

Chapter 2

Risk Measures and their Properties

We have seen that a risk measure can be thought of as a map from spaces of probability distributions to real numbers.

In this chapter we will provide a formal definition of risk measure and describe all the crucial properties it should satisfy.

2.1 Definition of risk measure

Given some "reference instrument", there is a natural way to define a measure of risk by describing how close or far a position is from acceptance by the regulator (Artzner et al. [5]).

Let Ω be the set of states of nature and assume it is finite.

Let $\mathcal{X}$ be the set of all risks, i.e. the set of all real-valued functions $X \in \mathcal{X}$, which represent the final net worth of an instrument, or of a portfolio of instruments, for each element of Ω .

Definition 2.1. A risk measure $\rho(X)$ is a mapping from $\mathcal{X}$ into $\mathbb{R}$.

A measure of risk allows us to express the riskiness of a position with just one number. Obviously, the riskier a position is, the higher its measure of risk will be. When positive, the number $\rho(X)$ assigned by the measure ρ to the risk X will be

interpreted as the amount of capital an agent has to add to the risky position X to make it an acceptable position. On the contrary, if $\rho(X) < 0$, the cash amount $-\rho(X)$ can be pulled out from the already being acceptable position and invested in a more profitable way.

Thus, it seems that the concept of measure of risk is strictly related to that of acceptability.

Indeed, Artzner et al. [5] state that sets of acceptable future net worths are of primary importance and need to be considered when describing allowance or rejection of a risky position.

An acceptance set $\mathcal{A}$ is a class of final net worths accepted by a regulator.

Depending on the degree of tolerance of a specific regulator, we can have different families of acceptable positions. Hence, each acceptance set identifies a risk measure returning acceptable positions, and each risk measure identifies an acceptance set of admissible positions.

Definition 2.2. Given the total rate of return r on a reference instrument, the risk measure associated with the acceptance set $\mathcal{A}$ is the mapping from $\mathcal{X}$ to $\mathbb{R}$, denoted by $\rho_{\mathcal{A},r}$, and defined by

$$\rho_{\mathcal{A},r}(X) = \inf \{m | m \cdot r + X \in \mathcal{A}\}$$

Definition 2.3. The acceptance set associated with a risk measure ρ is the set denoted by $\mathcal{A}_\rho$ and defined by

$$\mathcal{A}_\rho = \{X \in \mathcal{X} | \rho(X) \leq 0\}$$

From def.2.2, we see that a measure of risk of an unacceptable position is interpreted as the minimum extra capital (m) we need to invest in the reference instrument, for letting the future value of the modified position to be acceptable. From def.2.3, we observe that the risk measure of an acceptable position is less than or equal to zero.

Consequently, we can affirm that the risk measure of an (initially) unacceptable position is exactly the amount of cash m through which the position $X + m$ has

a null risk measure.

In the following sections, we will present and describe the three most popular risk measures in the financial environment.

2.2 Value at Risk

Value at Risk is probably the most widely used risk measure in finance. It has become the classic measure that financial executives use to quantify market risk. Indeed, when RiskMetrics announced Value at Risk as its measure of risk in 1996, the Basel Committee on Banking Supervision enforced financial institutions to meet capital requirements based on VaR estimates.

We will now provide a formal definition.

Consider a portfolio of risky assets and a fixed time horizon Δ .

Suppose we have estimated the loss distribution associated to this portfolio and denote by $F_L(l) = P(L \leq l)$ its distribution function.¹

We would like to define a statistic based on F_L which evaluates the level of risk associated to the holding of our portfolio over the time period Δ .

An noticeable applicant is the maximum possible loss, i.e. $\inf\{l \in \mathbb{R} : F_L(l) = 1\}$. Nonetheless, it is not difficult to believe that some models entail an unbounded support for F_L , meaning that the maximum possible loss would be infinity.

Value at Risk is a smart and straightforward extension of maximum loss: the idea is simply to replace "maximum loss" by "maximum loss which is not exceeded with a given high probability", the so called confidence level (McNeil et al. [37]).

Definition 2.4 (Value-at-Risk). Given some confidence level² $\alpha \in (0, 1)$, the VaR of our portfolio at the confidence level α is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $(1 - \alpha)$.

