QUANTIFYING REGULATORY CAPITAL FOR OPERATIONAL RISK

PAUL EMBRECHTS, HANSJÖRG FURRER, AND ROGER KAUFMANN

ABSTRACT. The proposed New Accord (Basel II) established by the Basel Committee on Banking Supervision calls for an explicit treatment of operational risk. Banks are required to demonstrate their ability to capture severe tail loss events. Value at risk is a risk measure that could be used to derive the necessary regulatory capital. Yet operational loss data typically exhibit irregularities which complicate the mathematical modeling. It is shown that traditional modeling approaches, including extreme value theory, reach their limits as the structure of operational loss data is barely in line with the modeling assumptions.

PRACTICAL APPLICATIONS

The aim of the present paper is to demonstrate potential difficulties on the way from operational loss data to a regulatory capital charge. Even though there exist a vast amount of loss modeling techniques, their applications to operational loss data is not straightforward. Commonly accepted stylized facts of operational loss data such as irregularities in the occurrence times or the presence of extremes for example require approaches beyond the traditional ones. When we review some of the classical loss models including the extreme value theory approach, the emphasis will be on the underlying modeling assumptions. We then indicate the direction into which these models could be amended in order to capture operational loss data.

1. INTRODUCTION

The Basel Committee on Banking Supervision (the Committee) established in its New Accord [3] (Basel II for short) a three pillar framework for risk management practices of financial institutions. Pillar 1 –minimum capital requirements– is devoted to risk measurement and the concomitant capital requirements serving as a cushion against unexpected losses. The second pillar –supervisory review of capital adequacy– calls for an effective framework to identify, assess, monitor,

Research supported by Credit Suisse Group, Swiss Re and UBS AG through RiskLab, Switzerland.

and control risks. Pillar 3 –public disclosure– finally requires public disclosure of loss data and management methods.

In the current accord as well as in the proposed New Accord the minimum capital requirements are based on a capital ratio where the numerator represents the total amount of capital a bank has available whereas the denominator consists of the risk-weighted assets. The resulting capital ratio shall not be less than 8%. Under the proposed New Accord, the definition of the numerator (i.e. regulatory capital) and the minimum ratio of 8% remain unchanged. The modifications apply to the risk-weighted assets, that is to the methods in place to measure the risks faced by a bank. One of these modifications concerns the explicit treatment of operational risk. Following the Committee's wording, we understand by operational risk "the risk of losses resulting from inadequate or failed internal processes, people and systems, or external events."

In this document we shed some light on the implementation of Pillar 1 in the context of operational risk losses. A major innovation of the proposed New Accord is the introduction of three different approaches for the calculation of operational risk. In this note, the emphasis will be on the implementation of the advanced measurement approach, see Section 2. The actuarial world provides a myriad of techniques how to analyze historical data. In Section 3.1 we briefly review some actuarial techniques designed to finding an overall model for loss severities. Yet we are more concerned with modeling the largest losses. Extreme value theory provides a number of sensible approaches to this problem, see Section 3.2. However, this presupposes that sufficient data are available above a certain high threshold value. We understand that the actuarial methods are eligible for operational risk data too *provided* those data are in line with the modeling assumptions. Section 3.3 is devoted to the applicability of actuarial models in general. In Section 3.3.3, we present a numerical example giving some insight into the number of observations needed to arrive at accurate quantile estimates. This example is of pedagogical nature only as it is based on independence and stationarity assumptions of the underlying data. Operational risk losses, however, reveal some facts which are barely in accordance with the modeling assumptions made in that example. Some concluding remarks can be found in Section 4.

2. Measuring operational risk

The launch of the New Accord has attracted great interest, not least because of the emphasis that is given to operational risk. In the wake of the New Accord, we have observed a number of articles, research papers and books addressing the issue on how to *quantify* operational risks. The proposed New Accord intends to introduce three distinct options for the calculation of operational risk. These approaches reflect different levels of risk sensitivity allowing banks to elect the approach that fits best to their operations. The three approaches for measuring operational risk are as follows:

(1) Basic Indicator Approach
(2) Standardised Approach
(3) Advanced Measurement Approach (AMA).

