ROYAL INSTITUTE OF TECHNOLOGY
MASTER'S THESIS

Robustness of Conditional Value-at-Risk (CVaR) when measuring market risk across different asset classes

Mattias Letmark
letmark@kth.se

March 2010

Tutor: Professor Boualem Djehiche

Acknowledgements

I would like to thank my tutor, professor Boualem Djehiche, for providing helpful feedback, support and valuable advice in writing this thesis. I would also like to thank my girlfriend, Monia Eddaoui, for love, support and understanding during the process

Abstract

We investigate the robustness of Conditional Value-at-Risk (CVaR) for market risk. The analysis is performed on different asset classes including stock indexes, bond indexes, exchange rates and individual stocks. We find that a robust CVaR measure can be constructed for almost all of these assets. The key issue is to choose appropriate parameters, such as confidence levels and ex ante window size for the CVaR estimator. However, in some cases, the CVaR measure is not robust, which happens primarily when measuring market risk on individual stocks.

Contents

1 Introduction

This section gives a background and an overview of previous work. Furthermore, the contribution and purpose of the thesis are presented.

1.1 Background

1.1.1 Market Risk Management

In light of the global financial crisis following the failure of Lehman Brothers, the measurement of market risk has become a primary concern for regulators and risk managers. Coordinated by regulatory authorities, such as the Basel Committee on Banking Supervision (the Basel Committee) [47] and by authorities of the EU member states, e.g. the Swedish Financial Supervisory Authority (Finansinspektionen) [48], banks are required to hold a certain amount of capital against adverse market movements. Specifically, banks must demonstrate that its capital is sufficient to cover losses 99.9% of the times over a one year holding period [11]. Such a risk capital is usually called Value-at-Risk (VaR).

An important milestone in the development of VaR models was J.P. Morgan's decision in 1994 to make its VaR system, RiskMetrics [40], available on the Internet. In the following years, the RiskMetrics system attained a de-facto standard status within the financial industry and a benchmark for measuring market risk. However, in the financial literature, additional measures of market risk besides VaR have been studied. Artzner et al. [8] highlighted some theoretical shortcomings of VaR as a measure of market risk. For example, it does not take into account the magnitude of losses when VaR is exceeded. VaR also fails to meet the characteristic of subadditivity (see section 2.2.3), i.e. the risk of a portfolio in terms of VaR may be larger than the sum of its components. Artzner et al. [8] proposed an alternative risk measure defined as the expected value of losses exceeding the VaR. This new risk measure has sounder theoretical properties, e.g. fulfills the subadditivity condition, and is usually called Conditional Value-at-Risk (CVaR).

There exist a large number of different risk measures, of which only a few have been mentioned here, all with its own characteristics, advantages and flaws. Triggered by the Basel Committee, VaR has been adopted as the main measure of market risk. However, VaR has attracted a lot of criticism as a risk measure. One reason is that the VaR concept can lead to perverse effects if used as a control mechanism. An example is shown in [15], where it is described how to earn \$1 million in one week with no initial capital. Other drawbacks with VaR are e.g. that using VaR as a risk measure may fail to stimulate diversification, due to its non-subadditivity characteristic (see e.g. [41]), and that VaR only provides a point-estimate of the loss distribution. The VaR estimate does not provide any information on the losses in the tail exceeding VaR, i.e. information on so-called *spike the firm* events (low probability, high loss) is not captured with the model. History and the recent global financial turmoil has shown that such events pose a real threat to e.g. the banking system (see [18] and [46]).

1.1.2 Asset Allocation and Portfolio Theory

Markowitz's portfolio theory has influenced academia and financial institutions since it was published in 1952 [33]. Markowitz proposed that a portfolio should be optimized in a mean-variance framework, i.e. maximizing the returns and at the same time keeping the risk under control. The definition of risk in this framework was defined as the overall portfolio variance. A more comprehensive description of portfolio theory and portfolio optimization is given in [14].

A drawback related to variance as a risk measure is that it penalizes upside (gains) and downside (losses) equally. As a complement to the mean-variance optimization model, not only relying on the variance as a risk measure, additional constraints can be added to control the risk. This is especially important as a tool for agency control. Alexander and Baptista [5] analyze the results from imposing VaR and CVaR constraints in the mean-variance framework. They show that in some cases such impositions may induce perverse effects, e.g. that risk averse agents select portfolios with larger standard deviations.

Instead of optimizing according to the mean-variance model, a portfolio can be optimized in other frameworks. Since VaR is one of the most popular risk measures in risk management, many studies have been performed on optimization in the mean-VaR framework. However, VaR is a non-convex and non-smooth function which has multiple local extrema [35]. Uryasev and Rockafellar [44] developed a mean-CVaR model using a linear optimization method and showed that VaR can be calculated as a by-product. Another advantage with the mean-CVaR model is that CVaR optimization seems more stable over different confidence levels, at least in the case of fixed income securities (see [34]). Olszewski [38] studied hedge funds and suggested that a more efficient portfolio can be constructed by optimization in the mean-CVaR domain compared to the classic mean-variance domain.

1.1.3 Contribution

In the financial literature it is often suggested that CVaR has more attractive properties as a risk measure than VaR. But is CVaR really better than VaR? At least, CVaR can solve some of the issues with VaR, e.g. the diversification problem stated in [41] (due to its sub-additivity property) and the situation where you could earn $1 million with no initial capital described in [15] (due to its consideration of high, low-probability losses). Furthermore, in [19] it is shown that different assets will be ranked in the same way in terms of risks measured as VaR and CVaR, respectively. This indicates that some useful properties of VaR are transferred to the properties of CVaR and that CVaR in terms of risk ranking seems to do an equally good job as VaR.

For a risk model to be considered robust, it should provide accurate risk forecasts across different assets, time horizons, and confidence levels within the same asset class. Surprisingly, risk forecast fluctuations have not been well documented. Nevertheless, fluctuations in risk forecasts have serious implications for the usefulness of a risk model. If a VaR value always fluctuates by 30% from one day to the next, it may be hard to sell risk modelling within the firm. Traders are not likely to be happy with routinely changing risk limits, and management does not like to change market risk capital levels too often. Moreover, since VaR is used to regulate market risk capital, a volatile VaR leads to costly fluctuations in capital if the firm keeps its capital at the predicted minimum level. This may severely hinder the adoption of risk models within a firm.

There has been extensive research on CVaR in terms of portfolio optimization, but not much in terms of robustness. The topic will be further investigated in this thesis. A similar study was also suggested as future work by Lambadiaris [29].

1.2 Purpose

The purpose of this thesis is to study the robustness of CVaR as a measure for market risk. The CVaR robustness will be analyzed and compared for different asset classes, ex ante window lengths and confidence levels.

1.3 Outline

The structure of this thesis is as follows: Section 1 provides a background to the subject and an overview of previous work. It also presents the contribution and purpose of the thesis. Section 2 introduces the theoretical framework, defines the two risk measures VaR and CVaR and explains the concept of robustness. Section 3 describes the selection of the data sample used in the thesis. Section 4 outlines methodology and the empirical measures and tests used in the analysis. Section 5 presents the empirical findings and analysis of them. Eventually, section 6 concludes the main results and provides suggestions for future research.

2 Theoretical Framework

This section introduces the theoretical framework for the thesis. The two risk measures VaR and CVaR are defined and some of their properties discussed. A brief overview of main estimation techniques is given, delta-normal approach and historical simulation. Moreover, backtesting procedures are described and eventually the concept of robustness is discussed.

2.1 Market Risk

Financial market participants face a risk of disastrous losses due to unexpected adverse movements in market factors. The risk of losses arising from movements in market prices is often referred to as *market risk*. The Basel Committee on Banking Supervision classifies the sources of market risk into four main categories: equities, interest rate related instruments, foreign exchange and commodities [10]. As we have seen over the last years, there has been an increasing instability in the financial environment, an increasing globalization of financial markets, a significant growth of trading activity, development of numerous new financial products, new enabling technologies and regulatory requirements. These are all factors contributing to an increasing interest in market risk.

There are two main approaches of measuring market risk, *statistical methods* and *scenario based methods*. Comprehensive risk managers combine the use of statistical risk measures with techniques such as stress testing, scenario analysis and visualization. Just as a single diagnostic such as body temperature is not a reliable measure of the health of a human being, risk managers should not rely solely on a single method to determine the health (or risk) of a portfolio.

The scope of this thesis covers only statistical risk measures.