¹Practitioners are often involved with the so called profit-and-loss (P&L) distribution. However in risk management we are mainly concerned with the probability of large losses and hence we often drop the P from P&L. Moreover, since actuarial risk theory is a theory of positive random variables, we will use the convention that losses are positive numbers and we will focus on the upper (right) tail of the loss distribution L.

² α denotes the confidence level, not to be confused with the significance level, which is equal to $1 - \text{confidence level}$. Hence, here α could assume values like 0.95 or 0.99, not 0.05 nor 0.01.

Formally,

$$VaR_\alpha(L) = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) > \alpha\} \quad (2.1)$$

In probabilistic terms, VaR is thus simply a quantile of the loss distribution.

We recall that, given some df L , the generalized inverse $F^{\leftarrow}$ is called the *quantile function* of L .

$$q_\alpha(L) := F^{\leftarrow}(L) = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}$$

Thusly, we get:

$$VaR_\alpha(L) = q_\alpha(L)$$

Typical values for α are $\alpha=0.95$, $\alpha=0.99$ or higher.

In market risk management the time horizon Δ is usually 1 or 10 days, while in credit risk management Δ is usually one year.

VaR has the advantage of being very intuitive and the great popularity this instrument has reached, is essentially due to its conceptual clarity.

However, VaR has two important drawbacks:

- i) it does not fulfill (in general) the property of sub-additivity, meaning that it does not award diversification benefits;
- ii) it is tail insensitive. It tells us that in the $\alpha \cdot 100\%$ of the cases the loss will not be greater than a certain level, but it does not give us any clue about the size of the loss in the remaining $(1 - \alpha) \cdot 100\%$ of the cases.

In Section 2.5 we will deepen the discussion about the properties a good risk measure should have.

2.3 Expected Shortfall

Risk professionals have been looking for a coherent alternative to Value at Risk for many years.

Expected Shortfall turns out to be a natural choice to appeal to, when VaR is unable to distinguish between portfolios embodying different levels of risk.

In fact, little imagination is necessary to construct portfolios with identical VaR and dangerously divergent levels of tail risk, namely the risk in the $(1 - \alpha)\%$ worst cases.

It seems that instead of asking what could be the minimum loss incurred in a presupposed percentage of worst cases, we should be concerned with the expected loss sustained in that portion of unfortunate possibilities.

It is effortless to figure out that, if the loss distribution function is continuous, then the statistical quantity which answers the questioning above, is simply given by the conditional expected value above the quantile, that is the *Tail Conditional Expectation*:

$$TCE_\alpha(L) = E\{L | L \geq VaR_\alpha\} \quad (2.2)$$

For more general distribution however, this statistic does not fit really well our purposes, since the event $\{L \geq VaR_\alpha\}$ may happen to have a probability larger than our set of selected worst cases.

Indeed TCE may violate the sub-additivity property on a general distribution.

In order to account for more general distribution we need to expand the definition of TCE. This leads to the following formulation.

Definition 2.5 (Expected Shortfall). For a loss L with $E(|L|) < \infty$ and distribution function F_L , the Expected Shortfall at confidence level $\alpha \in (0, 1)$ is defined as

$$ES_\alpha(L) = \frac{1}{1 - \alpha} \cdot (E(L; L \geq q_\alpha) + q_\alpha \cdot (1 - \alpha - P(L \geq q_\alpha))) \quad (2.3)$$

where q_α is the α -quantile of F_L .

Equation 2.3 may look complicated. Nevertheless, the concept it expresses is that cited above. Hence equation 2.3 is simply the mathematical translation of what we are looking for.

To have a more transparent view, we should clarify some notions: the term $q_\alpha \cdot (1 - \alpha - P(L \geq q_\alpha))$ has to be interpreted as the exceeding part which need to be added to the expected value $E(L; L \geq q_\alpha)$ when the event $\{L \geq q_\alpha\}$ has probability larger than $(1 - \alpha)$.

When, on the contrary, $P(L \geq q_\alpha) = 1 - \alpha$, as is always the case if the probability distribution is continuous, the term vanishes and equation 2.3 reduces to equation 2.2 .

