# Mathematics of Risk Measures

*And the measures of the Basel Committee*

Master's Degree Thesis

MARTINA F. VAN RIJN

University of Utrecht

**Supervisors:**

Prof. Dr. Roberto Fernández

Dr. Karma Dajani

Utrecht, The Netherlands - Januari 2015

Student number: 0410985

# Mathematics of Risk Measures

*And the measures of the Basel Committee*

by

Martina F. van Rijn

In partial fulfilment for the degree of
**Master of Science**
at the department of Mathematics, Utrecht University.

Supervisors: **Prof. Dr. Roberto Fernández**
*Utrecht University*

**Dr. Karma Dajani**
*Utrecht University*

# Abstract

Risk measurement continues to be of the utmost importance in practice. The axiomatic approach to risk measures of Artzner et all. [3] gave rise to a whole new theory round this topic. With their axiomatic approach the class of coherent risk measures was completely characterized. In later years it became evident that coherent risk measures as defined by these axioms do not posses all of the properties desired in practice, since they are not necessarily robust. Robustness of risk measures is important to properly determine the underlying loss distribution through methods like backtesting. In 2013 Gneiting [19] published an article in which he described elicitable risk measures, this class of risk measure is robust for (small) changes in the data. He showed that a necessary condition for elicitability is that of convex level sets, the question whether it is also sufficient still remains open.

In my thesis I shall compare the properties of these two different classes of risk measures, and argue that a risk measure should be both coherent and elicitable. Whereas there is a whole class of risk measures that are coherent or that are elicitable, there is only one such statistical functional that fits both criteria, the expectiles.

Furthermore I will compare special cases of these classes by means of a simple, random foreign exchange portfolio. Which will show not only the mathematical differences but also the practical implications of choosing a risk measure.

Since the practical implications where the key to further research on risk measures, I take a close look at the risk measures introduced by the Basel Committee on Banking Supervision and give a comparison to coherent, elicitable risk measures by means of the axioms they are based upon.

# Acknowledgements

This thesis was written as the final part of my master Mathematical Sciences at Utrecht University. I had the opportunity to start writing my thesis at the University of Malta as a visiting student.

First of all I would like to thank my supervisors, prof. dr. Roberto Fernández and dr. Karma Dajani, for their helpful comments and good advice. Without their constant support and guidance I would not have materialized this thesis.

I would also like to thank dr. Paul David Suda of the University of Malta for introducing me to the topic of risk measures. And for his help during the first steps of this project.

# Preface

Baring's Bank was the most prestigious bank of the United Kingdom, it collapsed in 1995 due to poor speculative investments in the Asian market done by one of its employers, Nick Leeson. Leeson falsified his losses, and reported to the bank and the British Tax Authorities that he was making substantial profits. He could no longer hide the losses when the Kobe earthquake sent the Asian financial markets into a tailspin. By the time the bank had discovered the losses made by Leeson, the total amount was more then $1.3 billion and the bank collapsed.

Orange County, a suburban in California USA, had to file for bankruptcy in 1994 after heavy borrowing and risky investments resulting in a loss of $1.6 billion. This massive loss was the result of the risky trading strategy of the treasurer at that time, Bob Citron. His trading strategy resulted in higher returns at first, but when the US government started a series of six consecutive interest rate hikes his strategy resulted in severe losses.

Because of events such as the ones stated above it became evident that the risky positions taken by banks should be monitored more accurately.

Since the financial crisis of 2007/2008 the research on risk measures and the measurement of risk has become increasingly popular and the task of risk measurement has become increasingly complex. As in all situations when modelling complex issues and translating these into a workable model in practice, the difficulty lies in making a sufficiently realistic, but simple model. To quote Einstein:

> "As simple as possible, but not simpler."

# Contents

# Part I

# Background

# 1 Financial Risks

## 1.1 Introduction

There are many kinds of financial risks against which institutions need to guard themselves. In general you could divide them in five groups.

**Market Risk** The risk of losses due to volatilities and movements of the financial market prices. These include interest rate risk, exchange rate risk, investment risk and many more.

**Credit Risk** The risk of losses due to changes in the credit quality of the counterparties. Counterparty default is the most extreme case, but losses can already occur when a counterparty's credit merely decreases.

**Liquidity Risk** The risk of losses due to travel-time of securities and assets. In other words, the risk that a given security or asset cannot be traded quickly enough in the market to prevent a loss.

**Operational Risk** The risk of losses due to failed internal processes, the use of a wrong pricing model for instance. But also the risk of losses due to fraud and human mistakes.

**Legal & Regulatory Risk** This type of risk includes losses due to changes in tax laws for instance. But also the risk of losses due to the lacking of appropriate licenses.

Hence it is important to monitor the amount of risk a company is at, this is done by appointing a measure to a risky position. These risk measures can be classified into two categories, internal risk measures and external risk measures. Internal risk measurement is used at a level of individual institutions, in this case any institution is free to choose a risk model that best fits their beliefs and are allowed to choose which data they use.

While external risk measurement is used for external regulations and the same models are imposed for all relevant institutions.

## 1.2 Introduction Risk Measurement

Since the 1997 article of Artzner et all. [2] risk measurement, and hence risk measures, have gained enormously in interest under economist, bank regulators and mathematicians, giving rise to a new theory.

In my thesis I shall assume that risk measurement is a decision problem, ie a problem of the YES or NO type as is also assumed in [2]. I find this approach most suitable since risk is ever present. The question therefore doesn't concern the magnitude of the risk as much as the acceptability of the risky position.

**Definition 1.2.1** *Risk is the future net worth of a position.*

At the beginning risk measurement was mainly focussed on the mathematical properties which reflect the underlying economical meaning, however in the last years the statistical properties have become of increasing interest. Nowadays it is obvious to all working with risk, be it in practice or theory, that the procedure of risk measurement in fact involves two steps.

(1) Estimating the loss distribution of the position.

(2) Constructing a risk measure that summarizes the risk of the position.

The position's loss distribution in practice is generally unknown, and therefore must be estimated from (historical) data. The estimation is essentially done by backtesting. Recall that backtesting is the procedure of periodically comparing the forecasted risk measure with realized values in the financial market.

Each one of the steps above should be regarded as equally important. Because risk measurement is of great practical importance, risk measures should be formalized with the regulations of the practical world in mind. For this reason risk measures are mostly considered to be single valued, as will I do in this thesis.

Taking a risk to be a single value can be problematic however, for instance a single number does not give any information about which risk within the position is problematic. But this is only the case when a risk is found to be unacceptable, than the portfolio should be rebalanced. If on the other hand the risk is found to be acceptable, these sort of problems do not play any part. Thus in this setting taking a single valued risk measure is justified.

Furthermore, I shall assume that the set of all states of the financial world is known at the time we want to measure risk. This set may consist out of all the prices of assets and securities and all exchange rates amongst others.

# 2 Axiomatic Approach to Risk Measures

The axiomatic approach to risk measurement first started with the papers by Artzner et all [2] and Delbaen [11]. Since these articles awareness has come to mathematicians as well as banks and regulators that it is of the utmost importance to clearly define what it means for a functional of portfolio dynamics to be called a "risk measure". Writing down axioms is a crucial step in translating the complexity of measuring risk into a mathematical formulation.

The space of all financial positions will be denoted $\mathscr{X}$ and every position $X \in \mathscr{X}$ is a real-valued measurable function on a measurable space $\left(\Omega, \Sigma\right)$ with finite first moments. Here $X$ is the final net worth of a position at the end of a trading period. Let $\mathscr{X}_+$ denote the cone of non-negative elements of $\mathscr{X}$ and $\mathscr{X}_-$ the negative of $\mathscr{X}_+$. Moreover, let $\mathscr{X}_{++}$ and $\mathscr{X}_{--}$ be the cone op positive elements in $\mathscr{X}$ and its negative respectively.

**Remark** Note that taking $X \in \mathscr{X}$ to be the final net worth of a position already takes into account the time-value of money.

## 2.1 Acceptance Sets

It is not enough to simply have a risk measure which gives a single value to a position. We need also to define the concept of whether this value is acceptable or not, in other words does it belong to a set of acceptable risks. In Artzner et all [3] the following axioms are given for such an acceptance set, which I shall adopt for this thesis.

**Axiom 2.1** *The acceptance set $\mathscr{A}$ contains $\mathscr{X}_+$.*

**Axiom 2.2** *The acceptance set $\mathscr{A}$ does not intersect the set $\mathscr{X}_{--}$.*

The interpretation of these first two axioms is that a final net worth that is always non-negative does not require extra capital, however a net worth that is always strictly negative does. The next axiom is a natural requirement on the set of all acceptable final net worths and comes from axiom 2.7 for risk measures.

**Axiom 2.3** *The acceptance set $\mathscr{A}$ is a positively homogeneous cone.*

And finally the fourth axiom reflects risk aversion.

**Axiom 2.4** *The acceptance set $\mathscr{A}$ is convex.*

Risk averse behaviour is mathematically described through utility functions, which are continuous, strictly increasing and strictly concave functions. The latter is the reason why axiom 2.4 is needed.

The acceptance set $\mathscr{A}$, defined through the axioms stated above, is an important object to be considered when describing acceptance or rejection of risk.

## 2.2 Risk Measures

Recall the definition of a risk measure from section 1.2. I shall now formally define a risk measure as a mapping defined on the set of all financial positions.

**Definition 2.2.1** *A measure of risk is a mapping, $\mu : \mathscr{X} \to \mathbb{R}$.*

In Acerbi [1] a statistical functional which does not satisfy all the axioms stated in this section is not called a risk measure at all. I do agree with this line of reasoning, but in order to avoid confusions I shall refer to those statistical functionals as risk measures and I shall refer to them as coherent and elicitable risk measures, or true risk measures, whenever they do satisfy all the axioms. What is meant by coherent risk measures and elicitiable risk measures will be defined in section 3.1 and 3.2 respectively.

**Axiom 2.5** *Translation invariance. For all $X \in \mathscr{X}$ and all real number $a \in \mathbb{R}$, we have*

$$\mu(X + a) = \mu(X) - a \tag{2.1}$$

In other words adding the sure amount $a$ to the initial position the risk decreases by the same amount $a$.

**Axiom 2.6** *Reverse monotonicity. For all $X_1, X_2 \in \mathscr{X}$ such that $X_1 \leq X_2$, we have*

$$\mu(X_2) \leq \mu(X_1) \tag{2.2}$$

**Axiom 2.7** *Positive homogeneity. For all $k \geq 0$ and all $X \in \mathscr{X}$*

$$\mu(kX) = k\,\mu(X) \tag{2.3}$$

This axiom says that if a position is increased in size than the risk of that position increases with the same factor.

**Axiom 2.8** *Relevance. For all $X \in \mathscr{X}$ with $X \leq 0$ and $X \neq 0$ we have*

$$\mu(X) > 0 \tag{2.4}$$

This axiom is necessary (but not sufficient) to rule out that concentration of risk remains undetected. Where $X \neq 0$ should be interpreted that the position $X$ is not risk free.

**Axiom 2.9** *Subadditivity. For all $X_1, X_2 \in \mathscr{X}$, the following inequality holds*

$$\mu(X_1 + X_2) \leq \mu(X_1) + \mu(X_2) \tag{2.5}$$

This property reflects the general assumption that diversification of assets within a portfolio leads to a lowering of risk.

So far I have followed the setting of Artzner et all [3] completely, however for this thesis I shall include a sixth axiom which is on loss distributions rather than positions.

**Axiom 2.10** *Convex level sets: For all estimated distributions, $F_1$ and $F_2$, of a position $X \in \mathscr{X}$, such that $\mu_{F_1}(X) = \mu_{F_2}(X)$ and all $\lambda \in [0,1]$ we have that*

$$\mu_{\lambda F_1 + (1-\lambda)F_2}(X) = \mu_{F_1}(X) = \mu_{F_2}(X) \tag{2.6}$$

I include this axioms because so far the axioms only take into account the second step of the risk measurement procedure. This last axiom 2.10 is necessary for a risk measure to give robust results when performing backtests, as shall be explained in section 3.2.

An additional advantage of including this sixth axiom is that the axiomatic approach to risk measures as done by Artzner et all is not restrictive enough to specify a unique risk measure, but rather characterizes a whole class of risk measures. Having this extended list of axioms does give enough restrictions to define a unique risk measure.

In the past years it has been pointed out in many articles ( [13] [15] [17] [26] ) that there is yet another desirable property of a risk measure; comonotonic additivity. This property should reflect the worst case scenario for the correlation between risks. The reason that

I do not include this property is that the mathematical requirements, that future net worths of different risky assets should be perfect substitutes, is in general not met in finance. Perfect substitutes in economics are two goods (assets) for which the utility of the one is the same as the utility of the other, for example one Euro form Company A which is a FOREX company is equal to one Euro from another FOREX company B. The following example makes it clear that this is generally not true in practice.

**Example 2.2.2** *Suppose we have two put-options on the same asset in our portfolio, each one with the same maturity and the same strike price. One of those put-options we have at Bank A and the other we have at bank B. Let de default probability of Bank A and Bank B be 0.005 and 0.007 respectively. The riskiness of these two put-options should be equal if they were to be perfect substitutes of each other, however the default probabilities of both banks are different and hence this is not the case.*

There is natural way to define a measure of risk. The number $\mu(X)$, assigned by the measure $\mu$ to the risk $X \in \mathscr{X}$ shall be interpreted as the minimum that needs to be added to the risky position in order to make it an acceptable risk. This is formalized in the next definition.

**Definition 2.2.3** *A risk measure that is associated with the acceptance set $\mathscr{A}$ is the mapping $\mu_{\mathscr{A}} : \mathscr{X} \to \mathbb{R}$ defined by*

$$\mu_{\mathscr{A}}(X) = \inf\Big\{m : m + X \in \mathscr{A}\Big\} \tag{2.7}$$

Thus if we have $\mu(X) < 0$ then the amount $-\mu(X)$ may be withdrawn from the position. In definition 2.2.4 a correspondence between acceptance sets and measures of risk are given.

**Definition 2.2.4** *An acceptance set associated with a risk measure $\mu$, denoted $\mathscr{A}_\mu$, is defined by*

$$\mathscr{A}_\mu = \Big\{X \in \mathscr{X} : \mu(X) \leq 0\Big\} \tag{2.8}$$

Which shows that the acceptance set defines the risk measures just as the risk measure defines the acceptance set.

# 3 Classes of Risk Measures

In this chapter I shall discuss two classes of risk measures, coherent risk measures and elicitable risk measures, and I shall show that these two classes could be regarded to represent one of the steps within the risk measurement procedure. The emphasis in this thesis is on risky positions of banks, and hence all risk measures discussed are so-called monetary risk measures.