2.2 Risk Measures

Since the pioneering work of Markowitz [33], where he introduced the modern portfolio theory, the variance has been the traditional risk measure in economics and finance. However, there are several shortcomings related to variance as a risk measure. For example, it penalizes upside (gains) and downside (losses) equally and mean-variance decisions are usually not consistent with the expected utility approach, unless returns are normally distributed or a quadratic utility function is used. Moreover, the variance does not account for fat tails of the underlying distribution and therefore is inappropriate to describe the risk of low probability events, such as default risks.

2.2.1 Value-at-Risk

In recent years, academics and practitioners have extensively studied a risk measure called Value-at-Risk (VaR). It was developed to respond to the need to aggregate the various sources of market risk into a single quantitative measure. VaR focuses on the downside risk of a portfolio and is defined as the maximum expected loss at a specific confidence level (e.g. 95%) over a certain time horizon[1] (e.g. ten days). For example, if VaR is -\$100 for a portfolio at a confidence level of 95% and a time horizon of one week we can state that "with 95% certainty we will not lose more than \$100 over the next week". In another example, consider a bank that calculates its VaR assuming a one-day holding period and a 99% confidence level. Then the bank can expect that, on average, trading losses will exceed the VaR on one occasion in one hundred trading days.

The choice of confidence level varies among different risk managers. For example, the Basel Committee recommends the 99.9% confidence level for capital adequacy purposes [11]. For internal use, lower confidence levels are often used. For example, J.P. Morgan [27] uses a 99% confidence level, Citibank [16] uses a level of 95.4% and Goldman Sachs [24] uses a 95% level.

Another parameter that varies among risk managers is the time horizon over which VaR is estimated. It is likely that the portfolio return changes more over a month than over a single day. The length of the holding period depends on the nature of the portfolio and typically ranges from one day to one month. The Basel Committee recommends a time horizon of ten days for most capital market transactions [11].

The mathematical definition of VaR is:

$$F(\alpha) = \int_{-\infty}^{\mathrm{VaR}_\alpha} p(r)\mathrm{d}r = 1 - \alpha \tag{1}$$

where the sample space Ω of the expected rates of return r on some arbitrary assets is represented by the set $\Re$. $p(r)$ is the probability density and we assume that the expected rates of return $r(t)$ with respect to the time horizon t of the investment and the confidence level $\alpha \in [0, 1]$ is a random variable determined by the distribution function $F : \Omega \to [0, 1]$.

Equivalently, this can be written

$$F(\alpha) = P[r \leq \mathrm{VaR}_\alpha] = 1 - \alpha \tag{2}$$

[1] Since VaR assumes no changes in the portfolio weights during the time horizon, the term *holding period* is often used instead of time horizon

A graphical interpretation of VaR using a confidence level of 95% is illustrated in Figure 1. VaR_α is the cut-off point separating the return distribu-

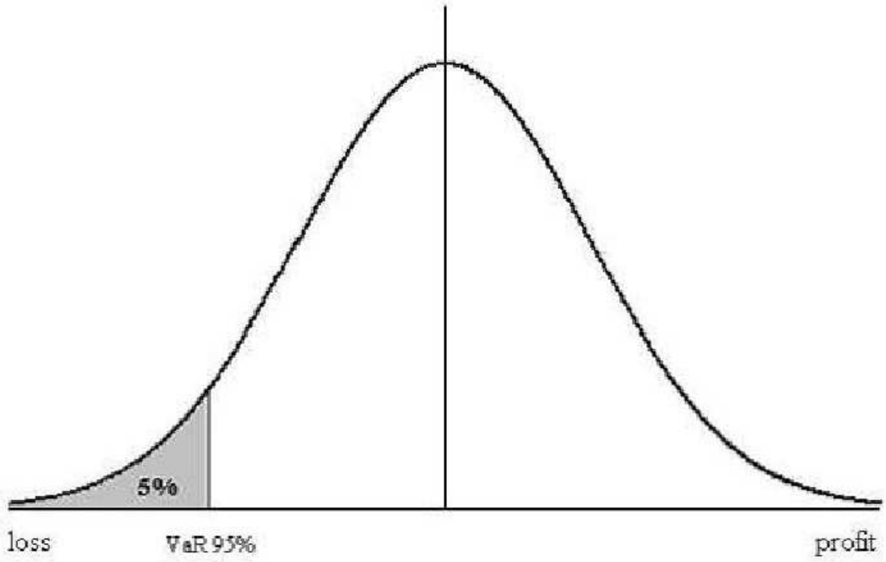

Figure 1: Graphical interpretation of VaR

tion from its 5% tail.

2.2.2 Coherent Risk Measure

Not until 1997, with the appearance of *Thinking Coherently* [7] by Artzner et al., it was defined in a clear way what properties a statistic should have in order to be considered a *coherent risk measure*. Artzner et al. (see [8] for a more technical presentation) formulated four axioms that have to be fulfilled by a coherent risk measure. X and Y denote portfolio returns, $\rho(X)$ and $\rho(Y)$ are their risk measures, respectively, and c is an arbitrary constant:

Translation invariance

$$\rho(X + c) = \rho(X) - c \tag{3}$$

Subadditivity

$$\rho(X + Y) \leq \rho(X) + \rho(Y) \tag{4}$$

Positive homogeneity

$$\rho(cX) = c\rho(X) \tag{5}$$

Monotonicity

$$\rho(X) \leq \rho(Y), \text{ if } X \geq Y \tag{6}$$

The translation invariance axiom (3) means that adding cash to the portfolio decreases the risk by the same amount. The axiom of subadditivity (4) ensures that the risk of the total portfolio is not larger than the sum of the risks of its components to reflect the effect of diversification and hedges. Positive homogeneity (5) means that the risk is scaled with the portfolio size. Finally, monotonicity (6) is required to ensure that if the payoff of portfolio X dominates the payoff of portfolio Y, then the risk of portfolio Y cannot be lower than the risk of portfolio X [43].

In simple words, the axioms defining a coherent risk measure means that whenever a portfolio is undoubtedly riskier than another one, it will always have a higher risk value as long as the risk measure is coherent. On the other hand, a measure not fulfilling all axioms might give wrong assessment of relative risks [1].

2.2.3 Criticism on VaR

Surprisingly, VaR, despite its wide acceptance, does not fulfill all axioms of coherence [2]. In fact, VaR fails to meet the characteristic of sub-additivity[2], i.e. the risk of a portfolio in terms of VaR may be larger than the sum of risks of its components. The sub-additivity condition plays a fundamental role in risk measurement. With non-subadditivity it could be the case that a well diversified portfolio require more regulatory capital than a less diversified portfolio. Thus, managing risk in terms of VaR prevents to add up the VaR of different risk sources and may fail to stimulate diversification (see e.g. [1], [7], [8] or [41]).

The non-subadditivity characteristic of VaR can be demonstrated by a simple example. Suppose that we have two short positions in out-of-the-money options. The specific details are shown in Table 1. Each of the options has a 4% probability of a payout of −\$100 and a 96% probability of a payout of zero. If we take the VaR at the 95% confidence level, then each of the positions has a VaR of zero. However, if we combine the two positions, the probability of a zero payout falls to less than 95%, and so the VaR of the combined portfolio is less than zero (in this case equal to −\$100, see Table 2). The VaR of the combined position is therefore greater than the sum of the VaRs of the individual components, so the VaR is clearly not sub-additive.

[2] However, VaR is a coherent risk measure when it is based on the standard deviation of normal distributions

Table 1: Non-subadditivity: Options positions considered separately

OPTION A		**OPTION B**	
Payout	**Probability**	**Payout**	**Probability**
-\$100	4%	-\$100	4%
0	96%	0	96%
VaR 95% = 0		VaR 95% = 0	

Table 2: Non-subadditivity: Options positions combined

COMBINED	
Payout	**Probability**
-\$200	0.16%
-\$100	7.68%
0	92.16%
VaR 95% = -\$100	

Another criticism on VaR is based on its non-convexity characteristic, which limits its use as a risk measure in optimal portfolio selection for investment purposes. It has been shown [5] that having embedded VaR into an optimization framework, VaR risk managers incur larger losses than non-risk managers in the most adverse states of the world. Moreover, Basak and Shapiro [9] show that an agent facing a VaR constraint may choose a larger exposure to risky assets than in the absence of the constraint. It is also shown in [35] and [36] that the problem of minimizing VaR of a portfolio of derivative contracts can have multiple local minimizers, which will lead to unstable risk ranking.

2.2.4 Conditional Value-at-Risk

VaR is often criticized for not taking into account the magnitude of losses when VaR is exceeded. CVaR is often proposed as an alternative to VaR. CVaR is also known as *expected shortfall* [1], *tail VaR* [7] or *mean shortfall* [35]. In the context of continuous distributions (which we assume for simplicity in this paper), for a given confidence level α and holding period t, CVaR is defined as the conditional expectation of the losses exceeding VaR. Hence, in contrast to VaR, CVaR provides additional information of the losses in the tail exceeding VaR.

