Mixing internal and external data for managing operational risk

Antoine Frachot* and Thierry Roncalli

Groupe de Recherche Opérationnelle, Crédit Lyonnais, France

This version: January 29, 2002

1 Introduction

According to the last proposals by the Basel Committee [1], banks are allowed to use the Advanced Measurement Approaches (AMA) option for the computation of their capital charge covering operational risks. Among these methods, the Loss Distribution Approach (LDA) is the most sophisticated one (see Frachot, Georges and Roncalli [2001] for an extensive presentation of this method). It is also expected to be the most risk sensitive as long as internal data are used in the calibration process and then LDA is more closely related to the actual riskiness of each bank. However it is now widely recognized that calibration on internal data only does not suffice to provide accurate capital charge. In other words, internal data should be supplemented with external data. The goal of this paper is to address issues regarding the optimal way to mix internal and external data regarding frequency and severity.

As a matter of fact, frequency and severity data must be treated differently as they raise rather different issues. Considering one specific bank, its internal frequency data likely convey information on its specific riskiness and the soundness of its risk management practices. For example, a bank whose past frequencies of events are particularly good with respect to its exposition should be charged a lower capital requirement than for an average bank. Unfortunately nothing ensures that this lower-than-average frequency does result from an outstanding risk management policy rather than from a "lucky" business history. If lower-than-average past frequencies happened by chance, charging a lower-than-average capital charge would be misleading. Comparing internal and external data is then a way to separate what could be attributed to a sound risk management practice (which should be rewarded by lower-than-average capital charge) and what comes from a lucky course of business (which has no reason to last in the future). As the last QIS exercise has highlighted the strong heterogeneity of capital allocations among banks, it is all the more important to decide whether a better-than-average track record of past frequencies results from sound risk management practices or should be suspected to result from a lucky course of business. We here propose a rigorous way to tackle this issue through a statistical model refered as Credibility Theory in the insurance literature.

In the same spirit, internal data on severity should be mixed with external data but for different reasons. We assume here that internal severity data do not convey any information on internal risk management practices. In this paper internal severity databases should be supplemented with external severity data in order to give a non-zero likelihood to rare events which could be missing in internal databases. Unfortunately mixing internal and external data altogether may provide unacceptable results as external databases are strongly biased toward high-severity events. A rigorous statistical treatment is developped in the sequel to make internal and external data comparable and to make sure that merging both databases results in unbiased estimates of the severity distribution.

*Address: Crédit Lyonnais – GRO, Immeuble ZEUS, 4$^{\text{è}}$ étage, 90 quai de Bercy — 75613 Paris Cedex 12 — France; E-mail: antoine.frachot@creditlyonnais.fr

2 Mixing internal and external frequency data

Mixing internal and external frequency data can be relied on the credibility theory which is at the root of insurance theory. This theory tells how one should charge a fair premium to each policyholder considering the fact that policyholders do not have experienced the same history of claims. Achieving a fair pricing requires to compute how to relate a past history of claims either to a lucky (or unlucky) history or to an intrinsically lower (or greater) riskiness. According to the credibility theory, one must consider that policyholder riskiness is a priori unobserved but is partially revealed by the track record of the policyholder. Then the past information recorded on one specific policyholder is viewed as a means to reduce the a priori uncertainty attached to the policyholder riskiness. As a result, there is a discrepancy between the a priori probability distribution of riskiness (before any observation) and the a posteriori probability distribution obtained by conditioning with respect to past information. This discrepancy between the two distributions is the basic foundation for a fair pricing. This idea is developped in the context of the banking industry.