**Definition 3.0.5** *A monetary risk measure is a function $\mu : \mathscr{X} \to \mathbb{R}$ such that for all positions $X \in \mathscr{X}$*

*(1) If $X \geq 0$ then $\mu(X) \leq 0$.*

*(2) For $a \in \mathbb{R}$ it holds that $\mu(X + a) = \mu(X) - a$.*

*(3) For $k \in \mathbb{R}$, such that $k \geq 0$ it holds that $\mu(kX) = k\,\mu(X)$.*

## 3.1 Coherent Risk Measures

This class of risk measures was first introduced by Artzner et all. [2]. And was constructed to possess all mathematical properties to properly reflect the economy. And hence the class of coherent risk measures takes the second step within the risk measurement procedure into account. A risk measure is called coherent if it satisfies the axiom 2.5-2.9. We have the following definition.

**Definition 3.1.1** *A coherent monetary risk measure is a mapping $\rho : \mathscr{X} \to \mathbb{R}$ such that $\rho$ is a monetary risk measure that is sub-additive. That is for all $X_1, X_2 \in \mathscr{X}$*

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2) \tag{3.1}$$

I shall often suppress the term monetary in the definition. A major advantage is that the properties above completely charaterize the class of coherent risk measures. And hence the following two propositions on the correspondence between a coherent risk measure $\rho$ and its associated acceptance set $\mathscr{A}_\rho$ hold.

**Proposition 3.1.2** *If a set $\mathscr{C}$ satisfies axioms 2.1-2.4 then the associated risk measure $\rho_{\mathscr{C}}$ is coherent. Moreover, the closure of $\mathscr{C}$ is the associated acceptance set.*

**Proof** From Rockafeller [34] axioms 2.2 and 2.4 ensure that for a convex function $h(\cdot)$ and a convex set which is bounded from below, $\mathscr{C}$, the infimum

$$\rho_{\mathscr{C}}(X) = \inf\{h(X) : X \in \mathscr{C}\} \tag{3.2}$$

is attained. Where we take $h(X) = X - m$. Hence we have $\rho_{\mathscr{C}}(X) < \infty$ for all $X \in \mathscr{C}$. Then from the equality

$$\inf\Big\{p : X + (a+p) \in \mathscr{C}\Big\} = \inf\Big\{q : X + q \in \mathscr{C}\Big\} - a \tag{3.3}$$

it follows that

$$\rho_{\mathscr{C}}(X + a) = \rho_{\mathscr{C}}(X) - a \tag{3.4}$$

And hence axiom 2.5 is satisfied. The sub-additivity of $\rho_{\mathscr{C}}$ follows from the fact that if

$$X_1 + b_1 \in \mathscr{C}, \quad \text{and} \quad X_2 + b_2 \in \mathscr{C} \tag{3.5}$$

then from axioms 2.3 and 2.4 we know that

$$X_1 + X_2 + (b_1 + b_2) \in \mathscr{C} \tag{3.6}$$

Now to tackle positive homogeneity. Let $c$ be such that $c \geq \rho_{\mathscr{C}}(X)$ then for all $k > 0$ we have

$$k \cdot X + k \cdot c \in \mathscr{C} \tag{3.7}$$

by axiom 2.3 and definition 2.2.4 proving that

$$\rho_{\mathscr{C}}(k \cdot X) \leq k \cdot c \tag{3.8}$$

If on the other hand we have that $c \leq \rho_{\mathscr{C}}(X)$ then for all $k > 0$ we have

$$k \cdot X + k \cdot c \notin \mathscr{C} \tag{3.9}$$

proving that

$$\rho_{\mathscr{C}}(k \cdot X) \geq k \cdot c \tag{3.10}$$

and hence we conclude equality,

$$\rho_{\mathscr{C}}(k \cdot X) = k \cdot \rho_{\mathscr{C}}(X) \tag{3.11}$$

Monotonicity of $\rho_{\mathscr{C}}$ follows from axiom 2.1 and 2.4 together with definition 2.2.4, since if we have that

$$X_1 \leq X_2 \quad \text{and} \quad X_1 + d \in \mathscr{C} \quad \Rightarrow \quad X_2 + d \in \mathscr{C} \tag{3.12}$$

And finally, for each $X \in \mathscr{C}$ we have $\rho_{\mathscr{C}}(X) \leq 0$ and hence $X \in \mathscr{C}$ then all the above together with definition 2.2.4 ensures that $\mathscr{C}$ is closed, which proves that

$$\mathscr{A}_{\rho_{\mathscr{C}}} = \bar{\mathscr{C}} \tag{3.13}$$

■

**Proposition 3.1.3** *If a risk measure* $\rho$ *is coherent, then the acceptance set* $\mathscr{A}_\rho$ *is closed and satisfies axioms 2.1-2.4.*

**Proof** Subadditivity and positive homogeneity ensure that the risk measure $\rho$ is convex on the set $\mathscr{A}_\rho$ and therefore continuous. Hence we see that the set

$$\mathscr{A}_\rho = \left\{X : \rho(X) \leq 0\right\} \tag{3.14}$$

is a closed, convex and homogeneous cone. Positive homogeneity also implies that $\rho(0) = 0$ and together with monotinicity this ensures that the acceptance set $\mathscr{A}_\rho$ contains the set $\mathscr{X}_+$. Then let $X \in \mathscr{X}_{--}$ such that $\rho_{\mathscr{A}}(X) < 0$. Axiom 2.6 on monotonicity than states that $\rho(0) < 0$, which is a contradiction. If we now take $X$ such that $\rho(X) = 0$, then we can find $a > 0$ such that

$$X + a \in \mathscr{X}_{--} \tag{3.15}$$

again this leads to a contradiction, since by axiom 2.5 on translation invariance we have that $-a > 0$. We conclude that we must have $\rho(X) > 0$, hence

$$X \notin \mathscr{A}_\rho \tag{3.16}$$

and we have proven axiom 2.2. For each $X$ let $b$ be an arbitrary number such that $\rho_{\mathscr{A}} < b$, then

$$X + b \in \mathscr{A}_\rho \tag{3.17}$$

and hence

$$\rho\,(X + b) \leq 0 \tag{3.18}$$

hence $\rho(X) \leq b$ which proves that

$$\rho(X) \leq \rho_{\mathscr{A}}(X) \quad \text{or equivalently} \quad \rho \leq \rho_{\mathscr{A}} \tag{3.19}$$

Also for each $X$ let $c$ be an arbitrary number such that $\rho < c$, then

$$X + c \in \mathscr{A}_\rho \tag{3.20}$$

and hence

$$\rho_{\mathscr{A}}\,(X + c) \leq 0 \tag{3.21}$$

hence $\rho_{\mathscr{A}}(X) \leq c$ which proves that

$$\rho_{\mathscr{A}}(X) \leq \rho(X) \quad \text{or equivalently} \quad \rho_{\mathscr{A}} \leq \rho \tag{3.22}$$

And we conclude equality and this proves the proposition.

■

## 3.2 Elicitable Risk Measures

Making forecasts about an uncertain future is essential for measuring risk. Surely these forecast are probabilistic in nature and take the form of probability distributions. To evaluate these forecasts, scoring functions may be used. A scoring function $S$ depends both on the forecasts and on the observations.

Let $F$ be an estimated loss distribution for the current position $X \in \mathscr{X}$. Denote by $\mathscr{F}$ the set of all estimated distributions. Furthermore, let $Y$ be a set of verifying (historical) risks. Assume that we have a monetary risk measure $\nu : \mathscr{X} \to \mathbb{R}$. Then a scoring function, $S : \mathbb{R}^2 \to [0, \infty)$, assigns a score to the accuracy of the estimate $t = \nu_F(X)$. Where $\nu_F(X)$ should be interpreted as the risk given that the distribution of the position is estimated by $F$.

Although scoring functions can be set-valued in this thesis it is enough to look at single-valued scoring functions, as only single-valued risk measures are considered. I shall adopt the following definition for a scoring function from Gneiting [19].

**Definition 3.2.1** *Given a position $X \in \mathscr{X}$, a scoring function $S : \mathbb{R}^2 \to [0, \infty)$ satisfies for any estimate $t = \nu_F(X)$ and observation $y \in Y$*

*(1)* $S(t, y) \geq 0$ *and* $S(t, y) = 0$ *if and only if* $t = y$.

*(2)* $S(t, y)$ *is continuous in* $t$.

*(3)* $\frac{\partial S(t, y)}{\partial t}$ *exists and is continuous in* $t$ *whenever* $t \neq y$.

There are many considerations to be made in choosing a suitable scoring function, it should be such that it can be used in practice but also have sound theoretical support. In this chapter the focus will be on scoring functions that make the corresponding statistical functional 'elicitable'. Since this thesis is about risk measures, I shall only focus on these statistical functionals.

Whenever an estimate receives a distributional feature from a risk measure it is important that the risk measure is robust for the estimate in the sense that the expected score is minimized.

**Definition 3.2.2** *Given a position $X \in \mathscr{X}$ the scoring function $S$ is weakly-robust for the risk measure $\nu$, relative to the class $\mathscr{F}$ if for all $F \in \mathscr{F}$*

$$\mathbb{E}_F S(t, Y) \leq \mathbb{E}_F S(x, Y) \tag{3.23}$$

*for all* $t = \nu_F(X)$ *and all* $x \in \mathbb{R}$. *Moreover, it is (strictly) robust if it is weakly robust and equality in equation (3.23) implies that* $x = \nu_F(X)$.

This definition of (weak) robustness is used in Gneiting [19], although it is there referred to as (strict) consistency. I changed the terminology because I believe the term robustness is more appropriate when describing sensitivity with respect to estimated distributions from a dataset. Also the term robustness is more accepted in practice.

**Definition 3.2.3** *The risk measure* $\nu$ *is elicitable relative to the class* $\mathscr{F}$ *if there exits a scoring function* $S : \mathbb{R}^2 \to [0, \infty)$ *that is robust for* $\nu$ *relative to the class* $\mathscr{F}$.

Suppose we wouldn't have a sstrict robust risk measure in the above sense, then the risk values given by different distribution could deviate largely from each othere. And a bank would be able to "choose" a distribution which returns the lowest risk measure. If Therefore from a regulatory point of view, robustness is significant.

In Krätschmer [27] it is argued that qualitative robustness is not an ideal way to represent robustness as it makes a division into 'robust' and 'non-robust' risk measures, they therefore suggest using a continuum of possible degrees of robustness. However, from my point of view the dichotomic classification of Gneiting fits very well with the scope of this thesis: that risk measurement is a decision problem of the YES or NO type.

**Proposition 3.2.4** *Let* $X \in \mathscr{X}$ *be given. If a mapping* $\nu : \mathscr{X} \to \mathbb{R}$ *is a monetary elicitable risk measure then it is a monetary risk measure such that, if* $F_1$ *and* $F_2$ *are estimated distribution functions wich* $\nu_{F_1}(X) = \nu_{F_2}(X)$, *we have for all* $\lambda \in [0,1]$ *it holds that*

$$\nu_{\lambda F_1 + (1-\lambda) F_2}(X) = \nu_{F_1}(X) = \nu_{F_2}(X) \tag{3.24}$$

**Proof** To proof property 3.24, suppose that the risk measure $\nu$ is elicitable relative to the class $\mathscr{F}$. Then there exists a scoring function $S : \mathbb{R}^2 \to [0, \infty)$ which is robust for it relative to $\mathscr{F}$. Also suppose that we have $F_1, F_2 \in \mathscr{F}$ and $t \in \mathbb{R}$ such that

$$t = \nu_{F_1}(X), \qquad \text{and} \qquad t = \nu_{F_2}(X) \tag{3.25}$$

If we let $x \in \mathbb{R}$ be arbitrary and $\lambda \in [0,1]$ such that equation (3.23) holds then for the observations $Y$ we have

$$\begin{aligned} \mathbb{E}_{F_\lambda} S(t, Y) &= (1-\lambda)\mathbb{E}_{F_1} S(t, Y) + \lambda \mathbb{E}_{F_2} S(t, Y) \\ &\leq (1-\lambda)\mathbb{E}_{F_1} S(x, Y) + \lambda \mathbb{E}_{F_2} S(x, Y) \\ &= \mathbb{E}_{F_\lambda} S(x, Y) \end{aligned} \tag{3.26}$$

And hence we have $x = \nu_{F_\lambda}(X)$.

■

This last property is that of convex level sets. Note that this is different from standard convexity which applies to sums, but a special case of quasi-convexity (see appendix A). Convex level sets apply to mixtures, whereas standard convexity applies to sums. The following is a practical proposition on elicitable risk measures from Osband [33].

**Proposition 3.2.5** *Let $f : \mathbb{R} \to \mathbb{R}$ be a one-to-one mapping. Then the following two statements hold.*

*(1) If $\nu$ is elicitable relative to the class $\mathscr{F}$, then the risk measure $f \circ \nu$ is elicitable relative to the class $\mathscr{F}$.*

*(2) If the scoring function $S : \mathbb{R} \to [0, \infty)$ is robust for the risk measure $\nu$, then the scoring function $S_f : \mathbb{R}^2 \to [0, \infty)$ defined by*

$$S_f(x, y) = S\big(f^{-1}(x), y\big) \tag{3.27}$$

*is robust for the risk measure $f \circ \nu$.*

**Proof** The theorem follows directly from the following inequality. Let $\tilde{t} = f \circ \nu(F)$ and $\tilde{x} \in \mathbb{R}$. then $\tilde{t} = f(\tilde{t})$ where $t = \nu_F(X)$ and $\tilde{x} = f(x)$, for $x \in \mathbb{R}$. Hence

$$\mathbb{E}_F S_f(\tilde{t}, Y) = \mathbb{E}_F S(t, Y) \leq \mathbb{E}_F S_f(x, Y) = \mathbb{E}_F S_f(\tilde{x}, Y) \tag{3.28}$$

where the inequality follows from the fact the $\nu$ is (weakly) robust. The theorem then follows from the fact that we have equality if and only if $x = \nu_F(X)$ or equivalently $\tilde{x} = f \circ \nu_F(X)$.

■

Note that the class of elicitable risk measures is less conservative then that of coherent risk measures, since in this case the associated acceptance set, $\mathscr{A}_\nu$, needs not to be convex, but we merely require that it is closed under convex level sets.

**Lemma 3.2.6** *If a risk measure $\nu$ is elicitable, then the acceptance set $\mathscr{A}_\nu$ is closed under convex level sets and satisfies axioms 2.1-2.3.*

**Proof** The fact that axiom 2.1 through axiom 2.3 are satisfied has already been done in the previous chapter. What is left to show is that $\mathscr{A}_\nu$ is closed under convex level sets. This follows from the fact that positive homogeneity and convex levels sets ensure that $\nu$ is continuous on all level sets.