Mathematically, CVaR is defined by:

$$\mathrm{CVaR}_\alpha = \frac{1}{1-\alpha}\int_{-\infty}^{\mathrm{VaR}_\alpha} rp(r)\mathrm{d}r \qquad (7)$$

or equivalently

$$\text{CVaR}_\alpha = E[x|x \leq \text{VaR}_\alpha] \tag{8}$$

where $p(r)$ is the probability density and $r(t)$ the expected rates of return with respect to the time horizon t of the investment and VaR is calculated over the same time horizon with confidence level $\alpha \in [0, 1]$.

A graphical interpretation of CVaR is illustrated in Figure 2. CVaR is the

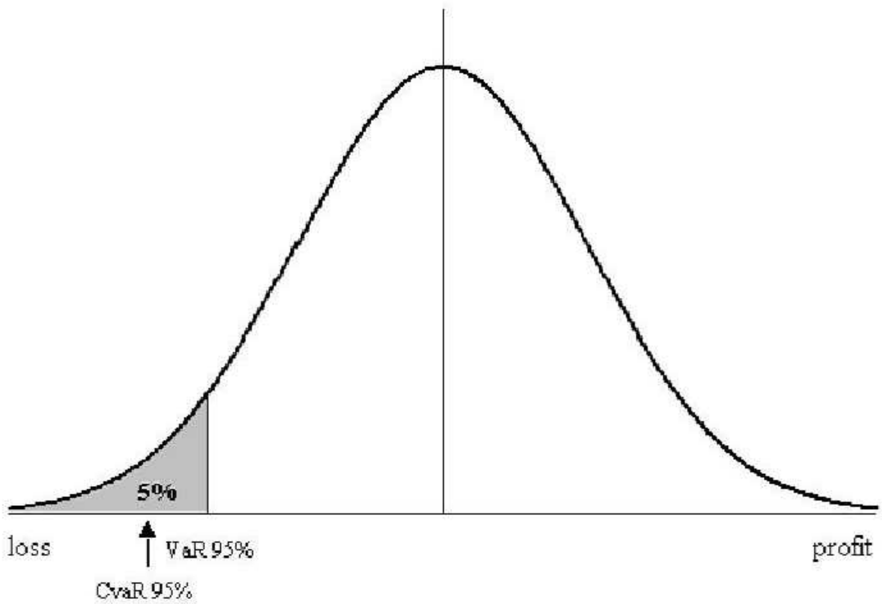

Figure 2: Graphical interpretation of CVaR

expected loss if a tail event does occur, and is therefore graphically located to the left of VaR.

Acerbi and Tasche [3] show that CVaR satisfies the four axioms in section 2.2.2 and, consequently, qualifies as a coherent risk measure. In fact, Acerbi [4] shows that any coherent risk measure can be represented as a convex combination of CVaRs with different confidence levels. In addition, CVaR is a convex function with respect to portfolio positions, allowing the construction of efficient optimizing algorithms. In particular, Uryasev and Rockafellar [44] show that CVaR can be minimized using linear programming techniques, which makes many large-scale calculations practical, efficient and stable.[3]

[3] In fact, the superintendent office of financial institutions in Canada has put in regulation for the use of CVaR to determine the capital requirement.

2.3 Estimating VaR and CVaR

There are many ways of estimating VaR (see Duffie and Pan [22] for a comprehensive overview). Given the return distribution, the calculation of VaR is straightforward and given VaR, the calculation of CVaR is straightforward. Therefore, the challenges of estimating VaR and CVaR are mainly related to the estimation of the return distribution. The approaches can be categorised to parametric and non-parametric methods. Parametric approaches make some assumptions about the return distribution, e.g. the assumption of normality (see section 2.3.1). The distribution assumptions imply model risk, i.e. the risk that there is a discrepancy between the assumed return distribution and the *true* underlying probability distribution [20]. Non-parametric methods base the VaR estimation solely on empirical distributions of returns. A disadvantage is that the estimates are completely dependent on a particular data set. The simplest non-parametric method is called *historical simulation method* (see section 2.3.2).

2.3.1 Delta-Normal Approach

The simplest parametric method is the delta-normal (analytic) approach. Following this approach it is assumed that all asset returns are normally distributed. As the portfolio return is a linear combination of normal variables, it is also normally distributed. The VaR of a portfolio is then calculated using historical (ex ante) means, variances and covariances of the portfolio components. More formally, this can be written as:

$$\mathrm{VaR}_\alpha = \mu - z_\alpha \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}} = \mu - z_\alpha \sigma_p \tag{9}$$

where w_i and w_j denote the weights of asset i and j in the portfolio of n assets, respectively. σ_{ij} denotes the covariance between returns of asset i and asset j, μ is the mean value of the returns of the portfolio and σ_p is the standard deviation of the total portfolio returns. The parameter z_α is the value of the cumulative normal distribution corresponding to the specific confidence level α, e.g. for the 95% confidence level $z_{95\%} = 1.64$ and for the 99% confidence level $z_{99\%} = 2.33$. Since the holding period is usually short (e.g. ten days) the assumption of a zero mean ($\mu = 0$) is often made. Thus, VaR of a portfolio is simply a multiple of the portfolio standard deviation. After calculating VaR, the calculation of CVaR is straightforward as the expected value of the portfolio losses exceeding VaR.

A major drawback with the delta-normal approach is the exposure to model risk. Even though normal distributions seem to describe the centre of *true* distributions rather well, problems arise when it comes to estimating the tails of distributions. Many empirical studies (see e.g. [17], [25], [26] and [32]) show that the assumption of normally distributed financial returns underestimates VaR. The underestimation becomes more significant when studying securities with heavy-tailed distributions and a high potential for large losses, i.e. that exhibit excess kurtosis [45]. In a similar fashion, Andersen et al. [6] show that accounting for heavy tails makes it possible to increase returns while lowering large risks. These empirical findings are intuitive since heavy tails mean that extreme outcomes are more frequent than what the use of a normal distribution would predict and therefore heavy tails lead to underestimated VaR measures.

Despite its drawbacks, the delta-normal approach is widely used among risk managers. For example, the RiskMetrics system is based on the parametric delta-normal model [40].

2.3.2 Historical Simulation

The most common and probably simplest non-parametric method to estimate VaR (and CVaR) is based on historical simulation. The main assumption is that trends of past price changes will continue in the future. The VaR (and CVaR) of a portfolio is then calculated using the percentile of the empirical distribution corresponding to the chosen confidence level. There is no need to estimate distribution parameters such as volatilities and correlation coefficients. The historical simulation method is relatively simple to implement, by keeping a historical record of past returns. The method is also free from model risk and makes it possible to accommodate the non-normal distributions with heavy tails that are often found in financial data [25].

The number of past observations to be included in the empirical distribution is often referred to as window size. The choice of window size has a significant impact on VaR measures, especially when using historical simulation [25]. A long window size may include observations that are not relevant to the current situation and may imply a fairly constant VaR measure. A short window size makes the calculations sensitive with respect to abnormal outcomes in the recent past and may imply high variance in VaR measures. Finansinspektionen recommends a window size of at least one year [23].

The historical simulation approach does not efficiently capture extreme events and pattern changes from historical behavior, but produces accurate results under normal market conditions [26]. Many large financial institutions and risk managers compute the VaR of their trading portfolios using the historical simulation approach, e.g. Goldman Sachs [24].

2.3.3 Alternative Approaches

Another widely used approach is the *Monte Carlo simulation*, where a future probability distribution is assumed and the behavior of asset prices is simulated by generating random price paths. The VaR (and CVaR) measures can then be determined from the distribution of simulated portfolio values. Monte Carlo frameworks have been shown to provide the best estimates for VaR (see e.g. [31] and [39]). However, at the same time, these models are extremely computer intensive and the additional information that these techniques provide is of most use for the analysis of complex options portfolios.

The *stress testing* method examines the effects of large movements in key financial variables on the portfolio value. The price movements are simulated in line with certain scenarios, such as movements of the yield curve, changes in exchange rates, etc. Portfolio assets are re-evaluated under each scenario and estimating a probability for each scenario allows to construct a distribution of portfolio returns, from which VaR (and CVaR) can be derived.

2.4 Backtesting

Since the *true* VaR measures cannot be observed, the evaluation of VaR models must be verified by backtesting. It means that, for a given backtesting period, the estimated VaR measures are compared to the observed returns [12]. The Basel Committee [10] defines backtesting as follows:

Backtesting is an ex-post comparison of the risk measure generated by the risk model against actual daily changes in portfolio value over longer periods of time, as well as hypothetical changes based on static positions.

