Parametric insurance for extreme risks: the challenge to properly cover severe claims

Olivier Lopez[1], Maud Thomas[1]

January 13, 2022

Abstract

Recently, parametric insurance has emerged as a convenient way to cover risks that may be difficult to evaluate. Through the introduction of a parameter that triggers compensation and allows the insurer to determine a payment without evaluating the true loss, these products simplify the compensation process, and provide easily tractable indicators to perform risk management. On the other hand, this parameter may sometimes deviate from its purpose, and may not always correctly represent the basis risk. In this paper, we provide theoretical results that investigate the behavior of parametric insurance products when they are confronted to large claims. These results, in particular, measure the difference between the true loss and the parameter in a generic situation, with a particular focus on heavy-tailed losses. Simulation studies that complete the analysis show the importance of nonlinear dependence measures to ensure a good protection over the whole distribution.

Key words: Parametric insurance; extreme value analysis; risk theory; copula theory.

Short title: Parametric insurance and extreme risks.

[1] Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, 4 place Jussieu, F-75005 Paris, France

1 Introduction

Parametric insurance (see for example Lin and Kwon, 2020) is a very elegant and efficient way to simplify risk management in situations when the evaluation of losses might be complex. Parametric solutions have been developed for example in the case of natural disasters (see for example Van Nostrand and Nevius, 2011; Horton, 2018). A typical example is the case of an hurricane striking some area. The damages of such an episode can be very complex to evaluate, leading to expertise costs and delays in the compensation process. The solution proposed by parametric insurance is to not directly cover the true losses, but to work with some "parameter", namely a quantity that is linked to the loss and is easily measurable. In the case of natural disasters, wind speed, precipitation level, or any index based on relevant physical quantities can be used. Figueiredo et al. (2018) described a detailed methodology in the example of parametric insurance in Jamaica against flooding. Since the parameter can be measured instantly (or in a short amount of time), payment can be performed in a faster way. Moreover, when it comes to evaluating the risk, the situation is considerably simplified if one works with an easily available quantity that can be tracked and model through standard actuarial methods. This explains the growing popularity of these solutions that are now widely promoted (see for example Prokopchuk et al., 2020; Broberg, 2020; Bodily and Coleman, 2021).

Nonetheless, parametric insurance is no miracle solution. Reducing the volatility of the outcome has a cost. One of the difficulties is to convince the policyholder that a guarantee based on a given parameter is relevant. The attractiveness of such contracts may be reduced by the fear of the customers that the policy does not correctly cover the risk itself, against which they want to be protected. This is especially true if the parameter is complex and may not be fully understandable. In such cases, the simplification of the compensation process is not necessarily worth the loss in terms of protection. Moreover, "calculative misfires" as those pointed by Johnson (2021) do exist, that may discourage policyholders. Johnson (2021) details some "Ex gratia repairs" that are sometimes activated to limit the impact of such inconvenience.

The aim of the present paper is to study, in a general simplified framework, under which conditions parametric insurance may still provide (or not) a good protection against the risk in case of large claims. By large claims, we mean that the true loss of the policyholder is large. These situations, which deviate from the central scenario which is expected to drive the calibration of the payoff based on the parameter, require particular attention because they correspond to situations that are maybe not the more likely, but that correspond to important preoccupations of the policyholders: if the potential customers feel like they are not properly covered in case of serious events—which occurrence is at the core of the decision to rely on insurance—they may be reluctant to buy such solutions.

In this work, we focus on two particular cases. In the first one, the loss variable has a Gaussian tail. In this situation, significant deviations from the central scenario are of small probability. Hence, simply working on the correlation between the parameter and the loss is enough to improve the coverage of the risk. On the other hand, losses with heavy-tail are more challenging. The results we derive show that extreme losses may be difficult to capture, except if the parameter is able to reflect this heaviness. We illustrate these properties through a simulation study inspired by the case of cyber risk, more precisely of insuring against data breaches. In this case, the volume of data that are breached can be related to an estimated cost, and using this indicator to design parametric insurance products would make sense.

The rest of the paper is organized as follows. In Section 2, we discuss the general framework that we consider to evaluate the performance of a parameter used in parametric insurance. We especially focus on the question of the dependence between the parameter and the true loss, which is key to hope to achieve satisfactory properties. Section 3 gathers some theoretical results to measure how the parametric solution is able to approximate the true loss when the amount of the latter is high. The simulation study illustrating these properties in the case of cyber insurance is described in Section 4. The proof of the technical results are listed in the appendix in Section 6.

2 Model for parametric insurance

In this section, we explain the general framework used to model the difference between the true loss and the payment made via the parametric insurance product. The general setting is described in Section 2.1. The key question of the dependence between the two variables (true loss versus payment) is discussed in Section 2.2 which introduces tools from copula theory.

2.1 Description of the framework

In the following, we consider a random variable X representing the true loss experienced by a policyholder. Parametric insurance relies on the fact that X may be difficult to measure. In case of a natural catastrophe, X may be for example the total cost of the event on a given area. It could take time to properly evaluate precisely this cost (if even possible), and the idea is to rather pay a cost Y which is not exactly X, but is supposed to reflect it. Y is supposed to be a variable which is must easier to measure.

For example, the precipitation level θ, or other meteorological variables, can be obtained instantly, and a payoff can be deduced from Y, that is, in this case, $Y = \phi(\theta)$ for a given non-decreasing function ϕ. We will use the term "parameter" to denote the random variable θ.

Ideally, we would like Y to be close to X. Another benefit that could be taken from this approach is the potential reduction of volatility: paying Y instead of X is interesting in terms of risk management if the variance σ_Y^2 of Y is smaller than the variance σ_X^2 of X. Of course, if the variance of Y is too small compared to X, the quality of approximation of X by Y can diminish, since the distribution of Y does not match with the one of X.

2.2 Dependence

For the parameter θ, or, more precisely, the payoff $Y = \phi(\theta)$, to describe accurately the risk, X and Y should be dependent. The most simple way to describe this dependence is through correlation. Namely, let ρ be the correlation coefficient of X and Y defined by $\rho = \mathrm{Cor}(X, Y) = \mathrm{Cov}(X, Y)\sigma_X^{-1}\sigma_Y^{-1}$, where $\sigma_X^2 = \mathrm{Var}[X]$ and $\sigma_Y^2 = \mathrm{Var}[Y]$. Considering the quadratic loss, we have

$$E[(X - Y)^2] = \sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X^2\sigma_Y^2 + (E[X^2] + E[Y^2] - 2E[X]E[Y]).$$

Hence, increasing this correlation reduces the loss.

However, correlation is known to be a measure of dependence which takes mostly into account the center of the distribution, but not the tail. When facing a large claim, that is when X is far from its expectation, correlation is not enough to ensure that Y stays close to its target.

To illustrate this matter, let us consider the case where we want to cover claims for which X exceeds a deductible x_0. If X is not observed and if the insurance product is based on the parameter θ, the insurance company will make errors: sometimes a payment will be initiated when $X < x_0$, and sometimes no payment will occur even if $X \geq x_0$. This is caused by the fact that θ is only a proxy to get to X : payment is in fact initiated when $\theta \geq t_0$, and the event $\{\theta \geq t_0\}$ does not exactly match with the event $\{X \geq x_0\}$. A good parameter should be such that $\pi_+(t_0, x_0) = \mathbb{P}(\theta \geq t_0 | X \geq x_0)$ and $\pi_-(t_0, x_0) =$

$\mathbb{P}(\theta < t_0|X < x_0)$ are close to 1. Maximizing π_+ is supposed to enhances the satisfaction of the policyholder: coverage is obtained for almost all claims that are significant. On the other hand, a high value of π_- ensures that the insurer will not pay for claims that were initially beyond the scope of the product.

