

A Practical Approach to Quantitative Model Risk Assessment

Carole Bernard,^{*}Rodrigue Kazzi,[†]and Steven Vanduffel^{‡§}

November 24, 2020

Abstract

Model-based decisions are highly sensitive to model risk that arises from the inadequacy of the adopted model. This paper reviews the existing literature on model risk assessment and shows how to use the theoretical results to develop a corresponding best practice. Specifically, we develop tools to assess the contribution to model risk of each of the assumptions that underpin the adopted model. Furthermore, we introduce new model risk measures and propose an intuitive formula for computing model risk capital. Some numerical examples and a case study illustrate our results.

Keywords: Value-at-Risk, Risk bounds, Model risk contribution, Model risk capital requirements, Parameter risk, Credibility factors, Generalized Pareto distribution.

^{*}Carole Bernard, Department of Accounting, Law and Finance, Grenoble Ecole de Management (GEM). Department of Economics and Political Sciences at Vrije Universiteit Brussel (VUB). (email: carole.bernard@grenoble-em.com).

[†]Rodrigue Kazzi, Department of Economics and Political Sciences at Vrije Universiteit Brussel (VUB). (e-mail: rodrigue.kazzi@vub.be).

[‡]Steven Vanduffel, Department of Economics and Political Sciences at Vrije Universiteit Brussel (VUB). (e-mail: steven.vanduffel@vub.be).

[§]The authors gratefully acknowledge funding from the 2019 Individual Grants Competition of the CAS and from the FWO.

1 Introduction

The financial and insurance industry relies heavily on the use of models in decision-making (e.g., in the determination of regulatory capital requirements, pricing). However, these models are subject to error. Specifically, model risk can be defined as the potential loss that can result from the misspecification or misuse of models (inspired by the definition in the Capital Requirements Directive (CRD) IV, Article 3.1.11). Model failures have had serious consequences in the past; e.g., in 1997, Long-Term Capital Management (LTCM) hedge fund lost around 4.5 billion dollars as a consequence of relying on normality assumptions while neglecting the importance of stress testing (Lowenstein 2008). In 2008, the investors poorly understood the Gaussian-copula in the pricing formula of David Li and yet they over-relied on it in pricing credit derivatives. This over-reliance was one of the drivers of the 2008 financial crisis (Salmon 2009). It is nowadays well-understood that model risk is a key concern and hence its quantification is of vital importance.

Recent regulations related to Model Risk Management: Calculating capital requirements based on models that are neither challenged nor back-tested is nowadays seen as non-credible by regulators. We cite from The Deloitte Center for Regulatory Strategy (2018), *“Supervisors will neither approve nor place reliance on the firm’s strategic and operational use of a model, including for risk assessment and capital planning, unless satisfied with a firm’s model risk management”*. In 2011, the first supervisory guidance on model risk management called SR 11-7 was published by the Office of the Comptroller of the Currency (OCC) and the Federal Reserve (Board of Governors of the Federal Reserve System (2011)). The key focus of the SR 11-7 is on ensuring an effective challenge of models. In 2013, the Basel Committee expressed some concerns about model uncertainty. It

conducted surveys that showed that risk-weightings for the same assets may differ among banks; this observation undermines the credibility of the models used by banks (Basel Committee (2013)). In their discussion paper on the review of specific items in Solvency II Regulation, the Actuarial Association of Europe insists on focusing more on model risk assessment (Actuarial Association of Europe (2017)). A recent supervisory statement of the Prudential Regulation Authority of the Bank of England highlights the necessity of understanding and accounting for the assessment of model uncertainties. This document also presents some Model Risk Management principles that are deemed important to apply when using stress test models (Prudential Regulation Authority (2018)). While accounting for model risk assessment in the decision-making process of financial institutions is a necessity that regulators are striving to pose, its application is still not obvious and, if performed, tends to be of a qualitative rather than quantitative nature. A major obstacle is that the quantitative methods developed by researchers so far are either not well suited to deal with the assessment of model risk in practice or are not yet recognized by the practitioners.

Current practices of Model Risk Assessment: Banks and insurance companies around the globe tend to use similar risk management frameworks that are proposed by local or international professional organizations. For example, insurance companies in the United States, Canada, and most of the countries of North and South America are known to follow the frameworks suggested by the Society Of Actuaries (SOA) and the Canadian Institute of Actuaries (CIA). Three main approaches to model risk assessment can be distinguished.

First, what is commonly demanded/applied in the market consists in assessing the model risk by so-called risk-rating schemes, commonly known as *Qualitative Scoring Method*

(Dionne and Howard (2017)). In other words, to assess a model, one evaluates various common inherent risk factors like the complexity of the model, the expertise of its users, the quality of its reporting, the frequency of its usage, its financial impact, and so on. The risk manager assigns a score to each factor and sums up the individual scores to get a comprehensive view of the model risk. Although this approach is easy to understand and to implement, it is also very imperfect. Indeed, this methodology may confuse the risk that results from improper implementation of the model with the risk that results from model uncertainty; the first risk is in principle already accounted for in the operational risk assessment, but the second one needs particular consideration. Moreover, the scores and weights attributed to each factor are highly subjective, as they depend heavily on expert judgment, which makes this approach far from robust. Finally, risk mitigation controls can be misleading. In order to decrease the model risk under this assessment methodology, one would aim to decrease the final score by, for example, decreasing the complexity of the model; however, a decrease in complexity does not necessarily decrease the model risk.

Second, a quantitative approach to model risk measurement is the *Modern Operational Risk Management* approach. Model risk is viewed as a type of operational risk and is measured by modeling the frequency and severity of model risk events. A presentation of this approach can be found in Samad-Khan (2008) and OpRisk Advisory and Towers Perrin (2010). Although this approach offers consistency with the way other types of risk are assessed and allows for interdependencies among risks, there are some serious drawbacks. Empirical data on model risk losses can be scarce and inaccurate. In addition, this approach in itself is subject to potential high model risk. These major limitations hinder the use of this approach in practice.

Third, a recent paper on model risk prepared by the Model Risk Working Party of the Institute and Faculty of Actuaries (Black et al. (2018)) states that a quantitative

assessment based on models' comparison is highly preferable but is unfortunately not ready for practical use; using their words "*Where alternative methodological choices to those employed in a model are plausible, the impact of method changes on model outputs can also be tested... However, ... methodological changes (e.g., a change in dependence structure or valuation method) are too time-consuming to implement for test purposes.*" A quantitative approach that reflects IFoA's view is the *Model Uncertainty Approach* (MUA). It is based on the principle of model risk sensitivity. This sensitivity is measured using benchmarking, back-testing, and comparison with alternative models (Jacobs Jr (2015)). In addition, this is a bottom-up approach to aggregating model risk in the sense that it consists in evaluating the model risk inherent in each individual modeling and then aggregating individual model risks using models' dependencies. As compared to the other approaches, MUA offers various advantages. It is much less subjective and provides a clear and traceable process with the ability to provide guidance on the level of capital that is needed to account for model risk. However, the expertise, time, and resources this approach requires make it less desirable by practitioners.

In this paper, we develop a quantitative approach to model risk assessment that aligns with the MUA. Our approach is based on the theory of risk bounds that studies the behavior of a model under worst-case and best-case scenarios. To compute risk bounds, one first selects the model assumptions that can be fully trusted, hence distrusting all other assumptions. Next, one determines a model (a worst-case scenario) that is consistent with these trusted assumptions and that leads to the highest possible value for the risk measure. Similarly, one determines the model that yields the lowest possible value. These two extreme values are the risk bounds. Our contributions are as follows. First, we define a new notion of model risk that allows to quantify the model risk contribution of each assumption by making use of the extensive literature on risk bounds. Second, we extend

the notion of risk bounds by assigning a “credibility score” to each assumption of interest instead of making a binary decision on whether it is fully trusted or distrusted. This results in tidier bounds that better reflect experts’ opinions and are more useful in practice. Our third contribution is to propose new measures that make it possible to establish the capital buffer for model risk.

The paper is organized as follows. In Section 2, we review the recent academic literature on risk bounds. To ensure that the paper is self-contained and to facilitate its use, we recall the main theorems on risk bounds in the extended Appendix C. In Section 3, we present our model risk assessment approach which incorporates the abovementioned contributions. Section 4 is a case study where we apply our approach to a real-world dataset, the SOA medical dataset (Grazier (1997)). Section 5 concludes.

2 Literature review on risk bounds

The quantitative approach to model risk assessment aims to assess the uncertainty arising from the choice of the probability model (e.g., the assumption that the loss distribution belongs to the exponential family of distributions), the adoption of parameter calibration techniques (since different techniques may actually lead to different parameter estimates), and the limitation of the collected data. Over the past decades, many researchers explored this topic. A standard approach to assess model risk consists in comparing the capital (or more generally speaking, the value of a risk measure) resulting from the model used with the one resulting from extreme models; extreme models are the models adopted in worst- and best-case scenarios. The extreme values taken by the adopted risk measure are referred to as risk bounds. In this approach of model risk assessment, a first step consists in specifying which assumptions we can be certain about and which ones we cannot. The

bounds thus clearly depend on the model assumptions that are considered as fully reliable. For instance, one might trust that the distributions of the portfolio components are fully known, but not the interdependence. Another example is that one might be certain that the loss distribution is unimodal and has a known mean and variance.

Dependence uncertainty bounds, i.e., risk bounds when the dependence structure is unknown but information on the marginals is available, have received special attention by researchers. This stream of the literature finds its pedigree in Rüschemdorf (1981) and Makarov (1982). In the homogeneous case, i.e., when the distribution functions of the marginals are identical, Wang and Wang (2011) and Puccetti and Rüschemdorf (2013) obtained sharp tail bounds in the case of monotone densities and concave densities, Wang et al. (2013) found explicit formulas for the worst Value-at-Risk when the marginal densities are monotone or tail-monotone, and Wang (2014) studied asymptotic bounds. In the inhomogeneous case, the analysis becomes more complicated, and approximations of bounds were needed. Therefore, a new algorithm, called Rearrangement Algorithm (RA), that can numerically approximate Value-at-Risk sharp bounds (i.e., attainable bounds) for the distribution of the aggregate risk was developed by Puccetti and Rüschemdorf (2012b) and Embrechts et al. (2013). Furthermore, Bernard et al. (2017) provided explicit (but non-sharp) upper and lower bounds of the Value-at-Risk of the portfolio loss when we only have information on the marginal distributions (see Theorem 1 in Appendix C).

Various attempts to include dependence information have been made in the literature. Puccetti and Rüschemdorf (2012a) offered an improvement on some existing bounds for the distribution function and the tail probabilities of portfolios by adding a positive dependence restriction on the dependence structure. Bigozzi et al. (2015) proved that an assumption of a negative dependence would mainly affect the upper bound of the Value-at-Risk but an assumption of a positive dependence would affect the lower bound (see Theorems 2,

3, and 4). Puccetti et al. (2016) considered the Value-at-Risk upper bounds in the case where positive dependence information is assumed in the tails or some central part of the distribution function. In Puccetti et al. (2017), independence among (some) subgroups of the marginal components is assumed, this fact leads to a considerable improvement to the Value-at-Risk bounds as compared to the case where only the marginals are known (see Theorems 5 and 6).

In practical situations, estimating the dependence structure can be very challenging and can lead to inaccurate results in most cases. By contrast, moment estimates can be performed with a reasonable degree of accuracy (note that the accuracy decreases with the increase in the order of the moment). This observation constitutes a motivation for the many papers that replaced the assumption on the dependence structure by a constraint on the variance as some source of dependence information. In fact, it is intuitive to see that adding variance and higher order moments constraints to a setting in which only the marginals are fully known is likely to improve the risk bounds since this addition captures information that cannot be represented by the marginals. Bernard et al. (2017) derived Value-at-Risk bounds based on the knowledge of the marginal distributions and the variance of the portfolio risk (see Theorem 7). Bernard et al. (2017) studied these bounds under the knowledge of higher order moments (skewness for instance). Interestingly, Bernard et al. (2018) provided evidence that replacing the knowledge of the marginal distributions by the knowledge of the collective mean does not cause a significant loss of information. In fact, a considerable number of papers have studied risk bounds in scenarios in which information on the mean and higher order moments of the portfolio risk is assumed instead of assuming knowledge on marginal distributions and dependence structure; we cite Kaas and Goovaerts (1986), Hürlimann (1998), Hürlimann (2002), De Schepper and Heijnen (2010), Zymmler et al. (2013), among others. Bertsimas et al. (2004) derived Value-

at-Risk bounds when only the mean and a maximum variance of the portfolio loss can be trusted (see Theorem 8). Moreover, Puccetti et al. (2017) derived bounds when information on the maximum variance of the portfolio loss is assumed in addition to the knowledge of the marginals and the independence among some subgroups of the marginals (see Theorem 9).