Roughly speaking, the basic indicator and the standardised approach require banks to hold a capital for operational risk losses equal to a fixed percentage of gross income. These two approaches are targeted to banks with moderate exposure to operational risk losses. Internationally active banks facing a substantial exposure to operational risk losses are expected to implement over time the more involved AMA. In developing concepts for the AMA, banks are given a great deal of flexibility as long as these approaches are consistent with the banks' activities. The Committee is not specifying distributional assumptions from which risk measures for regulatory capital purposes can be derived. However, a bank must be able do demonstrate that its approach captures potentially severe tail loss events, see Basel II [3] p. 126. To be more precise, banks should put an operational risk capital aside in line with the 99.9% or even higher confidence level over a one-year holding period. In general terms, the capital charge C can be derived as follows:

$$C = \sum_{i=1}^{s} \mathrm{VaR}_\alpha^i \ ,$$

where VaR_α^ℓ denotes the Value at Risk (VaR) at confidence level α of business line ℓ. Summation in the above definition is used as the Committee requires risk measures for different business lines and/or operational risk types to be added up. VaR can be defined as a statistical estimation of a portfolio loss with the property that, with a given (small) probability, we stand to incur that loss *or more* over a given (typically short) holding period. The holding period should be related to the liquidity of the assets: if a financial institution runs into difficulties, the holding period should cover the time necessary to raise additional funds for corrective actions. Formally, VaR is defined as follows. Note our convention to consider losses as positive and gains as negative.

Definition 1. Given some confidence level $\alpha \in (0,1)$, the VaR of a portfolio at the confidence level α is given by the smallest number ℓ such that the probability that the loss L exceeds ℓ is no greater than $(1-\alpha)$:

$$\mathrm{VaR}_\alpha(L) = \inf\{\ell \in \mathbb{R} \mid \mathbb{P}[L > \ell] \le 1-\alpha\} \ . \qquad \square$$

Formulated differently, $\mathrm{VaR}_\alpha(L)$ corresponds to the α-quantile q_α of the distribution function of L. If L has cumulative distribution function (cdf) F, this is also denoted as $q_\alpha = F^{\leftarrow}(\alpha)$. Typical values for α in the context of market risk management are $\alpha = 0.95$ or $\alpha = 0.99$. VaR techniques with confidence levels in the area of 99.9% and beyond at least become delicate for the simple reason that there is hardly enough repetitive data at hand to "predict" losses of such magnitude. Recall that $\alpha = 99.9\%$ corresponds to a one-in-thousand event. However, it is precisely the area of operational (and credit-) risk where such levels of α are ubiquitous.
We shall not address here the question whether or not VaR is actually a "good" risk measure. For example, it is known that VaR in general is *not* a coherent risk measure in the sense of Artzner et al. [1]. It lacks the property of subadditivity. Intuitively, subadditivity reflects the idea of diversification. Moreover, VaR does not tell anything about the potential size of the loss that exceeds it. To circumvent this problem, Artzner et al. [1] introduce the concept of expected shortfall instead. In mathematical terms, expected shortfall is the conditional expectation of L, given that $L \geq \mathrm{VaR}$. The amenities of expected shortfall as an alternative risk measure are twofold. Not only does expected shortfall provide information about the size of a loss, but it also falls into the class of coherent risk measures. (Strictly speaking, the latter property only holds for absolutely continuous loss distributions.)

3. Preconditions for fitting operational risk data

The AMA aims at being the most sophisticated approach for quantifying operational risk. However, the increased level of sophistication comes at the cost of a number of modeling assumptions which have to be fulfilled. Obviously, the accuracy in predicting future loss values depends on the volume and quality of the observed historical data. The Committee requires that operational risk measures are based on a minimum five-year observation period of internal loss data. When a bank first moves to the AMA, a three-year horizon will be accepted.
Actuarial science essentially provides the techniques to analyze the given data and to make inference about future losses. Applying those methods only makes sense if the assumptions underlying the actuarial models are fulfilled. Otherwise, erroneous conclusions may be drawn.
The purpose of this section is to briefly review those actuarial techniques which are also relevant for the banking industry when faced with capital adequacy requirement issues. To begin with, we recall some loss distribution fitting techniques which are widely used in the actuarial world. In essence, those methods are needed for VaR purposes too. There, the emphasis is on the inverse of the cdf, especially in the far end tail provided the VaR-confidence level α is high. Extreme

value theory (EVT) can offer a solution in this context. In particular, we discuss the preconditions under which standard actuarial methods complemented with EVT provide meaningful estimates of high quantiles. There is a vast amount of actuarial literature dealing with the analysis of aggregate loss distributions, see for instance Klugman et al. [13] or Panjer and Willmot [18] to mention just two of them. The standard reference book for EVT is Embrechts et al. [10].

