From concentration profiles to concentration maps. New tools for the study of loss distributions.

Andrea Fontanari*- TU Delft and CWI
Pasquale Cirillo†- TU Delft
Cornelis W. Oosterlee - CWI and TU Delft

Abstract

In this work we introduce a novel approach to risk management, based on the study of concentration measures of the loss distribution.

In particular we show that indices like the Gini index, especially when restricted to the tails by conditioning and truncation, represent an accurate way of assessing the variability of the larger losses–the most relevant ones–and the precision of common risk management measures like the Expected Shortfall. We then introduce the *Concentration Profile*, that is a sequence of truncated Gini indices that, we show, is able to characterize the loss distribution, providing interesting information about tail risks.

Combining concentration profiles and standard results from utility theory, we then develop a *Concentration Map*, which can be used to assess the risk attached to potential losses on the basis of the risk profile of the user, her beliefs and historical data.

Finally, we use the sequence of truncated Gini indices as weights for the expected shortfall, defining the so-called *Concentration Adjusted Expected Shortfall*, a measure able to capture interesting additional features of tail risk.

All tools are applied to empirical data to show how to use them in practice.

*A.Fontanari@tudelft.nl
†P.Cirillo@tudelft.nl

1 Introduction

The aim of this paper is to introduce a way of dealing with loss distributions, tail risk, and risk management in general, based on a new class of tools derived from concentration measures.

The object of study is the loss distribution, for which we assume data is available. At present, we are not interested in analyzing the type of risk leading to the possible losses; it can be market risk, credit risk, operational risk or any other type of risk. In this work we restrict our attention to a static framework, in which the loss distribution is given and does not change over time. An example can be the one-year credit loss distribution, commonly used by banks under the Basel framework [2, 3].

As known, Value-at-Risk (VaR) and Expected Shortfall (ES) represent two important risk measures in modern risk management [17]. However, despite their popularity, these measures are not really able to convey reliable information on how losses are dispersed in the tail. It is not difficult to image different distribution tails sharing the same VaR and ES, but with different risk profiles, because of the different "fatness". A measure of the dispersion of the losses beyond Value-at-Risk is therefore needed. A measure that plays the role of precision parameter for the Expected Shortfall, informing how dispersed the losses are beyond VaR and around ES.

This measure should be able to deal with the special features of the loss function, especially those related to the upper tail, not to be biased and reliable. Given the typical heavy-tailed nature of large losses, any measure based on the variance should for example be avoided. In fact, empirical studies [17, 20, 23] show that often the assumption of finite second moment, i.e. $\mathbb{E}(Y^2) < \infty$, is too stringent for the random variable Y representing losses.

Our proposal is to make use of concentration measures [14, 21, 32] as a way of measuring tail risk dispersion. In particular, we focus our attention on the Lorenz curve $L(x)$ [22] and the corresponding Gini index $G(x)$ [19], for which we derive a "risk management" interpretation. In particular we show how the Gini index does not only provide a robust measure for the precision of the Expected Shortfall, but it can also be used as an alternative measure for the fat-tailedness of the loss distribution itself.

Naturally, other concentration measures like the Pietra index [27] could be used. However, preliminary studies have shown that, in applications, there is no particular gain in substituting the Gini index with the Pietra index, while the mathematical complexity increases.

It is important to stress that the use of concentration measures in finance and, at least partially, in risk management, is not completely novel. An interesting application of the Gini index as a substitute of the more common standard deviation is for example the Mean-Gini efficient portfolio frontier developed by Shalit and Yitzhaki [31], in which the usual portfolio theory is rewritten using the Gini index as the reference risk measure. However, as we will soon show, we here present a new application of these concentration measures, also introducing new tools.

Our final goal is not to criticize VaR and ES per se, but to support them, improve them, compensating for some of their shortcomings. VaR and ES are there to stay, because of the financial regulations [2, 3]. The idea is to make them more reliable.

In this work we discuss how a proper sequence of truncated Gini indices, which we call *Concentration Profile*, can be used to characterize the loss distribution, trying not only to understand the possible parametric family (lognormal, Pareto, Weibull, etc.), but also grasping some features that are not immediately observable from data, like the actual behavior of risk in the right tail. For example,

when considering Paretian tails, the prototype of fat tails, the concentration profile tells us that risk does not vary, as it is evident from the constant concentration profile of the Pareto distribution (see below). Paretian losses maintain their riskiness notwithstanding the VaR level we might be interested in; and this risk is naturally higher the fatter the right tail.

We provide a full description of the use of the concentration profile, also offering quick heuristics for its use in everyday business. A possible use of the concentration profile as a goodness-of-fit tool is also discussed.

We then introduce the so-called *Risk Concentration Map*. This is a graphic tool identifying the main risk factors encoded in the concentration profile, mapping them into an easily readable plot in which, through the use of a risk and utility function approach [10], we are able to attach a synthetic risk score to a given concentration profile. This risk concentration map can be seen as a "traffic light" to have a quick-yet-comprehensive assessment of the risk of losses. Additionally, it can also be used to compare different loss distributions in terms of their concentration risk, enabling a comparison between portfolios with different scales and magnitudes, given that our approach is essentially scale-free.

Finally, to complement the information provided by the concentration profile, we introduce a *Concentration Adjusted Expected Shortfall (CAES)*, which is nothing but the Expected Shortfall at level α, multiplied by the corresponding truncated Gini index. This quantity proves to be very useful in characterizing tail risk. In fact, while the ES is bound to increase, for increasing confidence levels, thus moving right into the upper tail, the corresponding Gini could compensate this growth, informing that the tail is becoming less extreme, and losses less erratic above a given threshold. We provide guidelines for the use of the CAES.

The paper is structured as follows: in Section 2 we review some basic concepts about concentration measures. In Section 3 some common measures of risk used in risk management are analyzed and put in relation with the concentration measures. In Section 4, we introduce and study the concentration profile, while in Section 5 we describe the concentration map. In Section 6 some additional extension based on the concentration adjusted expected shortfall are discussed. In Section 7 empirical results on simulated and actual data are provided. Section 8 closes the paper. At the end of the paper, an Appendix contains some more technical details of our work, like proofs and explicatory computations.

2 Basic concentration quantities

2.1 The Lorenz Curve

At the beginning of the XX century, economists and social scientists started discussing about possible ways of measuring how wealth is distributed among individuals. In particular, the main question was if it were possible to deliver a compact measure–free of scale–that could allow for comparisons among different countries at different times. A relevant answer to this question was offered by Max Lorenz in 1905 with his pioneering work on the concentration of wealth [22]. The paper introduced what is now known as the Lorenz Curve, a fundamental building block for studying the distribution and concentration of wealth among individuals. The success of the Lorenz curve was so strong that

it soon started being used in other fields of science beyond economics, such as physics [15], survival analysis and biostatistics [4], and so on [13, 16].

Consider a positive continuous random variable $Y \in \mathbb{R}^+$, belonging to the $\mathcal{L}^1$ class, such that $\mu = \mathbb{E}(Y) < \infty$. Let $F(y) = \mathbb{P}(Y \leq y)$ be its cumulative distribution function. The quantile function of Y is $Q(u) = F^{-1}(u)$, where $F^{-1}(u) = \inf\{y : F(y) \leq u\}$ with $0 \leq u \leq 1$.

The Lorenz curve $L(x)$ is formally given by

$$L(x) = \frac{\int_0^x Q(u)du}{\int_0^1 Q(u)du},\ 0 \leq x, u \leq 1. \tag{1}$$

In terms of wealth, the Lorenz curve reads as follows: for a given $x \in [0,1]$, $L(x)$ tells us that $x \times 100\%$ of the population owns $L(x) \times 100\%$ of the total wealth. From this interpretation we see that the scale–the total amount of wealth–is not taken into consideration, whereas the way it is distributed among the individuals is the key information.

Statistically, the Lorenz curve $L : [0,1] \to [0,1]$ defined in equation (1) is a continuous, non-decreasing, convex function, differentiable almost everywhere in $[0,1]$, such that $L(0) = 0$ and $L(1) = 1$. It is bounded from above by the so called perfect equality curve, i.e. $L_{pe}(x) = x$, and from below by the perfect inequality curve, i.e.

$$L_{pi}(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1 \end{cases}$$

The perfect equality line L_{pe} indicates the purely theoretical situation in which everyone possesses the same amount of wealth in the economy. Conversely, the perfect inequality line L_{pi} states that only one individual owns all the wealth in the society. A visual representation of a possible Lorenz curve is given in Figure 1, where we also provide L_{pe} and L_{pi}.

Given its strong relation with the quantile function Q, the Lorenz curve is able to fully characterize a distribution [21]. If we know the exact functional form of $L(x)$, we can recover $F(y)$. However, graphically, it is not possible to discriminate among distributions just looking at their Lorenz curves [7]. A graph like the one in Figure 1 can give us a flavor of how unequal a society is, but it cannot tell us whether the underlying distribution is a Weibull or rather a lognormal, just to name two distributions.

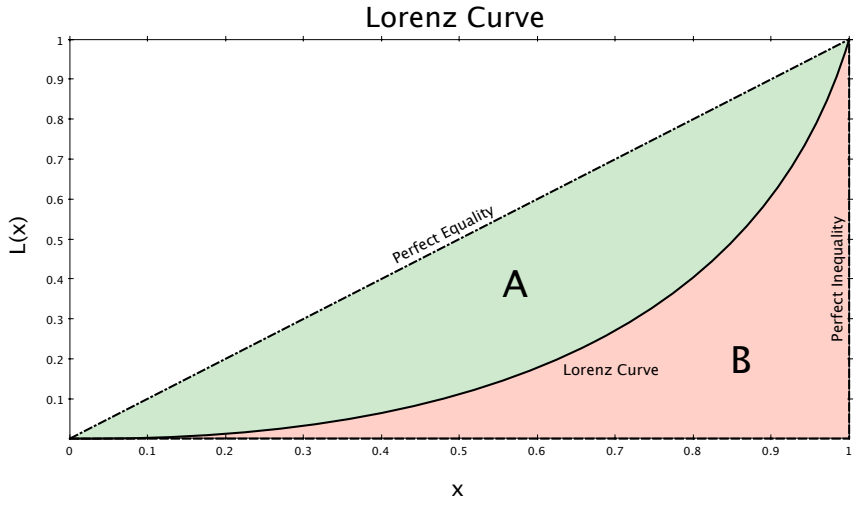

Figure 1: Graphical representation of a Lorenz Curve, of the perfect equality line and of the perfect inequality case. We also show the geometric interpretation of the Gini index as the ratio of area A over A+B.

2.2 The Gini Index

Strictly related to the Lorenz Curve are the so-called concentration indices, the most celebrated, used and abused one being the Gini index [19]. Naturally, many other indices based on the Lorenz curve have been introduced, e.g. the Pietra index [27] and the Generalized Gini [25]. We do not cover them here, but we refer to [14] for a review.

All concentration indices can be derived and computed in many different yet equivalent ways. In [32] more than fourteen different possibilities are given to express the Gini index alone. Here we will just focus on a few of them.

The first characterization is based on the Lorenz curve, in particular on the distance between the perfect equality line $L_{pe}(x)$ and the Lorenz curve $L(x)$. The idea is very simple: the closer these two lines are, the more equal the society is, and the concentration of wealth lower. The distance between these curves can be expressed as

$$D_p = \frac{\left\| L(x) - L_{pe}(x) \right\|_p}{\left\| L_{pi}(x) - L_{pe}(x) \right\|_p}, \qquad 0 \leq D_p \leq 1, \tag{2}$$

where the denominator is just a normalizing factor.

