Estimation of Operational Value-at-Risk in the Presence of Minimum Collection Thresholds ⋆

Anna Chernobai [a,*], Christian Menn [b], Svetlozar T. Rachev [c], Stefan Trück [d]

First draft: November 28, 2004
This version: September 30, 2005
Technical Report

[a] *University of California Santa Barbara*

[b] *Cornell University and Universität Karlsruhe, Germany*

[c] *Universität Karlsruhe, Germany and University of California Santa Barbara*

[d] *Universität Karlsruhe, Germany*

Abstract

The Basel II Capital Accord of 2004 sets guidelines on operational risk capital requirements to be adopted by internationally active banks by around year-end 2007. Operational loss databases are subject to a minimum recording threshold of roughly $10,000 (internal) and $1 million (external) – an aspect often overlooked by practitioners. We provide theoretical and empirical evidence that ignoring these thresholds leads to underestimation of the VaR and CVaR figures within the Loss Distribution Approach. We emphasize that four crucial components of a reliable operational loss actuarial model are: (1) non-homogenous Poisson process for the loss arrival process, (2) flexible loss severity distributions, (3) accounting for the incomplete data, and (4) robustness analysis of the model.

Key words: Operational Risk, Truncated Data, VaR/ES Estimation and Forecasting, Robust Methods

9 October 2005

Contents

⋆ We would like to thank The Journal of Banking and Finance and the organizers of the *International Summer School in Risk Management and Control*, Rome, June 2005, where parts of this paper were presented, and the participants for their valuable suggestions and helpful comments. Parts of the research were done when A.Chernobai was visiting University of Karlsruhe, Germany, in 2004. C.Menn's research was supported by the German Academic Exchange Service (DAAD). 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).

∗ Corresponding author. Address: Department of Statistics and Applied Probability, University of California Santa Barbara, CA 93106, USA

Email address: annac@pstat.ucsb.edu (Anna Chernobai).

1 Introduction

Operational risk has been acknowledged as a major contributor to banks' risk positions and is defined as *the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events* (BIS, 2004). Current estimates suggest that the allocation of total financial risk of a bank is roughly 60% to credit, 15% to market and liquidity, and 25% to operational risk (Jorion, 2000),[1] with internationally active banks allocating annually roughly \$2-\$7 billion for operational risk (De Fountnouvelle et al., 2003). Under the Basel II Capital Accord, banks must develop a methodology to estimate the operational risk capital charge which would allow to cover damages due to potential operational loss events (BIS, 2001a, 2004). The choice of the methodology depends on a bank's business structure complexity, risk exposure, and the ability to meet required criteria. Under the Loss Distribution Approach (LDA) banks compute separately the loss severity and frequency distribution functions for each business line and risk type combination. The total capital charge is then determined by the aggregation[2] of the one year Value-at-Risk (VaR) measures across all combinations, based on the compounded losses.[3]

[1] Cruz (2002) suggests 50%, 15%, and 35%, and Crouhy et al. (2001) suggest 70%, 10%, and 20%, respectively.
[2] A simple summation, suggested in BIS (2001a), assumes a perfect correlation between different business line/ event type combinations. To avoid over-estimation of the capital charge, correlation effects must be accounted for.
[3] BIS (2001b) suggested a compound Poisson process with Lognormally distributed loss amounts. We do not restrict our attention to the homogeneity of the arrival

An accurate estimation of the operational loss frequency and severity distributions is the key to determining an optimal operational capital charge. In an ideal scenario, the data collection process results in all loss events being detected and duly recorded. However, the data recording is subject to lower recording thresholds: roughly $1 million for external databases and $10,000 for the internal (BIS, 2003). Hence, the data available for estimation are left-truncated. Some existing empirical evidence suggests that the left-truncation of the data is ignored in estimations of the loss distribution. [4] Fitting unconditional distribution to the observed (incomplete) losses would lead to biased estimates of the parameters of both severity and frequency distributions, as shown in Chernobai et al. (2005b). The resulting VaR measure would be underestimated (Chernobai et al., 2005b). The magnitude of the effect is further dependent on the threshold level (assumed constant in this paper) and the underlying loss distribution. We emphasize that under the compound Poisson process model, the severity and frequency distributions of the operational risk are inter-related: if the fraction of missing data is estimated to be non-zero, then the frequency parameter(s) requires a proportional increase. A similar model has been applied to the natural catastrophe insurance data, and it was shown that the data misspecification leads to serious underestimation of the ruin probabilities (Chernobai et al., 2005a).

The aim of this paper is two-fold. We investigate the effects of missing data on loss severity and frequency distributions and then examine the impact on the operational risk capital charge that is determined by two alternatives: the VaR and Conditional VaR measures. We also test the adequacy of several distribu-

process and the Lognormal loss distribution.

[4] See e.g., Moscadelli (2005). Lewis and Lantsman (2005) use correct loss severity specification, but do not adjust the frequency.

tions to model losses arising from different types of operational risk. The paper is organized as follows. Section 2 introduces the operational loss data problem and discusses the general methodology of treating left-truncated data. Section 3 presents the results of an empirical study with external operational loss data over 1980-2002 and examines the effects of using wrongly and correctly specified loss distributions on the capital charge, using the classical approach. Further, goodness-of-fit tests are carried out to determine an optimal law for the loss severity and frequency. It is demonstrated that ignoring the missing data leads to seriously misleading (underestimated) capital charge estimates. Section 5 applies a robust approach to examine the behavior of the bulk of the data and the marginal effects of outlying high-magnitude observations on the capital charge and forecasts. Section 6 concludes and states final remarks.

2 Truncated Compound Poisson Model for Operational Risk

2.1 Compound Poisson Process Model

The Loss Distribution Approach assumes 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 $\{S_t\}_{t\geq 0}$:

$$S_t = \sum_{k=0}^{N_t} X_k, \quad X_k \overset{\text{iid}}{\sim} F_\gamma, \tag{1}$$

in which the random sequence of loss magnitudes $\{X_k\}$ follows the distribution function (cdf) F_γ and the density f_γ, and in which the counting process N_t is assumed to take a form of a homogeneous Poisson process (HPP) with intensity $\lambda > 0$ (or a non-homogeneous Poisson process (NHPP) with intensity $\lambda(t) > 0$). F_γ belongs to a sufficiently well-behaved parametric family of continuous probability distributions, and f_γ is defined on support $\mathbb{R}_{>0}$; γ can be estimated consistently by the Maximum Likelihood Estimation (MLE) method. Depending on the distribution, γ is a parameter vector or a scalar; for simplicity, we would refer to it as a parameter throughout the paper. Independence between frequency and severity distributions is assumed. The cdf of the compound Poisson process is given by:

$$P(S_t \leq s) = \begin{cases} \sum_{n=1}^{\infty} P(N_t = n)\ F_\gamma^{n*}(s) & s > 0 \\ \\ P(N_t = 0) & s = 0 \end{cases} \tag{2}$$

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

In practice, model (1) can be used to determine the required capital charge imposed by regulators. It is measured as the $(1-\alpha) \times 100^{\text{th}}$ quantile of the cumulative loss distribution (2) over a one year period, i.e., VaR, defined as the solution to:

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

An alternative risk measure, Conditional VaR (CVaR),[5] is defined by:

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

Given a sample $\mathbf{x} = (x_1, x_2 \ldots, x_n)$ containing n losses which have occurred during some time interval $\triangle t = T_2 - T_1$, under the imposed assumptions on the structure of F_γ, the task of estimating λ and γ can be performed with the MLE principle (or, in case of a NHPP, $\lambda(t)$ is estimated by directly fitting some deterministic function):

$$\hat{\lambda}_{\text{MLE}}(x) = \frac{n}{\triangle t} \quad \text{and} \quad \hat{\gamma}_{\text{MLE}}(x) = \arg\max_\gamma \sum_{k=1}^{n} \log f_\gamma(x_k). \tag{5}$$

[5] 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.

The task of operational loss data analysis is complicated by the presence of missing data falling below the left-truncation point (the minimum collection threshold). The question addressed in subsequent analysis is whether ignoring the missing data (we call it the 'naive' approach) has a significant impact on the estimation of the frequency parameter λ (or $\lambda(t)$) and the severity parameter γ. From the statistical point of view, the estimates of parameters involved in (5) would be misleading (biased). However, in practical applications an argument that small losses can not have a significant impact on VaR that is determined by the upper rather than lower quantiles of the loss distribution, can be used to wrongly justify ignoring the threshold. Another argument says that ignoring low-magnitude data may result in an upward bias in the capital charge; as an example, De Fountnouvelle et al. (2003) suggested that fitting raw loss severity distribution to the data may lead to overestimation of the capital charge, which is in contrast to the arguments and empirical findings that we present further in this paper. In the following section we review the methodology for consistent estimation of loss and frequency distributions, as suggested in Chernobai et al. (2005b).

2.2 Estimation of Complete-Data Severity and Frequency Distributions

In the presence of missing data, the observed operational losses follow a *truncated* compound Poisson process. We use notations similar to those in Chernobai et al. (2005b). The available data set collected in the time frame $[T_1, T_2]$ is incomplete due to the non-negative pre-specified threshold u that defines a partition on $\mathbb{R}_{>0}$ through the events $A_1 = (0, u)$ and $A_2 = [u, \infty)$. Realizations of the losses in A_1 do not enter the data sample, with neither the frequency nor the severity being recorded. Realizations in A_2 are fully reported, with both the frequency and the loss amounts being specified. Hence, observations in A_1 constitute the missing data, and those in A_2 the observed, left-truncated, data. The observed sample is of the form $\mathbf{z} = (n, \mathbf{x})$, where n is the number of observations in A_2 and $\mathbf{x}$ are the values of these concrete observations. Given that the total number of observations in the complete sample is unknown, one possible *joint* density specification of $\mathbf{z}$ (with respect to the product of counting and Lebesgue measures), consistent with the model in Equation (1), is given by the following expression:

$$g_{\lambda,\gamma}(\mathbf{z}) = \frac{(\triangle t\ \tilde{\lambda})^n}{n!} e^{-\triangle t\ \tilde{\lambda}} \cdot \prod_{k=1}^{n} \frac{f_\gamma(x_k)}{q_{\gamma,2}}, \tag{6}$$

where $q_{\gamma,j}$ denotes the probability for a random realization to fall into set $A_j,\ j = 1, 2$, $\tilde{\lambda}$ is the observed intensity [6] related to the complete-data intensity λ by $\tilde{\lambda} := q_{\gamma,2} \cdot \lambda$, and $\triangle t := T_2 - T_1$ is the length of the sample window. In the representation (6), the Poisson process $\tilde{N}_t$ with intensity $\tilde{\lambda}$ that counts only the observed losses exceeding u can be thus interpreted as a *thinning* of the original process N_t with intensity λ that counts all events in the complete data sample. Due to the assumption of independence between frequency and severity, the maximization of the corresponding log-likelihood function with respect to λ and γ can be divided into two separate maximization problems, each depending on only one parameter:

$$\hat{\gamma}_{\text{MLE}} = \arg\max_{\gamma} \log g_\gamma(\mathbf{z}) = \arg\max_{\gamma} \log \left(\prod_{k=1}^{n} \frac{f_\gamma(x_k)}{q_{\gamma,2}} \right), \tag{7}$$

$$\hat{\lambda}_{\text{MLE}} = \arg\max_{\lambda} \log g_{\lambda,\hat{\gamma}_{\text{MLE}}}(\mathbf{z}) = \frac{n}{\triangle t \cdot q_{\hat{\gamma}_{\text{MLE}},2}}. \tag{8}$$

The MLE estimation of γ can be done in two ways: performing direct numerical maximization (used in this paper), or using the Expectation-Maximization algorithm (which was highlighted in our previous related study, Chernobai et al. (2005b)) developed by Dempster et al. (1977).

[6] In case of a NHPP, $\lambda \triangle t$ of a HPP is replaced with $\Lambda(t) = \int_0^t \lambda(s) ds$.

	$\mu_0 = 4$	$\mu_0 = 5$	$\mu_0 = 6.5$
$\sigma_0 = 1.5$	0.48	0.23	0.04
$\sigma_0 = 2$	0.48	0.29	0.10
$\sigma_0 = 2.7$	0.49	0.34	0.17

Table 1
Fraction of missing data, $F_{\gamma_0}(u)$, for the Lognormal(μ_0, σ_0) example with nominal threshold of $u = 50$.

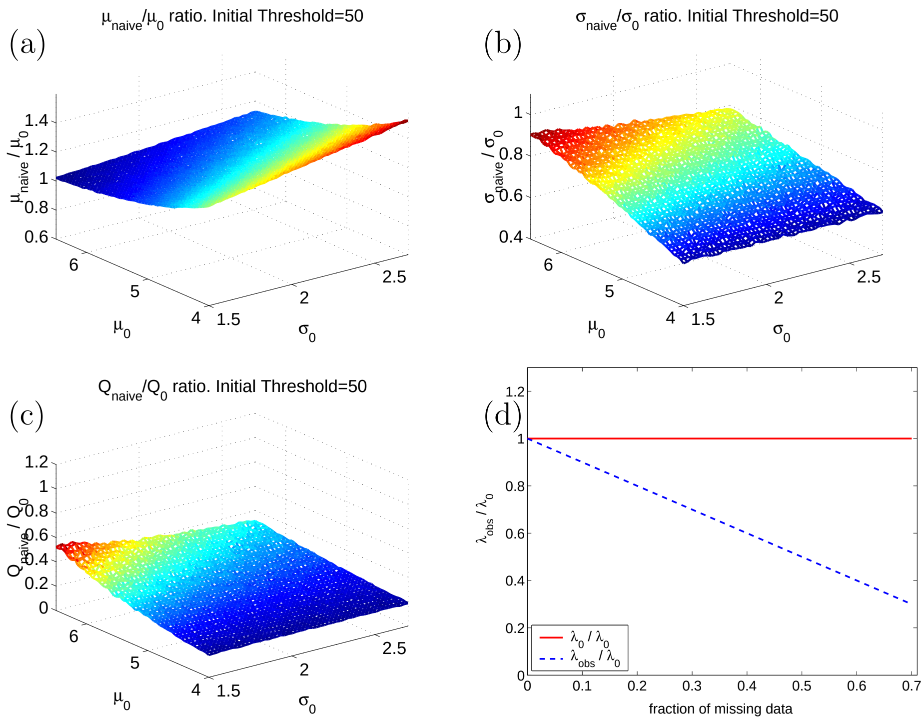

Fig. 1. Ratios of estimated parameters to the true (complete-data) parameter values for the Lognormal example under the 'naive' approach, $u = 50$.

The theoretical implications of a data misspecification on relevant parameters and VaR estimates have been discussed in Chernobai et al. (2005b). Here we limit ourselves to an illustrative example. We consider an exemplary model, defined by (1), with a Lognormal severity distribution of losses exceeding the threshold of $u = 50$ arriving as a homogeneous Poisson process. The corresponding fraction of missing data is depicted in Table 1. Figure 1 demonstrates the ratios of the estimated parameters – μ, σ, and λ – and estimated fraction of missing data, to the true parameters and fraction for a wide range of initial (complete-data) true values of μ and σ. The distance between the ratio and one represents the relative bias for each case. The ratio being equal to one indicates the absence of bias. Clearly, μ is overestimated, and σ and λ are underestimated under the 'naive' approach.

3 Application to Operational Risk Data

3.1 Purpose of Study and Data Description

The empirical section applies the model to operational risk data obtained from a major European operational loss data provider. The external database[7] is comprised of operational loss events throughout the world. The dataset used for the analysis covers losses in US$ for the time period between 1980 and 2002. The original dataset consists of five types of losses: *"Relationship"* (such as events related to legal issues, negligence, and sales-related fraud), *"Human"* (such as events related to employee errors, physical injury, and internal fraud), *"Processes"* (such as events related to business errors, supervision, security, and transactions), *"Technology"* (such as events related to technology and computer failure and telecommunications), and *"External"* (such as events related to natural and man-made disasters and external fraud). The loss amounts have been adjusted for inflation using the Consumer Price Index from the U.S. Department of Labor. In the main body of this paper, we present the results of the analysis using the *"External"* losses; all corresponding results for the remaining 4 loss types are presented in the Appendices

[7] We note that since the data are external, the estimates of the parameters and VaR and CVaR values may not be applicable to any particular bank. The purpose of the empirical study is to apply the model proposed in Section 2 and verify the effects of using wrong and correct approaches on the estimates of the capital charge. We recommend to the risk managers to apply the technique to their internal databases.

