



HAL
open science

Operational risk : A Basel II++ step before Basel III

Dominique Guegan, Bertrand Hassani

► **To cite this version:**

Dominique Guegan, Bertrand Hassani. Operational risk : A Basel II++ step before Basel III. 2011. halshs-00639484v1

HAL Id: halshs-00639484

<https://shs.hal.science/halshs-00639484v1>

Submitted on 9 Nov 2011 (v1), last revised 19 Feb 2013 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Documents de Travail du Centre d'Économie de la Sorbonne

C
E
S
W
o
r
k
i
n
g
P
a
p
e
r
s



Operational risk : A Basel II++ step before Basel III

Dominique GUEGAN, Bertrand K. HASSANI

2011.53



Operational risk: A Basel II++ step before Basel III

September 9, 2011

Authors

- Dominique Guégan: Université Paris 1 Panthéon-Sorbonne and CES UMR 8174, MSE 106 boulevard de l'Hopital 75647 Paris Cedex 13, France, e-mail: dguegan@univ-paris1.fr
- Bertrand K. Hassani: BPCE and CES-MSE, Université Paris 1 Panthéon-Sorbonne, 106 boulevard de l'Hopital 75647 Paris Cedex 13, France, phone: +33 (0) 158403397, e-mail: bertrand.hassani@bpce.fr

Abstract

Following Banking Committee on Banking Supervision, operational risk quantification is based on the Basel matrix which enables sorting incidents. In this paper we deeply analyze these incidents and propose strategies for carrying out the supervisory guidelines proposed by the regulators. The objectives are numerous.

- On the first hand, banks need to provide a univariate capital charge for each cell of the Basel matrix. That requires constructing Loss Distribution Function, which implies estimating a frequency and a severity distribution. We show that the choice of the theoretical distributions to build the LDFs has a tremendous impact on the capital charges, especially if we do not watch out for taking into account extreme losses.
- On the other hand, banks need also to provide a global capital charge corresponding to the whole matrix. We highlight the fact that a lack of consideration or a bad knowledge of the dependence structure may lead to wrong capital charges.
- Finally, we draw regulators and managers attention on two crucial points:
 1. the necessity of splitting information sets in two parts - the first covering small and medium losses, and the other one constituted of extreme ones - while estimating the severity distribution (this point induces granularity problem mentioned in the last Basel II guidelines),
 2. the choice of the risk measure which provides the capital amount. We underline that the expected shortfall measure enables a better anticipation of large incidents pertaining to operational risks.

1 Introduction

In 2001, the Basel Committee on Banking Supervision (BCBS) provided a set of principles (BCBS (2001)) for the effective management and supervision of operational risks, designed for banks and domestic authorities. According to these principles, banks could either use a basic, a standard or an advanced approach to calculate their capital charges. Carrying out the advanced measurement approach (AMA), banks developed internal systems.

However, at that time, operational risks management was a very young topic and due to its absence in Basel I accords (BCBS (1988)), the level of maturity was not the same as the one we had for credit risks. Therefore, Basel II initial requirements were very brief. They bound banks and financial institutions to use internal and external data, scenarios and qualitative criteria. They required banks to compute capital charges on a yearly basis and at a "99.9% confidence level". Therefore, at this time domestic authorities had no real experience of what would be a good approach to model the operational risks.

It is only after banks first proposals to comply regulatory requirements that the BCBS started to be more precise on the kind of models they were expecting (BCBS (2004; 2009; 2010)). New proposals have been done, they correspond to what we could refer as Basel II++ principles which has to be monitored inside the banks for the next few years. It is interesting to analyze the effects of these Basel II principles, to discuss the strategy before the deployment of the next Basel system and to anticipate the applications of these new principles with respect to the data on which these principles have been applied. It is the purpose of this article.

1.1 Initial principles in Basel II proposals

In the second part of the 80's when the regulators decided to establish rules to strengthen the banks, nothing concerned operational risks. It was only in 2001 that supervisors and the banking industry recognized the importance of operational risks in shaping the financial institutions risk profiles. Thus, the Basel Committee on Banking Supervision in their first Pillar about minimum regulatory capital charge for risks, required banks to take into account operational risks. There-

fore the Risk Management Group of the Basel Committee introduced a new regulatory capital approach based on banks' internal risk estimates called the "Advanced Measurement Approach" (AMA). The principle of this regulatory capital charge for operational risks was based on the following notion of operational risks: "risks of loss resulting from inadequate or failed internal processes, people and systems or from external events". This definition included legal risks, but not the strategic and reputational risks nor the systemic risks. The biggest challenge at this time to apply the recommended guidelines was to collect and to analyze loss data sets, because this implied efficient risk management devices and materials. This has been done in the past ten years. We will have the opportunity to discuss data collection and organization to get the minimum regulatory capital charge required from banks and insurance companies, in particular the problem of granularity.

The first Pillar proposal was established in close collaboration with the banking industry. It sorts operational risk exposures and losses into a series of standardized business lines and events types. We provided in Table 1 an example of the matrix on which the experts worked to provide the notorious Capital Requirement. The creation of this matrix could be responsible for errors in the calculation of the capital requirement due to the excessive aggregation of data collected. This is reflected in the guidelines for future Basel accords, and we will show in this paper the impact of the choice of the granularity of the matrix and the importance to take care of this point.

We discussed the data collection which is only use in the AMA to evaluate capital charges, and this lead us to step in another problem concerning the regulatory capital. The Committees's preliminary assessments of the possible level of operational risk regulatory capital was 20% of the current regulatory capital (CRM). Practitioners thought that a 20% figure should be lowered, and that 12% of minimum regulatory capital would provide a more reasonable cushion and produce required capital amounts more in line with the operational risks faced by large, complex banking organisations. However, it seems in reality that these first assesments (20% to 12%) were calibrated at a too low level. The first evaluation of operational risk capital requirements was used to define the Basic approach, and induced the weights of the standard approach. As banks incentive to adopt the AMA was a 25% capital allocation reduction compared to the standard approach, it naturally produced a modeling bias. One could answer that the capital reduction