From equation (2) it is clear that, by varying the type of distance $\|\cdot\|_p$, we can get different types of indices. The Gini index $G \in [0,1]$ is obtained by fixing $p = 1$. Formally,

$$\begin{aligned} G = D_1 &= \frac{\int_0^1 (L(x) - x)dx}{-\int_0^1 x dx} \\ &= 1 - 2\int_0^1 L(x)dx \end{aligned} \tag{3}$$

An alternative way to define the Gini index is purely geometrical. As also shown in Figure 1, G can be defined as the ratio between area A, the area between the Lorenz curve and the perfect equality line, and the total area $A + B$ lying below the perfect equality, i.e.

$$G = \frac{A}{A+B}.$$

Another interesting formulation is the so-called "stochastic" representation of the Gini index [14], for which

$$G = \frac{\mathbb{E}(\mid Y_1 - Y_2 \mid)}{2\mathbb{E}(Y)} = \frac{\mathbb{E}((Y_1 - Y_2)^+)}{\mathbb{E}(Y)} \in [0,1], \tag{4}$$

where Y_i, $i = 1,2$ are i.i.d copies of the same random variable $Y \sim F(y)$, and $(Y_1 - Y_2)^+ = \max\{Y_1 - Y_2; 0\}$.

Looking carefully at formula (4), we discover that the Gini index is nothing but the normalized average of the distance between two random independent observations taken from the underlying distribution $F(y)$. Thanks to the expression as expectation of $\mathcal{L}^1$ distances, the index can be rephrased as expectation of $\max\{Y_1 - Y_2, 0\}$. This formulation shows that the Gini index can be considered as an $\mathcal{L}^1$ alternative to the skewness, for measuring the asymmetry of the data, in particular to the right [14].

It may useful to compare the Gini index with some variance based measure like the variance-to-mean ratio defined as

$$VM = \frac{\mathbb{E}(Y - \mathbb{E}(Y))^2}{\mathbb{E}(Y)}.$$

The VM ratio measures the dispersion of data around the mean, without considering the direction of this dispersion, thus implicitly assuming them to be symmetric. However, as usual in risk management, this is not the case. In particular, if the distribution of losses is asymmetric, high values of VM are usually driven by the apparent "outliers" that occur because of asymmetry rather than for a real dispersion of the data. Additionally, since the distance is squared, this error is even magnified, while small deviations are compressed, thus leading to misleading values.

The Gini index, on the contrary, works well when asymmetries arise, since it implicitly assumes a "direction" for the dispersion of the observations: the right tail. Therefore it provides a better and more precise measure of their variability.

Another important feature which makes the Gini superior to variance-based measures is its formulation in terms of the $\mathcal{L}^1$ distance. The computation of the Gini index thus only requires the mean to be finite, while all higher moments can be infinite. This is not the case when dealing with the variance-to-mean ratio. Imagine our data come from a Pareto distribution whose c.d.f is given in Table 1 and assume $\rho \in (1,2]$. In this case, any measure which requires a squared expectation such as the variance cannot be meaningfully computed, because it is theoretically infinite. Conversely, the Gini index can be easily obtained.

Let's summarize the main properties of the Gini index:

- The Gini index is bounded between 0 and 1; $G = 0$ means that the Lorenz curve lies exactly on

the perfect equality line, while $G = 1$ means that the Lorenz curve corresponds to the "perfect inequality" line.

- The Gini index is an $\mathcal{L}^1$- measure, meaning that it can be computed for all random variables admitting a finite mean, without any further requirement. This is not true for measures like VM for which the second moment needs to be finite.

- The Gini index is a quasi-convex measure [5], and this has clear implications in terms of risk management. Quasi-convexity is indeed a more realistic relaxation of convexity, with important consequences for sub-additivity and risk diversification. In particular, quasi-convexity allows to deal better with distributions of risk that are not necessarily closed under convolution.

In Table 2 some examples of closed formulas for the Gini indices and the Lorenz curves for the same distributions of Table 1 are given.

Distribution	**c.d.f**	**p.d.f.**	**Support**	**Shape**
Weibull	$F(y) = 1 - e^{-(\frac{y}{\lambda})^\gamma}$	$f(y) = \frac{\gamma}{\lambda}(\frac{y}{\lambda})^{\gamma-1}e^{-(\frac{y}{\lambda})^\gamma}$	$y \in \mathbb{R}^+$	$\gamma \in \mathbb{R}^+$
Lognormal	$F(y) = \Phi(\frac{\ln(y)-\mu}{\sigma})$	$f(y) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(\ln(y)-\mu)^2}{\sigma^2}}$	$y \in \mathbb{R}^+$	$\sigma \in \mathbb{R}^+$
Exponential	$F(y) = 1 - e^{-\frac{y}{\lambda}}$	$f(y) = \lambda e^{-y\lambda}$	$y \in \mathbb{R}^+$	None
Pareto	$F(y) = 1 - (\frac{y}{y_0})^{-\rho}$	$f(y) = \frac{\rho y_0^\rho}{x^{\rho-1}}1_{\{y \geq y_0\}}$	$y \in [y_0, +\infty)$	$\rho \in \mathbb{R}^+$

Table 1: C.d.f., p.d.f., support and shape parameter of some of the most used distributions in risk management.

Distribution	**Gini Index**	**Lorenz Curve** $x \in [0,1]$
Weibull	$1 - 2^{-\frac{1}{\gamma}}$	$L(x) = 1 - \frac{\Gamma(-\log(1-x), 1+\frac{1}{\gamma})}{\Gamma(1+\frac{1}{\gamma})}$
Lognormal	$2\Phi(\frac{\sigma}{\sqrt{2}}) - 1$	$L(x) = \Phi(\Phi^{-1}(x) - \sigma)$
Exponential	$\frac{1}{2}$	$L(x) = x + x(-\log(1-x)) + \log(1-x)$
Pareto	$\frac{1}{2\rho-1}$	$L(x) = 1 - (1-x)^{\frac{\rho-1}{\rho}}$

Table 2: Gini index and Lorenz curve for the same distributions of Table 1 .

3 Basic concepts of Risk Management

From now on, $Y \in \mathbb{R}^+$ is the continuous random variable representing losses, and $F(y)$ is the corresponding loss distribution function.
Regarding the loss distribution, some stylized facts can be derived from empirical analysis, e.g. [21, 23], and it is useful to keep them in mind, when dealing with risk measures. In a nutshell:

- The loss distribution is asymmetric and right-skewed [21]. This typically means that the right tail has more mass than the left one, when losses are taken to be positive quantities and profits are negative.

- The loss distribution is usually fat tailed [17, 21]. Paretianity emerges on the right tail, often leading to the non-existence of higher-order moments. Very unpleasant things are likely to happen. This brings to the next point.
- The loss distribution is typically $\mathcal{L}^2$ integrable, but there are several cases in which it is only $\mathcal{L}^1$ integrable [23]. The practical implication is clear: measures requiring the existence of the moments that may not exist are not appropriate under fat-tailed losses[1].

In risk management, we are often not interested in studying the entire distribution of losses, but a part of it, usually the right tail, where the larger losses concentrate. Most Basel regulations [2, 3], but also Solvency II [29] for insurance companies, deal with the large unexpected losses, the few game-changers, not the many small negligible losses we can hedge. This requires us to deal with truncated random variables and distributions, rather than with the original ones.
Below we provide some basic quantities that we will use in the rest of the paper.

Definition 1. Given a random variable $Y \in \mathbb{R}^+$ with c.d.f. $F(y)$ and p.d.f. $f(y)$, its (left-) truncated counterpart Y_T has c.d.f. $F_T(y)$ at truncation level $a > 0$ defined as follows:

$$F_T(y) = \frac{F(y) - F(a)}{1 - F(a)},\ a \leq y \leq \infty,$$

and p.d.f

$$f_T(y) = \frac{f(y)}{1 - F(a)}.$$

The quantities $F_T(y)$ and $f_T(y)$ are known as exceedance distribution and exceedance density of the random variable Y respectively.

Definition 2. Given a confidence level $\alpha \in (0, 1)$, the Value-at-Risk (VaR_α) is the statistical quantile of the loss distribution function $F(y)$ defined as

$$VaR_\alpha = \inf\{y \in \mathbb{R} : P(Y \geq y) \leq 1 - \alpha\} = \inf\{y \in \mathbb{R} : F(y) \geq \alpha\}$$

Definition 3. Given a VaR_α, the Expected Shortfall ES_α, for $Y \in \mathbb{R}^+$ with c.d.f. $F(y)$, is given by

$$ES_\alpha = \mathbb{E}(Y \mid Y > VaR_\alpha)$$

Interestingly, the ES_α is nothing more than the mean of the truncated random variable Y_T, when the truncation occurs at truncation level $a = VaR_\alpha$. Just notice that

$$\mathbb{E}(Y \mid Y > a) = \int_a^{+\infty} y f_T(y) dy, \qquad a > 0.$$

VaR_α and ES_α are two fundamental measures of risk in modern risk management. It is well-known that while the expected shortfall is always coherent (positive homogeneous, monotone, translation invariant and sub-additive), the value-at-risk is not, unless we restrict our attention to elliptic

[1] In the case of operational risk, even the first moment, the mean, may be infinite [8, 24], but we ignore this radical case here. Naturally, in such a situation, the Gini index would not be defined.

loss distributions and comonotonic portfolios [17, 28]. This is why, in the last years, ES_α seems to be preferred by regulators [3, 29], even if both measures have been heavily criticized by experts [23], for their incapacity of dealing with the variability of losses in the tails.

Both VaR_α and ES_α are in fact unable to convey information on how losses are dispersed in the tail. As shown in Figure 3, two loss distributions with very different tails may share the same value-at-risk and the same expected shortfall. This is why a measure of dispersion of the losses in the tail, beyond the VaR_α level is needed. A measure that can be used to understand how reliable the ES_α is.

However, given the stylized facts we have listed above, this measure cannot rely on variances or higher moments, as some authors suggest [30] by assuming elliptically contoured loss distributions. It must be a measure able to deal with the special features of the loss function, and of the right tail, and not being biased by them. Our proposal is to look at the Gini index.

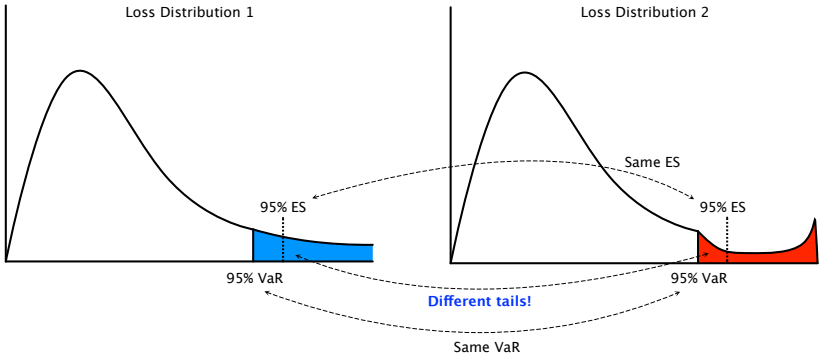

Figure 2: In this example two loss distributions with same $VaR_{0.95}$ and same $ES_{0.95}$ are shown. However, the structure of their tails is different meaning that we need an additional measure for the dispersion of the losses in the tail to grasp this difference.

By construction, a higher Gini index indicates that more losses are present far in the right tail–losses may play the role of the rich individuals in the distribution of wealth. A lower Gini, conversely, indicates a distribution of losses which is more concentrated around the same values. Therefore, two distributions sharing the same VaR and ES but different tails are likely to have different Gini indices.

However, the Gini index, in the formulation we have presented so far, is not optimal, as it takes into account all values of the distribution of losses. What we need is a truncated Gini index, which only measures dispersion above a certain threshold–the VaR, so that we can define a more precise measure of tail risk and ES precision. The next section, and in particular Subsection 4.1 is devoted to this.

4 The Concentration Profile

Despite being able of characterizing the distribution of losses, the Lorenz curve is not a viable tool in risk management applications. This is due to the fact that $L(x)$ does not provide a unique value, but rather a continuum of information. Moreover, a graphical inspection of the Lorenz curve does not allow the identification of any particular distribution [21].

Conversely, a concentration measure like the Gini index, while it contains important pieces of information about the distribution it is derived from, which it summarizes in a single value between 0 and 1, is not sufficient to completely characterize the tail of a distribution, given that it relies on the entire support of random variable Y. Moreover, and most importantly, notwithstanding the fact that the Gini index G is obtained from $L(x)$, there is no one-to-one mapping between them. The same value of G can correspond to an infinte numbers of Lorenz curves [32].