Said differently, for an integrable loss L with continuous distribution function F_L and any $\alpha \in (0, 1)$, we have that

$$ES_\alpha(L) = \frac{E(L; L \geq q_\alpha(L))}{1 - \alpha} = E(L|L \geq VaR_\alpha) \quad (2.4)$$

and hence

$$ES_\alpha(L) = TCE_\alpha(L) \quad (2.5)$$

The comprehensibility of the ES_α can be perceived making use of an equivalent definition, which handles the ES_α as a combination of expected values.

There exists an analogous representation to equation 2.3 which reveals the close link with the parameter α and the distribution function F_L .

Literally, employing the left generalized inverse function of F_L ,

$$F_L^{\leftarrow}(u) = \inf\{u \in \mathbb{R} : F_L(u) \geq \alpha\}$$

one can handily illustrate that ES_α can be expressed as the mean of $F^{\leftarrow}(L)$ on the significance level interval:

$$ES_\alpha(L) = \frac{1}{1 - \alpha} \cdot \int_\alpha^1 q_u(F_L) du. \quad (2.6)$$

Expected Shortfall is thus related to VaR by

$$ES_\alpha(L) = \frac{1}{1-\alpha} \cdot \int_\alpha^1 VaR_u(L) du. \quad (2.7)$$

This is the principal and most used formulation of ES_α .

Its mathematical manageability makes it especially suitable for studying the analytical properties of ES_α . For instance, one of the property that differentiate the ES_α from its competitors, specifically the continuity in α , is manifest in equation 2.7, while it is not in equation 2.3.

To sum up, Expected Shortfall comes out to be a coherent risk measure; however it has some weaknesses as well.

In particular we will see that ES is not an *elicitable* risk measure.

2.4 Expectiles

Given that Value at Risk is not coherent and Expected Shortfall is not elicitable, many authors have been searching recently, for a risk measure sharing both the properties of coherence and elicibility.

Possible candidates are Expectiles. In this section we are going to see what Expectiles are and why they are so interesting as a potential risk measure.

We will define Expectiles starting from the work of Newey and Powell [39].

We know that a quantile can be defined as the minimizer of an asymmetric loss function:

$$q_\alpha(L) := \arg \min_{l \in \mathbb{R}} E[\alpha \cdot (L - l)^- + (1 - \alpha) \cdot (L - l)^+]$$

Similarly Newey and Powell [39] have defined the Expectile as the minimizer of an asymmetrically weighted squared loss function.

Definition 2.6 (Expectiles). For $0 < \beta < 1$ and square-integrable L , the β - Expectile $e_\beta(L)$ is defined as

$$e_\beta(L) = \arg \min_{l \in \mathbb{R}} E[(1 - \beta) \max(L - l, 0)^2 + \beta \max(l - L, 0)^2] \quad (2.8)$$

Note that, as for the variance, the notion of Expectile requires finite second moment.

Expectiles can also be defined in terms of their first order condition (f.o.c henceforth). Indeed the β -Expectile, $e_\beta(L)$, is the unique solution l of the following equation:

$$(1 - \beta) E[\max(L - l, 0)] = \beta E[\max(l - L, 0)] \quad (2.9)$$

Consequently, $e_\beta(L)$ satisfies

$$e_\beta(L) = \frac{\beta E[L \mathbf{1}_{\{L < e_\beta(L)\}}] + (1 - \beta) E[L \mathbf{1}_{\{L \geq e_\beta(L)\}}]}{\beta P[L < e_\beta(L)] + (1 - \beta) P[L \geq e_\beta(L)]}$$

From all the above definitions we can understand that the Expectiles are specified by the weighted conditional expectations of the random variable L .

Strictly speaking Expectiles are a measure of the right tail and left tail expected values of the loss L . Thus, while VaR does not consider any of the two tails of the loss distribution and ES only examines one of them, Expectiles assess both the tails with different weights.

By way of explanation, they balance the left and right tail of the distribution, so that the ratio between the expected positive and negative deviations from e_β will be equal to a predetermined constant:

$$\frac{E[(L - e_\beta(L))^+]}{E[(L - e_\beta(L))^-]} = \frac{\beta}{(1 - \beta)}$$

Supposing that $\beta = 1 - \beta = \frac{1}{2}$, the ratio will be equal to 1, implying that the means of the left and the right deviations from e_β are equal.

We should mention that in Rossi [42] the risk measure associated to expectiles is

indicated as EVaR.