Another case of interest is the factor model in which each individual risk depends on a common risk factor. Many important models in risk management can be seen as factor models, e.g., the multivariate normal mean-variance mixture model. Bernard et al. (2017) derived risk bounds (mainly of the Value-at-Risk and the Tail-Value-at-Risk) when factor models are only partially specified (see Theorems 10 and 11).

Other sets of assumptions that were considered in the literature are the shape and the domain of the loss distribution. Bernard et al. (2017) derived the upper bound of the Value-at-Risk of a non-negative portfolio loss whose mean is the only assumption that can be fully trusted (Theorem 12). Bernard et al. (2018) derived risk bounds when the portfolio loss is bounded and information on marginals and maximum moments are available (see Theorems 13 and 14). Bernard et al. (2019) derived risk bounds for unimodal portfolio distributions, and considered the case of non-negative portfolios with possibly a theoretically infinite variance (see Theorems 15, 16, and 17). Additional results can be found in the literature such as in Li et al. (2018) where information on the shape (unimodality and symmetry) of either the individual risks or the total risk is taken into account.

3 Model risk assessment

An adopted model is composed of a set of adopted assumptions. The traditional way, as seen in the academic literature, to assess the model risk inherent in an adopted model

consists in following a three-step approach. In a first step, one specifies all assumptions that can be fully trusted. In general, there exist many models that are consistent with these assumptions, and the adopted model is merely one among these. In a second step, one maximizes and minimizes the risk measure over the set of all plausible models, that is, one determines the worst-case and best-case value (i.e., the bounds). Finally, in a third step, one compares the risk bounds with the value the risk measure takes under the adopted model.

Under the traditional approach, however, all the assumptions that are not fully credible are completely neglected. In this section, we work on improving the traditional approach by overcoming this drawback. Firstly, we will propose a method to assess the contribution of each assumption to the total model risk. This analysis will provide the modeler with insight as to how risky it is to adopt each of the assumptions in terms of model risk. Secondly, we present a novel approach to improve the risk bounds by making use of the assumptions that are not fully credible. Finally, we introduce a new measure for the model risk capital buffer.

3.1 Setting

Consider a portfolio of n individual risks $\{X_i\}_{1 \leq i \leq n}$, and denote the portfolio loss variable by $S = \sum_{i=1}^n X_i$. Note that, unless otherwise stated, we do not assume the portfolio to be homogeneous.

We write $X \sim F_X$ to express that F_X is the cumulative distribution of X . And we denote by $E[X]$, $\text{var}[X]$, and $\text{std}[X]$ the mean, variance, and standard deviation of X , respectively.

Let $\rho : \mathcal{M} \rightarrow \mathbb{R}$ be a risk measure where \mathcal{M} is the space of all real-valued random

variables defined on a probability space (i.e., \mathcal{M} is a set of measurable functions). Let F^* be the cumulative distribution function of the adopted model and let $\{a_i\}_{i \in \{1, \dots, n\}}$ be a set of assumptions that characterizes F^* , i.e, adopting the set of assumptions $\{a_i\}_{i \in \{1, \dots, n\}}$ is equivalent to adopting F^* . We define a decreasing sequence of sets $\{\mathcal{A}_k\}_{k \in \{1, \dots, n\}}$ with $\mathcal{A}_k = \{X \in \mathcal{M} \mid X \text{ respects assumptions } \{a_i\}_{i \leq k}\}$.

In a scenario where only the assumptions $\{a_i\}_{i \leq k}$ are adopted, we define the corresponding upper and the lower bounds of the risk measure ρ by

$$\bar{\rho}_k = \sup_{X \in \mathcal{A}_k} \rho(X), \quad (3.1)$$

and

$$\underline{\rho}_k = \inf_{X \in \mathcal{A}_k} \rho(X), \quad (3.2)$$

respectively.

Without loss of generality of the methods we propose, the risk measure that we consider in this paper is the Value-at-Risk (VaR). Note that VaR is indeed commonly used as the reference measure for computing the capital requirement in the industry. In fact, the Value-at-Risk at a probability level α represents the amount of capital necessary to ensure with a confidence level α that the insurance company or financial institution will not be technically insolvent after a specific period. Formally, VaR is defined as

$$\text{VaR}_\alpha(S) = \inf\{x \in \mathbb{R} \mid F_S(x) \geq \alpha\}, \quad \alpha \in (0, 1), \quad (3.3)$$

where F_S is the cumulative distribution function of the aggregate risk S .

3.2 Model risk allocation

A fully trusted model (i.e., when the assumptions leading to the model are all fully trusted) poses no model risk. However, such a model does not exist. Typically, only some of the assumptions can be fully trusted, say $\{a_i\}_{i \leq k}$. The risk bounds ($\bar{\rho}_k$ and $\underline{\rho}_k$) corresponding to those fully trusted assumptions reveal the uncertainty coming from the non-fully trusted assumptions ($\{a_i\}_{k < i \leq n}$); the wider the bounds, the more uncertainty we have in the non-fully trusted assumptions. This uncertainty is seen as a representation of the total model risk entailed by the non-fully trusted assumptions.

However, if we only calculate risk bounds in the one scenario where we split fully trusted versus non-fully trusted assumptions, we will be only assessing the uncertainty that arises from completely distrusting the whole set $\{a_i\}_{k < i \leq n}$. This approach is incomplete as it ignores the actual impact of each assumption on the total model risk.

In order to reveal the impact of a specific set of assumptions, say $\{a_i\}_{k < i \leq l}$, on the total model risk inherent in $\{a_i\}_{k < i \leq n}$, it would be intuitive to observe how sensitive the risk bounds are to the addition or removal of $\{a_i\}_{k < i \leq l}$ from the set of assumptions responsible of the total model risk.

In other words, in order to assess the model risk contribution of a subset of the non-fully trusted assumptions, we can assume knowledge of this subset and translate it into additional constraints in the maximization and minimization of the adopted risk measure. We can then compare the newly derived risk bounds with the risk bounds derived before adding the new constraints. The risk bounds, if changed, will get tighter as the uncertainty decreases when more information is added. The decrease in uncertainty will reveal the contribution this subset of non-fully trusted assumptions had to the total model risk. Following this reasoning, we can define a measure of the contribution of any set of assump-

tions to the total model risk conditional on the knowledge of another set of assumptions. This measure is introduced in mathematical terms in Definition 3.1.

Definition 3.1 (Conditional model risk contribution measure). Let $\rho : \mathcal{M} \rightarrow \mathbb{R}$ be a risk measure. Let F^* be the cumulative distribution function of the adopted model and let $\{a_i\}_{i \leq n}$ a set of assumptions that characterizes F^* . Let us define a decreasing sequence of sets $\{\mathcal{A}_k\}_{k \leq n}$ with $\mathcal{A}_k = \{X \in \mathcal{M} \mid X \text{ respects assumptions } \{a_i\}_{i \leq k}\}$. Then, for $k < l$, we can define a measure of the contribution of assumptions $\{a_i\}_{k+1 \leq i \leq l}$ to the total model risk given full knowledge of $\{a_i\}_{i \leq k}$ when using the risk measure ρ as follows,

$$\mathcal{C}(\rho, \mathcal{A}_k, \mathcal{A}_l) = 1 - \frac{\bar{\rho}_l - \underline{\rho}_l}{\bar{\rho}_k - \underline{\rho}_k}, \quad (3.4)$$

where $\bar{\rho}_j = \sup_{X \in \mathcal{A}_j} \rho(X)$ and $\underline{\rho}_j = \inf_{X \in \mathcal{A}_j} \rho(X)$.

Example Let us consider an adopted model $F^* = \mathcal{N}(10, 4)$ where the assumptions of interest are $\{a_i\}_{i \in \{1, 2, 3, 4\}} = \{a_1 = \text{the mean is 10}, a_2 = \text{the variance is equal to 4}, a_3 = \text{the distribution is unimodal}, a_4 = \text{the distribution is Normal}\}$. Let $\text{VaR}_{95\%}$ be the adopted risk measure. In order to assess how much the unimodality assumption contributes to the total model risk left after fully trusting that the mean and the variance are equal to 10 and 4 respectively, we calculate $\mathcal{C}(\text{VaR}_{95\%}, \mathcal{A}_2, \mathcal{A}_3)$. Using the bounds of Theorems 8 and 15, $\mathcal{C}(\text{VaR}_{95\%}, \mathcal{A}_2, \mathcal{A}_3) = 1 - \frac{28.75-10}{38.21-9.86} = 33.44\%$. Hence, given the knowledge of the mean and the variance, the unimodality assumption constitutes 33.44 % of the total model risk.

We can see that $\mathcal{C}(\rho, \mathcal{A}_k, \mathcal{A}_l)$ is a relative measure of model risk contribution whose value is between 0 and 1. Specifically, we observe $\mathcal{C}(\rho, \mathcal{A}_k, \mathcal{A}_l) = 0$ only when $\bar{\rho}_l = \bar{\rho}_k$ and $\underline{\rho}_l = \underline{\rho}_k$, which means that assuming $\{a_i\}_{k < i \leq l}$ after having already assumed $\{a_i\}_{i \leq k}$ does not bring any additional model risk. In the case of the other extreme, we observe

$\mathcal{C}(\rho, \mathcal{A}_k, \mathcal{A}_l) = 1$ when $\bar{\rho}_l = \underline{\rho}_l$, this means that assuming $\{a_i\}_{k < i \leq l}$ constitutes the whole model risk conditional to the knowledge of $\{a_i\}_{i \leq k}$.

Remark 3.1. We actually have, in several cases, explicit formulas for the risk bounds, which leads to an explicit formula for the conditional model risk contribution $\mathcal{C}(\rho, \mathcal{A}_I, \mathcal{A}_J)$. For instance, in the previous example, in order to calculate the model risk contribution of adding a unimodality assumption, given the information on the mean and the maximum variance when using the risk measure VaR_α ($0 < \alpha < 1$), we used the bounds of Theorems 8 and 15 and easily obtained an explicit form of \mathcal{C} . In this particular scenario, interestingly, the conditional model risk contribution \mathcal{C} only depends on the probability level α , i.e., if we assume knowledge of the mean and the maximum variance, the model risk contribution of adding the unimodality assumption is independent of the values of the mean and the maximum variance.

Remark 3.2. Note that, based on the difference between the sets \mathcal{A}_k and \mathcal{A}_l , different types of risk can be assessed. For example, if the difference in the sets results from adding an assumption on the parameter, then a parameter risk is being assessed and so on.

Furthermore, the measure in Definition 3.1 can be used sequentially in order to perform a so-called “model risk allocation.” This can be explained as follows.

- We look at the adopted model, and we try to disassemble the assumptions upon which it was built. This can be done in a backward order in the sense that we start with a set of assumptions that fully define the adopted model and then we start to remove one or a subset of assumptions at a time until we obtain a set of the most basic and credible assumptions (i.e., the set of the fully trusted assumptions).
- After arriving at the most basic scenario, a comparison of the distance between the upper and lower bounds among the different relevant scenarios can be performed.

Specifically, the comparison can be performed using $\mathcal{C}(\rho, \mathcal{A}_k, \mathcal{A}_l)$ by moving forward from the most basic scenario to the adopted model. This would reveal the marginal effect on model risk of adding each assumption conditional on the already adopted assumptions.

This idea will become clearer, as we present and explain the methodology of model risk allocation more thoroughly using diagrams in the following subsection.

3.3 Summary diagrams

The literature offers risk bounds in many scenarios of interest. We present hereafter a few of them with their interrelation in diagrams. Indeed, we present three diagrams that can be useful in performing the model risk allocation. Note that we use the VaR as the risk measure ρ . Each scenario of the first two diagrams corresponds to a theorem that stems from the literature. To ensure that the paper is self-contained and to facilitate its use, we recall each theorem in Appendix C using consistent notations. The first diagram (in Figure 1) presents particular assumptions that can be made for the dependence among the individual components of a portfolio, whereas the other two diagrams serve to deal particularly with assumptions made for the portfolio loss when seen as a univariate random variable.

