DATA COMBINATION UNDER BASEL II AND SOLVENCY 2: OPERATIONAL RISK GOES BAYESIAN

Dominik D. LAMBRIGGER[1]
Pavel V. SHEVCHENKO[2]
Mario V. WÜTHRICH[1]

ABSTRACT

Under the new regulatory standards Basel II and Solvency 2, many financial institutions adopt a Loss Distribution Approach (LDA) to estimate the operational risk capital charge.

Such an approach requires the combination of internal and external data with expert opinion in an adequate manner. In this article we present a consistent and unified way how this task can be fulfilled. The simultaneous consideration of the three different sources of information is done in a Bayesian inference model.

The main idea is to start with external market data which determines a prior estimate. This prior estimate is then modified according to internal observations and expert opinion leading to a posterior estimate. Risk measures as for instance Value-at-Risk and Expected Shortfall may then easily be inferred from this posterior knowledge.

RÉSUMÉ

En tenant compte des nouveaux standards de surveillance Basel II et Solvency 2, beaucoup d'assurances et de banques adoptent une *Loss Distribution Approach* (LDA) pour estimer le capital de risque opérationnel.

Une telle approche exige une combinaison raisonnable des données internes, externes et l'opinion d'experts. Dans cet article, nous allons présenter une méthode cohérente et unifiée pour accomplir cette tâche. La considération simultanée des trois sources d'information est atteinte en utilisant un modèle d'inférence bayésienne.

L'idée principale est de commencer par des données externes qui déterminent une estimation a priori. Cette estimation sera à modifier selon les observations internes et selon

[1] ETH Zurich, Department of Mathematics, 8092 Zurich, Switzerland.
[2] CSIRO Mathematical and Information Sciences, Sydney, Locked Bag 17, North Ryde, NSW, 1670, Australia; e-mail: Pavel.Shevchenko@csiro.au.

l'opinion d'experts menant ainsi à une estimation a posteriori. Les mesures de risque, par exemple la *Value-at-Risk* et l'*Expected Shortfall*, peuvent ainsi facilement être inférées de cette connaissance a posteriori.

Keywords: Advanced Measurement Approach, Basel II, Bayesian Inference, Loss Distribution Approach, Operational Risk, Quantitative Risk Management, Solvency 2.

1. INTERNAL DATA, EXTERNAL DATA AND EXPERT OPINION

The quantification of an operational risk capital charge under Solvency 2 or Basel II [3] is for all financial institutions a challenging task. Typically, many cells of the Basel II operational risk matrix contain very few internal data. This implies that it is difficult to find reliable risk estimates based on these observations solely. Therefore, there is a strong need for incorporating other sources of information such as expert opinion and relevant external data in order to achieve an adequate picture about the high severity, low frequency operational risk landscape.

The Basel Committee, for example, mentions this concern explicitly; see for instance BIS [3], paragraph 675: *"A bank must use scenario analysis of expert opinion in conjunction with external data to evaluate its exposure to high-severity events. This approach draws on the knowledge of experienced business managers and risk management experts to derive reasoned assessments of plausible severe losses. For instance, these expert assessments could be expressed as parameters of an assumed statistical loss distribution."*

In industry practice ad-hoc methods are used for the combination of the different sources of information. Often this does not lead to satisfactory results and to consistent answers. Therefore, it is still an open issue to combine internal data with external data and expert opinion. This has also been emphasized by different leading risk managers; see, e.g., Davis [4], an interview with four industry's top risk executives in September 2006: "*[A] big challenge for us is how to mix the internal data with external data; this is something that is still a big problem because I don't think anybody has a solution for that at the moment."*

In our opinion classical actuarial concepts give answers to these questions. Recently, a Bayesian inference model focused on operational risk losses has been developed in Lambrigger et al. [5]. That model allows for the combination of the mentioned three sources of risk information simultaneously. The goal of this paper is to review that method in an easily accessible overview.

First we introduce the methodology to combine the three types of knowledge in the context of operational risk. Then this framework is used to quantify the severity distribution of operational risk losses. Finally, an example illustrates the robustness of this quantitative approach.