BUSINESS UNITS	BUSINESS LINES LEVEL 1	(1) Internal Fraud	(2) External Fraud	(3) Employment Practices & Workplace Safety	(4) Clients, Products & Business Practices	(5) Damage to Physical Assets	(6) Business Disruption & System Failures	(7) Execution, Delivery & Process Management
INVESTMENT BANKING	(1) Corporate Finance (including Municipal/Gov.t Finance & merchant banking)							
	(2) Trading & Sales				F_1			
	(3) Retail Banking		F_8		F_2	F_5	F_9	
BANKING	(4) Commercial Banking							
	(5) Payment & Settlement		F_{10}		F_3		F_6	F_7
OTHERS	(6) Agency Services & Custody							
	(7) Asset Management							
	(8) Retail Brokerage				F_4			

Table 1: Basel Matrix for operational risks: each cell represents Business line/event types classification. $F_i, i \in [1, 10]$ denotes the distribution associated to each cell for which data is available.

is motivated by a better risk management, which is not entirely false, nevertheless the initial lack of experience led to too low forfaiary capital charges which were used as a benchmark for the AMA. Besides, and we would like to emphasize this point, loss data sets (internal or external) are not used in the basic and the standard approach to compute capital requirements. The regulatory capital in their case is evaluated as a fraction of the net banking income. Therefore, we cannot compare these approaches with the AMA. We show with empirical studies the difficulty to assess a correct level of charges. Therefore, after suggesting higher risk levels for the basic and the standard approach, we provide conservative and accurate solutions to model operational risks. The choice of methodologies is crucial and we detail those evidences at different steps of the calibration process.

1.2 The basic tools: LDA and VaR

A traditional way to deal with operational risks is the Loss Distribution Approach (LDA) (Frachot et al. (2001)). This one aims at building a Loss Distribution Function (LDF) (Guégan and Hassani (2009) for instance) G which is the mixture of a frequency distribution $p(k)$ and a severity distribution $F(x)$ (whose density is given by f), thus

$$G = \sum_{\gamma=1}^{\infty} p(k)F^{*\gamma}(x), \quad x > 0, \quad G_{b,e} = 0, \quad x = 0, \quad (1.1)$$

where $*$ denotes the convolution operator and $F^{*\gamma}$ the γ -fold convolution of F with itself. We denote g the density of G , and

$$g = \sum_{k=\gamma}^{\infty} p(k)f^{*\gamma}(x), \quad x > 0. \quad (1.2)$$

Experts working on operational risks agree on the fact that the frequency distribution has a negligible impact on the LDF's quantile which provides *in fine* the regulatory capital. However, the choice of the severity distribution has a tremendous impact on this one because the most pertinent information in terms of level of loss. The capital charge pertaining to G is obtained, at the request of the regulators, through a 99.9% Value-at-Risk (VaR) - a quantile computed from this LDF. We recall now the definition of the VaR measure,

Definition 1.1. *Given a confidence level $\alpha \in [0, 1]$, the VaR associated to a random variable X is given by the smallest number x such that the probability that X exceeds x is not larger than*

$(1 - \alpha)$

$$VaR_{(1-\alpha)\%} = \inf(x \in \mathbb{R} : P(X > x) \leq (1 - \alpha)). \quad (1.3)$$

Here, the random variable X is associated to a cell of the Basel matrix (Table 1). To compute $P(X > x)$ we need to determine the LDF associated with this random variable, thus the choice of the severity distribution which constitutes this LDF will have a tremendous impact on the final capital allocation. Figure 1 provides an example of this severity distribution.

Severity distribution	K-S test	VaR (95%)	ES (95%)	VaR (99.9%)	ES (99.9%)
Lognormal	<2.2 e-16	6 448 883	7 531 208	13 354 724	17 863 146
Weibull	<2.2 e-16	4 081 999	4 204 917	4 530 362	4 631 586
GPD	<2.2 e-16	33 512 000	192 902 521	1 009 472 708	4 865 819 191
Gumbel	<2.2 e-16	4 247 871	4 336 857	4 597 061	4 669 047
GB2	<2.2 e-16	9 321 772	19 497 566	76 326 374	253 422 675
g-and-h	<2.2 e-16	6 030 146	6 289 353	6 940 614	7 231 981

Table 2: Several theoretical distributions have been fitted on a data set representing the External Fraud events on the Retail Banking business line of the Banques Populaire perimeter. The table presents the goodness-of-fit tests results. These ones are very bad. In this case it is not possible to state which distribution is better than the others. The fifth column presents the capital charges we would have using these distributions, and the sixth the Expected Shortfall. These results are compared to 95% confidence level risk measures.

For instance to model the severities of a sample representing the External Fraud event on the Retail Banking business line¹, we have fitted and tested the following distributions (for which we provide the densities in appendix A): the lognormal distribution, the Weibull distribution, the Gumbel distribution, the Generalized Pareto distribution (GPD) (Pickands (1975), Embrechts et al. (1997)), the GB2 distribution (Gupta and Nadarajah (2004), Rigby and Stasinopoulos (2005)), and the g-and-h distribution (Hoaglin (1985), Buch-Kroman (2009)). For the choosen sample, we provide, for each distribution the corresponding capital requirement computed using the 95% and 99.9% Value-at-Risk (VaR) introduced in (1.3). The second column of Table 2

¹The results presented in this paper have been obtained using Caisse d'Epargne data sets (2006-2010) , except the results presented in Table 2, which have been obtained on a Banques Populaires data set (2008-2010).

provides the result of an adequation test based on Kolmogorov-Smirnov distance to check the adequation of these distributions to the data sets. They all failed, thus statistically we cannot accept any of these distributions. Nevertheless if we continue the exercise² we provide in columns 3 and 5 the capital computed using the VaR for two values of α . Columns 4 and 6 exhibit the capital requirements that the banks need if we use another measure named Expected Shortfall that we introduce below. Looking at these four columns we observe that the choice of the distribution has a tremendous impact on the capital allocation. Thus at this point, there is a trade-off between the choice of a distribution that minimizes the capital allocation (here the Gumbel one), or a distribution which provides a more conservative capital (here the GB2). We also observe large differences with respect to the choice of the so-called risk measure. Thus, besides the calibration problem, the choice of risk measure is crucial. Now, we come back in details to these points.

Therefore, in a first part - section 2 - we discuss estimation of the LDF for each cell of the matrix, and provide the corresponding regulatory capital with respect to the corresponding LDF and the chosen risk measure. In section 3, introducing a dependence structure between the cells and for the different risk measure, we do the same exercise. Finally a last section concludes with some important remarks focusing on the impact of the distribution families, the parameters estimation procedures and the risk measures on the regulatory capital.

