



## How to Avoid Over-estimating Capital Charge for Operational Risk?\*

Nicolas Baud, Antoine Frachot and Thierry Roncalli  
Groupe de Recherche Opérationnelle, Crédit Lyonnais, France<sup>†</sup>

This version: December 1, 2002

The present document reflects the methodologies, calculations, analyses and opinions of their authors and is transmitted in a strictly informative aim. Under no circumstances will the above-mentioned authors nor the Crédit Lyonnais be liable for any lost profit, lost opportunity or any indirect, consequential, incidental or exemplary damages arising out of any use or misinterpretation of the present document's content, regardless of whether the Crédit Lyonnais has been apprised of the likelihood of such damages.

Le présent document reflète les méthodologies, calculs, analyses et positions de leurs auteurs. Il est communiqué à titre purement informatif. En aucun cas les auteurs sus-mentionnés ou le Crédit Lyonnais ne pourront être tenus pour responsables de toute perte de profit ou d'opportunité, de toute conséquence directe ou indirecte, ainsi que de tous dommages et intérêts collatéraux ou exemplaires découlant de l'utilisation ou d'une mauvaise interprétation du contenu de ce document, que le Crédit Lyonnais ait été informé ou non de l'éventualité de telles conséquences.

---

\*We thank Giulio Mignola (Sanpaolo IMI) and Maxime Pennequin (Crédit Lyonnais) for stimulating discussions.

<sup>†</sup>address: Crédit Lyonnais - GRO, Immeuble Zeus, 4<sup>è</sup> étage, 90 quai de Bercy — 75613 Paris Cedex 12 — France;  
E-mail: antoine.frachot@creditlyonnais.fr or thierry.roncalli@creditlyonnais.fr

# 1 Introduction

The Basel Committee recently agreed to eliminate the separate floor capital requirement that had been proposed for the Advanced Measurement Approaches (AMA). As a result there is no more regulatory limit to the reduction of the capital charge which is obtained by using AMA in comparison with other methodologies (such as Basic Indicator Approach and Standard Approach). This represents a strong incentive for banks to develop internal models (via a Loss Distribution Approach or LDA) in order to get a correct grasp of their true risks and to compute more accurate capital requirements than with other one-size-fits-all methods.

Contrary to other methods which compute capital charges as a proportion of some exposure indicators (e.g. gross income), LDA takes its inspiration from credit risk or market risk internal models where frequency and severity distributions are compounded in order to evaluate the 99.9% quantile of the total loss amount. In practice, as the compounding process does not result in closed-form expressions, Monte-Carlo simulations are necessary for computing these quantiles. Most practitioners will agree on the idea that this is not the most difficult part of the process as Monte Carlo technology is now a standard skill among quantitative analysts. On the contrary, the calibration of loss distributions is the most demanding task because of the shortage of good-quality data.

First, risk managers do not have access to many data since most banks have started collecting data only recently. Therefore, internal loss data must be supplemented by external data from public and/or pooled industry databases. Unfortunately, incorporating external data is rather dangerous and requires careful methodology to avoid the now widely-recognised pitfalls regarding data heterogeneity, scaling problems and lack of comparability between too heterogeneous data. Our experience at Cr dit Lyonnais has taught us that incorporating external data directly into internal databases leads to totally flawed results.

As a matter of fact, the main problem lies in the data generating processes which underlies the way data have been collected. In almost all cases, loss data have gone through a truncation process by which data are recorded only when their amounts are higher some (possibly ill-defined) threshold. As far as internal data are concerned, these thresholds are defined by the global risk management policy. In practice, banks' internal thresholds are set in order to balance two conflicting wishes: collecting as many data as possible while reducing costs by collecting only significant losses. In the same spirit, industry-pooled databases try to enforce an explicit threshold. Finally, public database also pretend that they record losses only above some threshold (generally much higher than for industry-pooled data). Whichever type of databases (internal, industry-pooled, public) we talk about, they are all truncated with various cut-offs and then can not be compared with one another nor pooled together without any care. Furthermore, one may suspect actual thresholds to be rather different from stated thresholds:

- (Internal data) Even though thresholds have been imposed by the global risk management, they cannot always be made enforceable at business units level since small business units or under-staffed units may be unable to comply.
- (Industry-pooled data) As no enforcement process does really exist, nothing ensures that contributors actually comply with stated thresholds.
- (Public data) It is even worse for public database since they are fed with publicly-released losses with no guarantee that all losses are recorded in an homogeneous way and according to the same threshold.