Let us introduce the cumulative distribution function (c.d.f.) $F_\theta(t) = \mathbb{P}(T \leq t)$ (resp. $F_X(t) = \mathbb{P}(X \leq t)$) defining the distribution of θ (resp. of X), and the joint c.d.f. $F_{\theta,X}(t_1, t_2) = \mathbb{P}(\theta \leq t_1, X \leq t_2)$. A common and general way to describe the dependence structure between θ and X is through copulas. Copula theory is based on the seminal result of Sklar (1959), stating that

$$F_{\theta,X}(t_1, t_2) = \mathfrak{C}(F_\theta(t_1), F_\theta(t_2)), \tag{2.1}$$

where $\mathfrak{C}$ is a copula function, that is the joint distribution function of a two-dimensional variable on $[0, 1]^2$ with uniformly distributed margins. The decomposition is unique if θ and X are continuous, which is the assumption we make in the following. Hence, (2.1) shows that there is a separation between the marginal behavior of (θ, X), and the dependence structure which is contained in $\mathfrak{C}$. Many parametric families of copulas have been proposed to describe various forms of dependence (see for example Nelsen, 2007).

Let us write π_+ and π_- in terms of copulas. Introducing the survival functions $S_\theta(t) = 1 - F_\theta(t)$, $S_X(t) = 1 - F_X(t)$, and $S(t_1, t_2) = \mathbb{P}(\theta \geq t_1, X \geq t_2)$, we have

$$\begin{aligned} \pi_+(t_0, x_0) &= \frac{\mathfrak{C}^*(S_\theta(t_0), S_X(x_0))}{S_X(x_0)}, \\ \pi_-(t_0, x_0) &= \frac{S_\theta(t_0) - S(t_0, x_0)}{F_X(x_0)}, \end{aligned}$$

where $\mathfrak{C}^*$ is the survival copula associated with $\mathfrak{C}$, that is

$$\mathfrak{C}^*(v, w) = v + w - 1 + \mathfrak{C}(1 - x, 1 - w).$$

To make the link with classical dependence measures for π_+, let us consider the case where $S_\theta(t_0) = S_X(x_0) = u$. In this situation, a large value of π_+ means that the deductible on θ that we use (based on a quantile of the distribution of θ), has approximately the same effect as a deductible directly on X (based on the same quantile). π_+ close to 1 rewrites $\mathfrak{C}^*(u, u)/u \approx 1$. If u is small (which means that we are focusing on higher quantiles, that is large claims), π_+ becomes close to

$$\lambda = \lim_{u \to 0} \frac{\mathfrak{C}^*(u, u)}{u} = \mathbb{P}(\theta \geq S_\theta^{-1}(u)|X \geq S_X^{-1}(u)),$$

which is the upper tail dependence, see Nelsen (2007). From this, we see that if we are focusing on large claims, correlation is not sufficient if one wants to correctly represent the risk, tail dependence seems more relevant.

Through this discussion, we focus solely on what triggers the payment, that is when $\theta > t_0$. But another issue is the difference between the payment $\phi(Y)$ and the true loss X. This question is investigated in the next section.

Remark 2.1 *π_- can also be expressed in terms of copula, but is harder to link with classical dependence measure. Our scope being essentially to focus on the tail of the distribution and on the potential difference between what the customer expects and what she/he gets, we do not develop this point.*

3 Difference between the true loss and the payoff based on the parameter

We here provide theoretical results to help quantifying the difference between X and Y when a claim is large, that is when X is. The quantities that are measured are defined in Section 3.1. We next consider two types of distribution: Gaussian variables (Section 3.2) are used as a benchmark, while Pareto-type variables are considered in Section 3.3.

3.1 Measuring the difference

In the following we consider two different quantities to measure how far $\phi(Y)$ is from X for large claims, that is when X exceeds some high value s.

The first measure we focus on is $E[X - Y|X \geq s]$. The advantage of this measure is that it shows if Y tends to be smaller or larger than X. On the other hand, the conditional bias $E[X - Y|X \geq s]$ may be zero (if $E[Y|X] = X$) while the conditional variance may be large, leading to potentially huge gaps between X and Y in practice.

For this reason, we also consider a classical quadratic loss, that is $E[(X - Y)^2|X \geq s]$. Note that this quadratic loss may not be defined for distributions that have a too heavy tail (this is also the case for $E[X - Y|X \geq s]$ which may not be defined if the expectation is infinite, but assuming a finite variance restrains even more the set of distributions).

To understand how the approximation made by the parametric approach deteriorates when X is large, we will next derive asymptotic approximations of these quantities when s tends to infinity.

3.2 Gaussian losses

In this section, we assume that (X, Y) are Gaussian variables, with distribution

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix} \right). \tag{3.1}$$

Considering such a framework is not completely realistic, in the sense that X and Y might take negative values with non-zero probability. Nevertheless, if μ_X and μ_Y are large enough, the probability of such an event is quite small. The Gaussian case is considered here essentially because it gives us an example of variables that are highly concentrated near their expectation, in order to measure the difference with heavy-tailed variables of Section 3.3.

Moreover, another motivation for considering Gaussian variables is the Central Limit Theorem. If X consists of the aggregation of individual claims, that is $X = \sum_{i=1}^n Z_i$, where $(Z_i)_{1\leq i\leq n}$ are independent identically distributed losses, the Central Limit Theorem states that X is approximately distributed as a Gaussian random variable with mean $nE[Z_1]$ and variance $n^{1/2}\mathrm{Var}(Z_1)$, provided that n is large enough. A Gaussian limit can also be obtained under some weak forms of dependence for these aggregated losses. This requires of course that the variance of Z_1 is finite.

A specificity of Gaussian random vectors is that their dependence structure is solely determine by the correlation matrix. Here, the dependence is driven by the correlation coefficient ρ. As we already mentioned, this is somehow a way to define dependence in the central part of the distribution. Due to the particular structure of Gaussian variable, this quantity has also an effect on the tail, that is even looking at situations where $X \geq s$ with s large.

Proposition 3.1 and 3.2 provide explicit formulas for $E[X - Y|X \geq s]$ and $E[(X - Y)^2|X \geq s]$.

Proposition 3.1 *Consider a random vector distributed as (3.1). Then, as $s \to +\infty$,*

$$E[X - Y|X \geq s] \sim (\mu_X - \mu_Y) + \left(1 - \frac{\rho\sigma_Y}{\sigma_X}\right)(s - \mu_X), \tag{3.2}$$

where $\sim$ is the symbol for equivalence.

Let us note that if $E[Y|X] = X$, that is if $\rho\sigma_Y\sigma_X^{-1} = 1$ and $\mu_X = \mu_Y$, we retrieve that $E[X - Y|X \geq s] = 0$. Apart from this trivial case, we can decompose (3.2) into two parts.

First, the difference between the expectation of μ_X and μ_Y, which reflects how good Y is able to capture the central part of the distribution of X. The second term increases with s when $\rho\sigma_Y\sigma_X^{-1} < 1$. However, a large value of the correlation coefficient ρ tends to reduce this effect. Hence, we can rely on correlation between X and Y to improve the ability of the parametric insurance contract to provide good results even for large claims.

On the other hand, the case $\rho\sigma_Y\sigma_X^{-1} \geq 1$ is less interesting to study: it would correspond to a situation where $\sigma_Y \geq \sigma_X$, that is a payoff based on the parameter which is more volatile that the one we would have directly used X. Although this situation may occur, this is not the ideal case where parametric insurance is used to both facilitate the collect of information required to trigger claim payment, and reduce the uncertainty.

Similar observations apply from the result of Proposition 3.2 for the quadratic loss.

Proposition 3.2 *Consider a random vector distributed as (3.1). Then, as $s \to +\infty$,*

$$E[(X-Y)^2|X \geq s] \sim \left(1 - \rho\frac{\sigma_Y}{\sigma_X}\right)^2 \frac{s^2}{\sigma_X^2}. \tag{3.3}$$

The exact value for $E[(X-Y)^2|X \geq s]$ can also be computed for a vector distributed as (3.1). The formula can be obtained from the proof of Proposition 3.2, which is made in Section 6.1.2.