Sample Descriptive Statistics	
n	233
min ($ '000,000)	1.1
max ($ '000,000)	6384
mean ($ '000,000)	103.35
median ($ '000,000)	12.89
st.dev. ($ '000,000)	470.24
skewness	11.0320
kurtosis	140.8799

Table 2
Descriptive statistics of the *"External"* type loss data.

A (classical analysis) and B (robust analysis; see Section 5). The descriptive statistics for the *"External"* losses are presented in Table 2.

In the empirical study we focus on two scenarios. The first scenario we refer to as a 'naive' approach, in which no adjustments to the missing data are made in the analysis. The second scenario is the refined 'conditional' approach, in which the losses are modelled with truncated (conditional) distributions, given that the losses are larger than or equal $1 million (which is the lower threshold for external databases). The MLE estimates of the loss distribution parameters are obtained according to Equation (8). The frequency function's parameters of the Poisson counting process are adjusted according to Equation (7).

3.2 Operational Frequency Distributions

We consider two types of the counting process: a homogeneous (HPP) and a non-homogeneous (NHPP) Poisson process. The cumulative intensity for the HPP is computed by $\lambda \triangle t$, and the estimate for λ is obtained by simply averaging the annual total number of operational loss events. The cumulative intensity for the NHPP equals $\Lambda(t) = \int_0^t \lambda(s)ds$.

For this particular data set, visual inspection of the annually aggregated num-

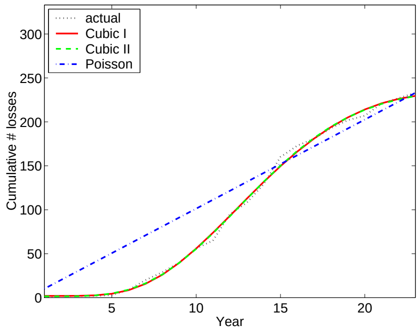

Fig. 2. Annual accumulated number of *"External"* operational losses, with fitted non-homogeneous and simple Poisson models.

ber of losses (Figure 2) suggests that the accumulation resembles a continuous cdf-like process. On the basis of this, we consider two following functions [8] for the NHPP, each with four parameters: [9]

(1) *Type I*: a Lognormal cdf-like process of form

$$\Lambda(t) = a + b \exp\left\{ -\frac{(\log t - d)^2}{2c^2} \right\} (2\pi)^{-1/2} c^{-1};$$

(2) *Type II*: a Log-Weibull cdf-like process of form

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

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

Table 3
Fitted frequency functions to the *"External"* type losses.

We obtain parameters a, b, c, d by minimizing the Mean Square Error (MSE). Table 3 demonstrates the estimated parameters and the MSE and the MAE for the cumulative intensities of the Type I, Type II, and a simple homogeneous Poisson processes with a constant intensity factor.

[8] Of course, asymptotically (as $t \to \infty$) such functions would produce a constant cumulative intensity. However, for this particular sample and this particular time frame the model appears to provide a good fit.

[9] Other deterministic functions were tried for the cumulative intensity (sinusoidal, tangent, etc.), but did not result in a good fit.

Figure 2 shows the three fits plotted together with the actual aggregated number of events. The two non-linear fits appear to be superior to the standard Poisson, and confirmed by the MSE and MAE error comparison from Table 3. In the subsequent analysis, we will assume the deterministic non-linear (Type I or II) forms for the operational loss frequency distributions, and will no longer consider the HPP case.

3.3 Operational Loss Distributions

We restrict our attention to the loss distributions that lie on the positive real half-line.[10] The following loss distributions are considered in the study:[11]

Lognormal	$\mathcal{LN}(\mu,\sigma)$	$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma x}\exp\left\{-\frac{(\log x-\mu)^2}{2\sigma^2}\right\}$ $x\geq 0,\ \mu,\sigma>0$
Weibull	$\mathcal{W}eib(\beta,\tau)$	$f_X(x)=\tau\beta x^{\tau-1}\exp\{-\beta x^\tau\}$ $x\geq 0,\ \beta,\tau>0$
Logweibull	$\log\mathcal{W}eib(\beta,\tau)$	$f_X(x)=\frac{1}{x}\tau\beta(\log x)^{\tau-1}\exp\{-\beta(\log x)^\tau\}$ $x\geq 0,\ \beta,\tau>0$
Generalized Pareto	$\mathcal{GPD}(\xi,\beta)$	$f_X(x)=\beta^{-1}(1+\xi x\beta^{-1})^{-(1+\frac{1}{\xi})}$ $x\geq 0,\ \beta>0$
Burr	$\mathcal{B}urr(\alpha,\beta,\tau)$	$f_X(x)=\tau\alpha\beta^\alpha x^{\tau-1}(\beta+x^\tau)^{-(\alpha+1)}$ $x\geq 0,\ \alpha,\beta,\tau>0$
log-αStable	$\log\mathcal{S}_\alpha(\beta,\sigma,\mu)$	$f_X(x)=\frac{g(\ln x)}{x},\ g\in\mathcal{S}_\alpha(\beta,\sigma,\mu)$ no closed-form density $x>0,\ \alpha\in(0,2),\ \beta\in[-1,1],\ \sigma,\mu>0$
Symmetric αStable	$\mathcal{S}_\alpha\mathcal{S}(\sigma)$	$f_X(x)\in\mathcal{S}_\alpha(0,\sigma,0)$, no closed-form density

[10] The exception is the symmetric αStable distribution, for which we symmetrized the data by multiplying the losses by -1 and adding them to the original data.
[11] See e.g., Samorodnitsky and Taqqu (1994) and Rachev and Mittnik (2000) on discussion and applications of the α-Stable distributions.

	γ,$F_\gamma(u)$	'Naive'	Conditional
$\mathcal{LN}$	μ	16.5789	15.7125
	σ	1.7872	2.3639
	$F_\gamma(u)$	0.0610	0.2111
$\mathcal{W}eib$	β	$1.1613\cdot10^{-4}$	0.0108
	τ	0.5175	0.2933
	$F_\gamma(u)$	0.1375	0.4629
$\log\mathcal{W}eib$	β	$3.1933\cdot10^{-12}$	$2.8169\cdot10^{-8}$
	τ	9.2660	6.2307
	$F_\gamma(u)$	0.1111	0.3016
$\mathcal{GPD}$	ξ	1.2481	1.5352
	β	$1.2588\cdot10^{7}$	$0.7060\cdot10^{7}$
	$F_\gamma(u)$	0.0730	0.1203
$\mathcal{B}urr$	α	0.0987	0.1284
	β	$2.5098\cdot10^{26}$	$3.2497\cdot10^{20}$
	τ	4.2672	3.3263
	$F_\gamma(u)$	0.0145	0.0311
$\log\mathcal{S}_\alpha$	α	1.8545	1.3313
	β	1	-1
	σ	1.1975	2.7031
	μ	16.6536	10.1928
	$F_\gamma(u)$	0.0331	0.9226
$\mathcal{S}_\alpha\mathcal{S}$	α	0.6820	0.5905
	σ	$1.1395\cdot10^{7}$	$0.7073\cdot10^{7}$
	$F_\gamma(u)$	0.0715	0.1283

Table 4
Estimated γ and $F_\gamma(u)$ values for the *"External"* type operational loss data.

Table 4 demonstrates the parameter γ values of the fitted distributions to the *"External"* data set and the estimated fraction of the missing data $F_\gamma(u)$, under the 'naive' and the conditional approaches. $F_\gamma(u)$ estimates in Table 4 indicate that under the truncated fit more weight is put on the lower magnitude losses, including the missing losses, than what is predicted by the 'naive' model. $F_\gamma(u)$ indicates the true 'information loss' due to data misspecification. We find that for all distibutions except the log-αStable the estimated fraction of data below the threshold is between 1.5 and four times higher under the conditional approach. For the log-αStable under the conditional approach the fraction of missing data is estimated to be approximately 92% and is 28 times higher than under the naive approach.

Further, under the conditional approach the location parameters (if relevant) are decreased, the scale parameters increased and the shape parameters (if relevant) decreased under the correct model, in most cases.[12] Furthermore, the change in the skewness parameter β of the log-αStable law from 1 to -1 indicates that the right tail of the loss distribution under the correct model has a near-exponential decay.

3.4 Goodness-of-Fit Tests for Operational Loss Distributions

The goal of this section is to determine which of the considered loss distributions fits the data sample best, based on the in-sample goodness-of-fit tests. We point out that the 'naive' approach produced high test statistic values and near-zero p-values for most distributions indicating the inadequacy of the methodology. Therefore, in this section we restrict ourselves only to the results on the correct conditional approach. We compare the empirical distribution function (EDF) with the fitted distribution functions. In particular, we wish to determine, whether (and to what degree) it is likely that the sample is drawn from a Lognormal, Weibull, or other considered distribution. We thus test a composite hypothesis that the EDF of the truncated sample belongs to a hypothesized truncated distribution. The null and alternative hypotheses are summarized as:

$$H_0: \; F_n(x) = \widehat{F}(x) \qquad H_A: \; F_n(x) \neq \widehat{F}(x), \tag{9}$$

where $F_n(x)$ is the empirical and $\widehat{F}(x)$ is the fitted cdf defined as:

$$\widehat{F}(x) = \begin{cases} \frac{\widehat{F}_{\theta^c}(x) - \widehat{F}_{\theta^c}(H)}{1 - \widehat{F}_{\theta^c}(H)} & x \geq H \\ 0 & x < H. \end{cases} \tag{10}$$

[12] For GPD, the shape parameter equals $1/\xi$.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	0.6504	1.2144	2.1702	316.20	0.5816	2.5993	0.0745
	[0.326]	[0.266]	[0.469]	[0.459]	[0.120]	[0.589]	[0.210]
$\mathcal{W}eib$	0.4752	0.9498	2.4314	4382.7	0.3470	5.3662	0.0337
	[0.852]	[0.726]	[0.384]	[0.108]	[0.519]	[0.164]	[0.781]
$\log \mathcal{W}eib$	0.6893	1.1020	2.2267	3130.6	0.4711	4.1429	0.0563
	[0.296]	[0.476]	[0.481]	[0.128]	[0.338]	[0.283]	[0.458]
$\mathcal{GPD}$	0.9708	1.8814	2.7742	151.94	1.7091	8.6771	0.2431
	[0.009]	[<0.005]	[0.284]	[0.949]	[<0.005]	[0.106]	[<0.005]
$\mathcal{B}urr$	1.3266	2.0385	2.8775	113.13	2.8954	15.4410	0.5137
	[0.050]	[0.048]	[0.328]	[0.989]	[0.048]	[0.064]	[0.048]
$\log \mathcal{S}_\alpha$	7.3275	7.4089	37.4863	4708.7	194.74	3132.6	24.3662
	[0.396]	[0.458]	[0.218]	[0.354]	[0.284]	[0.128]	[0.366]
$S\alpha S$	0.7222	1.4305	$1.1{\cdot}10^5$	$3.4{\cdot}10^{16}$	1.7804	$1.2{\cdot}10^{10}$	0.1348
	[0.586]	[0.339]	[0.990]	[0.797]	[0.980]	[0.841]	[0.265]

Table 5
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the loss distributions fitted to the *"External"* loss data. Smaller statistic values and greater p-values suggest better fit.

We consider seven statistics for the measure of the distance between the empirical and hypothesized cdf: Kolmogorov-Smirnov (D), Kuiper (V), supremum and quadratic Anderson-Darling (A and A^2), supremum and quadratic *"upper tail"* Anderson-Darling (A_{up} and A^2_{up}), and Cramér-von Mises (W^2). The A_{up} and A^2_{up} statistics are introduced and studied in Chernobai et al. (2005c) and are designed to put most of the weight on the upper tail. Scaling factors $\sqrt{n}$ and n were used for the supremum class and the quadratic class statistics, respectively, to make them comparable across samples of different size. The limiting distributions of the test statistics are not parameter-free, so the p-values and the critical values were obtained with Monte Carlo simulations, as described in Ross (2001). p-values suggest how likely it is that the data comes from a considered distribution, and were obtained following the four steps: (a) generate 1,000 samples from fitted distribution, of the same size as the original sample, (b) fit the distribution to each sample, (c) estimate the statistic value for each sample, and (d) find the proportion of time the statistic values from the simulated samples exceed the original statistic value.

The goodness-of-fit test statistics and the corresponding p-values for the conditional approach are presented in Table 5. Weibull and Logweibull show the best overall fit around the center of the data, based on the Kolmogorov-Smirnov and Kuiper tests, and more heavy-tailed distributions such as Burr, Pareto, and symmetric Stable, suggest the best fit around the right tail based on the "upper-tail" Anderson-Darling test. In general, the figures suggest that none of the data provides with the best overall fit: the data is best described by distributions with a moderate tail around the center and by those with a very heavy-tail around the upper tail.

3.5 Expected Loss, Value-at-Risk and Conditional Value-at-Risk

In this section, we estimate the expected aggregated loss (EL), VaR and CVaR and examine the impact of ignoring the missing data on the operational risk capital charge. We use a forward-looking approach, and use the functional form of the frequency and the parameters of the severity distribution, obtained from the historical data over the available 23 year period, to forecast EL, VaR, and CVaR one year ahead. We only consider the Type I case for the frequency, and rescale it for the conditional case using the procedure described earlier. Table 6 provides the estimates of expected loss (whenever applicable),[13] VaR, and CVaR (whenever applicable) estimates for the year 2003 for *"External"* type losses, and compares the figures obtained using the 'naive' approach and the conditional approach, obtained via 50,000 Monte Carlo samples.

[13] EL was calculated directly by $EL_t = \mathbb{E}X \times \mathbb{E}N_t$.

	EL $\times 10^{10}$	$\text{VaR}_{0.95}$ $\times 10^{10}$	$\text{VaR}_{0.99}$ $\times 10^{10}$	$\text{CVaR}_{0.95}$ $\times 10^{10}$	$\text{CVaR}_{0.99}$ $\times 10^{10}$
			$\mathcal{LN}$		
'Naive'	0.0157	0.0613	0.1697	0.1450	0.3451
Condit.	0.0327	0.1126	0.4257	0.3962	1.1617
			$\mathcal{W}eib$		
'Naive'	0.0151	0.0613	0.1190	0.0975	0.1628
Condit.	0.0208	0.0885	0.2494	0.2025	0.4509
			$\log \mathcal{W}eib$		
'Naive'	-	0.0611	0.1309	-	-
Condit.	-	0.0839	0.2489	-	-
			$\mathcal{GPD}$		
'Naive'	-	0.1190	0.8381	-	-
Condit.	-	0.2562	2.6514	-	-
			$\mathcal{B}urr$		
'Naive'	-	0.4072	8.7417	-	-
Condit.	-	0.7165	15.8905	-	-
			$\log \mathcal{S}_\alpha$		
'Naive'	-	0.1054	3.7687	-	-
Condit.	-	0.3879	0.8064	-	-
			$\mathcal{S}_\alpha\mathcal{S}$		
'Naive'	-	0.1730	1.8319	-	-
Condit.	-	0.4714	7.6647	-	-

Table 6
Estimates of EL, VaR, and CVaR for *"External"* type losses.

Table 6 indicates that EL figures appear underestimated in all cases,[14] and VaR and CVaR figures appear underestimated in most cases, whenever the 'naive' approach is used instead of the conditional. The figures also suggest that the effect is more severe for heavier-tailed distributions. In general, the EL, VaR, and CVaR estimates were 1.2 to 5 times higher under the conditional fit than under the 'naive' fit (except for log-αStable – the distribution under the conditional fit has a thinner right tail).[15] We conclude that ignoring minimum collection thresholds may lead to a substantial underestimation

[14] For some distributions (LogWeibull, GPD, Burr, log-αStable and symmetric αStable) the EL and CVaR figures were not available, because of the infinite mean due to the heavy-tailedness of the fitted distributions.
[15] Similar conclusions were made for the other four loss type categories. See Appendix A.

of the operational regulatory capital. It is strongly recommended to financial institutions to apply the correct conditional approach by using numerical maximization techniques or the Expectation-Maximization algorithm in cases when part of the data is missing.