■

# Part II

# Risk Measures

# 4 Value-at-Risk

Probably the most widely used risk measure at this time is Value-at-Risk, abbreviated VaR. This risk measure was developed in the 1990's as a response to financial disasters like Baring's Bank and Orange County mentioned in the preface of this thesis. In the 1990's there was a whole string of financial disasters like these and the VaR risk measure was developed as a tool to warn investors of the risks they are incurring. Although developed in the 1990's, the methodology behind VaR is not new, it can be traced back to 1952 to the basic mean-variance framework of Markowitz [30]. Moreover, the VaR principle was used in actuarial sciences long before it was reinvented for investment banking. Although, within actuarial sciences the more common phrase was the quantile risk measure as opposed to Value-at-Risk.

## 4.1 Definition Value-at-Risk

VaR is a statistical measure of downside risk based on current positions. The great advantage of VaR is that it is simple to compute and easy to understand. Informally, VaR can be defined as the worst loss over a target horizon such that with a pre-specified probability that the actual loss will be higher. The formal mathematical definition is the following:

**Definition 4.1.1** *Given a position $X \in \mathscr{X}$, and $\alpha \in [0,1]$ we define*

$$VaR_\alpha(X) = -q^{(\alpha)}(X) \tag{4.1}$$

*where and $q^{(\alpha)}(X)$ is the smallest $\alpha$-quantile, ie*

$$q^{(\alpha)}(X) = \inf\Big\{x : \mathbb{P}\left[X \leq x\right] \geq \alpha\Big\} \tag{4.2}$$

It is then easily seen that the acceptance set for $\mathrm{VaR}_\alpha$ is given by

$$\mathscr{A}_{\mathrm{VaR}_\alpha} = \Big\{X \in \mathscr{X} : \mathbb{P}\Big[X \leq 0\Big] \geq \alpha\Big\} \tag{4.3}$$

There is another way of defining the acceptance set for VaR, it is easily seen that the above representation is equivalent to

$$\mathscr{A}_{\mathrm{VaR}_\alpha} = \left\{ X \in \mathscr{X} : \frac{\mathbb{P}\left[X \geq 0\right]}{\mathbb{P}\left[X < 0\right]} \geq \frac{1-\alpha}{\alpha} \right\} \tag{4.4}$$

Since VaR is a quantile measure extra attention should be paid to discontinuities and intervals of the quantiles. Note that the only case in which the choice between using the largest or the smallest quantile becomes non-arbitrary is whenever $q^{(\alpha)} \neq q_{(\alpha)}$, where

$$q_{(\alpha)}(X) = \inf\left\{x : \mathbb{P}\left[X \leq x\right] \geq \alpha\right\} \tag{4.5}$$

is the smallest quantile. This can only be the case when we are dealing with a discrete distribution of the loss variable.

**Lemma 4.1.2** *VaR has the following four properties:*

*(1)* *If $X \geq 0$, then $VaR_\alpha(X) \leq 0$.*

*(2)* *If $X \geq Y$, then $VaR_\alpha(X) \leq VaR_\alpha(Y)$.*

*(3)* *For $\lambda \in \mathbb{R}$, such that $\lambda \geq 0$, then $VaR_\alpha(\lambda X) = \lambda VaR_\alpha(X)$.*

*(4)* *For $k \in \mathbb{R}$, it holds that $VaR_\alpha(X+k) = VaR_\alpha(X) - k$.*

**Proof** We may rewrite equation 4.1 as

$$\mathrm{VaR}_\alpha(X) = \inf\left\{m \in \mathbb{R} : \mathbb{P}\left[X + m < 0\right] \leq \alpha\right\} \tag{4.6}$$

Then all properties of the proposition follow either from this equation or from the definition of the $\alpha$-quantile.

∎

Thus VaR is a monotonic, homogeneous and translation invariant risk measure by definition. I shall now show that VaR is an elicitable risk measures, based on the following proposition from Gneiting [19] on quantiles.

**Proposition 4.1.3** *The $\alpha$-quantile is elicitable relative to the class $\mathscr{F}$.*

**Proof** In Gneiting 2008 [20] it is shown that if $g : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function, then the scoring function

$$S(x,y) = \left(\mathbb{1}_{\{x \geq y\}} - \alpha\right)\left(g(x) - g(y)\right) \tag{4.7}$$

is weakly-robust for $q_\alpha$ relative to the class $\mathscr{F}$. Moreover, if the expectation exists and is finite

$$\mathbb{E}_F g(Y) < \infty \tag{4.8}$$

for all $F \in \mathscr{F}$, then the scoring function $S : \mathbb{R}^2 \to [0,\infty)$ is robust relative to the class $\mathscr{F}$ and hence the risk measure $q_{(\alpha)}$ is then elicitable. Take $g : \mathbb{R} \to \mathbb{R}$ to be the function defined by

$$g(x) = \frac{1}{1+e^{-x}} \tag{4.9}$$

Then obviously $g(\cdot)$ is bounded and increasing and hence meets the conditions stated above. Thus the scoring function

$$S(x,y) = \left(\mathbb{1}_{\{x \geq y\}} - \alpha\right)\left(\frac{1}{1+e^{-x}} - \frac{1}{1+e^{-y}}\right) = \left(\mathbb{1}_{\{x \geq y\}} - \alpha\right)\frac{e^{-y} - e^{-x}}{(1+e^{-x})(1+e^{-y})} \tag{4.10}$$

is a robust scoring function which makes $q_{(\alpha)}$ into an elicitable risk measure.

■

**Theorem 4.1.4** *$\mathrm{VaR}_\alpha$ is elicitable relative to the class $\mathscr{F}$.*

**Proof** It follows directly from propositions 3.2.5 and 4.1.3 that the function

$$S_{\mathrm{VaR}_\alpha}(x,y) = \left(\mathbb{1}_{\{-x \geq y\}} - \alpha\right)\frac{e^{-y} - e^{x}}{(1+e^{x})(1+e^{-y})} \tag{4.11}$$

elicits $\mathrm{VaR}_\alpha$.

■

The scoring function defined in the proof of the theorem above is not one of the most widely used scoring functions in practice however. I found that in most literature it is recommended to use the so-called piecewise linear scoring function for the evaluation of VaR forecasts, or for quantile risk measure in general The piecewise linear scoring function, $S_{\mathrm{PL}} : \mathbb{R}^2 \to [0,\infty)$ is given by

$$S_{\mathrm{PL}}(x,y) = \left(\mathbb{1}_{\{x \geq y\}} - \alpha\right)\left(x - y\right) \tag{4.12}$$

Obviously the scoring function is of the form from equation (4.7), with $g : \mathbb{R} \to \mathbb{R}$ is taken to be the identity function

$$g(x) = x \tag{4.13}$$

which is not bounded and hence we cannot guarantee the existence of the expectation of equation (4.8). In practice this is not a problem since it is mostly assumed that

$$g : I \to [0, \infty) \tag{4.14}$$

where $I \subset \mathbb{R}$ is a bounded subset.

## 4.2 Limitations of Value-at-Risk

One of the raised limitations of VaR is that it does not say anything about the severity of the losses after the VaR-value. An even more serious problem with VaR is that it is not sub-additive. It is commonly thought of that diversification of risk leads to risk reduction. However, diversification of risk may lead to higher VaR values. In Acerbi [1] this is illustrated with the following two simple discrete examples.

**Example 4.2.1** *Suppose a bank loans Company A \$100.000,- and that this company will default on the loan with a probability of 0.8%. Suppose, furthermore, that Company A either defaults the entire loan, or not at all. Thus if we denote by $X_1$ the default amount of a portfolio with just this one loan we get*

$$X_1 = \begin{cases} -\$100.00,- & \text{if Company A defaults} \\ \$0,- & \text{otherwise} \end{cases} \tag{4.15}$$

*then the distribution of $X_1$ is discrete and follows*

$$\mathbb{P}\Big[X_1 = -\$100.000,-\Big] = 0.008, \qquad \text{and} \qquad \mathbb{P}\Big[X_1 = \$0,-\Big] = 0.992 \tag{4.16}$$

*And hence if $\alpha = 0.01$, then the VaR of this portfolio (consisting of this one single loan) satisfies*

$$VaR_{0.01}(X_1) = -\inf\Big\{x : \mathbb{P}\left[X_1 \le x\right] > 0.01\Big\} = 0 \tag{4.17}$$

*If the bank would have diversified this amount, say the bank loans \$50.000,- to Company A and another \$50.000 to Company B, where both companies have the same default probability of 0.8% and like in the first scenario both companies either default on their entire loan or not at all, then, if we denote the default amount in this case by $X_2$, this yields*

$$X_2 = \begin{cases} -\$100.00,- & \text{if Company A and Company B default} \\ -\$50.00,- & \text{if Company A or Company B defaults} \\ \$0,- & \text{otherwise} \end{cases} \tag{4.18}$$

*The distribution of $X_2$ is therefore given by*

$$\mathbb{P}\Big[X_2 = -\$100.000,-\Big] = 0.000064, \quad \mathbb{P}\Big[X_2 = -\$50.000,-\Big] = 0.016, \quad \mathbb{P}\Big[X_2 = 0\Big] = 0.983936 \tag{4.19}$$

*Showing that when taking* $\alpha = 0.01$ *the VaR for this portfolio is given by*

$$VaR_{0.01}(X_2) = -\inf\Big\{y : \mathbb{P}[X_2 \leq y] > 0.01\Big\} = \$50.000,\text{-} \tag{4.20}$$

□

Another important consequence of lacking the sub-additivity property is illustrated by the next example. Showing that the dangers of concentrating credit risk are also missed by VaR.

**Example 4.2.2** *Consider the issue of corporate bonds in a market with zero base rate, all corporate bond spreads equal 2% and the default by any company is set at 1%. Taking* $\alpha = 0.05$, *a loan of \$1.000.000,- invested in bonds with a single company thus gives the following VaR:*

$$VaR_{\alpha}(X) = -\$20.000 \tag{4.21}$$

*Note that this indicates that this loan is* $VaR_{0.05}$*-acceptable and there is no risk.*

*If we now consider the loan is placed in bonds issued independently by 100 companies at \$10.000,- each. Then the probability that two companies will default is*

$$\binom{100}{2}(0.01)^2(0.99)^{98} \approx 0.185 \tag{4.22}$$

*and hence in this case we see that* $VaR_{\alpha} > 0$. *And again diversification does not lead to lower risk according to VaR. Moreover the portfolio where the loan was invested in a single company resulted in an acceptable risk according to VaR, but the portfolio when the loan was spread over 100 companies gave an unacceptable value.*

□

From the above examples is it clear that VaR doesn't belong to the class of coherent risk measures. However, practitioners sometimes argue that the use of Value-at-Risk is justified by the following proposition which says that VaR is sub-additive under a normal distribution.

**Proposition 4.2.3** *If quantiles are computed under a normal distribution, then the quantiles do satisfy the property of sub-additivity as long as probabilities of exceedence are smaller than 0.5.*

**Proof** Denote by $\mathbb{V}\mathrm{ar}$ the variance, then indeed $\mathbb{V}\mathrm{ar}_{X+Y} \leq \mathbb{V}\mathrm{ar}_X + \mathbb{V}\mathrm{ar}_Y$ for each pair of random variables $(X,Y)$ such that they are jointly normally distributed. Since for a normal distributed random variable $Z$ we have the following representation for VaR,

$$\mathrm{VaR}_\alpha(Z) = -\left[\mathbb{E}Z + \Phi^{-1}(\alpha)\cdot\mathbb{V}\mathrm{ar}Z\right] \tag{4.23}$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution function and hence we have that $\Phi^{-1}(0.5) = 0$. If we thus take $Z = X + Y$ the above equation yields the result.

$$\begin{aligned}
\mathrm{VaR}_\alpha(X+Y) &= -\left[\mathbb{E}(X+Y) + \Phi^{-1}(\alpha)\cdot\mathbb{V}\mathrm{ar}(X+Y)\right] \\
&\leq -\left[\mathbb{E}X + \mathbb{E}Y + \Phi^{-1}(\alpha)\cdot(\mathbb{V}\mathrm{ar}X + \mathbb{V}\mathrm{ar}Y)\right] \\
&= -\left[\mathbb{E}X + \Phi^{-1}(\alpha)\cdot\mathbb{V}\mathrm{ar}X\right] - \left[\mathbb{E}Y + \Phi^{-1}(\alpha)\cdot\mathbb{V}\mathrm{ar}Y\right] \\
&= \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y)
\end{aligned} \tag{4.24}$$

■

# 5 Expected Shortfall

Expected shortfall, sometimes also referred to as Conditional-Value-at-Risk or CVaR, was at one time the most prominent risk measure as a potential replacement for the VaR risk measure since it does possess the property of sub-additivity.

## 5.1 Definition of Expected Shortfall

In chapter 4 its shown that VaR answers the question

$$\textit{What is the minimum loss in the } \alpha \cdot 100\% \textit{ worst cases of our portfolio?} \tag{5.1}$$

Due to the term "minimum loss" in the definition VaR is not sub-additive and hence is not a true risk measure. Moreover, VaR is indifferent of how serious the losses beyond the VaR-value really are. And hence, it is useful to modify the above question to the following

$$\textit{What is the expected loss incurred in the } \alpha \cdot 100\% \textit{ worst cases of our portfolio?} \tag{5.2}$$

Expected Shortfall was constructed in a bottom-up fashion so that it would be sub-additive. It is easily seen that whenever the profit-loss distribution is continuous the answers to question 5.2 is given by the conditional expected value beyond the lower $\alpha$-quantile. Define the risk measure

$$\mathrm{TCE}_\alpha(X) = -\mathbb{E}\left[X \,|\, X \leq q^{(\alpha)}(X)\right] \tag{5.3}$$

where TCE stands for tail conditional expectation and as before $q^{(\alpha)}(X)$ is the $\alpha$-quantile of $X$. For general distribution equation (5.3) does not hold anymore since it could be the case that the event $\left\{X \leq q^{(\alpha)}(X)\right\}$ may have a probability larger that $\alpha$. Moreover, TCE is a coherent risk measure only if we restrict ourselves to continuous distributions.