Specifically, for a loss L with $E(L^2) < \infty$ and distribution function F_L , we have

$$EVaR_\beta(L) = e_\beta(L)$$

Before going on and describing the very important properties of a risk metric, we would like to make a parallelism with Section (2.1) and define a risk measure using the notion of acceptance set (see def. 2.3).

This is done, mainly, to let the reader better understand the concept of EVaR. Hence, we can define our three risk measures through the concept of acceptance set.

For a generic random variable X , the VaR acceptance set is:

$$\mathcal{A}_{VaR_\alpha} = \{X \in \mathcal{X} : P(X < 0) \leq 1 - \alpha\}$$

For a generic random variable X , the ES acceptance set is:

$$\mathcal{A}_{ES_\alpha} = \left\{ X \in \mathcal{X} : \frac{1}{1 - \alpha} \cdot \int_0^{1-\alpha} q_u(F_X) du \geq 0 \right\}$$

For a generic random variable X , the EVaR acceptance set is:

$$\mathcal{A}_{EVaR_\beta} = \left\{ X \in \mathcal{X} : \frac{E[X^+]}{E[X^-]} \geq \frac{1 - \beta}{\beta} \right\}$$

Thus, when we refer to VaR, a position is included in the acceptance set if the probability of a loss does not exceed the fixed level $(1 - \alpha)$.

In the case we are dealing with ES, a position is said acceptable if the average loss in the worst $(1 - \alpha) \cdot 100\%$ cases is not greater than zero and, for this reason, we should say that ES is more conservative than VaR.

Finally, looking at EVaR, a position can be considered acceptable if its gain-loss ratio is greater than the prearranged value $\frac{1-\beta}{\beta}$.

2.5 Coherent risk measures

We are now going to outline the subset of coherent risk measures.

The notion of coherence was introduced by Artzner et al. [5] and currently, it is a fundamental concept related to the acceptability of a risk measure.

In reality, Acerbi and Tasche [3] state that if a measure turns out to be not coherent, then it simply cannot be named as *risk measure*:

"To avoid confusion, if a measure is not coherent we just choose not to call it a risk measure at all."

So, let us describe what *coherence* means.

Definition 2.7. A risk measure is *coherent* if it satisfies the following four axioms:

Axiom 1. Translation Invariance

For all $X \in \mathcal{X}$ and for all $m \in \mathbb{R}$, we have

$$\rho(X + m) = \rho(X) - m \quad (2.10)$$

Axiom 2. Sub-additivity

For all $X_1 \in \mathcal{X}$ and $X_2 \in \mathcal{X}$, we have

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2) \quad (2.11)$$

Axiom 3. Positive Homogeneity

For all $X \in \mathcal{X}$ and for all $\tau > 0$, we have

$$\rho(\tau \cdot X) = \tau \cdot \rho(X) \quad (2.12)$$

Axiom 4. Monotonicity

For all $X_1 \in \mathcal{X}$ and $X_2 \in \mathcal{X}$ with $X_1 \leq X_2$, we have

$$\rho(X_1) \geq \rho(X_2) \quad (2.13)$$

The axiom of *Translation Invariance* illustrates that adding (or deducting) a risk-free amount to (from) a portfolio and investing it in the reference instrument, results in a decrease (increase) of the risk of the position by exactly the same amount.

From this property we can derive a well-known fact in finance:

if $m = \rho(X)$, then

$$\rho(X + m) = \rho(X + \rho(X)) = \rho(X) - \rho(X) = 0$$

Hence, it is possible to hedge an underwritten risky position, by simply adding a certain amount of risk-free instruments in the portfolio.

The axiom of *Sub-additivity* reminds us of the diversification theory.

In fact, thanks to diversification benefits, a portfolio composed by several assets will be strictly less risky than a portfolio made up by a single instrument, provided that the correlation among the assets is different from 1.

In this sense, we can say that the sub-additivity property sets an upper bound to the risk of a portfolio and thus to the amount of regulatory capital we need to allocate. Only when there is a well-founded possibility that the sources of these risks may act altogether, the global risk of a portfolio will be the sum of the risks of its components.

According to Artzner et al. [5], the *Sub-additivity* axiom captures the essence of how a risk measure should behave, especially in the aggregation (disaggregation) of portfolios, and for this reason it is probably the key feature of a risk metric.