4 Forecasting

In this section, we conduct an out-of-sample backtesting of the models. We split our data sample into two parts: (1) the first sample consists of all data points in 1980-1995 and will be used for calibration, and (2) the second sample consists of the remaining data in 1996-2002. We use the first sample and the obtained truncated loss distributions' parameter estimates to analyze our models' predicting power regarding the data belonging to the second sample. We assume that our model has a one-step ahead predicting power, with one step equal to one year (due to a scarcity of data, it would be unreasonable to use smaller intervals). In the primary step we use the data from 1980 until 1995 to conduct the forecasting about 1996 losses.

First, we estimate the unknown parameters of truncated distributions. Next, to obtain the distribution of the annually aggregated losses we repeat the following a large number (10,000) of times: use the estimated parameters to simulate N losses exceeding the \$1 million threshold, where N is the random number of losses in the year that we perform forecasting on as dictated by the fitted frequency function, and aggregate them. At each forecasting step (seven steps total) we shift the window by one year forward and repeat the above procedure. In this way we test the model for both the severity and frequency distributions. We have observed that both Type I and II models fit the data very well. For simplicity, in this section we only focus on the Type I model. Since the observed data is incomplete, we are only able to compare the forecasting power regarding the truncated (rather than complete) data.

The analysis was carried out in two parts. In part one, we compared several quantiles (25, 50, 75, 95, 99, and 99.9) of the forecasted aggregated loss distribution with the corresponding bootstrapped (non-parametric) quantiles of the realized loss distribution.[16] Table 7 presents the MSE and MAE error estimates for the forecasted quantiles relative to the corresponding bootstrapped quantiles (left) and relative to the realized total loss (middle), and the errors of the simulated relative to the actual aggregate loss (right), for the *"External"* type losses. Errors around the 25th and 75th quantiles show the errors around the central bulk of the data, 50th quantile corresponds to the median, and the errors around the highest quantiles (95, 99, and 99.9) are the errors at the far right end of the distribution. Clearly the Weibull model provides the lowest estimates for the errors, followed by the Logweibull and log-αStable models.

In the second part of the analysis, we tested the severity distribution models (without checking for the frequency) via the Likelihood Ratio (LR) test suggested in Berkowitz (2000). While non-parametric tests like the Kolmogorov-Smirnov, Kuiper, or Cramér-von Mises are rather data-intensive (Crnkovic and Drachman, 1997), the LR test is especially useful for small data samples. It is based on the following methodology. Assume that we are interested in a stochastic process $x_t,\ t > 0$, which is being forecasted at time $t-1$. Let further the probability density of x_t be $f(x_t)$ and the associated distribution function be $F(x_t) = \int_{-\infty}^{x_t} f(u)du$. To conduct the test, we estimate the parameters of the loss distribution $\hat{F}$ from the historical data in the calibration period. If $\hat{F}$

[16] The use of bootstrapping and Monte Carlo was suggested by the Basel Committee (BIS, 2001b, 2004).

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0006	0.0222	0.0016	0.0353		
50	0.0016	0.0287	0.0016	0.0295		
75	0.0088	0.0731	0.0106	0.0836	0.6723	0.1351
95	0.2182	0.4012	0.2563	0.4449		
99	2.5172	1.4166	2.7096	1.4796		
99.9	60.6060	6.9971	62.0201	7.0857		
			$\mathcal{W}eib$			
25	0.0006	0.0219	0.0015	0.0348		
50	0.0015	0.0281	0.0016	0.0289		
75	0.0065	0.0619	0.0078	0.0724	0.0360	0.0841
95	0.0780	0.2426	0.1004	0.2862		
99	0.4396	0.6138	0.5223	0.6768		
99.9	3.1658	1.6437	3.4673	1.7296		
			$\log \mathcal{W}eib$			
25	0.0006	0.0218	0.0015	0.0350		
50	0.0015	0.0282	0.0015	0.0289		
75	0.0064	0.0608	0.0078	0.0713	0.0716	0.0899
95	0.0938	0.2569	0.1177	0.3007		
99	0.6237	0.7115	0.7186	0.7750		
99.9	7.4055	2.4166	7.8048	2.5032		
			$\mathcal{GPD}$			
25	0.0007	0.0230	0.0017	0.0359		
50	0.0028	0.0349	0.0028	0.0349		
75	0.0540	0.1674	0.0059	0.1779	$0.56 \cdot 10^{10}$	309.80
95	19.5035	3.2362	19.8077	3.2804		
99	6988.76	54.7280	6995.41	54.7908		
99.9	$3.1 \cdot 10^{7}$	3054.17	$3.1 \cdot 10^{7}$	3054.26		
			$\mathcal{B}urr$			
25	0.0008	0.0244	0.0017	0.0373		
50	0.0194	0.0714	0.0109	0.0713		
75	0.5632	0.5450	0.5812	0.5555	$40.0 \cdot 10^{10}$	6604.93
95	1286.89	25.4530	1289.38	25.4968		
99	$0.2 \cdot 10^{7}$	991.46	$0.2 \cdot 10^{7}$	991.52		
99.9	$4.5 \cdot 10^{10}$	$1.4 \cdot 10^{5}$	$4.5 \cdot 10^{10}$	$1.4 \cdot 10^{5}$		
			$\log \mathcal{S}_\alpha$			
25	0.0006	0.0223	0.0015	0.0353		
50	0.0016	0.0284	0.0016	0.0291		
75	0.0064	0.0626	0.0078	0.0731	0.0844	0.0937
95	0.0899	0.2662	0.1150	0.3099		
99	0.6948	0.7844	0.7980	0.8476		
99.9	8.5024	2.7599	8.9913	2.8457		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	0.0007	0.0223	0.0015	0.0354		
50	0.0038	0.0429	0.0038	0.0427		
75	0.0782	0.2175	0.0851	0.2280	$0.03 \cdot 10^{10}$	130.60
95	26.1532	4.1342	26.5507	4.1781		
99	6397.98	63.2193	6406.84	63.2823		
99.9	$2.8 \cdot 10^{7}$	3339.51	$2.8 \cdot 10^{7}$	3339.60		

Table 7
Average forecast errors for *"External"* type aggregated losses. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss.

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log \mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log \mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.4739	0.4617	0.4751	0.4682	0.4210	0.0156	0.4768

Table 8
Averaged p-values for *"External"* aggregated losses in the 7-year forecast period.

is the correct loss distribution, then based on the so-called Rosenblatt (1952) transformation

$$y_t = \int_{-\infty}^{x_t} \hat{f}(u)du = \hat{F}(x_t) \tag{11}$$

y_t are iid and distributed uniformly on $[0, 1]$. Further, an iid standard Normal (i.e. $N(0, 1)$) series z_t can be generated from the original data x_t with:

$$z_t = \Phi^{-1}(y_t) = \Phi^{-1}\Big(\int_{-\infty}^{x_t} \hat{f}(u)du\Big). \tag{12}$$

If $\hat{F}$ is correctly specified, z_t will be iid $N(0, 1)$. To test whether the obtained series z_t is independent across observations and standard Normal, we follow Berkowitz (2000). The used test statistic is $LR = -2(l_0 - l_1)$ where l_0 and l_1 are, respectively, the log-likelihood estimates under the null parameters ($\mu = 0$ and $\sigma = 1$) and under the parameters μ_{z_t} and σ_{z_t} estimated via MLE. The p-values are obtained by referring to the χ^2 distribution table and using 2 degrees of freedom. This LR test has a number of desirable statistical properties and can be considered a powerful test even for small sample sizes (Berkowitz, 2000). Thus, we consider this methodology as an adequate method to investigate whether the realized losses have come from a particular estimated distribution.

Table 8 presents the results for the *"External"* losses. The symmetric Stable shows the highest 7-year average p-values, and the Logweibull and Weibull distributions - that gave the lowest forecast errors using the bootstrap methodology - are only slightly worse. Note that the log-αStable provides the worst forecasting results in terms of the LR test. Overall we conclude that Log-weibull and Weibull seem to be most appropriate for forecasting considered *"External"* losses.

5 Extension: Robust Approach

In 2001 the Basel Committee made the following statement: *"...data will need to be collected and robust estimation techniques (for event impact, frequency, and aggregate operational loss) will need to be developed"* (BIS, 2001a, Annex 6, p. 26). "Outlier-resistant" and "distributionally robust" – so-called *robust* – statistical analysis has found application in regression analysis, where classical estimation routines (e.g., OLS) are very sensitive to seemingly minor departures from the model assumptions. In particular, classical procedures are highly sensitive to longtailedness in the data. Improved robustness of the model can be achieved by (a) cleaning the data by applying some reasonable procedure for outlier rejection, and then (b) using classical estimation and testing procedures on the remainder of the data (Huber, 1981). Early references on robust methods include Huber (1981); a well-known study from finance on stock return anomalies is Fama and French (1992). More recent references include Hampel et al. (1986), Rousseeuw and Leroy (1987), Lawrence and Arthur (1990), Knez and Ready (1997), Martin and Simin (2003), Hubert et al. (2004), and Olive (2005).

Applying robust methods to (operational) risk management seems merely a matter of time, and lack of applications is likely to be explained by the following paradox. On the one hand, the outliers are frequently the most important part of the data (as demonstrated by the large-scale banking failures in the last two decades) – hence, they cannot be blindly discarded and should be examined to see if they follow a pattern (Olive, 2005).[17] On the other hand, outliers are 'bad' data in the sense that they deviate from the pattern set by the majority of the data (Hampel et al., 1986) and tend to obscure its generic flow. The few outlying observations may lack explanatory power regarding the majority of the data, and classical methods will frequently fit neither the bulk of the data nor the outliers well. Analysis of the full data with classical approaches may also produce unsatisfactory forecasts (Olive, 2005).

We emphasize that the robust model and the classical model (applied earlier in this paper) are not competitors: we are not advocating the use of one instead of the other – rather, we encourage the use of both models as complements to each other. We believe that in operational risk modelling it is important to conduct the robust analysis in parallel with the classical, to see the behavior associated with the bulk of the data, as well as to give consideration to the influence imposed by the extreme losses. 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).

We note that Stress Tests widely applied in the operational risk modelling have a similar goal. Under Stress Testing, 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 from the dataset using formal or informal outlier detection method,[18] with the purpose of examining VaR, forecasting, and

[17] For example, the excessive losses may be examined using Extreme Value Theory and can be modelled by the Generalized Pareto Distribution.
[18] In order to determine whether to exclude a high loss from the data, it may be

	γ,$F_\gamma(u)$	Conditional
$\mathcal{LN}$	μ	15.8095
	σ	1.9705
	$F_\gamma(u)$	0.1558
$\mathcal{W}eib$	β	0.0012
	τ	0.4178
	$F_\gamma(u)$	0.3185
$\log \mathcal{W}eib$	β	$0.21 \cdot 10^{-9}$
	τ	7.9597
	$F_\gamma(u)$	0.2254
$\mathcal{GPD}$	ξ	1.1813
	β	$7.7 \cdot 10^{6}$
	$F_\gamma(u)$	0.1132
$\mathcal{B}urr$	α	1.1642
	β	$8.6 \cdot 10^{5}$
	τ	0.8490
	$F_\gamma(u)$	0.1451
$\log \mathcal{S}_\alpha$	α	2
	β	0.4377
	σ	1.3992
	μ	15.7960
	$F_\gamma(u)$	0.1593
$\mathcal{S}_\alpha\mathcal{S}$	α	0.6598
	σ	$0.68 \cdot 10^{7}$
	$F_\gamma(u)$	0.1208

Table 9
Estimated γ and $F_\gamma(u)$ values for the conditional distributions fitted to the *"External"* operational loss data under the *robust* approach.

other properties of the main bulk of the data in the absence of these potentially impossible events. Decisions on whether to include (Stress Testing) 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, is left up to the subjective judgement of the risk expert.

We here consistently exclude the highest 5% of the dataset over the entire 1980-2002 period. We expect the proportion of the number of outliers in the calibration period (16 years, 1980-1995) to that in the testing period (7 years, 1996-

useful to look at the background information of the loss, and exclude it if there is sufficient reason to believe that a similar event would not occur in the future.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	0.8005	1.5985	2.5289	89.9172	1.2314	10.3629	0.1581
	[0.074]	[0.017]	[0.331]	[>0.995]	[<0.005]	[0.080]	[0.013]
$\mathcal{W}eib$	0.8193	1.3842	2.1208	118.00	0.8992	6.3242	0.1149
	[0.074]	[0.108]	[0.469]	[>0.995]	[0.038]	[0.102]	[0.069]
$\log \mathcal{W}eib$	0.9288	1.5545	2.3070	114.96	1.1789	6.9924	0.1550
	[0.030]	[0.034]	[0.430]	[>0.995]	[0.005]	[0.124]	[0.017]
$\mathcal{GPD}$	1.0889	2.1497	3.2082	78.7580	2.4537	14.2314	0.3238
	[<0.005]	[<0.005]	[0.193]	[>0.995]	[<0.005]	[0.063]	[<0.005]
$\mathcal{B}urr$	1.0552	2.0537	2.8205	85.6792	2.1547	13.0326	0.2922
	[0.106]	[0.005]	[0.362]	[>0.995]	[<0.005]	[0.006]	[0.019]
$\log \mathcal{S}_\alpha$	0.8213	1.5891	2.5280	89.9499	1.2311	10.4590	0.1605
	[0.046]	[0.012]	[0.302]	[>0.995]	[<0.005]	[0.088]	[0.008]
$S\alpha S$	0.8182	1.6214	3.4638	67.2664	1.7561	17.3323	0.2046
	[0.076]	[0.034]	[>0.995]	[>0.995]	[0.850]	[>0.995]	[0.020]

Table 10
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the conditional loss distributions fitted to the *"External"* loss data, under the *robust* approach. Smaller statistic values and greater p-values suggest better fit.

EL $\times 10^{10}$	$\mathrm{VaR}_{0.95}$ $\times 10^{10}$	$\mathrm{VaR}_{0.99}$ $\times 10^{10}$	$\mathrm{CVaR}_{0.95}$ $\times 10^{10}$	$\mathrm{CVaR}_{0.99}$ $\times 10^{10}$
		$\mathcal{LN}$		
0.0154	0.0580	0.1642	0.1397	0.3334
		$\mathcal{W}eib$		
0.0088	0.0354	0.0715	0.0599	0.1066
		$\log \mathcal{W}eib$		
-	0.0395	0.0865	-	-
		$\mathcal{GPD}$		
-	0.0943	0.5604	-	-
		$\mathcal{B}urr$		
-	0.0676	0.3246	-	-
		$\log \mathcal{S}_\alpha$		
-	0.0570	0.1695	-	-
		$\mathcal{S}_\alpha\mathcal{S}$		
-	0.2234	2.4408	-	-

Table 11
Estimates of EL, VaR, and CVaR for *"External"* type losses, under the *robust* conditional approach.

2002) to be 0.7:0.3. For the *"External"* losses the proportion was 0.75:0.25. For the remaining 4 loss types the proportions were 0.58:0.42, 0.56:0.44, 0.67:0.33,

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log \mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log \mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
$\frac{\triangle \mathrm{VaR}_{0.95}}{\mathrm{VaR}_{0.95}^{class.}} \times 100$	48%	60%	53%	63%	91%	85%	53%
$\frac{\triangle \mathrm{VaR}_{0.99}}{\mathrm{VaR}_{0.99}^{class.}} \times 100$	61%	71%	65%	79%	98%	79%	68%

Table 12
Sensitivity of classical VaR to outliers for *"External"* type aggregated losses. Figures indicate incremental VaR as the percentage of classical VaR attributed to top 5% of data.