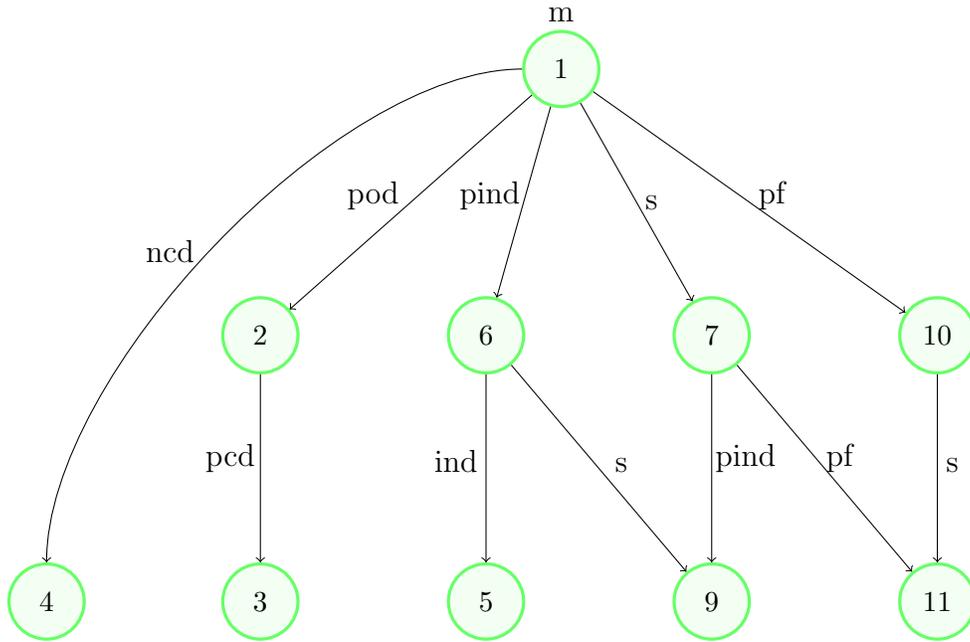


Figure 1: Diagram showing the link between scenarios adopted in the academic literature. The scenarios consist on assumptions made on the portfolio loss and its components on aggregate and marginal levels. Each number refers to a scenario for which the bounds are derived and presented in the appendix by a theorem of the same number. The notations are consistent with the ones in the theorems. The assumptions considered are the assumptions of knowledge of the marginals (m), sequential negative cumulative dependence (ncd), positive orthant dependence (pod), partial independence (pind), maximum standard deviation of the portfolio loss (s), having a partially specified factor model (pf), sequential positive cumulative dependence (pcd), and independent subgroups (ind). The symbols above the root node refer to the assumption adopted in the node, the notations on each arrow refers to the assumption added when following this arrow.

In order to perform the model risk allocation on an adopted model described by a specific probability distribution, we can start first by relaxing some of its assumptions while staying in the same family distribution, i.e., by studying the position of our model compared to its family of distributions (e.g., Diagram in Figure 3). Then we can make a bigger relaxation to drop the family of distribution assumptions and move backwardly to

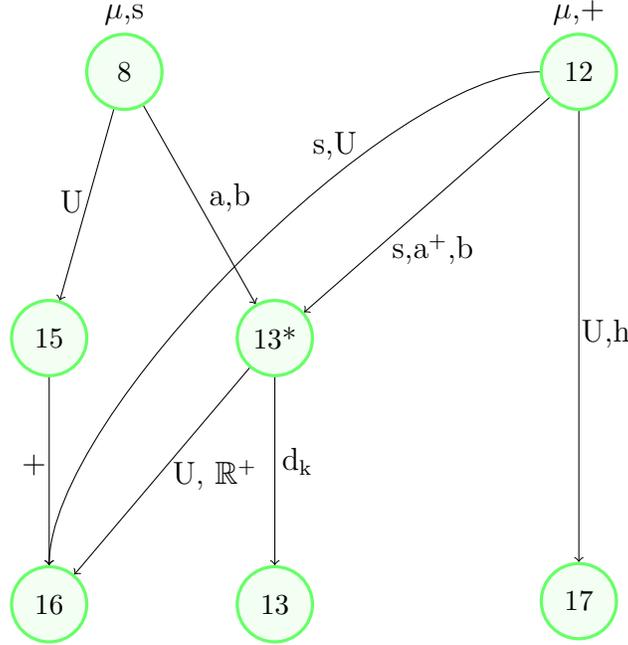


Figure 2: Diagram showing the link between scenarios adopted in the academic literature. The scenarios consist on assumptions made on the portfolio loss considered as a univariate random variable. Each number refers to a scenario for which the bounds are derived and presented in the appendix by a theorem of the same number. The notations are consistent with the ones in the theorems. The assumptions considered all refer to the portfolio loss random variable and are the assumptions of knowledge of the mean (μ), maximum standard deviation (s), non-negativity ($+$), unimodality (U), minimum portfolio loss value (a), maximum portfolio loss value (b), minimum positive portfolio loss value ($+$), infinite variance (h), setting the infimum and supremum of the portfolio loss to 0 and infinity respectively (\mathbb{R}^+), and adding $k-2$ higher moments than the second moment (d_k). The symbols above the root nodes refer to the set of assumptions adopted in the corresponding nodes, the notations on each arrow refers to the set of assumptions added when following this arrow. The asterisk on 13* refers to the fact that Theorem 13 was restricted to the case where only the maximum value of the second moment is known.

the more basic assumptions (e.g., Diagrams 1 and 2 respectively displayed in Figures 1 and 2). Note that the assumptions of at least one of the root scenarios should be respected during the whole backward path, this root scenario is often seen easy to fully trust. Note

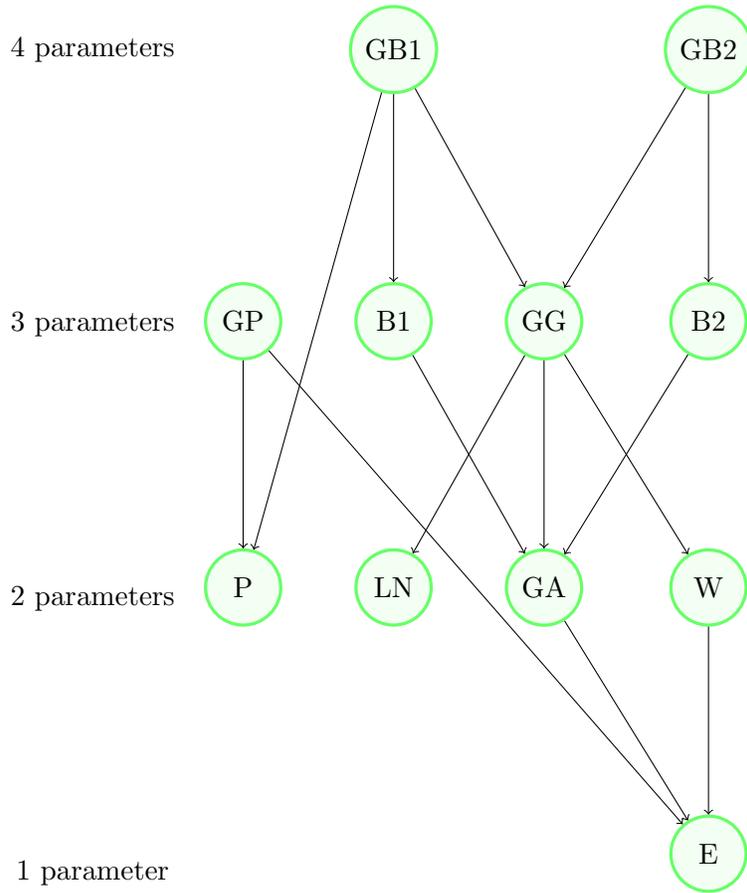


Figure 3: Diagram showing the link between some families of distributions. The position of two distributions on the two endpoints of an arrow refers to the fact that the one on the head of the arrow is a special case of the other distribution. The distributions set on the same horizontal line share the same number of parameters as depicted on the left. The distributions considered in the diagram are: the Generalized Beta of first and second kind (GB1 and GB2 respectively), the Beta of the first and second kind (B1 and B2 respectively), the Generalized Gamma (GG), the Generalized Pareto (GP), the Lognormal (LN), the Gamma (GA), the Weibull (W), the Pareto (P), and the Exponential (E) distributions.

also that the assumption of knowledge of the portfolio mean μ in Diagram 2 can be easily replaced by an assumption of knowledge of an interval of the mean; this would be calculated by maximizing/minimizing the bounds as functions of the mean.

Illustration We present a toy example to illustrate the methodology. Assume that a variable of interest X is modeled using an exponential distribution with mean $E[X] = 10$ and variance $\text{var}[X] = 100$, i.e., $X \sim \text{Exp}(\lambda = 0.1)$. Then the third quartile of X is equal to $\text{VaR}_{75\%}(X) = \frac{-\ln(1-0.75)}{0.1} = 13.86$. Assuming an exponential distribution with a specific parameter is rather a very strong and restrictive assumption. If the data allows being fully confident only about the fact that the mean belongs to the interval $[8, 12]$ and the variance is less than 196, then all other assumptions that led to the adopted model should be tested. Disassembling assumptions by walking backward in the graphs would lead to several paths. We provide a few examples for illustrative purposes.

- If we consider that we only trust the intervals of the mean and the variance, the risk bounds based on Theorem 8 are $(-0.08, 36.25)$. These bounds are very wide and very far from 13.86, which makes them, if used on their own, almost useless in practice.
- We start by choosing one simple backward path, say, $E \rightarrow \text{GA} \rightarrow 16 \rightarrow 15 \rightarrow 8$ (see Figures 2, and 3).
- At this step, we start the allocation of model risk. Moving from the bounds of 8 to the ones of 15, we compare the two bounds to assess how much model risk is allocated to the unimodality assumption assuming that we trust the information on the intervals of the mean and the variance. The bounds of 15 are in this case $(1.27, 27.87)$. Using the notations introduced in Definition 3.1, we have that $\rho = \text{VaR}_{75\%}$, $\mathcal{A}_2 = \{X \in \mathcal{M} \mid E[X] \in [8, 12], \text{var}[X] \in [0, 196]\}$, and $\mathcal{A}_3 = \{X \in \mathcal{M} \mid E[X] \in [8, 12], \text{var}[X] \in [0, 196], X \text{ is unimodal}\}$, and we obtain $\mathcal{C}(\rho, \mathcal{A}_2, \mathcal{A}_3) = 26.78\%$. Hence, assuming that we trust the intervals of the mean and the variance, the unimodality assumption constitutes 26.78 % of the total model risk.
- The bounds in 16 are $(1.27, 23.68)$. Since the lower bound is not sharp, the conditional model risk contribution of the non-negativity assumption, i.e., $\mathcal{C}(\rho, \mathcal{A}_3, \mathcal{A}_4)$, is at least

15.75%. However, for simplicity, we take here the exact value of 15.75% (for more details on the sharpness of bounds refer to Appendix A).

- It is interesting as well to study the model risk inherent in the choice of a bigger family of the exponential and gamma distributions. It is aimed to answer the question on whether the risky choice was the adoption of the gamma family or the choice of the exponential among the gamma family members. The VaR bounds for a gamma distributed random variable whose mean and variance respect the given intervals are (8.11, 16.64). The conditional model risk contribution of choosing a gamma distribution after having knowledge of the intervals on the mean and the variance and of the unimodality and non-negativity, i.e., $\mathcal{C}(\rho, \mathcal{A}_3, \mathcal{A}_4)$, is 61.94%.
- Let us now assess the parameter risk inherent in choosing an exponential distribution family. The lower and upper bounds of the $\text{VaR}_{75\%}(X)$ within the exponential distribution while respecting the intervals of the mean and variance are (11.09, 16.64). The conditional parameter risk contribution is 34.94%.

In order to formally summarize the results using the notations of Definition 3.1, we express the assumptions in Table 1, where the notations are as defined in Figures 2 and 3, but with μ and s particularly referring to $E[X] \in [8, 12]$ and $\text{var}[X] \in [0, 196]$ respectively.

i	1	2	3	4	5	6	7
a_i	μ	s	U	+	GA	E	Adopted model

Table 1: A sequence of assumptions inherent in the adopted model. The notations are as defined in Figures 2 and 3, but with μ and s particularly referring to $E[X] \in [8, 12]$ and $\text{var}[X] \in [0, 196]$ respectively.

We then display our estimates for the conditional risk contributions in Table 2. The first three columns of Table 2 reflect what was discussed previously. The last column of Table 2 presents how much the aggregate assumptions constitute of the total model risk. This measure is only conditional on the basic scenario, which we assume fully credible by default. Hence, $\mathcal{C}(\rho, \mathcal{A}_2, \mathcal{A}_{k+1})$ can be seen (in some sense) as the unconditional measure of model risk contribution of assumptions $\{a_i\}_{i \leq k+1}$.

k	$(\underline{\rho}_k, \bar{\rho}_k)$	$(\underline{\rho}_{k+1}, \bar{\rho}_{k+1})$	$\mathcal{C}(\rho, \mathcal{A}_k, \mathcal{A}_{k+1})$	$\mathcal{C}(\rho, \mathcal{A}_2, \mathcal{A}_{k+1})$
2	(-0.08, 36.25)	(1.27, 27.87)	26.78%	26.78%
3	(1.27, 27.87)	(1.27, 23.68)	15.75%	38.32 %
4	(1.27, 23.68)	(8.11, 16.64)	61.94 %	76.52 %
5	(8.11, 16.64)	(11.09, 16.64)	34.94 %	84.72%
6	(11.09, 16.64)	(13.86, 13.86)	100%	100%

Table 2: An application of the conditional model risk contribution measure defined in Definition 3.1 to the toy example with $\rho(X) = \text{VaR}_{75\%}(X)$.

