Modelling macroeconomic effects and expert judgements in operational risk : a Bayesian approach

Holger Capa Santos, Marie Kratz, Franklin Mosquera Munoz

Modelling macroeconomic effects and expert judgements in operational risk : a Bayesian approach

H. Capa Santos
M. Kratz
F.V. Mosquera Munoz

ESSEC
BUSINESS SCHOOL

Modelling macroeconomic effects and expert judgements in operational risk : a Bayesian approach

H. Capa Santos*,   M. Kratz [†]   and   F. V. Mosquera Muñoz [‡]

Abstract

This work presents a contribution on operational risk under a general Bayesian context incorporating information on market risk profile, experts and operational losses, taking into account the general macroeconomic environment as well. It aims at estimating a characteristic parameter of the distributions of the sources, market risk profile, experts and operational losses, chosen here at a location parameter. It generalizes under more realistic conditions a study realized by Lambrigger, Shevchenko and Wüthrich, and analyses macroeconomic effects on operational risk. It appears that severities of operational losses are more related to the macroeconomics environment than usually assumed.

---

*Prof. Holger Capa Santos, Escuela Politécnica Nacional (EPN), Ladrón de Guevara E11-253, Quito, Ecuador; Email: holger.capa@epn.edu.ec

[†]Prof. Marie Kratz, ESSEC Business School Paris, avenue Bernard Hirsch BP 50105, 95021 Cergy-Pontoise Cedex, France; Email: kratz@essec.edu.

[‡]Franklin Mosquera is student in engineering at EPN, Quito, Ecuador.

[0]Marie Kratz is also member of MAP5, UMR8145, Univ. Paris Descartes, France.

Introduction

The quantification of an operational risk capital charge under the regulatory standards Basel II and Solvency 2 is a challenging task; many financial institutions adopt a Loss Distribution Approach to estimate this risk capital charge (see eg, [Klugman *et al.*(2008)], [McNeil *et al.*(2005)]). Such approach requires combination of internal and external data, as well as expert opinions; indeed, this third component appeared recently as crucial. An interesting Bayesian scenario to estimate the parameters of the risk frequency and severity distributions has been proposed by Shevchenko and Wüthrich (see [Shevchenko *et al.*(2006)]) based on internal and external data, then by Lambrigger et al. (see [Lambrigger *et al.*(2007)], or [Lambrigger *et al.*(2008)] for an overwiew) when adding the expert opinion as a third component. They assumed an *a priori* market risk profile and various conditions such as the conditional independence given the market risk profile between the expert opinions and the internal observations. Here one aims to investigate other scenarios under more realistic conditions and taking into account the general macroeconomic environment. Indeed, if we refer to what recently happened at UBS or at Société Générale (with the Kerviel affaire), operational risks of this type might become rather serious during a financial crisis, hence showing some economic dependence. Therefore, using a more general Bayesian approach, we add in the prior modeling a hyper level representing the general macroeconomic environment. Moreover, this hierarchical extension helps increasing the robustness of the prior distribution (see [Robert(2007)] p. 143) and allows relating the *a priori* to any of the three main sources mentioned previously. Recall also that one of the reasons to adopt the Bayesian point of view, which is intrinsically conditional on a fixed number of observations, is to obtain estimators having good properties even with a fixed and maybe small sample size, and not only under asymptotic criteria (see [Robert(2007)], p. 48). It is rather important when considering operational losses, in (re)insurance companies or in banks. Indeed, the number of operations being limited in insurance companies (and even more in reinsurance companies), the number of operational losses is restricted, even if these losses are high on average. In banks, operations are quite numerous, which leads to a large number of small operational losses, but high operational losses are a few. Hence the need to have reliable estimators whatever the size of the operational loss sample is.

We will introduce our general Bayesian approach, with a prior variable U representing the general macroeconomic environment. Having a few information on this environment, we decided to chose it as an noninformative prior on the parameter. Nevertheless, the method would work as well with a specific a priori distribution as can be seen in the next section; it might be relevant when working in banking context, during a financial crisis.
We will illustrate our general approach considering first when the *a priori* is related to the market risk profile. Then we will consider the point of view of the company and relate the *a priori* directly to the operational severities.
Note that we will consider the same type of distributions as in [Lambrigger *et al.*(2007)]

to enable fair comparisons. Nevertheless, our study will present one example only among the various examples developed in [Lambrigger *et al.*(2007)], working with normal and lognormal distributions and choosing the parameter characterizing the distribution of the three sources as a location parameter, the main objective being to show how to generalize those previous works cited above and not to constitute a catalogue of examples. We will rather complete the study, looking at properties of the location estimator.

1 An *a priori* related to the market risk profile

Let us present our model.
The main actors/sources of risk information of the model are respectively:
- the operational losses (internal data) $X = (X_i; i = 1, \cdots, k)$ of the concerning institution;
- the market risk profile Z;
- the expert opinion $Y = (Y_j; j = 1, \cdots, n)$;
- an unknown parameter U that we introduce to take into account the general macroeconomic environment.
This parameter might characterize the location or the shape parameter of the distributions of the three sources X, Y and Z, with a prior distribution denoted by π_U. Here we will illustrate our method by choosing for instance U as the location parameter. Its noninformative prior distribution Π_U is assumed to be uniform, hence a so called improper distribution but 'that should be preferred to vague proper priors such as a normal distribution with very large variance; it would give indeed a false sense of safety owing to properness, while lacking robustness in terms of influence on the resulting inference' (see [Berger(2000)]).
Some information on the risk profile Z (that represents the location parameter) is provided as a benchmark for banks by the Basel committee, and as an index of the insurance market in the Solvency II standard formula by the EIOPA. We assume that the conditional distribution $\mathcal{L}(Z|U)$ given U is known, with conditional expectation U: $\mathbb{E}[Z \mid U] = U$. Note that it implies $\mathbb{E}[Z] = \mathbb{E}[U]$.
The operational losses X of the given company are described by internal data, and depend, as the operational risk profile $h(X)$, on the parameter U but of course also on the market risk profile Z. This dependence can be translated through the location parameter of the conditional distribution $\mathcal{L}(X|U, Z)$ of X given Z and U in a simple way, when taking for instance a linear approximation of Z and U, namely $\alpha Z + (1 - \alpha)U$, $\alpha \in [0, 1[$. The operational risk profile $h(X)$ might be chosen as an unbiased estimator of the location parameter of the conditional distribution of X given Z and U, and satisfies $\mathbb{E}[h(X)] = \mathbb{E}(Z) = \mathbb{E}(U)$. The conditional distribution of X given Z and U is also supposed to be known.
Finally, we consider the expert opinion Y based on the knowledge of the market index Z but also, to be realistic, on the internal data X of the company; it is also influ-

enced by the macroeconomic conditions, hence depends on the unknown parameter U as well. We suppose that Y has a known conditional distribution $\mathcal{L}(Y|X, Z, U)$ given X, Z and U, with conditional expectation U. Using the same linear approximation to take this dependence into account, we ask for the conditional expectation of Y to

satisfy $\mathbb{E}[Y \mid X, Z, U] = \beta_1 h(X) + \beta_2 Z + \beta_3 U$, with $\beta_3 \neq 0, \beta_i \in [0, 1] : \sum_{i=1}^{3} \beta_i = 1$.

