Various Faces of Risk Measures: Internal Model's Perspective

Min Wang
Åbo Akademi University*
and
China University of Geosciences†
E-mail: mwang@abo.fi

Lasse Koskinen
FIN–FSA‡
and
HSE§
E-mail: Lasse.Koskinen@bof.fi

March 20, 2009

Abstract

In this article, we consider several aspects of risk measures from the internal models' perspective. We critically review the most widely used classes of risk measures. Especially, we attempt to clear up some of the most commonly misconstructed aspects: the choice between risk measures, and practical data and forecasting issues, like the importance of robustness. As a new result, solvency capital requirement is optimized under variance premium principle. The use of tail conditional trimmed mean is proposed as a robust risk estimator. One objective of this article is to emphasize that one single risk measure or a specific axiomatic system is not appropriate for all purposes.

Keywords: Forecasting Risk, Outlier, Robustness, Solvency II

*Mathematics Department, Åbo Akademi University, Turku 20500, Finland
†Mathematics Department, China University of Geosciences, Wuhan 430074, China
‡Finnish Financial Supervisory Authority, P.O. Box 103, FIN-00101 Helsinki, Finland
§Helsinki School of Economics, P.O. Box 1210, FIN-00101 Helsinki, Finland

1 Introduction

An early attempt to extend actuarial methods to full scale internal modelling is given in Pentikäinen (1975)[28]: *"The conventional applications of the theory of risk concern many important sides of insurance business, e.g. evaluating the stability, estimating a suitable level for maximum net retention, safety loading or the funds. Whether or not the applications have been useful for practical management may have depend very much on how the risk theoretical treatments have been linked with the complexity of various other aspects involved with the actual decision making... Our purpose is to attack just this problem and to endeavour to build up a picture of the management process of the insurance business in its entirely (as far as possible).*"

In this paper we shall focus on risk measures from the internal models' viewpoint. The authors have been inspired by professor Teivo Pentikäinen's lifetime work on the modelling of insurance business. This article is a little attempt in the same direction.

The Groupe Consultatif Glossary on insurance terms defines the internal model as "Risk management system of an insurer for the analysis of the overall risk situation of the insurance undertaking, to quantify risks and/or to determine the capital requirement on the basis of the company specific risk profile."

Risk measurement and management has been thrust into the forefront of issues facing insurers and regulators alike. Supervisory authorities require insurers to practice risk management, and report risk measures to them. In addition to publicly mandated risk measuring, insurers measure and manage risk internally. Insurance company's approach to risk management ranges from a minimum compliance to regulations to comprehensive internal models.

Pentikäinen (ed.) (1982) [29], and Daykin, Pentikäinen and Pesonen (1994)[8] argued that the solvency issue consists of numerous sub-problems such as model building, and studies concerning variations in risk exposure (cyclical and other kinds of variation), catastrophic risks and other kinds of risk categories jeopardizing solvency.

One special challenge is to measure risk and assess profitable areas of business. Numerous risk measures and their implementation procedures are used in insurance and finance. Actuaries are left with the challenging task of having to choose a suitable one. Fundamentally, all statistical risk measuring techniques fall into one of three categories: a) fully parametric methods based

on modelling the entire distribution of returns; b) the nonparametric method of historical simulation; and c) parametric modelling of the tails of the return distribution. Another important classification is between static and dynamic risk measures.

All these methods have pros and cons, generally ranging from easy and inaccurate to difficult and, in theory, more precise. No method is perfect, and usually the choice of a technique depends on the application in question, and the availability of actuary to implement the procedures.

As the financial system becomes more complex, the need for complicated statistical models to measure risk and to price assets becomes greater. Unfortunately, the reliability of such models decreases with complexity, so when we need the models the most they tend to be least reliable. Indeed, the credit crunch, which started in the summer of 2007, shows that risk models are of somewhat lower quality than was generally believed (Danielsson (2008)[5]).

Further, Danielsson (2008)[5] claims that this does not suggest that statistical models should not be employed. On the contrary, they play a fundamental role in the internal risk management operations of financial institutions. The main problem is unrealistic expectations of what models can do. The recent financial crises and disastrous losses have increased the demand for reliable quantitative risk management tools.

When outliers, like crashes in financial market, are present, it is essential to accommodate them into risk model with e.g. a heavy-tailed distribution, to adjust them in some way or to use robust estimation method that is not affected too much by departures from the model assumptions. When one only cares about large movements in some random variable, it may not be optimal to model the entire distribution of the event with all available data. Instead, it may be better only to model the tails with tail events. However, this can not be done without cost. There is much larger risk for misspecified model since the heaviness of tail distribution is a controversial subject.

One important issue is the objective of choosing a risk measure. Heyde et al. (2007)[19] posed a question: Is the risk measure proposed for the interest of a firms equity shareholders, regulator or manager? They conclude by arguing that there does not exist a unique risk measure that fits the needs of these different parties.

In this paper we first review the main axiomatic systems, then we consider the importance of robustnes in risk measurement. Tail conditional trimmed mean is proposed as a compromise between value at risk and conditional tail expectation. Secondly, we consider risk measures from the regulatory per-

spective and optimize solvency capital requirement under variance premium principle. Final section concludes.

2 Theory of Risk Measure

Pentikäinen (ed.) (1982)[29]: *"The above reasoning was intended to confirm the fact, well-known by insurance experts, that there is no single indicator which can adequately describe the solvency of any insurer."* We think that this holds true especially for any risk measure.

2.1 Examples of Risk Measures

From Dhaene et al. (2008) [12], we consider a set Γ of real-valued random variables defined on a given measurable space $(\Omega, F, \mathcal{P})$. We will assume that $X, Y \in \Gamma$ implies that $X + Y \in \Gamma$, and also $aX \in \Gamma$ for any $a > 0$ and $X + b \in \Gamma$ for any real b. A functional

$$\rho : \Gamma \to \mathbb{R},$$

mapping every element of a loss (or profit) distribution in Γ to the real numbers, is called a *risk measure* (with domain Γ). The risk measure is assumed in some way to encapsulate the risk associated with a loss distribution.