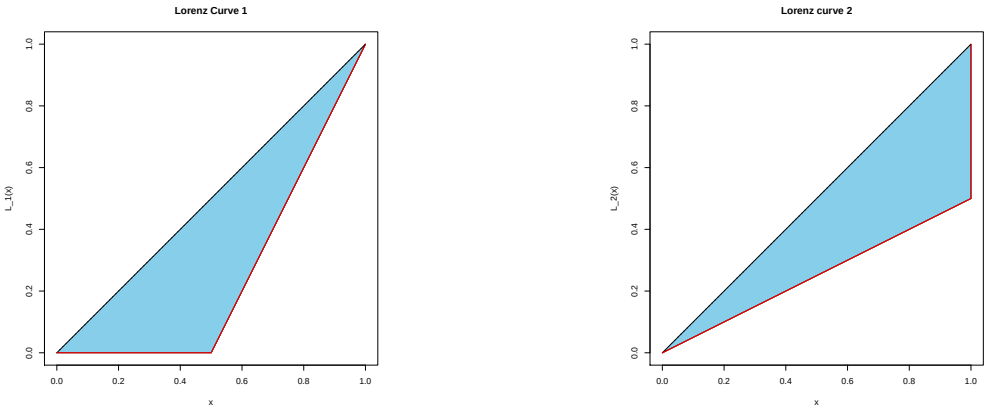

Figure 3: In the first graph the Lorenz curve $L_1(x)$ is shown, while in the second $L_2(x)$. The area shaded between the L_{pe} and the two Lorenz curves is the same, delivering the same Gini index, however the asymmetry is clearly different.

In Figure 3 two different Lorenz curves are shown. Their functional forms are

$$L_1(x) = (2x - 1)1_{\{x \geq 0.5\}}$$

for the one on the left, and

$$L_2(x) = 0.5x$$

for that on the right.

Fixing $x = 0.9$, gives $L_1(x) = 0.8$ and $L_2(x) = 0.45$. This means that, according to L_1, the top 10% losses account for 20% of the total loss. These numbers change under L_2, where the same top 10% accounts for 55% of all losses! In both cases, the Gini index is the same: $G_1 = \frac{1}{2} = G_2$. Graphically speaking, it is sufficient to notice that the two blue triangles in Figure 3 occupy the same area. This simple example thus shows that the Gini index alone is not able to capture the dissimilarity between the two loss distributions, despite their very different risk profiles.

The Gini index can be the basic ingredient of a more sophisticated tool, able to distinguish among

different risk profiles and distributions of losses. This tool is the *Concentration Profile* (CP), and it is built on a sequence of truncated Gini indices.

The basic idea is to order all observations in the sample of losses. Then, recursively, we exclude the points at the left-hand side, computing the Gini index for the remaining ones. In other words, we truncate the distribution of losses at every single loss, from left to right, obtaining a collection of exceedance distributions. For each of these, we compute the Gini index–the truncated Gini index with respect to the original Y. We then study the behavior of the Gini index as a function of the truncation level.

Before considering the mathematical details, we give an example, to show that this technique makes sense. Let's come back to the example of Figure 3. We truncate the original data at $x = 0.5$, i.e. ignore the first 50% of the two samples. The new Lorenz curves are given in Figure 4.

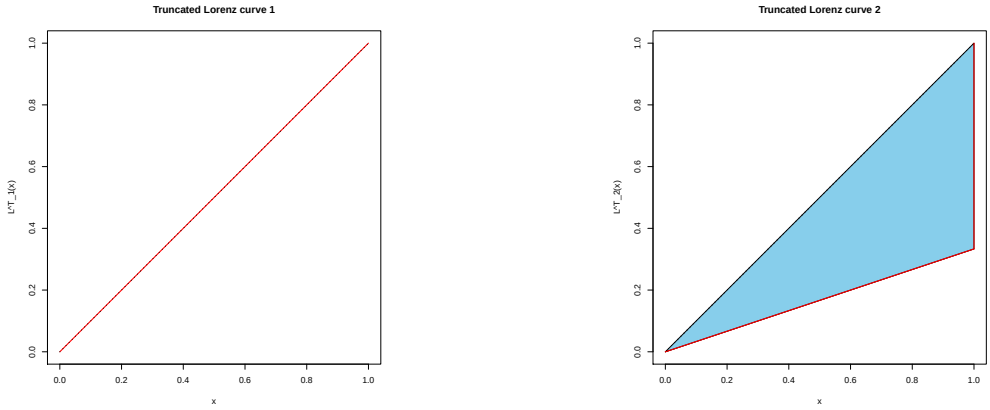

Figure 4: The rescaled versions of the Lorenz curves of Figure 3, where only the upper 50% of the domain is considered. Notice how the shaded area is absent in the left side meaning that the Gini of $L_1^{0.5}(x)$ is zero.

It is not difficult to prove that, analytically, the two curves are now $L_1^{0.5}(x) = x$ and $L_2^{0.5}(x) = \frac{1}{3}x$ (they can be obtained using the equation (5) we introduce later on).

It is clear that the second situation is riskier than the first one, given the higher inequality shown. This is also reflected by the new Gini indices, which now differ: $G_1^{0.5} = 0 < G_1$ and $G_2^{0.5} = \frac{2}{3} > G_2$. This clearly denotes a different risk concentration in the tails.

In a concentration profile, we perform the same computation as above at every point of the support of the loss distribution. We get a sequence of truncated Gini indices, which is rather informative about tail risk.

4.1 Mathematical construction

To rigorously define the concentration profile, we need some preliminary notation.

Definition 4. Consider a random variable $Y \in \mathbb{R}^+$ with distribution $F(y)$, and a confidence level $\alpha \in [0, 1)$. We call *risk subclass* at level α, the truncated support of $F(y)$ defined as

$$S_\alpha = \left[F^{-1}(\alpha), +\infty\right) = [Q(\alpha), +\infty)\,.$$

In risk management, if a VaR_α is defined, we clearly have

$$S_\alpha = [VaR_\alpha, +\infty)\,.$$

As $\alpha \nearrow 1$, the left point of S_α approaches the right-end point of the distribution $F(y)$, i.e. $y_F = \sup\{y : F(y) < 1\}$.

For every α, let's define the truncated Lorenz curve $L_\alpha(x)$, whose derivation is available in the Appendix to this paper. In formula,

$$L_\alpha(x) = \frac{L(\alpha + (1-\alpha)x) - L(\alpha)}{1 - L(\alpha)}, \qquad 0 \leq \alpha, x \leq 1, \tag{5}$$

where $L(\cdot)$ is the original Lorenz curve.

Similarly, we can define the truncated distance index

$$D^T_\alpha = \frac{\left\| L_\alpha(x) - L_{pe} \right\|_p}{\left\| L_{pi} - L_{pe} \right\|_p} \qquad 0 \leq D^T_\alpha \leq 1, \tag{6}$$

and, setting $p = 1$ as before, the truncated Gini index

$$G(\alpha) = 1 - 2\int_0^1 L_\alpha(x)dx = 1 - 2\int_0^1 \frac{L(\alpha + (1-\alpha)x) - L(\alpha)}{1 - L(\alpha)}dx. \tag{7}$$

In terms of VaR_α and ES_α, the previous equation can be stated as

$$G(\alpha) = \frac{\mathbb{E}(|Y_1 - Y_2| \mid Y_1 \wedge Y_2 > VaR_\alpha)}{2ES_\alpha}.$$

The truncated Gini index is thus a measure of the "absolute" dispersion of the losses in the subclass S_α. High values of $G(\alpha)$ imply that the losses in a given subclass S_α are far apart, resulting in a persistent thickness of the tail in the subclass.

Since the truncated Gini index derives from $L_\alpha(x)$, which depends on $F_{VaR_\alpha}(y)$, each truncated Gini conveniently lies in the interval $[0, 1]$. It is therefore natural to read the Gini index as a probability.

Using Markov inequality, we obtain the following non trivial bound:

$$\mathbb{P}(|Y_1 - Y_2| > 2ES_\alpha) \leq G(\alpha) \leq 1. \tag{8}$$

We now have all ingredients to define the concentration profile formally:

Definition 5. Given a sequence of risk subclasses $\{S_\alpha\}$, $\alpha \in [0, 1)$, we call *Concentration Profile* (*CP*) the sequence $\{G(\alpha)\}_{\alpha\in[0,1)}$, defined by equation (7).

Graphically speaking, a CP can be visualized by plotting $G(\alpha)$ against increasing values of α. The concentration profile is therefore contained in the $[0, 1] \times [0, 1]$ square making it easy to interpret and asses. Values close to 0 are points of minimum concentration (variability), while values close to 1 are points of maximum concentration (variably). Through the observation of the behavior of its CP, we can distinguish interesting features of the risk profile of the loss distribution under scrutiny.

In Figure 5 the theoretical *Concentration Profile* for four main classes of loss distributions is shown: Pareto, lognormal, Weibull and exponential (notice that an exponential is just a Weibull with shape parameter $\gamma = 1$).

Figure 5 already gives us an useful fact, which we further analyse in the following subsection: the Pareto distribution is the only distribution characterized by a constant, horizontal concentration profile.

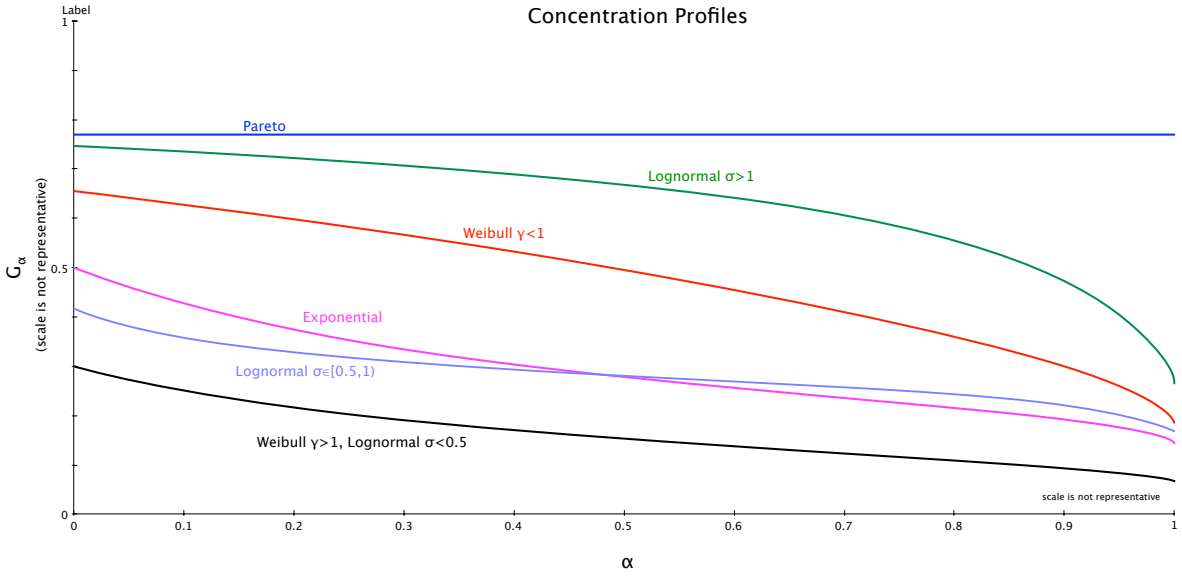

Figure 5: Theoretical concentration profile

4.2 Characterization of the Concentration Profile

Looking at Figure 5, the first natural question to ask is whether the concentration profile uniquely identifies a distribution. The answer is affirmative and is given by the following theorem, the proof of which can be found in [26]:

Theorem 6. *If $G(\alpha)$ is differentiable, then*

$$Q(\alpha) = \frac{(1-\alpha)G'(\alpha) - G(\alpha) + 1}{(1-\alpha)(1+G(\alpha))} \exp\{-\int \frac{(1-\alpha)G'(\alpha) - G(\alpha) + 1}{(1-\alpha)(1+G(\alpha))} d\alpha\}, \tag{9}$$

where $G'(\alpha) = \frac{dG(\alpha)}{d\alpha}$ and $Q(\alpha)$ is the quantile function of a general distribution.

A second question is whether different shapes of the CP define specific classes of distributions.

In order to answer this question we need to identify some limit distributions for light and fat tails.