Since in general internal data are scarce, the aim is to use U, X, Z and Y to improve an estimation of the location parameter.

Note that we introduced on purpose undefined parameters α and β_i (with $\alpha \neq 1$ and $\beta_3 \neq 0$, otherwise there is no point of introducing the r.v. U), so that their choice is made by the user according to the information he has. Nevertheless, the behavior of the estimator does not modify strongly when changing those parameters.

Suppose that the conditional operational losses $(X_i|Z, U)_{1 \leq i \leq k}$ are i.i.d., the experts opinion $(Y_i|X, Z, U)_{1 \leq i \leq n}$ are i.i.d., with parent variables denoted by $X|Z, U$ and $Y|X, Z, U$ respectively.

We will then consider as an example the following conditional distributions:

$$Z|U \stackrel{d}{\sim} \mathcal{N}(U, \sigma_Z^2) \qquad Y|X, Z, U \stackrel{d}{\sim} \mathcal{N}(\beta_1 \overline{\log X} + \beta_2 Z + \beta_3 U, \sigma_Y^2)$$

and $X|Z, U \stackrel{d}{\sim} LN(\alpha Z + (1 - \alpha)U, \sigma_X^2)$ with density function

$$f_{X|Z,U}(x|z, u) = \frac{1}{x \sigma_X \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_X^2}\big[\log x - (\alpha z + (1 - \alpha)u)\big]^2\right\} \text{ for } x > 0$$

where the variances σ_X^2, σ_Y^2 and σ_Z^2 are supposed to be known (or estimated on data when possible), $\sum_{i=1}^{3} \beta_i = 1$ and $\overline{\log X} = \frac{1}{n} \sum_{i=1}^{k} \log X_i$.

Note that $\mathbb{E}[X|Z, U] = \exp\{\alpha Z + (1 - \alpha)U + \sigma_X^2/2\}$ and $\mathbb{E}[\log X] = \mathbb{E}(U)$.

1.1 The posterior distribution and Bayes estimator of the location parameter

THEOREM 1.1 *The posterior distribution $\widehat{\Pi}_{U|X,Y,Z}$ is a Gaussian distribution $\mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$ where the parameters satisfy*

$$\hat{\sigma}^2 = \left(\frac{1}{\sigma_Z^2} + \frac{k(1-\alpha)^2}{\sigma_X^2} + \frac{n\beta_3^2}{\sigma_Y^2}\right)^{-1} \tag{1}$$

and

$$\hat{\mu} = \omega_X \overline{\log X} + \omega_Y \overline{Y} + \omega_Z Z \tag{2}$$

with credibility weights $\omega_\bullet$ given by

$$\omega_X = \hat{\sigma}^2\left(\frac{k(1-\alpha)}{\sigma_X^2} - \frac{\beta_1\beta_3 n}{\sigma_Y^2}\right) \ ; \ \omega_Y = \frac{\hat{\sigma}^2 \beta_3 n}{\sigma_Y^2} \ ; \ \omega_Z = \hat{\sigma}^2\left(\frac{1}{\sigma_Z^2} - \frac{k\alpha(1-\alpha)}{\sigma_X^2} - \frac{\beta_2\beta_3 n}{\sigma_Y^2}\right).$$

4

The weights depend on various parameters which may help to get closer to μ even for small k. Their sum is one.

Although data are often not so numerous, for completeness, note that as $k \to \infty$, we have

$$\omega_X \sim \frac{1}{1-\alpha} \quad ; \quad \omega_Y \sim 0 \quad ; \quad \omega_Z \sim -\frac{\alpha}{1-\alpha}$$

The variance $\hat{\sigma}^2$ is an increasing function of α and a decreasing function of β_3 and of the sample sizes k and n; it is then minimum, for fixed sample sizes, for $\alpha = 0$ and $\beta_3 = 1$, which corresponds to conditional independence of the three actors X, Y, Z given U.

Notice that $\hat{\mu}$ and $\hat{\sigma}^2$ do not depend on the parameter of the prior distribution Π_U chosen as uniform (noninformative).

If we would have chosen U as a normal r.v. $\mathcal{N}(m, \sigma_U^2)$, then we would have obtained that $\widehat{\Pi}_{U|X,Y,Z}$ is a Gaussian distribution $\mathcal{N}(\hat{\mu}_*, \hat{\sigma}_*^2)$ with

$$\hat{\mu}_* = \frac{\hat{\sigma}^2}{\sigma_U^2}m + \hat{\mu} \quad \text{and} \quad \hat{\sigma}_*^2 = \left(\frac{1}{\sigma_U^2} + \frac{1}{\sigma_Z^2} + \frac{k(1-\alpha)^2}{\sigma_X^2} + \frac{n\beta_3^2}{\sigma_Y^2} \right)^{-1}$$

This choice, or the choice of another specific a priori distribution, might be of relevance eg, when working in banking context during a financial crisis.

When considering a casi-noninformative prior distribution, U could be taken as a normal r.v. $\mathcal{N}(m, \sigma_U^2)$ with a very large variance, which would lead to

$$\hat{\mu}_* = \frac{\hat{\sigma}^2}{\sigma_U^2}m + \hat{\mu} \underset{\sigma_U^2 \to \infty}{\sim} \hat{\mu} \quad \text{and} \quad \hat{\sigma}_*^2 = \left(\frac{1}{\sigma_U^2} + \frac{1}{\sigma_Z^2} + \frac{k(1-\alpha)^2}{\sigma_X^2} + \frac{n\beta_3^2}{\sigma_Y^2} \right)^{-1} \underset{\sigma_U^2 \to \infty}{\sim} \hat{\sigma}^2$$

ie, with a close behavior to the one described in Theorem 1.1.

Proof of Theorem 1.1.
The proof relies on an application of the Bayes theorem that we recall for completeness. If $f_{A|B}$ denotes the conditional probability density function of the r.v. A given the r.v. B, the Bayes theorem provides the posterior density function $\widehat{\Pi}_{U|X,Y,Z}$ of U given X, Y, Z as

$$\widehat{\Pi}_{U|X,Y,Z}(u) = c\pi_U(u)f_{Z|U}(z|u)f_{X|Z,U}(x|z,u)f_{Y|X,Z,U}(y|x,z,u) \tag{3}$$

with c some normalizing constant independent of u.