3.3 Heavy-tail distributions

Heavy-tail random variables play an important role in Extreme Value Theory, (see for example Beirlant et al., 2004; Coles, 2001). Assuming that we are dealing with a i.i.d sample which cumulative distribution function F -satisfies the following property

$$F(t) = t^{-\gamma}\ell(t) \tag{3.4}$$

where ℓ is a slow-varying function, that is

$$\forall x > 0,\ \lim_{t\to\infty} \frac{\ell(tx)}{\ell(t)} = 1,$$

the fundamental result of Extreme Value Theory states that the normalized maximum M_n of the sample converges in distribution toward a non-degenerated distribution and that this distribution necessarily belongs to a parametric family of distributions, called the generalized extreme value distributions. More precisely, under assumption (3.4) when $\gamma > 0$, there exists normalizing constants a_n and $b_n > 0$ such that

$$\mathbb{P}\left(\frac{M_n - a_n}{b_n} \geq x\right) \underset{n\to\infty}{\longrightarrow} G_\gamma(x),$$

and $G_\gamma(x)$ is necessarily of the form, for x such that $1 + \gamma x > 0$,

$$G_\gamma(x) = \exp(-(1 + \gamma x)^{1/\gamma}).$$

from (Fisher and Tippett, 1928) and (Gnedenko, 1943). More precisely, Gnedenko (1943) showed that the survival function $1 - F(t)$ is necessarily of the form (3.4). The parameter γ which reflects the heaviness of tail of F is called the tail index. The higher γ, the heavier the tail of the distribution: X tends to take large values with a significant probability. Here, $\gamma > 0$, and X belongs to the heavy-tail domain. The tail of heavy-tail distribution decreases polynomially toward 0, their moments of order larger than $1/\gamma$ do not exist. For example, the Pareto, the Student, the log-normal and the Cauchy distributions are heavy-tailed.

Hence, Assumption 3.4 allows us to cover a large set of distributions. In the following, we thus assume that

$$S_X(t) = \ell_X(t)t^{-1/\gamma_X}, \quad \text{and} \quad S_Y(t) = \ell_Y(t)t^{-1/\gamma_Y}, \tag{3.5}$$

with $\gamma_X, \gamma_Y > 0$ and ℓ_X and ℓ_Y two slowing varying functions.

Let us also note that heavy-tailed variables can also be used to approximate sums of losses. Taking again the example of $X = \sum_{i=1}^n Z_i$, if Z_i are heavy-tailed i.i.d. random variables, Mikosch and Nagaev (1998) show that, when it comes to high quantiles, X can be approximated by a heavy-tailed variable.

We here do not provide exact values for $E[X-Y|X \geq s]$ and $E[(X-Y)^2|X \geq s]$, since these quantities depend on ℓ_X and ℓ_Y. Nevertheless, our results should general bounds for s large. We first consider the case of $E[X - Y|X \geq s]$ in Proposition 3.3. We recall that $E[X - Y|X \geq s]$ (resp. $E[(X - Y)^2|X \geq s]$) is defined only if γ_X and γ_Y are less than 1 (resp. less than 0.5).

Proposition 3.3 *Consider $X \geq 0$ and $Y \geq 0$ with survival functions as in* (3.5)*, with $\gamma_X > \gamma_Y$. There exists a constant $c > 0$ depending on ℓ_X and ℓ_Y and not on their dependence structure, such that, for s large enough,*

$$E[X - Y|X \geq s] \geq cs.$$

Proposition 3.3 shows that there is a linear increase in this difference for large values of s. However, the situation is quite different from the Gaussian case. In the Gaussian case, we could expect to reduce the slope by relying on a strong correlation between X and

Y. Here, improving the correlation would certainly have an effect, but not on the leading linear term. In fact, the distribution of X being heavier than Y, when s becomes large, X belongs to some areas which are unreachable by Y except for very low probability. In practice, the difference between γ_Y and γ_X also plays a role, but, again, only on smaller order terms.

We can obtain a more precise result under some additional assumptions, as we see in Proposition 3.4 below.

Proposition 3.4 *Consider $X \geq 0$ and $Y \geq 0$ with survival functions as in* (3.5) *with $\gamma_X > \gamma_Y$. Moreover, let*

$$\psi(x) = E[Y|X = x].$$

Assume that ψ is strictly non decreasing and such that the random variable $\psi(X)$ is heavy-tailed. Then

$$E[X - Y|X \geq s] = s - \psi(s) + o(s).$$

Under this additional assumption, we see that we even get a linear increase of $E[X - Y|X \geq s]$, since $\psi(s)$ is less than s since $\gamma_Y < \gamma_X$.

A similar result is obtained for the quadratic loss in Proposition 3.5, where we see that this quantity increases at rate s^2, no matter the dependence structure between X and Y.

Proposition 3.5 *Consider $X \geq 0$ and $Y \geq 0$ with survival functions as in* (3.5), *with $\gamma_X > \gamma_Y$.*

$$E[(X - Y)^2|X \geq s] = s^2 + o(s^2).$$

Through these theoretical results, we see that there tends to be a significant gap between the payoff Y and the true loss X when the variables are heavy-tailed. Here we assumed that $\gamma_X > \gamma_Y$, which seems to be the most interesting case since, in this situation, the risk taken by parametric insurance is less volatile than the original one. In the opposite situation, the parametric product would tend to over compensate the true loss. The situation $\gamma_X = \gamma_Y$ seems purely theoretical, since it would require a very particular situation in which both tail indexes are exactly the same.

Finally, let us mention that all the results of this section extends to the case where X is heavy-tailed, and Y is low-tailed. In this situation, the remainder terms in the asymptotic expansions are even smaller.

4 Illustration inspired by cyber risk

The purpose of this section is to illustrate the theoretical results and to go beyond the asymptotic approximations we gave. Our simulation setting is inspired from questions arising in cyber insurance. Cyber insurance is a field in which parametric insurance is increasingly mentioned as a promising tool to conceive contracts adapted to the complexity of the risk, see for example Dal Moro (2020) or Chapter 5 in OCDE (2017). Different indicators have been proposed to monitor the risk, either regarding frequency or severity of an event. We here build a simulation setting which is inspired (in terms of the distribution we use) from calibrations that have been done in the field of cyber, more precisely in the case of data breaches. This context offers a natural physical parameter which describes the severity. In Section 4.1, we give a short presentation of this context. The simulation settings we consider are explained in Section 4.2, with a focus on the copula models that we use in Section 4.3. The results and analysis are given in Section 4.4.

4.1 Description of data breaches through number of records

Among the various types of situations behind the concept of cyber risk, data breaches are probably the ones for which the cost related to such an event is relatively easy to evaluate. Indeed, the volume of data that has been breached (namely the "number of records") is a good indication of the severity. This quantity, which can be easily measured soon after occurrence of a claim, can be used as a parameter that should be able to give indications on the true loss.

Jacobs (2014) proposed a relationship between this number of records, say θ, and X, which can be taken as the formula defining the payoff Y. This relationship is of the following type,

$$\log Y = \alpha + \beta \log \theta, \tag{4.1}$$

where X and Y are expressed in dollars. The formula has been calibrated from data coming from the Ponemon institute (see Ponemon, 2018). Jacobs (2014) estimated $\alpha = 7.68$ and $\beta = 0.76$. Nevertheless, Farkas et al. (2021) pointed that the formula, computed from data collected in 2014, was not consistent with some so-called "mega-breache" observed afterwards. For example, two mega-breaches have been reported in the CODB report 2018 (see Ponemon, 2018). The first one, with 1 million breached records, would lead to an estimated cost of nearly 79 million dollars, while the true cost was approximately 39 millions. The biggest one, with 50 million breached records, would lead to $Y = 1.5$

billion dollars, far from the paid 350 millions. Hence, Farkas et al. (2021) proposed a (very rough) recalibration of the parameters, taking $\alpha = 9.59$ and $\beta = 0.57$.