2 Estimation of the severities in each cell: the importance of the tail distribution

Before providing a capital associated with the global matrix, we need to analyze the behavior of each cell. Here we need to know the correct distribution that characterizes each cell. It is now recognized and the example given in Table 2 confirms that for a proper statistical adjustment of the severity distribution, we must take into account the existence of extreme values in the analysis of the loss data sets. Different approaches can be considered, all are part of the extreme

²To continue this exercise is not unrealistic as soon as severity distributions like the lognormal distribution is commonly used in banks to compute the LDF

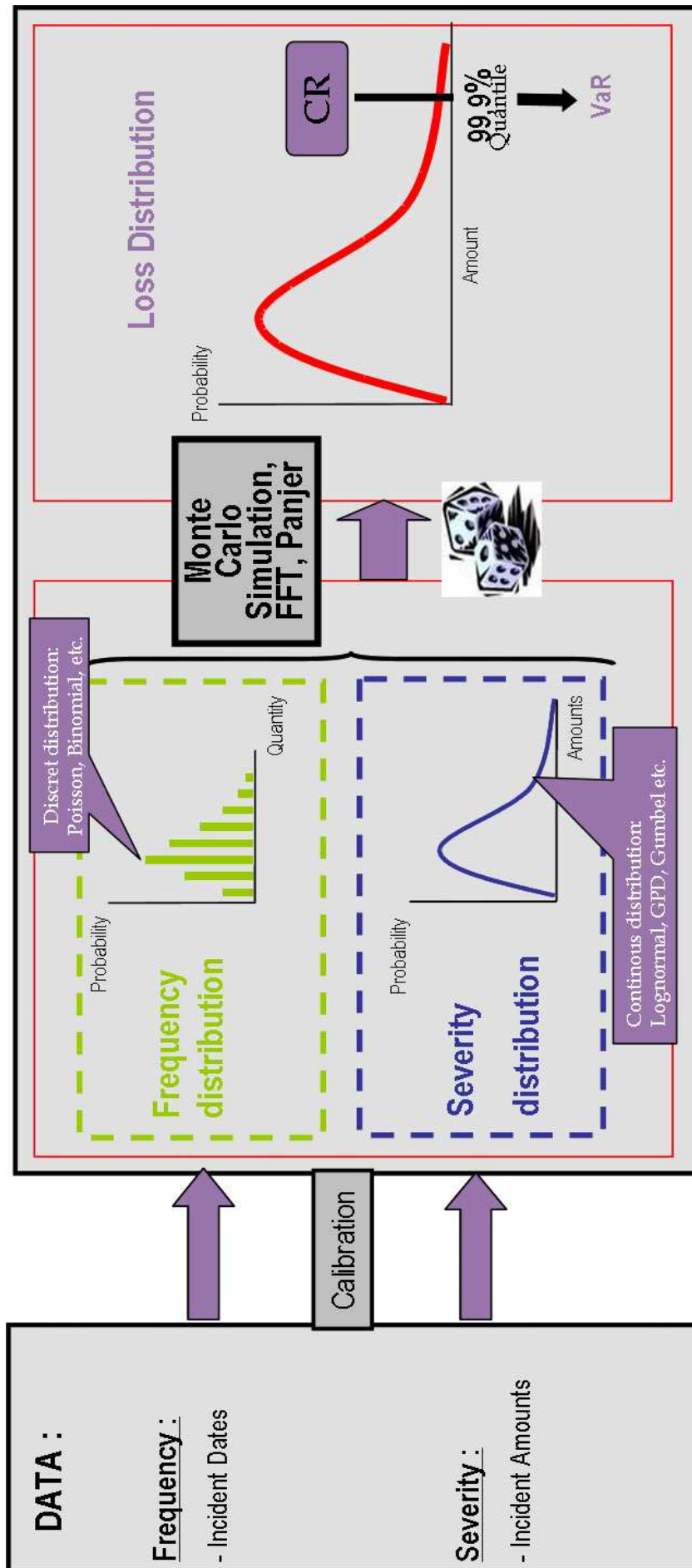


Figure 1: The Loss Distribution Approach.

value theory (Leadbetter and Rootzen (1988)). Following this theory in a recent paper, we have proposed a flexible approach (Guégan et al. (2010)) to model the severity distributions. We built it as a mix of two distributions: one characterizing the most important losses (the tail) using the Generalized Pareto Distribution (GPD), and another one modeling the remaining data (the central part of the distribution or corpus). We have used a GPD on the right tail for which we provided innovative theoretical and practical solutions, and fit a lognormal³ distribution on the remaining data via an EM algorithm. Figure 2 illustrates the approach. Then, to build the final LDF, we applied an adapted Monte Carlo algorithm. Second, we underlined the fact that once the threshold of the GPD has been found, the method chosen to estimate GPD's parameters tremendously impacts the VaR: it seems that the influence of the estimation procedure on the computation of the capital allocation has never been discussed before and we provide some information in the Appendix B. We follow here the same lines of the methodology developed in (Guégan et al. (2010)).

Thus, we define the severity distribution as a mixture of a lognormal distribution on the corpus, and a GPD on the right tail (cauda,) whose density $f(x; u, \beta, \xi)$ is :

$$f(x; u, \beta, \xi, \mu, \sigma) = \begin{cases} f^{(corpus)}(x; \mu, \sigma), & \text{if } x < u \\ f^{(cauda)} = \frac{1}{1 - \int_0^u f^{(center)}(x; \mu, \sigma) dx} \times f^{GPD}(x; u, \beta, \xi), & \text{if } x \geq u \end{cases}, \quad (2.1)$$

where, μ and σ are the lognormal distribution parameters, and f^{GPD} the density of the GPD given in appendix A.4. The importance of the estimation procedure for the parameters on the final capital is detailed in appendix B.

In Table 3, we provide the capital requirements computed using different distributions for the cell associated to the business line "Payment & Settlement" and the event type "Delivery, Execution and Process Management" for the year 2006. We observe that using only the lognormal distribution provides a very less conservative regulatory capital than using the POT method we introduced above. Note that when we adjust only a GPD on the data set, we do not obtain workable results because the parameters cannot be properly estimated, or we face an infinite mean model ($\xi > 1$).

³The lognormal assumption was the best on our data sets, but one can fit any other distributions.