3.4 Credibility-based bounds

In the literature, risk bounds are based on assigning either full or zero credibility to the assumptions on which the adopted model is built. As a result, the bounds are usually either very wide (because of trusting very few assumptions) or non-realistic (because of assigning full credibility to assumptions that cannot be fully trusted). A solution to this problem would be to assign partial credibility to the assumptions and calculate the bounds

accordingly. This can be expressed formally in the following analysis.

Our analysis is based on having a non-null set of assumptions $\{a_i\}_{i \leq r}$ that can be easily considered as almost sure, i.e., $P(\{a_i\}_{i \leq r} \text{ are correct}) = 1$. Hence, for an adopted model F^* and a set of assumptions $\{a_i\}_{i \leq n}$, we are able to compute $n - r + 1$ upper bounds, i.e., we can calculate the decreasing sequence of upper bounds estimates $\{\bar{\rho}_i\}_{n-r+1 \leq i \leq n}$ where $\bar{\rho}_k$ is the corresponding risk value for the set of assumptions $\{a_i\}_{i \leq k}$. We see these values as the possible realizations of a random variable UB and we aim to estimate the mean of UB. We find that

$$\begin{aligned}
 P(\text{UB} \leq \bar{\rho}_k) &= P(\{a_i\}_{i \leq k} \text{ are correct}) \\
 &= \begin{cases} P(a_k \text{ is correct} \setminus \{a_i\}_{i \leq k-1} \text{ are correct}) \\ \quad \times P(\{a_i\}_{i \leq k-1} \text{ are correct}) & \text{for } k \in \{r+1, \dots, n\}, \\ 1 & \text{for } k \in \{1, \dots, r\}, \end{cases} \\
 &= \begin{cases} \prod_{j=r+1}^k P(a_j \text{ is correct} \setminus \{a_i\}_{i \leq j-1} \text{ are correct}) & \text{for } k \in \{r+1, \dots, n\}, \\ 1 & \text{for } k \in \{1, \dots, r\}. \end{cases}
 \end{aligned}$$

Indeed, $P(a_j \text{ is correct} \setminus \{a_i\}_{i \leq j-1} \text{ are correct})$ can be seen as a conditional credibility factor and can be denoted as z_j . Thus, the cumulative distribution function of UB can be expressed as

$$P(\text{UB} \leq \bar{\rho}_k) = \begin{cases} \prod_{j=r+1}^k z_j & \text{for } k \in \{r+1, \dots, n\}, \\ 1 & \text{for } k \in \{1, \dots, r\}. \end{cases} \quad (3.5)$$

After specifying the credibility factors, the cumulative distribution function of UB can be calculated and therefore the $E[\text{UB}]$ can be determined; we denote it as the credibility-based upper bound (CUB). The same analysis can be performed to determine the credibility-based lower bounds (CLB). We provide explicit expressions for the CUB and CLB in the

following definition.

Definition 3.2 (Credibility-based upper and lower bounds). Let $\rho : \mathcal{M} \rightarrow \mathbb{R}$ be a risk measure. Let F^* be the cumulative distribution function of the adopted model and let $\{a_i\}_{i \leq n}$ a set of assumptions that characterizes F^* where $\{a_i\}_{i \leq r \leq n}$ can be fully trusted. Let us define a decreasing sequence of sets $\{\mathcal{A}_k\}_{k \leq n}$ with $\mathcal{A}_k = \{X \in \mathcal{M} \mid X \text{ respects assumptions } \{a_i\}_{i \leq k}\}$. Let us denote by z_j the credibility assigned to a_j given the knowledge of $\{a_i\}_{i \leq j-1}$. Then we can define the credibility-based upper bound CUB and lower bound CLB as follows,

$$\text{CUB}(\rho, \{\mathcal{A}_m\}_m, \{z_j\}_j) = \bar{\rho}_r + \sum_{m=r+1}^n \left(\prod_{j=r+1}^m z_j \right) (\bar{\rho}_m - \bar{\rho}_{m-1}), \quad (3.6)$$

and

$$\text{CLB}(\rho, \{\mathcal{A}_m\}_m, \{z_j\}_j) = \underline{\rho}_r + \sum_{m=r+1}^n \left(\prod_{j=r+1}^m z_j \right) (\underline{\rho}_m - \underline{\rho}_{m-1}), \quad (3.7)$$

where $\bar{\rho}_j = \sup_{X \in \mathcal{A}_j} \rho(X)$ and $\underline{\rho}_j = \inf_{X \in \mathcal{A}_j} \rho(X)$.

The credibility factors $\{z_j\}_j$ are to be assessed/specified differently according to each type of assumptions. For example, a specific statistical test can be used to show how confident we are about the unimodality property (e.g., the dip test of unimodality, see Hartigan et al. (1985) and Hartigan (1985)), another test can show how trustworthy a specific parameter estimation is (e.g., hypothesis testing), and yet another test is used to see how credible the normality assumption is (e.g., the numerous normality tests), and so on. Nevertheless, the assessment of credibility factors based on statistical tests or expert opinion is considered out of the scope of this paper.

Remark 3.3. The traditional approach to risk bounds can be seen as a particular case of the credibility-based approach where the credibility factors $\{z_j\}_j$ are dummy variables

(i.e., can take only the values 0 or 1).

Illustration Let us elaborate on our toy example introduced in Section 3.3. We specify the conditional credibility factors in Table 3. The sequence of fully trusted initial assumptions in this case is $\{a_1, a_2\}$, i.e., the intervals on the mean and the variance, and hence we have $r = 2$. The adopted model is reached after adding the assumption a_7 , we thus have $n = 7$. The credibility-based bounds would then be $\text{CLB} = A_7 = 5.69$ and $\text{CUB} = B_7 = 21.09$.

i	1	2	3	4	5	6	7
a_i	μ	s	U	+	GA	E	Adopted model
z_i	-	100%	90%	100%	50%	60%	90%

Table 3: Conditional credibility factors assigned to each of the assumptions of interest.

k	$(\underline{\rho}_k, \bar{\rho}_k)$	z_k	$\prod_{j=r+1}^k z_j$	$(\underline{\rho}_k - \underline{\rho}_{k-1}, \bar{\rho}_k - \bar{\rho}_{k-1})$	(A_k, B_k)
2	(-0.08, 36.25)	100%	-	-	-
3	(1.27, 27.87)	90%	90%	(1.35, -8.38)	(1.14, 28.71)
4	(1.27, 23.68)	100%	90%	(0.00, -4.19)	(1.14, 24.94)
5	(8.11, 16.64)	50%	45%	(6.84, -7.04)	(4.22, 21.77)
6	(11.09, 16.64)	60%	27%	(2.98, 0.00)	(5.02, 21.77)
7	(13.86, 13.86)	90%	24%	(2.77, -2.78)	(5.69, 21.09)

Table 4: An application of the credibility-based risk bounds defined in Definition 3.2. In this example, $A_k = \underline{\rho}_r + \sum_{m=r+1}^k \left(\prod_{j=r+1}^m z_j \right) (\underline{\rho}_m - \underline{\rho}_{m-1})$ and $B_k = \bar{\rho}_r + \sum_{m=r+1}^k \left(\prod_{j=r+1}^m z_j \right) (\bar{\rho}_m - \bar{\rho}_{m-1})$.

Remark 3.4. Table 4 helps to look at the reasoning behind Definition 3.2 from a different angle. In the first column, the bounds improve (i.e., become tighter) when assumptions are added (i.e., when k increases). However, this improvement (shown in the fourth column) cannot be realized when adding assumptions that cannot be fully trusted. A reasonable

way of incorporating these improvements is to adjust each improvement according to the credibility (shown in the third column) that the corresponding added assumption holds. This perspective automatically leads to the bounds (A_k, B_k) shown in the fifth column.

3.5 Model risk measures

One of the ultimate objectives of assessing model risk is to end up with a possible buffer that can be used to calculate the corresponding capital requirement. In this section, we propose an intuitive formula that can be used for that purpose.

Two popular measures of model risk are the absolute measure and the relative measure of model risk defined in Barrieu and Scandolo (2015). The absolute measure reflects the position of the value of the risk measure applied to the adopted model compared to the upper risk bound derived based on the scenario of fully trusted assumptions, whereas the relative measure reflects the position of the adopted model compared to both the lower and the upper risk bounds. In formal terms, if $X^* \sim F^*$, and $\{a_i\}_{i \leq r}$ is the set of fully trusted assumptions, then the two measures are defined as follows.

Definition 3.3 (Absolute measure of model risk, Barrieu and Scandolo (2015)).

$$\text{AM}(\rho, \mathcal{A}_r, X^*) = \frac{\bar{\rho}_r - \rho(X^*)}{\rho(X^*)}. \quad (3.8)$$

Definition 3.4 (Relative measure of model risk, Barrieu and Scandolo (2015)).

$$\text{RM}(\rho, \mathcal{A}_r, X^*) = \frac{\bar{\rho}_r - \rho(X^*)}{\bar{\rho}_r - \underline{\rho}_r}. \quad (3.9)$$

However, in practice, $\bar{\rho}_r$ and $\underline{\rho}_r$ are very different from $\rho(X^*)$ (i.e., the risk bounds are wide), meaning that in many cases of interest these two model risk measures not very

informative. We thus propose to use the credibility-based upper and lower bounds in order to extend these definitions to the credibility-based absolute and relative measures of model risk as follows.

Definition 3.5 (Credibility-based absolute measure of model risk).

$$\text{CAM}(\rho, \text{CUB}, \text{CLB}, X^*) = \frac{\text{CUB} - \rho(X^*)}{\rho(X^*)}. \quad (3.10)$$

Definition 3.6 (Credibility-based relative measure of model risk).

$$\text{CRM}(\rho, \text{CUB}, \text{CLB}, X^*) = \frac{\text{CUB} - \rho(X^*)}{\text{CUB} - \text{CLB}}. \quad (3.11)$$

The four measures are positive increasing functions of model risk with a value of 0 for no model risk. CRM, similarly to RM, is unitless and reaches 1 when the model risk is maximal. Indeed, CRM incorporates information on the worst-case and best-case models, on the adopted model, and on the credibility assigned to each assumption. Hence, it would be interesting to incorporate the CRM as a factor in the model risk capital formula. In addition, the difference between the CUB and the CLB represents the maximum model risk capital that could be required when using a model that adopts the corresponding assumptions and credibility factors. These ideas lead us to Definition 3.7.

Definition 3.7 (Model risk capital). For a continuous increasing function $f : [0, 1] \rightarrow [0, 1]$, with $f(0) = 0$ and $f(1) = 1$, we can define the model risk capital (MoRC) by

$$\text{MoRC}(\text{CRM}, \text{CUB}, \text{CLB}, f) = f(\text{CRM}) \times (\text{CUB} - \text{CLB}). \quad (3.12)$$

We can look at $f(\text{CRM})$ as the percentage of the maximum capital that can be allocated

to model risk, starting from 0 for no model risk and reaching 100% for full model risk. The regulator and the model risk manager decide on the degree of conservatism towards model risk. This can be translated into the choice of f , i.e., how the percentage of the maximum capital increases with the increase in model risk. One suggestion would be to use the convex function $f(x) = x^n$, for $n \geq 1$; then the higher the n , the less conservative the MoRC. Indeed, any continuous increasing function g can lead to an admissible function $f = \frac{g-g(0)}{g(1)-g(0)}$.

Illustration In our toy example, $\text{CAM} = \frac{21.09-13.86}{13.86} = 52.16\%$ and $\text{CRM} = \frac{21.09-13.86}{21.09-5.69} = 46.96\%$. If we choose $f(x) = x^2$, then $\text{MoRC} = 3.4$.

Sometimes it is interesting to compare the model risks of two possible models. This can be done by comparing the corresponding CAMs and CRMs. In addition, noting that the CUBs and CLBs are basically model specific, then it can be meaningful to compare the width of the credibility-based bounds of two different models adopted for the same dataset; a higher difference is an indicator of a higher uncertainty in the model.

Interestingly, the more assumptions are challenged and the more credibility a modeler can assign to his/her assumptions, the lower CAM and (CUB - CLB) are expected to be. This fact encourages the modeler to strive for higher credibility in his/her assumptions and to challenge as many assumptions as possible.

Remark 3.5. The measures presented in definitions 3.3, 3.4, 3.5, and 3.6 focus on the risk of underestimation of the risk measure. Even though this is the case that calls for a buffer in the capital requirement, complimentary measures can be easily constructed to reflect the risk of overestimation of the risk measure.