As a result, we have come to the conclusion that stated thresholds can not be taken for granted and should be considered as unknown parameters which have to be estimated. Secondly, actual thresholds are likely to be higher than stated thresholds for the very same reasons as given above. Therefore, loss data - especially industry-pooled and public data - may be severely biased towards high losses, resulting in over-estimated capital charges. This issue is generally addressed by saying that, in short,

internal data must be used for calibrating the main body of the severity distribution while external data should be related to the tail of the distribution. As far as our knowledge, this methodology has not yet received a rigorous description and has more to do with art than statistics. On the contrary, we have tried to build a sound methodology based on maximum-likelihood principles. Our main idea is to consider that the main source of heterogeneity comes from the different thresholds which underly data generating processes. As a consequence, thresholds should be taken into account explicitly in the calibration process in the form of additional parameters which have to be estimated along with other parameters of the loss distribution. Provided that thresholds are carefully managed, internal and external data are made comparable and the most part of heterogeneity is then eliminated. It is also worth mentioning that, since our methodology relies on maximum-likelihood principles, we can prove (in particular to our supervisors) that it is statistically sound and provides unbiased capital charges.

The paper is organized as follows. We first provide a classification of the different bias which come from the data generating process from which operational risk loss data are drawn. According to this classification, we then propose a rigorous statistical treatment to correct all biases. Finally we provide a real-life experiment to show how our methodology can be applied. In particular, we show that, if thresholds are ignored (as commercial softwares often do), then capital requirements are considerably over-estimated by up to 50 % or even more !

## 2 A typology of operational risk loss data

In this section, we discuss how external databases are built, which is a good starting point for assessing to what extent operational risk databases are biased. Two types of external databases are encountered in practice.

- The first type corresponds to databases which record publicly-released losses. In short these databases are made up of losses that are far too important or emblematic to be concealed away from public eyes. The first version of OpVar<sup>®</sup> Database pioneered by PwC is a typical example of these first-generation external databases.
- More recent is the development of databases based on a consortium of banks. It works as an agreement among a set of banks which commit to feed a database with their own internal losses, provided that some confidentiality principles are respected. In return banks which are involved in the project are of course allowed to use these data to supplement their own internal data. Gold of BBA (*British Bankers' Association*) is an example of consortium-based data.

The two types of database differ by the way losses are supposed to be truncated. In the first case, as only publicly-released losses are recorded, the truncation threshold is expected to be much higher than in the consortium-based data. For example, the OpVar<sup>®</sup> Database declares to record losses greater than USD 1 million while consortium-based data pretend to record all losses greater than USD 25.000 for **ORX** database (or USD 10.000 by 2003, see PEEMÖLLER [7]).

Furthermore public databases, as we name the first type of external databases, and industry-pooled databases differ not only by their stated threshold but also by the level of confidence one can place on it. For example, nothing ensures that the threshold declared by a industry-pooled database is the actual threshold as banks are not necessarily able to uncover all losses above this threshold even though they pretend to be so<sup>1</sup>. Rather one may suspect that banks do not have always the ability to meet this requirement yet. As a result, stated thresholds must be seen more like a long-term target than a strong commitment.

As said before, the same argument applies to internal databases in a lesser but significant extent. Business units inside a bank are supposed to report their losses according to some guidelines defined

---

<sup>1</sup>The **ORX** project seems more ambitious and proposes reporting control and verification. In particular, the financial institution must prove its capability to collect and to deliver data if it wants to be a member of the **ORX** consortium.

by the global risk management. In practice, business units do not always have the resources necessary to comply. As a result, internal database suffer from truncation bias as well.

As an example, following is the kind of data risk managers have to deal with:

- dataset 1 from business unit 1 which declares to report – and does report effectively – loss amounts above (say) 10.000 euros;
- dataset 2 from business unit 2 which is in the same position as business unit 1 but with a threshold of 20.000;
- dataset 3 from business unit 3 which pretends to report above 10.000 euros but whose quality of reporting channels does not ensure it really does;
- industry-pooled database which is fed by many contributors with different and unknown thresholds, or with thresholds that are suspected to be different from the stated threshold;
- etc.

Risk managers and quantitative analysts have to use such heterogeneous data generating processes. Unfortunately, as it will become obvious in the sequel, calibration is dramatically distorted and capital charges are severely over-estimated if these data are pooled together without any care.

