LUND UNIVERSITY
School of Economics and Management

Risk Measures – from theory to an empirical study over time

Author: Hampus Lingnardz

Supervisor: Frederik Lundtofte

Bachelor's Thesis

Financial Economics

Lund University

2012-01-20

Abstract

This thesis concerns risk measures in theory and an empirical study of their accuracy in predicting future risks, back-testing them using an out-of-sample study with a rolling window scheme. The theoretical part includes a general presentation of risk, covers various risk measures – dispersion measures and safety measures – and attempts to sort out their advantages and disadvantages. The coherent risk measure Conditional Value at Risk should, in theory, be an adequate tool to measure risk. However, the important thing is to know what to do and most likely, one needs to include several measures to get a complete picture of the risk.

The results of the empirical study indicate that the chosen safety risk measures and time periods do not succeed in predicting the amount of risk very well, even though they in some periods do perform quite well. However, if one calls a successful prediction a prediction which falls below the outcome, the methodology succeeds quite well. Also, especially for Maximum Draw-down, instable periods seem to affect the accuracy significantly. Future research might include other risk measures and other time periods – longer prediction windows than one year and shorter or longer estimation periods. Furthermore, one could also look at other time periods for the risk measure definitions (e.g. a week instead of a day).

Keywords

Risk Measures, Value-at-Risk, CVaR, Maximum Draw-down, Risk in Theory, Lower Partial Moments

Acknowledgements

I would like to sincerely thank Frederik Lundtofte at Lund University, Department of Economics, and Handelsbanken Capital Markets. I thank Frederik for his guidance and academic advice and clever words and input along the way. I thank Handelsbanken Capital Markets for providing me with the required data for the empirical part.

Contents

1. Introduction

1.1 Background

The task of measuring risk on an investment or money allocation has become increasingly targeted lately, and more measures have arisen and have been incorporated in regulators' as well as investment firms' daily work. Therefore, it has become more and more important for financial firms to know what to do and measure, as well as being able to perform meaningful risk calculations and estimations – both for internal purposes and to obey regulators. Given this fact, this paper aims at presenting risk measures thoroughly but easily graspable – going through their respective pros and cons – and studying how well some of the most interesting and used safety, i.e. "how-bad-can-things-get", measures perform in predicting future risk.

1.2 Purpose and Problem Formulation

The main purposes of the bachelor's thesis are to take a deeper look into different risk measures used to estimate the risk of a portfolio of assets or a single asset over time. Also, a study is performed on some risk measures, using a back-testing method with rolling estimation windows, on different types of time series data. The aim of the study is to test how well the chosen risk measures can be said to predict future risk.

Given the background, the following issues are defined as the formulation of the problem:

- Present a meaningful, thorough but easily graspable review of available risk measures, commenting on their advantages and disadvantages.
- Perform a study on some chosen risk measures (VaR, CVaR and Maximum Draw-down; see below) with regards to how well they predict risk, back-testing with rolling estimation windows followed by one year prediction periods. Here, I use three different kinds of underlying asset classes, and different approaches.
- Analyze the results of the study, hopefully coming to a conclusion as to how well the chosen risk measures can predict future risk. I also comment on possible future research on the topic and what could be enhanced in the methodology chosen for this study.

1.3 Methodology

For the theoretical part, cover and analyze the market's available risk measures. Present their formal definition, how they are measured and a review of their pros and cons.

For the empirical part, choose three different time series and calculate three different risk measures on them. Perform back-testing with rolling windows over five periods and analyze how well the estimates predict the future risk.

1.4 Limitations

The theoretical part does not include any performance measures – an important part to look at in conjunction with risk measures from a portfolio and investment perspective. The empirical study does not include all the risk measures reviewed in the theoretical part. Rather, the aim is to examine some of the most important and used measures as examples of means of predicting how bad things can get. Finally, the frequency of the time series and estimation procedure is kept at a daily scheme.

1.5 Target Audience

The target audience can be anyone (especially within finance, portfolio risk or financial engineering) interested in getting a thorough and understandable presentation of risk measures and how some of them succeed in predicting risk over time. One can also define an academic target group: students of engineering and/or financial management – even with only a limited beforehand knowledge in the area – are also likely an audience that can benefit from acquainting themselves with the thesis.

1.6 Disposition and Reader's Guide

After this first section, which introduces the thesis, the second section is intended to give the reader a relevant theoretical background including the task of measuring risk and what advantages and disadvantages different risk measures can be said to have. The third section presents the methodology of the empirical part including choosing of data and measuring risk. The fourth section presents the results of the empirical study. Also, analyses of the results are given in this section. Finally, the fifth section concludes and presents suggestions for future research on this topic.