4 Case study: SOA Medical dataset

In this section, we present an application of the ideas developed in this paper to an SOA Group Medical Insurance Large Claims Database described thoroughly in Grazier (1997). The data is collected from 26 insurers and covers the total claim amounts exceeding \$25 000 over the year 1991. We study the total claim amounts as a univariate variable, which makes it possible to challenge the two different univariate models suggested to this same dataset by two scientific papers, namely Cebrián et al. (2003) and Zisheng and Chi (2006).

4.1 Data and model description

The dataset is composed of 75 789 observations. The average total claim amount is \$58 413 and the largest observed total claim is \$4 518 420. The standard deviation among the total claims is \$66 005, and the Value-at-Risk at a probability level of 99.5% is \$406 190.

Both Cebrián et al. (2003) and Zisheng and Chi (2006) adopted an Extreme Value Theory (EVT) perspective and fit a Generalized Pareto Distribution (GPD), which is known as the “natural” distribution for modeling excess-of-loss over high thresholds.

Let S denote the random variable of the total claim and u the threshold after which the data is fit to a GPD, and let $G_{\zeta,\theta,\lambda}$ represent the Generalized Pareto distribution function with a shape parameter ζ , a location parameter θ , and a scale parameter λ . Then, one can easily prove that $S|S \geq u \sim G_{\zeta,u,\lambda}$ implies $\text{VaR}_\alpha(S) = u + \frac{\lambda}{\zeta} \left[\left(\frac{1-\alpha}{\hat{P}(S>u)} \right)^{-\zeta} - 1 \right]$ for $\alpha \geq 1 - \hat{P}(S > u)$ (see Appendix B for more details on the GPD). In order to fit the GPD model to the dataset, one has to choose a threshold u and then fits the GPD to the conditional distribution of the excesses above the threshold u . Typically, $\hat{P}(S > u)$ is calculated empirically.

Cebrián et al. (2003) found that the best choice for the threshold is $u_1 = 200\,000$, which

gives 2013 exceedances. The estimated parameters are $\zeta_1 = 0.314$ and $\lambda_1 = 93\,901$. The mean, standard deviation, and Value-at-Risk at 99.5% of S under the adopted model are $E_1[S] = 58\,405$, $\text{std}_1[S] = 66\,178$, and $\text{VaR}_{99.5\%,1} = 406\,161$.

On the other hand, Zisheng and Chi (2006) chose a threshold of $u_2 = 162\,402$, which gives 3083 exceedances and leads to the estimated parameters $\zeta_2 = 0.311962$ and $\lambda_2 = 82\,652.07$. The mean, standard deviation, and Value-at-Risk at 99.5% of S under this model are $E_2[S] = 58\,422$, $\text{std}_2[S] = 66\,110$, and $\text{VaR}_{99.5\%,2} = 406\,928$.

4.2 Model risk allocation

The first step in the model risk assessment is to disassemble the assumptions upon which the adopted model was built and try to assess the model risk contribution of each assumption of interest. Using the tools provided in the literature (many of which are stated in Figures 1, 2, and 3), some of the assumptions that are of interest to challenge are questioned as follows:

- Given the threshold, what is the parameter risk in the estimation of the scale and shape parameters of the GPD?
- Given a GPD is adopted, what is the parameter risk in choosing the threshold?
- Given mean, variance, non-negativity, and unimodality, how risky is it, in terms of model risk, to choose a GPD?
- How much does each of the assumptions on the moments, non-negativity, and unimodality contribute to the total model risk of the adopted model?

The diagram in Figure 4 shows some possible paths from two basic scenarios to the adopted models. Motivated by the large number of observations, we calculate the interval

on the mean and the maximum standard deviation based on the standard confidence interval procedure. Indeed, μ' refers to the assumption that the average total claim amount belongs to the interval $(58\,413 - 1.96 \frac{66\,005}{\sqrt{75\,789}}, 58\,413 + 1.96 \frac{66\,005}{\sqrt{75\,789}}) \simeq (57\,940, 58\,880)$, and s refers to the assumption that the standard deviation is lower or equal than 66 339 calculated based on the formula of the upper limit presented on pages 197-198 of Sheskin (2003).

The value of the maximum third moment is calculated by bootstrapping; we simulated 100 000 samples of 10 000 data points each taken from the set of 75 789 observations, we calculated the third moment of each bootstrap sample, we then calculated the third quartile of the set of third moment values and adopted the third quartile as the upper limit for the third moment. Indeed, d_3 refers to the assumption that the upper limit of the third moment of the portfolio loss random variable is equal to 4.78×10^{15} .

The feature of unimodality can be easily detected from the data and in the two adopted models when complemented by the empirical distribution for the values that are lower than the threshold (the GPD model is usually fitted to the tail of the distribution starting at the threshold, whereas the rest of the distribution is usually modeled empirically).

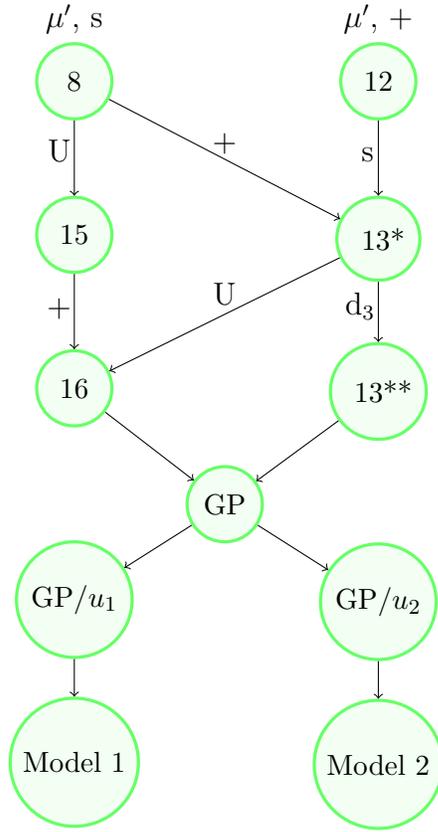
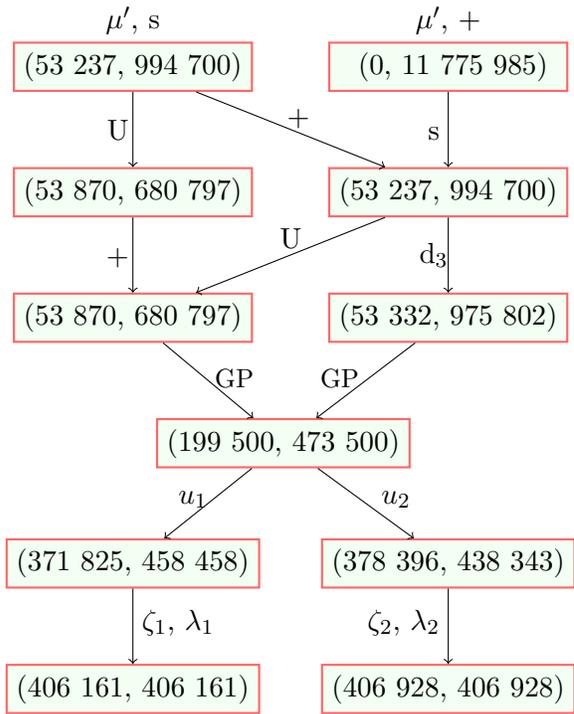


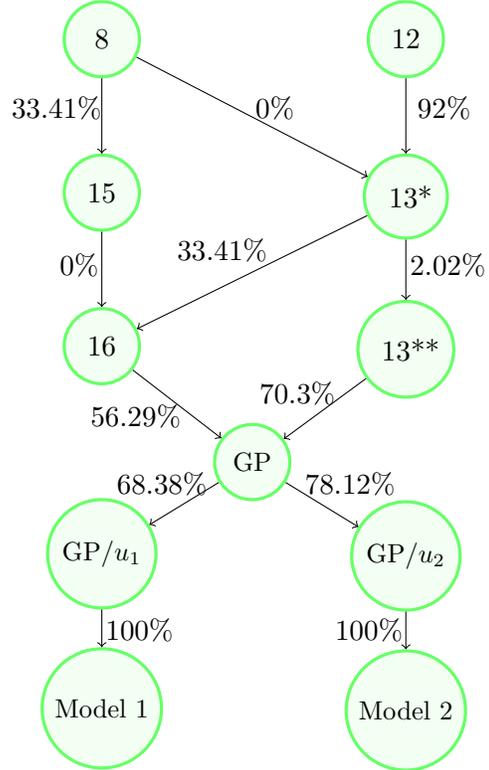
Figure 4: Diagram showing the link between the scenarios that underlie a GPD model. The scenarios consist in assumptions made on the portfolio loss considered as a univariate random variable. Each number refers to a scenario for which the bounds are derived and presented in the appendix by a theorem of the same number. The considered assumptions all refer to the portfolio loss random variable and are the assumptions of knowledge on: an interval of the mean (μ'), a maximum standard deviation (s), non-negativity ($+$), unimodality (U), positive portfolio loss value ($+$), a maximum value for the third moment (d_3), fitting a GPD (GP), choosing the threshold of Cebrián et al. (2003) for the GPD (GP/u_1), choosing the threshold of Zisheng and Chi (2006) for the GPD (GP/u_2), adopting the model of Cebrián et al. (2003) (Model 1), and adopting the model of Zisheng and Chi (2006) (Model 2). The symbols above the root nodes refer to the set of assumptions adopted in the corresponding nodes, the notations on each arrow refers to the set of assumptions added when following this arrow. The asterisks on 13^* and 13^{**} respectively refer to the fact that Theorem 13 was restricted to the case where only the maximum values of the second moment or the second and third moment are known.

Remark 4.1. It is important to note that the adopted methods in the estimation of the moments intervals are just a choice out of many others; indeed, this case study does not aim to adopt the best estimation methods but rather to solely assess the model risk.

The adopted risk measure in this case study is the one that is most used in the calculation of capital requirement in Solvency II, the Value-at-Risk at a probability level of 99.5% ($VaR_{99.5\%}$). In the diagrams of Figure 5, the risk bounds and the conditional model risk contributions under the various assumptions are presented. At this step, one can directly make several observations:



(a) Diagram showing the risk bounds in each of the scenarios presented in Figure 4. Each box contains the corresponding interval $(\underline{\rho}_k, \bar{\rho}_k)$. Each symbol on the arrows represents the added assumption(s) between a pair of scenarios, these symbols are consistent with the notation presented in this section and in Figure 4.



(b) Diagram showing the conditional model risk contributions of each added assumption between the scenarios presented in Figure 4. The percentages on the arrows represent the values of the conditional model risk contributions $\mathcal{C}(\rho, \mathcal{A}_k, \mathcal{A}_l)$ calculated based on Definition 3.1.

Figure 5: Diagrams showing the $\text{VaR}_{99.5\%}$ bounds and the corresponding conditional model risk contributions that are relative to the scenarios shown in Figure 4.

- The non-negativity assumption does not contribute to the model risk (0%).
- Assuming a maximum value for the third moment after already having information on the first two moments is not risky in terms of model risk (2.02%).
- Adding information on the variance after having already assumed an interval for the mean has a great effect on the model risk and should be done cautiously (92%).

- Being able to trust the unimodality feature of the data gives much more confidence in choosing the GPD compared to being able to trust some information on the third moment (56.29% vs 70.3%).
- Even when the information on the moments, non-negativity, and unimodality is trusted, the choice of GPD still contributes significantly in the total model risk (56.29%).
- Choosing a threshold after having already chosen to adopt a GPD model can have a relatively higher contribution to the model risk than choosing the GPD itself after having already trusted the unimodality property and some information on the moments (e.g., 78.12% vs 56.29%).
- The threshold chosen by Zisheng and Chi (2006) is a stronger assumption (in terms of model risk) than the one chosen by Cebrián et al. (2003) (78.12% vs 68.38%).

The conditional model risk contribution of the last assumption that leads to the adopted model is by definition 100% and hence is not very helpful in interpreting specific assumptions. For this case, one can use the relative measure of model risk (RM) defined in Definition 3.4. The parameter risk inherent in choosing the couple (ζ_1, λ_1) after having already chosen the threshold u_1 can be presented by $RM_1 = \frac{458}{458} \frac{458-406}{458-371} \frac{161}{825} = 60.37\%$, whereas in the case of Model 2 we have $RM_2 = 52.4\%$. This directly implies that the parameter risk in choosing the scale and the shape parameters of Model 1 after having already chosen the threshold of Model 1 is higher than the one in choosing the scale and shape parameters of Model 2 after having already chosen the threshold of Model 2.

4.3 Model risk measurement

The next step in our model risk assessment framework is to assess the model risk in the model as a whole. This can be done using the credibility-based bounds defined

in Definition 3.2. We first choose the assumptions (a_i) to be used in the calculation and assign the corresponding conditional credibility factors (z_i) . Based on the results obtained so far, a meaningful set of assumptions that leads to the adopted model is $\{a_i\}_i = \{(\mu', s), U, GP, u, (\zeta, \lambda)\}$.

