Should the advanced measurement approach be replaced with the standardized measurement approach for operational risk?

Research Paper

Should the advanced measurement approach be replaced with the standardized measurement approach for operational risk?

Gareth W. Peters,[1] Pavel V. Shevchenko,[2] Bertrand Hassani[3] and Ariane Chapelle[4]

[1]Department of Statistical Sciences, University College London, Gower Street, London WC1E 6BT, UK; email: gareth.peters@ucl.ac.uk

[2]CSIRO, PO Box 52, North Ryde, NSW 1670, Australia; emails: Pavel.Shevchenko@csiro.au, pavel.v.shevchenko@gmail.com

[3]Université Paris 1 Panthéon-Sorbonne, CES UMR 8174, 106 boulevard de l'Hôpital, 75647 Paris, Cedex 13, France; email: bertrand.hassani@univ-paris1.fr

[4]Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK; email: a.chapelle@ucl.ac.uk

ABSTRACT

Recently, the Basel Committee for Banking Supervision proposed to replace all approaches, including the advanced measurement approach (AMA), to operational risk capital with a simple formula referred to as the standardized measurement approach (SMA). This paper discusses and studies the weaknesses and pitfalls of the SMA, such as instability, risk insensitivity, super-additivity and the implicit relationship between the SMA capital model and systemic risk in the banking sector. We also discuss issues with the closely related operational risk capital-at-risk (OpCar) Basel Committee-proposed model, which is the precursor to the SMA. In conclusion, we advocate to maintain the AMA internal model framework and suggest as an alternative a number of standardization recommendations that could be considered to unify the internal modeling of operational risk. The findings and views presented

in this paper have been discussed with and supported by many OpRisk practitioners and academics in Australia, Europe, the United Kingdom and the United States, and recently at the OpRisk Europe 2016 conference in London.

1 INTRODUCTION

Operational risk (OpRisk) management is the youngest of the three major risk branches, with the others being market and credit risks within financial institutions. The term OpRisk became more popular after the bankruptcy of Barings bank in 1995, when a rogue trader caused the collapse of a venerable institution by placing bets in the Asian markets and keeping these contracts out of sight of management. At the time, these losses could be classified neither as market nor as credit risks, and the term OpRisk started to be used in the industry to define situations where such losses could arise. It was quite some time before this definition was abandoned and a proper definition was established for OpRisk. In these early days, OpRisk had a negative definition, "any risk that is not market or credit risk", which was not very helpful to assess and manage OpRisk. Looking back at the history of risk management research, we observe that early academics found the same issue of classifying risk in general, as Crockford (1982) noted:

> Research into risk management immediately encounters some basic problems of definition. There is still no general agreement on where the boundaries of the subject lie, and a satisfactory definition of risk management is notoriously difficult to formulate.

One thing is for certain: as risk management started to grow as a discipline, regulation also began to get more complex in order to catch up with new tools and techniques. It is not a stretch to say that financial institutions have always been regulated one way or another, given the risk they bring to the financial system. Regulation was mostly on a country-by-country basis and very uneven, which allowed for arbitrages. As financial institutions became more globalized, the need for more symmetric regulation that could level the way institutions would be supervised and regulated mechanically increased worldwide.

As a consequence of such regulations, in some areas of risk management, such as market risk and credit risk, there has been a gradual convergence or standardization of best practice, which has been widely adopted by banks and financial institutions. In the area of OpRisk modeling and management, such convergence of best practice is still occurring. This is due to multiple factors, such as many different types of risk

processes being modeled within the OpRisk framework, different influences and loss experiences in the OpRisk categories in different banking jurisdictions and the very nature of OpRisk being a relatively immature risk category compared with market and credit risk.

The following question therefore arises: how can one begin to induce a standardization of OpRisk modeling and capital calculation under Pillar I of the current banking regulation accords from the Basel Committee for Banking Supervision (BCBS)? It is stated under these accords that the basic objective of the Basel Committee's work has been to close gaps in international supervisory coverage in pursuit of two basic principles: that no foreign banking establishment should escape supervision, and that supervision should be adequate. It is this second note that forms the context for the new proposed revisions to simplify OpRisk modeling approaches. These have been brought out as two consultative documents:

- the standardized measurement approach (SMA), which was proposed in the Basel Committee consultative document "Standardized measurement approach for operational risk", issued in March 2016 for comments by June 3, 2016 (Basel Committee on Banking Supervision 2016); and

- the closely related OpRisk capital-at-risk (OpCar) model, which was proposed in the Basel Committee consultative document "Operational risk: revisions to the simpler approaches", issued in October 2014 (Basel Committee on Banking Supervision 2014).

In Basel Committee on Banking Supervision (2014, p. 1), it is noted that "despite an increase in the number and severity of operational risk events during and after the financial crisis, capital requirements for operational risk have remained stable or even fallen for the standardised approaches". Consequently, it is reasonable to reconsider these measures of capital adequacy and to decide if they need further revision. This is exactly the process undertaken by the Basel Committee in preparing the revised proposed frameworks that are discussed in this paper. Before getting to the revised framework of the SMA capital calculation, it is useful to recall the current best practice in Basel regulations.

Many models have been suggested for modeling OpRisk under the Basel II regulatory framework (Basel Committee on Banking Supervision 2006). Fundamentally, two different approaches are considered: the top-down approach and the bottom-up approach. A top-down approach quantifies OpRisk without attempting to identify the events or causes of losses explicitly. It can include the risk indicator models, which rely on a number of OpRisk exposure indicators to track OpRisks, and the scenario analysis and stress-testing models, which are estimated based on what-if scenarios. A bottom-up approach quantifies OpRisk on a micro-level, being based on identified

internal events. It can include actuarial-type models (referred to as the loss distribution approach (LDA)) that model the frequency and severity of OpRisk losses.

Under the current regulatory framework for OpRisk (Basel Committee on Banking Supervision 2006), banks can use several methods to calculate OpRisk capital: the basic indicator approach (BIA), the standardized approach (TSA) and the advanced measurement approach (AMA). Detailed discussion of these approaches can be found in Cruz *et al* (2015, Chapter 1). In brief, under the BIA and TSA, the capital is calculated as simple functions of gross income (GI):

$$K_{\text{BIA}} = \alpha \frac{1}{n} \sum_{j=1}^{3} \max\{\text{GI}(j), 0\}, \quad n = \sum_{j=1}^{3} 1_{\{\text{GI}(j)>0\}}, \ \alpha = 0.15, \tag{1.1}$$

$$K_{\text{TSA}} = \frac{1}{3} \sum_{j=1}^{3} \max\left\{ \sum_{i=1}^{8} \beta_i \, \text{GI}_i(j), 0 \right\}, \tag{1.2}$$

where $1_{\{\cdot\}}$ is the standard indicator symbol, which equals 1 if the condition in $\{\cdot\}$ is true, and 0 otherwise. Here, $\text{GI}(j)$ is the annual gross income of a bank in year j; $\text{GI}_i(j)$ is the gross income of business line i in year j; and β_i are coefficients in the range [0.12–0.18], specified by the Basel Committee for eight business lines. These approaches have a very coarse level of model granularity and are generally considered simplistic top-down approaches. Some country-specific regulators have adopted slightly modified versions of the BIA and TSA.

Under the AMA, banks are allowed to use their own models to estimate the capital. A bank intending to use AMA should demonstrate the accuracy of the internal models within Basel II-specified risk cells (eight business lines by seven event types) relevant to the bank. This is a finer level of granularity, more appropriate for a detailed analysis of risk processes in the financial institution. Typically, at this level of granularity, the models are based on bottom-up approaches. The most widely used AMA is the LDA based on modeling the annual frequency N and severities $X_1, X_2, \ldots$ of OpRisk losses for a risk cell, so that the annual loss for a bank over the d risk cells is

$$Z = \sum_{j=1}^{d} \sum_{i=1}^{N_j} X_i^{(j)}. \tag{1.3}$$

Then, the regulatory capital is calculated as the 0.999 value-at-risk (VaR), which is the quantile of the distribution for the next year's annual loss Z:

$$K_{\text{LDA}} = \text{VaR}_q[Z] := \inf\{z \in \mathbb{R}: \ \Pr[Z > z] \leqslant 1 - q\}, \quad q = 0.999; \tag{1.4}$$

this can be reduced by expected loss covered through internal provisions. Typically, frequency and severities within a risk cell are assumed to be independent.

For around ten years, the space of OpRisk has been evolving under this model-based structure. A summary of the the Basel accords over this period of time (Basel II–Basel III) can be captured as follows:

- they ensure that capital allocation is more risk sensitive;

- they enhance disclosure requirements that allow market participants to assess the capital adequacy of an institution;

- they ensure that credit risk, OpRisk and market risk are quantified based on data and formal techniques;

- they attempt to align economic and regulatory capital more closely to reduce the scope for regulatory arbitrage.

While the final Basel accord has at large addressed the regulatory arbitrage issue, there are still areas where regulatory capital requirements will diverge from the economic capital.

However, it was observed recently in studies performed by the BCBS and several local banking regulators that the BIA and TSA do not correctly estimate the OpRisk capital, ie, GI as a proxy indicator for OpRisk exposure appeared to be not a good assumption. Also, it appeared that capital under the AMA is difficult to compare across banks, due to the wide range of practices adopted by different banks.

So, at this point, two options are available to further refine and standardize OpRisk modeling practices: (1) to refine the BIA and TSA and, more importantly, converge within internal modeling in the AMA framework, or (2) to remove all internal modeling and modeling practice in OpRisk in favor of an overly simplified "one size fits all" SMA model (sadly, this is the option that has been adopted by the current round of Basel Committee consultations (Basel Committee on Banking Supervision 2016) in Pillar 1).

This paper is structured as follows. Section 2 formally defines the Basel-proposed SMA. The subsequent sections involve a collection of summary results and comments for studies performed on the proposed SMA model. Capital instability and sensitivity are studied in Section 3. Section 4 discusses the reduction of risk responsivity and incentivized risk-taking. Discarding key sources of OpRisk data is discussed in Section 5. The possibility of super-additive capital under SMA is examined in Section 6. Section 7 summarizes the Basel Committee procedure for the estimation of OpCar model and underlying assumptions and discusses the issues with this approach. The paper then concludes with suggestions in Section 8 relating to maintaining the AMA internal model framework with standardization recommendations that could be considered in order to unify the internal modeling of OpRisk.

2   BASEL COMMITTEE-PROPOSED STANDARDIZED MEASUREMENT APPROACH

This section introduces the new simplifications that are being proposed by the Basel Committee for models in OpRisk, starting with a brief overview of how this process was initiated by the OpRisk capital-at-risk (OpCar) model proposed in Basel Committee on Banking Supervision (2014), and then finishing with the current version of this approach, known as the SMA, proposed in Basel Committee on Banking Supervision (2016).

We begin with a clarification comment on the first round of the proposal, which is important conceptually to clarify for practitioners. In Basel Committee on Banking Supervision (2014, p. 1), it is stated that:

> Despite an increase in the number and severity of operational risk events during and after the financial crisis, capital requirements for operational risk have remained stable or even fallen for the standardized approaches. This indicates that the existing set of simple approaches for operational risk – the BIA and TSA, including its variant the alternative standardized approach (ASA) – do not correctly estimate the operational risk capital requirements of a wide spectrum of banks.

We agree that, in general, there are many cases in which banks will be under-capitalized for large crisis events, such as that which hit in 2008. Therefore, with the benefit of hindsight, it is prudent to reconsider these simplified models and look for improvements and reformulations that can be achieved in the wake of new information after the 2008 crisis. In fact, we would argue this is sensible practice in model assessment and model criticism, after new information regarding model suitability.

As we observed, the BIA and TSA make very simplistic assumptions regarding capital. Namely, that the GI can be used as an adequate proxy indicator for OpRisk exposure and, further, that a bank's OpRisk exposure increases linearly in proportion to revenue. Basel Committee on Banking Supervision (2014, p. 1) also makes two relevant points that "the existing approaches do not take into account the fact that the relationship between the size and the operational risk of a bank does not remain constant or that operational risk exposure increases with a bank's size in a nonlinear fashion".

Further, neither the BIA nor TSA approaches have been recalibrated since 2004. We believe that this is a huge mistake, that models and calibrations should be tested regularly and that each piece of regulation should come with its revision plan. As can be seen from experience, the model assumption has typically turned out to be invalid in a dynamically changing non-stationary risk management environment.

The two main objectives of the OpCar model proposed in Basel Committee on Banking Supervision (2014) were stated to be the following:

(i) to refine the OpRisk proxy indicator by replacing GI with a superior indicator;

(ii) to improve calibration of the regulatory coefficients based on the results of the quantitative analysis.

To achieve this, the Basel Committee argued – on practical grounds – that the model developed should be sufficiently simple to be applied with "comparability of outcomes in the framework", and that it should be "simple enough to understand, not unduly burdensome to implement, should not have too many parameters for calculation by banks and it should not rely on banks' internal models". However, they also claimed that such a new approach should "exhibit enhanced risk sensitivity" relative to the GI-based frameworks.

Additionally, such a one-size-fits-all framework "should be calibrated according to the OpRisk profile of a large number of banks of different size and business models." We disagree with this motivation, as many banks in different jurisdictions and for different bank size and different bank practice may indeed come from different population level distributions. In other words, the OpCar approach assumes all banks have a common population distribution from which their loss experience is drawn, and that this will be universal, no matter what your business practice, business volume or jurisdiction of operation. Such an assumption may lead to a less risk-sensitive framework, with poorer insight into actual risk processes in a given bank than a properly designed model developed for a particular business volume, operating region and business practice.

The background and some further studies on the precursor OpCar framework, which was originally supposed to replace just the BIA and TSA methods, are provided in Section 7. We explain how the OpCar simplified framework was developed, discuss the fact that it is based on an LDA model and a regression structure, and demonstrate how this model was estimated and developed. Along the way, we provide some scientific criticism of several technical aspects of the estimation and approximations utilized. It is important to still consider such aspects, as this model is the precursor to the SMA formula. That is, a single LDA is assumed for a bank, and single-loss approximation (SLA) is used to estimate the 0.999 quantile of the annual loss. Four different severity distributions were fitted to the data from many banks, and Poisson distribution is assumed for the frequency. Then, a nonlinear regression is used to regress the obtained bank capital (across many banks) to different combinations of explanatory variables from bank books, to end up with the OpCar formula.