The property of *Positive Homogeneity* underlines that, if (for instance) the exposure to a specific position doubles, then the risk measure related to that position doubles as well. However, in the case that the position size directly influences risk

(think for example of liquidity risk), we should account for any possible repercussion (e.g. difficulty in liquidate the position), and we might expect the risk will increase more than twice.

Hence, we could have $\rho(\tau \cdot X) \geq \tau \cdot \rho(X)$.

Nonetheless, from the *Sub-additivity* axiom we know that $\rho(\tau \cdot X) \leq \tau \cdot \rho(X)$.

Indeed:

$$\rho(\tau \cdot X) = \underbrace{\rho(X + X + \cdots + X)}_{\tau\text{-times}} \leq \underbrace{\rho(X) + \rho(X) + \cdots + \rho(X)}_{\tau\text{-times}} = \tau \cdot \rho(X)$$

Accordingly, the only possible solution is the equality as expressed in Axiom 3.

Finally, the *Monotonicity* axiom explains that if, in each state of the world, the position X_2 performs always better than position X_1 , then the risk associated to X_1 should be higher than that related to X_2 .

2.5.1 Coherence of VaR

As previously mentioned, the sub-additivity property fails to be valid for Value at Risk (*in general*), meaning that it is not a coherent risk measure.

Thus we have:

$$VaR_\alpha(L_1 + \cdots + L_d) > VaR_\alpha(L_1) + \cdots + VaR_\alpha(L_d) \quad (2.14)$$

We have seen that for a sub-additive measure, portfolio diversification always lead to risk reduction, while for metrics like VaR this fact may not hold anymore. Whenever we try to get the whole picture, we immediately realize that sub-additivity is a necessary property for capital adequacy requirements.

Consider a financial institution made of several businesses: if the regulatory capital is computed using a 'bottom-up' approach (and using a non sub-additive risk measure), how could we be sure that the final number we get, consistently estimates the whole risk of the enterprise?

Hence, a decentralization of risk management using VaR is troublesome, since we

cannot be certain that, by aggregating VaR numbers for different branches, we will end up with a bound for the overall risk of the company.

When looking at aggregated risk $\sum_{i=1}^n L_i$, the equation 2.14 becomes:

$$VaR_{\alpha}(\sum_{i=1}^n L_i) > \sum_{i=1}^n VaR_{\alpha}(L_i)$$

Previously, we have stated that VaR is not sub-additive *in general*.

Indeed, whether or not it is the case depends on the properties of the joint loss distribution and the standard situations in which our measure comes out to be sub-additive, are the following:

- i) The random variables are independent and identically distributed, as well as positively regularly varying.
- ii) The random variables have an elliptical distribution.
- iii) The random variables have an Archimedean survival dependence structure.

What the first context directly implies is that, when the loss distribution is normal, VaR does not have any problems.

Of course, it will not be a big issue to assess the tail risk since, with normality assumption, VaR is a scalar multiple of standard deviation and consequently also a scalar multiple of ES (which we know is good in providing information about tail losses).

Moreover, being a multiple of standard deviation, "Normal" VaR satisfies the sub-additivity axiom.

Howsoever, it should be noted that the real world has nothing to do with the Gaussian one, and also that in Gaussian world everything is proportional to standard deviation, so everything is sub-additive, and there is nothing special with VaR.

As regard the other two more complicated situations, the interested reader should refer to Danielson et al. [17], Embrechts et al. [19] and Embrechts et al. [20].

Furthermore Embrechts et al. [21] has introduced a numerical algorithm which

provides boundaries for the VaR of high-dimensional (inhomogeneous) portfolios, and Kratz [33] has studied the topic of the evaluation of VaR of aggregated heavy tailed risks.

2.5.2 Coherence of ES

Expected Shortfall fulfills all the four axioms above and so it is a coherent risk measure.

Thence, many authors believe that ES is an excellent substitute for VaR in risk management applications. Undoubtedly, some of the warnings that apply for VaR, such as the daintiness of the estimation procedure, the reliability of approximations and the consistency of the assumptions, should also be considered in this case. The only difference is that, if correctly estimated, the ES will give less misleading answers.

From the practical point of view, the goodness of Expected Shortfall is contingent on the stability of the estimation approach and on the finding of effective back-testing methodologies.