Sample Descriptive Statistics	
n	221
min ($ '000,000)	1.1
max ($ '000,000)	364.80
mean ($ '000,000)	39.7515
median ($ '000,000)	11.40
st.dev. ($ '000,000)	63.84
skewness	2.5635
kurtosis	10.0539

Table 13
Descriptive statistics of the *"External"* type loss data under the *robust* approach.

and 1:0, for *"Relationship"*, *"Human"*, *"Processes"*, and *"Technology"*, respectively. Descriptive statistics for the *"External"* losses are depicted in Table 13. Notice significant reduction in the mean, standard deviation, skewness and kurtosis coefficients – all indicate that the trimmed data set is likely to be less heavy-tailed.

We reproduce the results for the parameter estimates, goodness-of-fit tests, EL, VaR, and CVaR measures, [19] together with out-of-sample goodness-of-fit tests for the *"External"* type losses. All corresponding results for the other 4 types of losses can be found in Appendix B.

[19] Note that many of the EL and CVaR measures produce infinite figures (denoted by "-"). One approach may be to consider doubly-truncated loss distributions, further truncated from above at a fixed level. This ensures the existence of a finite mean and variance. A possible choice of the truncation point is the total value of assets, in consideration of the *limited liability* by banks. Such analysis in not included in this study.

It is clear from Table 9 that more weight is put on the low and medium-size losses, as can be concluded from decreased location and shape parameters and increased scale parameters, and a higher fraction of missing data. The log-Stable distribution's shape parameter has increased to 2, indicating a thinner-tailed law, comparable to that of Lognormal. From Table 10 we can conclude that medium-tailed distributions such as Lognormal and Weibull fit the trimmed data well, in contrast to heavy-tailed laws in the case when all data is included into the analysis. Table 11 indicates that the EL, VaR, and CVaR estimates have considerably dropped. The sensitivity of classical VaR to the highest 5% of the loss data in shown in Table 12: for a reasonable choice of the loss distribution,[20] roughly 53% to 60% of the classical 95% VaR, and roughly 65% to 71% of the classical 99% VaR, are attributed to the extreme observations.

Considering the forecasts of future losses, Table 14 demonstrates that the accuracy of the forecasts has remarkably improved. For all distributions forecasted quantiles of the loss distribution are much closer to the bootstrapped quantiles and actual losses. This is especially true for the high quantiles, as expected, since extreme losses were excluded from the analysis. We conclude that with the robust approach the general tendency of the losses is captured adequately. Further, the approach shows the vast influence of extreme losses on operational VaR and the sensitivity of the risk measures VaR and CVaR to the biggest losses in the data. Considering the choice of the right distribution, both tests (forecast error estimates, and the LR test) converge in their indication of the best model with the in-sample goodness of fit tests: the robust

[20] We have previously concluded (and confirm further by the robust methodology) that Logweibull and Weibull models are the most appropriate candidates for modelling *"External"* type losses.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0002	0.0143	0.0004	0.0173		
50	0.0004	0.0171	0.0004	0.0171		
75	0.0012	0.0256	0.0017	0.0326	0.2235	0.0486
95	0.0179	0.1146	0.0254	0.1437		
99	0.1615	0.3686	0.1946	0.4108		
99.9	2.6093	1.4848	2.7830	1.5420		
			$\mathcal{W}eib$			
25	0.0002	0.0145	0.0004	0.0169		
50	0.0003	0.0157	0.0003	0.0157		
75	0.0006	0.0199	0.0009	0.0238	0.0019	0.0285
95	0.0037	0.0484	0.0071	0.0752		
99	0.0141	0.1041	0.0244	0.1464		
99.9	0.0591	0.2240	0.0868	0.2810		
			$\log \mathcal{W}eib$			
25	0.0002	0.0142	0.0004	0.0169		
50	0.0003	0.0153	0.0003	0.0153		
75	0.0006	0.0195	0.0009	0.0232	0.0021	0.0285
95	0.0037	0.0480	0.0071	0.0760		
99	0.0149	0.1097	0.0256	0.1518		
99.9	0.0775	0.2561	0.1084	0.3130		
			$\mathcal{GPD}$			
25	0.0002	0.0143	0.0005	0.0182		
50	0.0006	0.0205	0.0006	0.0205		
75	0.0047	0.0486	0.0059	0.0577	$0.23 \cdot 10^5$	1.6344
95	0.6101	0.5945	0.6453	0.6236		
99	45.2582	5.1661	45.7041	5.2080		
99.9	$0.3 \cdot 10^5$	124.80	$0.3 \cdot 10^5$	124.85		
			$\mathcal{B}urr$			
25	0.0003	0.0161	0.0005	0.0183		
50	0.0019	0.0344	0.0019	0.0344		
75	0.0568	0.1692	0.0611	0.1785	$0.31 \cdot 10^{10}$	401.06
95	36.6345	4.1927	36.8775	4.2218		
99	$0.2 \cdot 10^5$	90.7056	$0.2 \cdot 10^5$	90.7477		
99.9	$2.8 \cdot 10^8$	$0.1 \cdot 10^5$	$2.8 \cdot 10^8$	$0.1 \cdot 10^5$		
			$\log \mathcal{S}_\alpha$			
25	0.0002	0.0141	0.0004	0.0176		
50	0.0004	0.0168	0.0004	0.0168		
75	0.0012	0.0253	0.0016	0.0307	0.0154	0.0416
95	0.0153	0.0972	0.0215	0.1262		
99	0.1256	0.2958	0.1510	0.3377		
99.9	1.7147	1.0895	1.8363	1.1464		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	0.0003	0.0176	0.0005	0.0174		
50	0.0014	0.0305	0.0014	0.0305		
75	0.0200	0.1080	0.0229	0.1171	$2.34 \cdot 10^7$	29.9384
95	3.5661	1.5601	3.6621	1.5892		
99	620.34	19.6938	622.08	19.7359		
99.9	$12.5 \cdot 10^5$	812.08	$12.5 \cdot 10^5$	812.14		

Table 14
Average forecast errors for *"External"* type aggregated losses, under the *robust* approach. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss.

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log\mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log\mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.4876	0.5131	0.5290	0.4730	0.4395	0.3272	0.4704

Table 15
Averaged p-values for *"External"* type aggregated losses in the 7-year forecast period under the *robust* approach.

approach confirms that the Logweibull distribution has the best forecasting power for the *"External"* type loss, with Weibull being the second best choice.

6 Conclusions

In this study we proposed and empirically investigated a methodology for consistent estimation of the loss and frequency distributions for the assumed actuarial model of operational losses in the presence of minimum collection thresholds. The analysis was conducted using five types of loss data (*"Relationship"*, *"Human"*, *"Processes"*, *"Technology"*, and *"External"*) obtained from an operational risk loss database.

Our findings demonstrated that ignoring such minimum thresholds leads to severe biases in corresponding parameter estimates whenever the thresholds are ignored. As a consequence, EL, VaR, and CVaR are underestimated and are generally 1.2 to 5 times higher under the conditional approach, in which truncated loss distributions were fitted to the loss data and frequency was adjusted to account for information loss. A variety of goodness-of-fit measures were used to test the adequacy of different loss distributions. For the *"External"* type losses the Logweibull and Weibull distributions showed the best overall fit, while more heavy-tailed distributions such as Burr, Pareto, and symmetric αStable, better fit the upper tail, supporting the conjecture that the operational loss data is severely heavy-tailed. The forecasts based on the classical approach supported moderately heavy-tailed distributions.

A robust statistics approach was introduced. Excluding outliers allows to see the behavior of the bulk of the data, and comparison with the classical approach allows to examine the sensitivity of risk measures and forecasts to the tail events. Roughly 53-60% of the 95% VaR, and 65-71% of the 99% VaR, is attributed to the highest 5% losses. The robust methodology also resulted in significantly smaller forecast errors for the trimmed data, suggesting that the general tendency of the losses is well captured by the approach. It further confirmed the choice of the loss distributions from the classical approach.

7 APPENDIX A: Classical Analysis

Sample Descriptive Statistics				
	Relationship	Human	Processes	Technology
n	849	813	325	67
min ($ ’000,000)	1.07	1.10	1.10	1.13
max ($ ’000,000)	6,480	23,630	13,334	830
mean ($ ’000,000)	89.86	138.47	285.55	77.43
median ($ ’000,000)	14.63	12.32	39.98	11.60
st.dev. ($ ’000,000)	360.45	901.51	955.52	136.65
skewness	11.6429	22.2416	9.1070	3.1761
kurtosis	169.9732	570.1188	112.5151	15.7230

Table 16
Descriptive statistics for the 4 loss types data. *Note:* due to the small sample size of *“Technology”* losses, all empirical results should be treated with caution.

loss/process					MSE	MAE
“Relationship”						
Cubic I	***a***	***b***	***c***	***d***		
	34.13	1364.82	0.63	3.32	76.57	7.05
Cubic II	***a***	***b***	***c***	***d***		
	930.29	896.17	0.0010	6.82	69.08	6.57
Poisson				λ		
				36.91	5907.45	65.68
“Human”						
Cubic I	***a***	***b***	***c***	***d***		
	33.49	1436.56	0.65	3.43	68.05	6.89
Cubic II	***a***	***b***	***c***	***d***		
	950.20	917.11	0.0008	6.80	61.59	6.60
Poisson				λ		
				35.35	6600.38	65.33
“Processes”						
Cubic I	***a***	***b***	***c***	***d***		
	9.44	2098.96	1.04	4.58	22.50	3.64
Cubic II	***a***	***b***	***c***	***d***		
	2034.25	2024.77	0.0007	4.79	23.06	3.65
Poisson				λ		
				14.13	1664.82	36.57
“Technology”						
Cubic I	***a***	***b***	***c***	***d***		
	0.79	120.20	0.58	3.47	3.71	1.28
Cubic II	***a***	***b***	***c***	***d***		
	137.68	138.39	0.0006	6.32	4.89	1.67
Poisson				λ		
				3.35	217.04	13.42

Table 17
Fitted frequency functions to the operational losses.

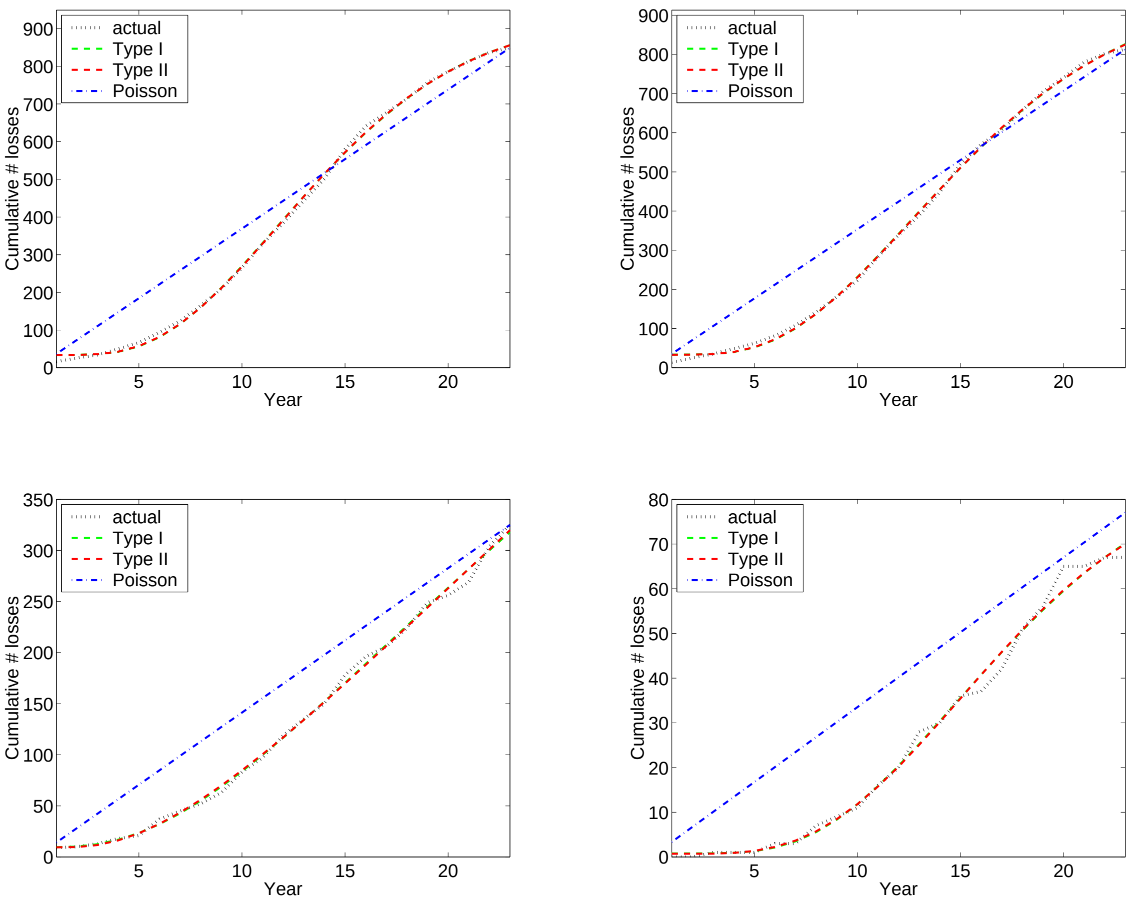

Fig. 3. Fitted frequency functions to the operational losses. Top left: *"Relationship"*, top right: *"Human"*, bottom left: *"Processes"*, bottom right: *"Technology"*.

	γ,$F_\gamma(u)$	'Naive'	Conditional
$\mathcal{LN}$	μ	16.6771	16.1911
	σ	1.6956	2.0654
	$F_\gamma(u)$	0.0457	0.1250
$\mathcal{W}eib$	β	$6.1038\cdot10^{-5}$	0.0032
	τ	0.5528	0.3538
	$F_\gamma(u)$	0.1189	0.3479
$\log\mathcal{W}eib$	β	$0.5128\cdot10^{-12}$	$0.2694\cdot10^{-8}$
	τ	9.8946	7.0197
	$F_\gamma(u)$	0.0938	0.2386
$\mathcal{GPD}$	ξ	1.0882	1.2852
	β	$1.5516\cdot10^{7}$	$1.0558\cdot10^{7}$
	$F_\gamma(u)$	0.0604	0.0855
$\mathcal{B}urr$	α	0.4817	5.1242
	β	$3.4832\cdot10^{9}$	$1.0221\cdot10^{4}$
	τ	1.4077	0.4644
	$F_\gamma(u)$	0.0365	0.2575
$\log\mathcal{S}_\alpha$	α	1.9097	1.9340
	β	1	-1
	σ	1.1584	1.5198
	μ	16.7182	15.9616
	$F_\gamma(u)$	0.0303	0.1742
$\mathcal{S}_\alpha\mathcal{S}$	α	0.7377	0.6592
	σ	$1.3695\cdot10^{7}$	$0.9968\cdot10^{7}$
	$F_\gamma(u)$	0.0558	0.0841

Table 18
Estimated γ and $F_\gamma(u)$ values for the *"Relationship"* type operational loss data.

	γ,$F_\gamma(u)$	'Naive'	Conditional
$\mathcal{LN}$	μ	16.5878	15.4627
	σ	1.8590	2.5642
	$F_\gamma(u)$	0.0679	0.2603
$\mathcal{W}eib$	β	0.0002	0.0240
	τ	0.4841	0.2526
	$F_\gamma(u)$	0.1501	0.5441
$\log\mathcal{W}eib$	β	$14.3254 \cdot 10^{-12}$	$30.7344 \cdot 10^{-8}$
	τ	9.8946	7.0197
	$F_\gamma(u)$	0.1221	0.3718
$\mathcal{GPD}$	ξ	1.3761	1.6562
	β	$1.1441 \cdot 10^{7}$	$0.6135 \cdot 10^{7}$
	$F_\gamma(u)$	0.0792	0.1344
$\mathcal{B}urr$	α	0.0938	0.0922
	β	$5.1819 \cdot 10^{27}$	$2.8463 \cdot 10^{27}$
	τ	4.4823	4.4717
	$F_\gamma(u)$	0.0131	0.0195
$\log\mathcal{S}_\alpha$	α	1.6294	1.4042
	β	1	-1
	σ	1.1395	2.8957
	μ	16.8464	10.5108
	$F_\gamma(u)$	0.0083	0.8793
$\mathcal{S}_\alpha\mathcal{S}$	α	0.6724	0.6061
	σ	$1.1126 \cdot 10^{7}$	$0.7143 \cdot 10^{7}$
	$F_\gamma(u)$	0.0742	0.1241

Table 19
Estimated γ and $F_\gamma(u)$ values for the *"Human"* type operational loss data.