2.4.1 Backtesting VaR

Regulatory requirements have motivated the development of backtesting theory for VaR models. However, there are several possible ways to backtest VaR models (see [28] and [30] for an overview). Typically, the number of times the portfolio loss exceeds VaR is calculated. For each backtesting period the number of violations are calculated. The number of violations divided by the number of observations in the backtesting period gives the violation rate, to be compared to the expected rate of violations. For example, VaR at the 95% confidence level has an expected rate of violations of 5%, and for VaR 99% the expected rate of violations is 1%.

The most widely used test is developed by Kupiec [28]. He examines whether the observed violation rate is statistically equal to the expected violation rate. Under the null hypothesis that the model is adequate, the appropriate likelihood ratio statistic is:

$$L = 2\ln\left(\left(1-\frac{n}{T}\right)^{T-n}\left(\frac{n}{T}\right)^{n}\right) - 2\ln\left((1-q)^{T-n}q^{n}\right) \sim \chi_1^2 \qquad (10)$$

where n is the number of days over a period T that a violation occurred and q is the expected violation rate. Therefore, the risk model is rejected if it generates too many or too few violations.

2.4.2 Backtesting CVaR

Despite the close theoretical relation to VaR, the backtests designed for VaR can not be used directly in backtesting of CVaR. Yamai and Yoshiba [45] consider backtesting of CVaR more complex than backtesting of VaR, and claim that it is one of the reasons for exclusion of CVaR from the Basel Committee framework.

To implement a backtesting procedure for CVaR, we specify a loss function ρ. A number of different loss functions have been suggested in the literature. An intuitive and easily interpreted loss function would compare the calculated CVaR to the actual return r in the cases where the r exceeds the VaR, i.e. the returns in the tail that are included in the calculation of CVaR:

$$\rho = \begin{cases} r & \text{if } r < \text{VaR} \\ 0 & \text{if } r \geq \text{VaR} \end{cases} \qquad (11)$$

Function (11) gives each tail-loss observation a weight equal to 1 and the corresponding benchmark is simply the CVaR.

Furthermore, we need a normalized measure of relative performance that can be used for backtesting over different asset classes. One of them is proposed by Blanco and Nihle [13] and Dowd [21]:

$$\rho = \begin{cases} \frac{r-\text{VaR}}{\text{VaR}} & \text{if } r < \text{VaR} \\ 0 & \text{if } r \geq \text{VaR} \end{cases} \qquad (12)$$

Function (12) gives each tail-loss observation a weight equal to the tail loss divided by the VaR, i.e. the measure is normalized for the VaR of the underlying asset. The benchmark is equal to the difference between CVaR and VaR divided by VaR. However, a potential problem with the loss function (12) is that VaR is in the denominator, and hence is not defined if VaR is zero. It may also give mischievous answers if VaR gets close to zero or changes sign.

2.5 Robustness

Most risk measures, such as VaR and CVaR, are defined as functions of the distribution of the considered return. However, since the probability measure describing market events is unknown the distinction between the theoretical risk measure and its estimator allows us to study the relation between the choice of the estimator and the specification of risk measures. In particular, it allows us to consider some natural requirements of the risk measurement procedure. For example, how robust is the result with respect to the data set or with respect to other parameters? Constructing and computing measures of sensitivity allows a quantification of the robustness of VaR and CVaR with respect to the data set and parameters used to compute them. However, VaR and CVaR have completely different properties: VaR is an estimate of a percentile in the distribution of returns, i.e. a single point in the distribution. CVaR, on the other hand, is the expected value of returns beyond the VaR percentile, i.e. an estimate that takes all points beyond the VaR percentile into account. Comparing them directly is like comparing apples and oranges.

For a risk model to be considered robust, it should provide accurate risk forecasts across different assets, time horizons, and confidence levels. Fluctuations in risk forecasts have serious implications for the usefulness of a risk model. There are various definitions of a *robust statistic* in the literature. However, strictly speaking, a robust statistic is resistant to errors in the results, produced by deviations from assumptions (e.g. of normality). This means that if the assumptions are only approximately met, the robust estimator will still have a reasonable accuracy and a reasonably small bias, as well as being asymptotically unbiased, i.e. having a bias tending towards zero as the sample size tends towards infinity.

In order to quantify the robustness of an estimator for the purpose of this thesis, it is necessary to define some measures. Rousseeuw and Croux [42] and Mizera and Müller [37] propose the characteristics *bias* and *interquartile range* for quantifying the robustness of an estimator.

The bias of an estimator is the distance between the average of the collection of estimates, and the single parameter being estimated. Mathematically, it is defined as

$$\mathrm{Bias}[\hat{\rho}] = \mathrm{E}[\hat{\rho}] - \rho = \mathrm{E}[\hat{\rho} - \rho] \tag{13}$$

where $\hat{\rho}$ is an estimator of parameter ρ
Furthermore, an estimator is said to be *unbiased* if its bias is equal to zero for all values of parameter ρ.

Statistical dispersion is the variability or spread in a variable or a distribution. A measure of statistical dispersion is a real number that is zero if all the data are identical, increases as the data becomes more diverse and cannot be less than zero. A common measure of statistical dispersion is the inter-quartile range (IQR), defined as the difference between two percentiles of a sample.

3 Data

In this section the selection and collection of the data sample used in the thesis are presented. A descriptive overview of the collected data is also given.

For the purpose of this thesis Datastream is used to gather time series data for equity indices, bond indices, exchange rates and individual stocks. Datastrem is a comprehensive online historical database service provided by Thompson Financial, which is a globally leading supplier of financial information. However, Thomson Financial gives no warranty as to its accuracy, completeness or correctness. Still, data contained in Datastream has been compiled by good faith from sources believed to be reliable and seems like an appropriate database for the type of information used in this study.

Daily prices are collected for the 12 year period from January 1, 1996 to December 31, 2007. The starting date is the earliest available that holds for all variables in the data set. Furthermore, Yamai and Yoshiba [45] concludes that both VaR and CVaR are less reliable during periods of market turmoil and under such circumstances tend to give overly optimistic results. Due to the global financial turmoil during 2008 and 2009, and to avoid unnecessary reliability issues and too optimistic results, the scope of this thesis excludes any data post December 31, 2007.

The data used in this thesis consists of a variety of major international equity and bond indices as well as major exchange rates and individual stocks. More specifically, 42 international equity indices from Europe (excluding Sweden), America (excluding the US), Asia-Pacific and Africa are used. Most indices were listed on Yahoo Finance [49] as *major world indices*. To complete the set, additional 13 US market equity indices and 39 Swedish market equity indices are used. Moreover, a data sample of 20 major bond indices, 17 major exchange rates and 15 individual international stocks (five stocks from the Stockholm Stock Exchange, five stocks from the MICEX index, which comprises the most liquid Russian stocks, and finally five stocks from the Dow Jones Industrial Average index in New York) are used. The analysis is restricted to simplest possible portfolios consisting of a single asset, i.e. equity or bond index, exchange rate or individual stock.

Equity prices are adjusted for dividends, share repurchases and share issues. Non-trading days are excluded from the data set.

An example of a portfolio return distribution over time as well as VaR and CVaR estimates is given for Affärsvärldens Generalindex in Figure (3). As expected CVaR is always less than or equal to VaR.

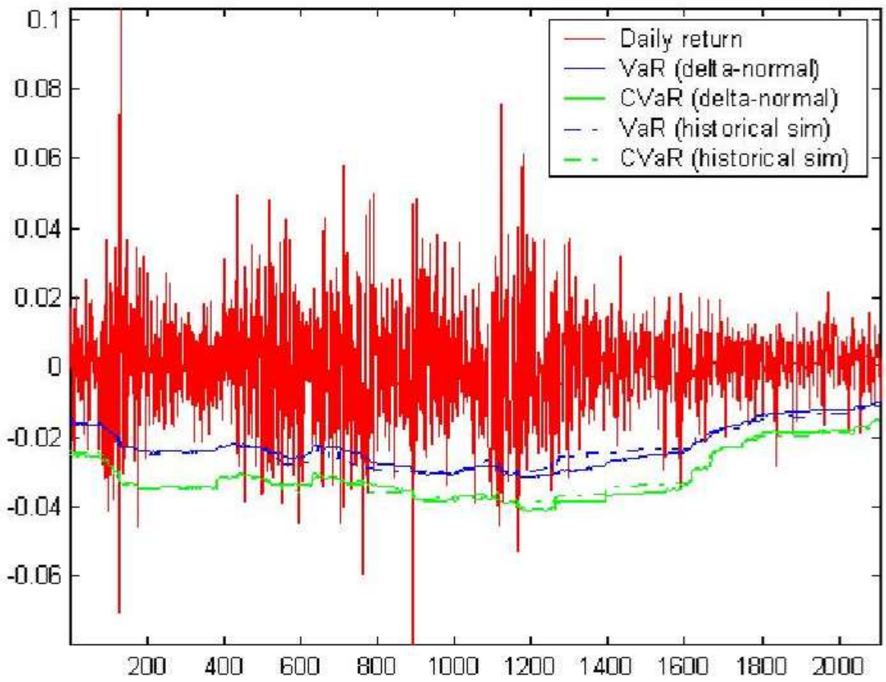

Figure 3: Affärsvärldens Generalindex

4 Analysis

In this section, the methodology is outlined, as well as empirical measures and tests that are used in the thesis. Furthermore, two hypotheses about the robustness of CVaR is presented. The first about the robustness within all examined asset classes and the second about the difference in robustness between different asset classes

4.1 Returns

Throughout the thesis the daily returns are calculated as:

$$r_t = \ln \frac{P_t}{P_{t-1}}$$

where r_t is the daily return, P_t is the closing price on day t and P_{t-1} is the closing price on day $t-1$. In other words, the standard in financial analysis is followed, using log-returns.