We will address this topic later on, placing particular emphasis on the robustness of risk measurement procedures and on the relationship between backtesting and a property called *elicitability*.

2.5.3 Coherence of Expectiles

Expectiles can be a coherent risk measure under some restrictions. In fact, they satisfy all the axioms, except the one of sub-additivity.

Actually, expectiles are sub-additive if $\beta \geq \frac{1}{2}$, although super-additive if $\beta \leq \frac{1}{2}$:

Sub-additivity. For two random variables $X_1, X_2 \in \mathcal{X}$, and for $\beta \geq \frac{1}{2}$, we have

$$e_\beta(X_1 + X_2) \leq e_\beta(X_1) + e_\beta(X_2)$$

Super-additivity. For two random variables $X_1, X_2 \in \mathcal{X}$, and for $\beta \leq \frac{1}{2}$, we have

$$e_\beta(X_1 + X_2) \geq e_\beta(X_1) + e_\beta(X_2)$$

Hence, in the moment we look at expectiles for risk measurement purposes, we should keep in mind that it is a coherent risk measure only in the case the confidence level is above 0.5.

What is worth mention is that expectiles, not only fulfill the other three axioms, but also comply with a stricter property:

Strong Monotonicity. For two random variables $X_1, X_2 \in \mathcal{X}$, and for $\beta \in (0, 1)$, if $X_1 \leq X_2$ and $P(X_1 < X_2) > 0$, we have

$$e_\beta(X_1) < e_\beta(X_2)$$

This is a more rigorous axiom than the monotonicity one.

Indeed strong monotonicity guarantees that, given $X_1 \leq X_2$, if the probability of the event " X_1 is strictly less than X_2 " is non-null, then the above inequality involving $e_\beta(X_1)$ and $e_\beta(X_2)$, always holds.

On the contrary, simple monotonicity could end up with $e_\beta(X_1) = e_\beta(X_2)$, even when there is a strictly positive probability for the event " $X_1 < X_2$."

2.6 Risk Measures: a deeper view

In this brief section we will provide some other significant properties a risk measure should comply with. In truth, we believe they reserve some care.

2.6.1 Convexity

The concept of convex measure of risk is an extension of that of coherent risk measure discussed formerly.

Following the reasoning of Föllmer and Schied [24], we start with a comprehensible characterization of a risk metric:

A coherent measure of risk ρ originates from some family $\mathcal{Q}$ of probability measures, by gauging the expected loss under $Q \in \mathcal{Q}$ and then considering the worst result as Q varies over $\mathcal{Q}$:

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E_Q[-X]$$

In real world, however, there exist a large number of situations in which, the risk borne in an investment, evolves in a nonlinear way with the size of a position. Recall the example (about liquidity risk) we made for explaining the property of positive homogeneity.

In this context, we should relax the assumption of sub-additivity and positive homogeneity and require the weaker property of *Convexity*:

$$\rho(\lambda X_1 + (1 - \lambda)X_2) \leq \lambda \rho(X_1) + (1 - \lambda)\rho(X_2)$$

for any $\lambda \in (0, 1)$ and any (risky positions) $X_1, X_2 \in \mathcal{X}$.

Also here, we can notice the direct link with the notion of diversification: the risk of a diversified portfolio, which in this case is $\lambda X_1 + (1 - \lambda)X_2$, is less or equal to the weighted average of individual risks.

Assume now that $0 \in \mathcal{X}$ and that $\mathcal{X}$ is closed under the addition of constants.

Definition 2.8 (Convex Measure of Risk). A map $\rho : \mathcal{X} \rightarrow \mathbb{R}$ will be called a *convex measure of risk* if it satisfies the condition of *convexity*, *monotonicity* and *translation invariance*.

When we deal with *normalized convex* measure of risk, i.e. $\rho(0) = 0$, the quantity $\rho(X)$ can be understood as a "margin requirement".

Remind that a margin is the minimum amount of capital we have to add to a risky position and invest in a risk-free asset, in order to make that position "acceptable". We conclude this subsection by enunciating the fundamental representation theorem for convex measures of risk.