A natural choice is to focus on the Generalized Pareto Distribution or GPD [9, 17]. This distribution is a well-known family used in extreme value statistics, and plays an important role in defining the properties of tails. Formally, $Z \geq 0$ follows a GPD if its distribution function H is such that

$$H(z) = \begin{cases} 1 - (1 + \xi z)^{-1/\xi} & \xi \neq 0 \\ 1 - e^{-z} & \xi = 0 \end{cases}. \tag{10}$$

This distribution can be shown to be the limit distribution of all distribution tails, provided that we move sufficiently far on the right [9, 17], setting high thresholds in defining what the tail is. This is known as the GPD approximation.

Here we focus on the case $\xi \geq 0$, because when $\xi < 0$ we are mainly dealing with bounded support distributions [17, 23]. A value of $\xi > 0$ usually indicates a fat tail. The higher ξ, the fatter. For us, anyway, ξ needs to be smaller than 1, given that for $\xi > 1$ the mean of the GDP is not finite, hence we cannot define the Gini index. The case $\xi = 0$ corresponds to thin tails, whose limit behavior is that of an exponential.

We state the following result, whose proof is to be found in the Appendix.

Theorem 7. *Consider the Generalized Pareto Distribution, its concentration profile is given by the following formula:*

$$G(\alpha) = \begin{cases} \frac{\xi}{2-(1-\alpha)^{\xi}(-2+\xi)(-1+\xi)-\xi} & \xi \in (0,1) \\ \frac{1}{2-2\log(1-\alpha)} & \xi = 0 \end{cases}. \tag{11}$$

for $\alpha \in [0, 1)$.

Looking at the two expressions in formula (35) we distinguish the following behavior, for every $\alpha \in [0, 1]$:

- For what concerns the light tail domain($\xi = 0$), $G(\alpha)$ exhibits a decreasing behavior.
- For the fat-tailed domain($\xi > 0$), as the shape parameter ξ grows to 1, the slope of the CP tends to zero. This means that the fatter the tail, the more constant the CP. For a purely fat-tailed distribution as the Pareto, and only for the Pareto, we get the horizontal line in Figure 5.

In case of a purely Pareto distribution, the GPD approximation is not an approximation, as there is a perfect overlap for the entire support of the distribution. In this case, the ξ of the GPD corresponds to $1/\rho$ of the Pareto distribution, following the definition of Table 1. Therefore, for $\rho \searrow 1$, we have $\xi \nearrow 1$, and the concentration profile translate towards 1, indicating maximum concentration for all values of α. As we show in Table 3, the Pareto distribution is the only distribution whose Gini index is immune to truncation, as it does not depend on any threshold value.

Conversely, if we look at Figure 5, the lognormal, the Weibull and the exponential distributions show a decreasing CP. Can we extrapolate some characterizing behavior?

We analyze the behavior of the exponential distribution (equivalently the Weibull distribution with shape parameter $\gamma = 1$), which is also a limiting case for the GPD approximation. Its theoretical concentration profile is given in Table 3 (as an example of derivation, we show the full procedure in the Appendix). Interestingly, it does not depend on any parameter. This means that all exponential distributions share the same CP, no matter the value of their intensity. The exponential distribution CP has a $G(0) = 0.5$, for $\alpha = 0$, and it decreases towards zero first convexly and successively concavely, with a point of flex at $\alpha = 0.63$. Any empirical loss distribution whose $G(0)$ is far away from 0.5 cannot be of the exponential type.

Since the Weibull distribution is a generalization of the exponential, there is an interesting relation between the two. Any Weibull distribution with shape parameter $\gamma > 1$ has a CP that always lies below the CP of the exponential distribution without any intersection. For $\gamma < 1$, the Weibull distribution CP lies above the exponential distribution CP. This behavior agrees with theoretical results pointing out a phase transition in the Weibull distribution between $\gamma < 1$ and $\gamma > 1$ [21].

In terms of risk, when the underlying distribution is of the Weibull type, we can read the values of $\gamma < 1$ as riskier than exponential. Vice versa for $\gamma > 1$.

Regarding the lognormal family, the CP has some relationship with the exponential distribution.

In particular when the lognormal distribution has a shape parameter $\sigma < 1$ (the standard deviation of the normal distribution of the logs), its starting point is $G(0) < 0.5$. On the contrary, when $\sigma > 1$, $G(0) > 0.5$.

Compared to the Weibull's, the lognormal CP is always flatter, indicating that risks tend to decrease more slowly in the tail–consistent with extreme value theory, where the lognormal distribution is considered riskier than the Weibull [23]. In order to find a lognormal CP that always lies below the exponential CP, without any intersection, we need $\sigma < 0.5$. On the opposite side, in order to find a lognormal CP which does not intersect the exponential CP from above, we have to raise the shape parameter up to $\sigma > 1$.

We can derive some general principle to asses the riskiness of a distribution using the CP. This issue will be discussed in the next Section, where we introduce the *Concentration Map*.

In Table 3 we provide the expression for the *Concentration Profile* for exponential, Weibull, Pareto and lognormal distributions.

While theoretically the CP may exhibit any type of behavior, we have verified that for the most used loss and size distributions (see [21] for a review), the behavior is always non-increasing. Additionally, it is important to notice that the distributions discussed in this work are somehow representative of many other useful size distributions. An example is given by the Gamma distribution, whose behavior can be represented by a combination of the Weibull and the exponential distribution.

Distribution	**Concentration Profile** $G(\alpha)$
Weibull	$G(\alpha) = \frac{-2^{-\frac{1}{\gamma}}\Gamma(1+\frac{1}{\gamma},-2\log(1-\alpha))-(-1+\alpha)\Gamma(1+\frac{1}{\gamma},-\log(1-\alpha))}{(1-\alpha)\Gamma(1+\frac{1}{\gamma},-\log(1-\alpha))}$
Lognormal	$G(\alpha) = \frac{\int_\alpha^1(-1+2x-\alpha)e^{-\sqrt{2}\sigma\mathrm{erfc}^{-1}(2x)}dx}{(1-\alpha)\int_\alpha^1 e^{-\sqrt{2}\sigma\mathrm{erfc}^{-1}(2x)}dx}$
Exponential	$G(\alpha) = \frac{1}{2-2\log(1-\alpha)}$
Pareto	$G(\alpha) = \frac{1}{(2\rho-1)}$ for $\rho \in (1,+\infty)$

Table 3: Theoretical Concentration Profile for some of the most famous loss distributions, here $\Gamma(\cdot,\cdot)$ is the incomplete Gamma function i.e. $\Gamma(a,b) = \int_b^{+\infty} e^{-x}x^{a-1}dx$ and $\mathrm{erfc}^{-1}(x)$ is the complementary error function.

5 The Concentration Map

Given a concentration profile, a natural way to asses the risk encoded in a distribution is to look at its behavior.

However, as the Gini index provides a single value for the Lorenz curve, can we define something similar for the concentration profile? Can we simplify the job of the risk manager by providing a flexible tool that summarizes the information of interest: the riskiness of the loss distribution based on its concentration structure?

The answer is provided by the so-called *Concentration Map*.

5.1 Risk Drivers

The first step for constructing a concentration map for the loss distribution, on the basis of its CP, is to identify what we define the "risk drivers" of the concentration profile.

The first risk driver r_1 is simple.

Definition 8. Given a Concentration Profile $\{G(\alpha)\}_{\alpha\in[0,1)}$, the first risk driver $r_1 \in [0,1]$ is given by the starting value: $r_1 = G(0)$.

The definition of r_1 follows from the fact that, at least for unimodal size distributions, the concentration profile is a non-increasing function of the truncation level. Therefore, given different CPs, we expect the one with a higher starting value $G(0)$ to be more risky.

For example, consider the CP of a Pareto distribution (Table 3). We know that it is a constant function of the parameter $\rho \in (1, +\infty)$, i.e. $G(\alpha) = G(0)\ \forall\alpha$. Additionally, after an easy computation, we obtain that $\frac{dG(0)}{d\rho} = -\frac{2}{(-1+2\rho)^2} < 0$, for every value of $\rho \in (1, +\infty)$. Therefore, the higher the starting value of the CP, the lower the value of ρ, and the fatter and riskier the loss distribution.

Regarding the second risk driver r_2, we state the following.

Definition 9. Given a concentration profile $\{G(\alpha)\}_{\alpha\in[0,1)}$, the second risk driver $r_2 \in [0,1]$ is given by the difference between the first risk driver $r_1 = G(0)$, and $G(\alpha)$ with $\alpha \nearrow 1$. Namely:

$$r_2 = \lim_{\alpha\to 1} |G(0) - G(\alpha)| .$$

From Theorem 7 we know that, for light-tailed distributions, the CP converges to zero as $\alpha \nearrow 1$. This is because, as we move towards the right endpoint of the domain, the number of observations in the far right tail decreases, since the mass under the tail of the p.d.f. must converge to zero, in order to ensure the existence of the moments. This does not necessarily happen when the loss distribution is fat-tailed.

For this reason, measuring the gap between the initial and final value of the CP is another quantitative way of assessing the fatness of the tail. In particular, the smaller the gap, the larger the mass present in the tail, and the higher the probability of experiencing extreme losses.

In Fig 6, we show a graphical representation of r_1 and r_2.

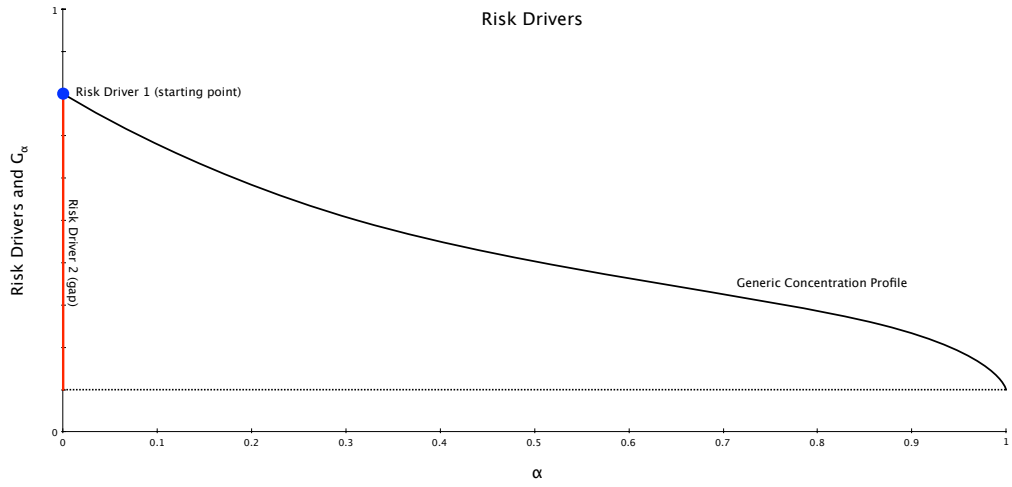

Figure 6: Risk drivers visualization with respect to a generic concentration profile

5.2 The Map

Once we have identified the main risk drivers, we need a way of combining them to rank different concentration profiles on the basis of the embedded risk.

We need a risk function $R(r_1, r_2) : [0,1] \times [0,1] \to [0,1]$ which assigns to the combination of the risk drivers of the concentration profile a number–w.l.o.g. in $[0,1]$–representing the risk level of the profile itself.

Trivially, this risk function must satisfy the constraints given in equations (12) and (13):

$$\frac{dR(r_1, r_2)}{dr_1} > 0, \tag{12}$$

and

$$\frac{dR(r_1, r_2)}{dr_2} < 0. \tag{13}$$

We denote by Γ the set of all functions $R(x, y) : [0,1] \times [0,1] \to [0,1]$ that respect the constraints in (12) and (13). Therefore, any function in this set is a candidate to assess the risk given by the CP.

If we additionally impose the constraint of a non-increasing concentration profile ($\frac{dG(\alpha)}{d\alpha} \leq 0$), which is the case for many unimodal distributions used in risk management, we can reduce the set Γ to contain functions whose domain is the $\mathbb{R}^2$ simplex with vertices $[0,0], [1,1], [1,0]$. This happens since, with a non-increasing concentration profile, $r_2 \leq r_1$.