Let $\left(X_i\right)_{i=1}^n$ be $n$ realizations of the loss random variable $X$, and we define the order statistics

$$X_{1:n}, X_{2:n}, \ldots, X_{n:n} \tag{5.4}$$

as the sorted values of the vector $(X_1,\ldots,X_n)$. Moreover, if we approximate the $\alpha\cdot 100\%$ elements in the sample by

$$w = \max\Big\{m : m \le n\alpha,\, m \in \mathbb{N}\Big\} \tag{5.5}$$

The set of $\alpha\cdot 100\%$ worst outcomes of $X_1,\ldots,X_n$ is thus represented by the least $w$ outcomes $\Big\{X_{1:n},\ldots,X_{w:n}\Big\}$. And a natural estimator for the $\alpha$-quantile $q^{(\alpha)}$ is therefore

$$\hat{q}^{(\alpha)}(X) = X_{w:n} \tag{5.6}$$

where the estimator is dependent on the sample size $n$. A natural estimator for the expected loss in the $\alpha\cdot 100\%$ of the worst cases then becomes

$$\hat{\text{ES}}_\alpha(X) = -\frac{\sum_{i=1}^{w} X_{i:n}}{w} \tag{5.7}$$

This shall be the $\alpha$-expected shortfall of the sample. Note that in this case we also have

$$\hat{\text{TCE}}_\alpha(X) = -\frac{\sum_{i=1}^{n} X_i \mathbb{1}_{\{X_i \le X_{w:n}\}}}{\sum_{i=1}^{n} \mathbb{1}_{\{X_i \le X_{w:n}\}}} \tag{5.8}$$

To show that sub-additivity is met, let $\Big(X_i, Y_i\Big)_{i=1}^{n}$ be simultaneous realizations then

$$\begin{aligned}
\hat{\text{ES}}_\alpha(X+Y) &= -\frac{\sum_{i=1}^{w}(X+Y)_{i:n}}{w} \\
&\le -\frac{\sum_{i=1}^{w}\left(X_{i:n}+Y_{i:n}\right)}{w} \\
&= \hat{\text{ES}}_\alpha(X) + \hat{\text{ES}}_\alpha(Y)
\end{aligned} \tag{5.9}$$

And hence if we understand this estimator, we are most likely to find a coherent risk measure. If we expand the definition of the estimator

$$\begin{aligned}
\hat{\text{ES}}_\alpha(X) &= -\frac{\sum_{i=1}^{w} X_{i:n}}{w} \\
&= -\frac{\sum_{i=1}^{n} X_{i:n}\mathbb{1}_{\{i\le w\}}}{w} \\
&= -\frac{1}{w}\left(\sum_{i=1}^{n} X_{i:n}\mathbb{1}_{\{X_{i:n}\le X_{w:n}\}} - \sum_{i=1}^{n} X_{i:n}\left[\mathbb{1}_{\{X_{i:n}\le X_{w:n}\}} - \mathbb{1}_{\{i\le w\}}\right]\right) \\
&= -\frac{1}{w}\left(\sum_{i=1}^{n} X_i\mathbb{1}_{\{X_{i:n}\le X_{w:n}\}} - X_{w:n}\sum_{i=1}^{n}\left[\mathbb{1}_{\{X_{i:n}\le X_{w:n}\}} - \mathbb{1}_{\{i\le w\}}\right]\right) \\
&= -\frac{n}{w}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\mathbb{1}_{\{X_i\le X_{w:n}\}} - X_{w:n}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{X_i\le X_{w:n}\}} - \frac{w}{n}\right]\right)
\end{aligned} \tag{5.10}$$

If, with probability one, we would have

$$\lim_{n\to\infty} X_{w:n} = q^{(\alpha)} \tag{5.11}$$

Then, also with probability one, we would have

$$\lim_{n\to\infty} \hat{\mathrm{ES}}_\alpha(X) = -\frac{1}{\alpha}\Big(\mathbb{E}\Big[X\mathbb{1}_{\{X\le q^{(\alpha)}(X)\}}\Big] - q^{(\alpha)}(X)\Big(\mathbb{P}\Big[X \le q^{(\alpha)}(X)\Big] - \alpha\Big)\Big) \tag{5.12}$$

While the limit in equation (5.11) does not hold when $q^{(\alpha)} \neq q_{(\alpha)}$, it is shown by Acerbi [**?** ] that the limit in equation (5.12) does hold in full generality. We can then formulate the following definition.

**Definition 5.1.1** *Let $X \in \mathscr{X}$ be given and let $\alpha \in (0,1)$ be a pre-specified probability level. The $\alpha$-expected shortfall of the position is then defined through*

$$ES_\alpha(X) = -\frac{1}{\alpha}\Big(\mathbb{E}\Big[X\mathbb{1}_{\{X\le q^{(\alpha)}\}}\Big] - q^{(\alpha)}\Big(\mathbb{P}\Big[X \le q^{(\alpha)}\Big] - \alpha\Big)\Big) \tag{5.13}$$

In this definition the term $q^{(\alpha)}\Big(\mathbb{P}[X \le \mathrm{VaR}_\alpha] - \alpha\Big)$ in equation (5.13) should be interpreted as the exceeding part to be subtracted from the expected value $\mathbb{E}\Big[X\mathbb{1}_{\{X\le q^{(\alpha)}\}}\Big]$ when the event $\Big\{X \le q^{(\alpha)}\Big\}$ has probability larger than $\alpha$. And hence it is easily seen that when we have a continuous distribution, ie $\mathbb{P}\Big[X \le q^{(\alpha)}\Big] = \alpha$, then the extra term vanishes and we have $\mathrm{ES}_\alpha(X) = \mathrm{TCE}_\alpha(X)$.

There is a more fundamental definition of equation (5.13) which better reveals the dependence on both the parameter $\alpha$ and the distribution function $F(x) = \mathbb{P}\Big[X \le x\Big]$. For this recall the generalized inverse function $F^{\leftarrow}(p)$,

$$F^{\leftarrow}(p) = \inf\Big\{x : F(x) \ge p\Big\} \tag{5.14}$$

The expected shortfall can then be expressed as minus the mean of $F^{\leftarrow}(p)$ for $p \in (0,\alpha]$, thus

$$\mathrm{ES}_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha F^{\leftarrow}(p)\,dp \tag{5.15}$$

In equation (5.15) also continuity in $\alpha$ is made clear directly. This distinguishing equation of the expected shortfall is not shared which TCE and VaR. We may also define the expected shortfall through the equation

$$\mathrm{ES}_\alpha(X) = \mathrm{TCE}_\alpha(X) + (\lambda - 1)\Big(\mathrm{TCE}_\alpha(X) - \mathrm{VaR}_\alpha(X)\Big) \tag{5.16}$$

from this equation it is easily seen that $\mathrm{ES}_\alpha(X) \ge \mathrm{TCE}_\alpha$ in general. Here $\lambda$ is defined to be

$$\lambda = \frac{\mathbb{P}[X \le q^\alpha]}{\alpha} \tag{5.17}$$

From the definition of Expected Shortfall we can derive its acceptance set $\mathscr{A}_{\mathrm{ES}_\alpha}$.

$$\mathscr{A}_{\mathrm{ES}_\alpha} = \left\{ X \in \mathscr{X} : \frac{1}{\alpha} \int_0^\alpha F^{\leftarrow}(p)\, dp \geq 0 \right\} \tag{5.18}$$

From this representation we see that Expected Shortfall is a more conservative risk measure than Value-at-Risk, since $\mathscr{A}_{\mathrm{ES}_\alpha} \subset \mathscr{A}_{\mathrm{VaR}_\alpha}$.

Expected Shortfall is a coherent risk measure and thus by definition is has the property of sub-additivity. This is confirmed when we look at the simple discrete example 4.2.1 again, but this time we shall use Expected Shortfall as the risk measure.

**Example 5.1.2** *Again a bank loans Company A \$100.000,- and that this company will default on the loan with a probability of 0.8%. Suppose, furthermore, that Company A either defaults the entire loan, or not at all. Thus if we denote by X the default amount we get*

$$X = \begin{cases} -\$100.00,- & \textit{if Company A defaults} \\ \$0,- & \textit{otherwise} \end{cases} \tag{5.19}$$

*We have seen that the* $VaR_{0.01}(X) = 0$*. In order to compute* $ES_{0.01}(X)$ *we need the conditional distribution*

$$\tilde{X} = X | X \leq q^{(\alpha)} = X | X \leq 0 \tag{5.20}$$

*Thus* $\tilde{X}$ *has the following distribution*

$$\mathbb{P}\left[\tilde{X} = -\$100.000,-\right] = 0.008, \qquad \mathbb{P}\left[\tilde{X} = \$0,-\right] = 0.992 \tag{5.21}$$

*Thus we get the following for the tail conditional expectation*

$$TCE_{0.01}(X) = -\mathbb{E}\left[\tilde{X}\right] = 0.008 \cdot \$100.000,- = \$800,- \tag{5.22}$$

*Moreover we have*

$$\lambda = \frac{\mathbb{P}\left[X \leq \$0,-\right]}{0.01} = \frac{1}{0.01} = 100 \tag{5.23}$$

*And hence the Expected Shortfall for portfolio X is given by*

$$\begin{aligned} ES_{0.01}(X) &= TCE_{0.01}(X) + 99 \cdot \left(TCE_{0.01} - VaR_{0.01}(X)\right) \\ &= \$80.000,- \end{aligned} \tag{5.24}$$

*Note that this means that* $X \notin \mathscr{A}_{ES_{0.01}}$.

*If we now look at the diversified position, ie the bank loans \$50.000,- to Company A and another \$50.000 to Company B, where both companies have the same default probability of*

*0.08% and like in the first scenario both companies either default on their entire loan or not at all, then, if we denote the default amount in this case by $Y$, this yields*

$$Y = \begin{cases} -\$100.00,\text{-} & \textit{if Company A and Company B default} \\ -\$50.00,\text{-} & \textit{if Company A or Company B defaults} \\ 0 & \textit{otherwise} \end{cases} \tag{5.25}$$

*We know that in this case the Value-at-Risk of the position is given by*

$$VaR_{0.01}(Y) = -\inf\left\{y : \mathbb{P}[Y \leq y] > 0.01\right\} = \$50.000,\text{-} \tag{5.26}$$

*If we take $\tilde{Y}$ to be the tail conditional distribution of this position, thus $\tilde{Y} = Y|Y \leq -\$50.000,\text{-}$*

$$\begin{aligned} \mathbb{P}\left[\tilde{Y} = -\$100.000,\text{-}\right] &= 0.000064 \\ \mathbb{P}\left[\tilde{Y} = -\$50.000,\text{-}\right] &= 0.016 \\ \mathbb{P}\left[\tilde{Y} = \$0,\text{-}\right] &= 0 \end{aligned} \tag{5.27}$$

*Then we may easily compute the tail conditional expectation*

$$\begin{aligned} TCE_{0.01}(Y) &= -\mathbb{E}\left[\tilde{Y}\right] \\ &= 0.008 \cdot \$100.000,\text{-} + 0.016 \cdot \$50.000,\text{-} \\ &= \$1.600,\text{-} \end{aligned} \tag{5.28}$$

*In this case we have the following value for $\lambda$*

$$\lambda = \frac{\mathbb{P}\left[Y \leq -\$50.000,\text{-}\right]}{0.01} = \frac{0.016064}{0.01} = 1.6064 \tag{5.29}$$

*Resulting in an Expected Shortfall*

$$ES_{0.01}(Y) = -\$29.349, 76 \tag{5.30}$$

*Now note that in this case the risk is acceptable, $Y \in \mathscr{A}_{ES_{0.01}}$.*

□

## 5.2 Limitations of Expected Shortfall

Although ES has the key property of sub-additiviy that is lacking in VaR, it is not a mathematically flawless risk measure. It is easily checked that Expected Shortfall does not belong the class of elicitiable risk measures.

**Theorem 5.2.1** *Expected Shortfall is not elicitable relative to any class $\mathscr{F}$ of probability distributions.*

**Proof** Let $a, b, c, d$ be constants such that $a < b < c < \frac{1}{2}(b+d)$ and define the two (estimated) probability measures $F_1$ and $F_2$ of a position $X \in \mathscr{X}$.

$$F_1 = \alpha\delta_\alpha + \frac{1}{2}(1-\alpha)(\delta_b + \delta_b) \tag{5.31}$$

$$F_2 = \alpha\delta_c + (1-\alpha)\delta_{\frac{b+d}{2}} \tag{5.32}$$

It is easily seen that for $\alpha \leq \frac{1}{3}$ we have

$$ES^{\alpha}_{F_1}(X) = ES^{\alpha}_{F_2} = \frac{b+d}{2} \tag{5.33}$$

The result follows from the fact that in this case we don't have convex level sets.

$$ES^{\alpha}_{\frac{F_1+F_2}{2}}(X) = \frac{1}{4}(b+c+2d) > \frac{b+d}{2} \tag{5.34}$$

■

There is yet another shortcoming of expected shortfall. That is that the accurate estimation of the tail of the distribution is especially important for expected shortfall. However, this estimation is quite tricky. For example: the correlation among asset prices observed in normal market conditions is often very different from the correlation observed in extreme market conditions. Such a correlation breakdown would make it near to impossible for the risk manager to estimate the tail distribution with the conventional estimation methods such as Monte Carlo.

# 6 Expectile Value-at-Risk

In Cont et all [8] and in Kou et all [26] it is shown that there is a conflict between robust (and hence elicitable) risk measures and coherent risk measures. However, in [8] a more-restrictive distribution-based approach was chosen and not an axiomatic approach to risk measurement. And in [26], where an axiomatic approach is taken, the extra property of law-invariance is added to the set of axioms.

**Definition 6.0.2** *Distribution-based risk measures, $\mu$, also referred to as law-invariant risk measures, are such that*

$$\mu(X_1) = \mu(X_2) \tag{6.1}$$

*if $X_1$ and $X_2$ have the same distribution.*

As mentioned at the beginning, in this setting such a conflict does not arise and there exists a unique coherent, elicitable risk measure.

## 6.1 The Expectiles

Newey and Powell first introduced the $\tau$-expectile functional in 1987. This functional is defined as the unique solution to asymmetric least squares minimization.

**Definition 6.1.1** *For $X \in \mathscr{X}$ and $\tau \in (0,1)$, the $\tau$-expectile risk measure is the unique solution $\nu_\tau = \nu_\tau(X)$ to the following equation*

$$\tau \int_{\nu_\tau}^{\infty} \left(y - \nu_\tau\right) dX(y) = (1-\tau) \int_{-\infty}^{\nu_\tau} \left(y - \nu_\tau\right) dX(y) \tag{6.2}$$

Note that since it is assumed that for all $X \in \mathscr{X}$ the first order expectation exists and is finite, a unique solution to equation (6.2) exists.

In Kuan et all [28] it is shown that from equation (6.2) we may conclude that

$$\nu_\tau(X) = \gamma \mathbb{E}\left[X \,|\, \nu_\tau < X\right] + (1-\gamma)\mathbb{E}\left[X \,|\, X \leq \nu_\tau\right] \tag{6.3}$$

where

$$\gamma = \frac{\tau\Big(1 - F_X\big[\nu(\tau)\big]\Big)}{\tau\Big(1 - F_X\big[\nu_\tau\big]\Big) + (1-\tau)F_X\big[\nu_\tau\big]} \tag{6.4}$$

Which then shows that the $\tau$-expectile is the weighted sum, balance, between the conditional upside mean, $\mathbb{E}\big[X \,|\, \nu_\tau < X\big]$ and the conditional downside mean, $\mathbb{E}\big[X \,|\, X \leq \nu_\tau\big]$. This is different from Value-at-Risk and Expected Shortfall, which only take the downside loss into consideration and hence the $\tau$-expectile is less conservative.