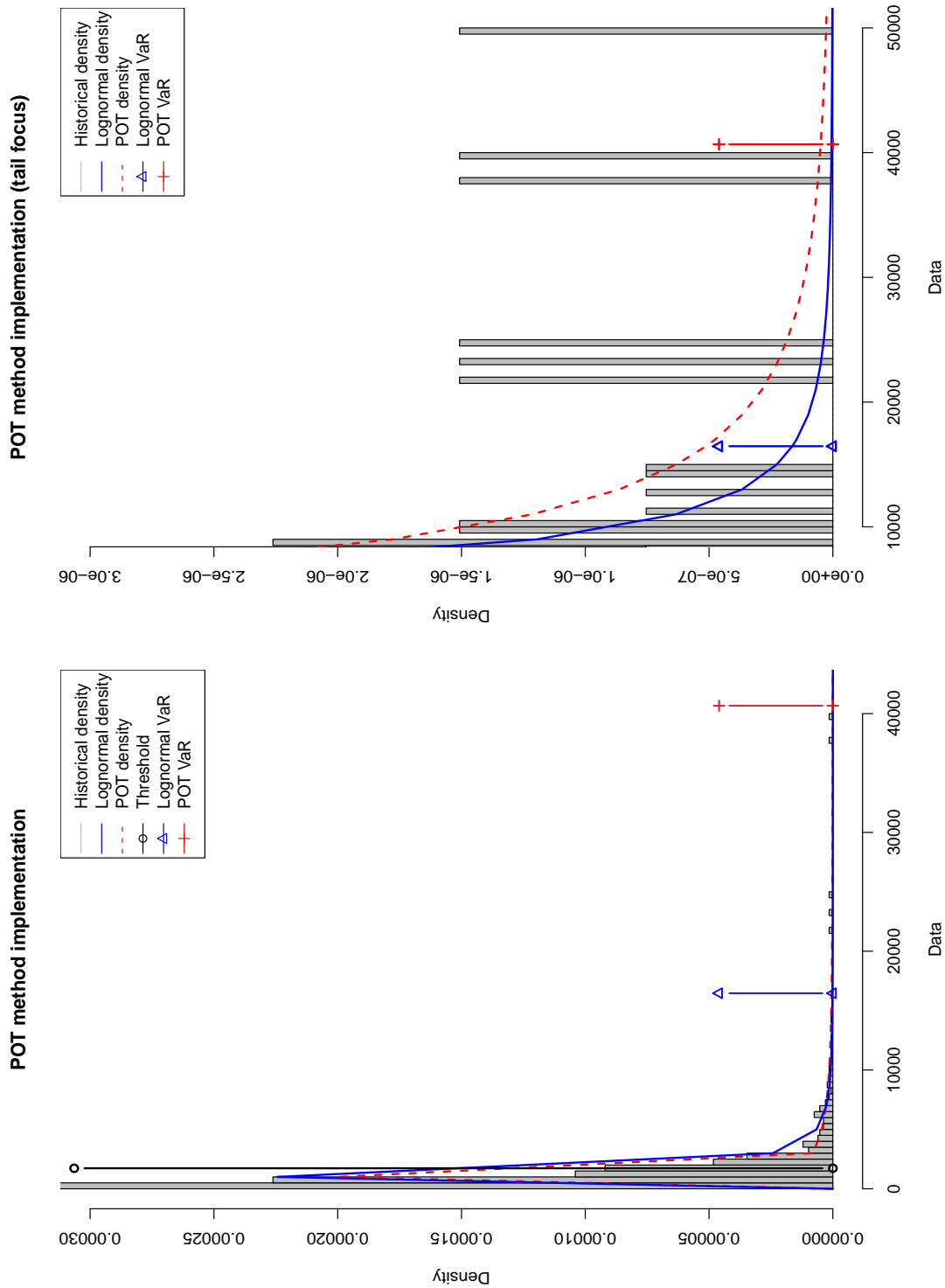


Figure 2: Method Illustration: This figure presents the histogram (in grey) of the Historical LDF. The black line corresponds to a LDF mixing a Poisson distribution and a lognormal severity. The associated 99.9% VaR is pointed out by Δ . The dash line represents a LDF mixing a Poisson distribution and a multiple pattern severity (lognormal-GPD). The corresponding VaR is represented by a $+$. The right graph focuses on the right tail of the LDFs described above. This figure highlights the fact that this method enables thickening up the right tail of the LDF.

Severity distribution	VaR (95%)	ES (95%)	VaR (99.9%)	ES (99.9%)
Lognormal	305 303	336 455	463 192	533 160
GPD	∞	∞	∞	∞
POT	1 388 655	3 070 808	15 627 945	30 166 707

Table 3: We provide risk measures on the data set which represents the severity of the business line "Payment & Settlement" and the event type "Delivery, Execution and Process Management" for the year 2006, comparing different hypothesis for the severity distribution. We cannot provide a GPD estimation on the whole data set as the shape parameter is superior to 1, and therefore we face an infinite mean model.

We also observe large amount differences with respect to the choice of the risk measure: the VaR or the ES. We analyze this last evidence in the following.

3 Capital charges associated to a coherent risk measure

The regulator imposed a confidence level of 99.9% for the capital charge, therefore if we use the VaR measure, we have $\alpha = 0.1\%$ in (1.3). Nevertheless, using this risk measure, we face several problems, especially while computing a global capital charge. Indeed, the sum of the VaRs may be lower than the VaR of the sum i.e. the VaR is known to be not subadditive is questionable as soon as we need to aggregate different VaRs. On the other hand, this measure -by building- does not take into account the large losses. However, the Conditional VAR (CVaR) or Expected Shortfall (ES) is a coherent measure (Artzner et al. (1999)) and therefore it is more appropriate to work with it. We recall its definition:

Definition 3.1. *Let η be the $VaR_{(1-\alpha)\%}$, and X be a random variable which represents losses during a prespecified period (such as a day, a week, or some other chosen time period) then,*

$$ES_{(1-\alpha)\%} = E(X|X > \eta) \quad (3.1)$$

The ES measure which is subadditive takes into account the information contained in the distribution tail contrary to the VaR measure. Therefore, extremal events exposure is contained in the ES measure. However in practice the ES value may be much higher than the corresponding VaR value, therefore, one could use a lower confidence level when we consider the ES measure

than when we use the VaR measure.

As illustrated in Table 2, it might be judicious to use a GB2 distribution to model the severity and to compute a 95% ES than to use a lognormal distribution at a 99.9% VaR. Thus, we see that, to take into account very large losses, it could be more appropriate to associate at the same time the use of extreme distributions and the ES measure: in that latter case we are close to the reality, and by the way banks could protect themselves against extremal events. In Tables 2 and 3, we illustrate these facts. we can observe a large difference between the capital computed with the VaR and the ES measures. Thus, a trade-off between a VaR at 99.9% and an ES at 95% could be considered .