We will interpret Ω as the set of all possible states of nature at the end of some fixed reference period, for instance one year. The set Γ will be interpreted as the extended set of financial losses under consideration at the end of the reference period, related to insurance and investment portfolios that a particular regulatory authority controls.

Concrete examples of risk measures are value at risk, tail conditional expectation and variance premium principle. The value at risk (VaR) is the minimum amount of losses on a trading portfolio over a given period of time with a certain probability:

$$\mathrm{VaR}_\alpha[X] = Q_\alpha[X],\ \alpha \in (0, 1). \tag{1}$$

where Q_α is α-quantile. The tail conditional expectation (TCE) is the conditional expected loss given that the loss exceeds its VaR:

$$\mathrm{TCE}_\alpha[X] = mean[X|X > Q_\alpha[X]],\ \alpha \in (0, 1). \tag{2}$$

The variance premium principle (VPP), determined by the expectation and variance for a given risk:

$$\mathrm{VPP}_\beta[X] = E[X] + \beta Var[X],\ \beta > 0. \tag{3}$$

2.2 Coherent and Convex Risk Measures

Here we briefly review the main axiomatic systems for risk measures and comment on their use. There are two main families of risk measures suggested in the literature. First, the coherent risk measures suggested by Artzner et al. [3], and its relaxed version, the convex risk measures proposed by Föllmer and Schied (2002) [14] and also by Frittelli and Gianin (2002) [15]. Second, the insurance risk measures in Wang et al. (1997) [34]. Each of these axiomatic systems represent different schools of thought. More extensive literature review can be found in the excellent textbook of Denuit et al. (2005)[11].

In Artzner et al. (1999) [1], a risk measure ρ is called a coherent risk measure, if it satisfies the following axioms: *monotonicity, positive homogeneity, translation invariance and subadditivity.* It is good to note that no set of axioms is universally accepted. Modifying axioms may lead to other 'coherent' risk measure.

Axiom 1 *Monotonicity: for any X and $Y \in \Gamma$ with $X \leq Y$, we have $\rho[X] \leq \rho[Y]$.*

This rules out the risk measure, $VPP_\beta[X] = E[X] + \beta Var[X],\ \beta > 0$.

Axiom 2 *Positive homogeneity: for any $\lambda > 0$ and $X \in \Gamma$, $\rho[\lambda X] = \lambda\rho[X]$.*

If position size directly influences risk (for example, if positions are large enough that the time required to liquidate them depends on their sizes) then we should consider the consequences of a lack of liquidity when computing the future net worth of a position.

Axiom 3 *Translation invariance: for any $X \in \Gamma$ and all real numbers b, we have $\rho[X + b] = \rho[X] + b$.*

This says that a sure loss of amount b simply increases the risk by b and it is an axiom for accounting-based risk measures. For many external risk measures, such as a margin deposit, the accounting-based risk measures seem to be reasonable. For internal risk measures, attitude-based measures may be preferred.

Axiom 4 *Subadditivity: for all X and $Y \in \Gamma$, $\rho[X+Y] \leq \rho[X] + \rho[Y]$.*

The subadditivity property, which reflects the diversification of portfolios (*see* Meyers (2000) [26]), or that 'a merger does not create extra risk,' is a natural requirement. This rules out the risk measure, $VaR_\alpha[X] = Q_\alpha[X],\ \alpha \in (0,1)$.

The Axiom *Subadditivity* implies that $\rho[nX] \leq n\rho[X]$ for $n = 1, 2, \ldots$. In the Axiom *Positive homogeneity* we have required equality for all positive $\lambda > 0$ to model what a government or an exchange might impose in a situation where no netting or diversification occurs, in particular because the government does not prevent many firms to all take the same position.

Artzner et al. (1999) [1] pointed out that Huber (1981) [20] showed that if Ω has a finite number of elements and Γ is the set of all real random variables, then a risk measure ρ is coherent if and only if there exists a family $\mathcal{Q}$ of probability measures on Ω, such that

$$\rho[X] = \sup_{Q \in \mathcal{Q}} E_Q[X], \qquad \forall X \in \Gamma,$$

where $E_Q[X]$ is the expectation of X under the probability measure $\mathcal{Q}$. Delbaen (2002) [9] extended the above result when Ω has infinite number of elements.

Getting a coherent risk measure amounts to computing the maximal expectation under different scenarios (different Q's), thus justifying scenarios analysis used in practice. Artzner et al. (1999) [1] and Delbaen (2002) [9] also presented an equivalent approach of defining the coherent risk measure through the specification of acceptance sets, the sets of financial positions accepted by regulators or investors.

Hence, to get a coherent risk measure, one first chooses a set of scenarios (different probability measures), and then computes the coherent risk measure as the maximal expectation of the loss under these scenarios.

Convex risk measures were proposed by Föllmer and Schied (2002) [14] and also by Frittelli and Gianin (2002) [15] independently. In many situations the risk of a position might increase in a nonlinear way with the size of the position. For example, an additional liquidity risk may arise if a position is multiplied by a large factor. This suggests relaxing the conditions of positive homogeneity and of subadditivity and to require the weaker property of

Axiom 5 *Convexity: $\rho[\lambda X + (1-\lambda)Y] \leq \lambda\rho[X] + (1-\lambda)\rho[Y]$, for any $X, Y \in \Gamma, \lambda \in [0,1]$.*

Convexity means that diversification does not increase the risk, i.e., the risk of a diversified position $\lambda X + (1-\lambda)Y$ is less or equal to the weighted average of the individual risks.

Definition 1 *A map $\rho : \Gamma \to \mathbb{R}$ will be called a* ***convex measure*** *of risk if it satisfies the axioms of convexity, monotonicity, and translation invariance.*

2.3 Insurance Risk Measure

Insurance risk premiums can also be viewed as risk measures, as they aim at using one number to summarize future random losses. Following Wang et al. (1997) [34], let Γ be the set of non-negative random variables which represent the random losses associated with insurance contracts. Thus, market premiums can be described as a functional ρ from the set of insurance risks to the extended non-negative real numbers. A risk measure ρ is said to be an insurance risk measure if it satisfies the following four axioms.