2. THE BAYESIAN APPROACH

The Basel Committee has established an operational risk matrix consisting of 8×7 risk cells (8 business lines, 7 risk types). In each of these 56 risk cells financial institutions model the corresponding loss frequency (e.g., by a Poisson distribution) and loss severity distribution (e.g., by a lognormal or Pareto distribution). Hereafter we concentrate only on one single risk cell and for the moment refrain from modeling the dependence structure between business lines and risk types.

After an appropriate choice of the frequency and severity distribution, the risk manager is required to estimate the parameters of these distributions in an appropriate way. We denote this (unknown) parameter vector for the insurance company under consideration by z, which is also referred to as the company's *risk profile*. The parameter vector $\boldsymbol{z}$ could for instance correspond to the location, shape or/and scale parameter of the severity distribution function. The company's true (but unknown) risk profile $\boldsymbol{z}$ needs to be estimated from the available (internal) information. If few internal data is available, a precise and robust estimation of $\boldsymbol{z}$ becomes difficult. Therefore, the estimate needs to include other sources of information (external data and expert opinion). This is a well-known problem in actuarial practice where, for example, certain lines of business have only a small volume or only few observations.

In a Bayesian context the unknown risk profile $\boldsymbol{z}$ is treated as a realization of a random vector $\boldsymbol{Z}$ illustrating that we do not have perfect knowledge about the true underlying parameters. In our setup the distribution of $\boldsymbol{Z}$ stems from market information (external data). That is, every company's risk profile can be viewed as a realization of the market's risk profile. $\boldsymbol{Z}$ is therefore a random vector with *known* distribution. It models (after a possible scaling) the risk profiles over the whole financial industry.

Before having any company specific information (internal data, expert opinion), we completely rely on the available industry data. The best prediction of our company specific risk profile $\boldsymbol{z}$ would hence be based on the belief in this external knowledge only, represented by the random vector $\boldsymbol{Z}$. The distribution of $\boldsymbol{Z}$ is called prior risk profile or prior distribution. The parameters of the prior distribution (so-called hyper-parameters) are estimated using industry (external) data or, if no industry data is available, defined by some "super expert" (e.g., regulator) that has an overview over the whole financial industry.

$\boldsymbol{Z}$	$\xrightarrow{\boldsymbol{X},\boldsymbol{\vartheta}}$	$\boldsymbol{z}$
• parameter representing the whole industry		• company specific parameter
• considers external (market) data only		• considers internal data $\boldsymbol{X}$ and expert opinion $\boldsymbol{\vartheta}$
• random variable		• realization of $\boldsymbol{Z}$, hence deterministic
• with known distribution		• unknown, estimated by $\mathbb{E}[\boldsymbol{Z} \mid \boldsymbol{X},\boldsymbol{\vartheta}]$

Table 1. Internal data $\boldsymbol{X}$ and expert opinion $\boldsymbol{\vartheta}$ transform the (prior) risk profile of the whole industry $\boldsymbol{Z}$ into an individual company specific (posterior) risk profile $\boldsymbol{z}$.

A priori, before assessing any expert opinion (for instance inferred from scenario analysis) and observing any internal data, all companies have the same prior risk profile $\boldsymbol{Z}$ stemming from market information only. As time passes, we gather internal experiences like internal operational risk events $\boldsymbol{X} = (X_1,\ldots,X_K)$ and expert opinion $\boldsymbol{\vartheta} = (\vartheta^{(1)},\ldots,\vartheta^{(M)})$. This additional information certainly influences our belief into the prior distribution of $\boldsymbol{Z}$, inferred from market data only, and therefore the prediction of the company specific parameter vector z is adjusted according to our internal observations; see Table 1.

The more internal information $\boldsymbol{X}$ and $\boldsymbol{\vartheta}$ we have, the better we are able to predict our company specific risk profile $\boldsymbol{z}$ and the less credibility we give to the market information. That is, the internal data $\boldsymbol{X}$ and the expert opinion $\boldsymbol{\vartheta}$ transform the risk profile of the overall market $\boldsymbol{Z}$ into a conditional distribution of $\boldsymbol{Z}$ given $\boldsymbol{X}$ and $\boldsymbol{\vartheta}$, formally denoted by $\boldsymbol{Z} \mid \boldsymbol{X},\boldsymbol{\vartheta}$. The natural question we have to answer is: How does this company specific information $\boldsymbol{X}$ and $\boldsymbol{\vartheta}$ change our view of the underlying parameter $\boldsymbol{Z}$, i.e., what is the explicit distribution of $\boldsymbol{Z} \mid \boldsymbol{X},\boldsymbol{\vartheta}$?