2.1 Computing the expected frequency of events

Considering one specific bank, let us focus on one business line and one type of events and note N_t the number of corresponding events at year t. Internal historical data (regarding frequency data) is represented by $\underline{N}_t$, that is the information set consisting of past number of events $\{N_t, N_{t-1}, \dots\}$. Now define EI the exposition indicator (i.e. gross income) and make the standard assumption that N_t is Poisson distributed with parameter $\lambda \times \mathrm{EI}$, where λ is a parameter measuring the (unobserved) riskiness of the bank. The key point is that supervisors should consider that the expected number of events for year t is $\mathbb{E}\left[N_t \mid \underline{N}_{t-1}\right]$ and not $\mathbb{E}\left[N_t\right]$. The latter term relates to the "pure premium" if we use insurance terminology while $\mathbb{E}\left[N_t \mid \underline{N}_{t-1}\right]$ is the expected number of events conditionally to the track record of the bank. As an example, let us assume that the bank has experienced lower-than-average N_s for $s < t$; this track record likely conveys information that this bank is intrinsically less risky than its competitors. As a result, $\mathbb{E}\left[N_t \mid \underline{N}_{t-1}\right]$ will be lower than an expectation which would ignore this information, that is $\mathbb{E}\left[N_t\right]$.

Regarding risk, banks differ from one another through parameter λ (and of course EI). Following credibility theory, we assume that λ is an unobservable random variable which ranges from low-risk banks (λ close to 0) to high-risk banks. In short, each bank has a specific unobserved riskiness λ and the internal data $\underline{N}_{t-1}$ conveys information on λ. Let us go one step further assuming that λ is distributed according to a Gamma law $\Gamma\left(a,b\right)$ among the banking industry[1]:

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

This means that all banks are not similar in terms of risk management and best practices but their respective riskinesses are unobserved. The choice of the Gamma distribution is rather arbitrary but has many appealing features. First, this class of distributions is large and flexible enough to capture the kind of heterogeneity one might expect. Secondly, this class of distributions permits easy and nice calculations and provides closed-form formulas. In particular, one can compute the unconditional expectation

$$\begin{aligned}
\pi_t^0 &= \mathbb{E}\left[N_t\right] \\
&= \mathbb{E}_\lambda\left[\mathbb{E}\left[N_t \mid \lambda\right]\right] \\
&= \mathrm{EI} \times \mathbb{E}\left[\lambda\right]
\end{aligned}$$

[1] $\Gamma\left(.\right)$ is the traditional Gamma function defined as:

$$\Gamma\left(x\right) = \int_0^{+\infty} u^{x-1} e^{-u} \, \mathrm{d}u$$

as well as conditional expectation

$$\begin{aligned}\pi_t &= \mathbb{E}\left[N_t \mid \underline{N}_{t-1}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[N_t \mid \lambda, \underline{N}_{t-1}\right] \mid \underline{N}_{t-1}\right] \\ &= \mathrm{EI} \times \mathbb{E}\left[\lambda \mid \underline{N}_{t-1}\right]\end{aligned}$$

As shown in Appendix A, computations can be made explicit:

$$\begin{aligned}\pi_t^0 &= a \times b \times \mathrm{EI} \\ \pi_t &= \omega \times \pi_t^0 + (1-\omega) \times \left[\frac{1}{t}\sum_{k=1}^{t} N_{t-k}\right]\end{aligned}$$

As a result, the expected number of events that relates best to the actual riskiness of this bank is a weighted linear combination between the unconditional expectation (which corresponds to the industry-wide expected number of events) and the historical average number of events experienced by the bank during its past course of business. ω is a parameter depending on the exposition indicator EI and the parameters a and b:

$$\omega = \frac{1}{1 + t \times b \times \mathrm{EI}}$$

As a consequence, the expected number of events to be considered is close to the average historical number of events experienced by the bank, if one of the two following conditions is satisfied:

- the length of track record is important, i.e. t is large;
- the exposure indicator EI is large.

The interpretation is straightforward: when a bank has a long history of frequencies of events and/or is highly exposed to operational risks, then supervisors should be confident enough to weigh significantly bank's internal data to assess its riskiness.

