics. Journal of the American Statistical Association 67, 422–424.

Hogg, R. V., 1974. Adaptive robust procedures: A partial review and some suggestions for future applications and theory. Journal of the American Statistical Association 69, 909–923.

Huber, P. J., 1964. Robust estimation of a location parameter. Annals of Mathematical Statistics 35, 73–101.

Huber, P. J., 1981. Robust Statistics. John Wiley & Sons, New York.

Hubert, M., Pison, G., Struyf, A., van Aelst, S. (Eds.), 2004. Theory and Applications of Recent Robust Methods. Birkhäuser, Basel.

Kendall, M. G., Stuart, A., 1977. The Advanced Theory of Statistics. Vol. 1. Charles Griffin and Co., London.

Kim, T.-H., White, H., 2003. On more robust estimation of skewness and kurtosis: Simulation and application to the S&P 500 Index. Tech. rep., University of California at San Diego.

Knez, P. J., Ready, M. J., 1997. On the robustness of size and book-to-market in cross-sectional regressions. Journal of Finance 52, 1355–1382.

Martin, R. D., Simin, T. T., 2003. Outlier resistant estimates of beta. Financial Analysts Journal 59, 56–69.

Medova, E. A., 2002. Operational risk capital allocation and integration of risks. Tech. rep., Center for Financial Research, University of Cambridge, UK.

Moors, J. J. A., 1988. A quantile alternative for kurtosis. The Statistician 37, 25–32.

Moscadelli, M., 2004. The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee. Tech. rep., Bank of Italy.

Neslehová, J., Embrechts, P., Chavez-Demoulin, V., 2006. Infinite mean models and the LDA for operational risk. Tech. rep., ETH Zurich.

Olive, D. J., 2005. Applied Robust Statistics. preprint, .

Perret-Gentil, C., Victoria-Feser, M., 2005. Robust mean-variance portfolio selection. Tech. Rep. 140, International Center for Financial Asset Management and Engineering.

Rousseeuw, P. J., Croux, C., 1993. Alternatives to the median absolute deviation. Journal of American Statistics Association 88, 1273–1283.

Rousseeuw, P. J., Leroy, A. M., 1987. Robust Regression and Outlier Detection. John Wiley & Sons, New York.

Simonoff, J. S., 1987a. The breakdown and influence properties of outlier-rejection-plus-mean procedures. Communications in Statistics: Theory and Methods 16, 1749–1769.

Simonoff, J. S., 1987b. Outlier detection and robust estimation of scale. Journal of Statistical Computation and Simulation 27, 79–92.

Stigler, S. M., 1973. Simon newcomb, percy daniell, and the history of robust estimation 1885-1920. Journal of the American Statistical Association 68, 872–878.