Behind this discussion, one sees that, even though we are facing an indicator (the number of records) that seems to be physically relevant to describe the magnitude of the event we considered, the payoff function Y may be a very rough approximation of the true loss, especially in the tail. Apart from the inherent variance of the error behind the calibration of (4.1), and the potential lack-of-fit of the model, we see that there is many uncertainties concerning the estimation of the parameters.

The examples we use in the following are inspired by this example, and the corresponding values of the parameters.

4.2 Simulation settings

In this section, we consider different settings inspired by the relationship (4.1) between the expected cost of a data breach and the number of records.

Main settings. First of all, to consider reasonable values that are connected to our example, we need a proper distribution for the parameter θ. We choose the simple Pareto distribution considered by Maillart and Sornette (2010), that is

$$\mathbb{P}(\theta \geq t) = \left(\frac{u}{t}\right)^b, \text{ for } t \geq u,$$

with $b = 0.7$ and $u = 7.10^4$. This heavy-tail distribution is consistent with the work of , or with the analysis of Farkas et al. (2021) based on more recent data (where a significant class of the data breaches has been identified to follow a distribution close to the one of (Maillart and Sornette, 2010)).

In this case, the variable Y inherits the heavy-tail property of θ if Y is defined according to (4.1). More precisely, we see that

$$\mathbb{P}\left(Y \geq t\right) = \left(\frac{u'}{t}\right)^{b/\beta}, \text{ for } t \geq u',$$

with $u' = u^\beta \exp(\alpha)$. The parameter β is of course the more important when we look at the tail of the distribution, since it is directly linked to the tail index $\gamma_Y = \beta/b$. We take $\alpha = 9.59$ and $\beta = 0.5$, that is slightly lower than the parameter β calibrated in Farkas et al. (2021).

Next, to simulate X, we consider

$$\log X = \alpha + \beta' \log(\theta'),$$

where

$$\mathbb{P}(\theta \geq t) = \left(\frac{u}{t}\right)^{b'}, \text{ for } t \geq u,$$

with $b' = b - 0.1$. The choice of $b' - 0.1$ is motivated by the fact that this corresponds to the error margin given by Maillart and Sornette (2010). The parameter θ' is simulated to be dependent from θ, and creates the dependence between X and Y, by considering different copula functions to describe the dependence structure. The copula families and how the corresponding parameters have been chosen is explained in Section 4.3.

Benchmark settings. To better understand the impact of the heaviness of the distributions and the impact of the difference between the values of the tail indexes. This settings are denoted B_1 to B_3 in the following.

- Setting B_1 : Y is simulated as in the main settings, but $X = Y + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma_1^2)$. The variance is taken so that X has the same variance as in the main settings.
- Setting B_2 : Y is simulated as in the main settings, but $\log X = \log Y + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma_2^2)$. Again, σ_2^2 is taken so that X has the same variance as in the main settings.
- Setting B_3 : (X, Y) is a Gaussian vector as in (3.1), with same mean and variance as in the main settings. We consider different values for the correlation coefficient, $\rho = 0.3$, $\rho = 0.5$ and $\rho = 0.7$.

All of these benchmark cases can be thought has "favorable cases": with B_1 and B_2, the tail index of X and Y is the same. In the first situation, we take an additive Gaussian error, which means that X is relatively concentrated around Y. In B_2, the errors are multiplicative, since the Gaussian error is applied to the logarithms. This typically corresponds to the optimistic case where $E[\log X|Y] = \alpha + \beta Y$: that is a standard linear regression model on the logarithm of X with no misspecification error. Finally, the benchmark B_3 is the more optimistic case, where X and Y have Gaussian tails.

4.3 Copula families

We consider three copula families, corresponding to different types of dependence structure. The copula functions are summarized in Table 1.

Copula family	Copula function	$\delta \in$	τ	λ_U
Clayton survival	$u+v-1+\left[\max\left(u^{-\delta}+v^{-\delta}-1,0\right)\right]^{-1/\delta}$	$\delta \geq -1$ and $\delta \neq 0$	$\frac{\delta}{\delta+2}$	$2^{-\delta}$
Gumbel	$\exp\left(-\left[(-\log(-u))^{\delta}-(\log(-v))^{\delta}\right]^{1/\delta}\right)$	$\delta \geq 1$	$\frac{\delta-1}{\delta}$	$2-2^{-\delta}$
Frank	$-\frac{1}{\delta}\log\left(1+\frac{(e^{-\delta u}-1)(e^{-\delta v}-1)}{e^{-\delta}-1}\right)$	$\delta \neq 0$	Implicit	0

Table 1: Copula families used in the different simulation settings.

The Frank copula family is classical and flexible, but does not allow us to take tail dependence into account. On the other hand, the two other families that we consider (Gumbel, and Clayton survival copula) have non-zero tail dependence.

To make things comparable, we consider values of parameters that provide a similar dependence according to an appropriate indicator. The dependence measure that we use is Kendall's tau coefficient, which is defined, for two random variables (θ, θ'), as

$$\tau = \mathbb{P}((\theta_1-\theta_2)(\theta'_1-\theta'_2)>0) - \mathbb{P}((\theta_1-\theta_2)(\theta'_1-\theta'_2)<0),$$

where (θ_1, θ'_1) and (θ_2, θ'_2) are independent copies of (θ, θ'). A simple relationship between the parameter τ and the classical parametrization of the copula families we consider.

We consider three values of the parameters for each copula family, corresponding respectively to $\tau = 0.3$, $\tau = 0.5$ and $\tau = 0.7$.

4.4 Simulation results

In Figure 1 and 2, we see the evolution of $E[(X-Y)|X \geq s]$ and $E[(X-Y)^2|X \geq s]$ in the different models used for the simulations. Let us recall that these models only differ because of the dependence structure between Y and X. In each case, X and Y are heavy-tailed.

We can observe that the evolution seems approximately linear for s large in the case of $E[(X-Y)|X \geq s]$, and approximately quadratic for $E[(X-Y)^2|X \geq s]$. Clearly, the dependence structure matters, allowing to reduce the slope (which is a property that was expected but not covered by Proposition 3.3 and 3.5). We also see in Figure 1 that the slope is close to 1 for Frank's copula, and is reduced for the families that allow tail dependence. Let us point out that a slope close to 1 is bad news: this means that the difference between Y and X is of the same magnitude as X itself (since $X \geq s$).

These results emphasize the need to use a parameter which is not only connected to the basis risk via a form of "central dependenc", but which can also integrate tail dependence.

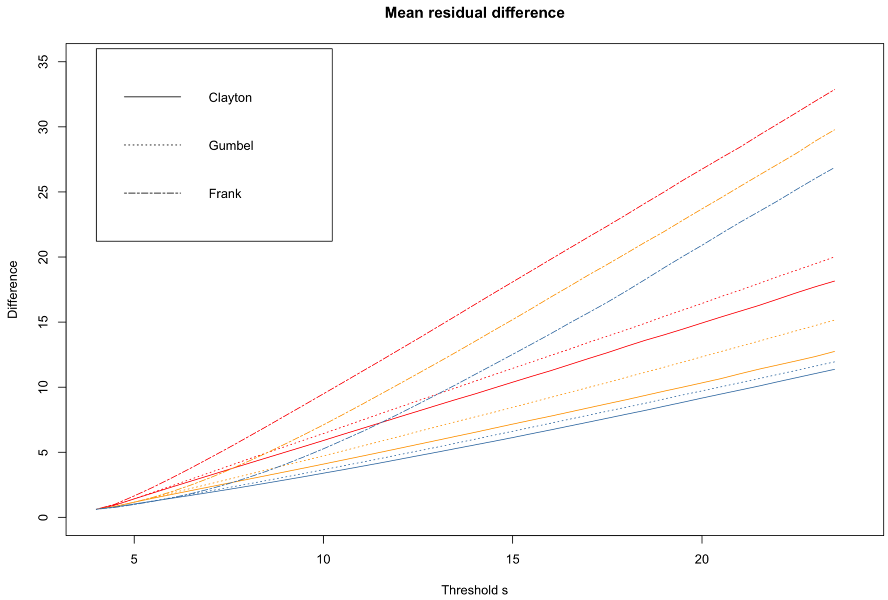

Figure 1: Evolution of $E[X - Y|X \geq s]$ with respect to the threshold s. The red lines correspond to a Kendall's tau coefficient $\tau = 0.3$, orange $\tau = 0.5$, blue $\tau = 0.7$.