Let denote $x = (x_1, ..., x_k)$ and $y = (y_1, ..., y_n)$. Under our conditions, the posterior

distribution (3) becomes

$$
\begin{aligned}
\widehat{\Pi}_{U|X,Y,Z}(u|x,y,z) &= C(x,y,z)\,\pi_U(u)\,f_{Z|U}(z|u)\prod_{i=1}^{k}f_{X_i|Z,U}(x_i|z,u)\prod_{j=1}^{n}f_{Y_j|X,Z,U}(y_j|x,z,u)\\
&= C(m,x,y,z)\exp\Bigg\{-\frac{1}{2}\Bigg(\frac{(z-u)^2}{\sigma_Z^2}+\frac{1}{\sigma_X^2}\sum_{i=1}^{k}(\log x_i-\alpha z-(1-\alpha)u)^2\\
&\qquad\qquad\qquad\qquad +\frac{1}{\sigma_Y^2}\sum_{j=1}^{n}(y_j-\beta_1\overline{\log x}-\beta_2 z-\beta_3 u)^2\Bigg)\Bigg\}\\
&= \frac{1}{\sqrt{2\pi\hat\sigma^2}}\exp\Big\{-\frac{1}{2\hat\sigma^2}(u-\hat\mu)^2\Big\}
\end{aligned}
$$

where $\hat\sigma^2$ and $\hat\mu$ are given in (1) and (2) respectively, and where C is a normalizing function independent of u, that may change from line to line. $\square$

We can deduce from Theorem 1.1 the properties of the Bayes estimator $\hat\mu$.

LEMMA 1.1 *The Bayes estimator $\hat\mu$ is an unbiased estimator of $\mathbb{E}(U)$ with variance given by $var(\hat\mu)=2\hat\sigma^2$, with $\hat\sigma^2$ defined in (1); it is then a consistent estimator of $\mathbb{E}(U)$.*

It appears that the variance of the estimator is linked to the parameters α and β_3 via the variance $\hat\sigma^2$ of the posterior distribution.

Proof of Lemma 1.1.
The sum of the credibility weights being 1, $\hat\mu$ is unbiased. Using that $var(X_1)=\mathbb{E}\big(var(X_1|X_2)\big)+var\big(\mathbb{E}(X_1|X_2)\big)$ for any r.v. X_1 and X_2, straightforward computations lead to the expression of the variance of $\hat\mu$. Finally, the consistency follows via the Bienaymé-Chebyshev inequality, since $var(\hat\mu)$ goes to 0 as $k\to\infty$ (for any $\alpha\in[0,1)$), $\hat\sigma^2$ being a decreasing function of k. $\square$

The following specific cases are also consequences of Theorem 1.1.

COROLLARY 1.1

1. *If X and Y are conditionally independent given U and Z, then $\beta_1=0$, $\beta_2+\beta_3=1$, and the credibility weights of (2) become*

$$
\omega_X=\hat\sigma^2\Big(\frac{k(1-\alpha)}{\sigma_X^2}\Big)\ ;\ \omega_Y=\frac{\hat\sigma^2\beta_3 n}{\sigma_Y^2}\ ;\ \omega_Z=\hat\sigma^2\Big(\frac{1}{\sigma_Z^2}-\frac{k\alpha(1-\alpha)}{\sigma_X^2}-\frac{\beta_3(1-\beta_3)n}{\sigma_Y^2}\Big)
$$

with $\hat\sigma^2$ given in (1).

2. If we assume $\alpha = \beta_2 = 0$ and $\beta_1 + \beta_3 = 1$ (set $\beta_3 = \beta$), ie, when considering $Z|U$, $X|U$ and $Y|X,U$, we obtain that $\widehat{\Pi}_{U|X,Y,Z}$ is $\mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$ with

$$\hat{\mu} \;=\; \hat{\sigma}^2 \left(\left(\frac{k}{\sigma_X^2} - \frac{n\beta(1-\beta)}{\sigma_Y^2} \right) \overline{\log X} \;+\; \frac{n\beta}{\sigma_Y^2} \overline{Y} \;+\; \frac{1}{\sigma_Z^2} Z \right)$$

$$and \quad \hat{\sigma}^2 \;=\; \left(\frac{1}{\sigma_Z^2} + \frac{k}{\sigma_X^2} + \frac{n\beta^2}{\sigma_Y^2} \right)^{-1} \;\; where \;\; \beta \in]0,1].$$

If we add the assumption of conditional independence between X and Y given U, ie, for $\beta = 1$, then

$$\hat{\mu} = \hat{\sigma}^2 \left(\frac{k}{\sigma_X^2} \overline{\log X} + \frac{n}{\sigma_Y^2} \overline{Y} + \frac{1}{\sigma_Z^2} Z \right) \quad and \quad \hat{\sigma}^2 = \left(\frac{1}{\sigma_Z^2} + \frac{k}{\sigma_X^2} + \frac{n}{\sigma_Y^2} \right)^{-1} \tag{4}$$

Replacing the r.v. Z by a constant μ_{ext} leads to the same Bayesian estimator $\hat{\mu}$ of the location parameter as in [Lambrigger et al.(2007)] (see Theorem 4.3, or Theorem 4.1 in [Lambrigger et al.(2008)]) and the same $\hat{\sigma}^2$.

1.2 Practical application and discussion

We consider the same data as the ones in [Lambrigger *et al.*(2008)], Figure 2, to be able to draw a fair comparison with our result.

The experts opinion is on average given by $\bar{y} = 6$, whereas the risk profile market by $z = 2$. Hence an estimate can be proposed from (2) in Theorem 1.1 as

$$\hat{\mu} = \omega_X \overline{\log x} + \omega_Y \bar{y} + \omega_Z z$$

where the weights $\omega_\bullet$, given in Theorem 1.1, are functions of the parameters k (operational loss sample size), n (experts number), $\alpha \in [0,1)$, $\beta_i \in [0,1]$ ($i = 1, 2, 3$, $\beta_3 \neq 0$, $\sum_{i=1}^{3} \beta_i = 1$), and of the known variances σ_X^2, σ_Y^2, σ_Z^2.

We choose, as in [Lambrigger *et al.*(2008)], k varying up to $K = 70$, $\sigma_X = 4$, $\sigma_Y = 3/2$ and $\sigma_Z = 1$. We consider one panel of dependent experts (with the corresponding standard deviation), *id est* we take $n = 1$.

To check about the convergence of the Bayesian estimator to the location parameter, we simulate the operational loss data $(x_i)_{i=1,\ldots,k}$ from a $LN(4, \sigma_X^2)$ distribution.

We represent on a same graph four estimates of $\mathbb{E}(U)$ computed in terms of the number k of operational loss data (with $1 \le k \le K = 70$), respectively the Maximum Likelihood estimator (MLE) (given by $\hat{\mu}_k^{ML} = \frac{1}{k} \sum_{i=1}^{k} \log X_i$), the Lambrigger et al. (LSW) estimator which corresponds to our estimator with $\alpha = 0 = \beta_1 = \beta_2$ and $\beta_3 = 1$ (assuming in particular the conditional independence of X and Y), the Shevchenko et al. (SW) estimator when having no experts ($n = 0$), and our estimator $\hat{\mu}$ defined by (2) in Theorem 1.1.