Before proceeding to the assessment of credibility factors, it should be noted that this case study does not aim to show the best way of assigning credibility factors but rather to give a simplistic illustration of how the framework works.

The interval on the mean and the maximum variance are both calculated based on a 95% confidence level, so it is not unreasonable to start with this information as the fully trusted basic assumptions.

The sample of observations clearly features unimodality; this can be even verified by performing some unimodality tests (e.g., the dip test of unimodality, see Hartigan et al. (1985) and Hartigan (1985)). Hence, one can confidently give a 95% credibility to the unimodality property.

Our risk measure is evaluated at the very end of the tail, at a probability level of 99.5%, which makes the GPD a good choice for the model, however, we have seen before that the choice of a GPD contributes efficiently to the total model risk even when we have knowledge on the moments and unimodality. Therefore, a 50% conditional credibility factor seems reasonable for the GPD assumption.

In order to choose the threshold, Cebrián et al. (2003) used the Gertensgarbe plot proposed in Gerstengarbe and Werner (1989) whereas Zisheng and Chi (2006) used the Goodness-of-fit test for the GPD developed in Choulakian and Stephens (2001). A statistician would have to compare the two tests and assign the corresponding conditional credibility factors. In this illustrative example, we chose to correlate the credibility of the

method with the number of times it was cited. The Goodness-of-fit test for the GPD is currently cited 390 times whereas the Gerstengarbe plot is cited only 24 times, and hence, we would consider Goodness-of-fit test more credible. The Gertensgarbe plot and the Goodness-of-fit test are respectively given 50% and 75% as conditional credibility factors.

The scale and shape parameters are estimated in both models using the maximum likelihood estimation. However, the parameters are estimated based on 2013 and 3083 data points in Cebrián et al. (2003) and Zisheng and Chi (2006), respectively. Hence, the estimates of the second model are more reliable, and we can give conditional credibility factors of 60% and 80% to the estimates of Model 1 and Model 2, respectively. A summary is presented in Table 5 and Table 6.

i	1	2	3	4	5
Assumption a_i	μ', s	U	GP	u_1	ζ_1, λ_1
Conditional credibility factor z_i	100%	95%	50%	50%	60%

Table 5: Conditional credibility factors assigned to each of the assumptions of Model 1.

i	1	2	3	4	5
Assumption a_i	μ', s	U	GP	u_2	ζ_2, λ_2
Conditional credibility factor z_i	100%	95%	50%	75%	80%

Table 6: Conditional credibility factors assigned to each of the assumptions of Model 2.

We can now apply the formula for the credibility-based bounds of Definition 3.2 and obtain: $CLB_1 = 168\ 833$, $CUB_1 = 587\ 001$, $CLB_2 = 194\ 876$, and $CUB_2 = 576\ 548$. Hence, the credibility-based absolute and relative measures of model risk for the two models are: $CAM_1 = 44.52\%$, $CRM_1 = 43.25\%$, $CAM_2 = 41.68\%$, and $CRM_2 = 44.44\%$. The widths of credibility-based bounds of the two models are $\Delta_1 = CUB_1 - CLB_1 = 418\ 168$ and $\Delta_2 = CUB_2 - CLB_2 = 381\ 672$. The two CRMs are very close, but the comparisons of the

CAMs and the Δ s show that Model 2 has the least model risk.

Finally, we calculate the suggested MoRC for each of the two models. To theoretically eliminate the model risk, one should add a buffer that, as a percentage of the risk value, is equal to the CAM. However, the CAM is quite high in our case and amounts to more than 40% of the value of the risk measure. Additionally, the CAM may misrepresent the model risk since it does not account for the position of the risk value compared to the CLB. A solution would be to adopt the MoRC in Definition 3.7 with a convex function f . The choice of f is based on how conservative the risk management team or the regulators are. If $f(x) = x^n$, for $n \geq 1$, then the higher n the less security is required. If we take $f(x) = x^3$ for example, we get $\text{MoRC}_1 = 33,821$ and $\text{MoRC}_2 = 33,500$, i.e., respectively 8.33% and 8.23% of the risk values of the two adopted models.

5 Conclusions

In this paper, we establish a practical framework for quantitative model risk assessment that builds on the literature of risk bounds (theory of model risk). Firstly, we disassemble an adopted model into a set of assumptions and use our novel model risk contribution measure in order to allocate the model risk to the various assumptions. By doing so, we aim to enlighten the modeler on how cautious he is expected to be when making every assumption in the model building process. Secondly, we acknowledge the fact that every single assumption can have its own level of credibility and we incorporate this information into the model risk assessment. Thirdly, we define new measures of model risk that can be used for model risk capital allocation. Lastly, we conduct a case study in which we apply our framework to a real-world dataset, the SOA Group Medical Insurance Large Claims Database.

Our framework incorporates previous findings from the literature on risk bounds and is built in a way to embrace future findings in this currently active research area.

Appendix A Definitions and notations

Basic notations We consider a portfolio of individual risks with a portfolio loss variable $S = \sum_{i=1}^n X_i$, where X_i is the i^{th} individual loss and n is the number of individual risks. We write $X_i \sim F_i$ to express that F_i is the marginal distribution of X_i , and we denote the vector of n individual risks by $\mathbf{X} = (X_1, X_2, \dots, X_n)$.

Our approach to model risk assessment is based on the derivation of upper and lower risk bounds in each of the scenarios of interest. Specifically, for the Value-at-Risk (VaR) defined in 3.3, and for $\alpha \in (0, 1)$, we denote the risk bounds as follows,

$$\overline{\text{VaR}}_\alpha^v = \sup_{S \in V} \text{VaR}_\alpha(S), \quad \underline{\text{VaR}}_\alpha^v = \inf_{S \in V} \text{VaR}_\alpha(S), \quad (\text{A.1})$$

where V stands for the set of random variables that respect a given set of assumptions denoted by v .

Sharpness of bounds A risk bound is called sharp if it is attainable, i.e., if there exists a random variable in the set of plausible random variables for which the risk measure reaches the value of the bound. In some scenarios, the sharp bounds $\overline{\text{VaR}}_\alpha^v$ and $\underline{\text{VaR}}_\alpha^v$ are hard to derive analytically and thus are replaced by non-sharp bounds. In order to underline the difference, we denote the not-necessarily-sharp bounds by $\overline{\overline{\text{VaR}}}_\alpha^v$ and $\underline{\underline{\text{VaR}}}_\alpha^v$.

Risk partitions Let I_1, \dots, I_n be mutually exclusive sets. We denote by $\bigcup_{i=1}^k I_i = \{1, \dots, n\}$ a partition and by $\mathbf{X}_{I_i} := (X_j, j \in I_i)$ the i^{th} subgroup of the risk vector \mathbf{X} .

Specific dependence structures We denote by $S^c = \sum_{i=1}^n X_i^c$ the comonotonic portfolio sum with $X_i^c \sim F_i$. We denote by $S^\perp = \sum_{i=1}^n X_i^\perp$ the independent portfolio sum with $X_i^\perp \sim F_i$. We denote by $S_i = \sum_{j \in I_i} X_j$ the sum of a subgroup I_i of individual losses, and by $S^{c,k} = \sum_{i=1}^k S_i^c$ the comonotonic version of the portfolio sum $S = \sum_{i=1}^k S_i$.

Additional risk measures Throughout the paper, we concentrate on assessing model risk using the Value-at-Risk. However, to do so, we need to refer to two other risk measures: the Tail Value-at-Risk (TVaR) and the Left Tail Value-at-Risk (LTVaR) that are defined

as follows,

$$\text{TVaR}_\alpha(S) = \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(S) du, \quad \alpha \in (0, 1), \quad (\text{A.2})$$

and

$$\text{LTVaR}_\alpha(S) = \frac{1}{\alpha} \int_0^\alpha \text{VaR}_u(S) du, \quad \alpha \in (0, 1). \quad (\text{A.3})$$

Dependence intensity A simple way to describe the positive dependence structure of \mathbf{X} is by the notion of positive orthant dependence (POD): \mathbf{X} is called POD if

$$\forall \mathbf{x} \in \mathbb{R}^d, \begin{cases} P(\mathbf{X} \leq \mathbf{x}) \geq \prod_{i=1}^n P(X_i \leq x_i), \\ P(\mathbf{X} > \mathbf{x}) \geq \prod_{i=1}^n P(X_i > x_i), \end{cases} \quad (\text{A.4})$$

(see for instance Puccetti and Rüschendorf (2012a)).

A stronger notion of positive dependence is the sequential positive cumulative dependence (PCD) defined in Rüschendorf (2017) as a sequential version of the initial PCD defined in Denuit et al. (2001). \mathbf{X} is sequential PCD if

$$P\left(\sum_{i=1}^{k-1} X_i > t_1 | X_k > t_2\right) \geq P\left(\sum_{i=1}^{k-1} X_i > t_1\right) \quad \forall k \in [2, n], \forall t_1, t_2 \in \mathbb{R}. \quad (\text{A.5})$$

Similarly, \mathbf{X} is sequential negative cumulative dependence (NCD) if

$$P\left(\sum_{i=1}^{k-1} X_i > t_1 | X_k > t_2\right) \leq P\left(\sum_{i=1}^{k-1} X_i > t_1\right) \quad \forall k \in [2, n], \forall t_1, t_2 \in \mathbb{R}. \quad (\text{A.6})$$

Factor model In a factor model, the individual risks X_i are expressed in a functional form as

$$X_i = f_i(Z, \epsilon_i), \quad 1 \leq i \leq n, \quad (\text{A.7})$$

where ϵ_i are idiosyncratic risk components and Z is a common risk factor taking values in a set $D \subset \mathbb{R}^d$. A partially specified factor model refers to the case where the marginal distributions F_i , the conditional distributions of X_i given $Z = z$ denoted by $F_{i|z}$, and the joint distribution of (X_i, Z) denoted by H_i , are all known. We denote by $S_z = \sum_{i=1}^n X_{i,z}$

the sum of the conditional marginal variables X_i given $Z = z$, and by S_z^c the corresponding comonotonic sum.

Appendix B Properties of the Beta and Generalized Pareto distributions

Beta distribution If $X \sim \text{Beta}(a, b)$ with $a > 0$ and $b > 0$, then the density function of X is given by

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \text{ for } 0 < x < 1,$$

where $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$ for $x > 0$. The first two central moments of X are $E[X] = \frac{a}{a+b}$ and $\text{var}[X] = \frac{ab}{(a+b)^2(a+b+1)}$.

Generalized Pareto distribution If $X \sim \text{GPD}(\zeta, \theta, \lambda)$ with $\zeta \in \mathbb{R}$, $\theta \in \mathbb{R}$ and $\lambda > 0$, then the cumulative distribution function of X is given by

$$G_{\zeta, \theta, \lambda}(x) = \begin{cases} 1 - \left(1 + \frac{\zeta(x-\theta)}{\lambda}\right)^{-\frac{1}{\zeta}} & \text{for } \zeta \neq 0, \\ 1 - \exp\left(-\frac{x-\theta}{\lambda}\right) & \text{for } \zeta = 0, \end{cases} \quad (\text{B.1})$$

for $x \geq \theta$ when $\zeta \geq 0$, and $\theta \leq x \leq \theta - \lambda/\zeta$ when $\zeta < 0$. The first two central moments of X are $E[X] = \theta + \frac{\lambda}{1-\zeta}$ for $\zeta < 1$, and $\text{var}[X] = \frac{\lambda^2}{(1-\zeta)^2(1-2\zeta)}$ for $\zeta < 1/2$.

Appendix C Value-at-Risk bounds

In this section, we present a review on the Value-at-Risk bounds derived in the literature. Although this section is designed to be read on its own, the theorems presented here are ordered based on their position in the literature review in Section 2.

In the banking and insurance practice, actuaries usually have better techniques and more data to estimate one-dimensional distributions compared to estimating an overall dependence structure. Hence, there is a tendency to give a full trust to the marginal distributions and distrust all additional assumptions. We first start by the bounds in this basic scenario where only the marginal distributions are known.

Theorem 1. (Bernard et al. (2017)) (VaR bounds with given marginals) Let $\alpha \in (0, 1)$ and $X_i \sim F_i$. Then,

$$\text{VaR}_\alpha(S) \leq \overline{\text{VaR}}_\alpha^m := \text{TVaR}_\alpha(S^c) = \sum_{i=1}^n \text{TVaR}_\alpha(X_i), \quad (\text{C.1})$$

$$\text{VaR}_\alpha(S) \geq \underline{\text{VaR}}_\alpha^m := \text{LTVaR}_\alpha(S^c) = \sum_{i=1}^n \text{LTVaR}_\alpha(X_i). \quad (\text{C.2})$$

The individual risks of portfolios faced in risk management do often show some positive dependence structure. We present in the following three theorems some bounds derived in Bigozzi et al. (2015) and Rüschendorf (2017).

