Operational risk modelled analytically

The current regulatory framework allows banks to compute their capital charge for operational risk under an internal model, which is often based on the loss distribution approach (LDA). Loss distributions are calibrated at the cell level (a cell is the elementary risk unit per business line and type of risk) and the bank’s capital charge is estimated by aggregating cell loss distributions under some dependence assumption (Chernobai, Rachev & Fassozzi 2007).

The Basel Committee provides some guidelines about how banks should appropriately reflect the risk profile in their internal model (Bank for International Settlements 2011). However, banks benefit from some flexibility in their modelling choices and this may lead to some discrepancies in capital charges for similar risk profiles. The broad range of practices that are observed among banks is the result of different distributional or dependence assumptions in their models.

Many studies have focused on the modelling of the tails of the severity distributions (Dutta & Perry 2007; Moscledeli 2004), but the bulk of the correlation problem is still unsolved and controversial. There is much debate about the choice of copula function for losses across cells due to the scarcity of data, but the regulators advise banks to determine sound correlations and to retain conservative assumptions. Some institutions have selected the simplest option and use equal correlations between cell losses. This assumption is questionable, of course, and may embed some model risk, but both regulators and practitioners have great difficulty in agreeing on realistic and conservative correlation levels. Some authors believe that correlations between cell losses are as low as 4% (Frachot, Roncalli & Salomon 2004).

Most of the knowledge we have about operational loss quantification comes from complex models and heavy Monte Carlo simulations, and as far as we know there is no analytical model that takes into account risk and correlation dispersion among cells. This article fills this gap. Under the asymptotic single risk factor assumption, we obtain new results for the bank’s capital charge sensitivity to the critical parameters of the model. In particular, we show the capital charge is not that sensitive to correlation dispersion, and that the constant correlation assumption is robust.

This new result is obtained for few specifications, and we conjecture that it remains valid at least qualitatively, for real bank portfolios that have a finite number of cells. We believe that our approach also provides a way to pioneer a new method for computing capital charges and challenging internal model assumptions, as exemplified in this paper.

This paper is organised as follows. First, we provide some real data evidence about cell loss distributions and correlations. Second, we solve the asymptotic single risk factor model with lognormal losses at the cell level, even when individual cells have varying risk profiles. Third, we solve the case of non-equal correlations between cells and provide some key results about the sensitivity of capital charge to the main critical parameters of the model.

Some empirical facts about cell loss distributions and correlations

In the LDA framework, the aggregate operational loss for cell number \( i \) is equal to the sum of individual losses:

\[
L_D = \sum_{n=1}^{N_i} X_n
\]

where \( L_D \) is the aggregate loss of cell number \( i \), \( N_i \) is the number of events over one year, and \( (X_i)_{1 \leq i \leq M} \) is the sequence of individual loss severities for cell number \( i \). The aggregate loss process is a compound Poisson process and, accordingly, the model is based on the following assumptions.

- The number of events and their severity are independent.
- Loss severities are independent and identically distributed random variables.
- Cell loss distribution parameters. There are many studies in the literature about individual loss distributions (see, for example, Dutta & Perry 2004; Moscledeli 2004), but there are very few empirical studies about aggregate cell losses.

We have conducted a study based on the SAS OpRisk Global Data database. As of November 2013, this database included 6,402 events that had occurred in financial firms since 2002, when financial institutions started to collect and report their operational losses systematically. We have calibrated the frequency of events and lognormal severity distributions for each of the 21 cells that have more than 30 losses. Direct calibration of the aggregate loss distribution from real data is, of course, impossible because there is only one observation per year. However, it is possible to assess the fit with the lognormal distribution of the aggregate loss distribution obtained through the LDA.