The currently proposed SMA for OpRisk capital in Basel Committee on Banking Supervision (2016) is the main subject of our paper. However, we note that it is based

on the OpCar formulation, which itself is nothing more than an LDA model applied in an overly simplified fashion at the institution top level.

In Basel Committee on Banking Supervision (2016), it was proposed that all existing BIAs, TSAs and AMAs would be replaced with the SMA, calculating OpRisk capital as a function of the so-called business indicator (BI) and loss component (LC). Specifically, denote $X_i(t)$ as the ith loss and $N(t)$ as the number of losses in year t. Then, the SMA capital K_{SMA} is defined as

$$
K_{\text{SMA}}(\text{BI}, \text{LC})
= \begin{cases}
\text{BIC} & \text{if Bucket 1,} \\
110 + (\text{BIC} - 110) \ln\left(\exp(1) - 1 + \dfrac{\text{LC}}{\text{BIC}}\right) & \text{if Buckets 2–5.}
\end{cases}
\tag{2.1}
$$

Here,

$$
\text{LC} = 7\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N(t)} X_i(t) + 7\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N(t)} X_i(t) 1_{\{X_i > 10\}}
$$
$$
+ 5\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N(t)} X_i(t) 1_{\{X_i(t) > 100\}}, \tag{2.2}
$$

where $T = 10$ years (or at least five years for banks that do not have ten years of good quality loss data in the transition period). Buckets and the business indicator component (BIC) are calculated as

$$
\text{BIC} = \begin{cases}
0.11 \times \text{BI} & \text{if BI} \leqslant 1000, \ \text{Bucket 1,} \\
110 + 0.15 \times (\text{BI} - 1000) & \text{if } 1000 < \text{BI} \leqslant 3000, \ \text{Bucket 2,} \\
410 + 0.19 \times (\text{BI} - 3000) & \text{if } 3000 < \text{BI} \leqslant 10\,000, \ \text{Bucket 3,} \\
1740 + 0.23 \times (\text{BI} - 10000) & \text{if } 10\,000 < \text{BI} \leqslant 30\,000, \ \text{Bucket 4,} \\
6340 + 0.29 \times (\text{BI} - 30000) & \text{if BI} > 30\,000, \ \text{Bucket 5.}
\end{cases}
\tag{2.3}
$$

BI is defined as a sum of three components: the interest, lease and dividend component; the services component; and the financial component. It is made up of almost the same profit and loss (P&L) items used for the calculation of GI but combined in a different way (for precise formula, see Basel Committee on Banking Supervision (2016)). All amounts in the above formulas are in € millions.

3 STANDARDIZED MEASUREMENT APPROACH INTRODUCES CAPITAL INSTABILITY

In our analysis, we observed that SMA fails to achieve the objective of capital stability. In this section, we consider several examples to illustrate this feature. In most cases, we show results for only lognormal severity; other distribution types considered in the OpCar model lead to similar or even more pronounced features.

3.1 Capital instability examples

Consider a simple representative model for a bank's annual OpRisk loss process, comprised of the aggregation of two generic loss processes: one high frequency with low severity loss amounts, and the other corresponding to low frequency and high severity loss amounts, given by Poisson–Gamma and Poisson–lognormal models, respectively. We set the BI constant to €2 billion at half way within the interval for Bucket 2 of the SMA. We kept the model parameters static over time and simulated a history of 1000 years of loss data for three differently sized banks (small, medium and large), using different parameter settings for the loss models to characterize such banks. For a simple analysis, we set a small bank corresponding to capital in the order of tens of millions of euros in average annual loss, a medium bank in the order of hundreds of millions of euros in average annual loss, and a large bank in the order of €1 billion in average annual loss. We then studied the variability that may arise in the capital under the SMA formulation, under the optimal scenario that models did not change, model parameters were not recalibrated and the business environment did not change significantly, in the sense that BI was kept constant. In this case, we observe the core variation that arises just from the loss history experience of banks of the three different sizes over time.

Our analysis shows that a given institution can experience the situation in which its capital more than doubles from one year to the next, without any changes to the parameters, the model or the BI structure (Figure 1). This also means that two banks with the same risk profile can produce SMA capital numbers differing by a factor of more than two.

In summary, the simulation takes the case of a BI fixed over time, and the loss model for the institution is fixed according to two independent loss processes given by Poisson(λ)– Gamma(α, β) and Poisson(λ)– lognormal(μ, σ). Here, Gamma(α, β) is the Gamma distribution of the loss severities, with mean $\alpha\beta$ and variance $\alpha\beta^2$; lognormal(μ, σ) is the lognormal distribution of severities with the mean of the log-severity equal to μ and the variance of the log-severity equal to σ^2.

The institution's total losses are set to be on average around 1000 per year, with 1% coming from the heavy-tailed loss process Poisson–lognormal component. We perform two case studies, one in which the shape parameter of the heavy-tailed loss

TABLE 1 Test case 1 versus test case 2.

	Test case 1		Test case 2	
Bank size	Mean annual loss (€ million)	Annual loss 99.9% VaR (€ million)	Mean annual loss (€ million)	Annual loss 99.9% VaR (€ million)
Small	15	260	21	772
Medium	136	1841	181	5457
Large	769	14 610	1101	41 975

Test case 1 corresponds to the risk process Poisson(10)–lognormal($\mu = \{10; 12; 14\}, \sigma = 2.5$) and Poisson(990)–Gamma($\alpha = 1, \beta = \{10^4; 10^5; 5 \times 10^5\}$). Test case 2 corresponds to Poisson(10)–lognormal($\mu = \{10; 12; 14\}, \sigma = 2.8$) and Poisson(990)–Gamma($\alpha = 1, \beta = \{10^4; 10^5; 5 \times 10^5\}$).

process component is $\sigma = 2.5$ and the other in which it is $\sigma = 2.8$. We summarize the settings for the two cases below in Table 1. The ideal situation that would indicate the SMA was not producing capital figures that were too volatile would be if each of parts (a) to (f) in Figure 1 were very closely constrained around 1. However, as we can see, the variability in capital from year to year in all size institutions can be very significant. Note that we used different sequences of independent random numbers to generate results for small, medium and large banks in a test case. Thus, caution should be exercised in interpreting the results of test case 1 (or test case 2) for the relative comparison of capital variability in different banks. At the same time, comparing test case 1 with test case 2, one can certainly observe that an increase in σ increases the capital variability.

3.2 Capital instability and BI when SMA matches AMA

As a second study of the SMA capital instability, we consider a loss process model Poisson(λ)– lognormal(μ, σ). Instead of fixing the BI to the midpoint of Bucket 2 of the SMA formulation, we numerically solve for the BI that would produce the SMA capital equal to the VaR for a Poisson–lognormal LDA model at the annual 99.9% quantile level, VaR$_{0.999}$.

In other words, we find the BI such that the LDA capital matches the SMA capital in the long term. This is achieved by solving the following nonlinear equation numerically via root search for the BI:

$$K_{\text{SMA}}(\text{BI}, \widetilde{\text{LC}}) = \text{VaR}_{0.999}, \tag{3.1}$$

where $\widetilde{\text{LC}}$ is the long-term average of the loss component (2.2) that can be calculated in the case of Poisson(λ) frequency as

$$\widetilde{\text{LC}} = \lambda \times (7\text{E}[X] + 7\text{E}[X \mid X > L] + 5\text{E}[X \mid X > H]). \tag{3.2}$$

FIGURE 1   Ratio of the SMA capital to the long-term average.
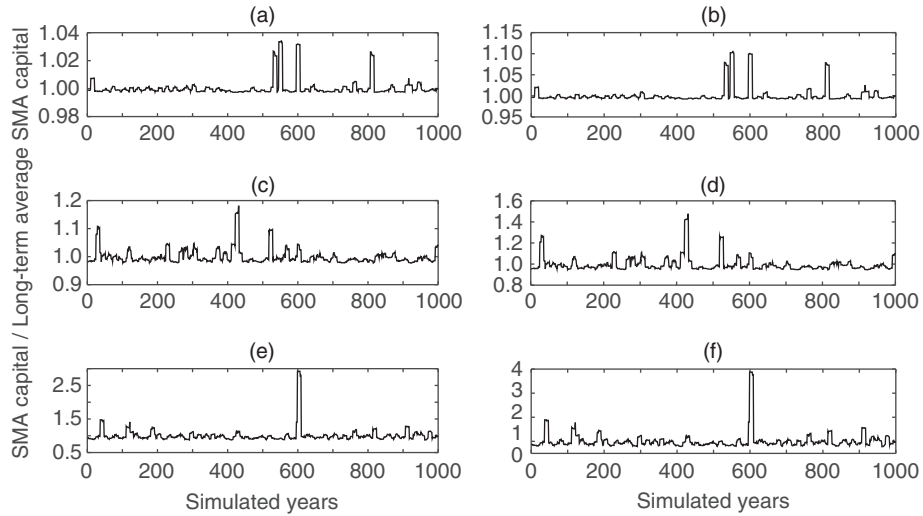
Test case 1 corresponds to $\sigma = 2.5$ (parts (a), (c) and (e)); Test case 2 corresponds to $\sigma = 2.8$ (parts (b), (d) and (f)). Other parameters are as specified in Table 1. Results for small (parts (a) and (b)), intermediate (parts (c) and (d)) and large (parts (e) and (f)) banks are based on different realizations of random variables in simulation.

In the case of severity X from lognormal(μ, σ), it can be found in closed form as

$$\widetilde{\mathrm{LC}}(\lambda, \mu, \sigma) = \lambda \mathrm{e}^{\mu + \frac{1}{2}\sigma^2} \left(7 + 7\Phi\left(\frac{\sigma^2 + \mu - \ln L}{\sigma} \right) + 5\Phi\left(\frac{\sigma^2 + \mu - \ln H}{\sigma} \right) \right), \tag{3.3}$$

where $\Phi(\cdot)$ denotes the standard normal distribution function, L is €10 million and H is €100 million, as specified by the SMA formula (2.2).

One can approximate the VaR$_{0.999}$ under the Poisson–lognormal model according to the so-called SLA (discussed in Section 7), given for $\alpha \uparrow 1$ by

$$\mathrm{VaR}_\alpha \approx \mathrm{SLA}(\alpha; \lambda, \mu, \sigma)$$

$$= \exp\left(\mu + \sigma \Phi^{-1}\left(1 - \frac{1 - \alpha}{\lambda} \right) \right) + \lambda \exp(\mu + \tfrac{1}{2}\sigma^2), \tag{3.4}$$

where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal distribution function. In this case, the results for the implied BI values are presented in Table 2 for $\lambda = 10$ and varied lognormal σ and μ parameters. Note that it is also not difficult to calculate VaR$_\alpha$ "exactly" (within numerical error) using Monte Carlo, Panjer recursion or fast Fourier transform (FFT) numerical methods.

TABLE 2 Implied BI in billions, $\lambda = 10$.

	σ						
μ	**1.5**	**1.75**	**2.0**	**2.25**	**2.5**	**2.75**	**3.0**
10	0.06	0.14	0.36	0.89	2.41	5.73	13.24
12	0.44	1.05	2.61	6.12	14.24	32.81	72.21
14	2.52	5.75	13.96	33.50	76.63	189.22	479.80

FIGURE 2 Ratio of the SMA capital to the long-term average.
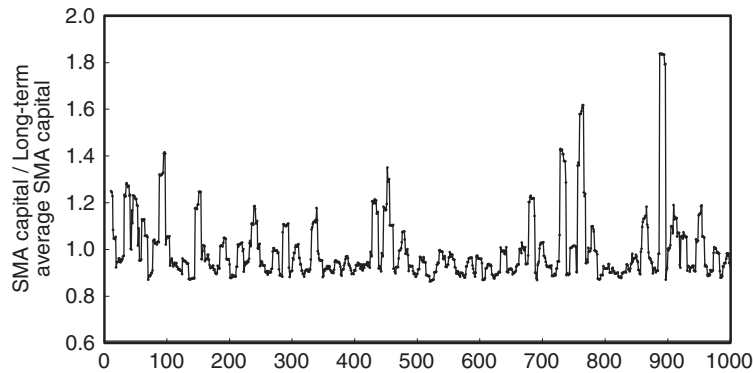
For a study of capital instability, we use the BI obtained from matching the long-term average SMA capital with the long-term LDA capital, as described above for an example generated by Poisson(10)–lognormal($\mu = 12, \sigma = 2.5$). We correspondingly found implied BI = €14.714 billion (Bucket 4). In this case, we calculate $VaR_{0.999}$ using Monte Carlo instead of an SLA (3.4); thus, the value of implied BI is slightly different from that in Table 2. In this case, the SMA capital based on the long-term average LC is €1.87 billion, which is about the same as $VaR_{0.999} = €1.87$ billion. The year-on-year variability in the capital with this combination of implied BI and Poisson–lognormal loss model is given in Figure 2. This shows that, again, we get capital instability with capital doubling from year to year compared with the long-term average SMA capital.

3.3 SMA is excessively sensitive to the dominant loss process

Consider an institution with a wide range of different types of OpRisk loss processes present in each of its business units and risk types. As in our first study above, we

FIGURE 3 Boxplot results for the ratio of the SMA capital to the long-term average for different values of the lognormal shape parameter σ, based on simulations over 1000 years.
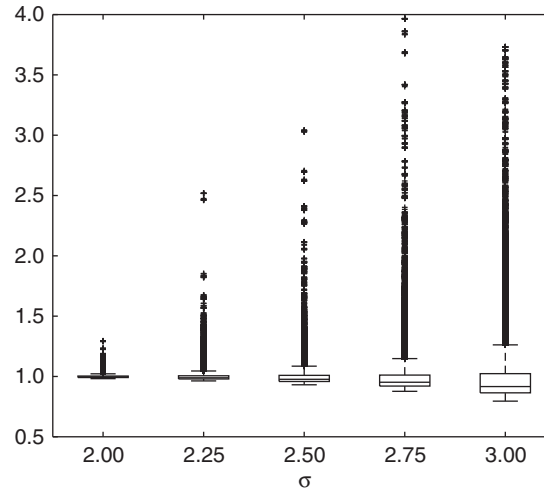
split these loss processes into two categories: high-frequency/low-severity types and low-frequency/high-severity types, given by $\text{Poisson}(990) - \text{Gamma}(1, 5 \times 10^5)$ and $\text{Poisson}(10) - \text{lognormal}(14, \sigma)$, respectively. In this study, we consider the sensitivity of SMA capital to the dominant loss process. More precisely, we study the sensitivity of SMA capital to the parameter σ, which dictates how heavy the tail of the most extreme loss process will be. Figure 3 shows boxplot results based on simulations performed over 1000 years for different values of $\sigma = \{2; 2.25; 2.5; 2.75; 3\}$.