3.1. **Standard actuarial methods.** In non-life insurance, the analysis of the total claim amount S of an insurer's portfolio has always been a key issue. Not only are most premium principles derived from S but also more modern concepts in financial risk management such as VaR for instance basically depend on aggregate loss distributions. In the framework of the *individual* risk model one assumes that the number of claims over a given time period is deterministic. Furthermore, it is assumed that the individual risks X_i, $i = 1, \ldots, n$, constituting a portfolio are independent random variables. It is not assumed, however, that the X_i's have identical distribution. Note that the X_i's are allowed to have positive mass at zero meaning that, with positive probability, policy i produces no claim . The independence assumption entails that the sum $S = \sum_{k=1}^{n} X_k$ can be calculated by means of convolution techniques. Yet this technique is quite laborious and explicit solutions for the cumulative distribution function of S can seldom be obtained. So there is a need for alternative methods. As S is the sum of independent random variables, it is tempting to approximate S by a normal law with the same mean and variance as S. Insurance as well as operational loss data, however, are often heavier tailed so this approximation in general does not work well. To compensate for the skewness, one can use an approximating distribution. The idea is to use the method of moments to estimate the parameters of the approximating distribution. The advantage of this method is that it is simple and easy to apply. Candidates for approximating distributions are:

- Translated gamma distribution
- Translated lognormal distribution.

Hence, the cdf of S is approximated by the cdf of $x_0 + Y$, say, where Y has a gamma or lognormal distribution with parameters α, β. Equating mean value, variance and skewness of $x_0 + Y$ with the corresponding moment estimates of S yields the unknown parameters α, β and x_0. Although this approach does a better job than the normal approximation the tail severity may still be underestimated.

In the *collective* risk model the portfolio is regarded as a collective that produces claims at random points in time. The number of claims is a random variable modeled by a counting process $N = (N_t)_{t \geq 0}$, where N_t denotes the number of claims in the time interval $[0, t]$. It

is assumed that the claim sizes $(X_i)_{i\in\mathbb{N}}$ are independent and identically distributed (iid) with common distribution function F_X and that N is independent of the sequence $(X_i)_{i\in\mathbb{N}}$. The most prominent example of a counting process is the homogeneous Poisson process with intensity λ. It is well-known that the mean value and the variance of a Poisson(λ)-distributed random variable are both equal to λ. If the number of claims exhibits a larger spread around the mean, one may use the negative binomial distribution instead. The negative binomial distribution arises naturally by assuming that the intensity λ of a Poisson process follows a gamma distribution; moving from a deterministic λ to a random intensity is referred to as mixing.
The compound distribution function $F_S(x) = \mathbb{P}[S \leq x]$ can be expressed in terms of convolutions of F_X. The computation of F_S is generally not an easy task, even in the simplest cases. Among the algorithms to calculate F_S we mention the following:

(A1) Approximation
(A2) Inversion methods
(A3) Recursive methods (Panjer recursion)
(A4) Simulation.

Approximation. Various approximation methods can be used to circumvent the difficulties in calculating F_S. The normal approximation based on the central limit theorem and the approximations by skewed random variables are also applicable in the collective model. Other methods of computing (tail-) probabilities for sums of random variables include the Edgeworth expansion or the saddlepoint approximation. These methods provide good approximations in various settings. For a thorough discussion of these techniques we refer to the actuarial literature given above; see also Jensen [11].

Inversion methods. Inversion methods are used to obtain numerically the probability function (density or mass function) from a known expression for a transform such as the characteristic function of the desired random variable. The fast Fourier transform (FFT) is an algorithm to invert the characteristic function $\varphi_S(z) = \mathbb{E}[e^{iSz}]$ to obtain densities of discrete random variables. Note that the FFT procedure requires a discretization of the severity distribution.

Recursive methods. The Panjer recursion calculates the probability of the event $\{S = k\}$ recursively in terms of the probabilities of $\{S = \ell\}$, $\ell = 0, 1, 2, \ldots, k-1$, see (2) below. The probabilities p_n of having n claims have to satisfy the following recursion relation for some real a and b:

$$p_n = \Big(a + \frac{b}{n}\Big) p_{n-1} \, , \quad n \geq 1 \, . \tag{1}$$

One can show that only the Poisson, negative binomial and the binomial distribution satisfy (1). For those claims frequency distributions with claim sizes defined on the positive integers, the following recursive formula for the distribution of total claims holds, see for instance Panjer and Willmot [18], Corollary 6.6.1.

$$\mathbb{P}[S=k]=\sum_{m=1}^{k}\Big(a+\frac{bm}{k}\Big)f_X(m)\mathbb{P}[S=k-m]\,,\quad k=1,2,3\ldots\,, \tag{2}$$

where f_X denotes the probability mass function of the claim severity, i.e. $f_X(m)=\mathbb{P}[X=m]$. In case the claim severity distribution is not discrete, one first has to discretize it. Commonly used discretization techniques are either the method of rounding (which is simplest to apply) or the method of moments matching which yields highly accurate results. Details of these discretization techniques can be found in Panjer and Willmot [18], Section 6.15.