In these two tables we have seen the impact of the choice of the risk measure. Now this impact can also be illustrated if we study the composition of the matrix and its granularity. Indeed, until now we use collected data organized into the Basel Matrix (BCBS (2001)). In its first level of granularity, this matrix is made up of 56 cases - 8 business lines ("b") \times 7 event types ("e")⁴. Nevertheless, each event type might be decomposed in several elements. For example, the "external fraud" event may be shared in two items - "Theft and Fraud" and "Systems Security" (second level of granularity). In a third level, the element "Theft and Fraud" may be split in several components: "Theft/Robbery", "Forgery" and "Check kiting". After a deep analysis, we observe that the kind of losses expected from a fraud with a credit card does not correspond to losses caused by someone hacking the system for instance. Nevertheless these two different kinds of losses are in the same cell. Therefore, considering the largest level of granularity, we could face multimodal empirical distributions. Consequently, the methods used to model the losses depend on the granularity level choice of the Basel matrix. This choice might have a tremendous impact on capital requirement computations. Besides, we face a trade-off between quantity of data and robustness of the estimations: indeed, if the quantity of data is not sufficient, we cannot go lower in the granularity level; on the other hand the ensuing empirical distribution is therefore an aggregate of various natures of data and the estimation of this last empirical distribution can

⁴The business lines are corporate finance, trading & sales, retail banking, commercial banking, payment and settlement, agency services, asset management and retail brokerage. The event types are internal fraud, external fraud, employment practices & workplace safety, clients, products & business practices, damage to physical assets, business disruption & system failures and execution, delivery & process management.

be source of unusable results.

For the moment, we cannot empirically illustrate these remarks because we do not have the appropriate information set. Nevertheless it seems reasonable, in order to be close to the reality, to introduce the computation of the distributions associated to second or even third level of granularity data set in order not to bias the LDF, as soon as this information exists.

4 Influence of the dependence structure between the cells on the regulatory capital computation

Banks need to have *in fine* an amount calculated using the global matrix. This means that, as soon as the LDFs have been determined on each cell, the question is to propose the best way to combine these distributions to provide a global regulatory capital i.e. corresponding to the whole matrix.

Traditionally the experts computed the global capital requirement summing the capital calculated in each cell. This procedure does not take into account the true dependence which exists within the cells. One way to bypass this problem is to use a copula distribution which is a multivariate distribution permitting to link a huge number of distributions. During the last years some experts used the Gaussian copula, to take these dependences into account, claiming inability to perform calculations using other copulas in high dimensions. Or the well-known Gaussian structure is not adapted to loss data sets (indeed, the Gaussian structure is elliptical, and does not capture tail dependence), and since 2005 it is now possible to work with copulas in high dimensions using nested copulas or vines (Aas et al. (2009)). Recently, we dealt with this last methodology and adapted it to compute the capital requirement associated with operational risks in high dimensions (Guégan and Hassani (2010)). Extending this work the main improvements for practitioners are the following:

- First, this methodology enables the use of numerous classes of copulas without restricting to the elliptic domain. One can consider copulas which focus on information contained in the tails, where we find the large losses.

- Second, this approach allows several combinations of margins (corresponding to the distributions computed for each cell) to derive robust adjustments in the statistical sense.
- Third, even working in the highest dimension, the procedure is easy to implement and is not too time consuming.
- Fourth, this method complies with the last Basel Committee (BCBS (2010)) requirements.

Now, we provide the results obtained considering 10 cells of the Basel matrix⁵. In Table 4 we introduce some notations corresponding to the losses we have studied.

BUSINESS UNITS	BUSINESS LINES LEVEL 1	Loss Distributions
INVESTMENT BANKING	(2) Trading & Sales	B_1
BANKING	(3) Retail Banking	B_2
	(5) Payment & Settlement	B_3
OTHERS	(8) Retail Brokerage	B_4

Table 4: Restricted Basel Matrix used to compute operational risk global capital allocation. B_1 , B_2 , B_3 , B_4 are the four loss distributions used in the vine methodology. B_2 and B_3 represent multivariate distribution respectively composed of F_2, F_5, F_8, F_9 and F_3, F_6, F_7, F_{10} (Table 1).

In a first exercise, we show how to compute the amount corresponding to these four cells. The diagram is provided in Figure 3. We begin by estimating the margins associated to the cells B_1 , B_2 , B_3 and B_4 . Thus, in a first step, we link the couples (B_1, B_2) , (B_1, B_4) and (B_3, B_4) with a copula. In a second step we link the copulas previously obtained by other copulas and so on. The choice of the cells to find the copula results from a sharp analysis of the data sets (cells). Nevertheless, some statistical studies enabling deciding which links are appropriate can be found in (Guégan and Maugis (2010) and Guégan and Hassani (2010)).

⁵Nevertheless the complete Basel matrix could contain more than 250 cells, and thus more research will be necessary to work with a so large matrix, mainly to limit the time of computation. Recent improvements have been proposed, using parallel computing, seem to provide interesting solutions to achieve this purpose (Brechmann et al. (2010))

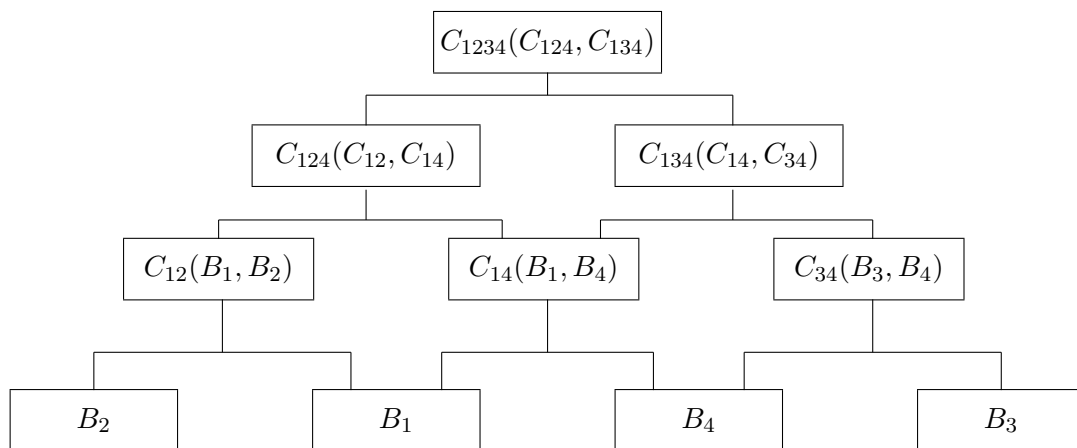


Figure 3: 4-dimensional Vine estimation to obtain the dependence structure for the whole Basel Matrix.