Let us assume that the loss distribution for cell \( i \) is lognormal with parameters \( \mu_i \) and \( \sigma_i \); the ratio between the expected value and any quantile depends only on the parameter \( \sigma_i \):

\[
\text{expected value} (i) = e^\mu_i - \frac{\sigma_i}{2} F^{-1}_q
\]

\[
\alpha_i = -F_{q} - \frac{1}{2} \ln \frac{\text{expected value} (i)}{\text{VaR}_q (i)}
\]

where \( \text{VaR}_q \) is the \( q \)-percentile of the lognormal distribution and \( F_{q} = N^{-1} (1-q) \). Inverting (2) leads to two different solutions, and we have chosen the one with a minus sign in front of the square root in (3) because we require the parameters \( \sigma_i \) to decrease with the ratio of expected value to quantile for all cells. We observe that broader distribution assumptions for cell losses in the model can naturally be taken into account by choosing the plus sign solution in (3).
To assess the robustness of these estimates, we compute the median of the observed values of \( \sigma_i \), which at 108.5% is very close to the average value. The med-med estimator (the median value of the spread with the median) is equal to 107%, and the observed standard deviation is equal to 42%.

### Cell loss correlations.

In most studies (see, for example, Aue & Kalkbrener 2007; Frachot, Roncalli & Salomon 2004) cell loss correlations are calculated from the dependence of the number of events between cells rather than the dependence of severities. Under the assumption of lognormal severity distributions (Kalkbrener 2007; Frachot, Roncalli & Salomon 2004), cell loss correlation between cells is given by:

\[
\text{corr}(L_1, L_2) = \text{corr}(N_1, N_2) e^{-r_1/2 - r_2/2} \tag{4}
\]

The correlation between the number of events, \( N_1 \) and \( N_2 \), is linked to the loss frequencies of cells 1 and 2. \( s_1 \) and \( s_2 \) are the standard deviation parameters of the two distributions. Bivariate Poisson variables are obtained by considering three independent Poisson variables \( Z, Y_1 \) and \( Y_2 \) with parameters \( \lambda_1 \) and \( \lambda_2 \) respectively, the variables \( N_i = Z + Y_i \) are also Poisson with intensities \( \lambda_i \), and their correlation is given by:

\[
\text{corr}(N_1, N_2) = \frac{1}{\sqrt{\lambda_1 \lambda_2}} \gamma \max(\lambda_1, \lambda_2) \tag{5}
\]

The upper bound \( R \) for the correlation comes from the inequalities \( \lambda_1 \geq \gamma \lambda_2 \) and \( \lambda_2 \geq \gamma \lambda_1 \). Wherever the bank's portfolio includes a large number of cells, the severities are distributed as a random variable. Internal data represents frequencies better than external data because it is specific to the bank, and it also includes the frequencies of rare but severe events that are taken into account by the scenario analyses in the model. Internal data and scenario analysis frequencies at Société Générale support the normal distribution assumption of the log-intensities of the Poisson processes (in particular, the skewness and normalised kurtosis are close to 0) with a standard deviation equal to \( \gamma = 2.35 \pm 0.35 \).

### Correlation parameters in the Gaussian copula model.

In the Gaussian copula framework with lognormal marginal cell losses, the correlation parameter \( \rho_{ij} \) between two cells is related to the cell loss correlation:

\[
\text{corr}(L_i, L_j) = \frac{e^{\rho_{ij}/\sqrt{\sigma_i \sigma_j}} - 1}{\sqrt{(e^{\sigma_i/2} - 1)(e^{\sigma_j/2} - 1)}} \tag{7}
\]

This formula with parameters \( \sigma_i = \sigma_j = 107\% \) and a conservative assumption for loss correlations of \( \text{corr}(L_i, L_j) = 4\% \) leads to a correlation parameter of the Gaussian copula equal to \( \rho_{ij} = 7.2\% \). External data supports the assumption of very low correlation parameters in the copula framework: much lower than 10%.
A class of solvable models with correlated risks

- A simplified LDA model. The rest of this article is dedicated to building a simple portfolio model for operational risk. We assess that the bank’s overall operational risk is composed of a portfolio of operational risks at cell level. We make the following assumptions.

  Lognormal distributions: the loss for cell number \( i \) is a lognormal random variable \( L_i \) with parameters \( \mu_i \) and \( \eta_i \). As discussed earlier in this article, we assume that the \( \eta_i \) have an expected value of \( \sigma = 10\% \) and, unless otherwise stated, a variance of \( \nu = 18\% \).

  Gaussian copula: pairwise correlations \( \rho_{ij} \) may be different to each other. For numerical estimations, we assume that the pairs of cells, we assume that the parameters \( \eta_i \) are not dependent on the number of cells \( N \).