Simulation. The recursion and inversion methods presented above assume that the claim sizes $(X_i)_{i\in\mathbb{N}}$ are iid and that N and $(X_i)_{i\in\mathbb{N}}$ are independent. Moreover, the true severity distribution has to be replaced with a discretized approximation. When the X_i's are iid it does not matter which loss is denoted by X_1, which one by X_2 and so on. However, because S is the aggregate loss over a time period, one should account for the time value of money. Simulation techniques such as bootstrapping for example can be useful in this context.

3.2. **Extreme value theory for risk management.** The techniques described in Section 3.1 are widely used in the actuarial world. However, they are not necessarily designed to make inference about the *tail* area of the distribution. Yet from a risk measurement viewpoint it is exactly the tail of the loss distribution which is of particular interest. Here, EVT can offer a solution. In a nutshell, EVT can help to fit a model to the tail distribution of a set of data using only the extreme event data. The *peaks-over-threshold* (POT) method thereby considers exceedances over high thresholds. Simple parametric formulae for measures of extreme risk such as VaR for high confidence levels α for example can be derived based on the so-called generalized Pareto distribution (GPD). As mentioned earlier, the standard reference book for EVT is Embrechts et al. [10]. In what follows, we also refer to McNeil [14] and McNeil and Saladin [16]. The aim of the latter is to investigate under what circumstances EVT yields accurate estimates of high quantiles. A collection of papers relevant for EVT applications to integrated risk management is Embrechts [8].

Originally, EVT evolves from the analysis of (standardised) maxima of a sequence of iid random variables $(X_i)_{i\in\mathbb{N}}$. An important theorem in the realm of extremes says that, for a certain class of distributions, the GPD appears as limiting distribution for the distribution of the excesses $X_i - u$, as the threshold u becomes large. More formally, one can find a positive function $\beta(u)$ such that $G_{\xi,\beta(u)}(x)$ approximates the unknown excess distribution $F_u(x) = \mathbb{P}[\,X - u \le x \,|\, X > u\,]$, where $G_{\xi,\beta}(x)$ is given by

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(\dfrac{1}{1+\xi x/\beta}\right)^{1/\xi} & \text{if } \xi \neq 0\,, \\ 1 - e^{-x/\beta} & \text{if } \xi = 0\,. \end{cases}$$

Here $\beta > 0$, and the support is $x \ge 0$ when $\xi \ge 0$ and $0 \le x \le -\beta/\xi$ when $\xi < 0$. The EVT parameter ξ categorises the various regimes: $\xi > 0$ (the heavy-tailed case), $\xi = 0$ (the medium-tailed case), $\xi < 0$ (the light-tailed case). It is important to stress that, though the $\xi > 0$ case is the most relevant one, EVT yields a methodology for estimating tails of distribution functions in all three cases. To proceed, it is then assumed that for risks X_i with common cdf F the excess distribution F_u follows *exactly* a GPD for some ξ and β, i.e.

$$F_u(x) = G_{\xi,\beta(u)}(x) \tag{3}$$

for a certain high threshold value u. Combining the historical simulation estimate $(n - N_u)/n$ and the tail estimate derived from equation (3), one arrives at the following estimator of F:

$$\hat{F}(x) = 1 - \frac{N_u}{n}\left(1 + \hat{\xi}\,\frac{x-u}{\hat{\beta}}\right)^{-1/\hat{\xi}}, \qquad x > u\,. \tag{4}$$

Note that N_u denotes the number of exceedances over the threshold value u and n the total sample size. It is important to realise that we only fit F beyond the (typically high) level u and that all estimates are a function of this chosen level. The fact that EVT does not fit a model below u is not a weakness, typically sufficient data are available in that range so that standard statistical models can be used. In risk management applications, it is the tail area $x > u$ (for large u) that matters and that is the region where EVT enters. By solving the equation $\hat{F}(q_\alpha) = \alpha$ for q_α and noticing that $q_\alpha = \mathrm{VaR}_\alpha$ we find the following estimate for the VaR at the confidence level α:

$$\widehat{\mathrm{VaR}}_\alpha = u - \frac{\hat{\beta}}{\hat{\xi}}\left(1 - \left(\frac{N_u}{n(1-\alpha)}\right)^{\hat{\xi}}\right)\,. \tag{5}$$

In statistical language, (5) is a quantile estimate, where the quantile is an unknown parameter of an unknown distribution. By means of a method known as profile likelihood it is possible to construct a confidence interval around $\widehat{\mathrm{VaR}}_\alpha$. The interval is asymmetric reflecting the asymmetry in the problem of estimating high quantiles for heavy-tailed data: it is easier to bound the interval from below than to bound it from above. The quantile estimate (5) together with its confidence interval is implemented in EVIS. The corresponding function is called `gpd.q`.