2. Theory

2.1 Risk Measures

2.1.1 Definition of Risk

Typically, risk is defined as the probability of a loss incurred by a choice of a given alternative. This alternative, within the scope of financial investments, is an investment in a certain asset or portfolio of assets. The amount of risk is set according to perception and preferences as well as estimation. Thus, investors may perceive the risk differently, partially depending on what risk measures they use and partially depending on their preferences.

The loss, moreover, does typically not have to be as simply defined as a negative outcome. It could be an outcome worse than a certain predefined target outcome or minimum accepted outcome, above which the outcomes are considered to be a gain. The predefined target is subject to choice by the investor and is of great importance. It has been shown that potential gains tend to decrease the perception of risk. It has not been determined how this works and it is a topic covered by behavioral finance, a modern area of finance that has emerged in later years. Nonetheless, it is important to remember that people tend to judge the riskiness of an investment in different ways, and in another way than when judging the attractiveness (Brachinger, 2002).

Although risk is usually meant to aim at the probability or possibility of a loss, it tends to be measured as the variations around, or discrepancies from, an anticipated average return. This is the case e.g. for the commonly used risk measure Volatility.

Also, risk is commonly divided into two dimensions; the amount of the possible loss and the probability of that loss occurring, both of which are subject to discussion.

Many risk measures have been presented over the years, intended to be applicable on as many as possible of the different risk sources that have been identified: market risk, credit risk, liquidity risk, operational risk etc. In recent years, risk management has become increasingly important, the financial crisis being perhaps the most important reason for the focus on it these days. One problem is how to measure risk in a “good” way in order to be able to quantify and optimize risk exposure. The following section defines, presents and discusses various risk measures from a general point of view.

2.1.2 Ways of Measuring Risk and Coherent Risk Measures

One can distinguish between two sorts of risk measures when going through the risk measures that have been introduced; the dispersion risk measures and the safety measures (Prigent, 2007).

The dispersion measures, which include Volatility, are increasing, positive, and positively homogeneous functions of the risk X. They try to capture the degree of deviations from a certain target level, typically the expected value.

The safety measures, that include the VaR, CVaR etc., involve the probability of the portfolio return becoming worse than a certain level.

Safety measures with certain desirable properties have been introduced as coherent risk measures by Artzner (Artzner, 1997). A coherent risk measure, ρ, should satisfy the following axioms:

- Translation invariance: $\rho(X + \alpha * r) = \rho(X) - \alpha$.
  If one invests a sure amount α in a reference asset with a constant total return r in addition to the risky asset, the total risk decreases by this sure amount. The variation of the risk measure is equal to α itself, which is in accordance with a monetary interpretation of the measure. In particular, $\rho(X + \rho(X)*r) = 0$.

- Subadditivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.
  The total risk of investing in two assets at the same time is less than or equal to the two separate risks on their own; the diversification principle.

- Positive homogeneity: $\rho(\lambda X) = \lambda\rho(X)$.
  The risk measure is a linear function of the size of the position.

- Monotonicity: $P(X_1 > t) \leq P(X_2 > t)$ for all real numbers t, implies $\rho(X_1) \geq \rho(X_2)$.
  The risk of an asset whose return distribution is constantly better than a second asset is less than that of the second asset. Thus, the reverse order is kept.

The arguments that a risk measure is relevant, rational or "good" if it is coherent are many. Primarily, the subadditivity property might be the most important one. It induces that a diversified portfolio should be regarded as less risky, which is a general standpoint among investors. Also, if this property is fulfilled, the risk can never be reduced by dividing the total position into smaller pieces, which is typically regarded as a desirable feature of a risk measure for regulation purposes. Otherwise, a firm could, for instance, break up into two separate affiliates if it needed to meet a requirement of extra capital (Artzner, 1998), not to mention the chance that e.g. a bank could in fact have a total risk a lot higher than the sum of many risks calculated for smaller parts of the bank (Acerbi, 2001) . There are a number of other factors that make the subadditivity property a natural requirement.

Furthermore, the properties of a coherent risk measures is connected to that the risk can be measured in cash. This is considered to be a major practical advantage. The translation invariance and the positive homogeneity properties induce that the risk measure can be easily interpreted. The monotonicity property relates and corresponds to the definition of risk as a probability of loss.

Mentionable is that all dispersion measures, including Volatility, as well as the most commonly used safety measure Value at Risk, are not coherent. However, Conditional Value at Risk (CVaR) is coherent.