Various graphs have been obtained when choosing for our estimator the coefficients α and β_i's varying in $[0,1]$, with a grid of step 0.1, providing more or less comparable

results; Figure 1 illustrates one of them where the conditional dependence appears clearly between the sources and is well distributed between them, with slightly more weight on U than Z ($\alpha = 0.3$, $\beta_1 = 0.3$, $\beta_2 = 0.2$, $\beta_3 = 0.5$).

Figure 1: Comparison of the various estimates of the location parameter μ

(A priori related to the market risk profile)
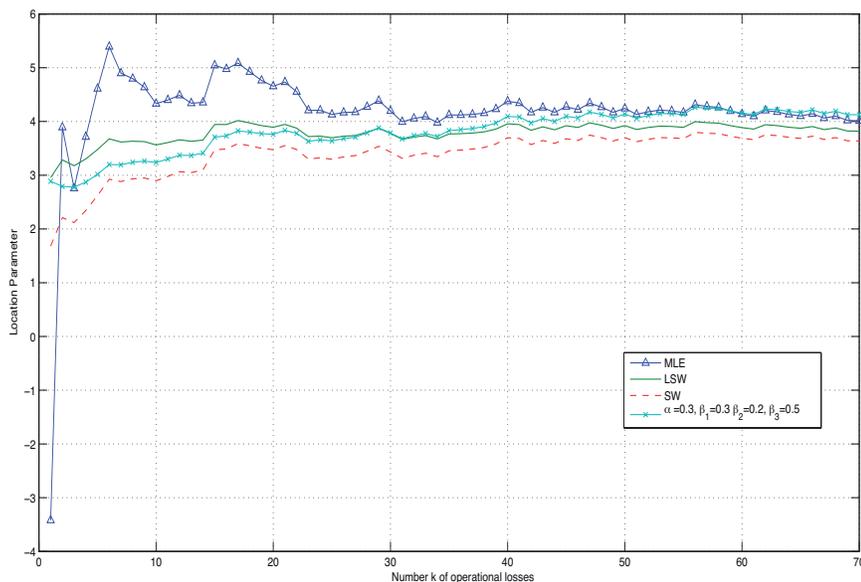
We can notice, as in [Lambrigger *et al.*(2008)], the high volatility of the ML estimator for very small k whereas the three other estimators appear more stable. Among the Bayesian estimators, we can confirm (whether we assume the conditional independence of X and Y or not) what noticed Lambrigger et al., namely that including the experts opinion improves the quality of the estimator. Under the assumption of conditional independence, the Bayesian estimator with parameters $\alpha = 0 = \beta_1 = \beta_2$; $\beta_3 = 1$ (denoted by LSW) behaves on this graph better for a very small number of data, whereas the general Bayesian estimator is closer to the true value for relatively small k and remains as such, presenting a behavior close to the one of the ML estimator whenever k is large (which is an advantage, even if we are not focusing on the behavior of the estimators as k becomes large). We can quantify this comparison using the mean square distance $d = d(\hat{\mu}, \mu)$ between the estimator and the true value $\mu = 4$. This distance d is smaller when using our estimator than the LWS estimator, whenever $k \geq 25$; eg, for $k = 25$ (ie, when summing from $k = 25$ to $K(=70)$), we obtained $d = 0.0561$ for the LWS estimator and $d = 0.0387$ for our estimator. Recall also that our model does not rely on the conditional independence of X and Y given Z.

1.3 Distribution of the risk market profile

It is important to understand how the company specific information in the Bayesian context modifies our view on the risk market profile Z. Indeed, it is interesting to know

how to modify the criteria on the risk market profile given by the control organisms (following the Basel or Solvency rules) whenever another information is provided by some experts (scientific academics or consulting companies) and/or by the operational loss data of the company, and/or under a given macroeconomics environment.

To measure this impact, we will compute the conditional distribution of the risk market profile Z given X, Y and U and compare it with the one of Z given U.

Let us compute the law $\mathcal{L}(Z|X, Y, U)$ of Z given X, Y and U. We can write

$$f_{Z|X,Y,U}(z|x, y, u) = \frac{f_{Z,X,Y|U}(z, x, y|u)}{\int f_{Z,X,Y|U}(z, x, y|u) \, dz}$$

Using that

$$f_{X_1}(z) f_{X_2}(z) = \frac{1}{\sqrt{2\pi(\gamma^2 + \sigma^2)}} \exp\left\{ -\frac{(a - b)^2}{2(\gamma^2 + \sigma^2)} \right\} f_{X_3}(z) \tag{5}$$

$$\text{where} \quad X_1 \overset{d}{\sim} \mathcal{N}(a, \sigma^2) \quad , \quad X_2 \overset{d}{\sim} \mathcal{N}(b, \gamma^2) \quad \text{and} \quad X_3 \overset{d}{\sim} \mathcal{N}\left(\frac{a\gamma^2 + b\sigma^2}{\gamma^2 + \sigma^2}, \frac{\gamma^2 \sigma^2}{\gamma^2 + \sigma^2} \right)$$

straightforward computations give

$$\begin{aligned}
f_{Z,X,Y|U}(z, x, y|u) &= \prod_{j=1}^{n} f_{Y_j|Z,U,X}(y_j|z, u, x) \prod_{i=1}^{k} f_{X_i|Z,U}(x_i|z, u) f_{Z|U}(z|u) \\
&= C(u, x, y) \exp\left\{ -\frac{1}{2\sigma^2}(z - a)^2 \right\} \exp\left\{ \frac{1}{2\sigma_z^2}(z - u)^2 \right\} \\
&= C(u, x, y) \exp\left\{ -\frac{1}{2\gamma_Z^2}(z - m_Z)^2 \right\}
\end{aligned}$$

where $C(u, x, y)$ denotes some constant depending on u, x and y which may change from line to line, and where

$$m_Z = \frac{a\sigma_Z^2 - u\sigma^2}{\sigma_Z^2 - \sigma^2} \quad ; \quad \gamma_Z^2 = \frac{\sigma^2 \sigma_Z^2}{\sigma_Z^2 - \sigma^2}$$

with

$$a := \sigma^2 \left[\left(\frac{1}{\sigma_Z^2} - \frac{k\,\alpha(1 - \alpha)}{\sigma_X^2} - \frac{n\beta_2\beta_3}{\sigma_Y^2} \right) u + \frac{\beta_2 n}{\sigma_Y^2} \, \bar{y} - \left(\frac{n\beta_1\beta_2}{\sigma_Y^2} - \frac{\alpha k}{\sigma_X^2} \right) \overline{\log x} \right]$$

$$\text{and} \quad \sigma^2 := \left(\frac{1}{\sigma_Z^2} + \frac{k\alpha^2}{\sigma_X^2} + \frac{n\beta_2^2}{\sigma_Y^2} \right)^{-1} \quad \text{(which satisfies } \sigma^2 < \sigma_Z^2\text{).}$$

We can then conclude to the following.