Note that EVIS is a suite of S-Plus functions for EVT developed at ETH Zurich by Alexander McNeil; see `www.math.ethz.ch/~mcneil`. Several authors have applied EVT to operational risk data, see for instance Cruz [7], King [12] and Medova [17].

3.3. **Applicability of standard actuarial and EVT methods.** Sections 3.1 and 3.2 provide the conceptual framework for fitting (aggregate) loss distributions to historical data. It is shown how VaR estimates can be derived by combining historical simulation and EVT techniques. Yet we know that statistical modeling faces a number of potential problems. First, there is the model risk which may lead to wrong results and conclusions when the model is calibrated on unreliable or sparse data. In this section we discuss the feasibility of the methods described above. In particular, we would like to describe the situations in which these techniques are (in-) appropriate. Our emphasis will mainly be on the assumptions underlying EVT since operational risk measurement is tantamount to dealing with extreme loss events. At first, we shall however make some comments on the stationarity or (time-) independence assumptions. Note that our reasoning will be more of a qualitative than a quantitative nature.

3.3.1. *Stationarity.* Many actuarial models, at least in their original form, are based on iid assumptions. In particular, this implies that the time aspect beyond correction for inflation is negligible and that there are no significant structural changes in the observed data as time evolves.

So far it has been hard to come by operational risk losses. One reason is no doubt the confidentiality, another the relatively short historical period over which historical data have been consistently gathered. Yet in the data which are available, we observe the following stylized facts (which seem to be accepted throughout the industry for several operational risk categories):

(a) Loss occurrence times are irregularly spaced in time
(b) Loss amounts very clearly show extremes.

The irregularity seems to go beyond randomness as for instance observed in a homogeneous Poisson process or even renewal process. The observed non-stationarity may in part be due to survival bias: operational losses more than some years ago have not "survived" in banks' databases. On the other hand, non-stationarity can be caused by business cycles, economic cycles, management interactions, regulation, etc. As a result of this serial dependence, volatility changes over time and large losses may tend to occur in clusters. We mentioned at the beginning of Section 3 that the Committee asks for a five-year horizon on which AMA models have to be calibrated on. Given the obvious structural changes of operational risk losses it is clear that sound analytical modeling is a major challenge and definitely can not get away without considering time aspects.

Concerning the (serial) dependence, we notice that the POT method has been generalized in various ways. Typically, excesses over high thresholds u are no longer assumed to occur on inter-arrival times of a homogeneous Poisson process. Moving away from the constant intensity assumption allows to bring in the desired time dependence of excesses over high thresholds. These so-called smoothing effects are discussed in Chavez-Demoulin and Embrechts [5] where further references on (weak-) dependence modeling are to be found. In a recent paper by McNeil and Frey [15], time-dependent VaR models are studied. Taking present and past information into account they consider conditional return distributions for financial return series, thereby incorporating the current volatility structure. Their method allows for determining conditional, i.e. time dependent VaR (and expected shortfall) figures. We stress however that first these non-stationarities have to be modeled before an appropriate EVT analysis can be made.

3.3.2. *(Non-) repetitiveness.* Banks may have a minimal gross threshold for internal loss data collection in place, say €10,000. Nevertheless a bank will track a lot of small to moderate operational risk losses which occur with a certain regularity, say on a monthly, weekly or even daily basis. Hence, the standard actuarial methods as described in Section 3.1 will prove sufficient to analyze them, see for instance Embrechts et al. [9]. As mentioned in Section 3.2, EVT can also be used at this level when the estimator of high quantiles is called for. The abundance of data will make such an analysis fairly standard.

The situation changes drastically if we consider non-repetitive losses. Here – nomen est omen – there is not sufficient data available on which any statistical model could be based on. We are thus in the outermost corner of low frequency-high severity claims. In extreme cases, a single non-repetitive loss alone may drive a financial institution into bankruptcy. The most conspicuous example is provided by the Barings Bank failure in 1995. Clearly, no economic capital charge would have been sufficient to prevent Barings from bankruptcy. The lack of non-repetitive historical data implies that sound high quantile estimates can not be obtained. Statistical modeling, including EVT, reaches its limits here, see Section 3.3.3 for more details on this.