These results can be interpreted to mean that banks with more extreme loss experiences as indicated by heavier-tailed dominant loss processes (increasing σ) tend to have significantly greater capital instability compared with banks with less extreme loss experiences. Importantly, these findings demonstrate how nonlinear this increase in SMA capital can be as the heaviness of the dominant loss process tail increases. For instance, banks with relatively low heavy-tailed dominant loss processes ($\sigma = 2$) tend to have a capital variability year on year of between 1.1 to 1.4 multipliers of long-term average SMA capital. However, banks with relatively large heavy-tailed dominant loss processes ($\sigma = 2.5, 2.75$ or 3) tend to have excessively unstable year-on-year capital figures, with variation in capital being as bad as three to four times multipliers of the long-term average SMA capital. Further, it is clear that when one considers each box plot as representing a population of banks with similar dominant

loss process characteristics, the population distribution of capital becomes increasingly skewed and demonstrates increasing kurtosis in the right tail as the tail heaviness of the dominant loss process in each population increases. This clearly demonstrates excessive variability in capital year on year for banks with heavy-tailed dominant loss processes.

Therefore, the SMA fails to achieve the claimed objective of robust capital estimation. Capital produced by the proposed SMA approach will be neither stable nor robust with worsening robustness as the severity of OpRisk increases. In other words, banks with higher severity OpRisk exposures will be substantially worse off under the SMA approach with regard to capital sensitivity.

4 REDUCED RISK RESPONSIVITY AND INDUCED RISK-TAKING

It is obvious that the SMA capital is less responsive to risk drivers and the variation in loss experience that is observed in a bank at granularity of the Basel II fifty-six business line/event type (BL/ET) units of measure.

This is due to the naive approach of modeling at the level of granularity assumed by the SMA, which only captures variability at the institution level, and not the intra-variability within the institution at business unit levels explicitly. Choosing to model at the institution level, rather than the units of measure or granularity of the fifty-six Basel categories, reduces model interpretability and reduces risk responsivity of the capital.

Conceptually, it relates to the simplification of the AMA under the SMA adopting a top-down formulation that reduces OpRisk modeling to a single unit of measure, as if all operational losses were following a single generating mechanism. This is equivalent to considering that earthquakes, cyber-attacks and human errors are all generated by the same drivers, and that they manifest in the loss model and loss history in the same manner as other losses that are much more frequent and have lower consequence, such as credit card fraud, when viewed from the institution level loss experience. It follows quite obviously that the radical simplification and aggregation of such heterogeneous risks in such a simplified model cannot claim the benefit of risk-sensitivity, even remotely.

Therefore, the SMA fails to achieve the claimed objective of capital risk sensitivity. Capital produced by the proposed SMA will be neither stable nor related to the risk profile of an institution. Moreover, the SMA induces risk-taking behaviors, and thus fails to achieve the Basel Committee objectives of stability and soundness of the financial institutions. Moral hazard and other unintended consequences include the following.

More risk-taking. Without the possibility of capital reduction for better risk management, in the face of increased funding costs due to the rise in capital, it is predictable

that financial institutions will raise their risk-taking to a level that is sufficient to pay for the increased cost of the new fixed capital. The risk appetite of a financial institution would mechanically increase. This effect goes against the Basel Committee objective of having a safe and secured financial system.

Denying loss events. While incident data collection involves a constant effort over a decade in every institution, large or small, the SMA is the most formidable disincentive to report losses. There are many opportunities to compress historical losses such as ignoring, slicing or transferring to other risk categories. The wish expressed in the Basel consultation that "banks should use ten years of good-quality loss data" is actually meaningless if the collection can be gamed. Besides, what about new banks, or BIA banks that do not currently have a loss data collection process?

Hazard of reduced provisioning activity. Provisions, which should be a substitution for capital, are vastly discouraged by the SMA, as they are penalized twice, counted both in the BI and the losses, and not accounted for as a capital reduction. The SMA captures both the expected loss and the unexpected loss, when the regulatory capital should only reflect the unexpected loss. We believe that this confusion might come from the use of the OpCar model as a benchmark, because the OpCar captures both equally. The SMA states in the definition of gross loss, net loss and recovery in Basel Committee on Banking Supervision (2016, Section 6.2, p. 10) under item (c) that the gross loss and net loss should include "provisions or reserves accounted for in the P&L against the potential operational loss impact". This clearly indicates the nature of the double counting of this component, since they will enter in the BI through the P&L and in the loss data component of the SMA capital.

Ambiguity in provisioning and resulting capital variability. The new guidelines on provisioning under the SMA framework follow a similar general concept to those that recently came into effect in credit risk with the International Financial Reporting Standard (IFRS 9). This was set forward by the International Accounting Standards Board (IASB), who completed the final element of its comprehensive response to the financial crisis with the publication of IFRS 9 Financial Instruments in July 2014. The IFRS 9 guidelines explicitly outline in Phase 2 an impairment framework, which specifies in a standardized manner how to deal with delayed recognition of (in this case) credit losses on loans (and other financial instruments). IFRS 9 achieves this through a new expected loss impairment model that will require more timely recognition of expected credit losses. Specifically, the new standard requires entities to account for expected credit losses from when financial instruments are first recognized, and it lowers the threshold for recognition of full lifetime

expected losses. However, the SMA OpRisk version of such a provisioning con-
cept for OpRisk losses fails to provide such a standardized and rigorous approach.
Instead, the SMA framework simply states that loss databases should now include:

> Losses stemming from operational risk events with a definitive financial impact,
> which are temporarily booked in transitory and/or suspense accounts and are not
> yet reflected in the P&L ("pending losses"). Material pending losses should be
> included in the SMA loss data set within a time period commensurate with the size
> and age of the pending item.

Unlike the more specific IFRS 9 accounting standards, under the SMA there is
a level of ambiguity. Further, this ambiguity can propagate now directly into the
SMA capital calculation, causing the potential for capital variability and instability.
For instance, there is no specific guidance or regulation requirement to standardize
the manner in which a financial institution decides what is to be considered a
"definitive financial impact" and what they should consider as a threshold for
deciding on existence of a "material pending loss". Also, the specific guidance
or rules about the time periods related to the inclusion of such pending losses in
an SMA loss data set, and, therefore, into the capital, are not stated. The current
guidance simply states that "material pending losses should be included in the SMA
loss data set within a time period commensurate with the size and age of the pending
item". This is too imprecise, and it may lead to the manipulation of provisions
reporting and categorization that will directly reduce SMA capital over the averaged
time periods in which the loss data component is considered. Further, if different
financial institutions adopt different provisioning rules, the capital obtained for
two banks with identical risk appetites and similar loss experiences could differ
substantially as a result of their provisioning practices.

Imprecise guidance on timing loss provisioning. The SMA guidelines also introduce
the topic of "timing loss provisioning", which they describe as follows:

> Negative economic impacts booked in a financial accounting period, due to oper-
> ational risk events impacting the cash flows or financial statements of previous
> financial accounting periods (timing losses). Material "timing losses" should be
> included in the SMA loss data set when they are due to operational risk events that
> span more than one financial accounting period and give rise to legal risk.

However, we would argue that for standardization of a framework there needs to be
more explicit guidance as to what constitutes a "material timing loss". Otherwise,
different timing loss provisioning approaches will result in different loss databases
and, consequently, differing SMA capital, just as a consequence of the provisioning
practice adopted. In addition, the ambiguity of this statement does not make it clear
whether such losses may be accounted for twice.

Grouping of losses. Under previous AMA internal modeling approaches, the unit of measurement or granularity of the loss modeling was reported according to the fifty-six BL/ET categories specified in the Basel II framework. However, under the SMA, the unit of measure is just at the institution level, so the granularity of the loss processes modeling and interpretation is lost. This has consequences when it is considered in light of the new SMA requirement that "losses caused by a common operational risk event or by related operational risk events over time must be grouped and entered into the SMA loss data set as a single loss." Previously, in internal modeling, losses within a given BL/ET would be recorded as a random number (frequency model) of individual independent loss amounts (severity model). Then, for instance, under an LDA model, such losses would be aggregated only as a compound process, and the individual losses would not be "grouped" except on the annual basis, and not on the per-event basis. However, there seems to be a marked difference in the SMA loss data reporting on this point. Under the SMA, it is proposed to aggregate the individual losses and report them in the loss database as a "single grouped" loss amount. This is not advisable from a modeling or an interpretation and practical risk management perspective. Further, the SMA guidance states that "the bank's internal loss data policy should establish criteria for deciding the circumstances, types of data and methodology for grouping data as appropriate for its business, risk management and SMA regulatory capital calculation needs." One could argue that if the aim of the SMA was to standardize OpRisk loss modeling in order to make capital less variable due to internal modeling decisions, then one can fail to see how this will be achieved with imprecise guidance, such as that provided above. One could argue that the above generic statement on criteria establishment basically removes the internal modeling framework of the AMA and replaces it with internal heuristic (non-model based, non-scientifically verifiable) rules to "group" data. This has the potential to result in even greater variability in capital than was experienced with non-standardized AMA internal models. At least under AMA internal modeling, in principle, the statistical models could be scientifically criticized.

Ignoring the future. All forward-looking aspects of risk identification, assessment and mitigation, such as scenarios and emerging risks, have disappeared in the new Basel consultation. This in effect introduces the risk of setting back the banking institutions in their progress toward a better understanding of threats; even though such threats may be increasing in frequency and severity, and the bank exposure to such threats may be increasing due to business practices, this cannot be reflected in the SMA framework capital. In that sense, the SMA is only backward looking.

5 STANDARD MEASUREMENT APPROACH FAILS TO UTILIZE RANGE OF DATA SOURCES OR PROVIDE RISK MANAGEMENT INSIGHT

As with any scientific discipline, OpRisk modeling is no different when it comes to developing a statistical modeling framework. In practical settings, it is therefore important to set the context with respect to the data and the regulatory requirements of Basel II when it comes to the data used in OpRisk modeling. In terms of the data aspect of OpRisk modeling, it has been an ongoing challenge for banks to develop suitable loss databases to record observed losses internally and externally, alongside other important information that may aid in modeling.

A key process in OpRisk modeling has not just been the collection itself, but, importantly, how and what to collect, as well as how to classify it. The first and key phase in any analytical process, certainly in the case of OpRisk models, is to cast the data into a form amenable to analysis. This is the very first task that an analyst faces when they set out to model, measure and even manage OpRisk. At this stage, there is a need to establish how the information available can be modeled to act as an input in the analytical process that would allow proper risk assessment processes to be developed. In risk management, and particularly in OpRisk, this activity is today quite regulated, and the entire data process, from collection to maintenance and use, has strict rules. In this sense, we see that qualitative and quantitative aspects of OpRisk management cannot be dissociated, as they act on one another in a causal manner.

Any OpRisk modeling framework starts by having solid risk taxonomy, so risks are properly classified. Firms also need to perform a comprehensive risk mapping across their processes to make sure that no risk is left out of the measurement process. This risk mapping is particularly important, as it directly affects the granularity of the modeling, the features observed in the data, the ability to interpret the loss model outputs and the ability to collect and model data. Further, it can affect the risk sensitivity of the models. It is a process that all large financial institutions have gone through at great cost of manpower and time in order to comply with Basel II regulations.

Under the Basel II regulations, there are four major data elements that should be used to measure and manage OpRisk: internal loss data, external loss data, scenario analysis, and business environment and internal control factors (BEICFs).

To ensure that data is correctly specified in an OpRisk modeling framework, one must undertake a risk mapping or taxonomy exercise, which basically encompasses the following: description, identification, nomenclature and classification. This is a very lengthy and time-consuming process that has typically been done by many banks at a fairly fine level of granularity with regard to the business structure. It involves going through, in excruciating detail, every major process of the firm. The outcome of this exercise would be the building block of any risk classification study. We also

observe that, in practice, this task is complicated by the fact that, in OpRisk settings, often when a risk materializes, and until it is closed, the loss process will continue to evolve over time – sometimes for many years, if we consider legal cases. In some cases, the same list of incidents taken at two different time points will not have the same distribution of loss magnitude. Here, it is important to bear in mind that a risk is not a loss: we may have risk and never experience an incident, and we may have incidents and never experience a loss. These considerations should also be taken into account when developing a classification or risk-mapping process.

There are roughly three ways that firms drive this risk taxonomy exercise: through cause, impact or events. The event-driven risk classification is probably the most common one used by large firms and has been the emerging best practice in OpRisk. This process classifies risk according to OpRisk events. This is the classification used by the Basel Committee, for which a detailed breakdown into event types at level 1, level 2 and activity groups is provided in Basel Committee on Banking Supervision (2006, pp. 305–307). Further, it is generally accepted that this classification has a definition broad enough to make it easier to accept/adopt changes in the process, should they arise. Besides, it is very interesting to note that a control taxonomy may impact the perception of events in the risk taxonomy, especially if the difference between inherent and residual risk is not perfectly understood. The residual risks are defined as inherent risk controls, ie, once we have controls in place, we manage the residual risks, while we may still be reporting inherent risks; this may bias the perception of the bank risk profile. The risk/control relationship (in terms of taxonomy) is not that easy to handle, as risk owners and control owners might be in completely different departments, preventing a smooth transmission of information. This we believe also needs further consideration in emerging best practice and governance implications for OpRisk management best practice.