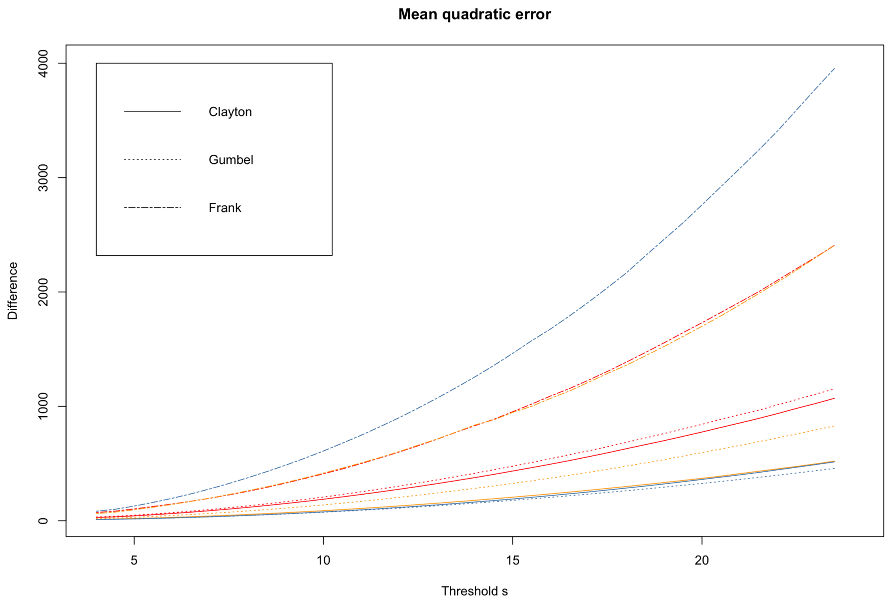

Figure 2: Evolution of $E[(X - Y)^2|X \geq s]$ with respect to the threshold s. The red lines correspond to a Kendall's tau coefficient $\tau = 0.3$, orange $\tau = 0.5$, blue $\tau = 0.7$.

Without this property, heavy tailed variables may not be approximated properly due to the significant proportion of events in the tail of the distribution.

Next, Figures 3 and 4 show comparisons between the Clayton case and the benchmarks. The conclusions for the other settings being similar, we postpone the figures to the appendix section (section 6.3).

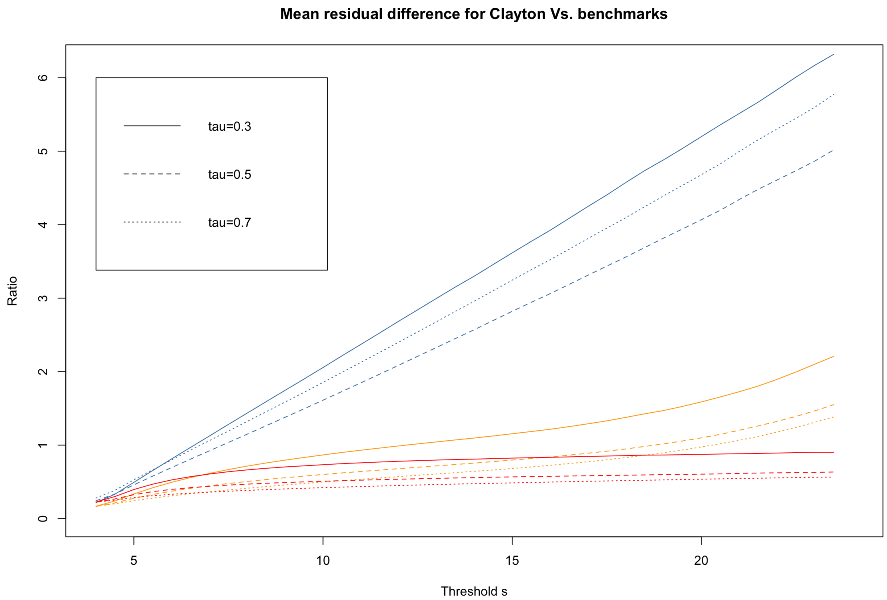

Figure 3: Evolution of the ratio of $E[(X-Y)|X \geq s]$ computed from the Clayton survival copula model, with respect to the value of $E[(X-Y)|X \geq s]$ obtained in the benchmark settings. The orange lines correspond to benchmark B_1, the red ones to scenario B_2, the blue one to scenario B_3.

From these figures, we see that the cases B_1 and B_2 are much favorable. In B_3, we are in the Gaussian case, that is X and Y have low tails, and the error is much smaller than in the case of heavy tails. In B_1 and B_2, X and Y should have the same tail, but we see that, although these cases seem very optimistic, the absence of tail dependence make the results poorer in some situations.

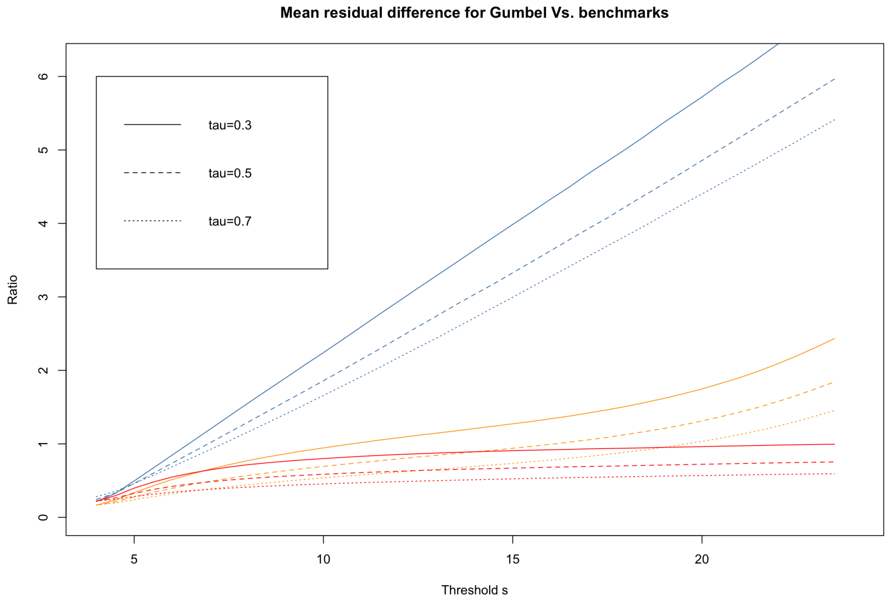

Figure 4: Evolution of the ratio of $E[(X-Y)|X \geq s]$ computed from the Gumbel copula model, with respect to the value of $E[(X-Y)|X \geq s]$ obtained in the benchmark settings. The orange lines correspond to benchmark B_1, the red ones to scenario B_2, the blue one to scenario B_3.