In Table 5 we provide the amount computed using VaR and ES measures. The results given in the first line correspond to the global amounts obtained summing the VaR provided by the four margins B_1, \dots, B_4 , considering different distributions to model the severities: 1 corresponds to a non parametric estimation procedure, 2 corresponds to the lognormal distribution estimated and 3 corresponds to the Gumbel adjustment. In the second line we provide the results obtained linking the margins with the Gumbel copula, 1, 2 and 3 correspond to the same distributions for the margins as before. Looking at the third column we observe that we do not have the same amount if we sum the four cells (methodology based on a univariate approach) or if we use a copula methodology. The amount of money is always bigger in the latter case. If we compare the third and the fourth column we observe that the capital charges are always bigger using the ES measure: this confirms some previous remarks. Now when we use a Gumbel distribution on the margins, the two methods are competitive: indeed, the results are nearly similar. The use of lognormal distribution provides erratic results. The use of a nonparametric modeling provides amount of money between the two previous situations. It seems to us that obtaining a very good fit on the margins is determinant to provide a realistic amount of capital. The modeling of the margins can be more determinant than the fit of the copula, nevertheless, using upper tail dependence copula structures permit to obtain more conservative capital charges for the banks.

Now we propose another exercise. In Table 1 we consider the cell F_9 corresponding to Business

$C_{1234}(B_1, B_2, B_3, B_4)$	Margins	VaR	ES
Univariate	1	77 018 239	80 800 510
	2	74 360 305	94 263 845
	3	48 871 039	49 085 839
Gumbel Copula	1	84 220 744	87 621 736
	2	180 619 174	234 616 218
	3	49 025 722	49 236 978

Table 5: This table provides the capital allocation (VaR) and the ES for the whole data set, considering three classes of severities (1 denotes the non parametric approach of the LDF, 2 the lognormal approach and 3 the Gumbel one.) and two classes of dependence. Univariate corresponds to the VaRs sum of each LDF, Gumbel corresponds to the Gumbel copula.

Disruption & System Failure events in the Retail Banking business unit and the distribution associated to the cell F_6 characterizing the same events in the Payment & Settlement business unit. For the distribution F_9 we estimate a Gumbel distribution or a lognormal distribution, for the distribution F_6 we estimate a Generalized Pareto distribution (GPD) or a lognormal distribution. Tables 6 and 7 provide the capital values when we link these two distributions with a Gumbel copula on one hand and with a Clayton copula on another hand. In columns 4 and 5 we provide the capital computed using the VaR and the ES measures respectively. Columns 2 and 3 gives the amount corresponding to each cell obtained by projection from the fourth column.

The results of these tables show that depending on the way we model the margins, we have tremendous differences between the VaRs. For example, we would have a VaR equal to 117 207 402 euros if F_9 is modeled with a lognormal distribution and F_6 with a GPD distribution versus a VaR equal to 2 037 655 euros if F_9 is modeled with a lognormal distribution and F_6 with a Gumbel one. Depending on the way we model the LDFs, the aggregated VaR may be multiplied by 57.52. The same behavior is observable when we project the corresponding values on the cells. For example, the multivariate VaR projection on F_9 is € 2 655 055 if F_6 is modeled using a lognormal distribution, and is equal to € 15 405 192 if F_6 is modeled using a GPD distribution. The peak for the VaR observed in that latter case is due to the capture of extreme events through

the choice of the margins: the Gumbel one. Finally, using at the same time, copula and severity distributions which take into account information in the tail provides very accurate results. Indeed, when we model F_6 using a GPD associated with a Gumbel copula, we provide a larger VaR than with the Clayton one. The differences are observed comparing the amounts € 105 422 356 with € 103 249 260 on the one hand, and € 117 207 402 with € 107 807 238 on the other.

In Table 6 we observe that the capital requirements obtained using a Gumbel copula are bigger than those obtained with a Clayton one, thus the choice of the dependence structure has also an impact on the computation of the capital charges. Now, if we use the VaR measure the difference is not significant for a bank (columns 4 in Tables 6 and 7), but if we compare the results obtained with the ES measure (columns 5 in Tables 6 and 7) the difference is tremendous (may reach almost a billion euros). Nevertheless, choosing the ES measure at a 99.9% confidence level induces much larger capital amount considering an upper tail dependence structure (Gumbel copula) than an lower tail dependence (Clayton copula).

Model	Gumbel Copula			
	LDF_9	LDF_6	VaR	ES
Gumbel-GPD	2 322 782	103 099 574	105 422 356	1 603 169 459
Gumbel-lognormal	1 471 343	566 312	2 037 655	3 091 139
lognormal-GPD	15 405 192	101 802 210	117 207 402	1 904 323 684

Table 6: For the LDF corresponding to F_9 and F_6 we provide the VaRs and the ES computed from a Gumbel copula for the year 2006. They are given respectively for three classes of severities. For instance, "Gumbel-GPD" means that we have chosen a Gumbel distribution to model F_9 and a mix of a lognormal and a GPD to model F_6 .

Finally, another exercise permits to see the influence of the dependence structure on the cells for which we need to know the capital allocation. Applying a Vine Approach, we obtained a Gumbel copula to model the dependence between several LDFs. Computing the corresponding multivariate VaR, we derived a global capital charge. In Table 10, we projected from the multivariate VaR given in column 7, the amount corresponding to each cell (axis). For example, for the second line corresponding to the "Retail Banking", we can provide the amounts pertaining

Model	Clayton Copula			
	LDF_9	LDF_6	VaR	ES
Gumbel-GPD	1 154 681	102 094 579	103 249 260	739 977 372
Gumbel-lognormal	1 455 693	649 164	2 104 857	2 637 980
lognormal-GPD	5 631 004	102 176 234	107 807 238	1 092 923 925

Table 7: For the LDF corresponding to F_9 and F_6 we provide the VaRs and the ES computed from a Clayton copula for the year 2006. They are given respectively for three classes of severities. For instance, "Gumbel-GPD" means that we have chosen a Gumbel distribution to model F_9 and a mix of a lognormal and a GPD to model F_6 .

independently to the "External Fraud", the "Clients, Products & Business Practices", the "Damage to Physical Assets" and the "Business Disruption & System Failures" event types (Table 1 line 3, column 2, 5, 8, 9). Our approach is interesting because it provides the capital for each cell through the dependence structure between the cells. This approach is totally different from the approach mainly used by practitioners who directly compute the capital associated to each cell without taking into account the information given by any other cell.