### 3 How to make data comparable ?

Our starting point says that the sample loss distribution is fundamentally different from the true loss distribution. In statistical terms, the sample distribution has more to do with conditional distributions, i.e. probability distribution conditionnally to losses higher than some thresholds. This is where maximum likelihood appears.

Maximum likelihood is an asymptotically efficient method provided that the likelihood is correctly specified, i.e. the sample distribution has been derived correctly. We note  $f(\cdot; \theta)$  the (true) loss distribution where  $\theta$  is a parameter characterizing this distribution. In the case of a log-normal distribution,  $\theta$  is no more than the mean and the variance of the logarithm of the losses.

Since data are recorded above a threshold that we shall denote by  $H$ , the sample loss distribution  $f^*(\cdot; \theta)$  is equal to the true loss distribution conditionnally to the loss exceeding  $H$ , that is:

$$f^*(\zeta; \theta) := f(\zeta; \theta | H = h) = \mathbf{1}\{\zeta \geq h\} \cdot \frac{f(\zeta; \theta)}{\int_h^{+\infty} f(x; \theta) dx}$$

Three cases may be encountered in practice:

- Threshold  $H$  is known for sure, i.e. actual threshold equals stated threshold. This is the ideal case but also the less likely one. As we discuss before, it is safer to consider that stated and actual thresholds may differ. As a result,  $H$  should be considered as unknown and calibrated along with  $\theta$ .
- Threshold  $H$  is unknown.
- There are more than one threshold. The multi-threshold case corresponds exactly to industry-pooled data since there are a priori as many thresholds as contributors are involved. In the limiting case of a public database, the number of contributors should be considered as almost infinite, meaning that  $H$  follows a continuous distribution.

As a result, in the multi-threshold case, the additional parameters are not only the thresholds  $h_1, \dots, h_n$  where  $n$  is the number of contributors, but also the weights  $p_1, \dots, p_n$  of each contributor (i.e. the number of loss data it has provided relatively to the total number of loss data). Finally, the likelihood must be based on the sample probability distribution:

$$f^*(\zeta; \theta, (h_i), (p_i)) \equiv f(\zeta; \theta \mid H_1 = h_1) \propto p_1 + \dots + f(\zeta; \theta \mid H_n = h_n) \propto p_n$$

Therefore, the total set of parameters is now:

$$\theta, h_1, \dots, h_n, p_1, \dots, p_n$$

Let us mention that most commercial softwares do not take into account the thresholds. With our notation, these softwares consider that  $n = 0$  and that the sample distributions of the different datasets are identically equal to the true distribution. As far as our knowledge, one commercial software considers the case of  $n = 1$  (with obviously  $p_1 = 1$ ). Its methodology is very close to what FRACHOT and RONCALLI [6] have proposed. In short,  $\theta$  is calibrated with different  $h$ , which provides the  $h \mapsto \theta(h)$  curve.  $H$  is then graphically determined like the inflexion point of the curve. It can be mathematically proved that this rule of thumb gives a correct answer when  $n$  is actually equal to 1. This methodology has however two pitfalls. First it is biased when more than one threshold are at stake (i.e. industry-pooled data). Secondly, even in the single-threshold case, it leads to severe loss of accuracy since all loss data lower than the inflexion point have to be dropped from the calibration process. As a result, the unbiasedness property is obtained at the expense of a loss of accuracy.

The multi-threshold case which is the most likely in practice is much harder to treat correctly. It requires high-level optimization algorithms we have imported from our past experience in internal market risk and credit risk models. The point is that we now have a tool which deals with the general case as exemplified in the following section.

## 4 Real-life capital computations

Calibration procedures are applied to a real-life example. Our tested procedures cover the whole spectrum ranging from “naive” calibration (i.e. ignoring any potential thresholds) to full-information maximum likelihood as just described.

### 4.1 Data description

As it is out of question to provide information about Crédit Lyonnais loss data, we have simulated several datasets and the different calibration procedures are applied on these datasets. Let us suppose we have to compute the capital requirement for one risk-type (say for example ‘External Fraud’). Data come from 3 different sources whose data generating process is the following:

- Dataset 1 (from business unit 1): losses are truncated above 10 k euros. Reporting processes have been fully audited concluding that business unit 1 complies with its stated threshold.
- Dataset 2 (from business unit 2): same but with a threshold equal to 15 k euros.
- Dataset 3 (from industry-pooled database): losses are supposed to be reported above 10 k euros but, in practice, contributors are not at the same level of compliance. So there are as many (unknown) thresholds as contributors. Moreover, since losses are anonymized, losses cannot be linked to any specific contributor. External data are thus drawn from a mixture of different data generating processes.