5.1 The elements of the OpRisk framework

The four key elements that should be used in any OpRisk framework are internal loss data, external loss data, BEICFs and scenario analysis.

In terms of OpRisk losses, typically, the definition means a gross monetary loss or a net monetary loss, ie, a net of recoveries but excluding insurance or tax effects, resulting from an operational loss event. An operational loss includes all expenses associated with an operational loss event except for opportunity costs, forgone revenue and costs related to risk management and control enhancements implemented to prevent future operational losses. These losses need to be classified using the Basel categories (and internal categories, if these are different from Basel's) and mapped to a firm's business units. Basel II regulation says that firms need to collect at least five years of data, but most decide not to discard any loss, even when these are older than

this limit. Losses are difficult to acquire, and most even pay to supplement internal losses with external loss databases. Considerable challenges exist in collating a large volume of data, in different formats and from different geographical locations, into a central repository, as well as in ensuring that these data feeds are secure and can be backed up and replicated in case of an accident.

There is also a considerable challenge with OpRisk loss data recording and reporting related to the length for resolution of OpRisk losses. For some OpRisk events, usually the largest, there will be a large interval between the inception of the event and final closure, due to the complexity of these cases. As an example, most litigation cases that came up from the financial crisis in 2007–8 were only settled by 2012–13. These legal cases have their own life cycle and start with a discovery phase, in which lawyers and investigators argue if the other party has a proper case to actually take the action to court or not. At this stage, it is difficult to even come up with an estimate for eventual losses. Even when a case is accepted by the judge, it might be several years until lawyers and risk managers are able to properly estimate the losses.

Firms can set up reserves for these losses (and these reserves should be included in the loss database), but they usually only do that a few weeks before the case is settled in order to avoid disclosure issues (ie, the counterparty eventually knowing the amount reserved and using this information to their advantage). This creates an issue for setting up OpRisk capital: since firms would know that a large loss is coming, but they cannot yet include it in the database, the inclusion of this settlement would cause some volatility in the capital. The same would happen if a firm set a reserve of, for example, US$1 billion for a case and then a few months later a judge decided in the firm's favor, and this large loss had to be removed. For this reason, firms need to have a clear procedure on how to handle those large, long-duration losses.

The other issue with OpRisk loss reporting and recording is the aspect of adding costs to losses. As mentioned, an operational loss includes all expenses associated with an operational loss event except for opportunity costs, forgone revenue and costs related to risk management and control enhancements implemented to prevent future operational losses. Most firms, for example, do not have enough lawyers on payroll (or of the expertise) to deal with all of the cases, particularly some of the largest or those that demand some specific expertise and whose legal fees are quite high. There will be cases in which the firm wins in the end, maybe due to external law firms, but the cost can reach tens of millions of dollars. In this case, even with a court victory, there will be an operational loss. This leads to the consideration of provisioning of expected OpRisk losses, which is unlike credit risk, where the calculated expected credit losses might be covered by general and/or specific provisions in the balance sheet. For OpRisk, due to its multidimensional nature, the treatment of expected losses is more complex and restrictive. Recently, with the issuing of IAS 37 by the International Accounting Standards Board IFRS 2012, the rules have become clearer

as to what might be subject to provisions (or not). IAS 37 establishes three specific applications of these general requirements, namely that

- a provision should not be recognized for future operating losses,

- a provision should be recognized for an onerous contract – a contract in which the unavoidable costs of meeting its obligations exceeds the expected economic benefits,

- a provision for restructuring costs should be recognized only when an enterprise has a detailed formal plan for restructuring and has raised a valid expectation in those affected.

The last of these should exclude costs, such as retraining or relocating continuing staff and marketing or investment in new systems and distribution networks; the restructuring does not necessarily entail that. IAS 37 requires that provisions should be recognized in the balance sheet when, and only when, an enterprise has a present obligation (legal or constructive) as a result of a past event. The event must be likely to call upon the resources of the institution to settle the obligation, and it must be possible to form a reliable estimate of the amount of the obligation. Provisions in the balance sheet should be at the best estimate of the expenditure required to settle the present obligation at the balance sheet date. IAS 37 indicates also that the amount of the provision should not be reduced by gains from the expected disposal of assets, nor by expected reimbursements (arising from, for example, insurance contracts or indemnity clauses). When and if it is virtually certain that reimbursement will be received, should the enterprise settle the obligation, this reimbursement should be recognized as a separate asset.

We also note the following key points relating to regulation regarding provisioning, capital and expected loss (EL) components in "Detailed criteria 669" (Basel Committee on Banking Supervision 2006, p. 151). This portion of the regulation describes a series of quantitative standards that will apply to internally generated OpRisk measures for purposes of calculating the regulatory minimum capital charge.

(a) Any internal operational risk measurement system must be consistent with the scope of operational risk defined by the Committee in paragraph 644 and the loss event types defined in Annex 9.

(b) Supervisors will require the bank to calculate its regulatory capital requirement as the sum of EL and unexpected loss (UL), unless the bank can demonstrate that it is adequately capturing EL in its internal business practices. That is, to base the minimum regulatory capital requirement on UL alone, the bank must be able to demonstrate to the satisfaction of its national supervisor that it has measured and accounted for its EL exposure.

(c) A bank's risk measurement system must be sufficiently "granular" to capture the major drivers of operational risk affecting the shape of the tail of the loss estimates.

Here, note that if EL was accounted for, ie, provisioned, then it should not be covered by capital requirements again.

With regard to BEICF data, in order to understand the importance of BEICF data in OpRisk practice, we discuss this data source in the form of key risk indicators (KRIs), key performance indicators (KPIs) and key control indicators (KCIs).

A KRI is a metric of a risk factor. It provides information on the level of exposure to a given OpRisk of the organization at a particular point in time. KRIs are useful tools for business lines managers, senior management and boards to help monitor the level of risk-taking in an activity or organization, with regard to their risk appetite.

Performance indicators, usually referred to as KPIs, measure performance or the achievement of targets. Control effectiveness indicators, usually referred to as KCIs, are metrics that provide information on the extent to which a given control is meeting its intended objectives. Failed tests on key controls are natural examples of effective KCIs.

KPIs, KRIs and KCIs overlap in many instances, especially when they signal breaches of thresholds: a poor performance often becomes a source of risk. Poor technological performance, such as system downtime, for instance, becomes a KRI for errors and data integrity. KPIs of failed performance provide a good source of potential risk indicators. Failed KCIs are even more obvious candidates for preventive KRIs: a key control failure always constitutes a source of risk.

Indicators can be used by organizations as a means of control to track changes in their exposure to OpRisk. When selected appropriately, indicators ought to flag any change in the likelihood or the impact of a risk occurring. For financial institutions that calculate and hold OpRisk capital under more advanced approaches, such as the previous AMA internal model approaches, KPIs, KRIs and KCIs are advisable metrics to capture BEICF. While the definition of BEICF differs from one jurisdiction to another, and in many cases is specific to individual organizations, these factors must

- be risk sensitive (here, the notion of risk goes beyond incidents and losses),

- provide management with information on the risk profile of the organization,

- represent meaningful drivers of exposure that can be quantified,

- be used across the entire organization.

While some organizations include the outputs of their risk and control self-assessment programs under their internal definition of BEICFs, indicators are an

appropriate mechanism to satisfy these requirements, implying that there is an indirect regulatory requirement to implement and maintain an active indicator program (see the discussion in Chapelle (2013)).

For instance, incorporating BEICFs into OpRisk modeling is a reflection of the modeling assumption that one can see OpRisk as a function of the control environment. If the control environment is fair and under control, large operational losses are less likely to occur and OpRisk can be seen as under control. Therefore, understanding the firm's business processes, mapping the risks on these processes and assessing how the controls implemented behave is the fundamental role of the OpRisk manager. However, the SMA does not provide any real incentive mechanism, first for undertaking such a process, and second for incorporating this valuable information into the capital calculation.

5.2 SMA discards 75% of OpRisk data types

Both the Basel II and Basel III regulations emphasize the significance of incorporating a variety of loss data into OpRisk modeling and, therefore, ultimately into capital calculations. As has just been discussed, the four primary data sources to be included are internal loss data, external loss data, scenario analysis and BEICF. However, under the new SMA framework, only the first data source is utilized; the other three are now discarded.

Further, even if this decision to drop BEICFs were reversed in revisions to the SMA guidelines, we argue that this would not be easy to achieve. In terms of using pieces of information such as BEICFs and scenario data, because under the SMA framework the level of model granularity is only at the institution level, it does not easily lend itself to the incorporation of these key OpRisk data sources.

To business line managers, KRIs help to signal a change in the level of risk exposure associated with specific processes and activities. For quantitative modelers, KRIs are a way of including BEICFs in OpRisk capital. However, since BEICF data does not form a component of required data for the SMA model, there is no longer a regulatory requirement or incentive under the proposed SMA framework to make efforts to develop such BEICF data sources. This reduces the effectiveness of the risk models through the loss of a key source of information. In addition, the utility of such data for risk management practitioners and managers is reduced, as this data is no longer collected with the same required scrutiny, including validation, data integrity and maintenance and reporting, that was previously required for AMA internal models using such data.

These key sources of OpRisk data are not included in the SMA and cannot easily be incorporated into an SMA framework, even if there were a desire to do so due to the level of granularity implied by the SMA. This makes capital calculations less risk

sensitive. Further, the lack of scenario-based data incorporated into the SMA model makes it less forward looking and anticipatory as an internal model-based capital calculation framework.

6 THE STANDARDIZED MEASUREMENT APPROACH CAN BE A SUPER-ADDITIVE CAPITAL CALCULATION

The SMA seems to have the unfortunate feature that it may produce capital at a group level compared with the institutional level in a range of jurisdictions, which has the property that it is super-additive. It might be introduced by the regulator on purpose to encourage the splitting of very large institutions, though this is not stated in the Basel Committee documents explicitly. In this section, we show several examples of super-additivity and discuss its implications.

6.1 Examples of SMA super-additivity

Consider two banks with identical BI and LC. However, the first bank has only one entity, while the second has two entities. The two entities of the second bank have the same BI and the same LC, and those are equal to both half the BI and half the LC of the first joint bank.

In case one, Table 3(a), we consider the situation of a bucket shift, where the SMA capital obtained for the joint bank is €5771 million, while the sum of the SMA capital obtained for the two entities of the second bank is only €5387 million. In this example, the SMA does not capture a diversification benefit; on the contrary, it assumes that the global impact of an incident is larger than the sum of the parts. Here, the joint bank is in Bucket 5, while the entities appear in Bucket 4. In the second case, Table 3(b), we consider no bucket shift between the joint bank (Bank 1) and the two-entity bank (Bank 2). Bank 1 is in Bucket 5, and the entities of Bank 2 are in Bucket 5 too. In this case, we see that the joint bank has an SMA capital of €11 937 million, whereas the two-entity bank has an SMA capital of €10 674 million. Again there is a super-additive property.

Of course, in the examples in Table 3, we set BI and LC somewhat arbitrarily. So, in the next example, we use BI implied by the 0.999 VaR of LDA. In particular, assume a bank with a Poisson(λ)–lognormal(μ, σ) risk profile at the top level. Then, calculate the long-term average LC using (3.3) and the 0.999 VaR of the annual loss using (3.4), and find the implied BI by matching SMA capital with the 0.999 VaR. Now, consider the identical Bank 2 that splits into two similar independent entities that will have the same LC and the same BI, both equal to half of the LC and half of the BI of the first bank, which allows us to calculate SMA capital for each entity. Also note that, in this case, the entities will have risk profiles Poisson($\frac{1}{2}\lambda$)–lognormal(μ, σ) each.

TABLE 3 Super-additivity examples (all amounts are in € million).

	(a) Bucket shift		
	Bank 1	**Bank 2**	
Component	**Group**	**Entity 1**	**Entity 2**
BI	32 000	16 000	16 000
BIC	6920	3120	3120
LC	4000	2000	2000
SMA	5771	2694	2694
		Sum of SMAs: 5387	

	(b) No bucket shift		
	Bank 1	**Bank 2**	
Component	**Group**	**Entity 1**	**Entity 2**
BI	70 000	35 000	35 000
BIC	17 940	7790	7790
LC	4000	2000	2000
SMA	11 937	5337	5337
		Sum of SMAs: 10 674	

(a) Bank 1 is in Bucket 5 and the entities of Bank 2 are in Bucket 4. (b) Bank 1 and the entities of Bank 2 are in Bucket 5.

REMARK 6.1   The sum of K independent compound processes Poisson(λ_i) with severity $F_i(x)$, $i = 1, \ldots, K$, is a compound process Poisson(λ), with $\lambda = \lambda_1 + \cdots + \lambda_K$ and severity

$$F(x) = \frac{\lambda_1}{\lambda} F_1(x) + \cdots + \frac{\lambda_K}{\lambda} F_K(x);$$

(see, for example, Shevchenko 2011, Section 7.2, Proposition 7.1).

The results in the case of $\lambda = 10$, $\mu = 14$, $\sigma = 2$ are shown in Table 4. Here, the SMA for Bank 1 is €2.13 billion, while the sum of SMAs of the entities of Bank 2 is €1.96 billion, demonstrating the sub-additivity feature. Note that, in this case, one can also calculate the 0.999 VaR for each entity, which is €1.47 billion, while SMAs for Entity 1 and Entity 2 are €983 million each. That is, the entities with SMA capital are significantly undercapitalized compared with the LDA economic capital model; this subject will be discussed more in Section 6.2.

TABLE 4 Super-additivity example (all amounts are in € million).

| | Bank 1 | Bank 2 | |
	Group	Entity 1	Entity 2
λ	10	5	5
μ	14	14	14
σ	2	2	2
BI	13 960	6980	6980
BIC	2 651	1166	1166
LC	1 321	661	661
SMA	2 133	983	983
LDA	2 133	1473	1473

BI for Bank 1 is implied by the 0.999 VaR of Poisson(λ)–lognormal(μ, σ) risk profile (LDA).

Next, we state a mathematical expression that a bank could utilize in business structure planning to decide, in the long term, if it will be advantageous under the new SMA framework to split into two entities (or more) or remain in a given joint structure, according to the cost of funding Tier I SMA capital.