Looking for a suitable risk function, we can rely on some well-known results from utility and risk theory. For example, we can choose a Cobb-Douglas risk function like

$$R(r_1, r_2) = r_1^a (1 - r_2)^b \qquad a, b \in \mathbb{R}^+. \tag{14}$$

In this function we can decide which risk driver is more relevant for risk managers, by acting on the parameters a and b. The values of these parameters can be based on historical data or expert

judgements. This allows for flexibility.
The choice of a Cobb-Douglas risk function can be motivated by the convenient properties that this function has in terms of risk-driver substitution, and the possibility of deriving isorisk curves that are easy to interpret [1, 10]. Naturally, other functions can be used, and it would be interesting to compare their actual performances.

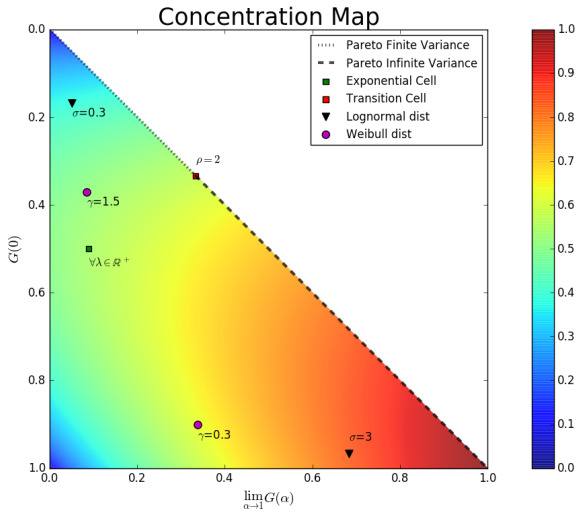

Figure 7: Concentration map with underlying Cobb-Douglas risk function of parameters $a = 0.5$, $b = 0.5$. The different points represent the theoretical concentration profiles of useful distribution in risk management: Pareto, Exponential, Weibull and lognormal. $\alpha = 0.99$ has been taken as empirical limit value for $\lim_{\alpha \to 1} G(\alpha)$

Figure 7 shows an example of the *Concentration Map* based on the Cobb-Douglas risk function. The map reads as following: warmer is the area where the points are scattered higher is the risk attached to the loss function. On the y-axis we can read the values of the first risk driver r_1, while on the x-axis the values of $\lim_{\alpha \to 1} G(\alpha)$ are given (in the picture, we choose $\alpha = 0.99$ to allow for a few observations in the tail). The second risk driver r_2 is then computed as the absolute gap, and inputed in the risk function.

The risk function,

$$R(r_1, r_2) = r_1^{0.5}(1 - r_2)^{0.5}$$

then induces an ordering on the concentration map. Clearly, different combination of values for the risk drivers can lead to the same output of the risk function. These are represented by the isorisk curves in the graph.

In Figure 8 we show how the concentration map of Figure 7 changes if, for example, different values of a and b are taken in consideration in equation 14.

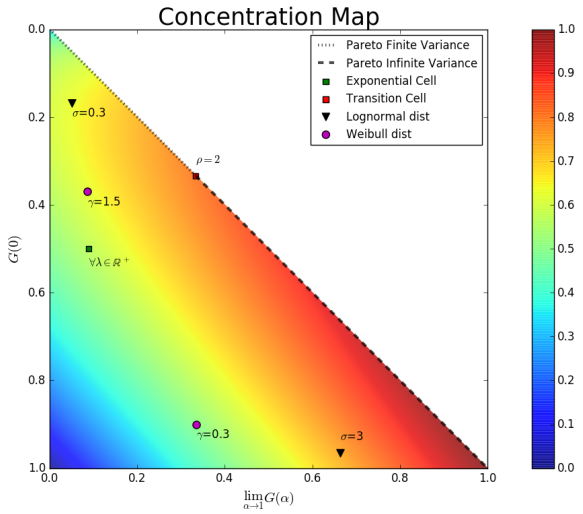

Figure 8: Concentration map, with underlying Cobb-Douglas risk function of parameters $a = 0.2$, $b = 0.8$. The different points represent the theoretical concentration profiles of useful distribution in risk management: Pareto, Exponential, Weibull and lognormal. $\alpha = 0.99$ has been taken as empirical limit value for $\lim_{\alpha \to 1} G(\alpha)$

5.3 Notable behavior on the map

For some common distributions, the concentration map provides an interesting interpretation.

The Pareto distribution with shape parameter $\rho \in (1, \infty)$ places itself on the hypothenuse of the *Concentration Map*. In particular $Pareto(\rho \searrow 1)$ corresponds to the point $(1, 1)$, while $Pareto(\rho \nearrow +\infty)$ tends to $(0, 0)$.

This is due the fact that the concentration profile associated to the Pareto distribution is constant. So that the second risk driver becomes $r_2 = 0$ (since $G(0) = \lim_{\alpha \to 1} G(\alpha)$), and the $x-$axis values for the *Concentration Map* are the same of those on the $y-$axis.

Given the fix concentration profile, $\forall \lambda \in \mathbb{R}^+$, all exponential distributions lie in the cell $(\frac{1}{2-2\log(1-\alpha)}, \frac{1}{2})$, which we can indicate as the exponential cell. The $x-$coordinate naturally depends on the chosen level of α. If for example $\alpha = 0.99$, the coordinates for the exponential become: $(0.09, 0.5)$.

Based on these two facts we can draw the conclusion that any color on the map that is equal to the one placed on the exponential cell or on the Pareto segment will share the same risk, due to the isorisk property of the Cobb-Douglas risk function [10].

For example, the points placed on the hypothenuse, from $(0.\overline{33}, 0.\overline{33})$ (what we call the transition cells in Figures 7 and 8) to $(1, 1)$, correspond to Pareto distributions with $\alpha \in (1, 2]$, whose variance is not finite. Therefore given a certain risk function, the areas of the map with the same color will share the same risk as that of a Pareto distribution with infinite variance.

6 Concentration Adjusted Expected Shortfall

As any measure of variability, the Gini index provides information about the ability of the mean to actually predict, to some degree, the outcome of a random variable: the so-called mean's precision. Recall also that, when dealing with truncated distributions in risk management, the truncated mean is the Expected Shortfall. The Gini index, when computed at a truncation level α, thus provides the precision of the ES_α computed on the risk subclass S_α.

In general, we expect the Expected Shortfall to depend on the truncation level. For example, it is known that for a Pareto distribution, the quantity $ES - u$ (also known as mean excess) increases linearly in the truncation level u, according to van der Wijk's law [7].

As we did for the truncated Gini index, we could compute the Expected Shortfall ES_α for all subclasses S_α, obtaining another "profile". However, this graph would show a trivial behavior, being ES_α an increasing function of α. And this is a pity, because the speed of growth certainly varies from distribution to distribution, and a faster growth is an indicator of a more risky loss function.

A natural solution is to weight ES_α by the corresponding G_α.

Recall that the truncated Gini index belongs to the interval $[0, 1]$, and, for $\alpha \nearrow 1$, it tends to zero for low risk light-tailed distributions. This indicates that for light tails the variability decreases quickly, while it is much more persistent for fat tails.

A natural question therefore arises: is the increase in the Expected Shortfall ruled out by an increase in its (Gini) precision? A high ES_α is not a problem per se, if it also becomes more and more precise and, in the limit for truncation level $\alpha \nearrow 1$, it becomes the certain loss.

The *Concentration Adjusted Expected Shortfall*, or $CAES_\alpha$, can be used in this framework.

Definition 10. Given a concentration profile $\{G(\alpha)\}_{\alpha\in[0,1)}$, and a sequence of expected shortfalls $\{ES_\alpha\}_{\alpha\in[0,1)}$, the concentration adjusted expected shortfall is defined as

$$CAES_\alpha = \{ES_\alpha \cdot G(\alpha)\}_{\alpha\in[0,1)}.$$

The $CAES_\alpha$ adjusts the ES_α for the variability of each subclass S_α where it is computed. Differently from the simple ES_α sequence, which is trivially increasing for every distribution, the $CAES_\alpha$ exhibits very different behavior according to different distributions. This allows us to better characterize the loss distribution.

We can distinguish three main types of behavior:

$$\lim_{\alpha\to 1} CAES_\alpha = \infty, \tag{15}$$

$$\lim_{\alpha\to 1} CAES_\alpha = c, \tag{16}$$

and

$$\lim_{\alpha\to 1} CAES_\alpha = 0, \tag{17}$$

with $c \in \mathbb{R}$.

The first case, equation (15), indicates that the increase in the Expected Shortfall is too large and

rapidly to be compensated by a decrease in the variability, or that the variability does not decrease fast enough. This behavior corresponds to the Pareto distribution with any value of $\rho \in (1, +\infty)$, the Weibull distribution with $\gamma < 1$ and lognormal distribution with $\sigma \geq 0.3$.

In the second and third cases instead, in equations (16) and (17), the decrease in the variability is sufficient to compensate for an increase of the ES_α. The value of the ES_α thus becomes more and more representative for the losses as $\alpha \nearrow 1$. The case represented by equation (16) only occurs for the exponential distribution, i.e. Weibull distribution with $\gamma = 1$; while that of equation (17) manifests itself for Weibull distributions with $\gamma > 1$ and lognormal distribution with $\sigma < 0.3$.

Finally, a remark is to be made for the specific case of lognormal distributions. Figure 10 shows how the lognormal $CAES_\alpha$ changes with σ. We can see that, for $\sigma \in (0.3, 0.7)$, it exhibits a u-shaped behavior. This feature can be used in practice to identify whether the loss function is lognormal rather than Weibull or exponential.

The combination $CP/CAES$ can thus be used to better discriminate among Weibull and lognormal distributions, also grasping the possible value of their shape parameter, by only looking at a limited number of graphs.

Figure 9 presents the general behavior of the $CAES_\alpha$ for the distributions presented in this work: Pareto, lognormal, Weibull and exponential distributions. For the lognormal only, further details are provided in Figure 10.

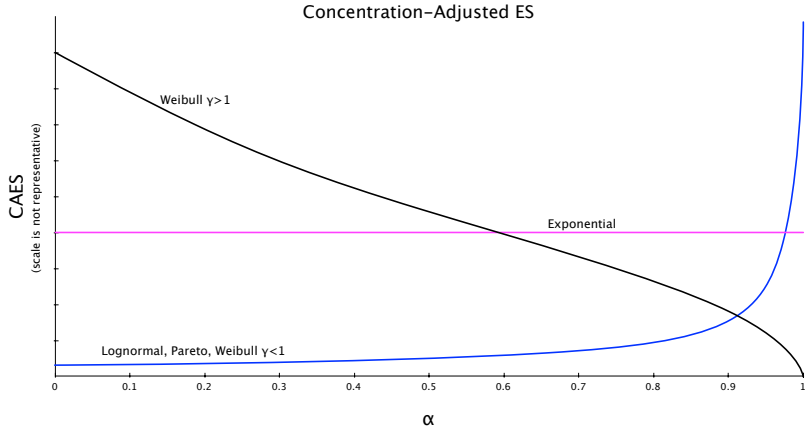

Figure 9: Theoretical Concentration Adjusted Expected Shortfall.

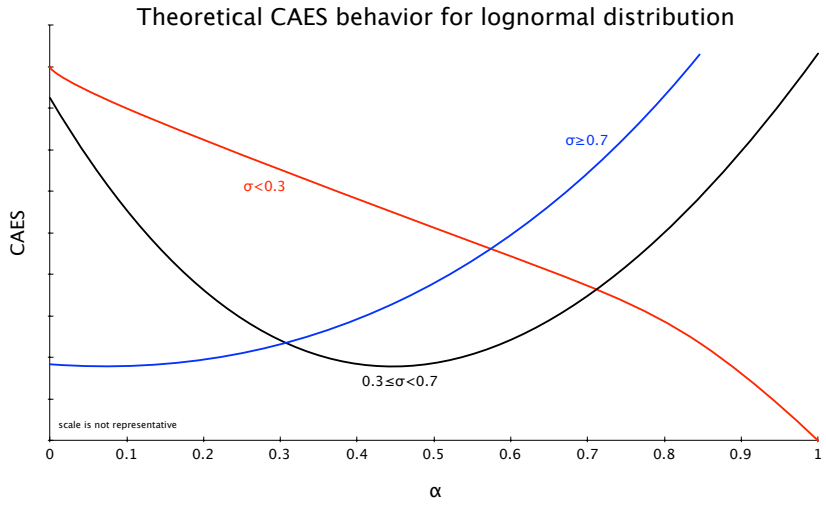

Figure 10: Concentration Adjusted Expected Shortfall for Lognormal distribution with different shape parameters.