In this framework, the annual loss for a cell can be written as the exponential function of a normal random variable that is a linear combination of the systemic and specific factors. For cell number \( i \) \((i = 1, \ldots, N)\) we obtain:

\[
L_i = \exp(\mu_i - \sigma_i(\beta_i + \sqrt{1 - \beta_i^2}\eta_i))
\]

The parameters \( \beta_i \) are linked to the pairwise correlations of the Gaussian copula: \( \rho_{ij} = \beta_i \). Because cells may have very different risk characteristics, and because correlations may be very different for different pairs of cells, we assume that the parameters \( \mu_i, \sigma_i \), and \( \beta_i \) are the observations of independent and identically distributed random variables called \( M, \Sigma \) and \( B \), respectively. In the limit \( N \to \infty \), the bank’s loss is equal to \( N \cdot L(F) \) and is a function of the common factor \( F \), as in Vasicek’s model for granulating homogeneous loss distributions (see Vasicek 2002).

\[
L(F) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} I_i(F) = \frac{\mathbb{E}[e^{M - \frac{X^2}{2}} + \sqrt{1 - \beta^2}(\eta_i)]}{\mathbb{E}[\eta_i^{M}]} \mathbb{E}[F]
\]

Without loss of generality, we assume that \( M = 0 \) because this simply rescales the bank’s loss by a constant factor \( \mathbb{E}[\eta_i^{M}] \) in the \( N \to \infty \) limit.

The stand-alone capital for cell number \( i \), called KSA0, is equal to the 99.9% percentile of the cell loss distribution and is given by

\[
\text{KSA}_0 = e^{-\bar{\mu}_{F_i}}, \quad \text{where } F_i = N^{-1}(0.1\%).
\]

The bank’s capital charge is equal to \( N \cdot L(F_i) \). The capital reduction coming from risk diversification is measured by the Diversification Index, which is defined as:

\[
\text{Diversification Index} = DI = \frac{N \cdot L(F)}{\sum_{i=1}^{N} \mathbb{E}[\eta_i^{M}]} \mathbb{E}[F]
\]

- Homogeneous risks. The simplest solvable model is obtained for homogeneous risks: the random variables \( L \) and \( B \) have constant values equal to \( \sigma \) and \( \sqrt{\nu} \), respectively, for all cells, and \( v = 0 \). In the limit \( N \to \infty \), the bank’s loss distribution remains lognormal:

\[
L_i = \exp(\mu_i - \sigma_i(\beta_i + \sqrt{1 - \beta_i^2}\eta_i))
\]

As in Markowitz’s portfolio theory, risk is not diversified away because of loss correlations between cells. The correlation parameter determines \( DI \):

\[
DI = \exp(\sigma(1 - \sqrt{\nu}))(1 - \rho)/2
\]

For \( \sigma = 10\% \) and \( \rho = 10\% \), we get \( DI = 17.5\% \). We note that \( DI \leq 1 \), which implies \( \nu > \frac{\sigma^2}{1 - \rho} \). The loss at cell level is:

\[
L_i = e^{-\bar{\mu}_{F_i} + \sqrt{\nu} \sqrt{\sigma}}
\]

In the limit \( N \to \infty \), the bank’s loss derived from (9) as a Gaussian integral:

\[
L(F) = \int_{-\infty}^{\infty} dx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left( 1 - \frac{x}{\sqrt{\nu}} \right)^{\nu/2} \left( \sigma \sqrt{\nu} + \frac{\sigma^2}{2} \right)
\]