Axiom 6 *Conditional state independence: $\rho[X] = \rho[Y]$, if X and Y have the same distribution.*

This means that for a given market condition, the insurance risk of a position is determined only by the loss distribution.

Axiom 7 *Monotonicity: for any X and $Y \in \Gamma$ with $X \leq Y$ a.s., we have $\rho[X] \leq \rho[Y]$.*

This axiom is the same with the one in coherent risk measure. It is reasonable because if a risk Y results in a larger insurance claim than risk X in a.s. every state of nature, then the insurance premium for Y should be greater than the premium for X.

Definition 2 *Let X and Y be random variables belonging to Γ. We say that X and Y are comonotonic if and only if the inequality*

$$[X(\omega_1) - X(\omega_2)][Y(\omega_1) - Y(\omega_2)] \geq 0$$

holds a.s. for ω_1 and $\omega_2 \in \Omega$.

Axiom 8 *Comonotonic additivity: If X and Y are comonotonic, $\rho[X+Y] = \rho[X] + \rho[Y]$.*

If X and Y are comonotonic, the outcomes of X and Y always move to the same direction (good or bad) as the state ω changes. For example, quota shares or excess-of-loss layers of the same risk are comonotonic. Thus there is no hedge. In Wang et al. (1997) [34], the axiom is based on the argument that the comonotonic random variables do not hedge against each other, leading to the additivity of the risks. However, Heyde et al. (2007)[19] thought that this is only true if one focuses on one scenario. Indeed, if one has multiple scenarios, a counterexample showed that comonotonic additivity fails to hold.

Axiom 9 *Continuity: For $X \in \Gamma$ and $d > 0$, the functional ρ satisfies $\lim_{d\to 0^+} \rho[(X-d)_+] = \rho[X]$ and $\lim_{d\to\infty} \rho[\min(X,d)] = \rho[X]$.*

In this axiom, the first condition says that a small truncation in the loss variable results in a small change in the premium; the second condition says that ρ can be calculated by approximating X by bounded variables. The continuity property has also been applied by Hurlimann (1994, Theorem 4.1) [21].

Consider a set function $\gamma : \mathcal{A} \to [0,\infty)$. Assume that the set function γ is finite, zero on the empty set, and monotone; that is, $\gamma[\varnothing] = 0, \gamma[\Omega] < \infty$, and if $A, B \in \mathcal{A}$ and $A \subset B$, then $\gamma[A] \leq \gamma[B]$.

For a non-negative, real-valued random variable X, the Choquet integral[1] of X with respect to γ can be evaluated as

$$\int X d\gamma = \int_0^\infty \gamma\left\{\omega : X(\omega) > t\right\} dt.$$

Definition 3 *Let P be a probability measure on a σ-algebra $\mathcal{A}$ in 2^Ω. For an increasing function g on [0, 1] with $g(0) = 0$ and $g(1) = 1$, the set function $\gamma = g \circ P$ is called a* ***distorted*** *probability and the function g is called a* ***distortion function****.*

[1]Let S be a set, and let $\mathcal{F}$ be any collection of subsets of S. Consider a function $f : S \to \mathbb{R}$ and a monotone set function $\nu : \mathcal{F} \to \mathbb{R}^+$. Assume that f is measurable with respect to ν, that is $\forall x \in \mathbb{R}$: $\{s|f(s) \geq x\} \in \mathcal{F}$. Then the Choquet integral of f with respect to ν is defined by:

$$\int f d\nu := \int_{-\infty}^0 (\nu(\{s|f(s) \geq x\}) - \nu(S))\, dx + \int_0^\infty \nu(\{s|f(s) \geq x\})\, dx$$

where the integrals on the right-hand side are the usual Riemann integral (the integrands are integrable because they are monotone in x). The detailed discussion of Choquet integration can be found in Denneberg (1994) [10].

Under the distorted probability $\gamma = g \circ P$, the Choquet integral of $X \in \Gamma$ can be evaluated as

$$\int X d\gamma = \int_0^\infty g \circ P \{\omega : X(\omega) > t\} \, dt = \int_0^\infty g(1 - F_X(t)) dt.$$

Wang et al. (1997) [34] proved that if Γ contains all the $Bernoulli(p)$ random variables, $0 \leq p \leq 1$, then risk measure ρ satisfies Axioms of Insurance Risk Measures and $\rho(1) = 1$ if and only if ρ has a Choquet integral representation with respect to a distorted probability.

Hence, to get an insurance risk measure, one fixes a distorted probability, and then computes the insurance risk measure as the expectation with respect to one distorted probability (only one scenario).

3 Need of Robustness

Experiences suggest that a simple model is more often robust than a complicated one. Beard, Pentikäinen and Pesonen (1969) [2] wrote: "*Furthermore various alternatives, lines and methods of presentation are possible. Our aim has been for simplicity.*"

3.1 Tail Estimation and Model Selection

In the estimation (calibration) step of internal models the role of data is crucial. The estimation results can be very sensitive to the chosen data set. Because the calibration target in Solvency II is a very rare event (once in 200 years), the available data set is usually too small, which means that the confidence intervals for the estimated parameters are large. Hence, the data problem is more severe in Solvency II than it is in Basel II. This makes great demand on the robustness of the used risk measure, i.e. small departure from model assumptions should have relative little effect on the value. Further, in many modelling areas the data have to be supplemented with additional assumptions.

In risk measurement the tail of the risk distribution is in focus. The standard error of quantile estimator typically increases when one goes further in the tail of the loss distribution (see e.g. Karuse (2003)[24]). When estimating 99.5 % quantiles for internal model this is a major problem in practice. The

first compliance arises from the fact that by definition only few observations are made at tails, and hence estimation errors are large.