Formally, this is described as follows. We denote the *prior* parameter density of $\boldsymbol{Z}$ by $\pi_{\boldsymbol{Z}}(\boldsymbol{z})$. Given our risk profile $\boldsymbol{Z} = \boldsymbol{z}$, the observations X_k have density $f_1(X_k \mid \boldsymbol{Z} = \boldsymbol{z})$ and the expert opinions $\vartheta^{(m)}$ have density $f_2(\vartheta^{(m)} \mid \boldsymbol{Z} = \boldsymbol{z})$. Under suitable independence assumptions, the *posterior* density of $\boldsymbol{Z} \mid \boldsymbol{X},\boldsymbol{\vartheta}$ can be calculated explicitly. Bayes' Theorem gives for the posterior density of $\boldsymbol{Z} \mid \boldsymbol{X}, \boldsymbol{\vartheta}$

$$\hat{\pi}_{\boldsymbol{Z}|\boldsymbol{X},\boldsymbol{\vartheta}}(\boldsymbol{z}) = c\, \pi_{\boldsymbol{Z}}(\boldsymbol{z}) \prod_{k=1}^{K} f_1(X_k \mid \boldsymbol{Z} = \boldsymbol{z}) \prod_{m=1}^{M} f_2(\vartheta^{(m)} \mid \boldsymbol{Z} = \boldsymbol{z}), \tag{2.1}$$

where c is the normalizing constant not depending on $\boldsymbol{z}$. Given the distribution of $\boldsymbol{Z} \mid \boldsymbol{X}, \boldsymbol{\vartheta}$, the company specific parameter $\boldsymbol{z}$ can then be estimated, e.g., by the posterior mean $\mathbb{E}[\boldsymbol{Z} \mid \boldsymbol{X}, \boldsymbol{\vartheta}] = \int \boldsymbol{z}\, \hat{\pi}_{\boldsymbol{Z}|\boldsymbol{X},\boldsymbol{\vartheta}}(\boldsymbol{z})\, d\boldsymbol{z}$.

3. A SIMPLE MODEL

As an example, we consider a so-called lognormal-normal-normal model for loss severities. We assume that operational losses of an individual company have a lognormal distribution $\mathrm{LN}(\Delta, \sigma_{\text{int}})$ with scale parameter σ_{int} and location parameter Δ. Δ plays the role of the unknown risk profile Z of our company with given prior distribution. Moreover, we assume that the expert opinion about the parameter Δ follows a normal distribution $\mathcal{N}(\Delta, \sigma_{\text{exp}})$. As soon as internal operational losses and expert opinion are available, we calculate the posterior of Δ under this additional information.

For illustrative purposes, we only consider the lognormal-normal-normal loss severity model in this article. Note that these ideas can easily be translated to other situations, such as the Poisson frequency model or the Pareto loss severity model. For a detailed outline of other distributional model assumptions we refer to Lambrigger et al. [5].

3.1 Model assumptions (Lognormal-normal-normal)

Let us assume the following loss severity model:

- **Market Profile:** Let $\Delta \sim \mathcal{N}(\mu_{\text{ext}}, \sigma_{\text{ext}})$ be a normally distributed random variable with parameters $\mu_{\text{ext}}, \sigma_{\text{ext}}$, which are estimated from (external) market data, i.e., $\pi_{\boldsymbol{Z}}(\boldsymbol{z})$ in (1) follows the density of $\mathcal{N}(\mu_{\text{ext}}, \sigma_{\text{ext}})$.
- **Internal Data:** The losses $k = 1, \ldots, K$ from the concerning institution are assumed to be conditionally (on Δ) i.i.d. lognormally distributed: $X_1, \ldots, X_K \mid \Delta \overset{\text{i.i.d.}}{\sim} \mathrm{LN}(\Delta, \sigma_{\text{int}})$, where σ_{int} is assumed to be known. That is, $f_1(\cdot \mid \Delta)$ in (1) corresponds to the density of a $\mathrm{LN}(\Delta, \sigma_{\text{int}})$ distribution.
- **Expert Opinion:** We assume that the company has M experts with opinion $\vartheta^{(m)}$, $1 \le m \le M$, about the parameter Δ with $\vartheta^{(1)}, \ldots \vartheta^{(M)} \mid \Delta \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\Delta, \sigma_{\text{exp}})$, where σ_{exp} denotes the expert uncertainty. That is, $f_2(\cdot \mid \Delta)$ corresponds to the density of a $\mathcal{N}(\Delta, \sigma_{\text{exp}})$ distribution.