4.2 Methodology

By using data in the ex ante time period, we can calculate VaR and CVaR using one of several possible methods, which are described in section 2.3. Thereafter, the accumulated return is observed, τ_H days after the end of the ex ante period, where τ_H is the VaR horizon or equivalently, the hypothesized holding period. If the return is less than VaR, the event is counted. This is later on used to evaluate VaR. On average, $100(1-\alpha)$ returns should be observed less than VaR, if VaR is a good measure and α is the confidence level of VaR. Furthermore, the empirical measures to evaluate CVaR are calculated. The calculation of these empirical measures are described in section 4.3.

The next step is to move the ex ante window one day forward and repeat the steps described above, starting with calculating updated values of VaR and CVaR for the new ex ante time period. This procedure is than repeated until the end of the date range. Thereafter, the algorithm is repeated for all the assets, the different methods of calculating VaR and the different parameters under study. The variable parameters are the VaR horizon τ_H, the ex ante window length and the confidence level.

4.3 Empirical Measures

In order to capture the behavior of VaR and CVaR as well as possible we defined appropriate empirical measures in section 2.4.1 and 2.4.2. Going forward, we will call the loss functions for CVaR backtesting *CVaR relative to return, CRR* (11) and *CVaR relative to VaR, CRV* (12).

4.3.1 CVaR Relative to Return (CRR)

For each sample of returns[4], the ex ante VaR and CVaR are calculated based on data in the ex ante period. By definition, CVaR is always less than VaR. As a second step, the actual return (loss) is examined. If the return is higher than VaR, nothing is recorded and we continue with the next step in the algorithm and form a new ex ante period. However, if the return is negative (i.e. a loss) and abnormal, i.e. less than VaR[5], the value is recorded. As a last step, the difference between the actual return and the calculated CVaR is computed. Thereafter, we form a new ex ante period one step forward in time and the algorithm is repeated from start until all samples of returns have been examined. Finally, the different values for each sample of returns for each asset are aggregated into a scalar measure, the mean value.

An example can be used to illustrate CRR. If the ex ante CVaR is -15% and the return is -14%, then CRR becomes $-14\% - (-15\%) = 1\%$. Hence, the return is 1% unit larger than CVaR.

An advantage of CRR compared to CRV is that its unit is more easily interpreted.

4.3.2 CVaR Relative to VaR (CRV)

Again, using the calculated ex ante VaR and CVaR, the value

$$\rho_1 = \frac{\text{CVaR} - \text{VaR}}{\text{VaR}}$$

which measures the magnitude of CVaR compared to VaR, is recorded. As a second step, the actual return (loss) is examined. If the return is higher than VaR, nothing is recorded and we continue with the next step in the algorithm and form a new ex ante period. However, if the return is negative (i.e. a loss) and abnormal, i.e. less than VaR, the value

$$\rho_2 = \begin{cases} \frac{\text{return} - \text{VaR}}{\text{VaR}} & \text{if return} < \text{VaR} \\ 0 & \text{if return} \geq \text{VaR} \end{cases} \tag{14}$$

[4] e.g. each trading day if the VaR horizon is one trading day

[5] inherent in the VaR and CVaR concepts are that returns are considered negative and abnormal when they are less than VaR for the relevant confidence level

which measures how much smaller the return is compared to VaR, is recorded. As a last step, the difference

$$\rho_3 = \rho_2 - \rho_1$$

is calculated, which is a measure of the difference between CVaR and the abnormal negative return. Thereafter, we form a new ex ante period one step forward in time and the algorithm is repeated from start until all samples of returns have been examined. The algorithm leaves us with a number of different ρ_3 for each asset. More specifically, there should be a number of ρ_3 equal to approximately $1 - \alpha$ times the number of samples. All ρ_3 related to a specific asset are aggregated into a scalar measure, the mean value.

An example can be used to illustrate CRV. If the ex ante VaR and CVaR are -10% and -15%, respectively, and the return is -14%:

$$\begin{aligned}
\rho_1 &= \tfrac{(-15)-(-10)}{(-10)} = 50\% \\
\rho_2 &= \tfrac{(-14)-(-10)}{(-10)} = 40\% \\
\rho_3 &= \rho_2 - \rho_1 = 40\% - 50\% = -10\%
\end{aligned}$$

The interpretation is that on this occasion, we see an abnormal negative return (since the return is less than VaR) and the return is 10% of VaR higher (less negative) than CVaR.

An advantage of CRV compared to CRR is that CRV by design is normalized for different volatilities of the different assets. The same measure was proposed by Blanco and Nihle [13] and Dowd [21].

4.4 Empirical Tests

Two different tests are performed in this thesis to evaluate CVaR. Let us call them the *CVaR robustness test* and the *Asset class difference test*.

4.4.1 CVaR Robustness Test

Hypothesis 1: CVaR is more robust than VaR
The hypothesis is that CVaR is a more robust risk measure than VaR. The reason is that all values in the tail of the return distribution are considered when estimating CVaR, compared to just the number of values for the case of VaR. For example, if the tail consists of the returns -8%, -10% and -12%, all these values are taken into account when estimating CVaR. When VaR is estimated, the most important feature about the tail is that it consists of three different values. It should be noted, that it is the tail of the distribution that is important for most risk measures, since the tail in some sense defines

the losses, and hence the risk. The variance of the estimate of a mean (e.g. CVaR) should, intuitively, be less than the variance of the estimation of a single point (e.g. VaR), and therefore the variance should be lower for CVaR compared to VaR. Hence CVaR should be a more robust measure of risk.

CVaR is estimated based on two different methods, the first assuming normally distributed returns (the delta-normal approach, see 2.3.1) and the second assuming that the distribution of returns in the ex ante period is representative for the distribution of future returns (the historical simulation method, see 2.3.2).

A formal test is performed where it is tested whether CVaR is an unbiased estimator[6] of the conditional return, given that the return is less than VaR. The test is performed by testing the null hypothesis H_0 that the empirical measure CRR is zero on average against the alternative hypothesis H_1 that it differs from zero. CRR is by design equal to zero if CVaR is an unbiased estimator of the conditional return.

To examine the variation of the empirical measures the inter-quantile range (IQR) is calculated, which measures the difference between two quantiles in the distribution of an empirical measure. E.g. $\mathrm{IQR}_{0.95-0.05}$ measures the difference between the five and ninety five percentile. If CVaR is a robust risk measure, IQR should be reasonably small. However, a formal statistical test based upon the variation or IQR is not captured within the scope of this thesis.

4.4.2 Asset Class Difference Test

Hypothesis 2: CVaR is robust over different asset classes
Since CVaR is estimated based on two different methods, the delta-normal approach and the historical simulation method, the asset class that, on average, has returns that are most similar to a normal distribution will seem to be the most robust asset class, when CVaR is estimated using the delta-normal method. On the other hand the asset class that, on average, is most constant over time, will seem to be the most robust asset class when CVaR is estimated using the historical simulation method. It is without further studies difficult to say which asset class that has returns that are more similar to a normal distribution or which asset class that has a pattern of returns that is the most constant over time.

The null hypothesis H_0 will be tested, that the different asset classes, pairwise, on average have the same value of the empirical measure CRV (cf. section 4.3) against the alternative hypothesis H_1 that they differ. If CVaR

[6]In this thesis, an estimator will be considered unbiased (and hence robust) if the bias is arbitrary close to zero.

is a robust measure of risk, there should be no difference over the different asset classes in terms of robustness (though the actual risk will of course differ) if CVaR is an unbiased risk measure for those particular asset classes.