	γ,$F_\gamma(u)$	'Naive'	Conditional
$\mathcal{LN}$	μ	17.5163	17.1600
	σ	2.0215	2.3249
	$F_\gamma(u)$	0.0336	0.0751
$\mathcal{W}eib$	β	0.0001	0.0021
	τ	0.4938	0.3515
	$F_\gamma(u)$	0.0923	0.2338
$\log\mathcal{W}eib$	β	$2.4894{\cdot}10^{-12}$	$0.1091{\cdot}10^{-8}$
	τ	9.1693	7.1614
	$F_\gamma(u)$	0.0687	0.1479
$\mathcal{GPD}$	ξ	1.4754	1.6147
	β	$2.9230{\cdot}10^{7}$	$2.2886{\cdot}10^{7}$
	$F_\gamma(u)$	0.0328	0.0413
$\mathcal{B}urr$	α	0.8661	14.3369
	β	$4.3835{\cdot}10^{6}$	$1.1987{\cdot}10^{4}$
	τ	0.8884	0.3829
	$F_\gamma(u)$	0.0405	0.2097
$\log\mathcal{S}_\alpha$	α	2.0000	2.0000
	β	0.9697	0.8195
	σ	1.4294	1.6476
	μ	17.5163	17.1535
	$F_\gamma(u)$	0.0336	0.0760
$\mathcal{S}_\alpha\mathcal{S}$	α	0.5902	0.5478
	σ	$2.7196{\cdot}10^{7}$	$1.9925{\cdot}10^{7}$
	$F_\gamma(u)$	0.0358	0.0536

Table 20
Estimated γ and $F_\gamma(u)$ values for the *"Processes"* type operational loss data.

	γ,$F_\gamma(u)$	'Naive'	Conditional
$\mathcal{LN}$	μ	16.6176	15.1880
	σ	1.9390	2.7867
	$F_\gamma(u)$	0.0742	0.3112
$\mathcal{W}eib$	β	$6.3668{\cdot}10^{-5}$	0.0103
	τ	0.5490	0.2938
	$F_\gamma(u)$	0.1177	0.4485
$\log \mathcal{W}eib$	β	$1.9309{\cdot}10^{-12}$	$11.0647{\cdot}10^{-8}$
	τ	9.4244	5.7555
	$F_\gamma(u)$	0.1023	0.3329
$\mathcal{GPD}$	ξ	1.5823	2.0925
	β	$1.0470{\cdot}10^{7}$	$0.3446{\cdot}10^{7}$
	$F_\gamma(u)$	0.0851	0.2029
$\mathcal{B}urr$	α	0.0645	0.0684
	β	$1.7210{\cdot}10^{35}$	$8.7406{\cdot}10^{20}$
	τ	5.8111	5.2150
	$F_\gamma(u)$	0.0227	0.8042
$\log \mathcal{S}_\alpha$	α	2.0000	2.0000
	β	0.7422	0.8040
	σ	1.3715	1.9894
	μ	16.6181	15.1351
	$F_\gamma(u)$	0.0747	0.3195
$\mathcal{S}_\alpha\mathcal{S}$	α	0.1827	0.1827
	σ	$0.1676{\cdot}10^{7}$	$0.1676{\cdot}10^{7}$
	$F_\gamma(u)$	0.3723	0.3723

Table 21
Estimated γ and $F_\gamma(u)$ values for the *"Technology"* type operational loss data.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	0.8056	1.3341	2.6094	875.40	0.7554	4.6122	0.1012
	[0.082]	[0.138]	[0.347]	[0.593]	[0.043]	[0.401]	[0.086]
$\mathcal{W}eib$	0.5553	1.0821	3.8703	$2.7{\cdot}10^4$	0.7073	13.8191	0.0716
	[0.625]	[0.514]	[0.138]	[0.080]	[0.072]	[0.081]	[0.249]
$\log\mathcal{W}eib$	0.5284	1.0061	3.0718	7332.1	0.4682	5.2316	0.0479
	[0.699]	[0.628]	[0.255]	[0.186]	[0.289]	[0.282]	[0.514]
$\mathcal{GPD}$	1.4797	2.6084	3.5954	374.68	3.7165	22.1277	0.5209
	[<0.005]	[<0.005]	[0.154]	[>0.995]	[<0.005]	[0.048]	[<0.005]
$\mathcal{B}urr$	1.3673	2.4165	3.3069	371.65	3.1371	22.0374	0.4310
	[0.032]	[<0.005]	[0.309]	[0.960]	[<0.005]	[0.019]	[0.011]
$\log\mathcal{S}_\alpha$	1.5929	1.6930	3.8184	1075.3	3.8067	10.1990	0.7076
	[0.295]	[0.295]	[0.275]	[0.041]	[0.290]	[0.288]	[0.292]
$S\alpha S$	1.1634	2.0695	$1.4{\cdot}10^5$	$5.0{\cdot}10^{16}$	4.4723	$2.6{\cdot}10^{14}$	0.3630
	[0.034]	[<0.005]	[>0.995]	[0.971]	[0.992]	[<0.005]	[<0.005]

Table 22
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the loss distributions fitted to the *"Relationship"* loss data. Smaller statistic values and greater p-values suggest better fit.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	0.8758	1.5265	3.9829	1086.2	0.7505	4.5160	0.0804
	[0.032]	[0.039]	[0.126]	[0.462]	[0.044]	[0.408]	[0.166]
$\mathcal{W}eib$	0.8065	1.5439	4.3544	$3.2{\cdot}10^4$	0.7908	8.6610	0.0823
	[0.103]	[0.051]	[0.095]	[0.068]	[0.068]	[0.112]	[0.188]
$\log\mathcal{W}eib$	0.9030	1.5771	4.1343	$1.1{\cdot}10^4$	0.7560	4.5125	0.0915
	[0.074]	[0.050]	[0.115]	[0.160]	[0.115]	[0.392]	[0.217]
$\mathcal{GPD}$	1.4022	2.3920	3.6431	374.68	2.7839	23.7015	0.3669
	[<0.005]	[<0.005]	[0.167]	[>0.995]	[<0.005]	[0.051]	[<0.005]
$\mathcal{B}urr$	2.2333	3.1970	4.7780	255.91	7.0968	46.3417	1.2830
	[0.115]	[0.115]	[0.174]	[>0.995]	[0.115]	[0.119]	[0.115]
$\log\mathcal{S}_\alpha$	9.5186	9.5619	36.2617	9846.3	304.61	4198.9	44.5156
	[0.319]	[0.324]	[0.250]	[0.354]	[0.312]	[0.215]	[0.315]
$S\alpha S$	1.1628	2.1537	$5.8{\cdot}10^5$	$4.3{\cdot}10^{17}$	11.9320	$3.3{\cdot}10^{11}$	0.2535
	[0.352]	[0.026]	[0.651]	[0.351]	[0.971]	[0.436]	[0.027]

Table 23
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the loss distributions fitted to the *"Human"* loss data. Smaller statistic values and greater p-values suggest better fit.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	0.6584	1.1262	2.0668	272.61	0.4624	4.0556	0.0603
	[0.297]	[0.345]	[0.508]	[0.768]	[0.223]	[0.367]	[0.294]
$\mathcal{W}eib$	0.6110	1.0620	1.7210	2200.7	0.2069	2.2340	0.0338
	[0.455]	[0.532]	[0.766]	[0.192]	[0.875]	[0.758]	[0.755]
$\log \mathcal{W}eib$	0.5398	0.9966	1.6238	658.42	0.1721	1.4221	0.0241
	[0.656]	[0.637]	[0.832]	[0.343]	[0.945]	[0.977]	[0.918]
$\mathcal{GPD}$	1.0042	1.9189	4.0380	148.24	2.6022	13.1082	0.3329
	[0.005]	[<0.005]	[0.128]	[>0.995]	[<0.005]	[0.087]	[<0.005]
$\mathcal{B}urr$	0.5634	0.9314	1.6075	364.08	0.2639	2.0133	0.0323
	[0.598]	[0.800]	[0.841]	[0.429]	[0.794]	[0.844]	[0.840]
$\log \mathcal{S}_\alpha$	0.6931	1.1490	2.0109	272.57	0.4759	4.0910	0.0660
	[0.244]	[0.342]	[0.534]	[0.786]	[0.202]	[0.361]	[0.258]
$S\alpha S$	1.3949	1.9537	$3.3{\cdot}10^5$	$2.5{\cdot}10^{17}$	6.5235	$6.8{\cdot}10^{14}$	0.3748
	[0.085]	[0.067]	[0.931]	[0.530]	[0.964]	[0.193]	[0.102]

Table 24
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the loss distributions fitted to the *"Processes"* loss data. Smaller statistic values and greater p-values suggest better fit.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	1.1453	1.7896	2.8456	41.8359	1.3778	6.4213	0.2087
	[<0.005]	[0.005]	[0.209]	[0.990]	[<0.005]	[0.067]	[<0.005]
$\mathcal{W}eib$	1.0922	1.9004	2.6821	52.5269	1.4536	4.8723	0.2281
	[<0.005]	[<0.005]	[0.216]	[0.944]	[<0.005]	[0.087]	[<0.005]
$\log \mathcal{W}eib$	1.1099	1.9244	2.7553	49.2373	1.5355	5.2992	0.2379
	[<0.005]	[<0.005]	[0.250]	[0.976]	[<0.005]	[0.114]	[<0.005]
$\mathcal{GPD}$	1.2202	1.8390	3.0843	33.4298	1.6182	8.8484	0.2408
	[<0.005]	[<0.005]	[0.177]	[>0.995]	[<0.005]	[0.067]	[<0.005]
$\mathcal{B}urr$	1.1188	0.9374	2.6949	28.4827	2.0320	10.5469	0.3424
	[0.389]	[0.380]	[0.521]	[>0.995]	[0.380]	[0.401]	[0.380]
$\log \mathcal{S}_\alpha$	1.1540	1.7793	2.8728	41.7454	1.3646	6.4919	0.2071
	[<0.005]	[0.007]	[0.208]	[0.995]	[<0.005]	[0.060]	[<0.005]
$S\alpha S$	2.0672	2.8003	$2.7{\cdot}10^5$	$3.6{\cdot}10^{16}$	19.6225	$7.2{\cdot}10^{10}$	1.4411
	[>0.995]	[>0.995]	[>0.995]	[>0.995]	[>0.995]	[>0.995]	[0.964]

Table 25
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the loss distributions fitted to the *"Technology"* loss data. Smaller statistic values and greater p-values suggest better fit.

	EL	$\mathrm{VaR}_{0.95}$	$\mathrm{VaR}_{0.99}$	$\mathrm{CVaR}_{0.95}$	$\mathrm{CVaR}_{0.99}$
			$\mathcal{LN}$		
'Naive'	0.1105	0.2832	0.5386	0.4662	0.8685
Condit.	0.1634	0.4662	1.0644	0.9016	1.9091
			$\mathcal{W}eib$		
'Naive'	0.1065	0.2203	0.2996	0.2700	0.3505
Condit.	0.1284	0.3187	0.5121	0.4430	0.6689
			$\log \mathcal{W}eib$		
'Naive'	-	0.2235	0.3193	-	-
Condit.	-	0.3332	0.5902	-	-
			$\mathcal{GPD}$		
'Naive'	-	0.8240	4.1537	-	-
Condit.	-	1.5756	11.3028	-	-
			$\mathcal{B}urr$		
'Naive'	-	2.8595	31.5637	-	-
Condit.	-	1.5713	11.5519	-	-
			$\log \mathcal{S}_\alpha$		
'Naive'	-	1.9124	7488.08	-	-
Condit.	-	0.4359	0.9557	-	-
			$\mathcal{S}_\alpha\mathcal{S}$		
'Naive'	-	2.1873	17.3578	-	-
Condit.	-	4.5476	56.2927	-	-

Table 26
Estimates of expected aggregated loss, VaR and CVaR (figures must be further scaled $\times 10^{10}$) for *"Relationship"* type losses. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

	EL	$\mathrm{VaR}_{0.95}$	$\mathrm{VaR}_{0.99}$	$\mathrm{CVaR}_{0.95}$	$\mathrm{CVaR}_{0.99}$
			$\mathcal{LN}$		
'Naive'	0.1981	0.4970	0.9843	0.8534	1.6652
Condit.	0.4171	1.2161	3.4190	3.3869	9.4520
			$\mathcal{W}eib$		
'Naive'	0.1993	0.4017	0.5507	0.4945	0.6456
Condit.	0.2881	0.7997	1.5772	1.3232	2.3746
			$\log \mathcal{W}eib$		
'Naive'	-	0.4174	0.6184	-	-
Condit.	-	0.8672	1.8603	-	-
			$\mathcal{GPD}$		
'Naive'	-	3.9831	33.5741	-	-
Condit.	-	12.1150	168.64	-	-
			$\mathcal{B}urr$		
'Naive'	-	85.5620	2690.44	-	-
Condit.	-	94.8281	3042.32	-	-
			$\log \mathcal{S}_\alpha$		
'Naive'	-	$1.9 \cdot 10^7$	$7.2 \cdot 10^{24}$	-	-
Condit.	-	2.2737	4.2319	-	-
			$\mathcal{S}_\alpha\mathcal{S}$		
'Naive'	-	6.2811	77.4762	-	-
Condit.	-	14.5771	203.24	-	-

Table 27
Estimates of expected aggregated loss, VaR and CVaR (figures must be further scaled $\times 10^{10}$) for *"Human"* type losses. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

	EL	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$
			$\mathcal{LN}$		
'Naive'	0.5622	1.5508	3.5665	3.1201	6.9823
Condit.	0.8457	2.5610	6.5625	5.7823	13.9079
			$\mathcal{W}eib$		
'Naive'	0.4170	0.8800	1.2102	1.0891	1.4311
Condit.	0.5131	1.2761	2.1308	1.8257	2.8578
			$\log \mathcal{W}eib$		
'Naive'	-	0.9611	1.4498	-	-
Condit.	-	1.4780	2.6511	-	-
			$\mathcal{GPD}$		
'Naive'	-	12.5930	131.25	-	-
Condit.	-	20.8700	262.52	-	-
			$\mathcal{B}urr$		
'Naive'	-	6.8569	52.0391	-	-
Condit.	-	1.7987	4.1859	-	-
			$\log \mathcal{S}_\alpha$		
'Naive'	-	1.5613	3.5159	-	-
Condit.	-	2.5394	6.7070	-	-
			$\mathcal{S}_\alpha\mathcal{S}$		
'Naive'	-	38.7627	529.99	-	-
Condit.	-	74.9073	1280.02	-	-

Table 28
Estimates of expected aggregated loss, VaR and CVaR (figures must be further scaled $\times 10^{10}$) for *"Process"* type losses. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

	EL	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$
			$\mathcal{LN}$		
'Naive'	0.0324	0.1202	0.3593	0.2970	0.7303
Condit.	0.0958	0.2898	1.2741	1.5439	5.4865
			$\mathcal{W}eib$		
'Naive'	0.0226	0.0798	0.1368	0.1159	0.1795
Condit.	0.0358	0.1454	0.3625	0.2958	0.6180
			$\log \mathcal{W}eib$		
'Naive'	-	0.0861	0.1683	-	-
Condit.	-	0.1670	0.4747	-	-
			$\mathcal{GPD}$		
'Naive'	-	0.4415	5.6954	-	-
Condit.	-	1.6249	54.4650	-	-
			$\mathcal{B}urr$		
'Naive'	-	2.8840	158.94	-	-
Condit.	-	9.0358	855.78	-	-
			$\log \mathcal{S}_\alpha$		
'Naive'	-	0.1222	0.3560	-	-
Condit.	-	0.2990	1.2312	-	-
			$\mathcal{S}_\alpha\mathcal{S}$		
'Naive'	-	$4.9{\cdot}10^5$	$3.2{\cdot}10^9$	-	-
Condit.	-	$7.1{\cdot}10^6$	$6.9{\cdot}10^{10}$	-	-