PROPOSITION 1.1 *The distribution of the market risk profile Z given the knowledge of the severity X, the expert opinion Y and the macroeconomic environment U, is a*

Gaussian distribution with mean M_Z equal to a linear combination of U, $\overline{Y}$ and $\overline{\log X}$ and with variance given in terms of the variances of X and Y:

$$Z|X,Y,U \;\sim\; \mathcal{N}\left(M_Z, \gamma_Z^2\right)$$

with

$$M_Z \;=\; \gamma_Z^2 \left[\left(\frac{\alpha k}{\sigma_X^2} - \frac{n\beta_1\beta_2}{\sigma_Y^2}\right)\overline{\log X} \;+\; \frac{\beta_2 n}{\sigma_Y^2}\,\overline{Y} \;-\; \left(\frac{k\;\alpha(1-\alpha)}{\sigma_X^2} + \frac{n\beta_2\beta_3}{\sigma_Y^2}\right)U\right] \quad (6)$$

and

$$\gamma_Z^2 = \left(\frac{k\alpha^2}{\sigma_X^2} + \frac{n\beta_2^2}{\sigma_Y^2}\right)^{-1} \quad (7)$$

Note that the conditional distribution of $Z|X,Y,U$ does not depend on the parameter σ_Z^2.

Although both conditional distributions of $Z|U$ and $Z|X,Y,U$ are normal, the parameters are quite different showing how the external risk market profile modifies for a given company. The variance γ_Z^2 of $Z|X,Y,U$ is a decreasing function in k, hence becomes soon or later smaller than the given variance σ_Z^2 of $Z|U$. The two conditional expectations $\mathbb{E}[Z|U] = U$ and $\mathbb{E}[Z|X,Y,U] = M_Z$ can not equal a.s., whatever is the choice of the coefficients α and β_i, $i = 1,2,3$, but both are normal, with the same expectation $\mathbb{E}[M_Z] = \mathbb{E}[U]$ that can be estimated by $\hat{\mu}$ given in (2).

Let us revisit our numerical example developed in §1.2. Consider the same values of parameters, namely $\bar{y} = 6$, $\sigma_X = 4$, $\sigma_Y = 3/2$, $\sigma_Z = 1$, $n = 1$, k varying up to $K = 70$, $(x_i)_{i=1,\dots,k}$ simulated from a $LN(4, \sigma_X^2)$ distribution, $(\alpha, \beta_1, \beta_2, \beta_3) = (0.3, 0.3, 0.2, 0.5)$, and choose eg, $u = 4.5$.

We deduce a possible realization of M_Z, given by

$$m_Z = \gamma_Z^2 \left[\left(\frac{\alpha k}{\sigma_X^2} - \frac{n\beta_1\beta_2}{\sigma_Y^2}\right)\overline{\log x} \;+\; \frac{\beta_2 n}{\sigma_Y^2}\,\bar{y} \;-\; \left(\frac{k\;\alpha(1-\alpha)}{\sigma_X^2} + \frac{n\beta_2\beta_3}{\sigma_Y^2}\right)u\right]$$

with $\gamma_Z^2 = \gamma_Z^2(k)$ defined in (7), decreasing function in k.

We represent it in Figure 2 below.

Since $Z|X,Y,U \;\sim\; \mathcal{N}\left(M_Z, \gamma_Z^2\right)$, let us represent also $m_Z \pm \gamma_Z$.

Figure 2: A realization m_Z of the conditional expectation $\mathbb{E}[Z|X,Y,U]$ of the risk profile Z for
α =0.3, β_1=0.3, β_2=0.2, β_3=0.5
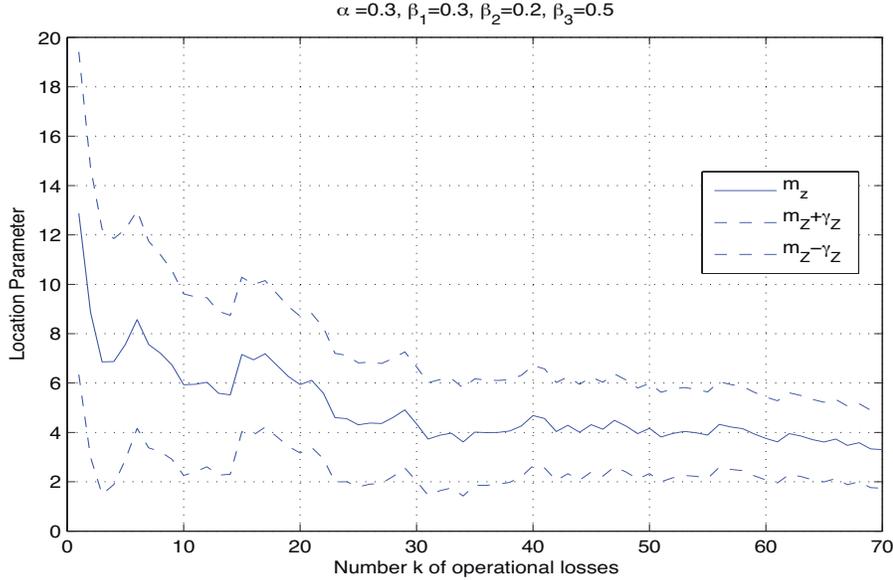
2 An *a priori* related to the severity

An alternative approach to handle our problem is to consider the point of view of the company and to relate the *a priori* directly to the severity of the operational loss, because of the few available data.

Suppose for instance that the severity X is log-normally distributed $LN(\beta_X, \sigma_X^2)$ and let us focus for instance on the position parameter β_X, fixing the shape parameter σ_X^2. Then, following our approach, we introduce a position parameter U having a non-informative prior distribution Π_U. As in the previous section, we choose Π_U to be uniform.

Suppose that the conditional operational losses $(X_i|U)_{1 \leq i \leq k}$ are i.i.d., the experts opinion $(Y_i|X,Z,U)_{1 \leq i \leq n}$ are i.i.d., with parent variables denoted by $X|U$ and $Y|X,Z,U$ respectively.

Consider the following conditional distributions:

$$X|U \overset{d}{\sim} LN(U, \sigma_X^2); \quad Z|U,X \overset{d}{\sim} \mathcal{N}(\lambda \overline{\log(X)} + (1-\lambda)U, \sigma_Z^2) \quad \text{with} \quad \lambda \in [0,1);$$

$$Y|X,Z,U \overset{d}{\sim} \mathcal{N}(\beta_1 \overline{\log X} + \beta_2 Z + \beta_3 U, \sigma_Y^2), \quad \text{with} \quad \sum_{i=1}^{3} \beta_i = 1, \beta_i \in [0,1], \beta_3 \neq 0,$$

where the variances σ_X^2, σ_Y^2 and σ_Z^2 are supposed to be known or estimated on data when possible.