2.2 Computing the probability distribution of future frequency

In practice, the expected number of events conditionnally to past experience does not suffice to compute a capital charge. As a matter of fact the entire conditional distribution of N_t is required. This distribution results from a classical result (see Klugman, Panjer and Willmot [1998]):

$$\Pr\left\{N_t = n \mid \underline{N}_{t-1}\right\} = \frac{\Gamma(\tilde{a}+n)}{\Gamma(\tilde{a})\, n!}\left(1+\tilde{b}\right)^{-\tilde{a}}\left(\frac{\tilde{b}}{1+\tilde{b}}\right)^{n} \tag{1}$$

where

$$\tilde{a} = a + \sum_{k=1}^{t} N_{t-k}$$

and

$$\tilde{b} = \frac{b \times \mathrm{EI}}{1 + t \times b \times \mathrm{EI}}$$

It means that the probability that a bank whose track record amount to $\underline{N}_{t-1}$ experiences n losses at time t is given by equation (1). This distribution must be compared with the unconditional distribution (in other words the industry-wide frequency distribution):

$$\Pr\left\{N_t = n\right\} = \frac{\Gamma(a+n)}{\Gamma(a)\, n!}\left(1+b^0\right)^{-a}\left(\frac{b^0}{1+b^0}\right)^{n} \tag{2}$$

which gives the a priori probability that one bank (whose track record is unknown) experiences n losses at time t. Note that $b^0 = b \times \mathrm{EI}$.

As a result, the a priori and a posteriori distributions are mathematically similar except that the "internal" parameters $\tilde{a}$ and $\tilde{b}$ (caracterizing the actual riskiness of the bank) are adjusted from their industry-wide counterparts a and b by taking into account past experiences.

2.3 Calibration

Parameters a and b have to be estimated on frequency data experienced by the banking industry. Let us assume that a future QIS-type exercise provides a set of observations of frequency data $\{N_t^i\}$ where i denotes the i^{th} bank. Then, a maximum likelihood procedure can be easily achieved with the distribution of (unconditional) numbers of events given by (2):

$$(a, b) = \arg\max \sum_i \ln\Gamma\left(a + N_t^i\right) - \ln\Gamma\left(a\right) - \left(a + N_t^i\right)\ln\left(1 + b \times \mathrm{EI}^i\right) + N_t^i \ln b$$

As a summary the optimal combination of internal and external data when dealing with frequency data is as follows:

- estimate the overall riskiness of the banking industry as reflected by (industry-wide) parameters a and b;
- adjust parameters a and b to take into account internal past frequencies (i.e. $\tilde{a}$ and $\tilde{b}$);
- use equation (1) to obtain the probability distribution corresponding to a specific bank i whose past frequencies amount to $\underline{N}_{t-1}^i$. If bank i has no historical data on its own frequencies, use (1) with $\tilde{a} = a$ and $\tilde{b} = b^0$, that is equation (2).

Even though the maximum likelihood approach is the most efficient strategy, alternative calibration procedures could be adopted in the extent that they are simpler to implement. For example, let us assume that the average ratio N/EI (i.e. the number of events relative to the exposition) is known in an industry-wide basis as well as its dispersion: these two numbers would be sufficient to obtain some first estimates of parameters a and b by solving the two equations relating $\mathbb{E}\left[N/\mathrm{EI}\right]$ et $\mathrm{var}\left[N/\mathrm{EI}\right]$ with a and b.

3 Mixing internal and external severity data

Mixing internal and external severity data is an almost impossible task because no one knows which data generating process external severity data are drawn from. As a matter of fact, external severity data are biased toward high severity events as only large losses are publicly released. Merging internal and external data together gives spurious results which tend to be over-pessimistic regarding the actual severity distribution.

In other words, as the threshold above which external data are publicly released is unknown, the true generating process of external data is also unknown making the mixing process an impossible (and misleading) task. One can not conclude however that external databases are useless for our purpose. Indeed, it only means that the threshold should be added to the set of parameters one has to calibrate. Let us assume that the true loss probability distribution is denoted $\ell\left(x;\theta\right)$ where θ is a set of parameters defining the entire severity distribution. Then internal data follow this severity distribution while external data are drawn from the same distribution but truncated by a (unknown) threshold H. If ξ_j (respectively ξ_j^*) denotes an internal (resp. external) single loss record, then:

$$\begin{array}{rcl} \xi_j & \sim & \ell\left(\cdot;\theta\right) \\ \xi_j^\star & \sim & \ell_{|H}\left(\cdot;\theta\right) \end{array}$$

where $\ell_{|H}(.;\theta)$ is the loss density conditionally to these losses being above threshold H:

$$\ell_{|H}(x;\theta) = \frac{\ell(x;\theta)}{\int_H^{+\infty} \ell(v;\theta)\ \mathrm{d}v}$$