Furthermore, we assume that expert opinion $\boldsymbol{\vartheta}$ and internal data $\boldsymbol{X}$ are conditionally independent given a realization of the risk profile Δ; that is, the distributional representation (1) holds. In order to get an accurate estimate for the true company specific risk profile, the market profile Δ is adjusted to the individual companies by internal data and expert opinion. Note that such a procedure is closely related to hierarchical credibility models discussed for instance in Bühlmann and Gisler [2].

The hyper-parameters μ_{ext} and σ_{ext} for the market profile distribution are estimated from external data, e.g., by maximum likelihood or by the method of moments. For $M \geq 2$, the parameter σ_{exp} is, e.g., estimated by the sample standard deviation of $\vartheta^{(m)}$:

$$\sigma_{\text{exp}} = \left(\frac{1}{M-1} \sum_{m=1}^{M} (\vartheta^{(m)} - \overline{\vartheta})^2 \right)^{1/2},$$

with averaged expert opinion $\overline{\vartheta} = \frac{1}{M} \sum_{m=1}^{M} \vartheta^{(m)}$, or can be defined externally by the regulator.

When the assumption of independent expert opinion is too restrictive, one could think of one panel of dependent experts or a pooling of scenarios. Consequently we would set $M = 1$ and the corresponding standard deviation σ_{exp} has to be calculated taking into account the dependence structure between experts. The assumption that expert opinion and internal data are conditionally independent given the risk profile Δ may be disputable if the experts are unable to specify their opinion regardless of the internal data observed. It is hence crucial that expert opinion is based on scenario analysis and stress tests independently from internal events.

4. CREDIBILITY WEIGHTED AVERAGE

Under Model Assumption 3 the posterior distribution can be calculated analytically. We have the following theorem.

Theorem 4.1: credibility weighted average

Under Model Assumptions 3 and with the notation $\overline{\log X} = \frac{1}{K} \sum_{k=1}^{K} \log X_k$, the posterior distribution $\Delta \,|\, \boldsymbol{X}, \boldsymbol{\vartheta}$ is a normal distribution $N(\widehat{\mu}, \widehat{\sigma})$ with parameters

$$\widehat{\sigma}^2 = \left(\frac{1}{\sigma_{\text{ext}}^2} + \frac{K}{\sigma_{\text{int}}^2} + \frac{M}{\sigma_{\text{exp}}^2} \right)^{-1}, \tag{4.1}$$

and

$$\widehat{\mu} = \mathbb{E}[\Delta \,|\, \boldsymbol{X}, \boldsymbol{\vartheta}] = \omega_1 \mu_{\text{ext}} + \omega_2 \overline{\log X} + \omega_3 \overline{\vartheta}, \tag{4.2}$$

where the so-called credibility weights are given by $\omega_1 = \hat{\sigma}^2 / \sigma_{\text{ext}}^2$, $\omega_2 = \hat{\sigma}^2 K / \sigma_{\text{int}}^2$ and $\omega_3 = \hat{\sigma}^2 M / \sigma_{\text{exp}}^2$. A proof is given in Lambrigger et al. [5].

Theorem 4.1 gives a consistent and unified way to combine the different sources of information. It shows how internal observations, relevant external data and expert opinion are weighted using credibility weights ω_1, ω_2 and ω_3. These are numbers between 0 and 1 that sum up to 1. Note that these credibility weights are provided by the model in a natural way, that is, there is no ad-hoc choice of the credibility weights. The less credible the information of one of the three data sources, the smaller the corresponding credibility weight ω_i, $i \in \{1,2,3\}$, in (4.2). If one information source is highly inaccurate (e.g., $\sigma_{\text{ext}}, \sigma_{\text{int}}$ or $\sigma_{\text{exp}} \to \infty$), then the corresponding credibility weight will be close to 0. If however, the information about one data source is very precise (e.g., many observations, small variance of expert opinion or small variation in the parameters of external data), then the corresponding ω_i, $i \in \{1,2,3\}$, will be close to 1.