Theorem 2.9. Suppose $\mathcal{X}$ is the space of all real-valued functions on a finite set Ω . Then, $\rho : \mathcal{X} \rightarrow \mathbb{R}$ is a convex measure of risk if and only if there exists a "penalty function" $\alpha : \mathcal{P} \rightarrow (-\infty, \infty]$ such that

$$\rho(X) = \sup_{Q \in \mathcal{P}} (E_Q[-X] - \alpha(Q))$$

The function α satisfies $\alpha(Q) \geq -\rho(0)$ for any $Q \in \mathcal{P}$, and it can be taken to be convex and lower semicontinuous on $\mathcal{P}$.

Observe that the structure theorem of coherent risk measure is a special case of the above. Accordingly, it is not surprising that any positive homogeneous and subadditive risk measure is also convex. Expected shortfall and expectiles are both convex risk measures. Obviously, Value at risk is not.

2.6.2 Comonotonic Additivity

Earlier, we have given considerable importance to the sub-additivity axiom.

Now we would like to present a similar result, which can be very useful for financial purposes. First of all we should define the concept of *comonotonicity*.

Definition 2.10. Two random variables X_1 and X_2 are said to be comonotonic if, given a third random variable Y , there exist two monotonic increasing function f_1 and f_2 , such that

$$X_1 = f_1(Y) \text{ and } X_2 = f_2(Y)$$

This simply means that, for instance, two risky positions X_1, X_2 are perfectly and also positively dependent on the same source of risk Y , i.e.

$$X_1 \uparrow \iff X_2 \uparrow$$

Obviously, the perfect positive dependence implies the maximum degree of correlation.

Given these clarifications, we can now turn to the concerned definition.

Definition 2.11 (Comonotonic Additivity). A risk measure $\rho(\cdot)$ is *comonotonically additive*, if for any comonotonic random variables X_1, X_2 , it holds that

$$\rho(X_1 + X_2) = \rho(X_1) + \rho(X_2)$$

The reason why this property extremely matters, is intuitive and again related to the notion of diversification.

As a matter of fact, if two different risky positions perfectly depend on the same risk factor, they should not benefit from diversification effects.

Thus, in risk management, we should always use comonotonic additive risk metrics. VaR and ES both comply with this property, while Expectiles do not.

2.6.3 Law Invariance

A risk measure is defined as *law invariant* if it depends entirely on the distribution of the random variable associated to it. More precisely:

Definition 2.12 (Law Invariance). Consider two random variables X_1, X_2 and their corresponding distribution functions F_{X_1}, F_{X_2} . A risk measure $\rho(\cdot)$ is a *law invariant* risk measure if

$$F_{X_1} = F_{X_2} \implies \rho(X_1) = \rho(X_2)$$

The importance of this property lies on the fact that, for assessing the risk level of a position, we apply the risk measure on the loss distribution, which is estimated from empirical data.

Hence, in the interest of making this approach accurate we need that, every time the variables follow identical distributions, the measure returns the same level of risk.

A direct consequence of this fact is that, whenever we deal with risk measures that are not law invariant, we could not evaluate the riskiness of a position through the loss distribution.

As underlined in Acerbi [1], the majority of risk measures used in finance (including VaR, ES and Expectiles) comply with this property.

2.6.4 Robustness

Another important issue when coping with risk measures is robustness.

A risk measure is said to be *robust* if it is quite insensitive to measurement errors. Cont et al. [16] observe that measuring the risk of a financial portfolio involves two steps: estimating the loss distribution from empirical data and computing the risk metric ρ which, as usual, is a map assigning a number to each random payoff. They also point out that, even if these two operations have always been studied separately, they are strictly related (at least in applications) and their connection is of great importance when opting for a risk measure, rather than another.

What could be really problematic is the sensitivity of risk measures to mis-specification errors in the portfolio loss distribution. We care so much about the portfolio loss distribution for the simple reason that VaR, ES and also Expectiles are directly estimated from it. In light of this, it could be useful to evaluate the risk estimators' sensitivity, for instance computing their relative change when a new observation in the data set is included.

The figure below, taken from Cont et al. [16], exhibits this measure of sensitivity of VaR and ES as a function of the size of a point added.

We can understand that the ES is much more sensitive than VaR to a change in the data set, especially when large observations are added.

Thus, ignoring robustness may lead to meaningless results, since small measurement errors in the estimation procedure can have a huge impact on our outcomes. Moreover, looking at figure 2 (Cont et al. [16]), we can observe that different estimation procedures for the same risk measure and the same portfolio display divergent sensitivities to a new outlier.