Consider the long-term SMA capital behavior averaged over the long-term history of the SMA capital for each case, both joint and disaggregated business models. Then, from the perspective of a long-term analysis regarding restructuring, the following expressions can be used to determine the point at which the SMA capital would be super-additive. If it is super-additive in the long term, this would indicate that there is therefore an advantage to splitting the institution in the long-run into disaggregated separate components. Further, the expression provided allows one to maximize the long-term SMA capital reduction that can be obtained under such a partitioning of the institution into m separate disaggregated entities.

PROPOSITION 6.2   (Conditions for super-additive SMA capital) *Under the LDA models with the frequency from* Poisson(λ_J) *and generic severity* $F_X(x; \theta_J)$, *the long-term average of the loss component* $\widetilde{\mathrm{LC}}$ *can be found using*

$$\widetilde{\mathrm{LC}} = \lambda \times (7\mathrm{E}[X] + 7\mathrm{E}[X \mid X > L] + 5\mathrm{E}[X \mid X > H]), \qquad (6.1)$$

ie, the short-term empirical loss averages in SMA formula (2.2) are replaced with the "long-term" average. We now denote the long-term average LC for a bank as $\widetilde{\mathrm{LC}}(\lambda_J, \theta_J)$ *and the long-term average LC for m bank entities after the split as* $\widetilde{\mathrm{LC}}(\lambda_i, \mu_i, \sigma_i)$, $i \in \{1, 2, \ldots, m\}$. *Therefore, the long-term SMA capital* $K_{\mathrm{SMA}}(\mathrm{BI}_J, \widetilde{\mathrm{LC}}(\lambda_J, \theta_J))$ *will be an explicit function of LDA model parameters* (λ_J, θ_J), *and the long-term SMA capital for the entity i is* $K_{\mathrm{SMA}}(\mathrm{BI}_i, \widetilde{\mathrm{LC}}(\lambda_i, \theta_i))$. *Hence, the*

SMA super-additive capital condition becomes

$$K_{\text{SMA}}(\text{BI}_{\text{J}}, \widetilde{\text{LC}}(\lambda_{\text{J}}, \boldsymbol{\theta}_{\text{J}})) - \sum_{i=1}^{m} K_{\text{SMA}}(\text{BI}_i, \widetilde{\text{LC}}(\lambda_i, \boldsymbol{\theta}_i)) > 0. \qquad (6.2)$$

The above condition is for a model-based restructuring, assuming each bank entity is modeled generically by an LDA model. Structuring around such a model-based assumption can be performed to determine optimal disaggregation of the institution to maximize capital cost reductions. Many severity distribution types will allow calculation of the long-term LC in closed form. For example, in the case of the Poisson–lognormal model, it is given by (3.3).

Of course, one can also use the above condition to perform maximization of the capital reduction over the next year by replacing $\widetilde{\text{LC}}$ with the observed LC calculated from empirical sample averages of historical data, as required in the SMA formula, avoiding explicit assumptions for severity and frequency distributions.

6.2 SMA super-additivity, macroprudential policy and systemic risk

In this section, we discuss the fact that the financial system is not invariant under observation, that is, banks and financial institutions will respond in a rational manner to maximize their competitive advantage. In particular, if new regulations allow and indeed provide incentives for banks to take opportunities to reduce, for instance, the cost of capital, they will generally act to do so. It is in this context that we introduce in brief the relationship between the new SMA capital calculations and the broader macroprudential view of the economy that the regulator holds.

It is generally acknowledged that the enhancement of the Basel II banking regulations by the additions that Basel III accords brought to the table were largely driven by a desire to impose greater macroprudential oversight on the banking system after the 2008 financial crisis. Indeed, the Basel III accords adopted an approach to financial regulation aimed at mitigating the "systemic risk" of the financial sector as a whole; we may adopt a generic high level definition of systemic risk as

> the disruption to the flow of financial services that is (i) caused by an impairment of all or parts of the financial system; and (ii) has the potential to have serious negative consequences for the real economy.

This view of systemic risk is focused on disruptions that arise from events such as the collapse of core banks or financial institutions in the banking sector, such as what happened after the Lehman Brothers collapse leading to the wider systemic risk problem of the 2008 financial crisis.

In response to reducing the likelihood of such a systemic risk occurrence, the Basel III regulation imposed several components that are of relevance to macroprudential financial regulation. Under Basel III, banks' capital requirements have been

strengthened, and new liquidity requirements, a leverage cap and a countercyclical capital buffer were introduced; these would remain in place under the new SMA guidelines.

In addition, large financial institutions, ie, the largest and most globally active banks, were required to hold a greater amount of capital, with an increased proportion of this Tier I capital being more liquid and of greater creditworthiness, ie, "higher-quality" capital. We note that this is consistent with a view of systemic risk reduction based on a cross-sectional approach. For instance, under this approach, the Basel III requirements sought to introduce systemic risk reduction macroprudential tools. These included the following:

(a) countercyclical capital requirements, which were introduced with the purpose of avoiding excessive balance-sheet shrinkage from banks in distress that may transition from going concerns to gone concerns;

(b) caps on leverage in order to reduce or limit asset growth through a mechanism that linked a banks' assets to their equity;

(c) time variation in reserve requirements with procyclical capital buffers as a means to control capital flows with prudential purposes.

In the United Kingdom, such Basel III macroprudential extensions are discussed in detail in the summary of the speech given at the 27th Annual Institute of International Banking Conference, Washington, by the Executive Director for Financial Stability Strategy and Risk in the Financial Policy Committee at the Bank of England (www .bankofengland.co.uk/publications/Documents/speeches/2016/speech887.pdf).

Further, one can argue that several factors can contribute to the systemic risk buildup in an economy, both locally and globally. One of these that is of relevance to discussions on AMA internal modeling versus SMA models is the risk that arises from the complexity of mathematical modeling being adopted in risk management and product structuring/pricing. One can argue from a statistical perspective that in order to scientifically understand the complex nature of OpRisk processes, and then respond to them with adequate capital measures and risk mitigation policies and governance structuring, it would be prudent to invest in some level of model complexity. However, with such complexity comes a significant chance of misuse of such models for gaming of the system to obtain competitive advantage via, for instance, achieving a reduction in capital. This can inherently act as an unseen trigger for systemic risk, especially if it is typically taking place in the larger, more substantial banks in the financial network, as is the case under AMA Basel II internal modeling. Therefore, we have this tension between reducing the systemic risk in the system due to model complexity and actually understanding the risk processes scientifically. We argue that the SMA

goes too far in simplifying the complexity of OpRisk modeling, rendering it unusable for risk analysis and interpretation. However, perhaps model complexity reduction could instead be reduced through AMA standardization of internal modeling practice, something we will discuss in the conclusions of this paper.

In this section, we ask what role the SMA model framework can play in the context of macroprudential systemic risk reduction. To address this question, we adopt the SMA's highly stylized view, as follows. We first consider the largest banks and financial institutions in the world. These entities are global and key nodes in the financial network; sometimes they have been referred to as "too big to fail" institutions. It is clear that the existence of such large financial institutions has both positive and negative economic effects. However, from a systemic risk perspective, they can pose problems for banking regulations both in local jurisdictions and globally.

There is, in general, an incentive to reduce the number of such dominant nodes in the banking financial network when viewed from the perspective of reducing systemic risk. So, the natural question that arises with regard to the SMA formulation is this: does the new regulation incentivize disaggregation of large financial institutions and banks, at least from the high-level perspective of the reduction of costs associated with obtaining, funding and maintaining Tier I capital and liquidity ratios, which is required under Basel III at present? In addition, if a super-additive capital is possible, is it achievable for feasible and practically sensible disaggregated entities? Finally, one could ask the following: does this super-additive SMA capital feature provide an increasing reward in terms of capital reduction as the size of the institution increases? We address these questions in the following stylized case studies, which illustrate that, in fact, the SMA can be considered as a framework that will induce systemic risk reductions from the perspective of providing potential incentives to reduce capital costs through disaggregation of large financial institutions and banks in the global banking network. However, we also observe that this may lead to significant under-capitalization of the entities after disaggregation; for example, in the case already discussed in Table 4 and considered in the next section (Figure 5).

6.3 SMA super-additivity is feasible and produces viable BI

For illustration, assume the joint institution is simply modeled by a Poisson–lognormal model Poisson(λ_J)– lognormal(μ_J, σ_J), with parameters sub-indexed by J for the joint institution and a BI for the joint institution denoted by BI$_J$. Further, we assume that if the institution had split into $m = 2$ separate entities for Tier I capital reporting purposes, then each would have its own stylized annual loss modeled by two independent Poisson–lognormal models: Entity 1, modeled by Poisson(λ_1)– lognormal(μ_1, σ_1), and Entity 2, modeled by Poisson(λ_2)– lognormal(μ_2, σ_2), with BI$_1$ and BI$_2$, respectively. Here, we assume that the disaggregation of the joint institution can occur in

such a manner that the risk profile of each individual entity may adopt more, less or equal risk aversion, governance and risk management practices. This means that there are really no restrictions on the parameters λ_1, μ_1 and σ_1, nor on the parameters λ_2, μ_2 and σ_2 from the perspective of λ_J, μ_J and σ_J.

In this sense, we study the range of parameters and BI values that will provide an incentive for large institutions to reduce systemic risk by undergoing disaggregation into smaller institutions. We achieve this through consideration of the SMA super-additivity condition in Proposition 6.2. In this case, it leads us to consider

$$K_{\text{SMA}}(\text{BI}_J, \widetilde{\text{LC}}(\lambda_J, \mu_J, \sigma_J)) - \sum_{i=1}^{2} K_{\text{SMA}}(\text{BI}_i, \widetilde{\text{LC}}(\lambda_i, \mu_i, \sigma_i)) > 0. \qquad (6.3)$$

Using this stylized condition, banks may be able to determine, for instance, if in the long term it would be economically efficient to split their institution into two or more separate entities. Further, they can use this expression to optimize the capital reduction for each of the individual entities, relative to the combined entities SMA capital. Hence, what we show here is the long-term average behavior, which will be the long-run optimal conditions for split or merge.

We perform a simple analysis below, where, at present, the joint institution is modeled by an LDA model, with frequency given by Poisson($\lambda_J = 10$) and severity given by lognormal($\mu_J = 12, \sigma_J = 2.5$), and with a BI implied by the 0.999 VaR of the LDA model, as detailed in Table 2, giving BI = €14.24 billion. We then assume the average number of losses in each institution if the bank splits into two is given by $\lambda_1 = 10$ and $\lambda_2 = 10$. In addition, the scale of the losses changes, but the tail severity of large losses is unchanged, such that $\sigma_1 = 2.5$ and $\sigma_2 = 2.5$, but μ_1 and μ_2 are unknown. We also calculate BI_1 and BI_2 implied by the LDA 0.999 VaR for the given values of μ_1 and μ_2. Then, we determine the set of values of μ_1 and μ_2, such that condition (6.3) is satisfied.

Table 5 shows the range of values for which μ_1 and μ_2 will produce super-additive capital structures and therefore justify disaggregation from the perspective of SMA capital cost minimization. We learn from this analysis that, indeed, it is plausible to structure a disaggregation of a large financial institution into a smaller set of financial institutions under the SMA capital structure. That is, the range of parameters μ_1 and μ_2 that produce super-additive capital structures under the SMA formulation setup are plausible, and the BI values are plausible for such a decomposition. In Table 5, we also show the BI values for Entity 1, implied in the cases satisfying the super-additive SMA capital condition.

We see from these results that the ranges of BI values that are inferred under the super-additive capital structures are also plausible ranges of values. This demonstrates that it is practically feasible for the SMA to produce incentive to reduce capital by disaggregation of larger institutions into smaller ones.

TABLE 5   Implied BI in €billions for Entity 1 (N/A indicates no super-additive solution).

	μ_2					
μ_1	**8**	**9**	**10**	**11**	**12**	**13**
8	0.301	0.301	0.301	0.301	N/A	N/A
9	0.820	0.820	0.820	0.820	N/A	N/A
10	2.406	2.406	2.406	2.406	N/A	N/A
11	5.970	5.970	5.970	5.970	N/A	N/A
12	N/A	N/A	N/A	N/A	N/A	N/A
13	N/A	N/A	N/A	N/A	N/A	N/A

6.4 Does SMA incentivize larger institutions to disaggregate more than smaller institutions?

Finally, we complete this section by addressing the following question: is there an increase in potential capital reductions through using super-additive SMA capital as a motivation to disaggregate into smaller firms as the size of the institution increases? As the tail of the bank loss distribution increases, the size of the institution as quantified through BI increases, and we are interested in seeing if there is an increase in incentive for larger banks to disaggregate to reduce capital charge.

To illustrate this point, we perform a study in which we use the following setup. We assume that we have a bank with a Poisson(λ)–lognormal(μ, σ) operational risk profile at the institution level. Now, we calculate $\widetilde{\text{LC}}$ and match the LDA VaR capital measure via the SLA at the 99.9% quantile to the SMA capital to imply the corresponding BI, BIC and SMA $= K_{\text{SMA}}(\text{BI}, \widetilde{\text{LC}})$ capital. Next, we consider disaggregating the institution into two similar independent components, which we denote as Entity 1 and Entity 2. This means the entities will have $\widetilde{\text{LC}}_1 = \widetilde{\text{LC}}_2 = \frac{1}{2}\widetilde{\text{LC}}$, $\text{BI}_1 = \text{BI}_2 = \frac{1}{2}\text{BI}$, $\lambda_1 = \lambda_2 = \frac{1}{2}\lambda$, $\mu_1 = \mu_2 = \mu$ and $\sigma_1 = \sigma_2 = \sigma$ (see Remark 6.1). Then, we calculate the SMA capitals $\text{SMA}_1 = K_{\text{SMA}}(\text{BI}_1, \widetilde{\text{LC}}_1)$ and $\text{SMA}_2 = K_{\text{SMA}}(\text{BI}_2, \widetilde{\text{LC}}_2)$ for Entity 1 and Entity 2, and we find the absolute super-additivity benefit from SMA capital reduction $\Delta = \text{SMA} - \text{SMA}_1 - \text{SMA}_2$ and relative benefit Δ/SMA. The results in the case $\lambda = 10$, $\mu = 14$ and varying σ are plotted in Figure 4. Note that the benefit is the non-monotonic function of σ, because in some cases the disaggregation process results in the entities shifting into a different SMA capital bucket compared with the original joint entity. One can also see that the absolute benefit from a bank disaggregation increases as the bank size increases, though the relative benefit drops. In the same figure, we also show the results for the case of bank disaggregation into ten similar independent entities, ie, $\lambda_1 = \cdots = \lambda_{10} = \lambda/10$, $\mu_1 = \cdots = \mu_{10} = \mu$ and $\sigma_1 = \cdots = \sigma_{10} = \sigma$.