5 Conclusion

In this paper, we investigated what are the factors that may disturb the project of parametric insurance to properly cover the basis risk in case of large claims. This question can be seen of the ability of a random variable to approximate another. We particularly emphasized the problem created by heavy tailed losses. In these cases, the difference between what is paid to the policyholder and the true may be quite large, particularly if the tail index is even slightly misevaluated. The ability of parametric insurance to reduce the variance will be diminished in this case—except if one accepts to provide a poorer coverage for large claims—since no reduction of the tail index compared to the basis risk leads to, more or less, the same kind of variability. Next, the dependence structure between the parameter and the true loss seems to be an important question to address. Tail dependence appears to be essential in order to obtain a proper approximation of the losses for large claims. Let us mention that designing a parameter tail dependent from the basis risk is a challenging task: analyzing "extreme" events requires lots of data, which pleads for a careful statistical analysis to properly define the proper metric. Alternatively, specific treatment for large claims could be anticipated, in order not to create a too huge gap between the expectations of the policyholders and the compensated amount.

6 Appendix

The Appendix section is organized in the following way. Section 6.1 is devoted to the results of Gaussian variables, while the results regarding variables with Pareto tail are gathered in Section 6.2. The additional comparisons with the benchmarks of the simulation study are shown in section 6.3.

6.1 Results on Gaussian variables

6.1.1 Proof of Proposition 3.1

First recall that X is distributed according to the distribution $\mathcal{N}(\mu_X, \sigma_X^2)$ so that

$$E[X|X \geq s] = \mu_X + \sigma_X^2 h(s|\mu_X, \sigma_X^2), \tag{6.1}$$

where $h(s|\mu_X, \sigma_X^2)$ is the hazard rate of a Gaussian random variable with mean μ_X and variance σ_X^2, that is

$$h(s|\mu_X, \sigma_X^2) = \frac{\exp\left(-\frac{(s-\mu_X)^2}{2\sigma_X^2}\right)}{\sqrt{2\pi\sigma_X^2}\bar{\Phi}\left(\frac{s-\mu_X}{\sigma_X}\right)} \underset{s\to+\infty}{\sim} \frac{s-\mu_X}{\sigma_X^2},$$

with $\bar{\Phi}$ the survival function of the standard Gaussian distribution $\mathcal{N}(0,1)$.

First of all, let $m(s) = E[(X-Y)\mathbf{1}_{X\geq s}]$. Since

$$E[Y|X] = \mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(X - \mu_X),$$

we have

$$m(s) = \left(1 - \frac{\rho\sigma_Y}{\sigma_X}\right) E\left[X\mathbf{1}_{X\geq s}\right] - \left(\mu_Y - \frac{\rho\sigma_Y}{\sigma_X}\mu_X\right)\mathbb{P}(X \geq s).$$

From (6.1), we get

$$\begin{aligned} E[X - Y|X \geq s] &= (\mu_X - \mu_Y) + \left(1 - \frac{\rho\sigma_Y}{\sigma_X}\right)\sigma_X^2 h(s|\mu_X, \sigma_X^2) \\ &\sim (\mu_X - \mu_Y) + \left(1 - \frac{\rho\sigma_Y}{\sigma_X}\right)(s - \mu_X), \end{aligned}$$

as $s \to \infty$.

6.1.2 Proof of Proposition 3.2

We have

$$E[(X-Y)^2|X\geq s]=E[X^2|X\geq s]+E[Y^2|Y\geq s]-2E[XY|X\geq s].$$

Moreover,

$$m_2(s)=E[X^2|X\geq s]=\left(\sigma_X^2+\mu_X^2\right)+\sigma_X^2h(s|\mu_X,\sigma_X^2)(s+\mu_X),$$

and

$$E[XY|X\geq s]=E[XE[Y|X]|X\geq s].$$

From (6.1),

$$E[XY|X\geq s]=\left(\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}\right)(\mu_X+\sigma_X^2h(s|\mu_X,\sigma_X^2))+\rho\frac{\sigma_Y}{\sigma_X}m_2(s).$$

Finally,

$$E[Y^2|X\geq s]=E[\mathrm{Var}(Y|X)|X\geq s]+E[E[Y|X]^2|X\geq s],$$

which leads to

$$\begin{aligned}E[Y^2|X\geq s] &= \sigma_Y^2(1-\rho^2)+\left(\mu_Y-\rho\mu_X\frac{\sigma_Y}{\sigma_X}\right)^2\\ &\quad+\rho^2\frac{\sigma_Y^2}{\sigma_X^2}m_2(s)+2\rho\frac{\sigma_Y}{\sigma_X}(\mu_Y-\rho\mu_X\frac{\sigma_Y}{\sigma_X})(\mu_X+\sigma_X^2h(s|\mu_X,\sigma_X^2)).\end{aligned}$$

Hence, we see that

$$\begin{aligned}E[(X-Y)^2|X\geq s] &= \left(1-\rho\frac{\sigma_Y}{\sigma_X}\right)^2m_2(s)+o(m_2(s))\\ &= \left(1-\rho\frac{\sigma_Y}{\sigma_X}\right)^2sh(s|\mu_X,\sigma_X^2)+o(sh(s|\mu_X,\sigma_X^2))\\ &\sim \left(1-\rho\frac{\sigma_Y}{\sigma_X}\right)^2\frac{s^2}{\sigma_X^2},\end{aligned}$$

as $s\to\infty$.

6.2 Results on heavy-tail variables

6.2.1 A preliminary result

We start with a preliminary result showing that the variable $X-Y$ has the same tail index as X, under the assumptions of Propositions 3.3 to 3.5. Lemma 6.1 consists in providing upper and lower bounds for the survival function of $X-Y$.

Lemma 6.1 *Let X, Y be as defined in Proposition 3.3, and let $Z = X - Y$. Then, introducing $S_Z(t) = \mathbb{P}(Z \geq t)$, for $t \geq 0$,*

$$h^{-1/\gamma_X} t^{-1/\gamma_X} \ell_X(th) - (h-1)^{-1/\gamma_Y} t^{-1/\gamma_Y} \ell_Y(t(h-1)) \leq S_Z(t) \leq t^{-1/\gamma_X} \ell_X(t),$$

with $h > 0$ fixed.

Proof. Since $Y \geq 0$ almost surely, $X - Y \geq t$ implies that $X \geq t$. Hence, we get $S_Z(t) \leq S_X(t)$, and the upper bound of Lemma 6.1 is obtained.

To obtain the lower bound, introduce the event, for $h > 0$ fixed,

$$E_h(t) = \{\{X \geq th\} \cap \{Y \leq t(h-1)\}\}.$$

We have $S_Z(t) \geq \mathbb{P}(E_h(t))$. Next,

$$\mathbb{P}(E_h(t)) \geq \mathbb{P}(X \geq th) - \mathbb{P}(Y \geq t(h-1)) = S_X(th) - S_Y(t(h-1)).$$

This shows the lower bound in Lemma 6.1. ■

We can now apply this lemma to obtain the following proposition.

Proposition 6.2 *Let X be a heavy tail variable with $\gamma_X > 0$. Let Y be a non negative variable with tail index $\gamma_Y < \gamma_X$. If $\gamma_X > \gamma_Y$, $Z = X - Y$ is heavy-tailed with tail index γ_X.*

Proof.