Simulations are specified in the following way:

- The true loss distribution is log-normal  $\mathcal{LN}(m, \sigma^2)$  with  $m = 8$  and  $\sigma = 2$ . All losses are independently drawn from this probability distribution. The mean loss is thus equal to 22 k euros.

- The number of losses are respectively 2000 for dataset 1, 2500 for dataset 2 and 5000 for dataset 3.
- Dataset 3 is made of 3 contributors. Contributors' actual thresholds are  $h_1 = 10$  k euros (1000 losses),  $h_2 = 20$  k euros (1500 losses),  $h_3 = 50$  k euros (2500 losses).

Following are some preliminary statistics. The true mean, variance and quantiles are given in brackets when relevant.

	Actual threshold	Nbr Losses	Sample $\mu$	Actual $\mu$	Sample $\sigma$	Actual $\sigma$
Dataset 1	10 k euros	2000	10.39	8.00	0.97	2.00
Dataset 2	15 k euros	2500	10.75	8.00	0.95	2.00
Dataset 3	10, 20, 50 k euros	5000	11.25	8.00	1.00	2.00

We see that truncation implies that the sample mean loss (resp. variance) is much higher (lower) than for the true distribution. This gives an example of how different the sample and the true distributions may be. Let us now compute the Capital-At-Risk that we would obtain if we used the sample  $\mu$  and  $\sigma$  corresponding to each dataset. Computations are performed under the assumption that the frequency distribution follows a Poisson distribution with a mean of 500 events per year.

	CaR (99.9%).	Actual CaR (99.9%)
Dataset 1	32.9 Mn euros	37.8 Mn euros
Dataset 2	45.9 Mn euros	37.8 Mn euros
Dataset 3	80.1 Mn euros	37.8 Mn euros

It is obvious that capital charge computations are totally flawed and may be dramatically over-estimated.

## 4.2 Calibration procedures

Considering these “real-life” datasets, we then test the following procedures:

- **Procedure 1:** merge the 3 datasets together, which gives one single merged dataset. Apply maximum likelihood ignoring any threshold effect, exactly as most commercial softwares would do.
- **Procedure 2:** merge the 3 datasets together again. Apply maximum likelihood principle under the assumption that all datasets share the same threshold ( $n = 1$ ). The implicit threshold is calibrated as in BAUD, FRACHOT and RONCALLI [4]. We solve (using obvious notations):

$$\hat{\theta} = \arg \max \begin{aligned} & \sum_{i \in dataset.1} \ln f^*(\zeta_i; \theta, h, p = 1) + \\ & \sum_{i \in dataset.2} \ln f^*(\zeta_i; \theta, h, p = 1) + \\ & \sum_{i \in dataset.3} \ln f^*(\zeta_i; \theta, h, p = 1) + \end{aligned}$$

for  $h$  ranging from 0 to 100 k euros.  $h$  is then estimated graphically as the inflexion point, that is the threshold above which the estimate  $\hat{\theta}$  stabilizes.

- **Procedure 3:** merge the 3 datasets together again. Apply maximum likelihood principle in the general case where the number of relevant thresholds is unknown. The maximum likelihood program is written as:

$$\left( \hat{\theta}, \hat{h}_1, \dots, \hat{h}_n, \hat{p}_1, \dots, \hat{p}_n \right) = \arg \max \begin{aligned} & \sum_{i \in dataset.1} \ln f^*(\zeta_i; \theta, (h_i), (p_i)) + \\ & \sum_{i \in dataset.2} \ln f^*(\zeta_i; \theta, (h_i), (p_i)) + \\ & \sum_{i \in dataset.3} \ln f^*(\zeta_i; \theta, (h_i), (p_i)) + \end{aligned}$$

- **Procedure 4:** do not merge the 3 datasets. Apply maximum likelihood by using the exact conditionnal distribution for datasets 1 and 2. This implicitly assumes that risk managers know the exact threshold for the two internal datasets. The maximum likelihood program is written as:

$$\left(\hat{\theta}, \hat{h}_1, \dots, \hat{h}_n, \hat{p}_1, \dots, \hat{p}_n\right) = \arg \max \begin{aligned} & \sum_{i \in \text{dataset.1}} \ln f^*(\zeta_i; \theta, h_i = 10000) + \\ & \sum_{i \in \text{dataset.2}} \ln f^*(\zeta_i; \theta, h = 20000) + \\ & \sum_{i \in \text{dataset.3}} \ln f^*(\zeta_i; \theta, (h_i), (p_i)) + \end{aligned}$$

**Procedure 1** poses no problem with standard statistical package we may find in commercial softwares. As said before, **Procedure 2** is being developed by one software. As far as our knowledge, **Procedures 3** and **4** have been implemented only at Crédit Lyonnais.