In this test, the different asset classes are ranked according to the empirical measure CRV. The reason that only CRV is considered in this test is that it is the only empirical measure that is adjusted for the volatility of the underlying asset (cf. section 4.3). By ranking the different assets, it should be possible to draw the conclusion whether CVaR is a robust risk measure by examining the difference in CRV between the different asset classes. If CVaR is a robust measure of risk, it should be transparent to the underlying type of asset, in the sense that a measure of robustness should not be different for different asset classes. Of course, the value of CVaR itself will vary significantly between different assets since e.g. T-bills are less risky than a stock in a mineral company listed on the Moscow Stock Exchange. A simple 2-sample t-test where the variances of the two populations are not assumed equal is performed to investigate the pair-wise difference in CRV between the different asset classes.

5 Empirical Findings

In this section the empirical findings and results of the study are presented. In the first part, all asset classes are treated jointly, whereas in the second part, they are treated individually. The different asset classes considered are stock indices, bond indices, exchange rates and individual stocks.

5.1 Parameters

As described in section 2.3, CVaR (and VaR) can be estimated by different methods. However, in all methods, some parameters always have to be decided beforehand. In our case, the parameters are the ex ante estimation window length, the time horizon and the confidence level.

The *ex ante estimation window length* is chosen to 250 or 500 trading days, corresponding to approximately one year and two years, respectively. The ex ante estimation window is used to estimate the model parameters, e.g. the standard deviation of the returns if that is an input to the model. Lambadiaris et al. [29] use an ex ante length of 100 or 252 trading days. They conclude that a longer estimation window is usually better. The reason that they do not use a longer ex ante window is a restriction in the number of samples. Finansinspektionen suggests an ex ante window length of at least one year [23].

The *confidence level* of VaR and CVaR is chosen to 95% or 99%. These are the most common levels in the literature. Furthermore, Finansinspektionen suggests a confidence interval of at least 99% [23].

The VaR and CVaR *time horizon* is chosen to be one day. This is a common interval in the literature. The time horizon is the same as the hypothetical holding period and hence the relevant return is the return during the VaR horizon period. In our case, where we study market risk, the VaR horizon is typically one trading day. However, Finansinspektionen suggests a VaR horizon of ten days [23]. On the other hand, they also say that it is equally good to perform all calculations assuming a one-day horizon and then as a final step calculate the final ten-day horizon VaR (or CVaR) value from the one-day horizon value through a simple transformation.

5.2 CVaR Robustness

The main empirical results are presented in Tables 3 - 8. The complete set of empirical findings is found in Appendix A.

The entries in Tables 3 - 4 should be interpreted as follows: Consider e.g. the second line, the first column tells us that ex ante window has a length of 500 trading days and the confidence level is 95%. The second column in Table 3 tells us that the average CRR over all assets is -0.008%. This means that if CVaR is e.g. -10%, the conditional return, given that the return is less than VaR, is -10.008% on average. Columns 3 - 6 in Table 3 and columns 2 - 3 in Table 4 are different measures of the deviation of CRR about its mean, where Q stands for quantile and IQR for inter-quantile range. The last two columns in Table 4 show the result of the test whether the mean CRR is equal to zero (accepted or rejected), and the associated t-statistic is given in column 4 in Table 4.

Table 3: Empirical results of CRR using the delta-normal method

ex ante length / confidence level	mean CRR	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
500 / 95%	-0.008	-0.250	-0.200	0.140	0.170
500 / 99%	-0.031	-0.900	-0.440	0.310	0.370
250 / 95%	-0.067	-0.300	-0.260	0.067	0.093
250 / 99%	-0.132	-1.150	-0.710	0.160	0.170

Table 4: (cont'd) Empirical results of CRR using the delta-normal method

ex ante length / confidence level	$IQR_{0.95-0.05}$	$IQR_{0.975-0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
500 / 95%	0.340	0.420	-0.083	acc	acc
500 / 99%	0.740	1.270	-0.068	acc	acc
250 / 95%	0.320	0.390	-0.430	acc	acc
250 / 99%	0.870	1.320	-0.297	acc	acc

Table 5: Empirical results of CRR using the historical simulation method

ex ante length / confidence level	mean CRR	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
500 / 95%	-0.032	-0.300	-0.230	0.110	0.140
500 / 99%	-0.287	-1.810	-1.520	0.170	0.240
250 / 95%	-0.110	-0.420	-0.300	0.002	0.009
250 / 99%	-0.425	-1.780	-1.300	-0.031	-0.004

We note that the hypothesis H_0 that the empirical measures are equal to zero, and hence that CVaR is an unbiased estimate of the conditional return, given that the return is less than VaR, is accepted in most cases. In fact, it is rejected only in two cases (see Appendix A) at a relatively low 90% confidence level. In both cases, the ex ante window length is 250 days, which

Table 6: (cont'd) Empirical results of CRR using the historical simulation method

ex ante length confidence level	IQR 0.95-0.05	IQR 0.975-0.025	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
500 / 95%	0.340	0.450	-0.295	acc	acc
500 / 99%	1.690	2.050	-0.419	acc	acc
250 / 95%	0.300	0.430	-0.918	acc	acc
250 / 99%	1.270	1.770	-0.725	acc	acc

is in line with the results of Lambadiaris et al. [29], who concluded that a longer ex ante window length gives a better estimate. It is also noticable, that both rejections occur when using the historical simulation method.

To further investigate whether CVaR is an unbiased measure, we study CRR of the individual assets and perform a similar test as above to test whether CRR is equal to zero or not. The test is two-sided and performed on a 95% confidence level. The summary of the results are shown in Table 7. CVaR seems to be an unbiased estimate even in the case of individual assets. The parameters of the CVaR estimator in this case are an ex ante window length of 500 trading days and a confidence level of 95%. The rationale behind this choice is that these parameters seem to render the most stable CVaR estimates, as shall be shown subsequently.

Table 7: Summary of the test whether CRR for individual assets are equal to zero. The table shows the number of assets (and % of times) in each asset class for which H_0 is rejected

Asset class	# assets	delta-normal	historical simulation
Stock indices	100	1 (1%)	4 (4%)
Stocks	14	0	0
Bonds	19	0	1 (5%)
Exchange rates	0	0	0

In Table 8 we perform a backtesting of VaR. In the table the proportions of returns less than VaR are stated as well as the corresponding p-values calculated from the binomial test described in section 2.4.1. As we can see, the null hypothesis of binomially distributed VaR values can be rejected at a 5% significance level in three of the cases. All rejections refer to the delta-normal approach. It seems like the historical simulation does a better job in estimating VaR values. In fact, the backtesting shows encouraging results when using the historical simulation method. A possible explanation could be that the assets in the data sample show fat-tail properties, and hence do not exhibit the normality characteristic assumed in the delta-normal approach.

Table 8: Back-testing of VaR: violation ratios and p-values

ex ante length / confidence level	500 / 95	500 / 99	250 / 95	250 / 99
Delta-normal method	4.1 **(0.028)**	1.5 **(0.023)**	4.5 (0.190)	1.6 **(0.004)**
Historical simulation	4.7 (0.460)	0.9 (0.676)	5.2 (0.620)	1.3 (0.214)

Furthermore, also the IQR indicates that CVaR is a robust measure of risk. The IQR increases as the confidence level increases, which is probably due to the fact that there are fewer samples to base the estimate on, compared to the lower confidence level. On average, there are five times as many samples of the return being less than VaR at the 95% confidence level compared to the 99% confidence level.

For illustrative purposes, consider the $\mathrm{IQR}_{0.95-0.05}$, the largest value is found for the historical simulation method with an ex ante length of 500 days and a confidence level of 99% (see Appendix A). The interpretation of this value would be as follows: the 5-percentile is -1.66% and the 95-percentile is 0.15%. If CVaR is, e.g. -10%, the true conditional return, given that the return is less than VaR would be between -11.66% and -9.85%, where *true* is defined as lying between the 5 and 95 percentiles. This may not appear as a narrow interval, but it should be considered that this is the worst outcome of all tests. Using a confidence level of 95% instead, the $\mathrm{IQR}_{0.975-0.025}$ is always less than 0.47%, no matter what estimation method and ex ante window length being used.

5.3 Asset Class Difference

Another way to test the robustness of CVaR is to compare the empirical measure CRV over different asset classes. If CVaR is robust, CRV should not differ over different asset classes. The test is performed as a difference-in-mean test where the null hypothesis H_0 that there is no difference between different asset classes is tested against H_1 that there is a difference in the mean of CRV between different asset classes. The number of degrees of freedom (DF) is approximated by the number of assets in the asset class with the lowest number of assets, minus one.