Note that $\mathbb{E}[\log(X)] = \mathbb{E}(U)$, $\mathbb{E}(Z) = \mathbb{E}(U)$, Z denoting the position parameter of the market risk profile, specific to the company, and that $\mathbb{E}(Y) = \mathbb{E}(U)$.

11

2.1 The posterior distribution and Bayes estimator

The Bayes theorem and straightforward computations on Gaussian distributions provide the posterior density function $\widehat{\Pi}_{U|X,Y,Z}$ of U given X, Y, Z as

$$
\begin{aligned}
\widehat{\Pi}_{U|X,Y,Z}(u) &= C\pi_U(u)f_{X|U}(x|u)f_{Z|X,U}(z|x,u)f_{Y|X,Z,U}(y|x,z,u) \\
&= C\,\pi_U(u)\,f_{Z|X,U}(z|x,u)\prod_{i=1}^{k}f_{X|U}(x_i|u)\prod_{j=1}^{n}f_{Y|X,Z,U}(y_j|x,z,u) \\
&= \frac{1}{\sqrt{2\pi\gamma^2}}\,\exp\left\{-\frac{1}{2\gamma^2}(u-\delta\gamma^2)^2\right\}
\end{aligned}
$$

with C some normalizing constant depending on x, y, u, and where

$$
\gamma^2 = \left(\frac{(1-\lambda)^2}{\sigma_Z^2}+\frac{n\beta_3^2}{\sigma_Y^2}+\frac{k}{\sigma_X^2}\right)^{-1} \tag{8}
$$

and

$$
\delta = \frac{n\beta_3}{\sigma_Y^2}\bar{y}+\left(\frac{1-\lambda}{\sigma_Z^2}-\frac{n\beta_2\beta_3}{\sigma_Y^2}\right)z+\left(\frac{k}{\sigma_X^2}-\frac{\lambda(1-\lambda)}{\sigma_Z^2}-\frac{n\beta_1\beta_3}{\sigma_Y^2}\right)\overline{\log x}. \tag{9}
$$

We obtain the following result.

THEOREM 2.1 *The posterior distribution $\widehat{\Pi}_{U|X,Y,Z}$ is a Gaussian distribution $\mathcal{N}(\tilde{\mu},\gamma^2)$ where the variance γ^2 is given in (8) and the mean $\tilde{\mu}$ satisfies*

$$
\tilde{\mu} = \omega_X\overline{\log X}+\omega_Y\overline{Y}+\omega_Z\,Z \tag{10}
$$

with the credibility weights $\omega_\cdot$ given by

$$
\omega_X = \gamma^2\left(\frac{k}{\sigma_X^2}-\frac{\lambda(1-\lambda)}{\sigma_Z^2}-\frac{n\beta_1\beta_3}{\sigma_Y^2}\right) \;;\; \omega_Y = \gamma^2\,\frac{n\beta_3}{\sigma_Y^2} \;;\; \omega_Z = \gamma^2\left(\frac{1-\lambda}{\sigma_Z^2}-\frac{n\beta_2\beta_3}{\sigma_Y^2}\right)
$$

It can be checked, from Theorem 2.1, that the Bayes estimator $\tilde{\mu}$ is an unbiased and consistent estimator of $\mathbb{E}(U)$.

When looking at specific values of the parameters α and β_i's, we notice that assuming $\lambda = \beta_1 = \beta_2 = 0$, $\beta_3 = 1$, ie, considering $Z|U$, $X|U$ and $Y|U$, provides

$$
\tilde{\mu} = \hat{\gamma}^2\left(\frac{k}{\sigma_X^2}\,\overline{\log X}+\frac{n}{\sigma_Y^2}\overline{Y}+\frac{1}{\sigma_Z^2}\,Z\right) \text{ and } \gamma^2 = \left(\frac{1}{\sigma_Z^2}+\frac{k}{\sigma_X^2}+\frac{n}{\sigma_Y^2}\right)^{-1}
$$

which corresponds also to the LWS estimator proposed in [Lambrigger *et al.*(2007)] when replacing the r.v. Z by a constant μ_{ext}.

Considering the same frame as in our example developed in §1.2, let us see how this new Bayesian estimator $\tilde{\mu}$ of the location parameter $\mathbb{E}(U)$ behaves.

Let $\bar{y} = 6$, $z = 2$, $n = 1$, $\sigma_X = 4$, $\sigma_Y = 3/2$ and $\sigma_Z = 1$, k varying up to $K = 70$, $(x_i)_{i=1,\dots,k}$ simulated from a $LN(4, \sigma_X^2)$ distribution.

Figure 3: Comparison of the various estimates of the location parameter μ

(A priori related to the severities)
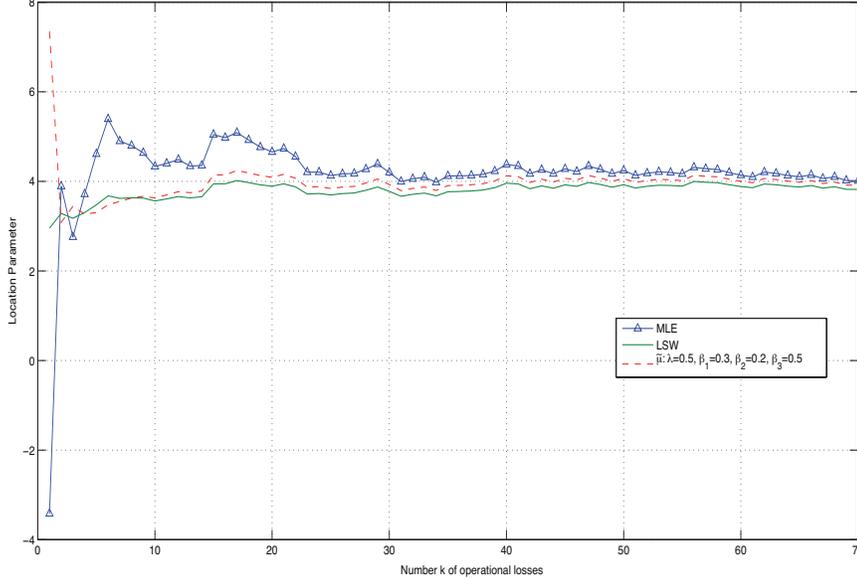
We represent in Figure 3 above an estimate of $\tilde{\mu}$ defined in (10) with distributed parameters $\lambda = 0.5$, $\beta_1 = 0.3$, $\beta_2 = 0.2$ and $\beta_3 = 0.5$, an estimate of $\tilde{\mu}$ with specific parameters $\lambda = \beta_1 = \beta_2 = 0$ and $\beta_3 = 1$ corresponding for fixed Z to the LWS estimator (with the conditional independence assumption), and also the MLE.