Table 29
Estimates of expected aggregated loss, VaR and CVaR (figures must be further scaled $\times 10^{10}$) for *"Technology"* type losses. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0038	0.0485	0.0302	0.1310		
50	0.0127	0.0823	0.0155	0.0891		
75	0.0260	0.1219	0.0155	0.1115	0.1357	0.1812
95	0.1342	0.3467	0.2140	0.4442		
99	0.8866	0.8618	1.4125	1.1760		
99.9	13.6999	3.6286	16.8095	4.0731		
			$\mathcal{W}eib$			
25	0.0018	0.0390	0.0017	0.0367		
50	0.0026	0.0446	0.0025	0.0439		
75	0.0039	0.0509	0.0055	0.0556	0.0052	0.0552
95	0.0083	0.0729	0.0170	0.1181		
99	0.0151	0.1069	0.0340	0.1733		
99.9	0.0288	0.1583	0.0667	0.2498		
			$\log \mathcal{W}eib$			
25	0.0036	0.0462	0.0295	0.1297		
50	0.0131	0.0822	0.0161	0.0890		
75	0.0278	0.1140	0.0120	0.0977	0.0397	0.1402
95	0.0800	0.2312	0.0766	0.2559		
99	0.2002	0.4232	0.3186	0.5532		
99.9	1.0261	0.9527	1.6881	1.2759		
			$\mathcal{GPD}$			
25	0.0034	0.0466	0.0272	0.1212		
50	0.0105	0.0807	0.0120	0.0828		
75	0.0892	0.2429	0.1195	0.3047	$3.1 \cdot 10^5$	6.9486
95	8.5160	2.7851	9.7921	2.9941		
99	475.35	21.3313	490.17	21.6451		
99.9	$1.8 \cdot 10^5$	404.99	$1.8 \cdot 10^5$	405.43		
			$\mathcal{B}urr$			
25	0.0031	0.0389	0.0240	0.1116		
50	0.0142	0.1049	0.0146	0.1069		
75	0.2266	0.4093	0.2875	0.4809	$7.9 \cdot 10^5$	13.1930
95	32.1950	5.2165	34.6466	5.4256		
99	2885.1	48.2596	2917.4	48.5730		
99.9	$2.5 \cdot 10^6$	1275.7	$2.5 \cdot 10^6$	1276.1		
			$\log \mathcal{S}_\alpha$			
25	0.0038	0.0488	0.0300	0.1308		
50	0.0129	0.0834	0.0158	0.0902		
75	0.0270	0.1208	0.0155	0.1104	0.1357	0.1746
95	0.1266	0.3286	0.1893	0.4083		
99	0.7405	0.7785	1.2092	1.0679		
99.9	8.3578	2.7464	10.7800	3.1893		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	0.0051	0.0471	0.0167	0.0969		
50	0.0483	0.1636	0.0463	0.1542		
75	0.7872	0.8014	0.9111	0.8727	$1.6 \cdot 10^9$	277.43
95	115.14	10.1901	120.01	10.3991		
99	$1.3 \cdot 10^4$	106.08	$1.3 \cdot 10^4$	106.40		
99.9	$2.6 \cdot 10^7$	4229.7	$2.6 \cdot 10^7$	4230.2		

Table 30
Average forecast errors for *"Relationship"* type aggregated losses. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0234	0.1082	0.1340	0.2867		
50	0.0552	0.1847	0.0696	0.2070		
75	0.0887	0.2432	0.0719	0.2314	35.7234	0.5449
95	0.9726	0.8818	1.9093	1.3553		
99	16.1004	3.9229	22.1948	4.6771		
99.9	518.30	21.9305	562.41	22.9933		
			$\mathcal{W}eib$			
25	0.0226	0.1077	0.1315	0.2842		
50	0.0594	0.1937	0.0748	0.2159		
75	0.1016	0.2524	0.0527	0.1937	0.2545	0.3286
95	0.2373	0.4028	0.5274	0.6930		
99	1.1885	1.0042	2.8841	1.6796		
99.9	9.9539	2.9937	16.8131	4.0560		
			$\log \mathcal{W}eib$			
25	0.0231	0.1074	0.1330	0.2862		
50	0.0585	0.1907	0.0738	0.2129		
75	0.0971	0.2429	0.0534	0.1970	0.4104	0.3584
95	0.3158	0.4950	0.7258	0.8181		
99	2.4188	1.4129	4.8112	2.1663		
99.9	32.7795	5.6141	45.4473	6.6756		
			$\mathcal{GPD}$			
25	0.0017	0.1070	0.1026	0.2493		
50	0.0062	0.1991	0.0604	0.2010		
75	1.8709	1.3121	2.3376	1.4844	$1.1{\cdot}10^{11}$	1996.5
95	755.17	26.8771	783.21	27.3695		
99	$2.0{\cdot}10^{5}$	430.82	$2.0{\cdot}10^{5}$	431.57		
99.9	$5.9{\cdot}10^{8}$	$2.1{\cdot}10^{4}$	$5.9{\cdot}10^{8}$	$2.1{\cdot}10^{4}$		
			$\mathcal{B}urr$			
25	0.0073	0.2000	0.0697	0.2146		
50	1.4999	1.0885	1.4432	1.0600		
75	65.7098	7.6095	68.3058	7.7819	$8.1{\cdot}10^{17}$	$4.2{\cdot}10^{6}$
95	$8.6{\cdot}10^{4}$	272.76	$8.6{\cdot}10^{4}$	273.25		
99	$6.4{\cdot}10^{7}$	7733.8	$6.4{\cdot}10^{7}$	7734.5		
99.9	$3.5{\cdot}10^{12}$	$1.5{\cdot}10^{6}$	$3.5{\cdot}10^{12}$	$1.5{\cdot}10^{6}$		
			$\log \mathcal{S}_\alpha$			
25	0.0230	0.1081	0.1322	0.2839		
50	0.0545	0.1865	0.0689	0.2090		
75	0.0876	0.2528	0.0683	0.2144	4.1107	0.4811
95	0.7972	0.7504	1.6497	1.2440		
99	11.7047	3.1938	16.8163	3.9486		
99.9	353.01	17.2799	389.68	18.3389		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	0.0190	0.1137	0.0964	0.2457		
50	0.0874	0.2288	0.0833	0.2202		
75	2.1481	1.3854	2.6382	1.5585	$2.5{\cdot}10^{8}$	248.42
95	620.66	24.3374	645.52	24.8316		
99	$9.3{\cdot}10^{4}$	298.99	$9.3{\cdot}10^{4}$	299.74		
99.9	$1.9{\cdot}10^{8}$	$1.4{\cdot}10^{4}$	$1.9{\cdot}10^{8}$	$1.4{\cdot}10^{4}$		

Table 31
Average forecast errors for *"Human"* type aggregated losses. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.1845	0.3767	0.7813	0.6932		
50	0.5499	0.5703	0.6100	0.6027		
75	1.0214	0.6848	0.4277	0.4759	4.8553	0.8211
95	2.5537	1.4027	2.1662	1.3522		
99	17.5320	3.9941	22.9917	4.5137		
99.9	372.16	17.7825	423.24	19.5002		
			$\mathcal{W}eib$			
25	0.1792	0.3718	0.7686	0.6856		
50	0.5613	0.5787	0.6226	0.6116		
75	1.1229	0.7448	0.4656	0.5072	0.6909	0.6418
95	2.2488	0.9302	0.5310	0.6656		
99	3.4129	1.3507	1.7793	1.2427		
99.9	7.3804	2.3555	8.5875	2.7014		
			$\log \mathcal{W}eib$			
25	0.1815	0.3737	0.7748	0.6895		
50	0.5588	0.5765	0.6212	0.6096		
75	1.1028	0.7324	0.4556	0.4948	0.8050	0.6653
95	2.2039	0.9722	0.6675	0.7635		
99	4.0523	1.7245	3.5640	1.7009		
99.9	19.5854	4.1441	26.7134	4.7749		
			$\mathcal{GPD}$			
25	0.1808	0.3723	0.7738	0.6878		
50	0.4797	0.5170	0.5314	0.5491		
75	1.2251	1.0102	1.0405	0.9499	$1.3 \cdot 10^{10}$	684.34
95	348.86	16.6486	371.07	17.4765		
99	$7.7 \cdot 10^{4}$	241.57	$7.8 \cdot 10^{4}$	242.82		
99.9	$4.7 \cdot 10^{8}$	$1.6 \cdot 10^{4}$	$4.7 \cdot 10^{8}$	$1.6 \cdot 10^{4}$		
			$\mathcal{B}urr$			
25	0.1912	0.3823	0.7968	0.7014		
50	0.5625	0.5791	0.6237	0.6119		
75	0.9859	0.6680	0.4165	0.4989	$3.3 \cdot 10^{4}$	2.6236
95	5.5129	2.1175	6.8680	2.2928		
99	204.42	12.6032	232.35	13.8325		
99.9	$3.3 \cdot 10^{4}$	162.51	$3.4 \cdot 10^{4}$	164.23		
			$\log \mathcal{S}_\alpha$			
25	0.1835	0.3758	0.7795	0.6919		
50	0.5532	0.5729	0.6136	0.6059		
75	1.0425	0.6940	0.4320	0.4572	3.1026	0.7709
95	2.3285	1.2779	1.6237	1.1215		
99	14.6743	3.2239	18.9137	3.7639		
99.9	327.02	13.9361	367.72	15.6366		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	0.1554	0.3542	0.7137	0.6587		
50	0.4121	0.4724	0.4522	0.4945		
75	3.4738	1.7482	3.8926	1.7777	$8.1 \cdot 10^{12}$	$1.4 \cdot 10^{4}$
95	1949.9	42.5619	2010.8	43.3890		
99	$7.0 \cdot 10^{5}$	785.20	$7.0 \cdot 10^{5}$	786.43		
99.9	$2.4 \cdot 10^{9}$	$4.7 \cdot 10^{4}$	$2.4 \cdot 10^{9}$	$4.7 \cdot 10^{4}$		

Table 32
Average forecast errors for *"Processes"* type aggregated losses. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0005	0.0146	0.0011	0.0205		
50	0.0008	0.0251	0.0008	0.0256		
75	0.0037	0.0514	0.0040	0.0534	29.1823	0.1734
95	0.1968	0.4199	0.2118	0.4381		
99	3.8728	1.8794	3.9690	1.9052		
99.9	154.56	12.1107	155.26	12.1468		
			$\mathcal{W}eib$			
25	0.0005	0.0149	0.0010	0.0202		
50	0.0007	0.0246	0.0008	0.0251		
75	0.0023	0.0440	0.0023	0.0422	0.0129	0.0537
95	0.0295	0.1628	0.0346	0.1812		
99	0.1884	0.4253	0.2090	0.4512		
99.9	1.1553	1.0500	1.2259	1.0858		
			$\log \mathcal{W}eib$			
25	0.0005	0.0148	0.0010	0.0204		
50	0.0007	0.0243	0.0008	0.0248		
75	0.0024	0.0454	0.0025	0.0432	0.0271	0.0612
95	0.0404	0.1924	0.0466	0.2106		
99	0.3151	0.5512	0.3425	0.5770		
99.9	3.2761	1.7834	3.3963	1.8192		
			$\mathcal{GPD}$			
25	0.0005	0.0140	0.0011	0.0206		
50	0.0008	0.0254	0.0008	0.0259		
75	0.0178	0.1108	0.0191	0.1176	$4.0 \cdot 10^{10}$	1599.5
95	44.4984	5.3895	44.6978	5.4077		
99	$6.9 \cdot 10^{4}$	185.52	$6.9 \cdot 10^{4}$	185.55		
99.9	$9.1 \cdot 10^{8}$	$2.2 \cdot 10^{4}$	$9.1 \cdot 10^{8}$	$2.2 \cdot 10^{4}$		
			$\mathcal{B}urr$			
25	0.0005	0.0138	0.0011	0.0205		
50	0.0010	0.0288	0.0010	0.2927		
75	0.1000	0.2734	0.1037	0.2803	$4.6 \cdot 10^{20}$	$9.7 \cdot 10^{7}$
95	1838.9	34.7495	1840.1	34.7676		
99	$2.2 \cdot 10^{7}$	3360.1	$2.2 \cdot 10^{7}$	3360.2		
99.9	$3.6 \cdot 10^{13}$	$3.9 \cdot 10^{16}$	$3.6 \cdot 10^{13}$	$3.9 \cdot 10^{6}$		
			$\log \mathcal{S}_\alpha$			
25	0.0005	0.0145	0.0011	0.0204		
50	0.0008	0.0251	0.0008	0.0256		
75	0.0037	0.0515	0.0041	0.0541	4.6680	0.1680
95	0.1965	0.4220	0.2117	0.4403		
99	4.4264	1.9974	4.5241	2.0233		
99.9	305.95	15.3729	306.88	15.4089		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	79.7415	4.1267	79.6492	4.1191		
50	$1.1 \cdot 10^{6}$	443.75	$1.1 \cdot 10^{6}$	443.75		
75	$1.0 \cdot 10^{11}$	$1.3 \cdot 10^{5}$	$1.0 \cdot 10^{11}$	$1.3 \cdot 10^{5}$	$4.1 \cdot 10^{63}$	$2.4 \cdot 10^{29}$
95	$4.9 \cdot 10^{20}$	$8.4 \cdot 10^{9}$	$4.9 \cdot 10^{20}$	$8.4 \cdot 10^{9}$		
99	$1.1 \cdot 10^{30}$	$4.0 \cdot 10^{14}$	$1.1 \cdot 10^{30}$	$4.0 \cdot 10^{14}$		
99.9	$1.4 \cdot 10^{45}$	$1.4 \cdot 10^{22}$	$1.4 \cdot 10^{45}$	$1.4 \cdot 10^{22}$		

Table 33
Average forecast errors for *"Technology"* type aggregated losses. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log\mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log\mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.4965	0.5922	0.5585	0.4217	0.4239	0.4130	0.4830

Table 34
Averaged p-values for *"Relationship"* aggregated losses in the 7-year forecast period.

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log\mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log\mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.4993	0.5310	0.5204	0.4530	0.4659	0.4115	0.0923

Table 35
Averaged p-values for *"Human"* aggregated losses in the 7-year forecast period.

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log\mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log\mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.2462	0.2392	0.2431	0.2526	0.2433	0.1339	0.2103

Table 36
Averaged p-values for *"Processes"* aggregated losses in the 7-year forecast period.

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log\mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log\mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.5238	0.5165	0.5107	0.5185	0.5210	0.3354	0.3247

Table 37
Averaged p-values for *"Technology"* aggregated losses in the 7-year forecast period.

8 APPENDIX B: Robust Analysis

Sample Descriptive Statistics				
	Relationship	Human	Processes	Technology
n	806	772	304	63
min ($ '000,000)	1.07	1.10	1.10	1.13
max ($ '000,000)	427.09	855.32	1,178	830.00
mean ($ '000,000)	39.63	45.47	113.31	74.40
median ($ '000,000)	13.50	11.12	33.61	11.60
st.dev. ($ '000,000)	59.78	85.41	188.66	134.77
skewness	2.4998	3.4703	2.6273	3.4060
kurtosis	10.1200	19.6686	10.5012	17.5666

Table 38
Descriptive statistics of the 4 types loss data, under the *robust* analysis in which highest 5% of data is excluded. *Note:* due to the small sample size of the *"Technology"* data, all results should be treated with caution.