If one adopts a maximum likelihood procedure, the following program has to be solved:

$$(\theta, H) = \arg\max \sum_{j\in\mathcal{J}} \ln \ell\left(\xi_j;\theta\right) + \sum_{j\in\mathcal{J}^*} \ln \ell_{|H}\left(\xi_j^*;\theta\right) \tag{3}$$

which should provide a non-zero threshold H as soon as external data are biased toward high severity losses. Contrary to what consultancy firms often propose, merging internal and external data should be performed under the condition that both data types are made comparable in nature. Threshold H is exactly the parameter which ensures this comparability.

Developing program (3) shows that maximizing a conventional likelihood which ignores truncation is totally misleading. Ignoring truncation is equivalent to work with the following incorrect log-likelihood function:

$$\mathbf{L}(\theta, 0) = \sum_{j\in\mathcal{J}} \ln \ell\left(\xi_j;\theta\right) + \sum_{j\in\mathcal{J}^*} \ln \ell\left(\xi_j^*;\theta\right)$$

meaning that the threshold is incorrectly set to zero:

$$\mathbf{L}(\theta, H) = \mathbf{L}(\theta, 0) - n^* \ln \int_H^{+\infty} \ell(v;\theta)\ \mathrm{d}v$$

where n^* is the number of external losses — $n^* = \operatorname{card} \mathcal{J}^*$. In the same spirit it can be adapted to other methods of calibration like the General Method of Moments, Indirect Inference, etc. In all cases, one must consider that the generating process of external data is truncated above a threshold which has to be calibrated itself along with the parameters of the severity distribution.

4 Computing capital requirements

Bringing together the results of the previous sections, we are now able to compute the capital to be charged to a bank when both internal and external data (or benchmarks) are available.

Let us define the time-t total loss amount experienced by one specific bank:

$$Z_t = \sum_{j=1}^{N_t} \xi_j$$

where $\left\{\xi_j\right\}$ is the set of losses experienced at year t. As usual, we shall assume that ξ_j are independent and identically distributed, and independent from the number of events N_t. From regulatory purposes, the capital charge CaR should be rigorousely defined as:

$$\Pr\left\{Z_t \leq \mathrm{CaR} \mid \underline{N}_{t-1}\right\} \geq \alpha \tag{4}$$

where α is for example equal to 99.9%. This actual capital charge may be significantly different from the one we would compute while ignoring past internal data. The latter that we call $\mathrm{CaR}_{\mathrm{bmk}}$ (where bmk stands for benchmark) is thus defined as:

$$\Pr\left(Z_t \leq \mathrm{CaR}_{\mathrm{bmk}}\right) \geq \alpha \tag{5}$$

where the information set has been dropped. Consequently, CaR for a specific bank may be lower or greater than $\mathrm{CaR}_{\mathrm{bmk}}$ depending on whether the bank has experienced (significantly) lower-than-average or greater-than-average number of events. Finally, assuming that parameters a and b have been consistently estimated as well as the parameters of the severity distribution, capital charge for one bank whose exposition indicator is CaR and whose past experience amounts to $\underline{N}_{t-1}$ should be equal to the solution of:

$$\alpha = \sum_{n=0}^{+\infty} \Pr\left\{N_t = n \mid \underline{N}_{t-1}\right\} \times \mathbf{F}^{n\star}(\mathrm{CaR})$$

where $\mathbf{F}^{n\star}$ denotes the cumulative probability function of $\sum_{j=1}^{n} \xi_j$. In short, the capital charge of a specific bank whose exposition indicator is EI can be expressed as:

$$\mathrm{CaR} = \mathrm{CaR}_{\mathrm{bmk}}(\mathrm{EI}) \times W\left(\mathrm{EI}, \underline{N}_{t-1}\right)$$

where $W\left(\mathrm{EI}, \underline{N}_{t-1}\right)$ is a (non-explicit) function of the exposition indicator and past frequencies of events, with $W\left(\mathrm{EI}, \underline{N}_{t-1}\right) = 1$ when $\underline{N}_{t-1} = \emptyset$.