Second problem is that the behaviour of quantile estimates is also critically influenced by the tails of the distributions. The distinctions between different distributions is exceedingly difficult for popular methods and particularly large samples are needed for clear discrimination. Using heavy tailed distributions and/or copulas makes the data issue even more critical. In the tail modelling subjectivity is unavoidable and this must be appropriately dealt with by the insurers and supervisors. For more information on these data problems, see e.g. Koskinen et al. (2009)[23].

3.2 Outliers

In practice other important practical problem is that with extensive data there will almost always be important unanticipated features. There may be isolated, very discrepant, observations whose inclusion or exclusion needs special consideration. These outliers are the most important observations. Statistics derived from data sets after wrong decision (eliminate or not) may be totally misleading.

One particular type of model robustness is robustness against outliers. Danielsson (2008)[5] underlines the importance of outliers for risk measurement. He argues that the quality of statistical risk models is much lower than often assumed. Such models are useful for measuring the risk of frequent small events but not for systemically important events.

It is clear that one area of actuarial attention should be the determination and investigation of the sources of outliers. There is a section, Outlier Detection in the Encyclopedia of Actuarial Science (2004) [32]. From there, in the statistics analysis of actuarial data, one often finds observations that 'appear to be inconsistent with the remainder of that set of data'. We can call these kinds of observations 'outliers'. They may have a large effect on the statistical analysis of the data and they can cause misleading results if only standard statistical procedures are used.

A possible approach to circumventing this problem is the use of robust statistical methods. These allow for statistical dispersion parameters, even when outliers themselves are the most interesting part of the data. Examples are an unexpectedly high claim against an insurance company or a surprisingly high (or low) stock return for a financial institution. Any statistical analysis of data should therefore include the identification of possible outliers.

An outlier is a observation that is unlikely under the assumed model distribution. To formalize this idea, Davies and Gather (1993) [7] introduced the concept of α-outliers. For example, if $P = N(\mu, \sigma^2)$ then

$$out(\alpha, N(\mu, \sigma^2)) = \{x \in \mathbb{R} : |x - \mu| > \sigma z_{1-\alpha/2}\},$$

which is just the union of the lower and the upper $\alpha/2$-tail region. Here, $z_{1-\alpha/2}$ denotes the $1 - \alpha/2$-quantile of the standard normal distribution. If $P = \exp(\lambda)$, an exponential distribution with scale parameter λ, then

$$out(\alpha, \exp(\lambda)) = \{x \in \mathbb{R} : x > -\lambda \ln \alpha\},$$

which is the upper α-tail region. Each point located in $out(\alpha, P)$ is called an α-outlier with respect to P, otherwise it is called an $\alpha-$inlier.

We can now formulate the task of outlier identification within the framework of α-outlier regions as follows. For a given sample, $x_n = (x_1, \ldots, x_n)$ of size n, which contains at least $[n/2] + 1$ i.i.d. observations coming from some distribution P, we have to find all those x_i that are location in $out(\alpha, P)$. The level α can be chosen depending on the sample size. If for some $\alpha^{'} \in (0.5, 1)$, we take

$$\alpha = \alpha_n = 1 - (1 - \alpha^{'})^{1/n},$$

then the probability of falsely detecting any outlier in a sample of size n coming i.i.d. from P is not larger that $\alpha^{'}$. As we assume only $P \in \mathcal{P}$ for some family of distributions $\mathcal{P}$ is usually unknown as well.

Since outliers may contain very important information, they should be investigated carefully. Robust algorithms may help not only to avoid distortion of the output, but to also determine outliers, which reflect unusual behavior and for which further investigation is necessary. However, the origin of some outliers is just data error, and these outliers are usually thrown away.

3.3 Compromise between Tail Dependence and Robustness

Heyde et al. (2007)[19] suggested the use of tail conditional median (TCM):

$$TCM_\alpha[X] = median[X|X > Q_\alpha[X]],\ \alpha \in (0, 1). \tag{4}$$

as a robust risk measure. However, the TCM is not a fundamentally new concept since TCM_α is equal to $Q_{\frac{1+\alpha}{2}}$ theoretically and the only differences come from the chosen estimator:

1. if X is continuous then

$$TCM_\alpha[X] = Q_{\frac{1+\alpha}{2}}[X],$$

2. and if X is discrete then the difference between $TCM_\alpha[X]$ and $Q_{\frac{1+\alpha}{2}}[X]$ depends on ways of estimating median.

To overcome these weaknesses we replace median by trimmed mean in the definition of tail conditional median. And get the new risk measure tail conditional trimmed mean (TCTM) at level α which is defined as (see, e.g. Huber (1981)[20])

$$TCTM_\alpha^k[X] = trimmed\ mean[X|X > Q_\alpha[X]],\ k \in Z^+, \alpha \in (0,1). \qquad (5)$$

where k is the 'trimming' parameter described below. A trimmed mean is calculated by discarding a certain percentage of the highest (or lowest) data and then computing the mean of the remaining scores. Here we discard the highest data. Let $x_1, x_2, \ldots, x_n$ be a set of real-valued data observations. More concretely, let $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ be the order statistics of the observations. The kth highest trimmed mean $\overline{x}_k$ is defined as:

$$\overline{x}_k = \frac{x_{(1)} + x_{(2)} + \ldots + x_{(n-k)}}{n-k} = \frac{1}{n-k}\sum_{i=1}^{n-k} x_{(i)}.$$

By ordering the original observations, and taking away the first k largest observations, the trimmed mean takes the arithmetic average of the resulting data. The idea of a trimmed mean is to eliminate outliers, or extreme observations that do not seem to have any logical explanations in calculating the overall mean of a population.

In contrast to the arithmetic mean, the trimmed mean is a robust measure of central tendency. For example, an outlier, or a small fraction of anomalous measurements with abnormally large deviation from the center may change the mean value substantially. At the same time, the trimmed mean is stable in respect to the presence of such abnormal extreme values, which get 'trimmed'(thrown) away.