7 Empirical Analysis

In this section we present some examples of the use of the tools introduced in this work.

In the first example we generate a data set from known distributions and we show how the concentration profile can be used in the identification of a parametric model.
In the second set of examples we use two well-known real data sets (freely available in the R package *"evir"*) as a basis for our analysis.

Finally, in Subsection 7.3, we show how the Concentration Profile finds application also in the field of Extreme Value Theory, when one is interested in defining the threshold above which fat-tails may manifest themselves.

7.1 Lognormal or Pareto distribution?

We generate a sample of size $n = 500$ from a lognormal distribution with parameters $\mu = 0$ and $\sigma = 1$ (finite mean, finite variance) and from a Pareto distribution with $\rho = 1.5$ (finite mean but infinite variance).
In Figure 11 the histograms are shown.

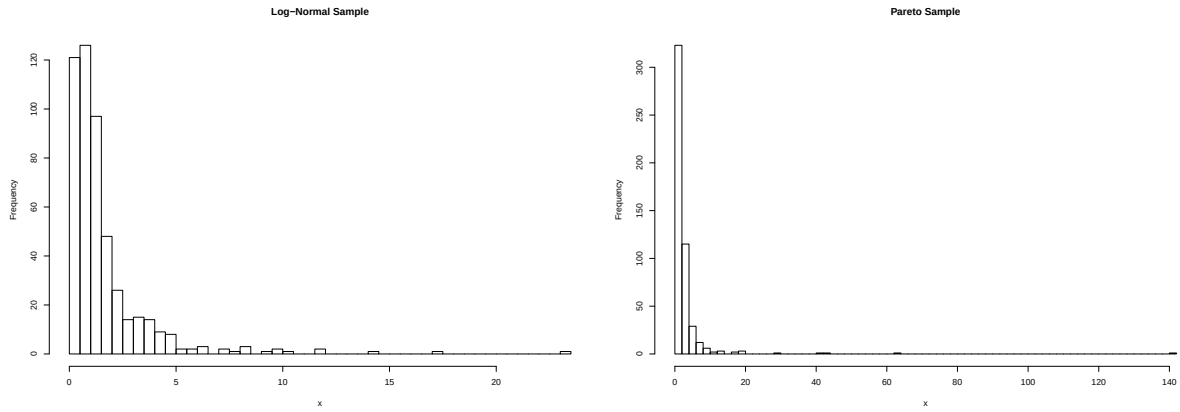

Figure 11: On the left side the histogram of a lognormal distributed data set with $\mu = 0$ and $\sigma = 1$. On the right side the histogram of a Pareto distributed dataset with $\rho = 1.5$.

Our aim is to show via the *Concentration Profile* that it is possible to fit the correct distribution. To do so we use the empirical version of the Concentration Profile presented in Section 4.

In order to estimate the CP, it is important to notice that the CP is basically a sequence of Gini indices computed on a decreasing number of ordered observations. Therefore, under normal circumstances, no additional effort must be done to estimate it with respect to the classical estimation of the Gini index, since no bias arises from the truncation procedure. This is because the left point of the domain where the truncation occurs can be successfully estimated with the empirical quantile $\hat{q}_{i:n}$.

The estimation of the Concentration Profile (CP) can be performed using any existing estimator for the Gini index $\hat{G}(i:n)$, with n being the total number of observations, and $i \in [1, n-1]$ being the number of data points to be taken into account in the estimation [32]. For example, one could use

$$\hat{G} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2n \sum_{i=1}^{n} x_i}. \tag{18}$$

After ordering the observations from the smallest to the largest, we obtain the CP by computing i times the Gini estimate, each time excluding the first i smallest observations.

Naturally, to get accurate estimates, it is important to leave a sufficient number of observations in the right tail. We define this number by k. The number of ordered observations to spare from the estimation cycle depends on the data generating process of the data set. In particular the fatter the tails, the more observations should be left outside the estimation cycle. Borrowing heuristics from extreme value theory [9, 17], we can set k between 1% and 5% of the original ordered data points.

When estimating the Gini index, attention must be paid to consistent estimates. According to [18], the estimator in equation (18) (but also several others) is consistent and asymptotically normal if and only if the second moment of the data generating process is finite. As already stressed out, this assumption can be too stringent in risk management, especially when dealing with Paretian tails.

In order to overcome this problem we propose the following solution.

Recall from Table 2 that the Gini index for the Pareto distribution is a continuous function of the shape parameter ρ. Estimators for the shape parameter of the Pareto distribution exist and do not

require any assumption on the finiteness of the second moment [17]. Therefore when we suspect an infinite variance, an estimator for the Gini index (and for the *Concentration Profile*) can be derived from the tail parameter ρ, by a plugging the estimated $\hat{\rho}$ into the theoretical expression of the Gini. The resulting estimator thus becomes:

$$\hat{G} = \frac{1}{2\hat{\rho} - 1}. \tag{19}$$

To obtain $\hat{\rho}$, one can use the maximum likelihood estimator or any other technique [17].

Confidence interval for the CP can be easily computed using resampling techniques like bootstrap, jackknife or k-jackknife[11, 12].

In Figure 12, a simple decision tree summarizes the previous considerations regarding the Gini estimation.

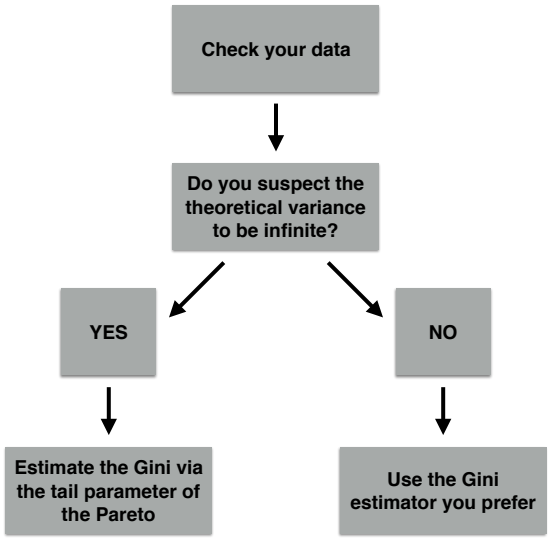

Figure 12: Decision tree for Gini index estimator

Figure 13 shows the results of experiments on simulated data. In both cases we can clearly observe that the theoretical concentration profile is contained in the 95% confidence interval obtained via bootstrapping the empirical CP. Both for the lognormal and the Pareto distributions, the parameters have been estimated via maximum likelihood estimator.

Therefore, exploiting Theorem 6, we can conclude that the samples are lognormal and Pareto, as expected.

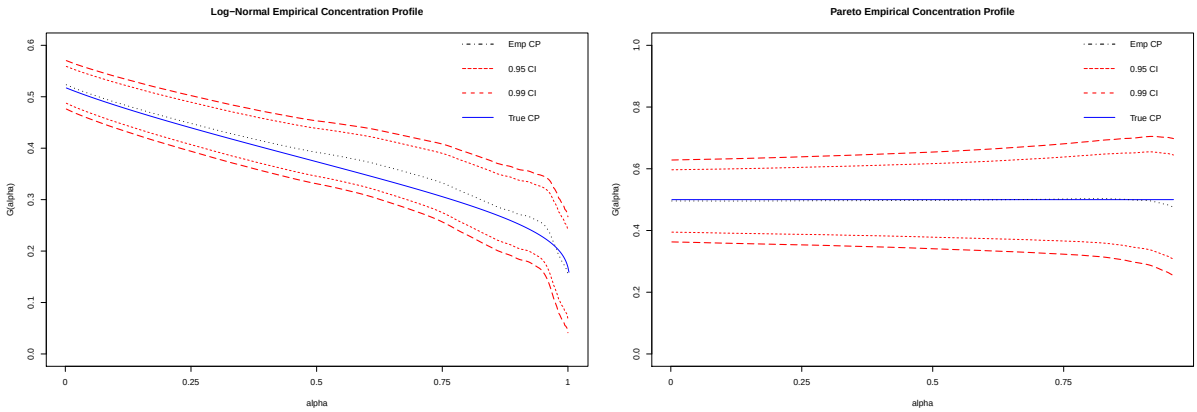

Figure 13: Empirical Concentration profile (black) with 95% bootstrapped confidence intervals (red) compared to the theoretical concentration profile (blue)

7.2 Real data study

Both data set we are going to use are freely available in the *evir* package of the R statistical language.

The first data set (name: danish) contains 2167 Danish fire insurance claims. The second one presents 2769 daily log losses on the BMW share price (name: BMW).

These very well-known data sets have been used as training data in many papers and books, e.g. [9, 17, 23]. In particular the first data set is a typical example of Pareto distribution with parameter ρ estimated to be smaller than $\rho = 2$ and bigger than $\rho = 1$, so that the distribution has finite mean but infinite variance.

The second data set is an example of lognormally distributed losses.

The empirical loss functions of the data are shown in Fig 14 as histograms.

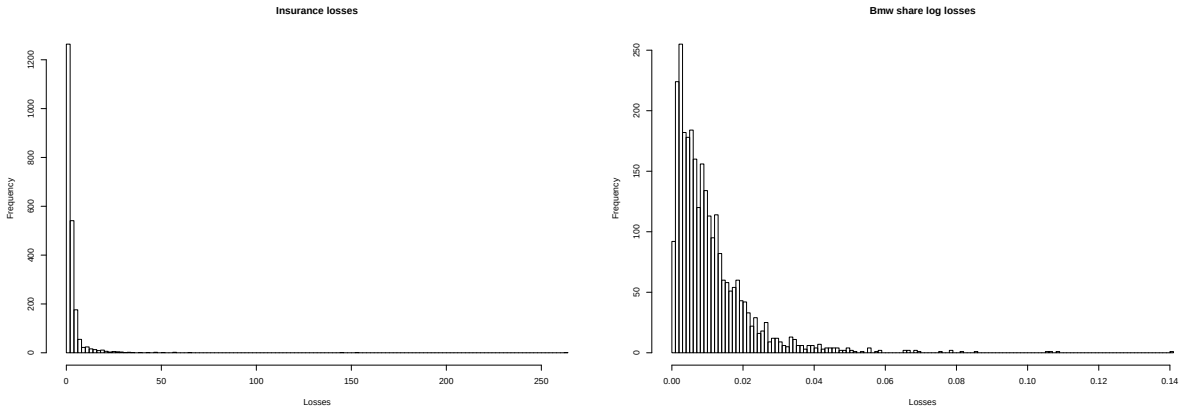

Figure 14: Empirical loss distributions shown in histogram form

Using the methodology explained above for obtaining the empirical concentration profile, we derive the *CPs* of the data sets, collecting them in Figure 15 together with their bootstrapped 95% and

99% confidence intervals.

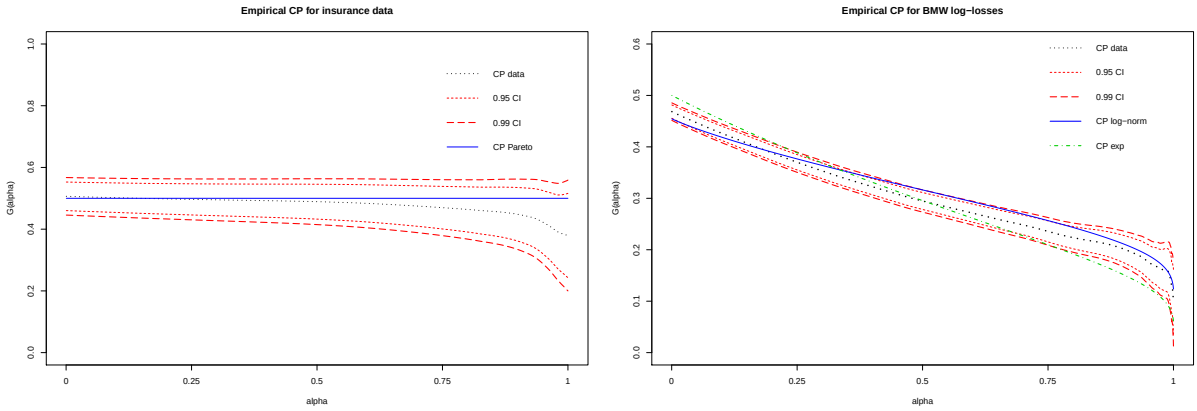

Figure 15: Empirical Concentration Profiles with attached 95% and 99% bootstrapped confidence intervals. In the left graph the empirical CP of the Danish insurance claims. In the right graph the CP for the BMW share log-losses.