## 6.2 Definition of Expectile Value-at-Risk

The risk measure based on the expectile functional is called Expectile Value-at-Risk and defined as follows.

**Definition 6.2.1** *Expectile Value-at-Risk is defined as*

$$EVaR_\tau(X) = -\nu_\tau(X) \tag{6.5}$$

*where $\nu_\tau(X)$ is the unique solution to the minimization problem of equation* (6.2).

If on the other hand we are looking at a sample rather than a distribution, the expectiles are quite similar to quantiles and the difference is that the sample expectiles are determined by tail expectations and not tail probabilities. For a given value of $\tau$ the sample expectile, $\tilde{\nu}_\tau$, is obtained by minimizing the function

$$\left|\tau - \mathbb{1}_{\{(x_i - \nu_\tau) < 0\}}\right| (x_i - \nu_\tau)^2 \qquad \text{for } 0 < \tau < 1 \tag{6.6}$$

It is shown in Bellini [5] that the acceptance set associated with Expectile Value-at-Risk is given by

$$\mathscr{A}_{\text{EVaR}} = \left\{ X \in \mathscr{X} : \frac{\mathbb{E}X^+}{\mathbb{E}X^-} \geq \frac{1-\tau}{\tau} \right\} \tag{6.7}$$

where $X^+ = \max\{X, 0\}$ and $X^- = \max\{-X, 0\}$. The similarity with VaR is also reflected by the acceptance sets, see equation (4.4). From the acceptance set it is clear that EVaR can be thought of as the expected gain-loss ratio which must exceed a certain threshold.

As mentioned at the beginning of this chapter a risk measure that is both coherent and elicitable will be defined. That EVaR is that risk measure is shown below.

**Theorem 6.2.2** *$EVaR_\tau(X)$ is a coherent risk measure.*

**Proof** Suppose that we a position with final net worth $X \geq 0$, then of course the $\tau$-expectile is positive and hence $\text{EVaR}_\tau(X) \leq 0$. Now to show that Expectile Value-at-Risk is translation invariant and positive homogeneous let $\tilde{X} = kX + a$, where $k \geq 0$. It is shown in Newey and Powell [32] that

$$\nu_\tau(\tilde{X}) = k\,\nu_\tau(X) + a \tag{6.8}$$

So that it follows that for the Expectile Value-at-Risk for $\tilde{X}$ it holds

$$\text{EVaR}_\tau\left(\tilde{X}\right) = \text{EVaR}_\tau(kX + a) = k\text{EVaR}_\tau(X) - a \tag{6.9}$$

If we have $X_1, X_2 \in \mathscr{X}$ such that $X_1 \leq X_2$ then evidently we have that $\nu_\tau(X_1) \leq \nu_\tau(X_2)$ and hence $\text{EVaR}(X_1) \geq \text{EVaR}(X_2)$, which proofs monotonicity.

To show that EVaR also satisfies the fourth property of a coherent risk measure, I follow Bellini [5] and rewrite the acceptance set $\mathscr{A}_\nu$

$$\mathscr{A}_\nu = \left\{X \in \mathscr{X} : \mathbb{E}[\delta X + (1-\delta)X^+] \geq 0\right\} \tag{6.10}$$

where $\delta = \dfrac{1-\tau}{\tau}$. It then follows that for $\tau \leq \dfrac{1}{2}$ the acceptance set is convex. By definitions 2.2.3 and 2.2.4 that this leads to a convex risk measure.

■

That $\text{EVaR}_\tau$ is also a elicitable risk measure, and hence a true risk measure in the sense that it satisfies the axioms 2.5-2.10 follows from the theorem below.

**Theorem 6.2.3** *EVaR is an elicitable risk measure relative to the class $\mathscr{F}$.*

**Proof** For the full proof I refer to Gneiting [19], but for completeness I shall give the main results. A scoring function $S : \mathbb{R}^2 \to [0,\infty)$ is weakly-robust for the $\tau$-expectile if and only if it is of the form

$$S(x,y) = \left|\mathbb{1}_{\{s\geq y\}} - \tau\right| \cdot \left(\phi(y) - \phi(x) - \phi'(x)(x-y)\right) \tag{6.11}$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function and $\phi'$ is its first derivative. Moreover, if $\phi$ is strictly convex such that

$$\mathbb{E}_F\phi(Y) < \infty, \quad \mathbb{E}_F\phi(Y) < \infty \tag{6.12}$$

exist and are finite for all $F \in \mathscr{F}$, then $S$ is robust for EVaR. Take $\phi$ to be defined by

$$\phi(y) = \frac{y^2}{1+|y|} \tag{6.13}$$

Note that in this case $\mathbb{E}\phi(Y)$ exists if and only $\mathbb{E}Y$ exists and if finite. An application of proposition 3.2.5 then gives the result.

$$S_{\text{EVaR}}(x,y) = \left|\mathbb{1}_{\{s\geq y\}} - \tau\right| \cdot \left[\frac{y^2}{1+|y|} - \frac{x^2}{1+|x|} - \left(\frac{2x}{1+|x|} - \frac{x^2}{(1+|x|)^2}\right)(x-y)\right] \tag{6.14}$$

■

Moreover, it is shown in Bellini [6] that $\text{EVaR}_\tau$ is the only risk measure that satisfies axioms 2.5-2.10.

# 7 Simulation Study: Random Foreign Exchange Portfolio

So far the mathematical and statistical properties of risk measures have been the main priority. This chapter is included to gain more insight in the practical implications of risk measures. On the basis of a randomly chosen foreign exchange portfolio the risk is calculated using Value-at-Risk, Expected Shortfall en Expectile Value-at-Risk.

## 7.1 Set-Up and Descriptive Statistics

To construct the random portfolio draw from a $U[-1,1]$ distribution to determine the total position taken in each currency. Where a draw of -1 stands for a short position of $100 mln in that currency, and likewise a draw of 1 means a position of $100 mln long in the corresponding currency. The results are shown in table 7.1 together with the equivalent positions in US Dollars.

The risk of this portfolio is calculated using VaR, ES and EVaR on December 23 2014.

| **Currency** | | **Position** | **USD Equivalence** |
|---|---|---|---|
| Euro | EUR | $-85$ mln | $-103,5$ mln |
| Britisch Pound | GBP | 41 mln | $63,6$ mln |
| Australian Dollar | AUD | $-32$ mln | $-25,9$ mln |
| Canadian Dollar | CAD | 74 mln | $63,7$ mln |
| Swiss Franc | CHF | 99 mln | $100,3$ mln |
| Japanese Yen | JPY | $-55$ mln | $-0,5$ mln |
| Chinese Yuan Renminbi | CNY | 40 mln | $6,4$ mln |
| Norwegian Kroner | NOK | $-13$ mln | $-1,8$ mln |
| Mexican Peso | MXN | $-22$ mln | $-1,5$ mln |

Table 7.1: *Weighted portfolio of nine currencies. Where the USD equivalence is based on the exchange rate on 23-12-2014, rounded to one decimal.*

In the case of foreign portfolio's the risk measurement procedure is quite simple and there are two main methods to do so. The first, and least complex, method is that of historical simulation. Within this method it is assumed that a bank's foreign exchange position will have the same distribution as it had in the past. The second is based on estimating a loss distribution. Here I shall analyse the risk measures by the second method, called LDA (Loss Distribution Approach). A time period needs to be set, for exchange-rate risk the typical so-called holding period is one day.

The descriptive statistics of the historical data can be found in tabel 7.2 and graphs 7.1 and 7.2 show the plots of the portfolio movements and a histogram of the verifying historical observations. In total there are 13.086 observations corresponding to the time period taken, that is for each of the nine currencies 1.455 observations. Any missing values, which accounted for less than 3%, were non-consecutive and were replaced by linear interpolation.


Figure 7.1: *Movements of the weighted portfolio, in USD · million, over the period 28-05-2009 through 30-10-2014.*

## Loss Distribution Approach

The compound loss distribution, $X$, of the foreign exchange rate position as in table 7.1 is the compound distribution of the loss arrival distribution and the loss/profit severity distribution. For the loss arrival distribution, $N$, I fitted a Bernoulli distribution, since

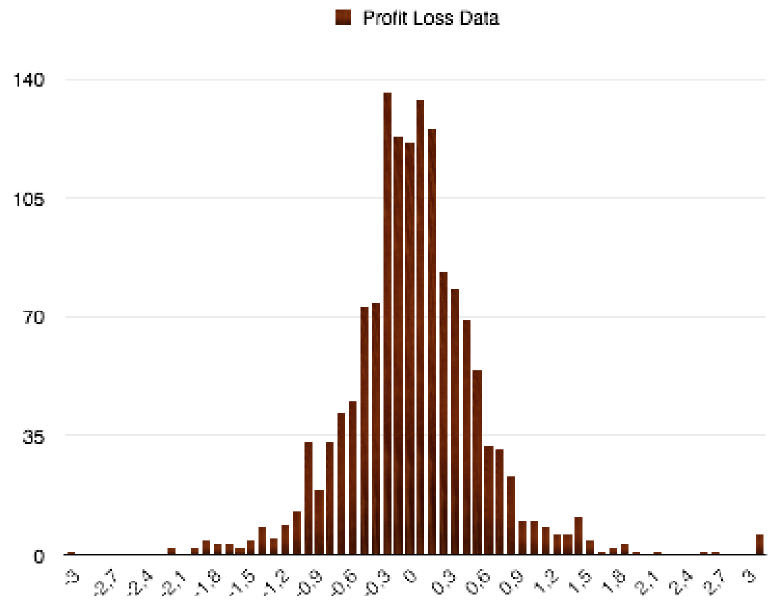



Figure 7.2: *Histogram of the verifying historical data over the period 28-05-2009 through 23-12-2014, in USD · million.*

| **Source** | **Federal Reserve Bank of St. Louis (Economic Research Division)** |
|---|---|
| **Frequency** | Daily |
| **Sample period** | May 28, 2009 - December 23, 2014 |
| **Sample size** | 1.454 observations |
| **Mean** | $ −0.014 mln |
| **Standard Deviation** | $ 0.62 mln |
| **Maximum** | $ 4.16 mln |
| **Minimum** | $ −5.29 mln |

Table 7.2: *Descriptive Statistics of the loss data.*

every new day can result in either a profit or a loss. For the loss/profit severity distribution, $L$ and $P$ respectively, I follow Mo and Zhou [31] and fit a two parameter Pareto distribution. The results are shown in table 7.3 and for the calculation of the maximum likelihood estimates of the Pareto distribution I refer to Appendix C.

Thus the compound loss distribution $X$ is given by

$$X_i = N_i \cdot P_i - (1 - N_i) \cdot L_i \tag{7.1}$$

| | | Parameter Distribution |
|---|---|---|
| **Loss Arrival Distribution** | $N \sim \text{Ber}(p)$ | $\hat{p} = 0,48$ |
| **Loss Severity Distribution** | $L \sim \text{Pareto}(\beta, k)$ | $\hat{\beta}_{\text{MLE}} = 0,482 \quad \hat{k}_{\text{MLE}} = 0,002$ |
| **Profit Severity Distribution** | $P \sim \text{Pareto}(\gamma, \ell)$ | $\hat{\gamma}_{\text{MLE}} = 0,297 \quad \hat{\ell}_{\text{MLE}} = 0,001$ |

Table 7.3: *Estimated distributions of the loss arrival process and the profit/loss severity processes.*

The compound loss distribution does not have a closed form expression, as is the mostly the case in finance. And hence $X$ is approximated by a Monte Carlo simulation, following Zhou [31] the minimal simulations required is 100 000.

## 7.2 Results

From the results, see table 7.4, it is directly evident that Expected Shortfall is the most conservative risk measure and Expectile Value-at-Risk is the least conservative, as was also seen from the accpetance sets. Given a pre-specified probability of 0.1%, the minimum amount that should be added to the position is \$ 3,44 mln when using Expected Shortfall opposed to \$2,27 mln when using Expectile Value-at-Risk. It is not hard to imagine that a difference in the minimal required extra capital of \$1,17 mln would make it difficult to switch from Expectile Value-at-Risk to Expected Shortfall for banks.

It is also obvious that the differences in risk values gets larger when the pre-specified probabilities, $\alpha$ and $\tau$, get smaller. Which is also what is expected since the risk measures are based on tail probabilities and tail expectations.

| $\alpha$ | 5 % | 1 % | 0,1 % |
|---|---|---|---|
| **Value-at-Risk** | 1,11 mln | 1,87 mln | 3,01 mln |
| **Expected Shortfall** | 1,59 mln | 2,35 mln | 3,44 mln |
| $\tau$ | 5 % | 1 % | 0,1 % |
| **Expectile Value-at Risk** | 0,84 mln | 1,39 mln | 2,27 mln |

Table 7.4: *Results based on Monte Carlo simulation (100 000 simulations).*

# Part III

# The Basel Accords

# 8 The Basel Committee

## 8.1 Disruptions Leading to the Establishment of the Basel Committee

**1973** The Bretton Woods System is an international basis for exchanging one currency for another and was established in 1944. The system itself collapsed in 1971, when President Nixon severed the link between the US dollar and gold. By 1973 most of the major world economies had allowed their currencies to float freely against the US dollar, which led to bank failures amongst others.

**June 1974** When West Germany's Federal Banking Supervisory Office withdrew the license of Bankhaus Herstatt in June of 1974, because the bank's foreign exchange exposures amounted to three times is own capital, foreign banks started a race on the remaining assets. As a result of liquidation they took heavy losses on their unsettled trades which made the debacle not only a German one, gave it an international dimension.

**October 1974** Just months later Franklin National Bank of New York also had to close its doors. Upto that time Franklin National Bank had always been one of the most profitable banks in the United States, but in 1972, when Michele Sindona (an Italian financier with suspected connections to the mafia and Vatican Banking) bought controlling interest in Franklin National Bank. Sindona, in an attempt to recuperate, led the bank into huge foreign exchange losses which caused the bank to close in 1974.

In response to the these disruptions in the international financial markets, as well as the ones stated in the preface, the central bank governors of 11 countries established a Committee on Banking Regulations and Supervisory Practices. This was later renamed

| | | | |
|---|---|---|---|
| **1** | Belgium | **7** | The Netherlands |
| **2** | Canada | **8** | Sweden |
| **3** | France | **9** | Switzerland |
| **4** | Germany | **10** | The United Kingdom |
| **5** | Italy | **11** | The United States of America |
| **6** | Japan | | |

Table 8.1: Members of the G-10.

as the Basel Committee on Banking Supervision.

The 11 countries who joined in on the establishment of the Basel Committee on Banking Supervision (BCBS), are known as the *Group of Ten* or G-10 for short. The G-10 have agreed to participate in the General Arrangement to Borrow (GAB), which was founded in 1962. It seems a bit odd that a group of eleven countries are referred as the G-10, the reason for this is that the eleventh member, Switzerland, didn't join the association until 1964 and the name of the association remained unchanged. The full list of the members can be found in tabel 8.1.