Note that Theorem 4.1 does not only provide us with the company's expected risk profile $\hat{\mu}$, but with the whole distribution $\Delta \mid X, \vartheta \sim N(\hat{\mu}, \hat{\sigma})$. Therefore, the parameter uncertainty and the model risk can be quantified.

Example 4.2

Assume that a financial company models its risk severities according to Model Assumptions 3 with scale parameter $\sigma_{\text{int}} = 4$ and the regulator provides external prior data with hyper-parameters $\mu_{\text{ext}} = 2$ and $\sigma_{\text{ext}} = 1$. Moreover, the company's internal expert opinion is $\overline{\vartheta} = 6$ with standard deviation $\sigma_{\text{exp}} = 3/2$ and we observe the internal operational risk losses (sampled from a $\text{LN}(\mu_{\text{int}} = 4, \sigma_{\text{int}} = 4)$ distribution) given in Figure 1. Note that the company under consideration does worse ($\mu_{\text{int}} = 4$) than the industry average ($\mu_{\text{ext}} = 2$). However, the company's experts even have a worse opinion about their own institution ($\overline{\vartheta} = 6$).

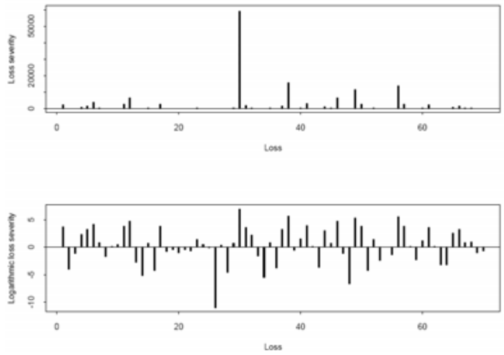

Figure 1. 70 loss severities (upper panel) and their corresponding logarithmic values centered around 0 (lower panel), sampled from a $LN(\mu_{\text{int}} = 4, \sigma_{\text{int}} = 4)$ *distribution.*

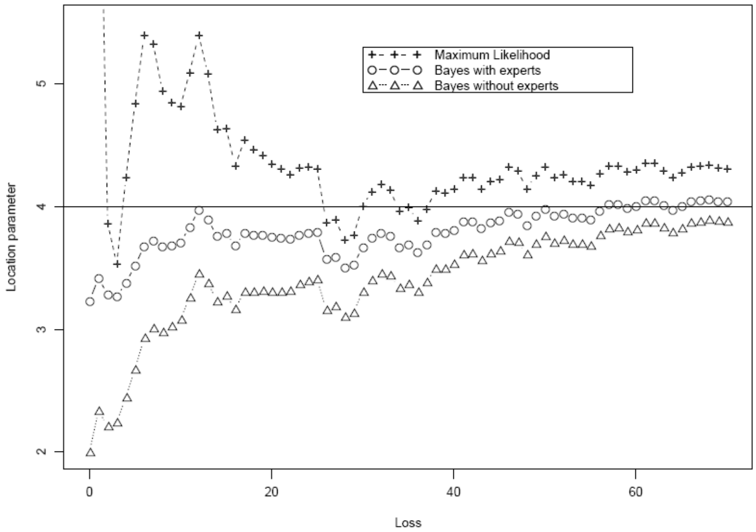

Figure 2. The Bayes estimator $\widehat{\mu}_k$ *including expert opinion (* $\circ$ *) is compared to the Bayes estimator* $\widehat{\mu}_k^{\text{SW}}$ *without expert opinion (* $\vartriangle$ *) and to the maximum likelihood estimator* $\widehat{\mu}_k^{\text{MLE}}$ *(* $+$ *). The straight line stands for the true company specific parameter* $\mu_{\text{int}} = 4$ *.*