where \( g(x) = e^{-x^2/2} / \sqrt{\nu} \). The bank’s loss follows a \( g \)-and-\( h \) distribution. Because we assume that the random variable \( \Sigma \) is normal and, strictly speaking, could take negative values, the resulting bank’s loss is a non-decreasing function of the systemic factor. However, this occurs when the systemic factor \( F \) is larger than \( F^* = \sigma / \sqrt{\nu} \), which is very unlikely in practice (for instance, \( F^* = 18.8 \) at \( \rho = 10\% \).
The bank’s capital charge, \( N \cdot L(F) \), increases with \( \rho \). For \( \rho = 10\% \), \( \sigma = 107\% \) and \( \rho = 18\% \), this increase is equal to +62\%. Unsurprisingly, the capital charge is very sensitive to risk dispersion measured by the parameter \( v \), which is a critical parameter of the model. Changing the value of \( \rho \) from 10\% to 20\% leads to a capital charge increase of +62\%. The average correlation level is also a critical parameter of the model. As the sum of the stand-alone capital charges is equal to \( N \cdot E[\beta_i \cdot \beta_j] \), the resulting DI is a decreasing function of \( \rho \). For equal parameters \( \sigma_i \), when included as uncorrelated additional risk, increase the capital charge but generate more diversification, as illustrated in figure 2.

**Uncertain correlations**

Correlations are not identical to each other, as discussed earlier, but estimating them from real data is a challenge from a statistical point of view. Data is scarce, limited as it is to only one observation per year for the aggregate loss. Estimation of the correlations between the number of events in each cell is no longer random. For the same reason, and severity correlations are only observable for cells that have a sufficient number of events per year. It is current practice to assume identical correlations among cells even if, in reality, correlations are unknown parameters. In what follows, we estimate the limit \( N \to \infty \) and correlation uncertainty is included in the model by assuming that the random variable \( B \) has an expected value equal to \( \beta = \sqrt{\sigma} \) and a variance equal to \( \sigma \). For the sake of clarity, we assume that individual risks are all equal among cells, so \( N \cdot \Sigma \) is a constant equal to \( \sigma \). From (9) we obtain, for the limit \( N \to \infty \):

\[
L(F) = \sum_{i=1}^{\infty} \frac{1}{N} \sum_{t=1}^{\infty} F_t = E[e^{-\beta F + \sigma^2 F^2 / 2 + F}] = \int_{-\infty}^{\infty} f(x) e^{-\beta F + \sigma^2 F^2 / 2} dx \tag{13}
\]

where the function \( f(x) \) is the density of the random variable \( B \). If we assume that the variable \( B \) is normally distributed, \( (B \sim N(\beta, \sigma)) \), we obtain:

\[
L(F) = \frac{1}{\sqrt{1 + \sigma^2 \beta^2}} \exp \left( -\beta \sigma F + \frac{(1 + \frac{\beta^2}{2}) \sigma^2}{\sqrt{1 + \sigma^2 \beta^2}} \right) \tag{14}
\]

If the variable \( B \) is uniformly distributed between \( \beta - \sqrt{\sigma} \) and \( \beta + \sqrt{\sigma} \) (the bounds are chosen to be the expected value and the variance are equal to \( \beta \) and \( \sigma \), respectively), we obtain:

\[
L(F) = \frac{1}{\sqrt{8 \pi \sigma^2 \beta^2}} \exp \left( -\beta \sigma F + \frac{(1 + \frac{\beta^2}{2}) \sigma^2}{\sqrt{8 \pi \sigma^2 \beta^2}} \right) \tag{15}
\]

As pairwise correlations are equal to \( \rho_{ij} = \rho_i \cdot \rho_j \), there is a direct link between the variance of \( \rho_{ij} \) and the variance of the sensitivity parameter \( \rho_i \). Because of the independence of the \( \rho_i \), we have:

\[
\text{var}(\rho_{ij}) = E[\beta_i^2 \cdot \beta_j^2] - E[\beta_i^2] \cdot E[\beta_j^2] = w(w + 2\beta^2) \tag{16}
\]

We solve this second-order equation to get the value of \( w \):

\[
w = \frac{-\beta^2 + \text{var}(\rho_{ij}) - \beta^2}{2} \tag{16}
\]

For \( \beta^2 = \rho = 10\% \) and \( \text{var}(\rho_{ij}) = 3\% \) (which is a conservative value compared with that measured from observed data; see above), we have \( w = 0.44\% \); that is, the standard deviation of the parameter \( \beta \) is equal to 6.6\%. The ratio of the capital charge including model risk \( (w > 0) \) to the capital charge without model risk \( (w = 0) \) measures the increase in capital due to dispersion or uncertainty on correlations.