FIGURE 4   Absolute (in € million) and relative super-additivity benefits from splitting a bank into two similar entities (and into ten similar entities) versus the lognormal severity shape parameter σ.
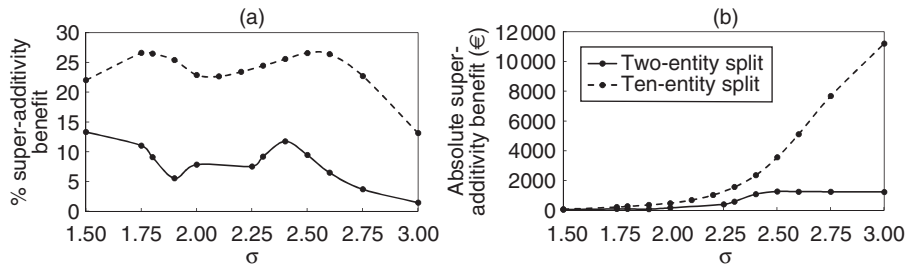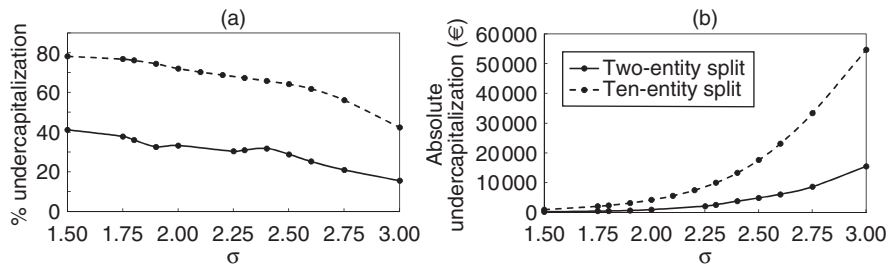

FIGURE 5   Absolute (in € million) and relative undercapitalization of the entities after bank disaggregation into two similar entities (and into ten similar entities) versus lognormal severity shape parameter σ.

We also calculate the 0.999 VaR of the Poisson–lognormal process, denoted as LDA_1 and LDA_2 for Entity 1 and Entity 2, respectively, using SLA (3.4). Then, we find undercapitalization of the entities $LDA_1 + LDA_2 - SMA_1 - SMA_1$, introduced by disaggregation, and corresponding relative undercapitalization, $(LDA_1 + LDA_2 - SMA_1 - SMA_1)/(LDA_1 + LDA_2)$. These results (and also for the case of bank disaggregation into ten similar entities) are shown in Figure 5. In this example, undercapitalization is very significant, and it increases for larger banks, although the relative undercapitalization gets smaller. Moreover, both the super-additivity benefit and undercapitalization features become more pronounced in the case of splitting into ten entities when compared with the two-entity split.

Thus, from one perspective, the capital calculated using the SMA formula encourages large banks to disaggregate (reducing the possibility of systemic risk from failure in the banking network); however, from another perspective, it introduces the significant undercapitalization of newly formed smaller entities, increasing their chances of failure (ie, increasing systemic risk in the banking system). In light of our analysis, it becomes clear that there is a downward pressure on banks to disaggregate, which reduces some aspects of systemic risk in the banking network. However, we also show that, at the same time, this very mechanism may undercapitalize the resulting smaller institutions, which would in turn increase systemic risk. The final outcome of the stability of the banking network will therefore depend largely on how aggressively larger banks choose to seek capital reductions at the risk of undercapitalizing their disaggregated entities, ie, their risk appetite in this regard will dictate the ultimate outcome. Such an uncertain future is surely not what the regulator had in mind in allowing for an incentive for disaggregation through the existence of super-additive capital measures.

7 OPCAR ESTIMATION FRAMEWORK

This section summarizes details of the OpCar model, which is the precursor to the SMA model and helped to form the SMA structure. First, we point out that the proposed OpCar model (Basel Committee on Banking Supervision 2014, Annex 2) is based on the LDA, albeit a very simplistic one that models the annual loss of the institution as a single LDA model formulation

$$Z = \sum_{i=1}^{N} X_i, \tag{7.1}$$

where N is the annual number of losses modeled as random variables from the Poisson distribution, Poisson(λ), ie, $\lambda = E[N]$, and X_i is the loss severity random variable from distribution $F_X(x; \boldsymbol{\theta})$, parameterized by vector $\boldsymbol{\theta}$. It is assumed that N and X_i are independent and $X_1, X_2, \ldots$ are independent too (note that modeling severities with autocorrelation is possible; see, for example, Guégan and Hassani (2013)). $F_X(x; \boldsymbol{\theta})$ is modeled by one of the following two-parameter distributions: Pareto, lognormal, log-logistic or log-gamma. Three variants of the Pareto model were considered, informally described in the regulation as corresponding to Pareto-light, Pareto-medium and Pareto-heavy. As a result, up to six estimates of the 0.999 VaR that were generated per bank were averaged to find the final capital estimate used in the regression model.

Next, we outline the OpCar fitting procedure and highlight potential issues.

7.1 Parameter estimation

The proposed OpCar model is estimated using data collected in a quantitative impact study (QIS) performed by the Basel Committee. Specifically, for each bank in the data set over $T = 5$ years corresponding to the 2005–9 period, the following data is used.

- $\tilde{n}_i$: the annual number of losses above $\tilde{u} = \text{€}10\,000$ in year $i \in \{1, \ldots, T\}$.

- n_i: the annual number of losses above $u = \text{€}20\,000$ in year $i \in \{1, \ldots, T\}$.

- S_i: the sum of losses above the level $u = \text{€}20\,000$ in year $i \in \{1, \ldots, T\}$.

- M_i: the maximum individual loss in year $i \in \{1, \ldots, T\}$.

The hybrid parameter estimation assumes that the frequency and severity distributions are unchanged over five years. Then, the following statistics are defined:

$$\lambda_u := \mathrm{E}[N \mid X \geqslant u] = \lambda(1 - F_X(u; \boldsymbol{\theta})), \tag{7.2}$$

$$\lambda_{\tilde{u}} := \mathrm{E}[N \mid X \geqslant \tilde{u}] = \lambda(1 - F_X(\tilde{u}; \boldsymbol{\theta})), \tag{7.3}$$

$$\mu_u := \mathrm{E}_{\boldsymbol{\theta}}[X \mid X \geqslant u], \tag{7.4}$$

which are estimated using observed frequencies n_i and $\tilde{n}_i$, and aggregated losses S_i, as

$$\hat{\lambda}_u = \frac{1}{T} \sum_{i=1}^{T} n_i, \qquad \hat{\lambda}_{\tilde{u}} = \frac{1}{T} \sum_{i=1}^{T} \tilde{n}_i, \qquad \hat{\mu}_u = \frac{\sum_{i=1}^{T} S_i}{\sum_{i=1}^{T} n_i}. \tag{7.5}$$

The conditional mean μ_u and severity distribution $F_X(\cdot; \boldsymbol{\theta})$ are known in closed form for the selected severity distribution types (see Basel Committee on Banking Supervision 2014, p. 26, Table A.5). Then, in the case of lognormal, log-gamma, log-logistic and Pareto-light distributions, the following two equations are solved to find severity parameter estimates $\hat{\boldsymbol{\theta}}$:

$$\frac{\hat{\lambda}_{\tilde{u}}}{\hat{\lambda}_u} = \frac{1 - F_X(\tilde{u}; \hat{\boldsymbol{\theta}})}{1 - F_X(u; \hat{\boldsymbol{\theta}})} \quad \text{and} \quad \hat{\mu}_u = \mathrm{E}_{\hat{\boldsymbol{\theta}}}[X \mid X \geqslant u], \tag{7.6}$$

which are referred to as the percentile and moment conditions, respectively. Finally, the Poisson λ parameter is estimated as

$$\hat{\lambda} = \frac{\hat{\lambda}_u}{1 - F_X(u; \hat{\boldsymbol{\theta}})}. \tag{7.7}$$

In the case of Pareto-heavy severity, the percentile condition in (7.6) is replaced by the "maximum heavy condition"

$$F_{X|X>\tilde{u}}(\hat{\mu}_M^{(1)}; \hat{\boldsymbol{\theta}}) = \frac{\tilde{n}}{\tilde{n} + 1}, \qquad \hat{\mu}_M^{(1)} = \max(M_1, \ldots, M_T); \tag{7.8}$$

in the case of Pareto-medium severity, the percentile condition is replaced by the "maximum medium condition"

$$F_{X|X>\tilde{u}}(\hat{\mu}_M^{(2)}; \hat{\boldsymbol{\theta}}) = \frac{\tilde{n}}{\tilde{n}+1}, \qquad \hat{\mu}_M^{(2)} = \frac{1}{T}\sum_{j=1}^{T} M_j. \qquad (7.9)$$

Here, $F_{X|X>\tilde{u}}(\cdot)$ is the distribution of losses conditional to exceed $\tilde{u}$. An explicit definition of $\tilde{n}$ is not provided in Basel Committee on Banking Supervision (2014, Annex 2), but it is reasonable to assume $\tilde{n} = (1/T)\sum_{i=1}^{T} \tilde{n}_i$ in (7.9) and $\tilde{n} = \sum_{i=1}^{T} \tilde{n}_i$ in (7.8).

These maximum conditions are based on the following result and approximation, stated in Basel Committee on Banking Supervision (2014). Denote the ordered loss sample $X_{1,n} \leqslant \cdots \leqslant X_{n,n}$, ie, $X_{n,n} = \max(X_1, \ldots, X_n)$. Using the fact that $F_X(X_i) = U_i$ is uniformly distributed, we have $\mathrm{E}[F_X(X_{k,n})] = k/(n+1)$ and, thus,

$$\mathrm{E}[F_X(X_{n,n})] = \frac{n}{n+1}. \qquad (7.10)$$

Therefore, when $n \to \infty$, one can expect that

$$\mathrm{E}[F_X(X_{n,n})] \approx F_X[\mathrm{E}(X_{n,n})], \qquad (7.11)$$

which gives conditions (7.8) and (7.9) when $\mathrm{E}(X_{n,n})$ is replaced by its estimators $\hat{\mu}_M^{(1)}$ and $\hat{\mu}_M^{(2)}$, and conditional distribution $F_{X|X>\tilde{u}}(\cdot)$ is used instead of $F_X(\cdot)$ to account for loss truncation below $\tilde{u}$. Here, we would like to note that, strictly speaking, under the OpRisk settings, n is random from Poisson distribution corresponding to the annual number of events that may have implications for the above maximum conditions. Also, note that the distribution of maximum loss in the case of Poisson-distributed n can be found in closed form (see Shevchenko 2011, Section 6.5).

We are not aware of results in the literature regarding the properties (such as accuracy, robustness and appropriateness) of estimators $\hat{\boldsymbol{\theta}}$ calculated in the above described way. Note that it is mentioned in Basel Committee on Banking Supervision (2014) that if the numerical solution for $\hat{\boldsymbol{\theta}}$ does not exist for a model, then this model is ignored.

The OpCar framework takes a five-year sample of data to perform sample estimation when fitting the models. We note that making an estimation of this type with only five years of data for the annual loss of the institution and fitting the model to this sample is going to result in very poor accuracy. This can then easily translate into non-robust results from the OpCar formulation if recalibration of the framework is performed in future.

At this point, we emphasize that a five-sample estimate is very inaccurate for these quantities; in other words, the estimated model parameters will have a very large

uncertainty associated with them. This is particularly problematic for parameters related to kurtosis and tail index in subexponential severity models. This is probably why the heuristic, practically motivated rejection criterion for model fitting was applied, in order to reject inappropriate fits (not in a statistical manner) that did not work due to the very small sample sizes used in this estimation.

To illustrate the extent of the uncertainty present in these five-year sample estimates, we provide the following basic case study. Consider a Poisson–lognormal LDA model with parameters $\lambda = 1000$, $\mu = 10$ and $\sigma = 2$ simulated over $m = 1000$ years. The $\text{VaR}_{0.999}$ for this model given by the SLA (3.4) is 4.59×10^8. Then, we calculate population statistics for each year n_i, $\tilde{n}_i$ and S_i, $i = 1, \ldots, m$, and form T-year non-overlapping blocks of these statistics from the simulated m years in order to perform estimation of distribution parameters for each block using percentile and moment conditions (7.6). Formally, for each block, we have to numerically solve two equations,

$$
\begin{aligned}
O_1 &:= \Phi\left(\frac{\hat{\mu} - \ln(\widetilde{u})}{\hat{\sigma}}\right)\left[\Phi\left(\frac{\hat{\mu} - \ln(u)}{\hat{\sigma}}\right)\right]^{-1} - \frac{\hat{\lambda}_{\tilde{u}}}{\hat{\lambda}_u} = 0, \\
O_2 &:= \exp\left(\hat{\mu} + \frac{\hat{\sigma}^2}{2}\right)\Phi\left(\frac{\hat{\mu} + \hat{\sigma}^2 - \ln(u)}{\hat{\sigma}}\right)\left[\Phi\left(\frac{\hat{\mu} - \ln(u)}{\hat{\sigma}}\right)\right]^{-1} - \hat{\mu}_u = 0,
\end{aligned}
$$

(7.12)

in order to find the severity parameter estimates $\hat{\mu}$ and $\hat{\sigma}$, which are then substituted into (7.7) to get the estimate $\hat{\lambda}$. Here, $\hat{\lambda}_{\tilde{u}}$, $\hat{\lambda}_u$ and $\hat{\mu}_u$ are the observed statistics (7.5) for a block.