In Figure 2 we compare the classical maximum likelihood estimator (corresponding to (4.2) with $M = 0$ and $\sigma_{\text{ext}} \to \infty$)

$$\widehat{\mu}_k^{\text{MLE}} = \frac{1}{k} \sum_{i=1}^{k} \log X_i, \quad 1 \le k \le K,$$

to the estimator proposed in Shevchenko and Wüthrich [6] without expert opinion (corresponding to (4.2) with $M = 0$)

$$\widehat{\mu}_k^{\mathrm{SW}} = \mathbb{E}[\Delta \mid X_1, \ldots, X_k], \quad 1 \le k \le K,$$

and to the Bayes estimator given by (4.2)

$$\widehat{\mu}_k = \mathbb{E}[\Delta \mid X_1, \ldots, X_k, \overline{\vartheta}] = \omega_1 \mu_{\mathrm{ext}} + \omega_2 \widehat{\mu}_k^{\mathrm{MLE}} + \omega_3 \overline{\vartheta}, \quad 1 \le k \le K.$$

Figure 2 shows the high volatility of the maximum likelihood estimator, for small numbers of observations k. It is very sensitive to newly arriving losses. However, the Bayes estimator discussed in this paper shows a much more stable behavior around the true value $\mu_{\mathrm{int}} = 4$, also when few data are available. This is due to the smaller variance of the Bayes estimator; see equation (4.1). Moreover, it performs better than the estimator $\widehat{\mu}_k^{\mathrm{SW}}$ due to the fact that the expert opinion has an additional smoothening effect.

In this example we see that relevant external data and well-specified expert opinion stabilize and smoothen the estimator, even when the input data (as for example the expert opinion) over- or underestimates the true company specific value. In that sense, Bayesian inference yields a suitable framework to combine several different data sources. For more numerical examples we refer to Lambrigger et al. [5].

5. CONCLUSION

To meet the operational risk regulatory requirements, one needs to incorporate internal data, relevant external data and expert opinion. We present a Bayesian framework that leads to a natural credibility weighted combination of the different sources of information.

To achieve this, we start with a general risk profile representing the whole financial industry (prior distribution) and then gradually incorporate the internal information based on loss data and expert opinion (yielding the posterior distribution). This is done by the specification of the underlying distribution and then by applying Bayes' theorem. A criticism often voiced against Bayesian statistics is that the choice of the prior distribution is somewhat arbitrary. In the present approach however, the choice of a prior distribution is based on statistically meaningful external market data (pure empirical Bayes approach).

The novelty of our approach in contrast to classical Bayesian inference is that we combine simultaneously three different risk information sources instead of only two. Under our model assumptions the posterior distribution is then calculated in an analytically closed form; see Theorem 4.1. There are various other examples that lead to closed analytical posterior distributions. If the posterior distribution can not be obtained in a closed form then one either applies linear credibility models (see, e.g., Bühlmann and Gisler [2]) or numerical methods like Markov chain Monte Carlo (MCMC) methods (see, e.g., Asmussen and Glynn [1]). MCMC methods have the advantage (over linear credibility methods) that they give information over the whole posterior distribution which allows for the calculation of any risk measure.

For a single risk cell of an individual company, risk measures as for instance VaR have to be inferred. One feature of our approach is that the parameter uncertainty and the model risk can be quantified, because the entire distribution function (and not only the expected value) of the parameters is known.

ACKNOWLEDGMENTS

The authors would like to thank Paul Embrechts for his valuable comments on a first draft of the paper.

REFERENCES

[1] ASMUSSEN, S. AND GLYNN, P. W. (2007) *Stochastic Simulation: Algorithms and Analysis*. Springer, Berlin.

[2] BÜHLMANN, H. AND GISLER, A. (2005) *A Course in Credibility Theory and its Applications*. Springer, Berlin.

[3] BIS (2005) Basel II: International Convergence of Capital Measurement and Capital Standards: a revised framework. Bank for International Settlements (BIS), www.bis.org .

[4] DAVIS, E. (2006) Theory vs Reality. *OpRisk and Compliance*. 1 September 2006, www.opriskandcompliance.com/public/showPage.html?page=345305 .

[5] LAMBRIGGER, D. D., SHEVCHENKO, P. V. AND WÜTHRICH, M. V. (2007) The quantification of operational risk using internal data, relevant external data and expert opinion. *Journal of Operational Risk* **2**(3), 3–27.

[6] SHEVCHENKO, P. V. AND WÜTHRICH, M. V. (2006) The structural modeling of operational risk via Bayesian inference: combining loss data with expert opinions. *Journal of Operational Risk* **1**(3), 3–26.