In the case of financial assets, one is interested in the return distribution to evaluate the risk. The return is normally defined as the percentage change in the value of the asset from time to the next or the logarithm of the present value divided by the last value. These two ways of defining returns converge to the same value if the returns are small, which tends to be the case for daily returns. The latter definition is commonly used for stock returns, since the assumption of the stock prices following a Geometric Brownian Motion implies that the log returns are normally distributed. As it is easier to interpret the first definition of returns, that is the one that will be used throughout this thesis.

2.1.3 Dispersion Measures

2.1.3.1 Volatility

Variance and Standard Deviation are defined and estimated as follows. Volatility is equivalent to Standard Deviation.

$$Var(R) = \int_{-\infty}^{+\infty} (r-\mu)^2 dF(r) \tag{1}$$

$$Sample\ Var(R) = \sqrt{\frac{\sum_{i=1}^{n}(r_i - \bar{r})}{n-1}}\ , where\ \bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i \tag{2}$$

$$StD(R) = \sigma = \sqrt{Var(R)} \tag{3}$$

where R is the random variable (returns) with probability function F(r), whose expected value is μ. There are n observations of R, r_i. Sample Variance is an unbiased estimate, although the square root of this is a biased estimate of Volatility. There exists no universal unbiased estimate of Volatility.

Variance is the second central moment of a distribution (the Mean is the first central moment) and it is defined as the squared deviations from the Mean summed up and weighted using the probabilities of the respective outcomes, as can be seen above. The square root of Variance is called Standard Deviation, and is typically named Volatility when discussing the risk of financial assets. Therefore, choosing between Variance and Volatility is just a question of scaling.

Volatility was the risk measure of choice when Modern Portfolio Theory and its Mean-Variance Portfolio Optimization were introduced in the 1950's. It is an easy way to get a number of the deviations from an average, i.e. a measure of how much the asset value tends to vary. It is very well known and widely used even today, although many do not regard it as a good measure of risk.

First of all, for the Variance measure to say anything about the riskiness of an asset, one must estimate it. This includes computing a sample Variance using a certain sample space that needs to be chosen in an appropriate way. The sample space needs to be big enough to capture enough data to avoid a bad estimate but small enough to give a good and relevant estimate of current and near-future Volatility.

However, many studies have shown that Volatility is not static; heteroskedasticity is a common feature, meaning that the Volatility changes over time. It is a well known phenomenon that Volatility typically increases when the underlying experiences negative shocks for instance (Schwert, 1990). There are several heteroskedastic models, such as the GARCH model, that try to capture the fluctuations in the Volatility, but this area is not covered in this thesis.

Second of all, there are practical issues of converting Volatility to different time scales. One most commonly refers to "yearly Volatility" when comparing assets through Volatility. Let us say that one has one year of monthly data and calculates Volatility for that. The Volatility will then be "monthly Volatility" based on that one has used monthly data, and needs to be converted to "yearly Volatility". The common practice is to then multiply the Volatility with $\sqrt{12}$ (or multiply the Variance by 12), to get "yearly Volatility". Scaling up and down in this way is only valid if the samples in question are i.i.d. (independent and identically distributed) (Diebold, 1996). High frequency (daily) financial data is typically not i.i.d. and scaling in this manner makes Variance/Volatility even worse as a risk measure. The only proper way of converting Volatility to different time scales is through assuming models of the Volatility structure, which itself incurs several problems, especially in the big scheme where all market players should calculate Volatility in the same way.

Furthermore, since an asset whose returns have a high Volatility is regarded as risky, a deviation from the Mean has to be regarded as something undesired. This can be the case for symmetric distributions, when negative and positive deviations will happen equally often. Numerous empirical results show that this is typically not the case. If the returns are not symmetrical, the deviations from the Mean summed up will include up-side and downside deviations, without making any distinction between them in terms of their impact of the total risk measure.

Finally, when dealing with distribution moments in a risk perspective, one needs to consider at least all lower moments for the evaluation to make sense. In the case of Variance, one also needs to consider the Mean to be able to fully evaluate the investment, but to get the whole picture higher order moments are required as well (see 2.1.3.3).

2.1.3.2 Beta

The Beta of an asset R with returns r_i to a portfolio of assets P with returns r_P is defined as:

$$\beta_{RP} = \frac{Cov(r_i, r_P)}{Var(r_P)} \tag{4}$$

Thus, the Beta value of an asset is the Covariance between the asset and the benchmark portfolio scaled by the Variance of the benchmark portfolio. Beta is also equivalent to the Correlation between the asset and the benchmark portfolio times the Variance of the asset. Therefore Variance plays a big part in Beta, and makes the measure show the same disadvantages as the Variance itself.