Approach	LDF	LDF_2	LDF_5	LDF_8	LDF_9	VaR	ES
Univariate	1	19 650 986	3 182 731	14 212 411	2 241 011	39 287 139	40 899 526
	2	6 240 984	2 151 627	11 676 534	4 049 831	24 118 976	32 957 873
	3	13 599 313	2 478 087	8 599 313	1 087 410	25 764 123	25 887 436
Gumbel Copula	1	20 578 056	3 300 471	15 191 828	2 386 106	41 456 461	43 024 128
	2	10 732 933	2 310 337	29 525 155	13 608 000	56 176 425	79 500 929
	3	13 617 419	2 486 453	8 603 916	1 095 587	25 803 375	25 927 487

Table 8: This table provides the VaRs and ES associated with each LDF of the set LDF_2 , LDF_5 , LDF_8 and LDF_9 when we decompose the dependence structure of the 4-dimensional set C_{2589} , considering three classes of severities (1 denotes the non parametric approach of the LDF, 2 the lognormal approach and 3 the Gumbel one.).

Table 10 highlights the fact that an upper tail dependence structure (line 3) always provides

larger capital charges than the sum of the univariate VaRs. We can also remark that the larger the tail of the theoretical distribution, the larger the gap between the ES and the VaR.

5 Conclusion: New Proposals

As discussed above, we offered a range of options for assessing Basel Pillar 1's capital charges for operational risk. We extended them and increased the sophistication and the precision in our measurement, management and control of operational risks. As we had the opportunity to experiment the methods suggested by the literature on real data sets, we faced several pitfalls and we proposed solutions to bypass them.

We suggested to use a Peak-over-Threshold method to thicken the right tail of the loss distribution function. Presenting this solution, we suggested an efficient way to obtain generalized Pareto distribution parameters which is accurate regarding goodness-of-fit tests, and therefore compliant regarding the regulator. Furthermore, our method have shown conservative results and quantitatively supported the idea that some data sets might be badly built.

We provided an innovative solution to compute aggregated risk measures (VaR and ES), dealing with dependences between Basel categories. This solution is based on nested structures and vine architectures. Carrying out this methodology, we were able to take into account specific dependences (upper tail etc.) between many margins with Archimedean and extreme value copulas. Besides, we studied the sensitivity of multivariate VaRs to modeled LDFs (margins), to dependence architectures and to copulas parameters.

In addition of these results, we also observed that estimating dynamically the parameter of the dependence structure creates important variations in the values of the Gumbel copula parameter. We illustrate this fact in Table 9. We have computed the parameter of the Gumbel copula linking the LDFs of the cells F_9 and F_6 . This parameter θ varies with respect of the information set used for its estimation : the parameters' value obtained using the year 2006, is different when we use the year 2007 or the year 2008, or the whole sample. We notice that the upper tail dependence is larger when we use this last data set. This will have an impact for

the computation of capital requirements. Thus, with respect to the information set we use, the capital requirement appears to be more or less conservative, and notion of dynamics inside the data need to be taken into account.

Year	θ	θ
2006	4.9202 (0.94)	
2007	3.7206 (0.75)	10.6610 (0.88)
2008	5.8490 (0.51)	

Table 9: Parameter estimation of Gumbel copulas estimated on F_9 and F_6 for each year 2006, 2007 and 2008 (second column). These parameters are compared to a Gumbel copula parameter estimated on the whole chronicle (third column). The corresponding standard deviation are provided in brackets.

Thus, we suggested to work dynamically and therefore measure the impact of the time passing on the distribution shapes. Besides, this idea led us to challenge the 5 years data sets required by the authorities. Indeed, these data sets may include outdated data - for example, a incident occurred in a department that does not exist anymore - or not long enough because some long memory process might be adapted.

Last but not least, we suggested to compute capital charges considering another risk measure than the VaR measure, say the expected shortfall. This measure is coherent and also enables taking into account the whole information contained in the tail. Computing a Capital allocation from a 99.9% ES is not realistic as it appears to be too conservative, but what about a lower critical threshold. For example one could consider a 95% ES instead of a 99.9% VaR (Table 2) as we illustrate in the previous exercises.

References

Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009), 'Pair copula constructions of multiple dependence.', *Insur. Math. Econ.* **44**, 182–198.

- Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999), ‘Coherent measures of risk’, *Math. Finance* **9** **3**, 203–228.
- BCBS (1988), ‘Basel committee: International convergence of capital measurement and capital standards’, *Bank for International Settlements* .
- BCBS (2001), ‘Working paper on the regulatory treatment of operational risk’, *Bank for International Settlements* .
- BCBS (2004), ‘International convergence of capital measurement and capital standards’, *Bank for International Settlements* .
- BCBS (2009), ‘Observed range of practice in key elements of advanced measurement approach (ama)’, *Bank for International Settlements* .
- BCBS (2010), ‘Basel iii: A global regulatory framework for more resilient banks and banking systems’, *Bank for International Settlements* .
- Brechmann, E., Czado, C. and Aas, K. (2010), ‘Truncated regular vines in high dimensions with applications to financial data.’, *Submitted preprint* .
- Buch-Kroman, T. (2009), ‘Comparison of tail performance of the champernowne transformed kernel density estimator, the generalized pareto distribution and the g-and-h distribution’, *Journal of Operational Risk* **4**.
- Danielsson, J., de Haan, L., Peng, L. and de Vries, C. (2001), ‘Using a bootstrap method to choose the sample fraction in tail index estimation’, *Journal of Multivariate Analysis* **76**, 226–248.
- Embrechts, P., Klüppelberg, C. and Mikosh, T. (1997), *Modelling Extremal Events: for Insurance and Finance*, Springer, Berlin.
- Frachot, A., Georges, P. and Roncalli, T. (2001), ‘Loss distribution approach for operational risk’, *Working Paper, GRO, Crédit Lyonnais, Paris* .
- Guégan, D. and Hassani, B. (2010), ‘ n -dimensional copula contributions to multivariate operational risk capital computations.’, *Working Paper, University Paris 1, n°2010.96 [halshs-00587706 - version 1]* .