We see that the estimate of $\tilde{\mu}$ with 'non specific' parameters fits the best the underlying real value μ starting at very small k. As in the previous section, various graphs have been obtained when taking the parameters on a grid; they provide more or less the same behavior, with a convergence to μ from a small k. We chose to present one example with balanced parameters. Using the mean square distance $d = d(\tilde{\mu}, \mu)$ between the estimator and the true value $\mu = 4$, we obtain $d = 0,0560$; $0,0173$; $0,0082$ whenever $k \geq 2$; 9; 20 respectively, whereas we obtain $d = 0,0678$; $0,0377$; $0,0288$ respectively for the LWS estimator, confirming the improvement when introducing the Bayesian estimator built with a general method.

We also notice that considering $\tilde{\mu}$ defined in (10) provides better results than the estimator $\hat{\mu}$ defined in (2). It would mean that it is more relevant to relate the *a priori*, namely the macroeconomics environment, to the severities rather than to the risk profile. Intuitively this result corresponds to what has been seen in the market with the phenomenal operational losses of Société Générale during the financial crises and UBS during the sovereign debt crises.

Let us compare in Figure 4 the estimates of $\tilde{\mu}$ and $\hat{\mu}$ computed on the same sample

13

$(X_i)_{i=1,\cdots,k}$, in the same frame, and for the α and β_i's previously selected.

Figure 4: Comparison of estimates of $\hat{\mu}$ and $\tilde{\mu}$
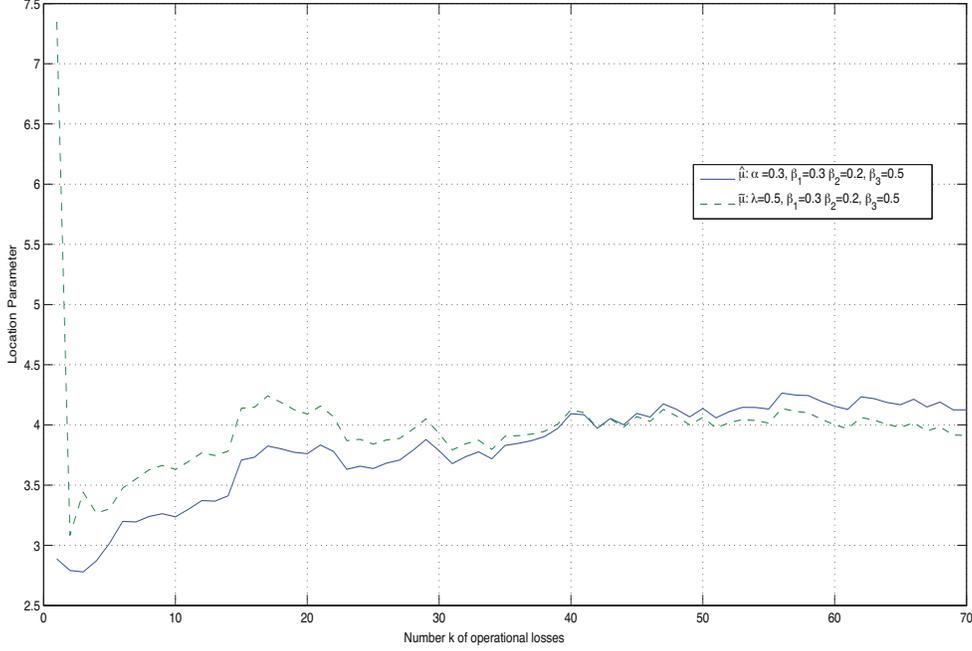
As already noticed, it appears clearly in Figure 4 that $\tilde{\mu}$ converges faster to μ, even for small values of k, than $\hat{\mu}$ does.

2.2   The conditional distribution of the risk market profile

Let us now compute the law $\mathcal{L}(Z|X,Y,U)$ of Z given X,Y and U, to measure the impact on the choice of the variable to which the *a priori* is directly related.
We can write

$$f_{Z|X,Y,U}(z|x,y,u) \;\; = \;\; \frac{f_{Y|X,U,Z}(y|x,u,z)f_{Z|X,U}(z|x,u)}{\int f_{Y|X,U,Z}(y|x,u,z)f_{Z|X,U}(z|x,u)dz}$$

which involves only Gaussian distributions. We obtain, using (5),

$$f_{Z|X,Y,U}(z|x,y,u) = \frac{1}{\sqrt{2\pi\Gamma_Z^2}} \; \exp\Big\{-\frac{1}{2\Gamma_Z^2}(z-m_Z^*)^2\Big\}$$

where

$$m_Z^* \;\; = \;\; \frac{n\beta_2\sigma_Z^2\,\bar{y} + (\lambda\sigma_Y^2 - n\beta_1\beta_2\sigma_Z^2)\,\overline{\log x} + ((1-\lambda)\sigma_Y^2 - n\beta_2\beta_3\sigma_Z^2)\,u}{\sigma_Y^2 + n\beta_2^2\sigma_Z^2}$$

$$and$$

$$\Gamma_Z^2 \;\; = \;\; \frac{\sigma_Y^2\sigma_Z^2}{\sigma_Y^2 + n\beta_2^2\sigma_Z^2}$$

14

We can conclude to the following proposition, showing how the company specific information in the Bayesian context modifies the view on the risk market profile Z.

PROPOSITION 2.1 *The distribution of the market risk profile Z given the knowledge of the a priori parameter U and of both the severity distribution X and the expert opinion Y, is a Gaussian distribution with mean M_Z^* equal to a linear combination of U, $\overline{Y}$ and $\overline{\log X}$ and with variance Γ_Z^2 given in terms of the variance of X and Y:*

$$Z|X,Y,U \ \sim \ \mathcal{N}\left(M_Z^*,\Gamma_Z^2\right)$$

with

$$M_Z^* \ = \ \Gamma_Z^2 \left[\left(\frac{\lambda}{\sigma_Z^2} - \frac{n\beta_1\beta_2}{\sigma_Y^2}\right) \overline{\log X} + \frac{n\beta_2}{\sigma_Y^2} \ \overline{Y} + \left(\frac{1-\lambda}{\sigma_Z^2} - \frac{n\beta_2\beta_3}{\sigma_Y^2}\right) U\right] \quad (11)$$

satisfying $\mathbb{E}(M_Z^*) = \mathbb{E}(U)$, and

$$\Gamma_Z^2 = \left(\frac{n\beta_2^2}{\sigma_Y^2} + \frac{1}{\sigma_Z^2}\right)^{-1} \quad (12)$$

Using the same sample $(x_i)_{i=1,\dots,k}$ (simulated from a $LN(4,\sigma_X^2)$ distribution) as in the previous section, let us represent in Figure 5 an estimate m_Z^* of M_Z^* with the values of λ and β_i previously selected, and with the same values for the other parameters as in the previous section when considering m_Z ($\bar{y} = 6$, $\sigma_X = 4$, $\sigma_Y = 3/2$, $\sigma_Z = 1$, $n = 1$, $u = 4.5$). The conditional variance is independent of the size of the operational losses and worths $\Gamma_Z^2 = 0.98$.