5 Concluding remarks

The Loss Distribution Approach has many appealing features since it is expected to be much more risk-sensitive than any other methods taken into consideration by the last proposals by the Basel Committee. Thus this approach is expected to provide significantly lower capital charges for banks whose track record is particularly good relatively to their exposures and compared with industry-wide benchmarks.

Unfortunately LDA when calibrated only on internal data is far from being satisfactory from a regulatory perspective as it could likely underestimate the necessary capital charge. This happens for two reasons. First if a bank has experienced a lower-than-average number of events, it will benefit from a lower-than-average capital charge even though its good track record happened by chance and does not result from better-than-average risk management practices. As a consequence, LDA is acceptable as long as internal frequency data are tempered by industry-wide references. As such, it immediately raises the issue of how to cope with both internal frequency data and external benchmarks. This paper proposes a solution based on credibility theory which is widely used in the insurance industry to tackle analogous problems. As a result, we show how to make the statistical adjustment to temper the information conveyed by internal frequency data with the use of external references.

Similarly if the calibration of severity parameters ignores external data, then the severity distribution will likely be biased towards low-severity losses since internal losses are typically lower than those recorded in industry-wide databases. Again from a regulatory perspective LDA cannot be accepted unless both internal and external data are merged and the merged database is used in the calibration process. Here again it raises the issue regarding the best way to merge these data. Obviously it cannot be done without any care since if internal databases are directly fuelled with external data, severity distributions will be strongly biased towards high-severity losses. This paper proposes also a statistical adjustment to make internal and external databases comparable with one another in order to permit a safe and unbiased merging.

References

[1] Basel Committee on Banking Supervision, Working Paper on the Regulatory Treatment of Operational Risk, september 2001

[2] Frachot, A., P. Georges and T. Roncalli [2001], Loss Distribution Approach for operational risk, Crédit Lyonnais, Groupe de Recherche Opérationnel, *Working Paper* (available from http://gro.creditlyonnais.fr)

[3] Klugman, S.A., H.H. Panjer and G.E. Willmot [1998], Loss Models: From Data to Decisions, *Wiley Series in Probability and Mathematical Statistics*, John Wiley & Sons, New York

A Computation of the conditionnal expected number of events

The expected number of events conditionnally to past information is equal to:

$$\begin{aligned}
\pi_t &= \mathbb{E}\left[N_t \mid \underline{N}_{t-1}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[N_t \mid \lambda, \underline{N}_{t-1}\right] \mid \underline{N}_{t-1}\right] \\
&= \mathbb{E}\left[\lambda \times \mathrm{EI} \mid \underline{N}_{t-1}\right]
\end{aligned}$$

Thus one has to compute the probability distribution of $\lambda \times \mathrm{EI}$ conditionnally to $\underline{N}_{t-1}$. This is a rather classical calculus one can find in Klugman, Panjer and Willmot [1998]: the conditionnal law is still a Gamma distribution but with different parameters. Denoting $f_{\lambda\times\mathrm{EI}}\left(\cdot \mid \underline{N}_{t-1}\right)$ the conditionnal probability density function (called also the posterior distribution):

$$f_{\lambda\times\mathrm{EI}}\left(\lambda \mid \underline{N}_{t-1}\right) = \frac{\lambda^{\tilde{a}-1}e^{-\lambda/\tilde{b}}}{\Gamma(\tilde{a})\tilde{b}^{\tilde{a}}}$$

with

$$\tilde{a} = a + \sum_{k=1}^{t} N_{t-k}$$

and

$$\tilde{b} = \frac{b \times \mathrm{EI}}{1 + t \times b \times \mathrm{EI}}$$

As a result, the conditionnal expected number of events is equal to:

$$\begin{aligned}
\pi_t &= \tilde{a} \times \tilde{b} \\
&= \omega \times \pi_t^0 + (1-\omega) \times \left[\frac{1}{t}\sum_{k=1}^{t} N_{t-k}\right]
\end{aligned}$$

with

$$\pi_t^0 = a \times b \times \mathrm{EI}$$

and

$$\omega = \frac{1}{1 + t \times b \times \mathrm{EI}}$$