The choice of how to deal with an outlier should depend on the case. The outlier should be excluded if it is misleading, otherwise it should be included. It is important to ensure that the choice of the risk measure is not allowed to

mislead any practical application in this respect. We propose the following procedure. First, detect outliers. Then decide:

a) If they are true, treat as extreme events and use TCE. For more information, see e.g. Neftci (2000)[27].

b) If they are data errors use a robust risk measure. 1) If the tail behavior is not important and the estimation error is small enough for the application in question, use VaR; 2) Otherwise, try to use TCTM with a suitable parameter k.

3.4 Forecasting Risk

Internal models should also serve as an early warning system (see, e.g. Ronkainen et al. (2007) [31]). Thus, risk measures should be forward-looking; The problem is to forecast the future based on past experience and available data. However, in order to properly analyze the data and to construct a satisfactory risk forecast, it is essential to understand the environment in which the data has been collected and in particular how the environment has changed in the past. Typically, the forecasting errors increase with the horizon h. Since in Solvency II there is one-year forecasting horizon, the forecasting error can be substantial.

Conditional on the information given up to time t, a risk forecast, i.e., predictive VaR for period $[t, t+h]$ is:

$$\mathrm{VaR}_{\alpha}^{h}[X_t] = Q_{\alpha}[X_{t+h}\,|F_t],\ \alpha \in (0,1), h > 0.$$

where $Q_{\alpha}[\bullet]$ is quantile, X_t loss distribution at presnt t and X_{t+h} at time $t+h$ (h is forecasting horizon), and filtration F_t represents the information available at time t.

In time series context the most common model for VaR forecasting is the GARCH class of models (Bollerslev (1986)[3]). It should be noted that more complicated models, while often giving a better historical fit, do not necessary produce better out-of-sample forecasts. Model's forecasts tend to have worse error than they should when used outside the period of fit. There are various reasons why (see, e.g. Chatfield (2001)[4]). Kuester et al. (2006)[25] compared the out-of-sample performance of several methods for predicting VaR in the case of equity index. They found a hybrid method, combining CARCH with an extre value theory based approach best, but most

methods performed indadequately. In general, the main sources of forecasting errors are:

a) Model parameters have to be estimated - This constrains the number of parameters in risk model;

b) Exogenous variables may also have to be forecasted - This is quite common problem for internal models;

c) The wrong model may be identified - Tail and dependence modelling are especially demanding tasks faced in risk modelling; and

d) The underlying model may change - This is important for risk forecasting since economic environment may change rapidly.

To conclude, robustness and related parsimony (simplicity) is also important for risk forecasting.

3.5 In Defence of Value at Risk

The VaR measure has been rightly criticized for being sometimes inadequate. The VaR is not unproblematic to use: it is not subadditive, its estimation is subject to large errors, the estimate is downward biased (see, Krause (2003)[24]). However, these shortcomings do not imply that VaR is not a useful tool in risk management. It has clear advantages:

a) First, it serves the need for non-technical parties to be able to understand the risk measure. As such, it is the minimum technical requirement from which other, more complicated, measures are derived.

b) More importantly, Value at Risk is robust risk measure.

On the other hand, the robustness of the coherent risk measures - like TCE - is questionable because its behavior is critically influenced by the tails of the distribution(s) which drive the models. Hence the estimation is very sensitive to outliers and distribution assumptions.

For instance, if it is accepted that stock returns have tails heavier than those of normal distribution, one school of thought believes tails to be exponential type and another believes power-type tails. Heyde and Kou (2004)[18] shows that it is very difficult to distinguish between exponential-type and

power-type tails with 5,000 observations (about 20 years of daily observations). This is mainly because the quantiles of exponential-type distributions and power-type distributions may overlap.

For example, surprisingly, an exponential distribution has larger 99 percentile than the corresponding t-distribution with degree of freedom 5. If the percentiles have to be estimated from data, then the situation is even worse, as we have to rely on confidence intervals which may have significant overlaps. Therefore, with ordinary sample sizes (e.g. 20 years of daily data), one cannot easily identify exact tail behavior from data. Hence, the tail behavior may be a subjective issue depending on people's modeling preferences.

The fact that diversification is not universally preferable makes it unreasonable to criticize VaR just because it does not have subadditivity universally. Although in the center of the distributions VaR may violate the subadditivity, Danielsson et al. (2005)[6] questioned whether the violation is merely a technical issue, at least if one focuses on the tail regions which are the most relevant regions for risk management. Indeed they showed that VaR is subadditive in the tail regions, provided that the tails in the joint distribution are not extremely fat (with tail index less than one). They also carried out simulations showing that VaR is indeed subadditive for most practical applications. Distributions with tail index less than one have very fat tails. They are hard to find and easy to identify. Danielsson et al. (2005)[6] argued that they can be treated as special cases in financial modeling.

To summarize, the conflict between the use of VaR and diversification is not as severe as is commonly believed:

a) When the risks do not have extremely heavy tails, diversification is preferred and VaR satisfies subadditivity in the tail region;

b) When the risks have extremely heavy tails, diversification may not be preferable and VaR may fail to have subadditivity;

c) Even if one encounters an extreme fat tail and insists on tail subadditivity, Garcia et al. (2007)[16] showed that, when tail thickness causes violation of subadditivity, a decentralized risk management team may restore the subadditivity for VaR by using proper conditional information.

4 The Required Solvency Capital

Pentikäinen et al. (1989)[30] wrote on the complexity of solvency assessment: "*It should be appreciated that the solvency situation and the financial strength of an insurer are affected by nearly all activities and decision-making process, such as rating, risk selection, reserves evaluation, reinsurance, investments, sales efforts, etc., and by external factors.*"

The choice of the risk measure should depend on the application in question. Heyde et al. (2007)[19] argue that when a regulator imposes a risk measure, it must be unambiguous, stable, and can be implemented consistently throughout all the companies. Otherwise, different companies using different risk models may report very different risk measures to the regulator. Further they conclude that from a regulator viewpoint, the risk measure should demonstrate robustness with respect to underlying models to maintain the stability of the regulation.

4.1 VaR as a Solvency Capital Requirement

In order to protect the policyholders, the regulatory authority in force will impose a solvency capital requirement. Although the regulator may want the solvency capital requirement to be as large as possible, clearly holding too much capital is costly.