Let $\ell_Z(t) = t^{1/\gamma_X} S_Z(t)$. Let us proof that ℓ_Z is slowly-varying. From Lemma 6.1, for $x > 1$ and for all $h > 10$,

$$\frac{\ell_Z(tx)}{\ell_Z(t)} \leq \frac{\ell_X(tx)}{h^{-1/\gamma_X} \ell_X(th) - (h-1)^{-1/\gamma_X} (t(h-1))^{1/\gamma_X - 1/\gamma_Y} \ell_Y(t(h-1))}, \quad (6.2)$$

and,

$$\frac{\ell_X(thx)}{h^{1/\gamma_X} \ell_X(t)} - \frac{x^{1/\gamma_X} (tx(h-1))^{-1/\gamma_Y} \ell_Y(tx(h-1))}{t^{-1/\gamma_X} \ell_X(t)} \leq \frac{\ell_Z(tx)}{\ell_Z(t)}. \quad (6.3)$$

Hence, for all $x > 0$ and $h > 1$, taking the limit of the right-hand side of (6.2),

$$\limsup_{t \to \infty} \frac{\ell_Z(tx)}{\ell_Z(t)} \leq h^{1/\gamma_X}.$$

To see that, let $\beta = \gamma_Y^{-1} - \gamma_X^{-1} > 0$. We have $t^{-\beta/2} l_Y(t(h-1)) \to_{t\to\infty} 0$, and $t^{\beta/2} l_X(tx) \to_{t\to\infty} \infty$.

Similarly, from (6.3), we get

$$\liminf_{t\to\infty} \frac{\ell_Z(tx)}{\ell_Z(t)} \geq \frac{1}{h^{1/\gamma_X}}.$$

This is valid for all $h > 1$. Next, let h tend to 1 in order to obtain that, for all x, $\lim_{t\to\infty} \ell_Z(tx)/\ell_Z(t) = 1$, leading to the result.

■

6.2.2 Proof of Proposition 3.3

From Proposition 6.2, $Z = X - Y$ is heavy-tailed with tail index γ_X, that is $S_Z(t) = \ell_Z(t)t^{-1/\gamma_X}$ with ℓ_Z slow-varying. Hence,

$$E[Z|Z \geq s] = s + o(s). \tag{6.4}$$

Next,

$$\begin{aligned} E[Z|X \geq s] &= E[Z|X \geq s, Z \geq s]\mathbb{P}(Z \geq s|X \geq s) \\ &\quad + E[Z|X \geq s, Z < s]\mathbb{P}(Z < s|X \geq s). \end{aligned}$$

Since $Z \geq s$ implies $X \geq s$, we have

$$E[Z|X \geq s, Z \geq s]\mathbb{P}(Z \geq s|X \geq s) = E[Z|Z \geq s]\mathbb{P}(Z \geq s|X \geq s).$$

Moreover,

$$\mathbb{P}(Z \geq s|X \geq s) = \frac{\mathbb{P}(\{Z \geq s\} \cap \{X \geq s\})}{\mathbb{P}(X \geq s)} = \frac{\mathbb{P}(Z \geq s)}{\mathbb{P}(X \geq s)}.$$

From this, we get

$$E[Z|X \geq s] \geq E[Z|Z \geq s] \times \frac{\ell_Z(s)}{\ell_X(s)} = (s + o(s)) \times \frac{\ell_Z(s)}{\ell_X(s)},$$

from (6.4). From Lemma 6.1, we see that, taking for example $x = 2$,

$$\frac{\ell_Z(s)}{\ell_X(s)} \geq \frac{1}{2^{1/\gamma_X}} \frac{\ell_X(2s)}{\ell_X(s)} - \frac{1}{s^\beta} \frac{\ell_Y(s)}{\ell_X(s)},$$

introducing $\beta = \gamma_Y^{-1} - \gamma_X^{-1} > 0$. Since ℓ_X and ℓ_Y are slow varying, $s^{\beta/2}\ell_X(s) \to \infty$, $s^{-\beta/2}\ell_Y(s) \to 0$ as s tends to infinity, leading to

$$\frac{1}{s^\beta} \frac{\ell_Y(s)}{\ell_X(s)} = o(1).$$

Moreover,

$$\lim_{s\to\infty} \frac{\ell_X(2s)}{\ell_X(s)} = 1,$$

showing that there exists a constant $c > 0$ such that, for s large enough,

$$E[Z|X \geq s] \geq cs.$$

6.2.3 Proof of Proposition 3.4

We have

$$E[X - Y|X \geq s] = E[X - \psi(X)|X \geq s],$$

where $\psi(X) = E[Y|X]$. Since X is heavy-tailed with tail index $\gamma_X > 0$, $E[X|X \geq s] = s + o(s)$. On the other hand,

$$E[\psi(X)|X \geq s] = E[\psi(X)|\psi(X) \geq \psi(s)],$$

since ψ is strictly non decreasing. Then, since $\psi(X)$ assumed to be heavy-tailed,

$$E[\psi(X)|\psi(X) \geq \psi(s)] = \psi(s) + o(\psi(s)),$$

which shows that

$$E[X - Y|X \geq s] = s - \psi(s) + o(s).$$

6.2.4 Proof of Proposition 3.5

Let $Z = X - Y$. We have

$$E[(X - Y)^2|X \geq s] = \frac{E[Z^2\mathbf{1}_{X\geq s}]}{\mathbb{P}(X \geq s)} \geq \frac{E[Z^2\mathbf{1}_{X\geq s}\mathbf{1}_{Z\geq s}]}{\mathbb{P}(X \geq s)}.$$

If $Z \geq s$, then, necessarily, $X \geq s$ since $Y \geq 0$. Hence

$$E[(X - Y)^2|X \geq s] \geq \frac{E[Z^2\mathbf{1}_{Z\geq s}]}{\mathbb{P}(X \geq s)} = E[Z^2|Z \geq s]\frac{\mathbb{P}(Z \geq s)}{\mathbb{P}(X \geq s)}.$$

Next,

$$E[Z^2|Z \geq s] = \frac{E[Z^2\mathbf{1}_{Z^2\geq s^2}]}{\mathbb{P}(Z \geq s)} - \frac{E[Z^2\mathbf{1}_{Z<-s}]}{\mathbb{P}(Z \geq s)}.$$

Moreover,

$$E[Z^2\mathbf{1}_{Z<-s}] \leq E[Y^2\mathbf{1}_{Y\geq s}] = E[Y^2|Y \geq s]\mathbb{P}(Y \geq s).$$

Combining this last equation with Proposition 6.2 leads to

$$\begin{aligned} E[Z^2|Z \geq s] &\geq E[Z^2|Z^2 \geq s^2]\frac{\mathbb{P}(Z^2 \geq s^2)}{\mathbb{P}(Z \geq s)} - \frac{\ell_Y(s)}{\ell_Z(s)}E[Y^2|Y \geq s]s^{1/\gamma_X - 1/\gamma_Y} \\ &\geq E[Z^2|Z^2 \geq s^2] - \frac{\ell_Y(s)}{\ell_Z(s)}E[Y^2|Y^2 \geq s^2]s^{1/\gamma_X - 1/\gamma_Y}, \end{aligned}$$

where the last line comes from the fact that, $\mathbb{P}(Z^2 \geq s^2) \geq \mathbb{P}(Z \geq s)$, and that, since $Y \geq 0$ almost surely, $E[Y^2|Y \geq s] = E[Y^2|Y^2 \geq s^2]$.

We have, since $Y \geq 0$ almost surely,

$$\mathbb{P}(Y^2 \geq t) = \mathbb{P}(Y \geq t^{1/2}) = \frac{\ell_Y(t^{1/2})}{t^{1/(2\gamma_Y)}},$$

where $t \to \ell_Y(t^{1/2})$ inherits the slow-varying property of ℓ_Y. Hence

$$E[Y^2 | Y^2 \geq s^2] = s^2 + o(s^2). \tag{6.5}$$

On the other hand, let $\beta = \gamma_Y^{-1} - \gamma_X^{-1} > 0$. Since $\ell_Y(s)s^{2-\beta/2} \to 0$ when s tends to infinity, and since $\ell_Z(s)s^{\beta/2} \to \infty$. Hence, using (6.5),

$$\frac{\ell_Y(s)}{\ell_Z(s)} E[Y^2 | Y^2 \geq s^2] s^{1/\gamma_X - 1/\gamma_Y} = s^{2-\beta/2} + o(s^{2-\beta/2}).$$

Since $E[Z^2 | Z^2 \geq s^2] = s^2 + o(s^2)$, the result follows.