This system of nonlinear equations may not have a unique solution, or a solution may not exist. Thus, to find an approximate solution, we consider two different objective functions.

- Objective function 1: a univariate objective function given by $O_1^2 + O_2^2$ that we minimize to find the solution.

- Objective function 2: a multi-objective function to obtain the Pareto optimal solution by finding the solution such that $|O_1|$ and $|O_2|$ cannot be jointly better off.

In both cases, a simple grid search over equally spaced values of $\hat{\mu}$ and $\hat{\sigma}$ was used to avoid any other complications that may have arisen with other optimization techniques. This leads to the most robust solution, which is not sensitive to gradients or starting points. A summary of the results for parameter estimates and the corresponding $\text{VaR}_{0.999}$ in the case of five-year blocks (ie, $[m/T] = 200$ independent blocks) is provided in Table 6.

TABLE 6 Mean of OpCar model parameter estimates over 200 independent five-year blocks if data is simulated from Poisson(1000)–lognormal(10,2).

Grid 1: $\hat{\mu} \in [8, 12]$, $\hat{\sigma} \in [1, 3]$, $\delta\hat{\mu} = 0.05$, $\delta\hat{\sigma} = 0.05$

Parameter estimate	Objective function 1	Objective function 2
$\hat{\mu}$	10.12(1.26)	9.35(1.03)
$\hat{\sigma}$	1.88(0.46)	2.16(0.28)
$\hat{\lambda}$	3248(5048)	959(290)
$\hat{\mathrm{VaR}}_{0.999}$	$10.6(13.4) \times 10^8$	$4.70(0.88) \times 10^8$

Grid 2: $\hat{\mu} \in [6, 14]$, $\hat{\sigma} \in [0.5, 3.5]$, $\delta\hat{\mu} = 0.05$, $\delta\hat{\sigma} = 0.05$

Parameter estimate	Objective function 1	Objective function 2
$\hat{\mu}$	9.79(1.94)	8.85(1.82)
$\hat{\sigma}$	1.88(0.71)	2.26(0.48)
$\hat{\lambda}$	$1.18(11.1) \times 10^8$	$1.5(21) \times 10^7$
$\hat{\mathrm{VaR}}_{0.999}$	$3.84(36.8) \times 10^{13}$	$4.34(61) \times 10^{12}$

The mean of sample statistics $\hat{\lambda}_u$, $\hat{\lambda}_{\tilde{u}}$ and $\hat{\mu}_u$ over the blocks used for parameter estimates are 654(11), 519(10) and $3.03(0.27) \times 10^5$, respectively, with the corresponding standard deviations provided in brackets next to the mean. The square root of the mean-squared error of the parameter estimator is provided in brackets next to the mean.

The results are presented for two search grids with different bounds but the same spacing between the grid points $\delta\hat{\mu} = 0.05$ and $\delta\hat{\sigma} = 0.05$. Here, $\mathrm{VaR}_{0.999}$ is calculated using SLA (3.4). These results clearly show that this estimation procedure is not appropriate. We note that, in some instances (for some blocks), both objective functions produce the same parameter estimates for common sample estimate inputs $\hat{\lambda}_{\tilde{u}}$, $\hat{\lambda}_u$ and $\hat{\mu}_u$; these estimates are close to the true values, although this is not systematically the case. We suspect that the reason for this is that, in some instances, there may be no solution in existence (or multiple solutions), and this then manifests in different solutions for the two objective functions (or inappropriate solutions). This is a serious concern for the accuracy of the findings from this approach; it suggests and is a key reason why other approaches to calibration at more granular levels (where more data is available) are often used – or, in cases where events are very rare, why alternative sources of data such as KRI, KPI, KCI and expert opinions should be incorporated into the calibration of the LDA model.

7.2 Capital estimation

The second stage of the OpCar framework is to take the fitted model parameters for the LDA model severity and frequency models at group or institution level and then calculate the capital. The approach to capital calculation under the OpCar analysis involves the so-called SLA. Here, we note that it would have been more accurate and easier to perform the capital calculations numerically, using methods such as Panjer recursion, FFT or Monte Carlo (see the detailed discussion of such approaches provided in Cruz *et al* (2015)).

Instead, the BCBS decided upon the following SLA to estimate the 0.999 quantile of the annual loss:

$$F_Z^{-1}(\alpha) \approx F_X^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right) + (\lambda - 1)\mathrm{E}[X], \qquad (7.13)$$

which is valid asymptotically for $\alpha \to 1$ in the case of subexponential severity distributions. Here, $F_X^{-1}(\cdot)$ is the inverse of the distribution function of the random variable X. It is important to point out that a correct SLA (in the case of finite mean subexponential severity and Poisson frequency) is actually given by a slightly different formula

$$F_Z^{-1}(\alpha) = F_X^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right) + \lambda\mathrm{E}[X] + o(1). \qquad (7.14)$$

This is a reasonable approximation, but it is important to note that its accuracy depends on distribution type and values of distribution parameters, and that further higher-order approximations are available (see the detailed discussion in Peters and Shevchenko (2015, Section 8.5.1) and the tutorial paper Peters *et al* (2013)).

To illustrate the accuracy of this first-order approximation for the annual loss VaR at a level of 99.9%, consider the Poisson(λ)–lognormal(μ, σ) model with parameters $\lambda = \{10, 100, 1000\}$, $\mu = 3$ and $\sigma = \{1, 2\}$; note that parameter μ is a scale parameter and will not affect relative differences in VaR. Then, we calculate the VaR of the annual loss from an LDA model for each of the possible sets of parameters, using a Monte Carlo simulation of 10^7 years that gives very good accuracy. We then evaluate the SLA approximation for each set of parameters. A summary of the findings is provided in Table 7.

The results indicate that though the SLA accuracy is good, the error can be significant for the parameters of interest for OpCar modeling. This can lead to a material impact on the accuracy of the parameter estimation in the subsequent regression undertaken by the OpCar approach and the resulting SMA formula.

TABLE 7 Accuracy of the SLA approximation of $VaR_{0.999}$ in the case of a Poisson(λ)–lognormal(μ, σ) model, with the scale parameter $\mu = 3$.

Parameters	MC VaR	SLA VaR	Δ VaR	$\epsilon(\%)$
$\lambda = 1000, \sigma = 1$	$3.88 \times 10^4 (0.02\%)$	3.54×10^4	-3.4×10^3	-8.7
$\lambda = 1000, \sigma = 2$	$4.24 \times 10^5 (0.28\%)$	4.19×10^5	-5.3×10^3	-1.3
$\lambda = 100, \sigma = 1$	$5.42 \times 10^3 (0.06\%)$	4.74×10^3	-6.8×10^2	-12.5
$\lambda = 100, \sigma = 2$	$1.17 \times 10^5 (0.4\%)$	1.17×10^5	-6.2×10^2	-0.52
$\lambda = 10, \sigma = 1$	$1.27 \times 10^3 (0.15\%)$	1.16×10^3	-1.1×10^2	-8.7
$\lambda = 10, \sigma = 2$	$3.57 \times 10^4 (0.52\%)$	3.56×10^4	-0.8×10^2	-0.23

Δ VaR is the difference between the SLA approximation and Monte Carlo estimate (the standard error of the Monte Carlo approximation is in brackets next to the estimate). $\epsilon(\%)$ is the relative difference between the SLA and Monte Carlo estimates.

7.3 Model selection and model averaging

The OpCar methodology attempts to fit six different severity models to the data, as described in the previous section, that generate up to six 0.999 VaR SLA estimates for each bank, depending on whether the models survived the imposed "filters". Then, the final estimate of the 0.999 VaR is found as an average of VaR estimates from the models that survived.

We would like to comment on the "filters" that were used to select the models to be used for a bank after the estimation procedures were completed. These heuristic ad hoc filters are not based on statistical theory and are determined by the following criteria:

- whether the proportion of losses above €20 000 is within a certain range (1–40%);

- whether the ratio between loss frequency and total assets is within a certain range (0.1–70 losses per billion euros of assets);

- whether the model estimation (outlined in Section 7.1) that was based on iterative solution converged.

We believe that in addition to practical considerations, other more rigorous statistical approaches to model selection could also be considered. For instance, Dutta and Perry (2006) discuss the importance of fitting distributions that are flexible but appropriate for the accurate modeling of OpRisk data. They focussed on the following five simple attributes in deciding on a suitable statistical model for the severity distribution.

- Good fit: statistically, how well does the model fit the data?

- Realistic: if a model fits well in a statistical sense, does it generate a loss distribution with a realistic capital estimate?

- Well specified: are the characteristics of the fitted data similar to the loss data and logically consistent?

- Flexible: how well is the model able to reasonably accommodate a wide variety of empirical loss data?

- Simple: is the model easy to apply in practice, and is it easy to generate random numbers for the purposes of loss simulation?

Further, in Cruz *et al* (2015, Chapter 8) there is a detailed description of appropriate model selection approaches that can be adopted in OpRisk modeling. These have been specifically developed for the setting that involves rigorous statistical methods and avoids heuristic methods.

The results of the heuristic filters applied in OpCar excluded a large number of distribution fits from the models for each institution. It was stated in Basel Committee on Banking Supervision (2014, p. 22) that the "OpCar calculator was run and validated on a sample of 121 out of 270 QIS banks which were able to provide data on operational risk losses of adequate quality" and "four of the distributions (lognormal, log-gamma, Pareto-medium and Pareto-heavy) were selected for the final OpCar calculation around 20% of the time or less".

7.4 OpCar regression analysis

Finally, we discuss aspects of regression in OpCar methodology based on the OpCar parameter and VaR estimation outlined in the previous section. Basically, for each bank's approximated capital, the regression is performed against a range of factors in a linear and nonlinear regression formulation. Given the potential for significant uncertainty in all aspects of the OpCar framework presented above, we argue that results from this remaining analysis may be spurious or biased, and we would recommend further study of this aspect. Further, these approximation errors will propagate into the regression parameter estimation in a nonlinear manner, making it difficult to directly determine the effect of such inaccuracies.

The regression models considered come in two forms: linear and nonlinear regressions. Given the sample of J banks on which to perform the regression, consider $(Y_j, X_{1,j}, \ldots, X_{20,j})$, $j = 1, \ldots, J$, where Y_j is the jth banks capital (dependent variable) obtained from the two-stage procedure described in Sections 7.2 and 7.1. Here, $X_{i,j}$ is the ith factor or covariate (independent variable) derived from the balance sheet and income statement of the jth bank. OpCar methodology considered twenty potential covariates. Only one observation of dependent variable Y_j per bank was used, and thus only one observation for each independent variable $X_{i,j}$ was needed and approximated by the average over QIS reporting years.

Then, the linear regression model considered in OpCar was specified by the model

$$Y_j = b_0 + \sum_{i=1}^{20} b_i X_{i,j} + \epsilon_j, \qquad (7.15)$$

with independent and identically distributed (iid) errors $\epsilon_1, \ldots, \epsilon_J$, which are from normal distribution with zero mean and the same variance.

In this particular, practical application, it is unlikely that the regression assumption of homoscedastic error variance is appropriate. We comment that the failure of this assumption, which is likely to be true, given the heterogeneity of the banks considered, would have caused spurious results in the regression analysis. Note that "outliers" were removed in the analysis, but no indication of what were considered to be outliers was mentioned. This again would have biased the results. We would have suggested not arbitrarily trying to force homoscedastic errors but instead considering weighted least squares if there was concern about outliers.

The regression analysis undertaken by the OpCar analysis and then utilized as a precursor to the SMA formulation implicitly assumes that all bank capital figures, used as responses (dependent variables) in the regression analysis against the factors (independent variables), such as BI for each bank, come from a common population. This is demonstrated by the fact that the regression is done across all banks jointly, and thereby a common regression error distribution type is assumed. This would probably not be the case in practice. We believe that other factors, such as banking volume, banking jurisdiction, banking practice and banking risk management governance structures, can have significant influences on this potential relationship. To some extent, the BI is supposed to capture aspects of this, but it cannot capture all of these aspects. Therefore, it may be the case that the regression assumptions about the error distribution, typically its having iid zero mean homoscedastic (constant) variance and normal distribution, would probably be questionable. Unfortunately, this would then directly affect all the analysis of significance of the regression relationships, the choice of the covariates to use in the model, etc.

It is very odd to think that, in the actual OpCar analysis, the linear regression model (7.15) was further reduced to the simplest type of linear regression model in the form of a simple linear model subfamily, given by

$$Y_j = b_0 + b_i X_{i,j} + \epsilon_j, \qquad (7.16)$$

ie, only simple linear models and not generalized linear model types were considered. This is unusual, as the presence of multiple factors and their interactions can often significantly improve the analysis, and it is simple to perform estimation in such cases. It is surprising to see this very limiting restriction in the analysis.

The second form of regression model considered in OpCar was nonlinear. It is given by the functional relationships

$$R(x) = xF(x),$$
$$\frac{\mathrm{d}}{\mathrm{d}x}R(x) = F(x) + xF'(x),$$
$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}R(x) = 2F'(x) + xF''(x), \tag{7.17}$$

where x is the proxy indicator, $R(x)$ represents the total OpRisk requirement (capital) and $F(x)$ is the functional coefficient relationship for any level x. It is assumed that $F(\cdot)$ is twice differentiable. The choice of function $F(x)$ selected is given by

$$F(x) = \theta\frac{(x-A)^{1-\alpha}}{1-\alpha}, \tag{7.18}$$

with $\alpha \in [0, 1]$, $\theta \geqslant 0$ and $A \leqslant 0$.

This model and the way it is described in the Basel consultative document is incomplete in the sense that it fails to adequately explain how multiple covariates were incorporated into the regression structure. It seems that, again, only single covariates are considered, one at a time; we again emphasize that this is a very limited and simplistic approach to performing such an analysis. There are standard software R packages that would have extended this analysis to multiple covariate regression structures, which we argue would have been much more appropriate.