Consider a portfolio with future loss X. To protect the policyholders from insolvency, the regulatory authority imposes a solvency capital requirement, risk measure $\rho[X]$, which means that the available capital in the company has to be at least equal to $\rho[X]$. This capital can be employed when premiums and provisions together with the investment income, turn out to be insufficient to cover the policyholders' claims. In principle, $\rho[X]$ will be chosen so that one can be 'fairly sure' that the event $X > \rho[X]$ will not occur.

The regulator wants the solvency capital requirement related to X to be sufficiently large, to ensure that the shortfall risk is sufficiently small. Following Dhaene et. al (2008) [12],We suppose that, to reach this goal, the regulator introduces a risk measure for the shortfall risk, which we will denote by φ:

$$\varphi[(X - \rho[X])_+].$$

One can see that two different risk measures are involved in the process

of setting solvency capital requirements: the risk measure ρ that determines the solvency capital requirement and the risk measure φ that measures the shortfall risk.

We will assume that φ satisfies the following condition:

$$\rho_1[X] \leq \rho_2[X] \Rightarrow \varphi[(X - \rho_1[X])_+] \geq \varphi[(X - \rho_2[X])_+], \tag{6}$$

which means that an increase of the solvency capital requirement implies a reduction of the shortfall risk as measured by φ. A sufficient condition for equation (6) to hold is that φ is monotonic.

Assumption equation (6) implies that the larger the capital, the smaller $\varphi[(X - \rho[X])_+]$. And the regulator wants $\varphi[(X - \rho[X])_+]$ to be sufficiently small. However, we know that holding a capital $\rho[X]$ involves a capital cost $\rho[X]i$, where i denotes the required excess return on capital. To avoid imposing an excessive burden on the insurer, the regulator should also consider this capital cost. For a given X and solvency capital requirement $\rho[X]$, we define the cost function

$$C(X, \rho[X]) = \varphi[(X - \rho[X])_+] + \rho[X]\varepsilon, \qquad 0 < \varepsilon < 1,$$

which considers the shortfall risk and the capital cost at the same time. And ε can be interpreted as a measure for the extent to which the capital cost is taken into account. Of course the regulatory authority can choose ε as company-specific or risk-specific. The optimal capital requirement $\rho[X]$ can be determined as the smallest amount d that minimizes the cost function $C(X, d)$.

Obviously increasing the value of ε means that the regulator increases the relative importance of the cost of capital. This will result in a decrease of the optimal capital requirement.

For example, let $\varphi[X] = E[X]$. This choice of φ satisfies condition (6). In this case, the shortfall risk measure can be interpreted as the net stop-loss premium that has to be paid to reinsurer the insolvency risk. We state the following result (*see* Dhaene, et al. (2008) [12]):

Theorem 1 *The smallest element in the set of minimizers to the cost function*

$$C(X, d) = E[(X - d)_+] + d\varepsilon, \quad 0 < \varepsilon < 1, \tag{7}$$

is given by

$$\rho[X] = Q_{1-\varepsilon}[X]. \tag{8}$$

Remark 1 *The minimal value of the cost function in (7) can be expressed as*

$$C(X, Q_{1-\varepsilon}[X]) = E[(X - Q_{1-\varepsilon}[X])_+] + Q_{1-\varepsilon}[X]\varepsilon = \varepsilon TVaR_{1-\varepsilon}[X].$$

Theorem 1 provides a theoretical justification for the use of Value-at-Risk to set solvency capital requirements. Hence, to some extent the theorem supports the current regulatory regime for banking supervision established by the Basel II Capital Accord and insurance supervision established by Solvency II, which has put forward a Value-at-Risk-based capital requirement approach.

Furthermore, following Dhaene et. al (2008) [12], we consider two portfolios with respective future losses X_1 and X_2. We assume that the solvency capital requirement imposed by the regulator in force is represented by the risk measure ρ. We say that the portfolios are merged when they are jointly liable for the shortfall of the aggregate loss $X_1 + X_2$. The solvency capital requirement imposed by the supervisory authority will, in this case, be equal to $\rho[X_1 + X_2]$. When each of the portfolios is not liable for the shortfall of the other portfolio, we will say that they are stand-alone portfolios. In this case, the solvency capital requirement for each portfolio is given by $\rho[X_j]$.

Throughout, we assume that the losses X_1 and X_2 remain the same, regardless of whether or not the portfolios are merged, and that only the (legal) liability construction changes. In practice, merging or splitting portfolios may change management, business strategy, cost structure, and so on, and may therefore change the losses under consideration.

The subadditivity condition is often imposed on solvency capital principles. Important to notice is that the requirement of subadditivity implies that

$$(X_1 + X_2 - \rho[X_1 + X_2])_+ \geq (X_1 + X_2 - \rho[X_1] - \rho[X_2])_+,$$

and consequently, for some realizations (x_1, x_2) we may have that

$$(x_1 + x_2 - \rho[X_1 + X_2])_+ > (x_1 - \rho[X_1])_+ + (x_2 - \rho[X_2])_+.$$

Hence, when applying a subadditive risk measure in a merger, one could end up with a larger shortfall than the sum of the shortfalls of the stand-alone entities. Therefore, the regulatory authority needs to restrict the subadditivity in order to avoid that merging leads to a riskier situation.

4.2 Optimization under Variance Premium Principle

In the earlier article by Dhaene et al. (2003) [13], there is another expression of cost function. An interpretation of the solvency margin based on an economic reasoning could be obtained as follows. The price of reducing the total risk is the sum of the risk measure of the remaining risk added to the cost of the available capital to be paid to the shareholders. For instance, in case one transfers the risk by a stop-loss insurance with a loaded premium, the cost is $(1+\alpha)E[(X-d)_+]+\varepsilon d$. In this particular case, the optimal capital, minimizing the total cost, can be shown to be given by $d = F_X^{-1}(1-\varepsilon/(1+\alpha))$. In this case we get as a risk measure a particular quantile, where the probability is not arbitrary but can be determined from economic parameters.