	$\gamma, F_\gamma(u)$	Conditional			
		Relationship	Human	Processes	Technology
$\mathcal{LN}$	μ	16.1722	15.6905	17.0090	15.0313
	σ	1.7476	2.0691	1.9917	2.8285
	$F_\gamma(u)$	0.0887	0.1824	0.0544	0.3336
$\mathcal{W}eib$	β	0.0003	0.0030	0.0003	0.0120
	τ	0.4952	0.3679	0.4671	0.2870
	$F_\gamma(u)$	0.2259	0.3845	0.1567	0.4697
$\log \mathcal{W}eib$	β	$1.0{\cdot}10^{-11}$	$1.8{\cdot}10^{-9}$	$9.0{\cdot}10^{-12}$	$7.7{\cdot}10^{-8}$
	τ	9.1858	7.2258	8.8672	5.8818
	$F_\gamma(u)$	0.1585	0.2673	0.1056	0.3253
$\mathcal{GPD}$	ξ	0.9352	1.2808	1.1848	2.1207
	β	$1.1{\cdot}10^7$	$0.7{\cdot}10^7$	$2.4{\cdot}10^7$	$0.3{\cdot}10^7$
	$F_\gamma(u)$	0.0803	0.1262	0.0394	0.2233
$\mathcal{B}urr$	α	2.6845	0.3288	48.4907	0.1643
	β	$4.1{\cdot}10^5$	$1.6{\cdot}10^{11}$	$4.2{\cdot}10^5$	$0.2{\cdot}10^5$
	τ	0.7242	1.7551	0.5125	2.2048
	$F_\gamma(u)$	0.1315	0.0621	0.1272	0.9650
$\log \mathcal{S}_\alpha$	α	2	2	2	2
	β	0.9936	-0.3944	-0.1606	-0.4694
	σ	1.2392	1.4700	1.4096	2.0239
	μ	16.1656	15.6746	17.0067	14.9627
	$F_\gamma(u)$	0.0905	0.1865	0.0551	0.3456
$\mathcal{S}_\alpha\mathcal{S}$	α	0.7532	0.6750	0.6087	0.1773
	σ	$9.6{\cdot}10^6$	$6.7{\cdot}10^6$	$19.9{\cdot}10^6$	$5.7{\cdot}10^6$
	$F_\gamma(u)$	0.0781	0.1201	0.0469	0.2931

Table 39
Estimated γ and $F_\gamma(u)$ values for the conditional distributions fitted to 4 types of the operational loss data under the *robust* approach.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	1.3111	2.1614	4.5274	209.12	2.8289	34.5294	0.3485
	[<0.005]	[<0.005]	[0.094]	[>0.995]	[<0.005]	[0.031]	[<0.005]
$\mathcal{W}eib$	1.0407	1.7907	3.1282	327.36	1.5822	17.8496	0.2144
	[0.005]	[0.007]	[0.241]	[>0.995]	[<0.005]	[0.051]	[<0.005]
$\log \mathcal{W}eib$	1.0827	1.9746	3.35808	298.45	2.1510	20.5534	0.2989
	[0.005]	[<0.005]	[0.209]	[>0.995]	[<0.005]	[0.061]	[<0.005]
$\mathcal{GPD}$	1.6949	3.1270	4.8998	186.58	6.4187	43.6995	0.8247
	[<0.005]	[<0.005]	[0.072]	[>0.995]	[<0.005]	[0.024]	[<0.005]
$\mathcal{B}urr$	1.4346	2.6549	4.1987	251.92	4.3188	30.0690	0.5892
	[<0.005]	[<0.005]	[0.091]	[>0.995]	[<0.005]	[<0.005]	[<0.005]
$\log \mathcal{S}_\alpha$	1.3409	2.1544	4.5217	209.27	2.8492	34.7768	0.3579
	[<0.005]	[<0.005]	[0.078]	[>0.995]	[<0.005]	[0.006]	[<0.005]
$S\alpha S$	1.4187	2.7793	6.3995	144.74	5.5682	59.9109	0.6432
	[<0.005]	[<0.005]	[>0.995]	[>0.995]	[0.444]	[>0.995]	[<0.005]

Table 40
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the conditional loss distributions fitted to the *"Relationship"* loss data, under the *robust* approach. Smaller statistic values and greater p-values suggest better fit.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	1.2655	2.0577	3.8877	263.20	1.9615	25.5394	0.2088
	[<0.005]	[<0.005]	[0.129]	[>0.995]	[<0.005]	[0.032]	[0.005]
$\mathcal{W}eib$	1.1172	1.9160	3.9489	400.50	1.4831	16.1407	0.1682
	[<0.005]	[<0.005]	[0.124]	[0.991]	[<0.005]	[0.044]	[0.2702]
$\log \mathcal{W}eib$	1.1910	2.0574	3.7219	375.73	2.1002	17.9457	0.2712
	[<0.005]	[<0.005]	[0.156]	[>0.995]	[<0.005]	[0.065]	[<0.005]
$\mathcal{GPD}$	1.4888	2.7433	4.8677	191.47	4.5564	39.6902	0.5588
	[<0.005]	[<0.005]	[0.091]	[>0.995]	[<0.005]	[0.019]	[<0.005]
$\mathcal{B}urr$	2.0257	3.4016	6.1192	149.34	7.6553	57.4970	1.1822
	[0.006]	[0.006]	[0.062]	[>0.995]	[0.006]	[0.011]	[0.006]
$\log \mathcal{S}_\alpha$	1.2826	2.0270	3.8906	262.57	1.9443	25.8006	0.2071
	[<0.005]	[<0.005]	[0.122]	[>0.995]	[<0.005]	[0.032]	[<0.005]
$S\alpha S$	1.0613	2.0876	5.1613	258.41	3.3279	38.9658	0.3708
	[0.032]	[<0.005]	[>0.995]	[>0.995]	[0.990]	[>0.995]	[<0.005]

Table 41
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the conditional loss distributions fitted to the *"Human"* loss data, under the *robust* approach. Smaller statistic values and greater p-values suggest better fit.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	0.9080	1.5288	2.9178	107.11	1.4125	14.0184	0.1825
	[0.034]	[0.040]	[0.249]	[>0.995]	[<0.005]	[0.070]	[<0.005]
$\mathcal{W}eib$	0.5271	1.0392	1.8810	163.49	0.5123	6.6792	0.0572
	[0.707]	[0.582]	[0.624]	[>0.995]	[0.201]	[0.137]	[0.377]
$\log \mathcal{W}eib$	0.5576	1.1087	2.1457	146.86	0.7221	8.0808	0.0811
	[0.593]	[0.436]	[0.499]	[>0.995]	[0.069]	[0.096]	[0.162]
$\mathcal{GPD}$	1.1004	2.1675	4.5813	96.7443	3.7234	18.1558	0.4483
	[<0.005]	[<0.005]	[0.097]	[>0.995]	[<0.005]	[0.046]	[<0.005]
$\mathcal{B}urr$	0.8113	1.4379	2.3716	185.52	1.0810	6.6638	0.1363
	[0.118]	[0.109]	[0.336]	[0.959]	[0.021]	[0.039]	[0.050]
$\log \mathcal{S}_\alpha$	0.9352	1.5433	2.9087	107.42	1.4381	14.0480	0.1915
	[0.028]	[0.026]	[0.218]	[>0.995]	[<0.005]	[0.056]	[0.008]
$S\alpha S$	1.1692	2.2717	4.3264	77.5942	2.7614	26.4799	0.3115
	[0.016]	[<0.005]	[>0.995]	[>0.995]	[0.964]	[>0.995]	[0.016]

Table 42
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the conditional loss distributions fitted to the *"Processes"* loss data, under the *robust* approach. Smaller statistic values and greater p-values suggest better fit.

	D	V	A	A_{up}	A^2	A^2_{up}	W^2
$\mathcal{LN}$	1.0796	1.7451	2.7127	41.1440	1.3364	5.9777	0.1978
	[<0.005]	[0.005]	[0.217]	[0.989]	[<0.005]	[0.080]	[<0.005]
$\mathcal{W}eib$	1.0368	1.8359	2.7551	51.8632	1.4171	4.5168	0.2150
	[<0.005]	[<0.005]	[0.210]	[0.929]	[<0.005]	[0.087]	[<0.005]
$\log \mathcal{W}eib$	1.0358	1.9068	2.9926	50.0593	1.6058	4.7010	0.2408
	[0.005]	[<0.005]	[0.179]	[0.971]	[<0.005]	[0.097]	[<0.005]
$\mathcal{GPD}$	1.1362	1.7691	2.8950	32.5285	1.5441	8.3021	0.2295
	[<0.005]	[<0.005]	[0.213]	[>0.995]	[<0.005]	[0.070]	[<0.005]
$\mathcal{B}urr$	1.1179	1.8744	2.5384	28.0818	1.8242	9.7598	0.3061
	[0.356]	[0.344]	[0.522]	[>0.995]	[0.345]	[0.361]	[0.346]
$\log \mathcal{S}_\alpha$	1.0877	1.7429	2.7385	40.9965	1.3202	6.0542	0.1961
	[<0.005]	[0.006]	[0.222]	[0.990]	[0.006]	[0.068]	[0.006]
$S\alpha S$	2.8693	2.9544	6.0980	33.9452	19.9170	28.8310	3.7892
	[0.918]	[0.990]	[>0.995]	[>0.995]	[>0.995]	[>0.995]	[0.084]

Table 43
Goodness-of-fit test statistics and corresponding p-values (in square brackets) for the conditional loss distributions fitted to the *"Technology"* loss data, under the *robust* approach. Smaller statistic values and greater p-values suggest better fit.

EL	$\mathrm{VaR}_{0.95}$	$\mathrm{VaR}_{0.99}$	$\mathrm{CVaR}_{0.95}$	$\mathrm{CVaR}_{0.99}$
		$\mathcal{LN}$		
0.0826	0.2068	0.3947	0.3450	0.6560
		$\mathcal{W}eib$		
0.0638	0.1307	0.1766	0.1604	0.2090
		$\log \mathcal{W}eib$		
-	0.1355	0.1924	-	-
		$\mathcal{GPD}$		
0.3013	0.3627	1.4156	5.2570	23.7687
		$\mathcal{B}urr$		
0.0732	0.1715	0.3333	0.3131	0.6709
		$\log \mathcal{S}_\alpha$		
-	0.2106	0.4006	-	-
		$\mathcal{S}_\alpha\mathcal{S}$		
-	1.5339	12.2508	-	-

Table 44
Estimates of expected aggregated loss, VaR, and CVaR (figures must be further scaled $\times 10^{10}$) for *"Relationship"* type losses, under the conditional *robust* approach. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

EL	$\mathrm{VaR}_{0.95}$	$\mathrm{VaR}_{0.99}$	$\mathrm{CVaR}_{0.95}$	$\mathrm{CVaR}_{0.99}$
		$\mathcal{LN}$		
0.1497	0.3953	0.8625	0.7443	1.5569
		$\mathcal{W}eib$		
0.1095	0.2377	0.3504	0.3096	0.4385
		$\log \mathcal{W}eib$		
-	0.2527	0.3958	-	-
		$\mathcal{GPD}$		
-	1.7932	13.4743	-	-
		$\mathcal{B}urr$		
-	10.6753	161.59	-	-
		$\log \mathcal{S}_\alpha$		
-	0.4074	0.8983	-	-
		$\mathcal{S}_\alpha\mathcal{S}$		
-	4.7612	47.7160	-	-

Table 45
Estimates of expected aggregated loss, VaR, and CVaR (figures must be further scaled $\times 10^{10}$) for *"Human"* type losses, under the conditional *robust* approach. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

EL	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$
		$\mathcal{LN}$		
0.3188	0.8921	1.9222	1.6858	3.5673
		$\mathcal{W}eib$		
0.2121	0.4439	0.6272	0.5568	0.7447
		$\log \mathcal{W}eib$		
-	0.4911	0.7287	-	-
		$\mathcal{GPD}$		
-	2.5224	15.4268	-	-
		$\mathcal{B}urr$		
0.1987	0.4053	0.5646	0.5043	0.6731
		$\log \mathcal{S}_\alpha$		
-	0.8842	1.8991	-	-
		$\mathcal{S}_\alpha\mathcal{S}$		
-	$2.04{\cdot}10^{11}$	$28.9{\cdot}10^{11}$	-	-

Table 46
Estimates of expected aggregated loss, VaR, and CVaR (figures must be further scaled $\times 10^{10}$) for *"Processes"* type losses, under the conditional *robust* approach. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

EL	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$
		$\mathcal{LN}$		
0.0921	0.2617	1.1838	1.3261	4.5793
		$\mathcal{W}eib$		
0.0334	0.1336	0.3326	0.2787	0.5962
		$\log \mathcal{W}eib$		
-	0.1514	0.4355	-	-
		$\mathcal{GPD}$		
-	$1.47{\cdot}10^{10}$	$45.1{\cdot}10^{10}$	-	-
		$\mathcal{B}urr$		
-	$7.73{\cdot}10^{10}$	$801{\cdot}10^{10}$	-	-
		$\log \mathcal{S}_\alpha$		
-	0.2802	1.2870	-	-
		$\mathcal{S}_\alpha\mathcal{S}$		
-	$5.89{\cdot}10^{17}$	$5.60{\cdot}10^{21}$	-	-

Table 47
Estimates of expected aggregated loss, VaR, and CVaR (figures must be further scaled $\times 10^{10}$) for *"Technology"* type losses, under the conditional *robust* approach. Figures are based on 50,000 Monte Carlo samples. Where applicable, EL was computed directly as $\mathbb{E}N \times \mathbb{E}X$.

	$\mathcal{LN}$	$\mathcal{W}eib$	$\log \mathcal{W}eib$	$\mathcal{GPD}$	$\mathcal{B}urr$	$\log \mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Relationship							
$\frac{\triangle \text{VaR}_{0.95}}{\text{VaR}_{0.95}^{class.}} \times 100$	56%	59%	59%	77%	89%	52%	66%
$\frac{\triangle \text{VaR}_{0.99}}{\text{VaR}_{0.99}^{class.}} \times 100$	63%	66%	67%	88%	97%	58%	78%
Human							
$\frac{\triangle \text{VaR}_{0.95}}{\text{VaR}_{0.95}^{class.}} \times 100$	68%	70%	71%	85%	89%	82%	67%
$\frac{\triangle \text{VaR}_{0.99}}{\text{VaR}_{0.99}^{class.}} \times 100$	75%	78%	79%	92%	95%	79%	77%
Processes							
$\frac{\triangle \text{VaR}_{0.95}}{\text{VaR}_{0.95}^{class.}} \times 100$	65%	65%	67%	88%	78%	65%	-
$\frac{\triangle \text{VaR}_{0.99}}{\text{VaR}_{0.99}^{class.}} \times 100$	71%	71%	73%	94%	87%	72%	-
Technology							
$\frac{\triangle \text{VaR}_{0.95}}{\text{VaR}_{0.95}^{class.}} \times 100$	10%	8%	9%	-	-	6%	-
$\frac{\triangle \text{VaR}_{0.99}}{\text{VaR}_{0.99}^{class.}} \times 100$	7%	8%	8%	-	-	-	-

Table 48
Sensitivity of classical VaR to outliers for the 4 types of operational losses. Figures indicate incremental VaR as the percentage of classical VaR attributed to top 5% of data. *Note:* for the *"Technology"* type losses, excluding the outliers in the data resulted in heavier-tailed distributions than under the classical analysis. This may be explained by the estimation inaccuracy due to the small dataset (63 points).