In Figure 15 we notice the following patterns: the CP relative to the fire insurance claim data is almost constant. Recalling the characterization of the Pareto CP provided in Section 4, we can conclude that the data are Pareto distributed.

An additional check to the Paretianity of the data set can be done through the $CAES_\alpha$. In fact, if this is the case, we expect again a diverging $CAES_\alpha$. The expected result is presented in Figure 16.

For the BMW log-losses, we notice that the empirical CP has the following characteristics: it starts around the point 0.5, it decreases quite rapidly towards zero, and exhibits a change in the slope being slightly convex at the beginning and ending up concave.

By the results presented in Section 4, we can exclude the Pareto distribution since the CP is not constant. But what about the other three cases: Weibull, lognormal and exponential distributions?

Since the starting point of the empirical CP is around $G(0) = 0.5$, and recalling the fact that the only Weibull distribution starting at this point coincides with the exponential, we can exclude Weibull with parameters $\gamma \neq 1$ from the set of candidates. What is left to choose is either the exponential distribution or the lognormal distribution with shape parameter $\sigma \simeq 1$.

In order to discriminate between these two options we can use the $CAES_\alpha$.

Recall that the shape of the $CAES_\alpha$ for these two distributions is different. In particular for the exponential distribution it shows a constant behavior. For the lognormal distribution instead, the $CAES_\alpha$ exhibits the more complex behaviors of Figure 10.

The empirical behavior of the $CAES_\alpha$: if it is constant it means that the data are exponentially distributed, otherwise the lognormal is probably the right candidate.

The results are shown in Figure 16. We can see how the $CAES_\alpha$ diverges. This suggests the possibility of lognormally distributed data.

In order to double check this hypothesis, we fit a lognormal distribution to the data, using a method of moments estimator for its shape parameters. For σ we get $\sigma = 0.88$ (with a s.e. of 0.0065). We then use this value to generate a theoretical lognormal CP, and we compare it with the empirical

one.

The results of this fitting procedure are given in Figures 15 and 17, showing a close fit between data and the lognormal CP.

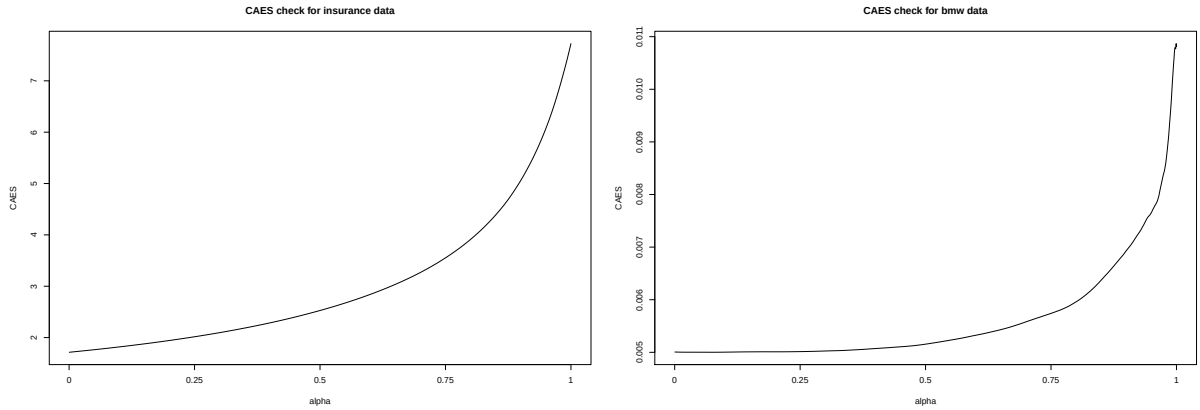

Figure 16: Concentration Adjusted ES_α for the danish insurance claim and BMW share daily log losses.

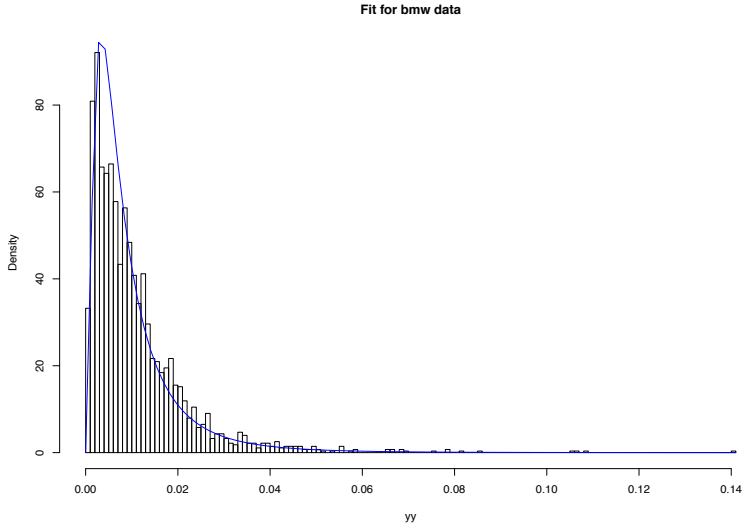

Figure 17: Lognormal fit (blue line) plotted against the histogram of the empirical data.

Clearly Pareto-like data are much riskier than light-tailed ones. However, comparing the VaR_α and the ES_α of the two data sets is useless since they measure different phenomena with different scales. Additionally, as stated before, VaR_α and ES_α alone are not very indicative for the potential losses since they do not provide information about the dispersion.

Computing the truncated Gini index $G(\alpha)$ appears to be a solution for both problems. It provides information about the precision of the ES_α and the length of the tail after VaR_α. Moreover, being a scale free measure, it can be used to compare completely different phenomena.

Table 4 compares the VaR_α, the ES_α and the $G(\alpha)$ with $\alpha \in \{0.9, 0.95, 0.99\}$ for both data sets. The first data set, the one about Danish losses, is riskier, because of the higher tail dispersion.

	Danish				BMW		
α	VaR_α	ES_α	$G(\alpha)$		VaR_α	ES_α	$G(\alpha)$
0.9	3.979	11.078	0.45		0.021	0.034	0.203
0.95	4.826	13.591	0.449		0.027	0.043	0.179
0.99	5.940	17.081	0.443		0.042	0.062	0.16

Table 4: Comparison of the main risk measures for both the datasets

In Figure 18, we show a comparison based on a concentration map. Once again, the Danish fire claims are confirmed to be riskier.

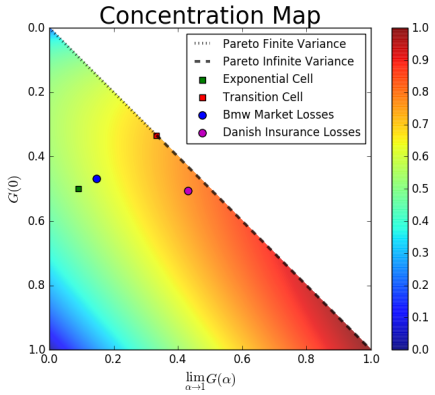

Figure 18: Risk evaluation carried via the risk concentration map. Cobb-Douglas parameters: $a = 0.3$, $b = 0.7$.

7.3 Extreme Value Theory Application

In the context of extreme value theory, one of the most important issues involves the estimation of the threshold above which the possible Paretian tail starts. This has applications in the estimators of the tail index, the tail quantile and many other relevant quantities [9, 17].

So far no unique methodology has been developed to estimate the right threshold. However, many different tools are used to provide hints about the best value, see for example [7].

Here we propose the use of the *Concentration Profile*.

We know that the concentration profile is the sequence of truncated Gini indices computed at an increasing truncation level, progressively discharging the previous observations.

Therefore, recalling the characterization of the Pareto concentration profile (Section 4), we would expect that if the data exhibit a Paretian right tail, then from a certain α-level onward the CP should be constant.

An example performed on simulated data will clarify this idea.

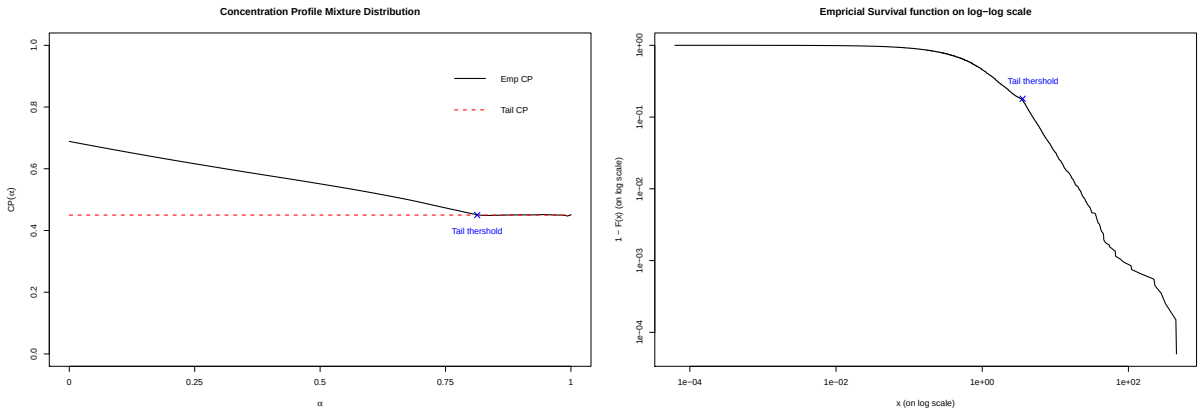

Figure 19: On the left hand side the concentration profile of the mixture distribution is shown. A dashed red line has been plotted to ease the reading to the plot and to better underline the constant behavior of the concentration profile at the tail level. On the right hand side a log-log is produced, where any Pareto-like behavior appears as a straight line with slope the tail of Pareto.

In Figure 19, at the left-hand side, the concentration profile of the following mixture is plotted:

$$Y_{mix} = \beta Y + (1 - \beta)Z, \tag{20}$$

where Y is a standard exponential distribution, Z is a Pareto distribution with $\rho = 1.5$, and the mixing parameter is $\beta = 0.85$.

It is therefore clear that by construction this dataset exhibits Paretian tails. In particular according to our methodology the Paretian behavior starts around the quantile $Q(0.825) = 3.57$ (denoted by a blue cross in Figure 19).

To asses the goodness of our findings we can compare our methodology with other techniques used in EVT to detect the tail threshold (see [7, 17] for a review). On the right-hand side of figure 19, we report the empirical survival function $S(x) = 1 - F(x)$ plotted in a log-log scale [7]. From this plot we can see how the straight line behavior which characterizes the Paretian tail is present from approximately the same level identified via the concentration profile.

8 Conclusions

In this work we have introduced a novel approach to risk management, based on the study of concentration measures of the loss distribution.

In particular we have shown that a truncated sequence of the Gini indices–the concentration profile–represents an accurate way of assessing the variability of the larger losses, the precision of common risk management measures like the Expected Shortfall, and tail risk in general.

Combining concentration profiles and standard results from utility and risk theory, we have developed a *Concentration Map*, which can be used to assess the risk attached to potential losses on the basis of the risk profile of a risk manager.

Finally, we have used the sequence of truncated Gini indices as weights for the expected shortfall defining the so-called *Concentration Adjusted Expected Shortfall*, a measure able to capture interesting additional features of tail risk.

Using simulated and empirical data, we have shown how to use our methods in practice, from risk management to extreme value theory.