Now, in the nonlinear regression model, if one takes $R(x_i) = Y_i$, ie, the ith bank's capital figure, this model could in principle be reinterpreted as a form of quantile regression model, such as those discussed recently in Cruz *et al* (2015). In this case, the response is the quantile function, and the function F would have been selected as a transform of a quantile error function, such as the class of Tukey transforms discussed in Peters *et al* (2016).

The choice of function $F(x)$ adopted by the modelers was just a translated and scaled power quantile error function of the type discussed in Dong *et al* (2015, Equation 14). When interpreting the nonlinear model in the form of a quantile regression, it is documented in several places (see the discussion in Peters *et al* (2016)) that least squares estimation is not a very good choice for parameter estimation for such models. Yet, this is the approach adopted in the OpCar framework.

Typically, when fitting quantile regression models, one would instead use a loss function corresponding to minimizing the expected loss of $Y - u$ with respect to u, according to

$$\min_u \mathrm{E}[\rho_\tau(Y-u)], \tag{7.19}$$

where $\rho_\tau(y)$ is given by

$$\rho_\tau(y) = y(\tau - 1_{\{y<0\}}). \tag{7.20}$$

Since the OpCar framework is only defined at the institutional level, this means that, effectively, the modeling of the regression framework cannot easily incorporate BEICFs such as KRI, KCI and KPI in a natural way. That is because these measures are typically recorded at a more granular level than the institution level. Instead, new OpRisk proxies and indicators were created under OpCar methods based on balance sheet outputs.

In fact, twenty proxy indicators were developed from the BCBS 2010 QIS data balance sheets and income statements of the participating banks selected. These included refinements related to GI, total assets, provisions, administrative costs and alternative types of factors. Unfortunately, the exact choice of twenty indicators used was not explicitly stated in detail in the OpCar document Basel Committee on Banking Supervision (2014, Annex 3); this makes it difficult to discuss the pros and cons of the choices made. Nor was there a careful analysis undertaken of whether such factors selected could have produced possible collinearity issues in the regression design matrix. For instance, if you combine one covariate or factor with another factor derived from this one, or one strongly related to it (as seems to be suggested in the cases with the GI-based factors), then the joint regression modeling with both factors will lead to an increased variance in parameter estimation and misguided conclusions about significance (statistical and practical) with regard to the factors selected in the regression model.

8 PROPOSITION: A STANDARDIZATION OF THE ADVANCED MEASUREMENT APPROACH

SMA cannot be considered as an alternative to AMA models. We suggest that the AMA is not discarded but instead improved by addressing its current weaknesses. It should be standardized! Details of how a rigorous and statistically robust standardization can start to be considered, with practical considerations, are suggested below.

Rather than discarding all OpRisk modeling, as allowed under the AMA, the regulator could instead make a proposal to standardize the approaches to modeling based on the accumulated knowledge to date of OpRisk modeling practices. We propose one class of models that can act in this manner and allow one to incorporate the key features offered by AMA LDA-type models, which involve internal data, external data, BEICFs and scenarios, with other important information on factors that the SMA method and OpCar approaches have tried to achieve but failed. As has already been noted, one issue with the SMA and OpCar approaches is that they try to model all OpRisk processes at the institution or group level with a single LDA model and

simplistic regression structure. This is bound to be problematic due to the very nature and heterogeneity of OpRisk loss processes. In addition, it fails to allow for the incorporation of many important OpRisk loss process explanatory information sources, such as BEICFs, which are often no longer informative or appropriate to incorporate at the institution level compared with the individual BL/ET level.

A standardization of the AMA internal models will remove the wide range of heterogeneity in model type. Here, our recommendation involves a bottom-up modeling approach, where for each BL/ET OpRisk loss process we model the severity and frequency components in an LDA structure. It can be comprised of a hybrid LDA model with factor regression components; these allow us to include the factors driving OpRisks in the financial industry at a sufficient level of granularity, while also utilizing a class of models known as the generalized additive models for location, shape and scale (GAMLSS) in the severity and frequency aspects of the LDA framework. The class of GAMLSS models can be specified to make sure that the severity and frequency families are comparable across institutions, allowing for both risk-sensitivity and capital comparability. We recommend in this regard the Poisson and generalized Gamma classes for the family of frequency and severity models, as these capture all typical ranges of loss models used in practice over the last fifteen years in OpRisk, including Gamma-, Weibull-, lognormal- and Pareto-type severities.

Standardizing recommendation 1. This leads us to the first standardizing recommendation relating to the level of granularity of modeling in OpRisk. The level of granularity of the modeling procedure is important to consider when incorporating different sources of OpRisk data, such as BEICFs and scenarios. This debate has been going on for the last ten years, with much discussion on bottom-up- versus top-down-based OpRisk modeling (see the overview in Cruz *et al* (2015) and Peters and Shevchenko (2015)). We advocate that a bottom-up-based approach be recommended as the standard modeling structure, as it will allow for greater understanding and more appropriate model development of the actual loss processes under study. Therefore, we argue that sticking with the fifty-six BL/ET structure of Basel II is best for a standardizing framework, with a standard aggregation procedure to institution level/group level. We argue that alternatives such as the SMA and OpCar approaches, which are trying to model multiple different featured loss processes combined into one loss process at the institution level, are bound to fail, as they need to capture high-frequency events as well as high-severity events. This, in principle, is very difficult if not impossible to capture with a single LDA model at the institution level, and it should be avoided. Further, such a bottom-up approach allows for greater model interpretation and incorporation of OpRisk loss data, such as BEICFs.

Standardizing recommendation 2. This brings us to our second recommendation for standardization in OpRisk modeling. Namely, we propose to standardize the modeling class to remove the wide range of heterogeneity in model type. We propose a standardization that involves a bottom-up modeling approach, where for each BL/ET level of the OpRisk loss process we model the severity and frequency components in an LDA structure that is comprised of a hybrid LDA model with factor regression components. The way to achieve this is to utilize a class of GAMLSS regression models for the severity and frequency model calibrations. That is, two GAMLSS regression models are developed, one for the severity fitting and the other for the frequency fitting. This family of models is flexible enough in our opinion to capture any type of frequency or severity model that may be observed in practice in OpRisk data, while incorporating factors such as BEICFs (KRIs, KPIs and KCIs) naturally into the regression structure. This produces a class of hybrid factor regression models in an OpRisk LDA family of models that can easily be fit, simulated from and utilized in OpRisk modeling to aggregate to the institution level. Further, as more years of data become available, the incorporation of time series structure in the severity and frequency aspects of each loss process modeling can be naturally incorporated into a GAMLSS regression LDA framework.

Standardizing recommendation 3. The class of models considered for the conditional response in the GAMLSS severity model can be standardized. There are several possible examples of such models that may be appropriate (Chavez-Demoulin *et al* 2015; Ganegoda and Evans 2013). However, we advocate for the severity models that the class of models be restricted in regulation to one family, the generalized Gamma family of models, where these models are developed in an LDA hybrid factor GAMLSS model. Such models are appropriate for OpRisk, as they admit special members that correspond to the lognormal, Pareto, Weibull and Gamma. All of these models are popular OpRisk severity models used in practice and represent the range of best practice by AMA banks, as observed in the recent survey by Basel Committee on Banking Supervision (2009). Since the generalized Gamma family contains all of these models as special sub-cases, this means that banks would only have to ever fit one class of severity model to each BL/ET LDA severity profile. Then, the most appropriate family member would be resolved in the fitting through the estimation of the shape and scale parameters in such a manner that, if a lognormal model was appropriate, it would be selected, whereas if a Gamma model were more appropriate, it would also be selected from one single fitting procedure. Further, the frequency model could be standardized as a Poisson GAMLSS regression structure, as the addition of explanatory covariates, and time varying and possible stochastic intensity allows for a flexible enough frequency model for all types of OpRisk loss processes.

Standardizing recommendation 4. The fitting of these models should be performed in a regression-based manner in the GAMLSS framework, which incorporates truncation and censoring in a penalized maximum likelihood framework (see Stasinopoulos and Rigby 2007). We believe that by standardizing the fitting procedure to one that is statistically rigorous, well understood in terms of the estimator properties and robust when incorporating a censored likelihood appropriately, we will remove the range of heuristic practices that has arisen in fitting models in OpRisk. The penalized regression framework, based on the L1 parameter penalty, will also allow for shrinkage methods to be used in order to select the most appropriate explanatory variables in the GAMLSS severity and frequency regression structures.

Standardizing recommendation 5. The standardization in form of Bayesian- versus Frequentist-type models should be left to the discretion of the bank, who can decide which version is best for their practice. However, we note that, under a Bayesian formulation, one can adequately incorporate multiple sources of information, including expert opinion and scenario-based data (see discussions in Cruz *et al* (2015), Peters *et al* (2009) and Shevchenko and Wüthrich (2006)).

Standardizing recommendation 6. The sets of BEICFs and factors to be incorporated into each BL/ET LDA factor regression model for severity and frequency should be specified by the regulator. There should be a core set of factors to be incorporated by all banks that include BEICFs and other factors to be selected. The following types of KRI categories can be considered in developing the core family of factors (see Chapelle 2013).

- Exposure indicators: any significant change in the nature of the business environment and its exposure to critical stakeholders or critical resources. Flag any change in the risk exposure.

- Stress indicators: any significant rise in the use of resources by the business, whether human or material. Flag any risk rising from overloaded humans or machines.

- Causal indicators: metrics capturing the drivers of key risks to the business. The core of preventive KRIs.

- Failure indicators: poor performance and failing controls are strong risk drivers. Failed KPIs and KCIs.

In this approach, a key difference is that instead of fixing the regression coefficients for all banks (as is the case for SMA and OpCar), pretending that all banks

have the same regression relationship as the entire banking population, one should standardize the class of factors. Specify explicitly how they should be collected as well as the frequency, and then specify that they should be incorporated in the GAMLSS regression. This will allow each bank to then calibrate the regression model to their loss experience through a rigorous penalized maximum likelihood procedure, with strict criteria on cross-validation-based testing on the amount of penalization admitted in the regression when shrinking factors out of the model. This approach has the advantage that banks will not only start to better incorporate the BEICF information into OpRisk models in a structured and statistically rigorous manner, but they will also be forced to better collect and consider such factors in a principled manner.

9 CONCLUSIONS

In this paper, we discussed and studied the weaknesses of the SMA formula for OpRisk capital recently proposed by the Basel Committee to replace the AMA and other current approaches. We also outlined the issues with the closely related OpCar model, which is the precursor of the SMA. There are significant potential problems with the use of the SMA, such as capital instability, risk insensitivity and capital super-additivity as well as serious concerns regarding the estimation of this model. We advocate standardization of the AMA rather than its complete removal, and we provide several recommendations based on our experience with OpRisk modeling.

DECLARATION OF INTEREST

ACKNOWLEDGEMENTS

REFERENCES

Basel Committee on Banking Supervision (2006). International convergence of capital measurement and capital standards: a revised framework, comprehensive version. Report, June, Bank for International Settlements.

Basel Committee on Banking Supervision (2009). Observed range of practice in key elements of Advanced Measurement Approaches (AMA). Report, July, Bank for International Settlements.

Basel Committee on Banking Supervision (2014). Operational risk: revisions to the simpler approaches. Report, October, Bank for International Settlements.

Basel Committee on Banking Supervision (2016). Standardized measurement approach for operational risk. Report, March, Bank for International Settlements.

Chapelle, A. (2013). The importance of preventive KRIs. *Operational Risk and Regulation* **58**(April).

Chavez-Demoulin, V., Embrechts, P., and Hofert, M. (2015). An extreme value approach for modeling operational risk losses depending on covariates. *Journal of Risk and Insurance*, forthcoming (http://doi.org/bm2q).

Crockford, G. N. (1982). The bibliography and history of risk management: some preliminary observations. *Geneva Papers on Risk and Insurance* **7**(23), 169–179 (http://doi.org/bppj).

Cruz, M., Peters, G., and Shevchenko, P.(2015). *Fundamental Aspects of Operational Risk and Insurance Analytics: A Handbook of Operational Risk*. Wiley (http://doi.org/bppk).

Dong, A. X., Chan, J. S., and Peters, G. W. (2015). Risk margin quantile function via parametric and non-paramteric Bayesian approaches. *Astin Bulletin* **45**(3), 503–550 (http://doi.org/bppm).

Dutta, K., and Perry, J. (2006). A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital. Working Paper 06-13, Federal Reserve Bank of Boston (http://doi.org/fzcjcw).

Ganegoda, A., and Evans, J. (2013). A scaling model for severity of operational losses using generalized additive models for location scale and shape (GAMLSS). *Annals of Actuarial Science* **7**(1), 61–100 (http://doi.org/bppn).

Guégan, D., and Hassani, B.(2013). Using a time series approach to correct serial correlation in operational risk capital calculation. *The Journal of Operational Risk* **8**(3), 31–56 (http://doi.org/bppq).

Peters, G., and Shevchenko, P. (2015). *Advances in Heavy Tailed Risk Modeling: A Handbook of Operational Risk*. Wiley (http://doi.org/bppr).

Peters, G., Shevchenko, P. V., and Wüthrich, M. V. (2009). Dynamic operational risk: modeling dependence and combining different sources of information. *The Journal of Operational Risk* **4**(2), 69–104 (http://doi.org/bprz).

Peters, G. W., Targino, R. S., and Shevchenko, P. V. (2013). Understanding operational risk capital approximations: first and second orders. *Journal of Governance and Regulation* **2**(3), 58–78.

Peters, G. W., Chen, W. Y., and Gerlach, R. H.(2016). Estimating quantile families of loss distributions for non-life insurance modelling via L-moments. *Risks* **4**(2) (http://doi.org/bmqv).

Shevchenko, P. V. (2011). *Modelling Operational Risk Using Bayesian Inference*. Springer (http://doi.org/bxsvwq).

Shevchenko, P. V., and Wüthrich, M. V. (2006). The structural modelling of operational risk via the Bayesian inference: combining loss data with expert opinions. *The Journal of Operational Risk* **1**(3), 3–26 (http://doi.org/bpqz).

Stasinopoulos, D. M., and Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software* **23**(7), 1–46 (http://doi.org/bpps).