The Beta is a relative measure arising from CAPM, telling the investor the extent to which the asset is correlated to a chosen benchmark portfolio. As one distinguishes systematic risk and idiosyncratic risk in CAPM, beta is a risk measure only measuring systematic risk as this is assumed to be the only risk affecting the value of the investment. The benchmark portfolio is often chosen to be a broad market index. According to CAPM, the benchmark is supposed to be the market portfolio; a perfectly weighted portfolio of all risky assets on the market.

All in all, the Beta value contains good information for an investor that believes that historical correlation of an asset with a certain benchmark is going to continue to be on the same level in the future, and information about the riskiness in comparison to a benchmark. A Beta of more than 1, for example, suggests that the asset is more risky than the benchmark. This is true if risk is defined as Variance. The Beta can say something about the risk relative to a benchmark index for instance, and could therefore be regarded as a relative risk measure, but not an absolute one.

The validity of the Beta relies on the assumptions of CAPM, of which some are obviously non-realistic. Especially, the correlation is, according to several empirical studies, not stable over time, implying that the beta measure does not necessarily say anything about the future – at least the distant one. This is even more obvious for assets that have a non-trivial and non-static correlation to any benchmark portfolio.

2.1.3.3 Higher Order Moments

Once first and second order central moments for the return distribution, the Mean and the Variance, have been identified, one can look at higher order moments to learn more about the properties of the distribution. The definition of the central moment of degree N, μ_N of the random variable R with pdf F(r) and expected value μ as:

$$\mu_N = \int_{-\infty}^{+\infty} (r-\mu)^N dF(r)\,, \tag{5}$$

Usually one refers to the *normalized* third and fourth central moment as Skewness and Kurtosis respectively, defined and estimated as (σ is the standard deviation of F(r)):

$$Skewess(X) = \int_{-\infty}^{+\infty} (r-\mu)^3 dF(r)/\sigma^3 \tag{6}$$

$$Sample\ Skewness(X) = \frac{\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2\right)^{3/2}} \tag{7}$$

$$Kurtosis(X) = \int_{-\infty}^{+\infty} (r - \mu)^4 dF(r)/\sigma^4 \tag{8}$$

$$Sample\ Kurtosis(X) = \frac{\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2\right)^{2}} \tag{9}$$

Both these estimates are biased. Although there are corrections for these biases, they do not hold for in general cases.

The Skewness contains information about the extent to which the distribution is non-symmetrical. For a distribution with a positive Skewness, which is also called a right-skewed distribution, the lower tail is thinner whereas the upper tail is fatter. This is typically seen as a good property by investors. A left-skewed distribution, or a negative Skewness, has a fatter lower tail and a thinner upper tail, which is typically not a sought-after property.

The advantage of using the Skewness to classify risk is that it adds valuable information of distributions that are non-symmetrical. For example, two distributions that have the same Mean and Volatility, would from a Mean-Volatility perspective be perceived as equally risky. By using the Skewness, the investor may be able to conclude that one is more attractive than the other.

The robustness of the Skewness measure can be affected if the sample is too small, resulting in extreme values contributing to an incorrectly high Skewness by the cubic function. This can be avoided by using some robust variant of the measure, such as Pearson's Skewness coefficient which is the difference between the Mean and the Median divided by the Standard Deviation and Bowley's Skewness coefficient which is defined as a ratio using the 25, 50 and 75% quantiles (Sulewski, 2008). These measures can sometimes be too robust, ignoring extreme values that should affect the Skewness value.

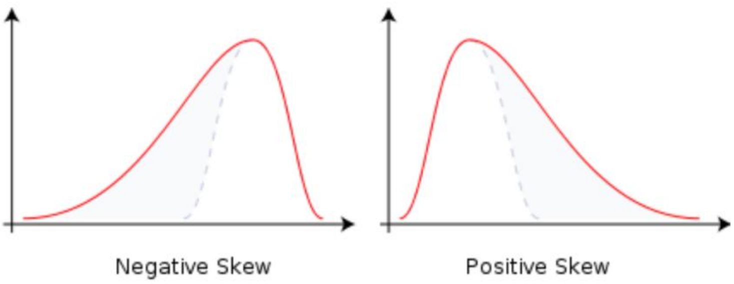

Figure 1 Probability density functions with negative (left) and positive (right) skews.