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0022	0.0417	0.0018	0.0394		
50	0.0040	0.0509	0.0039	0.0502		
75	0.0092	0.0753	0.0120	0.0921	0.0274	0.0859
95	0.0431	0.1916	0.0629	0.2368		
99	0.1704	0.3952	0.2283	0.4617		
99.9	1.1851	1.0771	1.3920	1.1685		
			$\mathcal{W}eib$			
25	0.0018	0.0390	0.0017	0.0367		
50	0.0026	0.0446	0.0025	0.0439		
75	0.0039	0.0509	0.0055	0.0556	0.0052	0.0552
95	0.0083	0.0729	0.0170	0.1181		
99	0.0151	0.1069	0.0340	0.1733		
99.9	0.0288	0.1583	0.0667	0.2498		
			$\log \mathcal{W}eib$			
25	0.0017	0.0384	0.0016	0.0361		
50	0.0025	0.0442	0.0024	0.0436		
75	0.0040	0.0512	0.0056	0.0572	0.0055	0.0561
95	0.0090	0.0782	0.0183	0.1234		
99	0.0182	0.1227	0.0394	0.1892		
99.9	0.0475	0.2101	0.0951	0.3017		
			$\mathcal{GPD}$			
25	0.0023	0.0425	0.0019	0.0403		
50	0.0056	0.0563	0.0054	0.0557		
75	0.0205	0.1200	0.0249	0.1368	10.6607	0.2400
95	0.3315	0.5390	0.3838	0.5842		
99	5.3368	2.1998	5.6442	2.2663		
99.9	345.46	17.1380	348.73	17.2293		
			$\mathcal{B}urr$			
25	0.0018	0.0384	0.0016	0.0362		
50	0.0033	0.0465	0.0031	0.0459		
75	0.0090	0.0781	0.0120	0.0948	1.7318	0.1225
95	0.0855	0.2745	0.1132	0.3196		
99	0.9772	0.9112	1.1087	0.9777		
99.9	31.5765	4.9530	32.5110	5.0443		
			$\log \mathcal{S}_\alpha$			
25	0.0006	0.0223	0.0015	0.0353		
50	0.0016	0.0284	0.0016	0.0291		
75	0.0064	0.0626	0.0078	0.0731	0.0844	0.0937
95	0.0899	0.2662	0.1150	0.3099		
99	0.6948	0.7844	0.7980	0.8476		
99.9	8.5024	2.7599	8.9913	2.8457		
			$\mathcal{S}_\alpha \mathcal{S}$			
25	0.0072	0.0643	0.0053	0.0586		
50	0.0326	0.1493	0.0321	0.1478		
75	0.2309	0.4306	0.2462	0.4474	$1.0 \cdot 10^5$	6.7593
95	13.7673	3.4414	14.0927	3.4866		
99	869.34	27.2307	873.16	27.2970		
99.9	$5.1 \cdot 10^5$	674.49	$5.1 \cdot 10^5$	674.58		

Table 49
Average forecast errors for *"Relationship"* type aggregated losses, under the *robust* approach. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0023	0.0356	0.0029	0.0468		
50	0.0050	0.0479	0.0046	0.0465		
75	0.0161	0.1099	0.0221	0.1364	0.1233	0.1392
95	0.1414	0.3660	0.1997	0.4399		
99	0.8455	0.9129	1.0574	1.0231		
99.9	10.0861	3.1153	11.0224	3.2675		
			$\mathcal{W}eib$			
25	0.0021	0.0357	0.0030	0.0477		
50	0.0032	0.0413	0.0031	0.0408		
75	0.0061	0.0558	0.0090	0.0785	0.0123	0.0789
95	0.0221	0.1322	0.0455	0.2062		
99	0.0587	0.2302	0.1190	0.3403		
99.9	0.2154	0.4556	0.3734	0.6083		
			$\log \mathcal{W}eib$			
25	0.0022	0.0362	0.0030	0.0481		
50	0.0034	0.0411	0.0032	0.0406		
75	0.0069	0.0589	0.0100	0.0829	0.0144	0.0828
95	0.0269	0.1468	0.0526	0.2207		
99	0.0820	0.2745	0.1529	0.3846		
99.9	0.3523	0.5741	0.5514	0.7270		
			$\mathcal{GPD}$			
25	0.0033	0.0361	0.0029	0.0428		
50	0.0171	0.1118	0.0163	0.1085		
75	0.1576	0.3784	0.1783	0.4050	$4.9 \cdot 10^{6}$	15.7218
95	10.6869	3.1887	11.1722	3.2626		
99	736.10	26.7003	742.06	26.8102		
99.9	$2.6 \cdot 10^{5}$	455.70	$2.6 \cdot 10^{5}$	455.85		
			$\mathcal{B}urr$			
25	0.0174	0.1131	0.0116	0.0828		
50	0.2081	0.4244	0.2052	0.4212		
75	4.0533	1.9078	4.1594	1.9344	$3.1 \cdot 10^{11}$	4224.2
95	1583.8	38.2283	1589.6	38.302		
99	$5.9 \cdot 10^{5}$	721.19	$5.9 \cdot 10^{5}$	721.30		
99.9	$4.5 \cdot 10^{9}$	$6.0 \cdot 10^{4}$	$4.5 \cdot 10^{9}$	$6.0 \cdot 10^{4}$		
			$\log \mathcal{S}_\alpha$			
25	0.0024	0.0364	0.0031	0.0484		
50	0.0044	0.0431	0.0042	0.0427		
75	0.0111	0.0726	0.0151	0.0993	0.0430	0.0994
95	0.0651	0.2067	0.1008	0.2807		
99	0.3245	0.4564	0.4411	0.5663		
99.9	3.5175	1.3713	3.9992	1.5242		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	0.0090	0.0731	0.0059	0.0489		
50	0.0619	0.2265	0.0604	0.2232		
75	0.6697	0.7782	0.7125	0.8048	$5.3 \cdot 10^{7}$	54.9360
95	71.5082	8.1565	72.7595	8.5327		
99	7551.45	85.2174	7570.7	2661.2		
99.9	$7.6 \cdot 10^{6}$	2661.0	$7.6 \cdot 10^{6}$	54.9360		

Table 50
Average forecast errors for *"Human"* type aggregated losses, under the *robust* approach. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0331	0.1494	0.0516	0.1785		
50	0.0392	0.1687	0.0400	0.1704		
75	0.0538	0.1963	0.0485	0.1893	0.2313	0.2548
95	0.2980	0.4759	0.3717	0.5666		
99	1.9817	1.3538	2.3595	1.5047		
99.9	22.4811	4.5575	24.1880	4.7621		
			$\mathcal{W}eib$			
25	0.0332	0.1492	0.0519	0.1783		
50	0.0406	0.1697	0.0416	0.1714		
75	0.0499	0.1893	0.0385	0.1697	0.0537	0.1869
95	0.0791	0.2440	0.0750	0.2330		
99	0.1437	0.3271	0.1833	0.3760		
99.9	0.3465	0.5078	0.5118	0.6844		
			$\log \mathcal{W}eib$			
25	0.0335	0.1498	0.0524	0.1789		
50	0.0405	0.1702	0.0415	0.1720		
75	0.0499	0.1893	0.0390	0.1706	0.0593	0.1931
95	0.0877	0.2569	0.0924	0.2660		
99	0.2033	0.3856	0.2751	0.4826		
99.9	0.7410	0.7835	1.0267	0.9885		
			$\mathcal{GPD}$			
25	0.0033	0.0361	0.0029	0.0428		
50	0.0171	0.1118	0.0163	0.1085		
75	0.1576	0.3784	0.1783	0.4050	$4.9 \cdot 10^6$	15.7218
95	10.6869	3.1887	11.1722	3.2626		
99	736.10	26.7003	742.06	26.8102		
99.9	$2.6 \cdot 10^5$	455.70	$2.6 \cdot 10^5$	455.85		
			$\mathcal{B}urr$			
25	0.0174	0.1131	0.0116	0.0828		
50	0.2081	0.4244	0.2052	0.4212		
75	4.0533	1.9078	4.1594	1.9344	$3.1 \cdot 10^{11}$	4224.2
95	1583.8	38.2283	1589.6	38.302		
99	$5.9 \cdot 10^5$	721.19	$5.9 \cdot 10^5$	721.30		
99.9	$4.5 \cdot 10^9$	$6.0 \cdot 10^4$	$4.5 \cdot 10^9$	$6.0 \cdot 10^4$		
			$\log \mathcal{S}_\alpha$			
25	0.0024	0.0364	0.0031	0.0484		
50	0.0044	0.0431	0.0042	0.0427		
75	0.0111	0.0726	0.0151	0.0993	0.0430	0.0994
95	0.0651	0.2067	0.1008	0.2807		
99	0.3245	0.4564	0.4411	0.5663		
99.9	3.5175	1.3713	3.9992	1.5242		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	0.0090	0.0731	0.0059	0.0489		
50	0.0619	0.2265	0.0604	0.2232		
75	0.6697	0.7782	0.7125	0.8048	$5.3 \cdot 10^7$	54.9360
95	71.5082	8.1565	72.7595	8.5327		
99	7551.45	85.2174	7570.7	2661.2		
99.9	$7.6 \cdot 10^6$	2661.0	$7.6 \cdot 10^6$	54.9360		

Table 51
Average forecast errors for *"Processes"* type aggregated losses, under the *robust* approach. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	Forecasted quantiles vs. bootstrapped quantiles		Forecasted quantiles vs. actual loss		Overall error: forecasted vs. actual loss	
%	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MSE ($\times 10^{20}$)	MAE ($\times 10^{10}$)	MAE ($\times 10^{20}$)	MSE ($\times 10^{10}$)
			$\mathcal{LN}$			
25	0.0005	0.0144	0.0011	0.0206		
50	0.0008	0.0247	0.0008	0.0252		
75	0.0036	0.0512	0.0039	0.0516	1.2010	0.1449
95	0.1906	0.4121	0.2055	0.4303		
99	3.9387	1.8837	4.0368	1.9095		
99.9	186.07	12.7354	186.96	12.7711		
			$\mathcal{W}eib$			
25	0.0005	0.0146	0.0011	0.0206		
50	0.0007	0.0241	0.0008	0.0246		
75	0.0020	0.0427	0.0020	0.0392	0.0131	0.0523
95	0.0277	0.1572	0.0326	0.1754		
99	0.1928	0.4277	0.2130	0.4534		
99.9	1.2291	1.0954	1.3054	1.1315		
			$\log \mathcal{W}eib$			
25	0.0005	0.0147	0.0010	0.0204		
50	0.0007	0.0241	0.0008	0.0246		
75	0.0021	0.0429	0.0021	0.0401	0.0256	0.0590
95	0.0368	0.1830	0.0429	0.2012		
99	0.3080	0.5413	0.3367	0.5670		
99.9	3.1803	1.7301	3.3058	1.7656		
			$\mathcal{GPD}$			
25	0.0005	0.0140	0.0011	0.0206		
50	0.0008	0.0255	0.0009	0.0260		
75	0.0205	0.1151	0.0220	0.1219	$2.1 \cdot 10^{13}$	$2.4 \cdot 10^{4}$
95	54.4298	5.6635	54.6411	5.6817		
99	$1.1 \cdot 10^{5}$	228.25	$1.1 \cdot 10^{5}$	228.27		
99.9	$2.9 \cdot 10^{9}$	$3.5 \cdot 10^{4}$	$2.9 \cdot 10^{9}$	$3.5 \cdot 10^{4}$		
			$\mathcal{B}urr$			
25	0.0005	0.0138	0.0011	0.0204		
50	0.0010	0.0293	0.0011	0.2979		
75	0.1533	0.3034	0.1576	0.3103	$1.0 \cdot 10^{21}$	$1.3 \cdot 10^{8}$
95	2873.8	39.8687	2875.3	39.8869		
99	$3.1 \cdot 10^{7}$	3789.3	$3.1 \cdot 10^{7}$	3789.3		
99.9	$1.8 \cdot 10^{13}$	$2.4 \cdot 10^{16}$	$1.8 \cdot 10^{13}$	$2.4 \cdot 10^{6}$		
			$\log \mathcal{S}_\alpha$			
25	0.0005	0.0144	0.0011	0.0206		
50	0.0008	0.0244	0.0008	0.0249		
75	0.0035	0.0505	0.0037	0.0509	4.4229	0.1518
95	0.1749	0.4035	0.1896	0.4219		
99	3.6596	1.8473	3.7559	1.8731		
99.9	186.17	12.0747	187.10	12.1103		
			$\mathcal{S}_\alpha\mathcal{S}$			
25	570.47	13.0921	570.22	13.0862		
50	$4.9 \cdot 10^{8}$	$1.2 \cdot 10^{4}$	$4.9 \cdot 10^{8}$	$1.2 \cdot 10^{4}$		
75	$1.7 \cdot 10^{16}$	$6.9 \cdot 10^{7}$	$1.7 \cdot 10^{16}$	$6.9 \cdot 10^{7}$	$7.3 \cdot 10^{78}$	$1.0 \cdot 10^{37}$
95	$1.4 \cdot 10^{31}$	$2.0 \cdot 10^{15}$	$1.4 \cdot 10^{31}$	$2.0 \cdot 10^{15}$		
99	$4.4 \cdot 10^{45}$	$3.0 \cdot 10^{22}$	$4.4 \cdot 10^{45}$	$3.0 \cdot 10^{22}$		
99.9	$1.3 \cdot 10^{65}$	$1.6 \cdot 10^{32}$	$1.3 \cdot 10^{65}$	$1.6 \cdot 10^{32}$		

Table 52
Average forecast errors for *"Technology"* type aggregated losses, under the *robust* approach. *Left panel:* errors between corresponding quantiles; *middle panel:* errors of forecasted quantiles relative to realized loss; *right panel:* overall error between forecasted and realized loss. Figures are based on 50,000 Monte Carlo samples.

	$\mathcal{LN}$	$\mathcal{Weib}$	$\log \mathcal{Weib}$	$\mathcal{GPD}$	$\mathcal{Burr}$	$\log \mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.5080	0.5260	0.4975	0.4246	0.4410	0.3679	0.5180

Table 53
Averaged p-values for *"Relationship"* aggregated losses in the 7-year forecast period, under the *robust* approach.

	$\mathcal{LN}$	$\mathcal{Weib}$	$\log \mathcal{Weib}$	$\mathcal{GPD}$	$\mathcal{Burr}$	$\log \mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.4821	0.4947	0.4774	0.4401	0.4730	0.1443	0.4841

Table 54
Averaged p-values for *"Human"* aggregated losses in the 7-year forecast period, under the *robust* approach.

	$\mathcal{LN}$	$\mathcal{Weib}$	$\log \mathcal{Weib}$	$\mathcal{GPD}$	$\mathcal{Burr}$	$\log \mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.2190	0.2008	0.2060	0.2526	0.2433	0.1339	0.2103

Table 55
Averaged p-values for *"Processes"* aggregated losses in the 7-year forecast period, under the *robust* approach.

	$\mathcal{LN}$	$\mathcal{Weib}$	$\log \mathcal{Weib}$	$\mathcal{GPD}$	$\mathcal{Burr}$	$\log \mathcal{S}_\alpha$	$\mathcal{S}_\alpha\mathcal{S}$
Ave. p-value	0.3373	0.3266	0.3278	0.3303	0.3428	0.3392	0.2148

Table 56
Averaged p-values for *"Technology"* aggregated losses in the 7-year forecast period, under the *robust* approach.

References

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

Berkowitz, J., 2000. Testing density forecasts with applications to risk management. Tech. rep., University of California Irvine.

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`.

Chernobai, A., Burnecki, K., Rachev, S., Trück, S., Weron, R., 2005a. Modelling catastrophe claims with left-truncated severity distributions. forthcoming in The Journal of Computational Statistics.

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

Chernobai, A., Rachev, S., Fabozzi, F., 2005c. Composite goodness-of-fit tests for left-truncated loss samples. Tech. rep., University of California Santa Barbara.

Crnkovic, C., Drachman, J., 1997. Quality Control in VAR: Understanding and Applying Value-at-Risk. Risk Books, London.

Crouhy, M., Galai, D., Mark, R., 2001. Risk Management. McGraw-Hill, New York.

Cruz, M. G., 2002. Modeling, Measuring and Hedging Operational Risk. John

Wiley & Sons, New York, Chichester.

Davis, E. (Ed.), 2005. Operational risk: Practical Approaches to Implementation. Risk Books, London.

De Fountnouvelle, P., DeJesus-Rueff, V., Jordan, J., Rosengren, E., 2003. Using loss data to quantify operational risk. Tech. rep., Federal Reserve Bank.

Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B (Methodological) 39 (1), 1–38.

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

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

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.

Jorion, P., 2000. Value-at-Risk: the New Benchmark for Managing Financial Risk, 2nd Edition. McGraw-Hill, New York.

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.

Lawrence, K. D., Arthur, J. L. (Eds.), 1990. Robust Regression: Analysis and Applications. Marcel Dekker, Inc., New York and Basel.

Lewis, C. M., Lantsman, Y., 2005. What is a fair price to transfer the risk of unauthorised trading? a case study on operational risk. In: Davis (2005), pp. 315–355.

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

Moscadelli, M., 2005. The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee. In: Davis (2005), pp. 39–106.

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

Rachev, S. T., Mittnik, S., 2000. Stable Paretian Models in Finance. John Wiley & Sons, New York.

Rosenblatt, M., 1952. Remarks on a multivariate transformation. Annals of Mathematical Statistics 23, 470–472.

Ross, S., 2001. Simulation, 3rd Edition. Academic Press, Boston.

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

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