Hence, depending on whether the loss distribution is estimated directly from historical data or approximated through a parametric model, the sensitivity of risk measures varies.

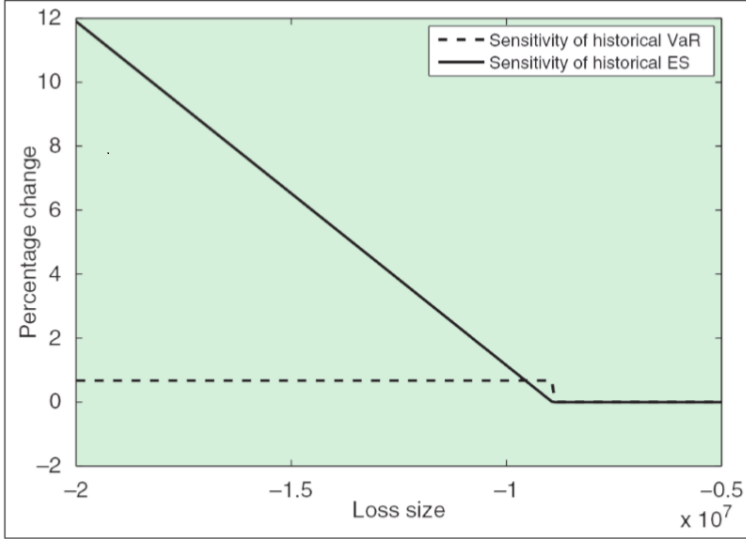


FIGURE 2.1: Empirical sensitivity (in percentage) of the historical VaR and historical ES, as in Cont et al. [16]

Concluding, Cont et al. [16] show us two important results:

- i) While Expected Shortfall has the advantage of being a coherent risk measure, it seems to lack robustness. On the contrary, VaR appears to be quite insensible to data modifications.
- ii) The choice of the estimation method has a considerable impact on the sensitivity of the risk measure.

These findings motivate the claim according to which there exists a conflict between coherence (or better sub-additivity) of a risk measure and the robustness of its statistical estimators.

Without going any further on the subject, we shall mention that robustness is usually investigated in terms of continuity (either with respect to the weak topology or considering the Wasserstein distance).

For the sake of conformity, we should report that Expectiles are not robust with respect to the weak topology but, as proved in Bellini et al. [9], they are robust

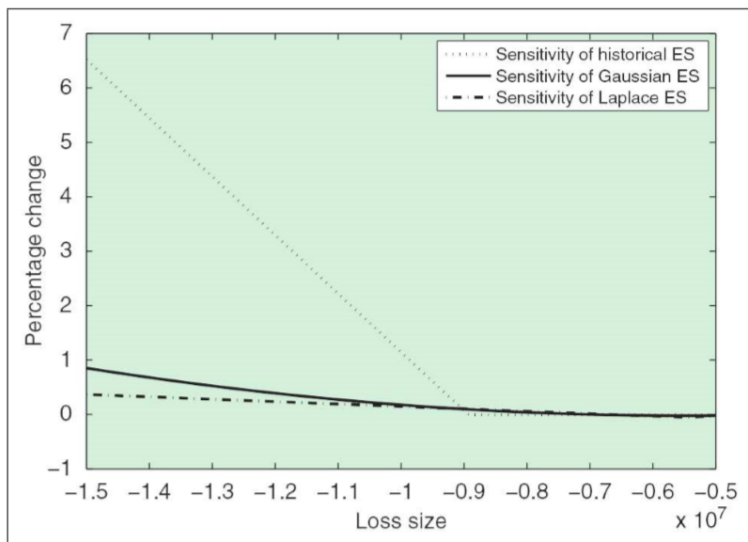


FIGURE 2.2: Empirical sensitivity (in percentage) of the $ES_{0.01}$ estimated with different methods, as in Cont et al. [16]

with respect to the Wasserstein distance.

The interested reader should refer to Cont et al. [16] and Stahl et al. [46] for further insights.

Backtesting Value at Risk and Expected Shortfall

Roccioletti, S.

2016, XIX, 145 p. 45 illus., Softcover

ISBN: 978-3-658-11907-2