Theorem 2. (VaR bounds with given marginals and positive orthant dependence) Let $\alpha \in (0, 1)$ and $X_i \sim F_i$, and \mathbf{X} is positive orthant dependent (POD). Then,

$$\underline{\text{VaR}}_\alpha^{m,pod} := \max_{1 \leq i \leq n} F_i^{-1}(\alpha) \leq \text{VaR}_\alpha(S) \leq \sum_{i=1}^n F_i^{-1}(\alpha^{\frac{1}{n}}) =: \overline{\text{VaR}}_\alpha^{m,pod}. \quad (\text{C.3})$$

If we replace the POD by the stronger notion of positive dependence, the sequential positive cumulative dependence (PCD), we get the following bounds.

Theorem 3. (VaR bounds with given marginals and sequential positive cumulative dependence) Let $\alpha \in (0, 1)$ and $X_i \sim F_i$, and \mathbf{X} is sequential positive cumulative dependent (PCD). Then,

$$\underline{\text{VaR}}_\alpha^{m,pcd} := \text{LTVaR}_\alpha(S^\perp) \leq \text{VaR}_\alpha(S) \leq \sum_{i=1}^n \text{TVaR}_\alpha(X_i) = \overline{\text{VaR}}_\alpha^m \quad (\text{C.4})$$

Some types of insurance products can show some negative dependence, e.g., weather related insurance products. Hence, an assumption of negative dependence may be realistic.

Theorem 4. (VaR bounds with given marginals and sequential negative cumulative dependence) Let $\alpha \in (0, 1)$ and $X_i \sim F_i$, and \mathbf{X} is sequential negative cumulative dependent (NCD). Then,

$$\underline{\text{VaR}}_\alpha^m = \sum_{i=1}^n \text{LTVaR}_\alpha(X_i) \leq \text{VaR}_\alpha(S) \leq \text{TVaR}_\alpha(S^\perp) =: \overline{\text{VaR}}_\alpha^{m,ncd} \quad (\text{C.5})$$

Additionally, some subgroups of the portfolio can prove to be independent. This relation can be easily checked statistically and by intuition. In the following theorems we present VaR bounds derived in Puccetti et al. (2017) in scenarios where all or some of the subgroups of the portfolio are assumed independent.

Theorem 5. (VaR bounds with given marginals and independent subgroups) Let $\alpha \in (0, 1)$, $X_i \sim F_i$, and $\bigcup_{i=1}^k I_i = \{1, \dots, n\}$ a partition. If the risk subvectors $\mathbf{X}_{I_1}, \dots, \mathbf{X}_{I_k}$ are independent, then,

$$\underline{\text{VaR}}_\alpha^{m,ind} := \text{LTVaR}_\alpha(S^{c,k}) \leq \text{VaR}_\alpha(S) \leq \text{TVaR}_\alpha(S^{c,k}) =: \overline{\text{VaR}}_\alpha^{m,ind}. \quad (\text{C.6})$$

Note that $\underline{\text{VaR}}_\alpha^{m,ind}$ and $\overline{\text{VaR}}_\alpha^{m,ind}$ are TVaR and LTVaR functions of independent sums, and hence can be obtained by Monte Carlo simulation.

The following theorem considers the case where only some of the risk subvectors are independent.

Theorem 6. (VaR bounds with given marginals and partial independent substructures) Let $\alpha \in (0, 1)$, $X_i \sim F_i$, $\bigcup_{i=1}^k I_i = \{1, \dots, n\}$ a partition, $H \subset \{1, \dots, k\}$, and $Y_i = \sum_{j \in I_i} X_j$. If the risk subvectors $\mathbf{X}_{I_i}, i \in H$ are independent, then,

$$\text{VaR}_\alpha(S) \leq \overline{\text{VaR}}_\alpha^{m,pind} := \sum_{i \notin H} \sum_{j \in I_i} \text{TVaR}_\alpha(X_j) + \text{TVaR}_\alpha\left(\sum_{i \in H} S_i^c\right), \quad (\text{C.7})$$

$$\text{VaR}_\alpha(S) \geq \underline{\text{VaR}}_\alpha^{m,pind} := \sum_{i \notin H} \sum_{j \in I_i} \text{LTVaR}_\alpha(X_j) + \text{LTVaR}_\alpha\left(\sum_{i \in H} S_i^c\right). \quad (\text{C.8})$$

Again, the bounds above can be obtained by Monte Carlo simulation since $\sum_{i \in H} S_i^c$ is an independent sum.

The following theorem considers the case where we can trust the marginal distributions of the individual risks and a maximum variance of the portfolio loss. The fact that, instead of taking an exact value, an upper bound is considered for the variance allows to capture the presence of statistical uncertainty in the second moment estimation.

Theorem 7. (Bernard et al. (2017)) (VaR bounds with given marginals and maximum

variance) Let $\alpha \in (0, 1)$, $X_i \sim F_i$ and $\text{var}[S] \leq s^2$. Then,

$$\text{VaR}_\alpha(S) \leq \overline{\overline{\text{VaR}}}_\alpha^{m,s} \min \left(\mu + s \sqrt{\frac{\alpha}{1-\alpha}}, \overline{\overline{\text{VaR}}}_\alpha^m \right), \quad (\text{C.9})$$

$$\text{VaR}_\alpha(S) \geq \underline{\underline{\text{VaR}}}_\alpha^{m,s} := \max \left(\mu - s \sqrt{\frac{1-\alpha}{\alpha}}, \underline{\underline{\text{VaR}}}_\alpha^m \right). \quad (\text{C.10})$$

Replacing the information on the marginals by the information on the collective mean alongside having information on the maximum variance simplifies the scenario with no considerable loss of information for the calculation on the bounds, especially for not “too large” maximum variances (Bernard et al. (2018)). The bounds derived in this scenario are known as the Cantelli bounds Bertsimas et al. (2004).

Theorem 8. (Bertsimas et al. (2004)) (VaR bounds with given mean and maximum variance) Let $\alpha \in (0, 1)$, $E[S] = \mu$ and $\text{var}[S] \leq s^2$. Then,

$$\underline{\underline{\text{VaR}}}_\alpha^{\mu,s} := \mu - s \sqrt{\frac{1-\alpha}{\alpha}} \leq \text{VaR}_\alpha(S) \leq \mu + s \sqrt{\frac{\alpha}{1-\alpha}} =: \overline{\overline{\text{VaR}}}_\alpha^{\mu,s}. \quad (\text{C.11})$$

Going back to the independent subgroups, a maximum variance assumption can be added leading to the following theorem derived in Puccetti et al. (2017).

Theorem 9. (VaR bounds with given marginals, partial independent substructures, and variance information) Let $\alpha \in (0, 1)$, $X_i \sim F_i$, $\bigcup_{i=1}^k I_i = \{1, \dots, n\}$ a partition, $H \subset \{1, \dots, k\}$, $Y_i = \sum_{j \in I_i} X_j$, and $\text{var}[S] \leq s^2$. If the risk subvectors $\mathbf{X}_{I_i}, i \in H$ are independent, then

$$\text{VaR}_\alpha(S) \leq \overline{\overline{\text{VaR}}}_\alpha^{m,pind,s} := \min \left(\mu + s \sqrt{\frac{\alpha}{1-\alpha}}, \overline{\overline{\text{VaR}}}_\alpha^{m,pind} \right), \quad (\text{C.12})$$

$$\text{VaR}_\alpha(S) \geq \underline{\underline{\text{VaR}}}_\alpha^{m,pind,s} := \max \left(\mu - s \sqrt{\frac{1-\alpha}{\alpha}}, \underline{\underline{\text{VaR}}}_\alpha^{m,pind} \right). \quad (\text{C.13})$$

Bernard et al. (2017) offers TVaR-based bounds for the Value-at-Risk of the portfolio risk in the partially specified factor model with and without variance information. Indeed, the TVaR-based representation of the bounds significantly simplifies the simulation process.

Theorem 10. (VaR bounds in a partially specified factor model) Let $\alpha \in (0, 1)$, $X_i \sim F_i$,

$X_{i,z} \sim F_{i|z}$, and $(X_i, Z) \sim H_i$. Then,

$$\underline{\underline{\text{VaR}}}_\alpha^{pf} := \text{VaR}_\alpha(\text{LTVaR}_V(S_Z^c)) \leq \text{VaR}_\alpha(S) \leq \text{VaR}_\alpha(\text{TVaR}_V(S_Z^c)) =: \overline{\overline{\text{VaR}}}_\alpha^{pf}, \quad (\text{C.14})$$

where $V \sim U(0, 1)$.

Theorem 11. (VaR bounds in a partially specified factor model with variance information) Let $\alpha \in (0, 1)$, $X_i \sim F_i$, $X_{i,z} \sim F_{i|z}$, $(X_i, Z) \sim H_i$, and $\text{var}[S|Z = z] \leq s^2(z)$. Then,

$$\text{VaR}_\alpha(S) \leq \overline{\overline{\text{VaR}}}_\alpha^{pf,s} := \text{VaR}_\alpha \left(\min \left(\mu(z) + s(z) \sqrt{\frac{V}{1-V}}, \text{TVaR}_V(S_Z^c) \right) \right), \quad (\text{C.15})$$

$$\text{VaR}_\alpha(S) \leq \underline{\underline{\text{VaR}}}_\alpha^{pf,s} := \text{VaR}_\alpha \left(\max \left(\mu(z) - s(z) \sqrt{\frac{1-V}{V}}, \text{LTVaR}_V(S_Z^c) \right) \right), \quad (\text{C.16})$$

where $V \sim U(0, 1)$ is independent of Z and $\mu(z) = E[S|Z = z]$.

In most situations, we can be sure about the range of values of the portfolio loss, e.g., we know that the portfolio loss is non-negative or we have information on the minimum and maximum loss. We start by presenting the VaR bounds for a non-negative portfolio loss of which we can trust only the mean (Bernard et al. (2017)). These bounds are straightforward to prove and are very wide; hence they are only useful for comparative purposes.

Theorem 12. (VaR bounds for non-negative random variables with given mean) Let $\alpha \in (0, 1)$, $E[S] = \mu$, and S is non-negative. Then,

$$\underline{\underline{\text{VaR}}}_\alpha^{+,\mu} := 0 \leq \text{VaR}_\alpha(S) \leq \frac{\mu}{1-\alpha} =: \overline{\overline{\text{VaR}}}_\alpha^{+,\mu}. \quad (\text{C.17})$$

The following theorem, derived in Bernard et al. (2018), presents the VaR bounds in a scenario where we know the range of values of the portfolio loss, the collective mean, and the maximum value of some higher moments. The idea of considering a maximum value for the moments instead of the exact values accounts for the statistical uncertainty of the moments estimation.

Theorem 13. (VaR bounds with given range of portfolio values, mean, and maximum moments.) Let $\alpha \in (0, 1)$, S valued in $[a, b]$, $E[S] = \mu$, and $E[S^k] \leq d_k$ for $k = 2, 3, \dots, r$.

Then,

$$\underline{\text{VaR}}_{\alpha}^{ab,\mu,d_k} := \mu - \psi \frac{1-\alpha}{\alpha} \leq \text{VaR}_{\alpha}(S) \leq \mu + \psi =: \overline{\text{VaR}}_{\alpha}^{ab,\mu,d_k}, \quad (\text{C.18})$$

where ψ is the maximum value in the interval $\left[0, \min \left\{ b - \mu, (\mu - a) \frac{\alpha}{1-\alpha} \right\} \right]$ such that the inequalities

$$(\mu + \psi)^k (1 - \alpha) + \left(\mu - \psi \frac{1-\alpha}{\alpha} \right)^k \alpha \leq d_k \quad (\text{C.19})$$

hold for $k = 2, \dots, r$.

In many practical situations, the first two moments of the individual risks can be trusted. Incorporating this information leads to the bounds derived in Bernard et al. (2018) and presented in the following theorem.

Theorem 14. (VaR bounds with given range of portfolio values, mean, maximum portfolio loss moments, and first two individual moments.) Let $\alpha \in (0, 1)$, S valued in $[a, b]$, $E[S] = \mu$, $E[S^k] \leq d_k$ for $k = 2, 3, \dots, r$, $E[X_i] = \mu_i$, and $\text{var}[X_i] = v_i^2$. Then,

$$\text{VaR}_{\alpha}(S) \leq \overline{\text{VaR}}_{\alpha}^{ab,\mu,d,im} := \min \left(\mu + \psi, \mu + \sum_{i=1}^n v_i \sqrt{\frac{\alpha}{1-\alpha}} \right), \quad (\text{C.20})$$

$$\text{VaR}_{\alpha}(S) \geq \underline{\text{VaR}}_{\alpha}^{ab,\mu,d,im} := \max \left(\mu - \psi \frac{1-\alpha}{\alpha}, \mu - \sum_{i=1}^n v_i \sqrt{\frac{1-\alpha}{\alpha}} \right), \quad (\text{C.21})$$

where ψ is the maximum value in the interval $\left[0, \min \left\{ b - \mu, (\mu - a) \frac{\alpha}{1-\alpha} \right\} \right]$ such that the inequalities

$$(\mu + \psi)^k (1 - \alpha) + \left(\mu - \psi \frac{1-\alpha}{\alpha} \right)^k \alpha \leq d_k \quad (\text{C.22})$$

hold for $k = 1, \dots, r$.