### 4.3 Empirical results

**Procedure 1** ignores all thresholds and truncation biases. The parameters of the loss distribution are estimated as:

$$\hat{\mu} = 10.94 \text{ and } \hat{\sigma} = 1.04$$

which have to be compared with the true parameters,  $\mu = 8$  and  $\sigma = 2$ . It is clear that **Procedure 1** is totally flawed although there still exist consultants who propose this procedure in their commercial offers. Regarding capital charge<sup>2</sup>, the procedure gives a 99.9% capital-at-risk of 61.8 Mn euros while the true capital charge is 37.8 Mn euros, that is a more than 50% over-estimation !

**Procedure 2** assumes one single threshold which is calibrated graphically. We can locate the threshold at approximately  $h = 50$  which corresponds to the highest threshold in our data. Parameters are found to be equal to:

$$\hat{\mu} = 8.41 \text{ and } \hat{\sigma} = 1.90$$

This is not so bad but it requires to drop almost one half of available data (i.e. all data less than 50 k euros). In particular, it implies a severe loss of accuracy. If we had performed our computations with a less favorable context (fewer internal data, higher thresholds for external data), the results would have been strongly inaccurate.

**Procedure 3** considers the general case. Our procedure finds 4 thresholds:

	$h_1$	$h_2$	$h_3$	$h_4$	$p_1$	$p_2$	$p_3$	$p_4$
Calibrated	10.02	15.00	20.09	49.61	32.9%	30.3%	11.4%	25.3%
Actual	10	15	20	50	31.6%	26.3%	15.8%	26.3%

where the actual weights are obtained as follows:

$$\begin{aligned} \text{Total number of losses} &= 2000 + 2500 + 5000 = 9500 \\ p_1 &= \frac{2000 + 1000}{9500} = 31.6\% \\ p_2 &= \frac{2500}{9500} = 26.3\% \\ p_3 &= \frac{1500}{9500} = 15.8\% \\ p_4 &= \frac{2500}{9500} = 26.3\% \end{aligned}$$

From maximum likelihood properties, procedure 3 gives a consistent estimate of the thresholds and their associated weights. In our sample, the main parameters are estimated as:

$$\hat{\mu} = 8.49 \text{ and } \hat{\sigma} = 1.88$$

---

<sup>2</sup>We remind that the frequency distribution is taken as a Poisson process whose mean equals 500 events per year.

Finally, regarding capital requirements, we obtain:

$$\text{CaR} = 40.0 \text{ Mn euros}$$

which is much closer to the true capital-at-risk (37.8 Mn euros).

**Procedure 4** is more demanding in terms of information since it requires to know that datasets 1 and 2 do not result from a mixture of different distributions with different thresholds, contrary to dataset 3. Instead, each of both datasets is associated to one single threshold. This is a valuable information, which in turn should improve the accuracy of our estimates. We obtain

$$\hat{\mu} = 7.95 \text{ and } \hat{\sigma} = 2.00$$

and

	$h_1$	$h_2$	$h_3$	$p_1$	$p_2$	$p_3$
Calibrated	10.03	19.9	50.1	16.9%	31.0%	50.1%
Actual	10	20	50	20%	30%	50%

Capital charge is then equal to:

$$\text{CaR} = 36.8 \text{ Mn euros}$$

Even though one cannot conclude with only one trial, our results confirm that due to maximum likelihood properties, **Procedure 4** estimators seem more accurate than with any other procedure. However one cannot recommend **Procedure 4** because it relies on the idea that business units comply perfectly to the stated thresholds, which at least should be confirmed by statistical tests. As a conclusion, **Procedure 3** is less efficient but it is able to deal with any data, no matter they fall into the single-threshold or multi-thresholds category.