We can expand the previous case by considering a risk business facing a net loss $X = Y - E[Y]$ at the end of the period. At the beginning of the period, the economic capital d is made available to the portfolio by the shareholders at a price of i per unit. To minimize the cost of capital $(i-r)d$ above the risk-free interest rd, we should take d as small as possible, but to minimize the insolvency risk $E[(X-(1+r)d)_+]$, the capital d should be large. Just like in the previous case, we have to minimize the total of these two cost components, hence

$$E[(X-(1+r)d)_+] + (i-r)d.$$

Note that a stop-loss premium can be expressed as follows:

$$E[(X-d)_+] = \int_d^\infty [1-F_X(x)]dx.$$

To explain the riskiness of the tail, we use Yaari's dual theory, introducing a distortion function g with $g(0)=0,\ g(1)=1,\ g(x)$ increasing and $g(x) \geq x$, *see*, e.g., Wang et al. (1996) [33]. Then we can compute the 'cost of avoiding insolvency' by

$$E[(X-d)_+] = \int_{(1+r)d}^\infty g(1-F_X(x))dx.$$

Therefore, instead of simply the total cost as above, we minimize the expression

$$C(X,d) = \int_{(1+r)d}^\infty g(1-F_X(x))dx + (i-r)d.$$

The optimal solution is given by

$$d = \frac{1}{1+r} F_X^{-1}\left(1 - g^{-1}\left(\frac{i-r}{1+r}\right)\right).$$

Assuming $r = 0$ and $i = 0.1$, without distortion, hence with $g(x) \equiv x$, the optimal working capital d equals the 90% percentile of X, but assuming $g(0.01) = 0.1$, it is the 99% percentile. Hence, the optimal threshold depends on i and on the way that we magnify the tail. It should be noted that the percentile is not the risk measure, but in fact it is the value of d corresponding to the (minimized) total cost.

Now we will more generically specify $\varphi[\cdot]$ as a variance premium principle, i.e.

$$\varphi[X] = E[X] + \beta Var[X],\ \beta > 0.$$

The interested reader is referred to Goovaerts et al. (1984) [17] for an elaborate treatment of the variance principle. Now the cost function $C(X, d)$ given by

$$C(X, d) = \varphi[(X-d)_+] + d\varepsilon = E[(X-d)_+] + \beta Var[(X-d)_+] + d\varepsilon.$$

We consider the capital allocation problem given by

$$\min_d E[(X-d)_+] + \beta Var[(X-d)_+] + d\varepsilon.$$

We differentiate $C(X, d)$ with respect to d:

$$\frac{\partial C(X,d)}{\partial d} = \frac{\partial E[(X-d)_+]}{\partial d} + \beta \frac{\partial Var[(X-d)_+]}{\partial d} + \varepsilon$$

The derivative of $E[(X-d)_+]$ is given *see* e.g. Section 1.4 in Kaas et al. (2001) [22]

$$\frac{\partial E[(X-d)_+]}{\partial d} = F_X(d) - 1.$$

We know that $Var[(X-d)_+] = E[(X-d)_+^2] - E[(X-d)_+]^2$, then

$$E[(X-d)_+^2]) = \int_d^\infty (x-d)^2 f_X(x)dx$$

Let's differentiate $E[(X-d)_+^2])$ with respect to d:

$$\frac{\partial E[(X-d)_+^2]}{\partial d} = \int_d^\infty \frac{\partial (x-d)^2}{\partial d} f_X(x)dx = -\int_d^\infty 2(x-d) f_X(x)dx = -2E[(X-d)_+].$$

So

$$\frac{\partial Var[(X-d)_+]}{\partial d} = -2E[(X-d)_+] - 2E[(X-d)_+][F_X(d)-1] = -2E[(X-d)_+]F_X(d).$$

Let

$$\frac{\partial C(X,d)}{\partial d} = F_X(d) - 1 - 2\beta E[(X-d)_+]F_X(d) + \varepsilon = 0,$$

get d satisfying the equation

$$F_X(d)(1 - 2\beta E[(X-d)_+]) = 1 - \varepsilon$$

makes the cost function minimal.

5 Conclusions

Pentikäinen (1975)[28] emphasized the advantage of theory and modelling over the rules of thumb: *"Often many 'experienced' managers take a suspicious and deprecatory stand on theoretical considerations, which they easily pass over with short comments on their lack of practical value. However, neglecting to clearly formulate the problems and the principles of policies does not mean that the manager in question does not, in fact, follow some strategy. On the contrary, every way of decision making, even neglecting to make a decision, is some kind of strategy. The difference is only that the strategy of 'practical men' can be a random product of old traditions, more or less reliable institutions, etc. without any clear formulation and analysis of various policies. A discussion on theoretical aspects and on the theoretical point of view, even if the direct numerical results are of little value, may anyway direct attention to the statement and restatement of problems and to a conscious analysis of the facts and possibilities*".

Another sound perspective on risk modelling is presented in Danielsson (2008)[5]. He claims that the current financial crisis took everybody by surprise in spite of all the sophisticated models and all the stress testing. Further, the financial institutions that are surviving this crisis best are those with the best management, not those who relied on models to do the managements job. He concluded that risk models do have a valuable function in the risk management process so long as their limitations are recognized.

For calculating the solvency capital requirement under Solvency II regime one has to consider the following aspects. First, the data series should be

sufficiently long to cover cycles, clusters and atypical observations that the random variable in question may produce in the future. Second, forecasting horizon is one year. That long horizon causes severe problems with forecast accuracy and backtesting.

The standard error of quantile estimator typically increases when one goes further in the tail of the loss distribution. When estimating 99.5 % quantiles for internal model this is a major problem in practice.

The behavior of quantile estimates is also critically influenced by the tails of the distributions. The distinctions are exceedingly difficult for popular methods and particularly large samples are needed for clear discrimination. Using heavy tailed distributions and/or copulas makes the data issue even more critical. In the tail modelling subjectivity is unavoidable and this must be appropriately dealt with by the firms and supervisors. While many risk measures may be suitable for internal risk management, robustness is especially important consideration for a regulatory risk measure.