In 2009 the BCBS expanded their memberships and nowadays it includes 27 jurisdictions, a list of which can be found in appendix C.1. The committee now reports an oversight body: the Group of Central Bank Governors and Heads of Supervision (GHOS).

## 8.2 The Aim of the Basel Committee

The BCBS was designed as a forum on banking supervisory matters amongst the member countries. Its aim was, and still is, to enhance financial stability by increased supervisory knowhow and better supervision of banks worldwide. The BCBS wants to achieve this aim by setting minimum standards. These aims where conducted on three area's:

- ⋆ Improving the effectiveness of techniques for supervising international banking business.
- ⋆ Exchanging information on national supervisory arrangements.
- ⋆ Engage with challenges presented by diversified financial conglomerates.

From the beginning, one of the most important aims of the committee was to close the gaps that existed in international supervisory coverage to establish that

(1) No foreign banking organ would escape supervision,

(2) Supervision would be adequate and sufficient across al the 27 members.

## 8.3 The Future of The Basel Committee

Even before the implementation of Basel III, there was already rumour that the groundwork for Basel IV was being set ([14] [16]). Moreover, the implementation of Basel III was not planned until 2019. This points out that the Basel Committee shall keep on updating the regulatory frameworks to fit the changing financial world.

In the 84$^{th}$ annual report the BIS stated that one of the key initiatives of the Basel Committee is to further examine the balance between simplicity, comparability and risk sensitivity in the regulatory framework. This has a great deal to do with the risk measures that is used. Since there has been a lot of debate about the limitations of VaR, which the Basel Committee at first set aside but now say are highly relevant, presumably the Committee will take these limitations into account for Basel IV.

# 9 Risk Measures of the Basel Accords

The risk measures used in the Basel Accord are used for setting capital requirements for the banking- and trading books of financial institutions. The risk measures of the Basel Accords lead to important regulations, naturally there is a lot of debate about which risk measure should be used.

## 9.1 Basel I: the Basel Capital Accord

In the 1988 Basel Accord the rule was that the capital charge on commercial loans should be a uniform 8% of the loan face value. This rule was only risk sensitive in the sense that certain countries, banks and classes of loans. The foundations of this percentage were not mathematical, but rather seemed about right on average.

After Value-at-Risk made its appearance in the 1990's, the percentage rule of 1988 was no longer the benchmark for risk measurement (see [24]). Gordy [21] showed that under five assumptions on the different exposures within a position the rule of Basel I is asymptotically equivalent to $\text{VaR}_{0.01}$.

For what I could find there is no direct link between the Quantile Risk Measure as it was first used in actuarial science and the method of measurement used in the first Basel Accord. Seemingly these two measures where used separately.

This Basel Accord was valid upto 2004, hence VaR had been the standard for risk measurement for almost 20 years. Understandably this led to situation that it was very hard to change this standard, since all people working with it were so used to the measure by this time.

## 9.2 Basel II: the New Capital Framework

Basel II was initially published in 2004, in order to create an international standard for risk regulation. By that time there had already been done a large amount of discussion about the risk measure VaR. Expected Shortfall became the number one candidate to replace VaR, because it is a coherent risk measure.

However, the Basel Committee decided not to take Expected Shortfall since backtesting is of the utmost importance in practice and ES is not elicitable. Instead the risk measure used in Basel II is the so-called VaR - With Situation Analysis. This risk measure specifies that the capital charge for the trading book on a particular day $t$ for banks using the internal model approach should be calculated by the formula

$$\mathrm{II}_t(X) := \max\left\{\mathrm{VaR}_\alpha(X_{t-1}), \frac{k}{60}\cdot\sum_{i=1}^{60}\mathrm{VaR}_\alpha(X_{t-i})\right\} \tag{9.1}$$

where $k \geq 3$ is a constant. It is easily seen that the measure $\mathrm{II}_t(X)$ as stated above is not a true risk measure in the scope of this thesis.

**Theorem 9.2.1** *The measure used in the Basel II is not a true risk measure, since it is not a coherent risk measure.*

**Proof** The proof is trivial, since it is the maximum of two Value-at-Risk measures. And it was shown in chapter 4 that VaR is not sub-additive.

■

Even though the Basel II risk measure is not a true risk measure since it lacks one of the properties, note that it remains an elicitable risk measure based on proposition 3.2.5 and the quantile functional.

## 9.3 Basel III: the Liquidity Coverage Ratio & Liquidity Risk Monitoring Tools

The banking crisis which started in 2007/2008 made it politically difficult to implement Basel II and therefore the negiotiations for Basel III started early on. Again the rumour had it that Value-at-Risk would be replaced by the sub-additive measure Expected Shortfall. And, again, at the last moment it was decided to cling to VaR with Situation Analysis. They did however introduce a new version of this measure because there was debate on the sensitivity to procyclicality with the risk measure of Basel II.

$$\text{III}_t(X) := \max\left\{\text{VaR}_\alpha(X_{t-1}), \frac{k}{60}\sum_{i=1}^{60}\text{VaR}_\alpha(X_{t-i})\right\} + \max\left\{s\text{VaR}_\alpha(X_{t-1}), \frac{\ell}{60}\sum_{i=1}^{60}s\text{VaR}_\alpha(X_{t-i})\right\} \tag{9.2}$$

Where $s\text{VaR}_\alpha$ is called stressed VaR. This is the Value-at-Risk risk measure computed under the scenario that the financial market is under stress, as was the case during the recent financial crisis. It was added to compensate for procyclicality. Due to this extra term, this risk measure is mostly referred to as a stress test.

**Theorem 9.3.1** *The Basel III risk measure is not a true risk measure since it is neither coherent nor elicitable.*

**Proof** Since VaR is not sub-additive then clearly the Basel III risk measure is not sub-additive since it is the sum of two VaRs. Moreover the sum of two distinct quantiles is in general not robust.

The following example illustrates that in general the sum of two distinct quantiles, like VaR and sVaR, do not have convex level sets.

Take $\alpha = 0.01$, I will supress the $\alpha$ in the notation from this point on. Suppose the position $X \in \mathscr{X}$ is estimated by the two discrete distributions $F_1, F_2 \in \mathscr{F}$ such that

$$\mathbb{P}_1 = \Big[X = -\$100.000,\text{-}\Big] = 0,008 \qquad \mathbb{P}_1 = \Big[X = 0\Big] = 0,992 \tag{9.3}$$

$$\mathbb{P}_1^S\Big[X = -\$100.000,\text{-}\Big] = 0,05 \qquad \mathbb{P}_1^S\Big[X = 0\Big] = 0,95 \tag{9.4}$$

where $\mathbb{P}_1$ stands for the the standard distribution under $F_1$ and $\mathbb{P}_1^S$ denotes the estimated distribution for a stressful market under $F_1$. Likewise we have

$$\mathbb{P}_2 = \Big[X = -\$100.000,\text{-}\Big] = 0,000064 \qquad \mathbb{P}_2 = \Big[X = -\$50.000,\text{-}\Big] = 0,03$$

$$\mathbb{P}_2 = \Big[X = 0\Big] = 0,969936 \tag{9.5}$$

$$\mathbb{P}_2^S\Big[X = -\$100.000,\text{-}\Big] = 0,0025 \qquad \mathbb{P}_2^S = \Big[X = -\$50.000,\text{-}\Big] = 0,1$$

$$\mathbb{P}_2^S\Big[X = 0\Big] = 0,9875 \tag{9.6}$$

Then the mixture distribution $F_\lambda \in \mathscr{F}$, for $\lambda = \frac{2}{3}$ is given by

$$\mathbb{P}_\lambda = \Big[X = -\$100.000,\text{-}\Big] = 0,0061 \qquad \mathbb{P}_\lambda = \Big[X = -\$50.000,\text{-}\Big] = 0,01$$

$$\mathbb{P}_\lambda = \Big[X = 0\Big] = 0,9839 \tag{9.7}$$

$$\mathbb{P}_{\lambda}^{S}\Big[X = -\$100.000,\text{-}\Big] = 0,0338 \qquad \mathbb{P}_{\lambda}^{S} = \Big[X = -\$50.000,\text{-}\Big] = 0.0067$$
$$\mathbb{P}_{\lambda}^{S}\Big[X = 0\Big] = 0,9595 \tag{9.8}$$

If we assume that

$$\begin{aligned} \mathrm{VaR}_{\alpha}(X_{t-1}) &= \max\left\{\mathrm{VaR}_{\alpha}(X_{t-1}), \frac{k}{60}\sum_{i=1}^{60}\mathrm{VaR}_{\alpha}(X_{t-i})\right\} \\ s\mathrm{VaR}_{\alpha}(X_{t-1}) &= \max\left\{s\mathrm{VaR}_{\alpha}(X_{t-1}), \frac{\ell}{60}\sum_{i=1}^{60}s\mathrm{VaR}_{\alpha}(X_{t-i})\right\} \end{aligned} \tag{9.9}$$

Then in this situation we have $\mathrm{III}_{F_1}(X) = \mathrm{III}_{F_2}(X) = \$100.000,\text{-}$. And

$$\begin{aligned} \mathrm{III}_{F_\lambda}(X) &= \mathrm{VaR}_{F_\lambda}(X) + \mathrm{sVaR}_{F_\lambda}(X) \\ &= \$50.000,\text{-} + \$100.000,\text{-} \end{aligned} \tag{9.10}$$

And we see that there is a $\lambda \in [0,1]$, such that we don't have convex level sets

$$\mathrm{III}_{F_1}(X) = \mathrm{III}_{F_2}(X) \neq \mathrm{III}_{F_\lambda}(X) \tag{9.11}$$

■

# 10 Natural Risk Statistics

As stated before, one of the four main goals of the Basel Committee is to further investigate the simplicity, comparabiltiy and sensitivity in the regulatory framework. In order to say anything about booked success in this area I shall make a comparison with the theory. In order to do so in this axiomatic setting of risk measures, the risk measures introduced by the Basel Committee should also be fitted into such a setting. This is exactly what Kou, Peng and Heyde [26] have done, they have set up an axiomatic approach to define a class that includes the Basel Risk Measures.

## 10.1 Axiomatic Approach to Natural Risk Statistics

The risk measure of the Basel III Accords discussed previously do not fit into a class of measures treated, however Kou, Peng and Heyde [26] formulated a set of axioms which classifies this risk statistics, the so-called natural risk statistics. Not only the Basel III risk measure belongs to this class, but the Basel II risk measure does as well.

**Definition 10.1.1** *A risk statistic is a mapping* $\hat{\mu} : \mathbb{R}^n \to \mathbb{R}$.

A risk statistic is a data-based risk measure. Where a risk measure uses a measurable function $X$ to define risk, a risk statistic uses $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ to represent $X$.

**Definition 10.1.2** $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ *and* $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ *are scenario-wise comonotonic if for all* $1 \leq j, k \leq n$ *it holds that*

$$\left(x_j - x_k\right)\left(y_j - y_k\right) \geq 0 \tag{10.1}$$

If $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ represent the random losses of position $X$ and $Y$ respectively, then if $\mathbf{x}$ and $\mathbf{y}$ are scenario-wise comonotonic means that $X$ and $Y$ move in the same direction.

**Axiom 10.1** *Positive homogeneity and translation scaling:*

$$\hat{\mu}\left(k\mathbf{x} + a\right) = k \cdot \hat{\mu}(\mathbf{x}) + a \tag{10.2}$$

*for all* $\mathbf{x} \in \mathbb{R}^n$, $k \geq 0$, $a \in \mathbb{R}$.

**Axiom 10.2** *Monotonicity:*

$$\hat{\mu}(\mathbf{x}) \leq \hat{\mu}(\mathbf{y}), \quad \textit{if} \quad \mathbf{x} \leq \mathbf{y} \tag{10.3}$$

*where* $\mathbf{x} \leq \mathbf{y}$ *means that* $x_i \leq y_i$ *for* $i = 1, \ldots, n$.

Note that these first two axioms are the counterparts of the axioms 2.5 and 2.7 of monetary risk measures.

**Axiom 10.3** *Scenario-wise comonotonic sub-additivity:*

$$\hat{\mu}(\mathbf{x} + \mathbf{y}) \leq \hat{\mu}(\mathbf{x}) + \hat{\mu}(\mathbf{y}) \tag{10.4}$$

*for any* $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ *that are scenario-wise comonotonic.*

**Axiom 10.4** *Empirical law-invariance: for any permutation* $(p^{i,1}, \ldots, p^{i,n_i})$ *of* $(1, 2, \ldots, n)$ *and* $i = 1, \ldots, m$ *we have*

$$\hat{\mu}(x_1, \ldots, x_m) = \hat{\mu}(x_1^{p_1,1}, \ldots, x_1^{p_1,n_1}, x_2^{p_2,1}, \ldots, x_2^{p_2,n_2}, \ldots, x_m^{p_m,1}, \ldots, x_m^{p_m,n_m}) \tag{10.5}$$

This last axiom is the risk statistic counterpart of law invariance, see definition 6.0.2.

That the risk measure of the Basel II Accord given by equation (9.1) and the risk measure of the Basel III Accord given by equation (9.2) are special cases of the class of natural risk measures I refer to [26].

## 10.2 Risk Measures vs Risk Statistics

Kou, Peng and Heyde [26] have three main arguments on why they use risk statistics rather than risk measures.

(1) Risk statistics directly measure risk from the available dataset, which greatly reduces model misspecification errors since they do not require specifying subjective models.

(2) Risk statistics can include data subsets generated by models based on forward-looking views or prior knowledge.

(3) Risk statistics can incorporate multiple prior probabilities which reflect multiple beliefs about the probabilities of occurrence of different scenario's.

| $\alpha$ | 5 % | 1 % | 0,1 % |
|---|---|---|---|
| **Value-at-Risk** | 0,88 mln | 1,57 mln | 4,82 mln |
| **Expected Shortfall** | 1,90 mln | 2,66 mln | 5,14 mln |
| $\tau$ | 5 % | 1 % | 0,1 % |
| **Expectile Value-at Risk** | 0,71 mln | 1,36 mln | 3,34 mln |

Table 10.1: *Results based on historical data simulation over the period May 28, 2009 through December 23, 2014.*

To make a comparison between risk measures and risk statistics I also calculated the three main risk measures of this thesis from a risk statistics point of view, hence I calculated them based solely on the dataset. The results are displayed in table 10.1.

From the results it appears that risk statistics are more sensitive to the tail probabilities and tail expectations. With smaller pre-specified probabilities, the differences with the loss distribution approach get more significant. Which is not surprising since by definition risk statistics depend solely on the data en hence the are sensitive to changes. Also there is a limited dataset that may be used in the analysis. The near history gives a better picture of the present situation than the distant history. Using a limited dataset makes that the outliers have a larger contribution to the results.