6.3 Additional comparisons with benchmarks

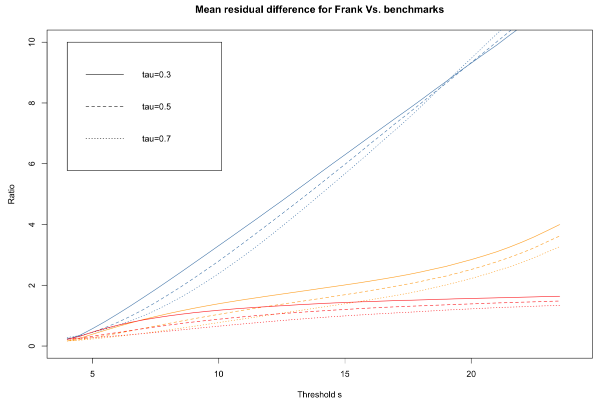

Figure 5: Evolution of the ratio of $E[(X-Y)|X \geq s]$ computed from the Frank copula model, with respect to the value of $E[(X-Y)|X \geq s]$ obtained in the benchmark settings. The orange lines correspond to benchmark B_1, the red ones to scenario B_2, the blue one to scenario B_3.

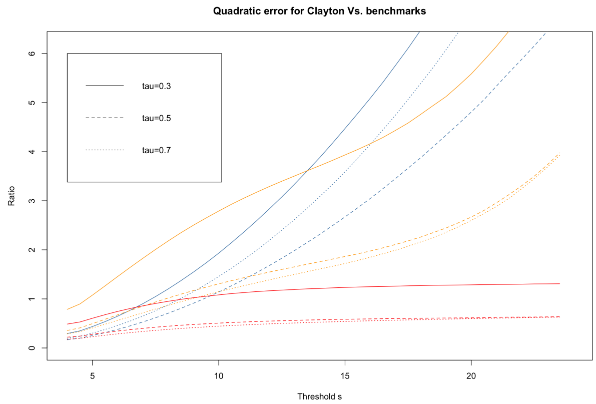

Figure 6: Evolution of the ratio of $E[(X-Y)^2|X \geq s]$ computed from the Frank copula model, with respect to the value of $E[(X-Y)|X \geq s]$ obtained in the benchmark settings. The orange lines correspond to benchmark B_1, the red ones to scenario B_2, the blue one to scenario B_3.

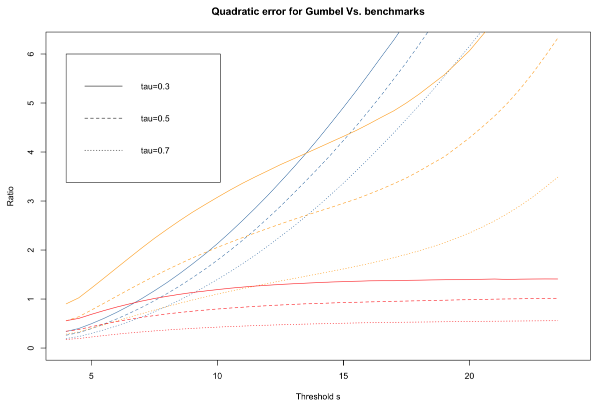

Figure 7: Evolution of the ratio of $E[(X-Y)^2|X \geq s]$ computed from the Frank copula model, with respect to the value of $E[(X-Y)|X \geq s]$ obtained in the benchmark settings. The orange lines correspond to benchmark B_1, the red ones to scenario B_2, the blue one to scenario B_3.

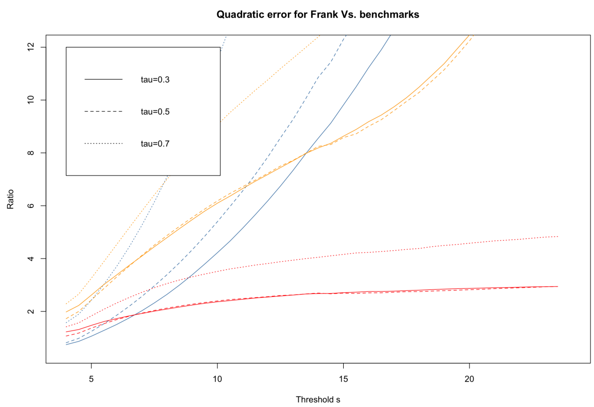

Figure 8: Evolution of the ratio of $E[(X-Y)^2|X \geq s]$ computed from the Frank copula model, with respect to the value of $E[(X-Y)|X \geq s]$ obtained in the benchmark settings. The orange lines correspond to benchmark B_1, the red ones to scenario B_2, the blue one to scenario B_3.

Acknowledgment: The authors acknowledge funding from the project *Cyber Risk Insurance: actuarial modeling*, Joint Research Initiative under the aegis of Risk Fundation, with partnership of AXA, AXA GRM, ENSAE and Sorbonne Université.

References

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. L. (2004). *Statistics of extremes: theory and applications*, volume 558. John Wiley & Sons.

Bodily, S. E. and Coleman, D. M. (2021). A portfolio of weather risk transfer contracts efficiently reduces risk. *Climate Risk Management*, page 100332.

Broberg, M. (2020). Parametric loss and damage insurance schemes as a means to enhance climate change resilience in developing countries. *Climate Policy*, 20(6):693–703.

Coles, S. (2001). An introduction to statistical modeling of extreme values.

Dal Moro, E. (2020). Towards an economic cyber loss index for parametric cover based on it security indicator: a preliminary analysis. *Risks*, 8(2):45.

Farkas, S., Lopez, O., and Thomas, M. (2021). Cyber claim analysis using generalized pareto regression trees with applications to insurance. *Insurance: Mathematics and Economics*, 98:92–105.

Figueiredo, R., Martina, M. L., Stephenson, D. B., and Youngman, B. D. (2018). A probabilistic paradigm for the parametric insurance of natural hazards. *Risk Analysis*, 38(11):2400–2414.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical proceedings of the Cambridge philosophical society*, volume 24, pages 180–190. Cambridge University Press.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, pages 423–453.

Horton, J. B. (2018). Parametric insurance as an alternative to liability for compensating climate harms. *Carbon & Climate Law Review*, 12(4):285–296.

Jacobs, J. (2014). Analyzing Ponemon Cost of Data Breach.

Johnson, L. (2021). Paying ex gratia: Parametric insurance after calculative devices fail. *Geoforum*, 125:120–131.

Lin, X. and Kwon, W. J. (2020). Application of parametric insurance in principle-compliant and innovative ways. *Risk Management and Insurance Review*, 23(2):121–150.

Maillart, T. and Sornette, D. (2010). Heavy-tailed distribution of cyber-risks. *The European Physical Journal B*, 75(3):357–364.

Mikosch, T. and Nagaev, A. V. (1998). Large deviations of heavy-tailed sums with applications in insurance. *Extremes*, 1(1):81–110.

Nelsen, R. B. (2007). *An introduction to copulas.* Springer Science & Business Media.

OCDE (2017). *Enhancing the Role of Insurance in Cyber Risk Management.*

Ponemon, L. (2018). Cost of a data breach study: global overview. *Benchmark research sponsored by IBM Security independently conducted by Ponemon Institute LLC.*

Prokopchuk, O., Prokopchuk, I., Mentel, G., and Bilan, Y. (2020). Parametric insurance as innovative development factor of the agricultural sector of economy. *AGRIS on-line Papers in Economics and Informatics*, 10(665-2021-565):69–86.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Van Nostrand, J. M. and Nevius, J. G. (2011). Parametric insurance: using objective measures to address the impacts of natural disasters and climate change. *Environmental Claims Journal*, 23(3-4):227–237.