The result of the test is presented in Tables 9 - 10. In the tables, the t-statistic and DF are presented. The critical values are 2.1009 and 2.1604 at the 95% confidence level for 18 and 13 df, respectively. In the tables, a t-statistic above the critical value at the 95% confidence level is indicated with a boldface typesetting.

In fact, there seems to be a significant difference in CRV between different

Table 9: Difference-in-mean test. Delta-normal method for certain ex ante lengths (days) and confidence levels (%).

Assets	t-stat 500 / 95	t-stat 500 / 99	t-stat 250 / 95	t-stat 250 / 99	DF
stock indices - bonds	-1.483	-1.205	-0.016	-0.534	18
stock indices - fx rates	-0.191	-0.787	**5.455**	**2.317**	13
bonds - fx rates	1.426	0.614	1.233	1.546	13
stocks - fx rates	1.764	1.763	**3.125**	**2.488**	13
stocks - stock indices	1.883	2.126	0.517	1.559	13
stocks - bonds	-0.414	1.283	0.202	0.777	13

Table 10: Difference-in-mean test. Historical simulation method for certain ex ante lengths (days) and confidence levels (%).

Assets	t-stat 500 / 95	t-stat 500 / 99	t-stat 250 / 95	t-stat 250 / 99	DF
stock indices - bonds	-1.939	-0.706	**3.790**	-1.187	18
stock indices - fx rates	-0.671	-0.041	**4.214**	1.375	13
bonds - fx rates	0.687	0.632	-0.696	2.006	13
stocks - fx rates	0.894	1.843	**3.092**	2.450	13
stocks - stock indices	1.569	2.011	0.375	1.885	13
stocks - bonds	-0.250	1.187	**3.110**	1.039	13

asset classes in the case of an ex ante window length of 250 days. On the other hand, when the ex ante window length increases to 500 days, most of the differences seem to disappear. This can be interpreted as the robustness of CVaR increases when the ex ante window length increases, which is in line with previous findings. The CVaR robustness seems sensitive to a short ex ante window.

Furthermore, the significant difference between the asset classes might be interpreted as there is a bias in the CVaR estimator. However, by analysing the data, we find that one part of the explanation to the difference is that there are a few individual stocks with extreme movements, mainly from the Moscow stock exchange. These stocks have a significant impact on the average CRV for the stock category, since the number of individual stocks in this thesis is rather limited. Hence, another interpretation of the results is that CVaR is not a good risk measure for certain asset classes, in this case individual stocks, due to the return distribution of that particular asset class. This is also in line with the previous findings where we saw that CVaR sometimes seem to be a biased estimate for individual stocks. However, if there is a bias with a 250 day ex ante window length, it seems to go away as the ex ante window length increases to 500 days.

6 Concluding Remarks

This section summarizes the findings and provides suggestions for future research.

In the recent days with turbulence on every major stock exchange, it is evident that controlling the risks in investment strategies is an important issue for the entire global economy. Perhaps there is no such thing as a golden rule on how to manage a portfolio, but history shows that focusing too much on the return is risky business. In the end of last decade, a risk-measure called Conditional Value-at-Risk (CVaR) was introduced to the market. It was the successor of a measure called Value-at-Risk (VaR), which caught the interest of the market, but had faced problems not being sub-additive, which is an important feature in the financial world. CVaR, with more attractive theoretical properties, has therefore been gaining ground in the last few years. CVaR is being used by insurance companies, mutual funds and other participants in the financial market who have the need of evaluating their risks.

Firstly, it seems plausible that CVaR is an unbiased estimate of the conditional return, given that the return is less than VaR, with the possible exclusion of individual stocks. The empirical findings support the hypothesis that CVaR is a robust risk measure.

Secondly, the results indicate that it is possible to construct a robust CVaR by estimating the input parameters carefully. This means choosing a confidence level that is not too high, since a certain number of samples is needed in the ex ante window to estimate the model parameters accurately. It also means that the ex ante window has to be chosen long enough, probably due to the same reason.

As the confidence level of VaR increases, the robustness seems to decrease. This is a problem, since a high confidence level of VaR is usually desirable. A 95% confidence level would mean that we consider approximately one trading day per month as being abnormal. A 99% confidence level would mean that we consider two to three trading days per year as being abnormal. When the confidence level increases, the number of relevant samples in the ex ante window decreases so that the CVaR estimate gets worse. This can probably to some extent be compensated by extending the ex ante window length, but this may be difficult for practical reasons, since there is often a lack of relevant historical data. Even if there is enough historical data available, it may not be representative due to its age.

6.1 Suggestions for Future Research

The results in this thesis indicate that the robustness of CVaR increases as the ex ante window length increases and the confidence level decreases. This might be due to the fact that the number of historical samples in the ex ante window increases, which renders a better CVaR estimate. It would be interesting to find out approximately how many samples are needed for a good CVaR estimate and if a longer ex ante window length can be directly traded for a higher confidence level in terms of robustness. Or is old historical data less useful than more recent data? What is the trade-off between recent and old data?

Inherent in the CVaR estimation process is the estimation of VaR. This might transfer some of the robustness issues of VaR onto CVaR. It would be interesting to try to isolate the evaluation of the CVaR robustness from VaR. In our study, we evaluate CVaR every time the return is less than VaR, but since there are issues with VaR, this might not be the best thing to do. One possible alternative might be to evaluate CVaR for e.g. the worst five percent of the returns, if the confidence level is 95 %.

References

[1] Acerbi, C; Nordio, C; Sirtori, C. Expected shortfall as a tool for financial risk management. *Technical report*, 2001.

[2] Acerbi, C; Tasche, D. Expected shortfall: A natural coherent alternative to Value at Risk, *Working paper*, 2001.

[3] Acerbi, C; Tasche, D. On the coherence of expected shortfall, *Working Paper*, Department of Mathematics, TU-Munchen, 2001.

[4] Acerbi, C. Spectral measures of risk: A coherent representation of subjective risk aversion, *Working Paper*, 2002.

[5] Alexander, G.J; Baptista A.M. A Comparison of VaR and CVaR Constraints on Portfolio Selection with the Mean-Variance Model. *Management Science*, Vol 50, No 9, September 2004, pp 1261-1273.

[6] Andersen, J; Sornette, D. Have your cake and eat it too: increasing return, *Working Paper*, 1999.

[7] Artzner, P; Delbaen, F; Eber, J.M; Heath, D. Thinking coherently, *Risk*, Vol 10, 1997, pp 68-71.

[8] Artzner, P; Delbaen, F; Eber, J.M; Heath, D. Coherent measures of risk, *Math Finance*, 9(3), 1999, pp 203-228.

[9] Basak, S; Shapiro, A; Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices. *Review of Financial Studies*, Vol 14, pp. 371-405, 2001.

[10] Basel Committee on Banking Supervision, Amendment to the capital accord to incorporate market risks, *Basel Committee Publications*, Bank for International Settlements, January 1996, p 1.

[11] Basel Committee on Banking Supervision, Analysis of the trading book quantitative impact study, *Basel Committee Publications*, Bank for International Settlements, October 2009, p 12.

[12] Best, P; *Implementing Value At Risk*, John Wiley & Sons, 1999.

[13] Blanco, C; Ihle, G. How good is your VaR? Using backtesting to assess system performance, *Financial Engineering News*, August 1998.

[14] Bodie; Kane; Marcus. *Investments*, 6th edition, McGraw-Hill, New York, 2005.

[15] Boyle, P; Hardy, M; Vorst, T; Life after VaR, *The Journal of Derivatives*, Fall 2005, pp 48-55.

[16] Citibank, Annual Report, 2008, p 38.

[17] Danielsson, J; de Vries, C. Value-ar-Risk and Extreme Returns, *FMG Discussion Paper*, No 273, Financial Markets Group, London School of Economics, 1997.

[18] Daníelsson, J; Keating, C; Shin, H.S; Muennich, F; Goodhart, C. An Academic Response to Basel II. *FMG Special Papers*, Sp 130, Financial Markets Group, 2001.

[19] Daníelsson, J; de Vries, C.G; Jorgensen, B.N; Mandira, S. Comparing Downside Risk Measures for Heavy Tailed Distributions, *FMG Discussion Papers*, dp 551, Financial Markets Group, 2005.

[20] Dowd, K; Beyond Value At Risk: The New Science of Risk Management, John Wiley & Sons, 1998.

[21] Dowd, K. Retrospective assessment of Value-at-Risk, John Wiley & Sons, 1999.

[22] Duffie, D; Pan, J. An Overview of Value at Risk, *Journal of Derivatives*, Vol 4, 1997, pp 7-49.