Most of the distributions used in financial and insurance risk modeling are unimodal by nature, as unimodality is a very common feature in real-world insurance datasets. Bernard et al. (2019) derives VaR bounds for unimodal distribution with given mean and variance.

Theorem 15 (VaR upper bounds for unimodal distributions with given mean and vari-

ance). Let $\alpha \in (0, 1)$, $E[S] = \mu$, $\text{var}[S] \leq s^2$, and S is unimodal. Then,

$$\text{VaR}_\alpha(S) \leq \overline{\text{VaR}}_\alpha^{u,\mu,s} := \begin{cases} \mu + s\sqrt{\frac{4}{9(1-\alpha)}} - 1 & \text{for } \alpha \in [5/6, 1), \\ \mu + s\sqrt{\frac{3\alpha}{4-3\alpha}} & \text{for } \alpha \in (0, 5/6), \end{cases} \quad (\text{C.23})$$

$$\text{VaR}_\alpha(S) \geq \underline{\text{VaR}}_\alpha^{u,\mu,s} := \begin{cases} \mu - s\sqrt{\frac{1-\alpha}{1/3+\alpha}} & \text{for } \alpha \in (1/6, 1), \\ \mu - s\sqrt{\frac{4}{9\alpha}} - 1 & \text{for } \alpha \in (0, 1/6]. \end{cases} \quad (\text{C.24})$$

Taking into account the non-negativity of the portfolio loss, Bernard et al. (2019) derives the following analytical bounds.

Theorem 16 (VaR upper bounds for non-negative unimodal distributions with given mean and variance). Let us denote by α_m the probability level of the mode. Let $\alpha \in [\alpha_m, 1)$, $E[S] = \mu$, $\text{var}[S] \leq s^2$, and S is a non-negative unimodal random variable. Then,

$$\underline{\underline{\text{VaR}}}_\alpha^{+,u,\mu,s} := \max(\underline{\text{VaR}}_\alpha^{u,\mu,s}, 0) \leq \text{VaR}_\alpha(S) \leq \overline{\text{VaR}}_\alpha^{+,u,\mu,s}, \quad (\text{C.25})$$

where

$$\overline{\text{VaR}}_\alpha^{+,u,\mu,s} := \begin{cases} \frac{\mu}{2(1-\alpha)} & \text{for } (\alpha, s) \in A_1, \\ \frac{3}{8}\mu(3\alpha+1) + \frac{3s^2}{4\mu}(3\alpha-1) + \frac{9s^4}{8\mu^3}(\alpha-1) & \text{for } (\alpha, s) \in A_2, \\ \mu + s\sqrt{\frac{4}{9(1-\alpha)}} - 1 & \text{for } (\alpha, s) \in A_3, \\ \mu + s\sqrt{3}(2\alpha-1) & \text{for } (\alpha, s) \in A_4, \\ \mu & \text{for } (\alpha, s) \in A_5, \end{cases} \quad (\text{C.26})$$

$$\text{in which } A_1 = \left(\frac{1}{2}, 1\right) \times \left[\mu\sqrt{\frac{\alpha-1/3}{1-\alpha}}, +\infty\right),$$

$$A_2 = \left(\frac{2}{3}, 1\right) \times \left(\mu\sqrt{\frac{\alpha-5/9}{1-\alpha}}, \mu\sqrt{\frac{\alpha-1/3}{1-\alpha}}\right) \cup \left(\frac{1}{2}, \frac{2}{3}\right] \times \left(\frac{\mu}{\sqrt{3}}, \mu\sqrt{\frac{\alpha-1/3}{1-\alpha}}\right),$$

$$A_3 = \left(\frac{2}{3}, 1\right) \times \left[0, \mu\sqrt{\frac{\alpha-5/9}{1-\alpha}}\right], \quad A_4 = \left(\frac{1}{2}, \frac{2}{3}\right] \times \left[0, \frac{\mu}{\sqrt{3}}\right], \quad A_5 = \left(0, \frac{1}{2}\right] \times [0, +\infty).$$

In heavy-tailed portfolios, the variance is sometimes large enough to be considered theoretically infinite. In order to treat this particular scenario, Bernard et al. (2019)

proved the following theorem.

Theorem 17 (VaR upper bounds for non-negative unimodal distributions with given mean and infinite variance). Let $\alpha \in (0, 1)$, $E[S] = \mu$, $\text{var}[S]$ is infinite, and S is a non-negative unimodal random variable. Then,

$$\underline{\text{VaR}}_{\alpha}^{+,u,\mu,h} := 0 \leq \text{VaR}_{\alpha}(S) \leq \overline{\text{VaR}}_{\alpha}^{+,u,\mu,h} := \begin{cases} \frac{\mu}{2(1-\alpha)} & \text{for } \alpha \in (\frac{1}{2}; 1), \\ \mu & \text{for } \alpha \in (0; \frac{1}{2}]. \end{cases} \quad (\text{C.27})$$

References

- Actuarial Association of Europe (2017). Comments template on discussion paper on the review of specific items in the solvency ii delegated regulation. Technical report.
- Barrieu, P. and G. Scandolo (2015). Assessing financial model risk. *European Journal of Operational Research* 242(2), 546–556.
- Basel Committee (2013). The Regulatory Framework: Balancing Risk-Sensitivity, Simplicity and Comparability. Technical report, Discussion paper, Basel: BIS.
- Bernard, C., M. Denuit, and S. Vanduffel (2018). Measuring portfolio risk under partial dependence information. *Journal of Risk and Insurance* 85(3), 843–863.
- Bernard, C., R. Kazzi, and S. Vanduffel (2019). Risk bounds for partially specified unimodal distributions. Available at SSRN 3412158.
- Bernard, C., L. Rüschendorf, and S. Vanduffel (2017). Value-at-Risk bounds with variance constraints. *Journal of Risk and Insurance* 84(3), 923–959.
- Bernard, C., L. Rüschendorf, S. Vanduffel, and R. Wang (2017). Risk bounds for factor models. *Finance and Stochastics* 21(3), 631–659.
- Bernard, C., L. Rüschendorf, S. Vanduffel, and J. Yao (2017). How robust is the Value-at-Risk of credit risk portfolios? *European Journal of Finance* 23(6), 507–534.
- Bertsimas, D., G. J. Lauprete, and A. Samarov (2004). Shortfall as a risk measure: properties, optimization and applications. *Journal of Economic Dynamics and control* 28(7), 1353–1381.
- Bignozzi, V., G. Puccetti, and L. Rüschendorf (2015). Reducing model risk via positive and negative dependence assumptions. *Insurance: Mathematics and Economics* 61, 17–26.
- Black, R., A. Tsanakas, A. D. Smith, M. B. Beck, I. D. MacLugash, J. Grewal, L. Witts, N. Morjaria, R. Green, and Z. Lim (2018). Model risk: illuminating the black box. *British Actuarial Journal* 23, 1–58.
- Board of Governors of the Federal Reserve System (2011). Supervisory guidance on model risk management.
- Cebrián, A. C., M. Denuit, and P. Lambert (2003). Generalized pareto fit to the society of actuaries’ large claims database. *North American Actuarial Journal* 7(3), 18–36.
- Choulakian, V. and M. A. Stephens (2001). Goodness-of-fit tests for the generalized pareto distribution. *Technometrics* 43(4), 478–484.
- De Schepper, A. and B. Heijnen (2010). How to estimate the Value-at-Risk under incomplete information. *Journal of Computational and Applied Mathematics* 233(9), 2213–2226.
- Denuit, M., J. Dhaene, and C. Ribas (2001). Does positive dependence between individual risks increase stop-loss premiums? *Insurance: Mathematics and Economics* 28(3), 305–308.

- Dionne, P. and B. Howard (2017). Use of models.
- Embrechts, P., G. Puccetti, and L. Rüschendorf (2013). Model uncertainty and VaR aggregation. *Journal of Banking & Finance* 37(8), 2750–2764.
- Gerstengarbe, F.-W. and P. Werner (1989). A method for the statistical definition of extreme-value regions and their application to meteorological time series. *Zeitschrift für Meteorologie* 39(4), 224–226.
- Grazier, K. L. (1997). *Group medical insurance large claims database collection and analysis*. Society of Actuaries.
- Hartigan, J. A., P. M. Hartigan, et al. (1985). The dip test of unimodality. *Annals of Statistics* 13(1), 70–84.
- Hartigan, P. (1985). Computation of the dip statistic to test for unimodality: Algorithm as 217. *Applied Statistics* 34(3), 320–5.
- Hürlimann, W. (1998). Extremal moment methods and stochastic orders. *Application in Actuarial Science. Monograph manuscript (available from the author)*.
- Hürlimann, W. (2002). Analytical bounds for two Value-at-Risk functionals. *ASTIN Bulletin: The Journal of the IAA* 32(2), 235–265.
- Jacobs Jr, M. (2015). The quantification and aggregation of model risk: perspectives on potential approaches. *International Journal of Financial Engineering and Risk Management* 2(2), 124–154.
- Kaas, R. and M. J. Goovaerts (1986). Extremal values of stop-loss premiums under moment constraints. *Insurance: Mathematics and Economics* 5(4), 279–283.
- Li, L., H. Shao, R. Wang, and J. Yang (2018). Worst-case Range Value-at-Risk with partial information. *SIAM Journal on Financial Mathematics* 9(1), 190–218.
- Lowenstein, R. (2008). Long-term capital management: It’s a short-term memory.
- Makarov, G. (1982). Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory of Probability & its Applications* 26(4), 803–806.
- OpRisk Advisory and Towers Perrin (2010). A New Approach for Managing Operational Risk: Addressing the Issues Underlying the 2008 Global Financial Crisis. *Sponsored by Joint Risk Management, Section Society of Actuaries, Canadian Institute of Actuaries & Casualty Actuarial Society*.
- Prudential Regulation Authority (2018). Model risk management principles for stress testing.
- Puccetti, G. and L. Rüschendorf (2012a). Bounds for joint portfolios of dependent risks. *Statistics & Risk Modeling with Applications in Finance and Insurance* 29(2), 107–132.
- Puccetti, G. and L. Rüschendorf (2012b). Computation of sharp bounds on the distribution of

- a function of dependent risks. *Journal of Computational and Applied Mathematics* 236(7), 1833–1840.
- Puccetti, G. and L. Rüschendorf (2013). Sharp bounds for sums of dependent risks. *Journal of Applied Probability* 50(1), 42–53.
- Puccetti, G., L. Rüschendorf, and D. Manko (2016). VaR bounds for joint portfolios with dependence constraints. *Dependence Modeling* 4(1), 368–381.
- Puccetti, G., L. Rüschendorf, D. Small, and S. Vanduffel (2017). Reduction of Value-at-Risk bounds via independence and variance information. *Scandinavian Actuarial Journal* 2017(3), 245–266.
- Rüschendorf, L. (1981). Sharpness of fréchet-bounds. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 57(2), 293–302.
- Rüschendorf, L. (2017). Risk bounds and partial dependence information. In *From Statistics to Mathematical Finance*, pp. 345–366. Springer.
- Salmon, F. (2009). Recipe for disaster: The formula that killed Wall Street.
- Samad-Khan, A. (2008). Modern operational risk management. *Emphasis* 2, 26–29.
- Sheskin, D. J. (2003). *Handbook of parametric and nonparametric statistical procedures*. crc Press.
- The Deloitte Center for Regulatory Strategy (2018). Model Risk Management, Building Supervisory confidence.
- Wang, B. and R. Wang (2011). The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis* 102(10), 1344–1360.
- Wang, R. (2014). Asymptotic bounds for the distribution of the sum of dependent random variables. *Journal of Applied Probability* 51(3), 780–798.
- Wang, R., L. Peng, and J. Yang (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics* 17(2), 395–417.
- Zisheng, O. and X. Chi (2006). Generalized pareto distribution fit to medical insurance claims data. *Applied Mathematics-A Journal of Chinese Universities* 21(1), 21–29.
- Zymler, S., D. Kuhn, and B. Rustem (2013). Worst-case Value-at-Risk of nonlinear portfolios. *Management Science* 59(1), 172–188.