In my opinion using risk statistics rather than risk measures could be a great practical advantage when dealing with internal risk management, since they can include different beliefs and forward-looking views of banks. However, including subjective beliefs and/or forward views increases the chances of model misspecification and nullifies the main advantage. From an external (regulatory) point of view, risk statistics are sensitive to the 'choice' of the data set and hence in this case the preferred method would be risk measures.

## 10.3 True Risk Measures Axioms vs Natural Risk Statistic Axioms

Including convex level sets to the axioms of risk measures was one of the main objectives of this thesis. A risk measure should belong to both the class of coherent measures and the elicitable risk measures. Risk statistics are not distribution based and therefore are sensitive to changes in the data set. The axioms for natural risk statistics thus do not take both steps of the risk measurement procedure into account.

An another major difference between the axioms of true risk measures and those formulated by Kou, Peng and Heyde is scenario-wise comonotonicity. As mentioned in chapter 2 I do not think that the comonotonicity property is a suitable mathematical representation in practice. See example 2.2.2. The example illustrates that the assumption that the future net worth of different risky financial instruments should be perfect substitutes is incorrect. The example also applies to scenario-wise comonotonicity.

If you do accept (scenario-wise) comonotonic sub-additivity as an axiom for risk measures, than of course the axiomatic approach taken by Kou, Peng and Heyde has an advantage over coherent risk measures. The relaxation of the sub-additivity property to scenario-wise comonotonic sub-additivity then makes the class of risk measures far less restrictive. But from a regulatory point of view, coherent risk measures are less suitable than elicitable risk measures. And thus I believe that for the risk measures of the Basel Accord elicitability is property of interest.

The axiom on empirical law invariance should take into account the estimation step within the risk procedure. Since it is data based, there is not estimation done. If you consider the distribution based equivalent of empirical law invariance, law invariance. It was shown that including this axiom causes a conflict to arise between coherent risk measures and elicitability.

In my opinion, the best way to approach the measurement of risky positions remains the axiomatic approach of Artzner et all. extended with convex level sets which is necessary for the exclusion of non-elicitable risk measures.

# Part IV

# Overview

# 11 Summary

The procedure of measuring risky positions consists of two equally important parts.

(1) Estimating the loss distribution of the position.

(2) Defining a risk measure that summarizes the risk of the position.

Roughly speaking, the first part of the procedure is represented by the class of elicitable risk measures, while the second part is dealt with by the class of coherent risk measure. I therefore argue that a *true* risk measure should be both coherent and elicitable.

The most widely known, and used, risk measure is the so-called Value-at-Risk measure. VaR is defined as minus the $\alpha$-quantile of the loss distribution;

$$\mathrm{VaR}_\alpha(X) = -q^{(\alpha)}(X) \tag{11.1}$$

where $q^{(\alpha)}(X)$ is the largest $\alpha$-quantile. Since it is based on the quantile functional VaR is an elicitable risk measure but not a coherent risk measure. It lacks the property of sub-additivity. Another shortcoming of Value-at-Risk is that it doesn't take into account the severity of the risk beyond the VaR-value.

After the article of Artzner et all [3] *Thinking Coherently*, the most likely candidate to replace VaR as a practical monetary risk measure was the coherent risk measure Expected Shortfall.

$$\mathrm{ES}_\alpha(X) = -\frac{1}{\alpha}\Big(\mathbb{E}\Big[X\mathbb{1}_{\{X\leq q^{(\alpha)}\}}\Big] - q^{(\alpha)}\Big(\mathbb{P}\Big[X\leq q^{(\alpha)}\Big]-\alpha\Big)\Big) \tag{11.2}$$

Expected Shortfall also has it limitations, the most important being that it is not an elicitable monetary risk measure. Due to the lack of this property, the Basel Committee decided not to use Expected Shortfall as the coherent alternative to VaR. That ES is not elicitable can easily be shown by a simple counterexample which shows that this measure does not have convex level sets, a necessary condition for elicitability. Another shortcoming of Expected Shortfall is that it is highly sensitive to the tail loss distribution.

It has been thought for some time that there exists a conflict between elicitability and the coherent properties. However, Gneiting (amongst others) found that there is indeed a monetary risk measure that is both coherent and elicitable. This is the risk measure known as Expectile Value-at-Risk defined by

$$\mathrm{EVaR}_\tau(X) = -\nu_\tau(X) \tag{11.3}$$

where $\nu_\tau(X)$ is the unique solution tot the minimization problem of the following equation,

$$\tau \int_{\nu_\tau}^{\infty} \left(y - \nu_\tau\right) dX(y) = (1-\tau) \int_{-\infty}^{\nu_\tau} \left(y - \nu_\tau\right) dX(y) \tag{11.4}$$

Risk measures are of practical importance since they are used in regulatory frameworks such as the Basel Accords. The first Basel Accord dates back to 1988. This at first was not a real risk measure but rather a percentage of a position's face value. However since the introduction of VaR within risk management of banks it quickly became the standard. Moreover, it was shown by Gordy that the percentage of the first Basel Accord is asymptotically equivalent to the 99.9% Value-at-Risk.

Regardless of the comments and limitations the Basel Committee continued using VaR as the standard risk measure in the second and third framework. Below the risk measure of Basel II and Basel III, denoted by $\mathrm{II}_t(X)$ and $\mathrm{III}_t(X)$ respectively.

$$\mathrm{II}_t(X) := \max\left\{\mathrm{VaR}_\alpha(X_{t-1}), \frac{k}{60} \cdot \sum_{i=1}^{60} \mathrm{VaR}_\alpha(X_{t-i})\right\} \tag{11.5}$$

$$\mathrm{III}_t(X) := \max\left\{\mathrm{VaR}_\alpha(X_{t-1}), \frac{k}{60} \sum_{i=1}^{60} \mathrm{VaR}_\alpha(X_{t-i})\right\} + \max\left\{s\mathrm{VaR}_\alpha(X_{t-1}), \frac{\ell}{60} \sum_{i=1}^{60} s\mathrm{VaR}_\alpha(X_{t-i})\right\} \tag{11.6}$$

It is widely known that the risk measures used by the BCBS are not coherent, since they use VaR. Moreover, I showed in this thesis that the risk measure used in the third Basel Accord is neither coherent nor elicitable, because it is the sum of two VaR's.

# 12 Conclusion

Based on the axiomatic approach used in this thesis, I conclude that a true risk measure which takes into account both the desired mathematical and statical properties should belong the class of coherent risk measures as well as the class of elicitable risk measures. And hence, for as far is known at this moment, is a risk measure based on the expectile functional.

Based on the mathematical properties Expected Shortfall to me is the least suitable risk measure. From a regulatory point of view it is important that a risk measure is elicitable. Moreover, in my conversations with risk managers and risk modellers I found that it is preferrable that a risk measure isn't conservative. The reason being that a conservative risk measure leads to more complex situations in practice because of the additional restrictions on the required capital.

Like Expected Shortfall the risk measure Expectile Value-at-Risk also takes the tail loss distribution into consideration, however is it weighted against the tail profits. This is a more natural manner of looking at risky positions. In most cases positions are entered according to the risk appetite of banks, the amount and type of risk a bank is willing to take in order to meet their objectives. And hence earnings should also be included in the decision making process.

From a mathematical point of view the Basel Committee made a measured decision when they formulated the measure of the second Basel Accord and they didn't replace Var with Expected Shortfall. The realization of the risk measure in the third Basel Accord cannot be explained by the mathematics of this thesis. The measure $\mathrm{III}_t(X)$ does not belong to the class of coherent risk measures nor is it elicitable. Moreover, the axioms that do classify this risk measure are not a valid reflection of the underlying economic meaning in my opinion.

Hence, if I were to rank the risk measures, from most favorable to least favorable, the list would be as the one below.

1. Expectile Value-at-Risk
2. Risk measure of the Basel II Accord
3. Value-at-Risk
4. Expected Shortfall
5. Risk measure of the Basel II Accord

From an implementation point of view it would not be too problematic for banks to switch from VaR tot Expectile VaR, since the latter is the lesser conservative risk measure. Unfortunately I found that Expectile Value-at-Risk is virtually unkown in practice. The only downside to EVaR that I have found so far is that the expectiles are not as simple to use and easy to understand as quantiles are. Most of the practioners are not mathematicians and a proper understanding of risk measures is necessary in order to avoid (costly) mistakes.

# 13 Recommendations

It still remains an open question in risk theory whether it is possible to make a classification for elicitable risk measures. Following Gneiting [19] and Bellini [5] for this thesis I tried to prove that the property of convex level sets is not only necessary for elicitable risk measure, but it is also sufficient. However, I was not able to prove this hypothesis, or find a counterexample.

**Hypothesis 13.0.1** *If a monetary risk measure, $\nu : \mathscr{X} \to \mathbb{R}$ has convex level sets in the following way, let $F_1, F_2 \in \mathscr{F}$, such that $\nu_{F_1}(X) = \nu_{F_2}(X)$. Define*

$$\mathscr{F} \ni F_\lambda = \lambda \cdot F_1 + (1-\lambda) \cdot F_2 \tag{13.1}$$

*then for all $\lambda \in [0,1]$ it holds that*

$$\nu_{\lambda F_1 + (1-\lambda) F_2}(X) = \nu_{F_1}(X) = \nu_{F_2}(X) \tag{13.2}$$

*Then $\nu$ is an elicitable monetary risk measure.*

Based on the article of Gneiting [19] I find that it is plausible to assume that the hypothesis is correct, in this article he shows that besides quantiles and expectiles, als expectations and ratio's of expectations (all of which have convex level sets) are elicitable. Moreover he shows that the mode is asymptotically elicitable, whether or not the mode is also elicitable in the 'normal' sense he wasn't able to show at that time.

The fact that finding a scoring function that makes the mode elicitable is already a difficult task may indicate that the hypothesis is not true. The mode is a relatively simple statistical functional and much more complex functionals, with the property of convex level sets, could be constructed. However, if the hypothesis is true, then the proof will be very analytical since it then depends on finding a scoring function that makes an arbitrary monetary risk measure with convex level sets elicitable. But if it is possible to do so, a counterpart of theorem 2.2.4 exists.

**Theorem 13.0.2** *If a set $\mathscr{E}$ satisfies axioms 2.1-2.4 then the associated risk measure $\nu_{\mathscr{E}}$ satisfies properties* **??**-*3.24 of elicitable monetary risk measures. Moreover, we have that the closure of $\mathscr{E}$ is the associated acceptance set, ie $\bar{\mathscr{E}} = \mathscr{A}_\nu$.*

**Proof** The axioms 2.5 through 2.8 have already been covered in the previous section. What is left is to proof that the risk measure $\nu_{\mathscr{E}}$ satisfies property 3.24 from definition 3.2.4.

Suppose $X_1, X_2 \in \mathscr{E}$ are such that $\nu_{\mathscr{E}}(X_1) = \nu_{\mathscr{E}}(X_2)$, then

$$\begin{aligned} \nu_{\mathscr{E}}(X_1) &= \inf\Big\{a_1 : X_1 + a_1 \in \mathscr{E}\Big\} \\ &= \inf\Big\{a_2 : X_2 + a_2 \in \mathscr{E}\Big\} \\ &= \nu_{\mathscr{E}}(X_2) \end{aligned} \tag{13.3}$$

By the definition of $\nu_{\mathscr{E}}$ and the properties of the infimum we get

$$\begin{aligned} \nu_{\mathscr{E}}\Big(\lambda X_1 + (1-\lambda)X_2\Big) &= \inf\Big\{a : \lambda X_1 + (1-\lambda)X_2 + a \in \mathscr{E}\Big\} \\ &\geq \lambda \cdot \inf\Big\{a_1 : X_1 + a_1 \in \mathscr{E}\Big\} + (1-\lambda)\cdot \inf\Big\{a_2 : X_2 + a_2 \in \mathscr{E}\Big\} \\ &= \lambda \nu_{\mathscr{E}}(X_1) + (1-\lambda)\nu_{\mathscr{E}}(X_2) \\ &= \nu_{\mathscr{E}}(X_1) \quad = \quad \nu_{\mathscr{E}}(X_2) \end{aligned} \tag{13.4}$$

On the other hand by axiom 2.4 and 2.3, if

$$X_1 + a_1 \in \mathscr{E} \quad \text{and} \quad X_2 + a_2 \in \mathscr{E} \tag{13.5}$$

Then by convexity, for $\lambda \in [0.1]$

$$\lambda\Big(X_1 + a_1\Big) + (1-\lambda)\Big(X_2 + a_2\Big) \in \mathscr{E} \tag{13.6}$$

By definition of the infimum, for given $\varepsilon > 0$, there exist $x_1 \in \{a_1 : X_1 + a_1 \in \mathscr{E}\}$ and $x_2 \in \{a_2 : X_2 + a_2 \in \mathscr{E}\}$ such that

$$\begin{aligned} x_1 &< \inf\{a_1 : X_1 + a_1 \in \mathscr{E}\} + \frac{\varepsilon}{2} \\ x_2 &< \inf\{a_2 : X_2 + a_2 \in \mathscr{E}\} + \frac{\varepsilon}{2} \end{aligned} \tag{13.7}$$

Hence, for all $\varepsilon > 0$ there exists a summation $x_1 + x_2$ such that

$$x_1 + x_2 < \inf\{a_1 : X_1 + a_1 \in \mathscr{E}\} + \inf\{a_2 : X_2 + a_2 \in \mathscr{E}\} + \varepsilon \tag{13.8}$$

Showing that, together with positive homogeinity, for all $\lambda \in [0, 1]$

$$\begin{aligned} \nu_{\mathscr{E}}\Big(\lambda \cdot X_1 + (1-\lambda)\cdot X_2\Big) &\leq \lambda \cdot \nu_{\mathscr{E}}(X_1) + (1-\lambda)\cdot \nu_{\mathscr{E}}(X_2) \\ &= \nu_{\mathscr{E}}(X_1) \quad = \quad \nu_{\mathscr{E}}(X_2) \end{aligned} \tag{13.9}$$

And we conclude equality.

■

The main advantage of characterizing elicitable monetary risk measures is that you can construct a risk measure that is elicitable bottom-up, as in the case of Expected Shortfall. And hence *any* characterization of the class of elicitable risk measures, and not just by convex level sets, would a great contribution to risk theory.

# Bibliography

[1] C. Acerbi and D. Tasche. Expected Shortfall: A Natural Coherent Alternative to Value-at-Risk. *Economic Notes*, (31):379–388, 2002.

[2] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Thinking Coherently. Risk, 10(11):68–71, 1997.

[3] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent Measures of Risk. Mathematical Finance, 9(3):203–228, July 1999.

[4] Basel Committee on Banking Supervision. A Brief History of the Basel Committee. Technical report, Bank For International Settlements, 2013.