It is worth stressing that concentration profiles and maps could also be obtained by substituting the Gini index with other measures of concentration, like the Pietra index [27]. However, preliminary studies made us notice that the increase in the mathematical complexity is not compensated by an improvement in the applicability. Using an Occam's razor principle, we therefore prefer the simpler (and better known) Gini index.

References

[1] E. Barucci, *Financial Markets Theory*, Springer (2003).

[2] Basel Committee on Banking Supervision - BCBS, *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version*, Bank for International Settlements (2006).

[3] Basel Committee on Banking Supervision - BCBS, *Basel III: A global regulatory framework for more resilient banks and banking systems*, Bank for International Settlements (2011).

[4] M. Bonetti, C. Gigliarano, P. Muliere,*The Gini concentration test for survival data,* Lifetime Data Analysis 15 (2009) 493-518.

[5] A. Cambini, L. Martein, *Generalized Convexity and Optimization*, Springer (2009).

[6] J-J. Cai, J.H.J. Einmahl, L. de Haan, C. Zhou, *Estimation of the marginal expected shortfall: the mean when a related variable is extreme*, Journal of the Royal Statistical Society B 77 (2015) 417-442.

[7] P. Cirillo, *Are your data really Pareto distributed?*, Physica A: Statistical Mechanics and its Applications 392 (2013) 5947-5962.

[8] P. Cirillo, N.N. Taleb, *Expected shortfall estimation for apparently infinite-mean models of operational risk*, Quantitative Finance 16 (2016) 1-10.

[9] L. de Haan, A. Ferreira, *Extreme Value Theory: An Introduction*, Springer (2006).

[10] W. Edwards , *Utility Theories: Measurements and Applications*, Springer (1992).

[11] B. Efron, R.J. Tibshirani, *An Introduction to the Bootstrap*, Chapman et Hall Monographs on Statistics and Applied Probability (1994).

[12] B. Efron, *The jackknife, the bootstrap, and other resampling plans*, Philadelphia, Pa: Society for Industrial and Applied Mathematics, (1982).

[13] I. Eliazar, M.H. Cohen, *On social inequality: Analyzing the rich-poor disparity*, Physica A 401 (2014) 148-158.

[14] I. Eliazar, I.M. Sokolov, *Measuring statistical evenness: A panoramic overview,* Physica A 391 (2012) 1323-1353.

[15] I. Eliazar, I.M. Sokolov, *Maximization of statistical heterogeneity: From Shannon's entropy to Gini's index*, Physica A 389 (2010) 3023-3038.

[16] I. Eliazar, I.M. Sokolov, *Gini characterization of extreme-value statistics*, Physica A 389 (2010) 4462-4472.

[17] P. Embrechts, C. Klüppelberg, T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer, (2003).

[18] J. L. Gastwirth, *The Estimation of the Lorenz Curve and Gini Index*, The Review of Economics and Statistics Vol. 54, No. 3 (1972).

[19] C. Gini, *Variabilità e mutabilità* (1912), Reprinted in :Variabilita e Mutabilita E. Pizetti and T. Salvemini, Memorie di Metodologica Statistica, Libreria Eredi Virgilio Veschi, Rome, 1955.

[20] J. Hull, *Risk Management and Financial Institutions*, Wiley (2012).

[21] C. Kleiber, S.Kotz, *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley (2003).

[22] M.O. Lorenz, *Methods of measuring the concentration of wealth* Publications of the American Statistical Association. Vol. 9 (New Series, No. 70) 209-219. (1905).

[23] A. McNeil, R. Frey, P. Embrechts, *Quantitative Risk Management*, second edition, Princeton University Press (2015).

[24] M. Moscadelli, *The modelling of operational risk: Experience with the analysis of the data collected by the Basel committee*, Technical Report 517, Bank of Italy, (2004).

[25] P. Muliere, M. Scarsini, *A note on stochastic dominance and inequality measures*, Journal of Economic Theory (1989),49, 314-323.

[26] N. U. Nair, P. G. Sankaran, B. V. Kumar, *Characterization of distributions by properties of truncated Gini index and mean difference*, Metron 70 (2002).

[27] G. Pietra, *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, tomo LXXIV, parte II, 1914?15, p. 775.

[28] A. Resti, A. Sironi, *Risk Management and Shareholders' Value in Banking: From Risk Measurement Models to Capital Allocation Policies*, *Wiley* (2007).

[29] Solvency II, European Union Directive 2009/138/CE (2009).

[30] E. A. Valdez, *Tail conditional variance for elliptically contoured distributions*, Belgian Actuarial Bulletin 5 (2005) 26-36.

[31] H. Shalit, S. Yitzhaki, *The mean-Gini efficient portfolio frontier,* Journal of Financial Research 28 (2005) 59-75.

[32] S. Yitzhaki, E. Schechtman, *The Gini Methodology* , Springer (2013).

Appendix

Derivation for truncated Lorenz curve

In order to derive the Lorenz curve for truncated random variables, i.e. equation (5), recall that the truncated p.d.f. of a random variable X truncated on $[a, b]$ is given by:

$$f^T(x) = \frac{f(x)}{F(b) - F(a)}, \tag{21}$$

where $f(x)$ is the original p.d.f. of X.

In our case we focus only on left-truncated random variables so that equation (21) simplifies as

$$f^T(x) = \frac{f(x)}{1 - F(a)}.$$

Let's define the Cumulative Mean Function $M(t)$ for the truncated random variable, i.e.

$$M(t) := \frac{\int_a^t \frac{xf(x)}{1-F(a)} dx}{\int_a^{+\infty} \frac{xf(x)}{1-F(a)} dx}, \tag{22}$$

with $t \in [a, +\infty)$, which becomes

$$M(t) := \frac{\int_a^t xf(x)dx}{\int_a^{+\infty} xf(x)dx}.$$

Note that the Lorenz curve is related to the Cumulative Mean Function via a change of variable $t = F^{-1}(x)$. Applying this change of variable to equation (22), we get the Lorenz curve associated to the truncated random variable X_T on $[a, +\infty]$,

$$L^*_\alpha(x) = \frac{\int_\alpha^x F^{-1}(u)du}{\int_\alpha^1 F^{-1}(u)du}, \tag{23}$$

where α is the confidence level associated to truncation in a, i.e. $F(a) = \alpha$. The curve $L_\alpha(x)$ is therefore defined on the risk subclass S_α.

Now, set $z = \frac{x-\alpha}{1-\alpha}$, and plug it in equation (23), so that

$$L_\alpha(z) = L^*_\alpha(z(1-\alpha) + \alpha). \tag{24}$$

It is straightforward to notice that equation (24) can be re-written as

$$L_\alpha(z) = L^*_\alpha(z(1-\alpha) + \alpha) = \frac{\int_\alpha^{z(1-\alpha)+\alpha} F^{-1}(u)du}{\int_\alpha^1 F^{-1}(u)du}. \tag{25}$$

Recalling that the Lorenz curve associated by a random variable with domain in $\mathbb{R}^+$ is given by equation (1), we can adjust the right-hand side of equation (25) so that

$$L_\alpha(z) = \frac{\int_\alpha^{z(1-\alpha)+\alpha} F^{-1}(u)du}{\int_\alpha^1 F^{-1}(u)du} = \frac{\int_0^{z(1-\alpha)+\alpha} F^{-1}(u)du - \int_0^\alpha F^{-1}(u)du}{\int_0^1 F^{-1}(u)du - \int_0^\alpha F^{-1}(u)du} \tag{26}$$

Now, by dividing the right-hand side by $\frac{\int_0^1 F^{-1}(u)du}{\int_0^1 F^{-1}(u)du}$, and by recalling that $L(\alpha) = \frac{\int_0^\alpha F^{-1}(u)du}{\int_0^1 F^{-1}(u)du}$ and that $L(1) = 1$, we finally obtain

$$L_\alpha(z) = \frac{L(z(1-\alpha)+\alpha) - L(\alpha)}{1 - L(\alpha)}. \tag{27}$$

Derivation for the Exponential distribution Concentration Profile

We here provide the derivation of the CP for the exponential distribution, for convenience. The derivation of the CP for others distributions only involves the computation of some more cumbersome integrals. For example for the lognormal distribution the actually-usable solution can only be obtained numerically.

Our aim is to use formula (7) to obtain a workable form for the CP.

Recalling Table 2, the Lorenz curve for the exponential distribution is given by:

$$L(x) = x + x(-\log(1-x)) + \log(1-x).$$

In order to get its truncated version, we exploit formula (5), so that

$$L_\alpha(x) = \frac{\alpha + (1-\alpha)x + (\alpha + (1-\alpha)x)(-\log(1-(\alpha+(1-\alpha)x)) + \log(1-(\alpha+(1-\alpha)x)) - (\alpha + \alpha(-\log(1-\alpha)) + \log(1-\alpha))}{1 - (\alpha + \alpha(-\log(1-\alpha)) + \log(1-\alpha))}, \tag{28}$$

with $\alpha \in [0,1]$.

After some algebra we get

$$L_\alpha(x) = \frac{-x + \log(1-\alpha) + (-1+x)\log((-1+x)(-1+\alpha))}{-1 + \log(1-\alpha)}. \tag{29}$$

We can now exploit equation (29) in formula (7), and we get

$$G(\alpha) = 1 - 2\int_0^1 \frac{-x + \log(1-\alpha) + (-1+x)\log((-1+x)(-1+\alpha))}{-1 + \log(1-\alpha)} dx. \tag{30}$$

The integral in (30) can be solved directly, thus getting the formula provided in Table 3.

Proof of Theorem 7

The quantile function of a GPD is given by

$$F^{-1}(u) = \frac{(1-u)^{-\xi} - 1}{\xi}. \tag{31}$$

From this, using equation (1), we get

$$L(x) = \frac{\int_0^x \frac{(1-u)^{-\xi}-1}{\xi} du}{\int_0^1 \frac{(1-u)^{-\xi}-1}{\xi} du}. \tag{32}$$

Solving (32), we obtain

$$L(x) = \frac{1-(1-x)^{(1-\xi)}-x+\xi x}{\xi}, \tag{33}$$

which is the closed form solution for the Lorenz curve of the GPD.

We now obtain the truncated Lorenz curve by evaluating (33) in $\alpha + (1-\alpha)x$, that is

$$L_\alpha(x) = \frac{-(1-\alpha)^\xi((\alpha-1)(x-1))^{-\xi} + x(1-\alpha)^\xi\left(((\alpha-1)(x-1))^{-\xi}+\xi-1\right)+1}{\xi(1-\alpha)^\xi-(1-\alpha)^\xi+1}. \tag{34}$$

Using equations (7) and (34), we then observe

$$G(\alpha) = 1 - 2\int_0^1 \frac{-(1-\alpha)^\xi((\alpha-1)(x-1))^{-\xi} + x(1-\alpha)^\xi\left(((\alpha-1)(x-1))^{-\xi}+\xi-1\right)+1}{\xi(1-\alpha)^\xi-(1-\alpha)^\xi+1} dx,$$

from which we derive

$$G(\alpha) = \frac{\xi}{2-(1-\alpha)^\xi(-2+\xi)(-1+\xi)-\xi}. \tag{35}$$

In order to prove that, in the case of a light-tailed distribution, the CP is always decreasing, we have to show that $\frac{dG(\alpha)}{d\alpha} < 0$ for every $\alpha \in [0,1)$. We easily observe

$$\frac{dG(\alpha)}{d\alpha} = \frac{1}{2(\alpha-1)(\log(1-\alpha)-1)^2}, \tag{36}$$

which is always negative.

What is left to show is that the CP of a fat-tailed distribution gets flatter as $\xi \nearrow 1$, with limiting case the constant.

Also in this case it is sufficient to show that: $\frac{d(\frac{dG(\alpha)}{d\alpha})}{d\xi} < 0$ and $\lim_{\xi\to 1}\frac{dG(\alpha)}{d\alpha} = 0$.

$$\frac{dG(\alpha)}{d\alpha} = -\frac{(\xi-1)\xi^2(1-\alpha)^{\xi-1}}{(\xi-2)\left((\xi-1)(1-\alpha)^\xi+1\right)^2}, \tag{37}$$

whose limit for $\xi \to 1$ goes to 0.