In Table 1, taken from Crouhy et al. [6], we have listed some typical types of operational risks. It is clear from this table that some risks are difficult to quantify (like incompetence under people risk) whereas others lend themselves much easier to quantification (as for instance

			Repetitive?	
1.	**People risk**	• Incompetency		no
		• Fraud		no
2.	**Process risk**			
	(a) Model risk	• Model/methodology error	(yes)	no
		• Mark–to–model error	(yes)	no
	(b) Transaction risk	• Execution error	yes	
		• Product complexity	(yes)	no
		• Booking error	yes	
		• Settlement error	yes	
		• Documentation/contract risk	yes	(no)
	(c) Operational control risk	• Exceeding limits	yes	(no)
		• Security risks	yes	(no)
		• Volume risk	yes	(no)
3.	**Technology risk**	• System failure	yes	
		• Programming error	yes	
		• Information risk	yes	(no)
		• Telecommunications failure	yes	(no)

TABLE 1. Types of operational risks (Crouhy et al. [6], the last column is added).

execution error under transaction risk). We added a column to this table indicating whether a certain risk typically creates repetitive or non-repetitive losses. Sometimes the distinction between "repetitive" and "non-repetitive" is not quite clear. If there is ambiguity (according to our judgment), we use the notion "yes (no)" for example to indicate that this type of risk is *mostly* repetitive. Yet we can think of situations where such losses do not obey a repetitive pattern.

3.3.3. *Number of exceedances.* The choice of a suitable threshold value u above which an EVT analysis can be calibrated is crucial, even for standard, repetitive iid data. A question we are often asked by practitioners is "how many observations in the tail are needed". This is tantamount to asking "at what level do we have to set u". There is no easy answer to these questions beyond "it all depends". For a relevant paper on the choice of the threshold value u, see Victoria-Feser and Dupuis [21] and references therein. Users of EVT, especially in operational risk, should be aware of the various caveats surrounding an EVT analysis; a first glimpse of these caveats can be found in Embrechts et al. [10] under the title "Hill horror plot". Our comments below should be seen as constructive warnings.

The material presented below mainly summarizes the findings of a simulation study worked out by McNeil and Saladin [16]. They compare the estimated quantiles (5) with the corresponding theoretical ones for known distributions for which (high) quantiles can be calculated explicitly. By the rejection method, independent random numbers with cdf F are generated until N_u excesses over u have been detected. They consider data sets of $N_u \in \{25, 50, 100, 200\}$ exceedances. 25 exceedances over thresholds are considered as the minimum number one would work with; estimates from less data would be too unreliable. 200 exceedances would be ideal and 50 to 100 realistic situations. For comparison purposes, McNeil and Saladin introduce a somewhat arbitrary, albeit natural definition of a "good" estimate. Their criteria in assessing the estimate (5) is based on the empirical bias and standard error.

The simulation study was performed for three types of loss distributions, all of which occur within operational (and also credit) risk literature.

(i) Medium-tailed (e.g. Lognormal, gamma).
(ii) Heavy-tailed with infinite moments of order greater than or equal to two.
(iii) Heavy-tailed with infinite moments of order greater than or equal to one.

Tables 2 to 4 below give a qualitative summary of the simulation results to be found in McNeil and Saladin [16]. For illustrative purposes, we present here the case of lognormally distributed claims representing class (i) ($\xi = 0$ in the EVT notation) and claim sizes following a Pareto distribution with shape parameters $\theta = 2$, $\theta = 1$, respectively as characteristic examples of classes (ii) and (iii). The Pareto is a commonly used heavy-tailed loss distribution with cdf $F(x) = 1 - (a/x)^\theta$, $x \geq a$ ($\xi = 1/\theta > 0$). Our choice is motivated by the relevance of these distributions in actuarial and financial modeling. The results in McNeil and Saladin [16] extend to other types of distributions such as the t-distribution for example. Note that the tail severity increases as we move from class (i) to class (iii). Again, we would like to stress the fact that the figures given below represent an ideal situation where losses are assumed to be iid, hence highly repetitive. As such, these values could be regarded as a bottom line (a best case analysis) when it comes to modeling operational loss data. Recall that the stylized facts of several, important operational risk classes are not consistent with the modeling assumptions made in this example.

Lognormal distribution:

$u = F^{\leftarrow}(q)$	α	Goodness of $\widehat{\mathrm{VaR}}_\alpha$
$q = 0.7$	0.99	A minimum number of 50 exceedances (corresponding to 167 observations) is required to ensure accuracy wrt bias *and* standard error.
	0.999	A minimum number of 100 exceedances (corresponding to 333 observations) is required to ensure accuracy wrt bias *and* standard error.
$q = 0.9$	0.99	Full accuracy can be achieved with the minimum number 25 of exceedances (corresponding to 250 observations).
	0.999	Full accuracy can be achieved with the minimum number 25 of exceedances (corresponding to 250 observations).

TABLE 2. Accuracy of estimating high quantiles by means of the POT method. Lognormally distributed claims.