[5] F. Bellini, editor. Some Remarks on Elicitable Risk Measures, Generalized Quantiles and Expectiles. Parmenides Foundation, June 2014.

[6] F. Bellini and V. Bignozzi. Elicitable Risk Measures. Quantitative Finance, 2013. Available at SSRN 2334746.

[7] V. Chavez-Demoulin, A.C. Davison, and A.J.McNeil. Estimating value-at-risk: a point process approach. Quantitative Finance, 5(2):227–234, 2005.

[8] R. Cont, R. Deguest, and G. Scandolo. Robustness and sensitivity analysis of risk measurement procedures. Quantitative Finance, 10(6):593–606, June/July 2010.

[9] T. Copeland, T. Koller, and J. Murrin. Valuation. Measuring and Managing the Value of Companies. John Wiley & Sons, 2nd edition, 1995.

[10] G. de Rossi and A. Harvey. Quantiles, Expectiles and Splines. Faculty of Economics, Cambrige University, February 2007.

[11] F. Delbaen. Monetary utility functions. ETH Zurich, 2007.

[12] D. Duffie and J. Pan. An overview of value-at-risk. The Journal of Derivatives, 4(3):7–49, 1997.

[13] I. Ekeland, A. Galichon, and M. Henry. Comonotonic measures of multivariate risks. Mathematical Finance, 22(1):109–132, 2012.

[14] P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj. An academic response to basel 3.5. Risks, 2(1):25–48, February 2014.

[15] S. Emmer, M. Kratz, and D. Tasche. What is the best risk measure in practice? a comparison of standard measures. Working paper ESSEC Research Center., 2013.

[16] J. Anderson et all. Basel 4: Emerging from the Mist. Technical report, KPMG Financial Services, September 2013.

[17] H. Föllmer and T. Knispel. Convex Risk Measures: Basic Facts, Law-Invariance and Beyond, Asymptotics for Large Portfolio's, chapter Part II, pages 507–554. World Scientific, 2013.

[18] J.B.G Frenk and G. Kassay. Introduction to Convex and Quasiconvex Analysis. Erasmus Research Institute of Management, August 2004.

[19] T. Gneiting. Making and evaluating point forecasts. Journal of the American Statistical Association, (106):746–762, 2001.

[20] T. Gneiting. Quantiles as optimal point predictors. Technical Report 538, Department of Statistics University of Washington, August 2008. Available at http://www.stat.washington.edu/research/reports/2008/tr538.pdf.

[21] M.B. Gordy. A risk-factor model foundation for ratings-based bank capital rules. Financial Intermed, (12):199–232, 2003.

[22] F.R. Hampel. A general qualitative definition of robustness. The Annals of Mathematical Statistics, 42(6):1887–1896, 1971.

[23] R. Henrion. Some remarks on value-at-risk optimization. International Journal of Management Science and Engineering Management, 1(2):111–118, October 2006.

[24] P. Jorion. Value-at-Risk: The New Benchmark for Managing Financial Risk. McGraw & Hill, 3rd edition, 2007.

[25] J. Keppo, L. Kofman, and X. Meng. Unintended consequences of the market risk requirement in banking regulation. Working paper.

[26] S. Kou, X. Peng, and C.C. Heyde. External risk measures and basel accords. Mathematics of Operations Research, 38(3):393–417, March 2013.

[27] V. Krätschmer, A. Schied, and H. Zähle. Comparitative and qualitative robustness for law-invariant risk measures. Finance and Stochastics, 18(2):271–295, January 2014.

[28] C.M. Kuan, J.H. Yeh, and J.C. Hsu. Assessing Value-at-Risk with CARE, the Conditional Autoregressive Expectile Models. Journal of Econometrics, 2(150):261–270, 2009.

[29] N.S. Lambert. Elicitation and evaluation of statistical forecasts. Working paper Stanford University, July 2013.

[30] H. Markowitz. Portfolio Selection. The Journal of Finance, 7(1):77–91, March 1952.

[31] J-M Mo and Z.-F. Zhou. Optimal Selection of Loss Severity Distribution Based on LDA. Networked Computing and Advanced Information Management, 2:570–574, September 2008.

[32] W.K. Newey and J.L. Powell. Assymetric Least Squares Estimation and Testing. Econometrica, 55(4):819–847, July 1987.

[33] K. Osband and S. Reichelstein. Information Eliciting Compensation Schemes. Journal of Public Economics, (27):107–115, 1985.

[34] R.T. Rockafeller. Convex Analysis. Number 28 in Princeton landmarks in mathematics and physics. Princeton University Press, 1997.

[35] D. Tasche, editor. Risk Measures: Another Search for the Holy Grail. Financial Services Authority, Isaac Newton Institute for Mathematical Sciences, March 2013.

[36] Y. Yamai and T. Yoshiba. On the validity of value-at-risk: Compararive analysis with expected shortfall. Monetary and Economic Studies, 20(1):57–86, 2002. Bank of Japan.

# Part V

# Appendices

# A Appendix

## A.1 Quasi-Convexity

A function $f : A \to \mathbb{R}$ defined on a convex subset $C \subset A$ is said to be quasi-convex if for all $x, y \in C$ and $\lambda \in [0,1]$ it holds that

$$f\Big(\lambda x + (1-\lambda)y\Big) \leq \max\Big\{f(x), f(y)\Big\} \tag{A.1}$$

Note that all convex functions are quasi-convex, but not all quasi-convex functions are convex. And hence quasi-convexity can be seen as the generalization of convexity.

## A.2 Comonotonicity

The measurable function $X_1$ and $X_2$ are said to be comonotonic if and only if

$$\forall \omega_1, \omega_2 \in \Omega \quad \Rightarrow \quad \Big(X_1(\omega_1) - X_1(\omega_2)\Big)\Big(X_2(\omega_1) - X_2(\omega_2)\Big) \geq 0 \tag{A.2}$$

## A.3 MLE for Pareto Distribution

Recall that the probability density function of the two parameter Pareto distribution is given by

$$f(x|\beta, k) = \frac{\beta \cdot k^{\beta}}{x^{\beta+1}} \tag{A.3}$$

where $\beta, k > 0$ and $k \leq x$. The likelihood function, $\mathcal{L}$, then has the form

$$\mathcal{L}(\beta, k|x) = \prod_{i=1}^{N} \frac{\beta \cdot k^{\beta}}{x_i^{\beta+1}} \tag{A.4}$$

Without any further calculations we see that the maximum likelihood estimator for $k$ must be

$$\hat{k}_{\text{MLE}} = \min\{x_i\} \tag{A.5}$$

since $\mathcal{L}$ get large when $k$ increases but the restriction on $k$ is that it isn't larger than the smallest sample value. Note that the likelihood function is non-negative and hence it is easier to look at the log-likelihood function.

$$\log(\mathcal{L}) = N \cdot \log\beta + N \cdot \beta \log k - (\beta - 1)\sum_{i=1}^{N} \log x_i \tag{A.6}$$

Taking the partial derivative to $\beta$ and setting it equal to zero yields

$$\frac{\partial \log(\mathcal{L})}{\partial \beta} = \frac{N}{\beta} + N \cdot \log k - \sum_{i=1}^{N} \log x_i \tag{A.7}$$

And hence, replacing the parameter k by its maximum likelihood estimator, we arrive at the maximum likelihood estimate

$$\hat{\beta}_{\text{MLE}} = \frac{N}{\sum_{i=1}^{N} \log\left[\frac{x_i}{\hat{k}_{\text{MLE}}}\right]} \tag{A.8}$$

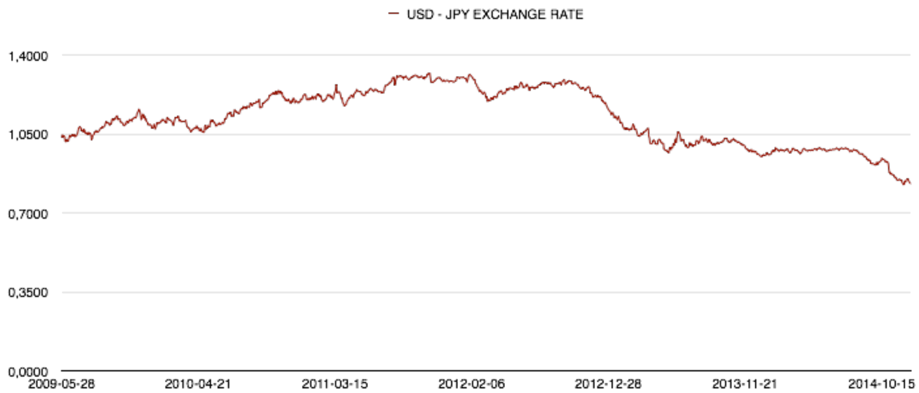
# Appendix B

## B.1 Currency Graphs

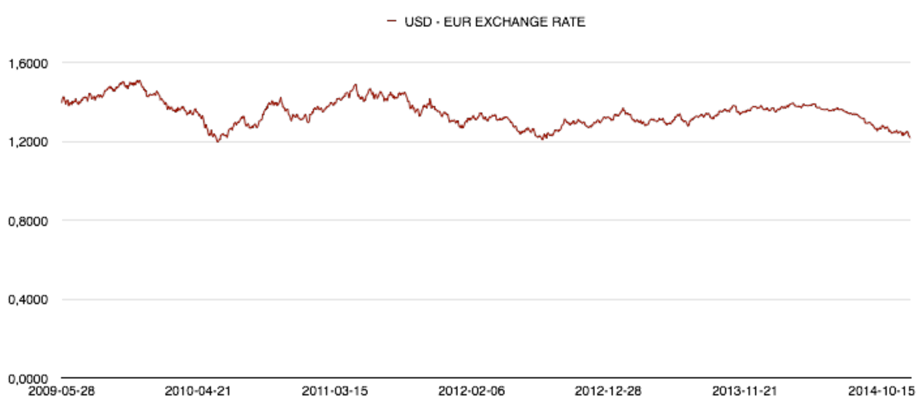
The data of the graphs in this appendix are taken from the Federal Reserve Bank of St. Louis, Economic Research Department. All graphs run from June 28 2013 through October 30 2014.

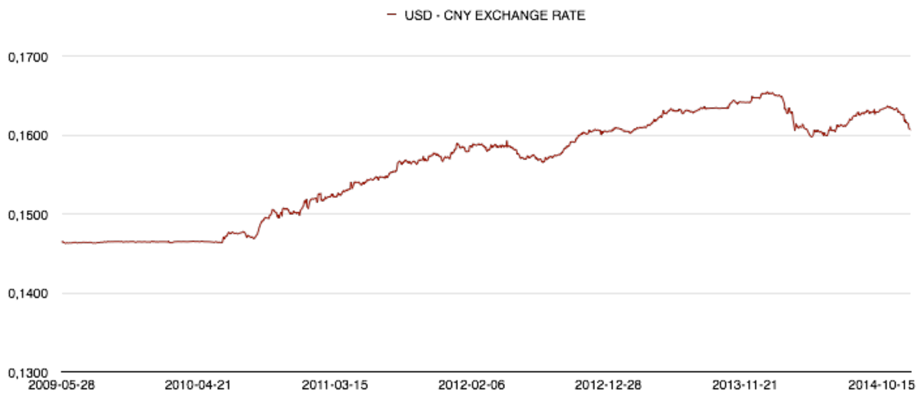


USD - MXN EXCHANGE RATE
0,9000
0,6750
0,4500
0,2250
0,0000
2009-05-28
2010-04-21
2011-03-15
2012-02-06
2012-12-28
2013-11-21
2014-10-15

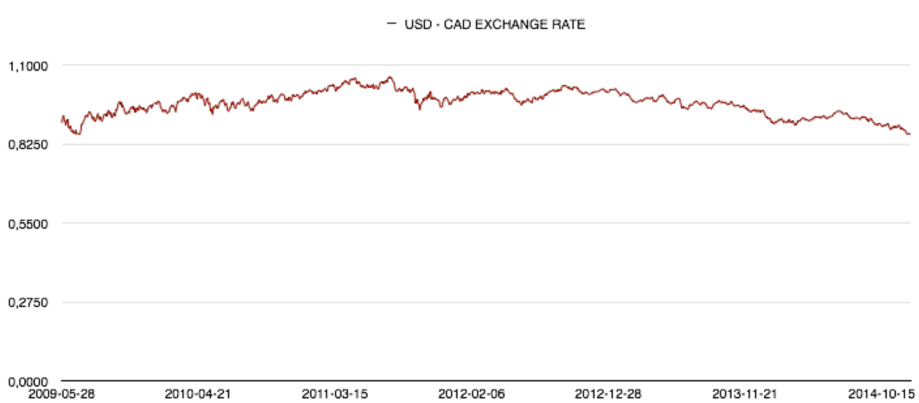

USD - JPY EXCHANGE RATE
1,4000
1,0500
0,7000
0,3500
0,0000
2009-05-28
2010-04-21
2011-03-15
2012-02-06
2012-12-28
2013-11-21
2014-10-15


USD - GBP EXCHANGE RATE
1,8000
1,6750
1,5500
1,4250
1,3000
2009-05-28
2010-04-21
2011-03-15
2012-02-06
2012-12-28
2013-11-21
2014-10-15


USD - EUR EXCHANGE RATE
1,6000
1,2000
0,8000
0,4000
0,0000
2009-05-28
2010-04-21
2011-03-15
2012-02-06
2012-12-28
2013-11-21
2014-10-15


USD - CNY EXCHANGE RATE
0,1700
0,1600
0,1500
0,1400
0,1300
2009-05-28
2010-04-21
2011-03-15
2012-02-06
2012-12-28
2013-11-21
2014-10-15


USD - CHF EXCHANGE RATE
1,4000
1,0500
0,7000
0,3500
0,0000
2009-05-28
2010-04-21
2011-03-15
2012-02-06
2012-12-28
2013-11-21
2014-10-15


USD - CAD EXCHANGE RATE
1,1000
0,8250
0,5500
0,2750
0,0000
2009-05-28
2010-04-21
2011-03-15
2012-02-06
2012-12-28
2013-11-21
2014-10-15

# C Appendix

## C.1 List of All Jurisdictions Included in the BCBS

| | | | |
|---|---|---|---|
| 1 | Argentina | 14 | Luxembourg |
| 2 | Australia | 15 | Mexico |
| 3 | Belgium | 16 | The Netherlands |
| 4 | Brazil | 17 | Russia |
| 5 | Canada | 18 | Saudi Arabia |
| 6 | China | 19 | Singapore |
| 7 | France | 20 | South Africa |
| 8 | Germany | 21 | Spain |
| 9 | Honk Kong SAR | 22 | Sweden |
| 10 | India | 23 | Switzerland |
| 11 | Indonesia | 24 | Turkey |
| 12 | Italy | 25 | The United Kingdom |
| 13 | Japan | 26 | The United States of America |
| 14 | Korea | | |

Table C.1: List of all jurisdictions included in the BCBS.