[23] *Intern VaR-modell för beräkning av kapitalkrav för marknadsrisk*, Finansinspektionen, April 2004.

[24] Goldman Sachs, Annual Report, 2008, p 51.

[25] Goorbergh, R; Vlaar, P. Value-at-Risk Analysis of Stock Returns Historical Simulation, Variance Techniques or Tail Index Estimation?, *De Nederlandsche Bank Staff Reports*, No 40, March 1999.

[26] Hendricks, D. Evaluation of value-at-risk models using historical data, *Economic Policy Review*, Federal Reserve Bank of New York, April 1996.

[27] JPMorgan, Annual Report, 2008, p 76.

[28] Kupiec, P.H. Techniques for Verifying the Accuracy of Risk Measurement Models, *Journal of Derivatives*, Vol 3, 1995, pp 73-84.

[29] Lambadiaris, G; Papadopoulou, L; Skiadopoulos, G; Zoulis, Y. VaR: History or simulation. *Risk Magazine* September 2003, pp 123-126.

[30] Lopez, J.A. Methods for Evaluating Value-at-Risk Estimates. Federal Reserve Bank of New York, *Economic Policy Review*, Vol 2, 1999, pp 3-17.

[31] Lucas, A; Klaassen, P. Extreme returns, downside risk, and optimal asset allocation. *Journal of Portfolio Management*, Vol 25, pp. 71-79, 1998.

[32] Mahoney, J. *Empirical-based versus Model-based Approaches to Value-at-Risk: An Examination of Foreign Exchange and Global Equity Portfolios*, Proceedings of a Joint Central Bank Research Conference, Board of Governors of the Federal Reserve System, 1996, pp 199-217.

[33] Markowitz, H.M. Portfolio Selection, *Journal of Finance*, Vol 7, No 1, pp 77-91, 1952.

[34] Martín Mato, M.A. Classic and modern measures of risk in fixed-income portfolio optimization. *The Journal of Risk Finance*, Vol 6, No 5, 2005, pp 416-423.

[35] Mausser, H; Rosen, D. Beyond VaR: From measuring risk to managing risk, *ALGO Research Quarterly*, Vol 1-20, 1999.

[36] McKay, R; Keefer, T.E. *VaR is a dangerous technique.* Corporate Finance Searching for Systems Integration Supplement, September:30, 1996.

[37] Mizera, I.; Müller, C.H. Location-scale depth, *Journal of the American Statistical Association*, Vol 99, 2004, pp 949–966.

[38] Olszewski, Y. Building a Better Fund of Hedge Funds: A Fractal and α - Stable Distribution Approach, *Working Paper*, draft #5, December 2005.

[39] Pritsker, M. Evaluating value at risk methodoligies: Accuracy versus computational time, *Journal of Financial Services Research*, Vol 12, 1997, pp 201-242.

[40] J.P. Morgan / Reuters, *RiskMetrics - Technical Document*, New York, 1994.

[41] Rootzén, H; Klüppelberg, C. A single number can't hedge against economic catastrophes. *Ambio*, 28(6), pp 550-555, Royal Swedish Academy of Sciences, 1999.

[42] Rousseeuw, P.J.; Croux, C. Alternatives to the Median Absolute Deviation, *Journal of the American Statistical Association*, Vol 88, December 1993, pp 1273–1283.

[43] Smithson, C. Beyond VAR, *Risk*, Vol 13, No 12, , December 2000, pp 85-87.

[44] Uryasev, S; Rockafellar, R.T. Optimization of Conditional Value-at-Risk, *The Journal of Risk*, Vol 2, No 3, 2000, pp 21-41.

[45] Yamai, Y; Yoshiba, T. Comparative Analysis of Expected Shortfall and Value-at-Risk: Their Validity under Market Stress, *Working Paper*, October 2002.

[46] Zikovic, S; Filer, R.K. Hybrid historical simulation VaR and ES: Performance in developed and emerging markets, *CESinfo Working Paper*, No 2820, Category 12: Empirical and theoretical methods, October 2009.

[47] Basel II: Revised international capital framework, Bank for International Settlements
http://www.bis.org/publ/bcbsca.htm

[48] Kapitaltäckning (Basel II), Swedish Financial Supervisory Authority (Finansinspektionen)
http://www.fi.se/Templates/ListPage____2914.aspx

[49] Yahoo Finance, Major world indices
http://uk.finance.yahoo.com/m2

A Empirical Findings - CVaR Robustness

In this appendix, we present the numerical results discussed in section 5.

Table 11: Empirical results using the delta-normal method. Ex ante length = 500 days, confidence level = 95%

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	1.32	-5.52	-4.58	7.98	10.95
CRR	-0.01	-0.25	-0.20	0.14	0.17

Table 12: (cont.) Empirical results using the delta-normal method. Ex ante length = 500 days, confidence level = 95%

	$IQR_{0.95\text{-}0.05}$	$IQR_{0.975\text{-}0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	12.60	16.50	0.26	acc	acc
CRR	0.34	0.42	-0.08	acc	acc

Table 13: Empirical results using the historical simulation method. Ex ante length = 500 days, confidence level = 95%

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	1.98	-6.78	-5.28	8.93	11.32
CRR	-0.04	-0.30	-0.23	0.11	0.14

Table 14: (cont.) Empirical results using the historical simulation method. Ex ante length = 500 days, confidence level = 95%

	$IQR_{0.95\text{-}0.05}$	$IQR_{0.975\text{-}0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	14.20	18.10	0.33	acc	acc
CRR	0.34	0.45	-0.30	acc	acc

Table 15: Empirical results using the delta-normal method. Ex ante length = 500 days, confidence level = 99%

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	0.59	-9.85	-8.67	10.74	15.20
CRR	-0.03	-0.90	-0.44	0.31	0.37

Table 16: (cont.) Empirical results using the delta-normal method. Ex ante length = 500 days, confidence level = 99%

	$IQR_{0.95-0.05}$	$IQR_{0.975-0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	19.40	25.10	0.09	acc	acc
CRR	0.74	1.27	-0.07	acc	acc

Table 17: Empirical results using the historical simulation method. Ex ante length = 500 days, confidence level = 99%

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	6.80	-7.76	-6.55	24.28	26.45
CRR	-0.29	-1.81	-1.52	0.17	0.24

Table 18: (cont.) Empirical results using the historical simulation method. Ex ante length = 500 days, confidence level = 99%

	$IQR_{0.95-0.05}$	$IQR_{0.975-0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	30.80	34.20	0.73	acc	acc
CRR	1.69	2.05	-0.42	acc	acc

Table 19: Empirical results using the delta-normal method. Ex ante length = 250 days, confidence level = 95%.

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	3.98	-3.82	-1.64	10.25	13.60
CRR	-0.07	-0.30	-0.26	0.07	0.09

Table 20: (cont.) Empirical results using the delta-normal method. Ex ante length = 250 days, confidence level = 95%.

	$IQR_{0.95-0.05}$	$IQR_{0.975-0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	11.90	17.40	0.76	acc	acc
CRR	0.32	0.39	-0.43	acc	acc

Table 21: Empirical results using the historical simulation method. Ex ante length = 250 days, confidence level = 95%.

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	5.41	-0.17	0.39	10.78	12.43
CRR	-0.11	-0.42	-0.30	0.00	0.01

Table 22: (cont.) Empirical results using the historical simulation method. Ex ante length = 250 days, confidence level = 95%.

	$IQR_{0.95-0.05}$	$IQR_{0.975-0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	10.40	12.60	1.65	acc	**rej**
CRR	0.30	0.43	-0.92	acc	acc

Table 23: Empirical results using the delta-normal method. Ex ante length = 250 days, confidence level = 99%

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	4.59	-4.51	-3.75	20.15	21.29
CRR	-0.13	-1.15	-0.71	0.16	0.17

Table 24: (cont.) Empirical results using the delta-normal method. Ex ante length = 250 days, confidence level = 99%

	$IQR_{0.95-0.05}$	$IQR_{0.975-0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	23.90	25.80	0.65	acc	acc
CRR	0.87	1.32	-0.30	acc	acc

Table 25: Empirical results using the historical simulation method. Ex ante length = 250, confidence level = 99%

	mean	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.95}$	$Q_{0.975}$
CRV	13.15	3.81	4.25	28.16	31.31
CRR	-0.43	-1.78	-1.30	-0.03	-0.00

Table 26: (cont.) Empirical results using the historical simulation method. Ex ante length = 250, confidence level = 99%

	$IQR_{0.95-0.05}$	$IQR_{0.975-0.025}$	t-stat	$H_0^{95\%}$	$H_0^{90\%}$
CRV	23.90	27.50	1.80	acc	**rej**
CRR	1.27	1.77	-0.73	acc	acc