Pareto distribution with $\theta = 2$:

$u = F^{\leftarrow}(q)$	α	Goodness of $\widehat{\mathrm{VaR}}_\alpha$
$q = 0.7$	0.99	A minimum number of 100 exceedances (corresponding to 333 observations) is required to ensure accuracy wrt bias *and* standard error.
	0.999	A minimum number of 200 exceedances (corresponding to 667 observations) is required to ensure accuracy wrt bias *and* standard error.
$q = 0.9$	0.99	Full accuracy can be achieved with the minimum number 25 of exceedances (corresponding to 250 observations).
	0.999	A minimum number of 100 exceedances (corresponding to 1000 observations) is required to ensure accuracy wrt bias *and* standard error.

TABLE 3. Accuracy of estimating high quantiles by means of the POT method. Pareto distributed claims with shape parameter $\theta = 2$.

Pareto distribution with $\theta = 1$:

$u = F^{\leftarrow}(q)$	α	Goodness of $\widehat{\mathrm{VaR}}_\alpha$
$q = 0.7$	0.99	For all number of exceedances up to 200 (corresponding to a minimum of 667 observations) the VaR estimates fail to meet the accuracy criteria.
	0.999	For all number of exceedances up to 200 (corresponding to a minimum of 667 observations) the VaR estimates fail to meet the accuracy criteria.
$q = 0.9$	0.99	A minimum number of 100 exceedances (corresponding to 1000 observations) is required to ensure accuracy wrt bias *and* standard error.
	0.999	A minimum number of 200 exceedances (corresponding to 2000 observations) is required to ensure accuracy wrt bias *and* standard error.

TABLE 4. Accuracy of estimating high quantiles by means of the POT method. Pareto distributed claims with shape parameter $\theta = 1$.

It readily follows from Tables 2 to 4 that the heavier the tails the larger the sample sizes should be to obtain the desired accuracy. Also it pays to have sufficient data far in the tail, i.e. q close to 1. To further clarify matters, we illustrate how the above results could be used in practice provided there is strong evidence that the data fulfil the modeling assumptions. To begin with, note that the simulation study starts with a fixed number N_u of exceedances. Random losses are then generated until N_u exceedances are detected. In reality though, this approach has to be "inverted" meaning that one has to work with a sample of n losses occurred in a pre-specified time interval from which the number of exceedances N_u can then be determined.

For illustrative purposes, we consider the following example. Let us assume that a total number of $n = 287$ losses were observed over a time period of one year, say. Moreover, suppose that there is strong evidence that the loss data are of class (ii)-type, e.g. Pareto distributed with shape parameter $\theta = 2$ (to justify the latter assumption one can plot the empirical tail distribution on a log-log scale and observe that the data points follow a straight line with slope -2). Also note that this assumption is realistic for several operational risk classes. It then follows from Table 3 that the VaR at confidence level $\alpha = 0.99$ can be estimated with the targeted accuracy. Note that it does not really matter whether we place the threshold value u at the 90%-level or at the 70% level as the corresponding minimal numbers are about of the same order. It is also evident that one should abstain from estimating the VaR at the 99.9%-level. The figures in Table 3 show that $n = 287$ observations are far too less: For a sound 99.9%-VaR estimate under these idealistic assumptions notably, one should have at least some 670 or more losses at one's disposal. If, however, the data were of class (i)-type, then estimating the VaR at the 99.9% level with $n = 287$ observed data points could be justified, see Table 2. Moving in the other direction towards class (iii)-type loss distributions, we notice that 287 observations, assumed to be iid and repetitive, are even not enough to estimate the VaR at the 99% level.

4. Conclusion

Under the proposed New Accord (Basel II), operational risk has to be treated explicitly. It is envisaged that internationally active banks with a substantial exposure to operational risk losses adopt the advanced measurement approach (AMA). Banks are given a great deal of flexibility in how to implement the AMA. However, the Committee requires banks to demonstrate their ability in capturing potentially severe tail loss events. To be more precise, banks should put an operational risk capital aside in line with the 99.9% confidence level over a one-year holding period. We have shown how actuarial techniques in principle could be used for estimating (high) quantiles of unknown loss distributions. Special emphasis was given to extreme value theory (EVT). EVT can help to fit a model to the tail distribution of a set of data using only the extreme event data. Still, estimation of high quantiles is an inherently difficult problem. As we have seen, 200 or more excesses over the 90%-quantile would often be necessary to obtain VaR estimates with the targeted accuracy. Note that these figures remain of indicative nature though. We have seen that the stylized facts of historical operational risk losses in general are not in accordance with iid modeling assumptions. Yet we do not want to insist that this inconsistency applies to every loss type category. Only preliminary explanatory analysis will show for which business line and event type actuarial techniques can readily be applied. In most cases, however, one will observe structural changes in operational risk data as time evolves. To take this serial dependence of the data into account, one could refine the POT method towards smoothing techniques.