Figure 5: **A realization m_Z^* of the conditional expectation $\mathbb{E}[Z|X,Y,U]$ of the risk profile for**


15
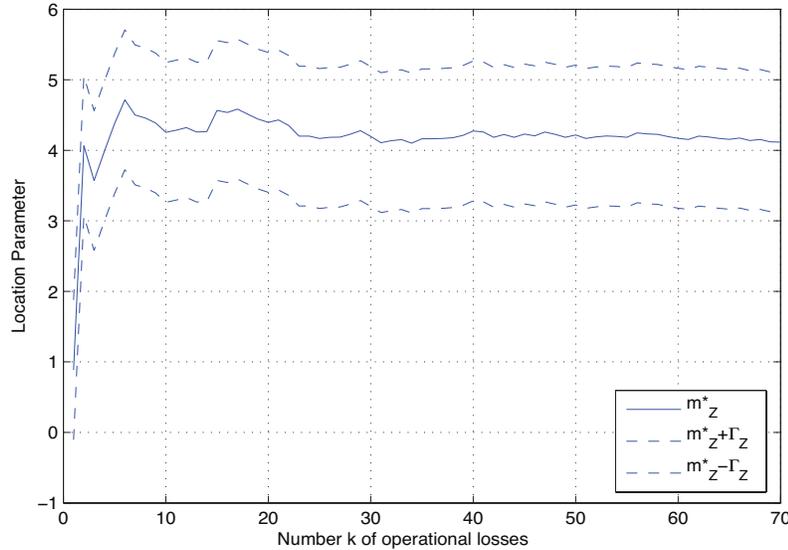
Finally, we may compare the distribution of Z given X, Y, U when considering the *a priori* on the unknown parameter of a given family of distributions for the severities and for the risk profile, respectively. As can be seen in Figure 6 below, the behavior of m_Z^* appears much more stable than the one of m_Z. It can be explained by the fact that m_Z relies directly on the operational loss index k, whereas m_Z^* depends on k through the empirical mean $\overline{\log x}$ only.

Figure 6: Comparison of m_Z and $m*_Z$ for
$\lambda = 0.3$, $\alpha = 0.5$, $\beta_1 = 0.3$, $\beta_2 = 0.2$, $\beta_3 = 0.5$
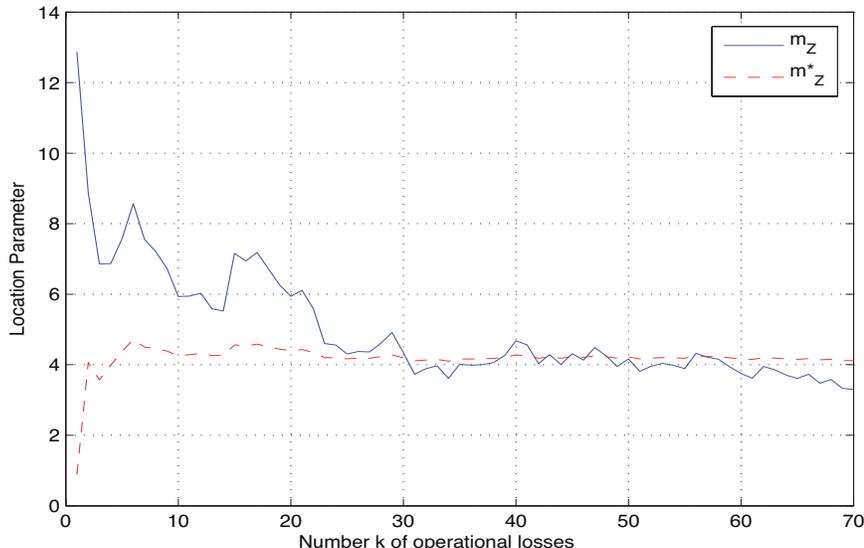
Conclusion

This paper shows how to include in the modeling of operational risk, macroeconomic effects and expert judgements through a Bayesian approach. Such an approach is chosen to obtain a good behavior of estimators even in the case of small sample of losses. To illustrate our approach, we assumed a uniform *a priori* distribution on a location parameter and considered normal or lognormal conditional distributions of the three factors, operational losses, market risk profile and experts opinion. It may be applied to a different parameter or with other types of distributions, although our choice appears reasonable in the insurance context where extreme operational losses have not been seen. It would be interesting to introduce a heavy tail distribution, for instance a Fréchet distribution, when treating operational losses in banks that have experienced extreme ones. This study has been developed when relating the *a priori*, namely the macroeconomics environment, first to the market risk profile, then to the operational losses. In both situations, our generalized estimator of the location parameter is given as a weighted average of the three mentioned factors. Properties of this estimator have been deduced and simulation performed to see how it behaves. It appears that

introducing a hyper level in the prior modeling provides a quite competitive estimator when compared with the maximum likelihood estimator and the Lambrigger et al.'s one. The latter can be deduced as a particular case (under a strong independence condition) of our general estimator. Finally we notice also that relating the *a priori* to operational losses rather than to the market risk profile leads to a much more stable behavior close to the true value even when few data are available. So it seems that macroeconomics environment has a direct impact on the severities of operational losses. It would deserve further empirical investigation on the influence of macroeconomics on operational risk. It might be undertaken in some future work.

References

[Berger(2000)] J. O. BERGER, Bayesian analysis: a look at today and thoughts of tomorrow. *J. Amer. Statist. Assoc.* **95**, (2000) 1269–1276.

[Klugman *et al.*(2008)] S. KLUGMAN, H. PANJER, G. WILLMOT, *Loss models: From Data to Decisions.* Wiley (2008).

[Lambrigger *et al.*(2007)] D. LAMBRIGGER, P. SHEVCHENKO, M.WÜTHRICH, The quantification of operational risk using internal data, relevant external data and expert opinion. *J. Operational Risk* **2**(3), (2007) 3-27.

[Lambrigger *et al.*(2008)] D. LAMBRIGGER, P. SHEVCHENKO, M.WÜTHRICH, Data combination under Basel II and Solvency 2: operational risk goes bayesian. *Bulletin Français d'Actuariat* **8**(16), (2008) 4-13.

[McNeil *et al.*(2005)] A. MCNEIL, R. FREY, P. EMBRECHTS, *Quantitative Risk Management.* Princeton (2005).

[Robert(2007)] C.P. ROBERT, *The Bayesian Choice.* Springer (2007, 2nd Ed.).

[Shevchenko *et al.*(2006)] P. SHEVCHENKO, M.WÜTHRICH, The stuctural modeling of operational risk via Bayesian inference: combining loss data with expert opinions. *J. Operational Risk* **1**(3), (2006) 3-26.

Informations

ESSEC
BUSINESS SCHOOL