- Guégan, D. and Hassani, B. K. (2009), ‘A modified panjer algorithm for operational risk capital computation’, *The Journal of OPERational Risk* **4**, 53 – 72.
- Guégan, D., Hassani, B. and Naud, C. (2010), ‘A efficient peak-over-threshold implementation for operational risk capital computation’, *Working Paper, University Paris 1, n°2010.96 [halshs-00544342 - version 1]*.
- Guégan, D. and Maugis, P.-A. (2010), ‘New prospects on vines.’, *Insurance Markets and Companies: Analyses and Actuarial Computations* **1**, 4–11.
- Gupta, A. K. and Nadarajah, S. (2004), *Handbook of Beta Distribution and Its Applications.*, New York: Marcel Dekker.
- Hall, P. (1990), ‘Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems.’, *Journal of multivariate analysis* **32**, 177–203.
- Hill, B. M. (1975), ‘A simple general approach to inference about the tail of a distribution’, *Ann. Statist.* **3**, 1163 – 1174.
- Hoaglin, D. C. (1985), ‘Summarizing shape numerically: the g-and-h distributions’, *John Wiley & Sons* pp. 461–513.
- Leadbetter, M. and Rootzen, H. (1988), ‘Extremal theory for stochastic processes.’, *Ann. Probab.* **16**, 431–478.
- Luceno, A. (2006), ‘Fitting the generalized pareto distribution to data using maximum goodness-of-fit estimators’, *Computational statistics and data analysis* **51**, 904 – 917.
- Pickands, J. (1975), ‘Statistical inference using extreme order statistics’, *annals of Statistics* **3**, 119–131.
- Rigby, R. A. and Stasinopoulos, D. M. (2005), ‘Generalized additive models for location, scale and shape,(with discussion).’, *Appl. Statist.* **54**, 507–554.

A Distributions for the severities

We provide the densities of the main severity distributions we used along this paper.

- lognormal distribution:

$$f_{b,e}(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}, \quad (\text{A.1})$$

for $x > 0$, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^{+*}$.

- Weibull distribution:

$$f_{b,e}(x; \beta, \xi) = \frac{\xi}{\beta} \frac{x^{\xi-1}}{\beta} e^{-\left(\frac{x}{\beta}\right)^\xi}, \quad (\text{A.2})$$

for $x > 0$, $\beta > 0$, $\xi > 0$.

- Gumbel distribution:

$$f_{b,e}(x; u, \beta) = \frac{1}{\beta} e^{\frac{x-u}{\beta}} e^{-e^{\frac{x-u}{\beta}}}, \quad (\text{A.3})$$

with $u \in \mathbb{R}$ and $\beta > 0$.

- Generalized Pareto distribution (GPD) (Pickands (1975), Embrechts et al. (1997)):

$$f_{b,e}(x; u, \beta, \xi) = \frac{1}{\beta} \left(1 + u \frac{(x-\xi)}{\beta}\right)^{-1-\frac{1}{u}}, \quad (\text{A.4})$$

with $1 + u \frac{(x-\xi)}{\beta} > 0$, $\beta > 0$ and $\xi \neq 0$ (or $f_{b,e}(x; u, \beta, \xi) = \frac{1}{\beta} \left(-\frac{(x-\xi)}{\beta}\right)$ if $u = 0$).

- GB2 distribution (Gupta and Nadarajah (2004), Rigby and Stasinopoulos (2005)):

$$f_{b,e}(x; \alpha, \beta, p, q) = \frac{\alpha x^{\alpha p-1}}{\beta^{\alpha p} B(p, q) \left[1 + \left(\frac{x}{\beta}\right)^{p+q}\right]} \quad (\text{A.5})$$

where $\alpha, \beta, p, q, x >$, $B(u, v) = \Gamma(u)\Gamma(v)\Gamma(u+v)$ is the Beta function, and $\Gamma(\cdot)$ is the Gamma function.

- g-and-h distribution (Hoaglin (1985), Buch-Kroman (2009)):

$$f_{b,e}(x; g, h) = \exp^{((g*x)^{-1})} * \exp^{\frac{\left(\frac{(h*p^2)}{2}\right)}{g}} \quad (\text{A.6})$$

when $g = 0$ and $h = 0$, the g-and-h distribution reduces to a standard normal distribution.

B Influence of estimation methods on the amount of regulatory capital

We show the influence of the estimation procedures of GPD's parameters used in section 2 on capital requirements. Assuming a bootstrap method to estimate the threshold (Hall (1990), Danielsson et al. (2001)), we estimate the remaining ξ and β parameters of the GPDs defined in (2.1) using the method introduced by (Luceno (2006)), denoting this method M1. We also consider three other alternative estimation methods to estimate these parameters in order to check their impact on VaR computations. These ones are respectively the Pickands method (M2) (Pickands (1975)), the Hill method (M3) (Hill (1975)), and the Maximum Likelihood method (M4). We provide in Table 10 the estimations of GDPs parameters (with their standard deviation in brackets) obtained for both sets, using these four methods and the corresponding capital charges.

Note that a shape parameter $\xi > 1$ in (2.1) induce an infinite mean model which naturally provides very high VaRs. Nevertheless, non-parametric estimator such as Pickands might provide this kind of value and by the way unusable models.

Method	β	ξ	VaR	ES
M1	932.854 (83.71)	0.767 (0.101)	15 700 112	34 215 896
M2	682.615 (160.70)	1.144 (0.266)	538 480 990	∞
M3	1007 (214.36)	0.66 (0.228)	5 725 341	14 780 214
M4	904.087 (92.31)	0.827 (0.097)	27 944 558	125 019 034

Table 10: We provide risk measures on the data set which represents the severity of the business line "Payment & Settlement" and the event type "Delivery, Execution and Process Management" for the year 2006, given estimations of the GPD's parameters ξ and β (A.4) using four methods for $\hat{u} = 179$. M1 is the method introduced by (Luceno (2006)), M2 is the Pickands method, M3 is the Hill method, and M4 is the Maximum Likelihood method. We provide in brackets the standard deviations computed by bootstrapping. The third column gives capital allocations pertaining to these estimates ($\mu = 3.593098$, $\sigma = 1.510882$). The fourth column presents the corresponding expected shortfall.