For these reasons, we argue that Pillar 1 in the operational risk management framework should not be overemphasized. For repetitive and stationary losses the standard actuarial methods and their refinements can be employed to derive capital charges. The crux, however, pertains the non-repetitive and non-stationary case. And it is exactly the losses of the latter category which jeopardize the existence of financial institutions. VaR estimates, even though complemented by stress testing and scenario analysis, can never be viewed as a "stand-alone" risk management tool. Keeping in mind that most serious operational risk losses can not be judged as mere accidents, it becomes obvious that the only way to gain control over operational risk is to improve the quality of control over the possible sources of huge operational losses. It is exactly here that Pillar 2, and to a less extent Pillar 3, becomes extremely important.

Acknowledgement

The authors would like to thank Harry Stordel for fruitful discussions which led to the current version of the paper. This paper was written while Paul Embrechts was Centennial Professor of Finance at the Department of Accounting and Finance of the London School of Economics.

References

[1] Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance.* Vol. 9, No. 3. 203-228.
[2] Basel Committee on Banking Supervision. *Sound Practices for the Management and Supervision of Operational Risk.* December 2001, revised July 2002. BIS, Basel, Switzerland
[3] Basel Committee on Banking Supervision. *The New Basel Capital Accord.* April 2003. BIS, Basel, Switzerland.
[4] Basel Committee on Banking Supervision. *Overview of The New Basel Capital Accord.* Consultative Document. April 2003. BIS, Basel, Switzerland.
[5] Chavez-Demoulin, V., and Embrechts, P. (2003). Smooth extremal models in finance and insurance. To appear in *Journal of Risk and Insurance.*
[6] Crouhy, C., Galai, D., and Mark, R. (2000). *Risk Management.* McGraw–Hill, New York.
[7] Cruz, M. G. (2002). *Modelling, Measuring and Hedging Operational Risk.* Wiley.
[8] Embrechts, P. (2000). (Editor) *Extremes and Integrated Risk Management.* Risk Waters Group, London.
[9] Embrechts, P., Kaufmann, R., and Samorodnitsky, G. (2002). Ruin theory revisited: stochastic models for operational risk. Submitted.
[10] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.
[11] Jensen, J. L. (1997). *Saddlepoint Approximations.* Oxford University Press.
[12] King, J. L. (2001). *Operational Risk.* Wiley.
[13] Klugman, S. A., and Panjer, H., and Willmot, G. E. (1998). *Loss Models: From Data to Decisions.* Wiley.
[14] McNeil, A. J. (1999). Extreme value theory for risk managers. *Internal Modelling and CAD II.* Published by *RISK Books*, 93-113.
[15] McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, Vol. 7, 271-300.

[16] McNeil, A. J., and Saladin, T. (1997) The peaks over thresholds method for estimating high quantiles of loss distributions. *Proceedings of XXVIIth International ASTIN Colloquium*, Cairns, Australia, 23-43.
[17] Medova, E. (2000). Measuring risk by extreme values. *Risk.* November 2000, 20-26.
[18] Panjer, H., and Willmot, G. (1992). *Insurance risk models.* Society of Actuaries, Schaumberg, IL.
[19] Pézier, J. (2002). A Constructive Review of Basel's Proposals on Operational Risk. *ISMA Discussion Papers in Finance. 2002-20.* To appear in *Mastering Operational Risk.* FT-Prentice Hall, 2003.
[20] Pézier, J. (2002). Operational Risk Management. *ISMA Discussion Papers in Finance. 2002-21.* To appear in *Mastering Operational Risk.* FT-Prentice Hall, 2003.
[21] Victoria-Feser, M. P., and Dupuis, D. J. (2003). A prediction error criterion for choosing the lower quantile in Pareto index estimation. Preprint, University of Geneva.

PAUL EMBRECHTS is Professor of Mathematics at the ETH Zurich, Centennial Professor of Finance at the London School of Economics and currently director of RiskLab.

HANSJÖRG FURRER holds a PhD in mathematics from the ETH Zurich an has worked for Swiss Re. He is currently working at RiskLab.

ROGER KAUFMANN is a PhD student in mathematics at the ETH Zurich. In 1999 he received his mathematics diploma from the ETH Zurich. From 1999 until 2003, additionally to writing his PhD thesis, he was a researcher at the financial and insurance mathematics institute RiskLab.

DEPARTMENT OF MATHEMATICS
ETH ZURICH
CH–8092 ZURICH
SWITZERLAND