In our opinion, the old advice "keep it sophisticatedly simple (KISS)" is also wise principle in the risk measurement context.

6 Acknowledgement

We are grateful for financial support from the China Scholarship Council and the Actuarial Society of Finland. In addition, we would like to thank Dr. Laura Koskela for helping us with Latex program.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.27.1799`.

[2] R. Beard, T. Pentikäinen, and E. Pesonen. *Risk Theory.* Methuen and Co LTD, London, 1969.

[3] T. Bollerslev. Generalized autoregressive conditional heteriskedasticity. *Journal of Econometrics*, (51):307–327, 1989.

[4] C. Chatfield. *Time Series Forecasting.* Chapman and Hall /CRC, 2001.

[5] J. Danielsson. Blame the models. *Journal of Financial Stability*, Forthcoming 2008.

[6] J. Danielsson, B. Jorgensen, M. Sarma, and C. de Vries. Sub-additivity re-examined: the case for value at risk. *Working paper*, February 2005.

[7] P. L. Davies and U. Gather. The identification of multiple outliers (invited paper with discussion and rejoinder). *Journal of the American Statistical Association*, 88:782–792.

[8] C. Daykin, T. Pentikäinen, and M. Pesonen. *Practical Risk Theory for Actuaries*. Chapman and Hall, London, 1994.

[9] F. Delbaen. Coherent risk measures on general probability spaces. *Advances in Finance and Stochastics*, pages 1–37, 2002. URL `http:citeseer.ist.psu.edu/delbaen00coherent.html`.

[10] D. Denneberg. *Non-additive measure and integral*. Kluwer Academic Publishers, Boston, 1994.

[11] M. Denuit, J. Dhaene, M. Goovaerts, and R. Kaas. *Actuarial Theory for Dependent Risks: Measures, orders and Models*. Wiley, Wiltshire, 2005.

[12] J. Dhaene, R. J. Laeven A., S. Vanduffel, G. Darkiewicz, and J. Goovaerts M. Can a coherent risk measure be too subadditive? *The Journal of Risk & Insurance*, 75(2):365–386, June 2008. URL `http://dx.doi.org/10.1111/j.1539-6975.2008.00264.x`.

[13] J. L. M. Dhaene, M. J. Goovaerts, and R. Kaas. Economic capital allocation derived from risk measures. *North American Actuarial Journal*, 7:44–59, 2003. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.62.5311`.

[14] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002. URL `http://www.springerlink.com/content/59tkugt7vtnetrbv/fulltext.pdf`.

[15] M. Frittelli and E. Rosazza Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486, July 2002. URL `http://ideas.repec.org/a/eee/jbfina/v26y2002i7p1473-1486.html`.

[16] R. Garcia, E. renault, and G. Tsafack. Proper conditioning for coherent var in portfolio management. *Management Science*, (53):483–494, 2007.

[17] M. Goovaerts, F. De Vylder, and J. Haezendonck. Insurance premiums. *North Holland Publishing, Amsterdam.*, 1984.

[18] C. Heyde and S. Kou. On the controversy over tailweight distributions. *Operations Research Letters*, pages 399–408, 2004.

[19] C. C. Heyde, S. G. Kou, and X. H. Peng. What is a good risk measure: Bridging the gaps between data, coherent risk measures, and insurance risk measures. *Columbia University*, July 2007. URL `http://www.fdic.gov/bank/analytical/cfr/2007/apr/KOU`$_h kpv1.pdf$.

[20] P. J. Huber. *Robust Statistics.* Wiley, New York, 1981.

[21] W. Hurlimann. Splitting risk and premium calculation. *Bulletin of the Swiss Association of Actuaries*, pages 169–179, 1994.

[22] R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit. *Modern Actuarial Risk Theory.* Kluwer Academic, 2001.

[23] L. Koskinen, L. Koskela, and V. Ronkainen. Data and internal models. *FIN-FSA Research Papers*, (2), Forthcoming 2009.

[24] A. Krause. Exploring the limitations of value at risk: How good is it in practice? *The Journal of Risk Finance*, pages 19–28, Winter 2003.

[25] K. Kuester, S. Mittnik, and M. Paolella. Value-at-risk predictions: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4:53–59, 2006.

[26] G. Meyers. Coherent measures of risk: An exposition for the lay actuary. *Insurance Services Office, Inc.*, pages 1–11, March 2000. URL `http://www.gloriamundi.org/picsresources/gmcmr.pdf`.

[27] S. N. Neftci. Value at risk calculations, extreme events, and tail estimation. *The Journal of Derivatives*, Spring 2000. URL `http://www.neftci.com/papers/papers/VaR and Extremes.pdf`.

[28] T. Pentikäinen. A model of stochastic-dynamic prognosis: An application of risk theory to business planning. *Scandinavian Actuarial Journal*, (24):29–53, 1975.

[29] T. Pentikäinen. *(ed.), Solvency of insurers and equalization reserves, Vol I.* Insurance publishing Company Ltd, Helsinki, 1982.

[30] T. Pentikäinen, H. Bonsdorff, M. Pesonen, J. Rantala, and Ruohonen. *Insurance solvency and financial strength.* Finnish Insurance Training and Publishing Company Ltd, Helsinki, 1989.

[31] V. Ronkainen, L. Koskinen, and R. Berglund. Topical modelling issues in solvency 2. *Scandinavian Actuarial Journal*, (2):135–146, 2007.

[32] J. L. Teugels and B. Sundt. *Encyclopedia of Actuarial Science.* Wiley, 1st edition, Nov. 2004. URL `http://www.wiley.co.uk/eoas/`.

[33] S. Wang. Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, 26:71–92, 1996. URL `http://www.casact.org/library/astin/vol26no1/71.pdf`.

[34] S. S. Wang, V. R. Young, and H. H. Panjer. Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, 21(2):173–183, November 1997. URL `http://ideas.repec.org/a/eee/insuma/v21y1997i2p173-183.html`.