We plot this quantity in figure 3 as a function of the mean deviation of the correlation parameter \( (\sqrt{\sigma}) \); we show that the impact of the mean deviation of the correlation parameter is lower than 2\% for \( \sqrt{\sigma} = 6.6\% \) for both the normal and uniform assumptions. Additionally, as the curves are very close to each other, we conclude that the shape of the correlation distribution function is not a driver of the capital charge: this validates the choice of the normal law for distribution functions that we have made throughout this article.

Even with a much more conservative choice for the individual cell risk parameter \( \sigma = 200\% \), the impact of correlation dispersion on the bank’s capital charge would be around +5\%. Our conclusion is that correlation dispersion (measured by parameter \( w \)) is nowhere near as critical as the other parameters of operational risk models (average cell risk \( \sigma \), cell risk dispersion \( \rho \) and average correlation level \( \beta = \sqrt{\sigma} \)).

**Conclusion**

This article pioneers analytical models for computing banks’ operational risk capital charges and provides some new results on the correlation problem. These simplified models are fairly realistic as they
incorporate dispersion in individual cell risks and correlation levels. Our main result is that uniform correlation is a robust assumption for capital charge modelling. This result is important because it means that model risk associated with the value of correlations is not a major issue for capital measurement. The impact of the choice of the copula function and of the average correlation value is much more significant, albeit that calibration suffers because of the scarcity of observed data. At the end of the day, dependence appears to be a subjective choice that determines the diversification benefit at the bank level, and cell loss distribution functions remain the main driver of the capital charge. We emphasise that our approach can straightforwardly be extended to other cell loss distribution or copula functions (Student t for example) in the one-factor framework. The extensions of our analytical approach should, as a second step, focus on broad distribution functions for cell losses and a smaller number of cells. This is left for future research.

Vivien Brunel is head of risk and capital modelling at Société Générale in Paris. The author would like to thank two anonymous referees for their comments on this paper. Special thanks to Pavel Shevchenko for the very stimulating correspondence. This article reflects the author’s opinions and not necessarily those of her employers.

Email: vivien.brunel@socgen.com

REFERENCES

Aue F and M Kalkbrener, 2007 LDA at work: Deutsche Bank’s approach to quantifying operational risk Journal of Operational Risk 1, pages 49–93


Frachot A, T Roncalli and E Salomon, 2004 The correlation problem in operational risk Groupe de Recherche Opérationnelle, Crédit Agricole, France

Moscadelli M, 2004 The modelling of operational risk: experience with the analysis of the data collected by the Basel committee Technical Report 517, Banca d’Italia


Vasicek O, 2002 Loan portfolio value Risk December, pages 160–162

GUIDELINES FOR THE SUBMISSION OF TECHNICAL ARTICLES

Risk welcomes the submission of technical articles on topics relevant to our readership. Core areas include market and credit risk measurement and management, the pricing and hedging of derivatives and/or structured securities, and the theoretical modelling and empirical observation of markets and portfolios. This list is not exhaustive.

The most important publication criteria are originality, exclusivity and relevance – we attempt to strike a balance between these. Given that Risk technical articles are shorter than those in affiliated academic journals, clarity of expression is another yardstick for publication. Once received by the technical editor and his team, submissions are logged and checked against these criteria. Articles that fail to meet the criteria are rejected at this stage.

Articles are then sent without author details to one or more anonymous referees for peer review. Our referees are drawn from the research groups, risk management departments and trading desks of major financial institutions, in addition to academia. Many have already published articles in Risk. Authors should allow four to eight weeks for the refereeing process. Depending on the feedback from referees, the author may attempt to revise the manuscript. Based on this process, the technical editor makes a decision to reject or accept the submitted article. His decision is final.

Submissions should be sent to the technical team, preferably by email addressed to technical@incisivemedia.com. LATEX is the preferred format for submissions (your LATEX file should be accompanied by a PDF of the output from it), although Word documents are also acceptable. It is helpful for graphs and figures to be submitted as separate Excel, postscript or EPS files. Vector graphics are strongly preferred, but rasterised images with a minimum resolution of 300dpi are sometimes acceptable.

The maximum recommended length for articles is 3,500 words, with some allowance for charts and/or formulas. We expect all articles to contain references to previous literature. We reserve the right to cut accepted articles to satisfy production considerations.