## 5 Future development

Previous calculations have been performed with the tool we have developed. All previous methodologies have been implemented for log-normal, exponential and Weibull distributions. It is remarkable that it permits us to achieve the previous calculations in few minutes through a very user-friendly Excel-based interface. Following are the further minor developments that will be added soon.

Our methodology is fully efficient for uncovering parameters of the true loss distribution. It is a direct consequence of maximum likelihood properties and then are likely to be accepted by supervisors. However, we are aware from our past experience that capital-at-risk calculations are quite sensitive to these parameters. Therefore, supervisors will certainly demand banks to be able to bound their capital-at-risk estimates into a confidence interval. This point is already mentioned in the last Basel II paper [2].

Theoretically, this task is quite complicated because here confidence interval should aggregate three sources of uncertainty: from parameters  $\mu$ ,  $\sigma$ , from parameters  $h$  and  $p$ , and from the fact that capital-at-risk are computed by Monte Carlo simulations. This latter source of uncertainty is unimportant since it can be reduced to almost zero provided that a sufficient number of simulations are drawn. It takes computing time but time is not at stake for capital charge calculations. The uncertainty surrounding thresholds' estimates and their weights is probably not a cause of concern because capital-at-risk does not seem to be very sensitive to small errors on these parameters (provided that parameters  $\mu$ ,  $\sigma$  have been consistently estimated, i.e. procedure 3 or 4 have been performed). Nonetheless, this point remains to be proved theoretically but it would be rather difficult as a rapid inspection of the log likelihood reveals that it is not differentiable with respect to these parameters. As far as we are aware, the derivation of interval confidence for such kind of parameters (with respect to which the log-likelihood behaves badly) results from quite recent and complex econometric papers. Our intuition is that it is not worth investigating this point any further. Finally, the main source of uncertainty



probably comes from parameters  $\mu, \sigma$ . Contrary to thresholds parameters, the derivation of confidence intervals is easy and requires simply the second derivative of the log-likelihood (with respect to  $\mu, \sigma$ ) which is a by-product of the estimation process.

A second issue would be worth being investigated. It concerns the implementation of rigorous statistical tests to decide whether some weights are equal to zero and then should be removed from the estimation process. In practice, we increase the number of possible thresholds  $n$  as long as the log-likelihood keeps growing. However it is quite easy to implement a statistical test giving the optimal  $n$  above which no (statistically) significant increase of the likelihood is to be expected. Contrary to the computation of confidence interval, it does need neither the second derivative nor the first one. The so-called likelihood ratio test should provide the answer with minimum computations.

## 6 Concluding Remarks

Intense reflections are being conducted at the moment regarding the way to pool heterogeneous data coming from both banks' internal systems and industry-pooled databases. We propose here a sound methodology. As it relies on maximum likelihood principle, it is thus statistically rigorous and should be accepted by supervisors. We believe that it solves the most part of data heterogeneity and scaling issues.

## References

- [1] Basel Committee on Banking Supervision, Working Paper on the Regulatory Treatment of Operational Risk, September 2001
- [2] Basel Committee on Banking Supervision, Quantitative Impact Study 3. Technical Guidance, October 2002
- [3] BAUD, N., A. FRACHOT, and T. RONCALLI [2002], An internal model for operational risk computation, Crédit Lyonnais, Groupe de Recherche Opérationnelle, *Slides* of the conference “Seminarios de Matemática Financiera”, Instituto MEFF – Risklab, Madrid (<http://gro.creditlyonnais.fr>)
- [4] BAUD, N., A. FRACHOT, and T. RONCALLI [2002], Internal data, external data, consortium data: how to mix them for measuring operational risk, Crédit Lyonnais, Groupe de Recherche Opérationnelle, *Working Paper* (<http://gro.creditlyonnais.fr>)
- [5] FRACHOT, A., P. GEORGES and T. RONCALLI [2001], Loss Distribution Approach for operational risk, Crédit Lyonnais, Groupe de Recherche Opérationnelle, *Working Paper* (<http://gro.creditlyonnais.fr>)
- [6] FRACHOT, A. and T. RONCALLI [2002], Mixing internal and external data for managing operational risk, Crédit Lyonnais, Groupe de Recherche Opérationnelle, *Working Paper* (<http://gro.creditlyonnais.fr>)
- [7] PEEMÖLLER, F.A. [2002], Operational risk data pooling, Deutsche Bank AG, *Presentation at CFSforum – Operational Risk*, Frankfurt/Main