The Kurtosis measures the extent to which the distribution is peaked and contains information about how fat the tails of the distribution are. One usually puts the sample Kurtosis in relation to the Kurtosis of the normal distribution which is equal to three. Therefore sample Kurtosis is usually normalized by subtracting three. A large (positive) Kurtosis implies a thinner and higher distribution close to the Mean and fatter tails – typically a negative property.

Just like for the Skewness measure, the Kurtosis measure suffers from lack of robustness to extreme values. There exist several Kurtosis measures that are aimed to be more robust than the classical Kurtosis. These include Moors's measure, defined as a ratio using six different octiles (Moors, 1995), and the Crow/Siddiqui measure, using four quantiles (White, 2008).

Figure 5 shows the meaning of Kurtosis and states what the three typical kinds of Kurtosis, K, where K=0 corresponds to the Kurtosis of the standard normal distribution – i.e. K is the excess Kurtosis; the Kurtosis in excess of that of the normal distribution.

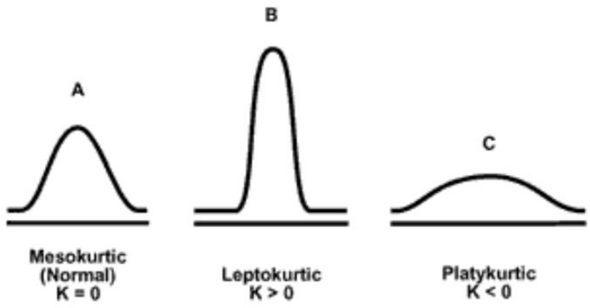

Figure 2 Probability density functions showing mesokurtic (left), leptokurtic (middle) and platykurtic (right) Kurtosis.

Taking higher order moments into account result in the question of how all measures should be weighted and summarized to give one measure of risk. This is highly non-trivial, and is a topic subject to continuous research. Also, especially for non-symmetrical distributions associated with non-linear products, one needs to be careful when making assumptions with respect to one moment.

2.1.3.4 Lower Partial Moments (LPMs)

Lower Partial Moments give a measure of the downside risk by capturing returns falling below a certain target return t, which could be set at any target return: zero, the risk free return or the Mean – the Mean being the most common one.

The N:th order LPM is defined and estimated as follows:

$$LPM_N(t;F) = \int_{-\infty}^{t} (t-r)^N dF(r) \tag{10}$$

$$Sample\ LPM_N(t;F) = \frac{1}{n}\sum_{i=1}^{n} \max(t-r_i,0)^N \tag{11}$$

The most commonly used LPMs are:

$$SemiVariance = LPM_2(\bar{r};F) \tag{12}$$

$$(Standard)SemiDeviation = \sqrt{LPM_2(\bar{r};F)} \tag{13}$$

If t is set to the VaR value and N=1, the LPM gets very similar to CVaR (see 2.1.4.2/2.1.4.3) (Witt, 2009). Also, the target return t may vary over the integration/summation area.

Lower Partial Moments as risk measures were introduced in the late 1970's. The argument for the use of LPMs as risk measures rather than the Volatility is that they only takes negative deviations into account – investors tend to be more worried about them than positive deviations. Another advantage is that they are more suitable to look at negative deviations if the return distributions are non-symmetrical, which tends to be the case. Nonetheless, they still haven't become as popular as the Volatility measure. One reason for that is that they are not additive – if one knows e.g. the semi deviations of all the constituents of a portfolio separately the total Semi-Deviation cannot be easily calculated using that information, which is the case for Volatility (knowing the Covariance matrix).

Furthermore, the choice of target return has a huge impact on the evaluation, and needs to be chosen with care consequently. Also, the number itself, regardless of what target return one chose, doesn't necessarily say anything, even though it can be a good way of comparing many assets to each other, in which case it is crucial to use the same target return.

2.1.4 Safety Measures

2.1.4.1 Maximum Drawdown

Maximum Drawdown (MDD) is defined as the largest drop in asset value along a specified time period. It is usually the maximum relative drawdown in percentage, from a “peak” to a “valley”, i.e. the maximum loss that an investor could have experienced within the time period.

MDD is a measure of sustainability of the investment (Magdon-Ismail, 2004), and has become increasingly popular in for example the hedge fund industry. Apart from calculating the value, one sometimes also looks at the uninterrupted drawdown which is defined as the length and the severity of an uninterrupted drop. Furthermore, the recovery time states the time to recover from the draw down, i.e. the time it takes before the asset value is back at the original level which was the asset value before the drawdown started.

The advantage of MDD is that it is very easily understood and indeed says something about a ceiling for losses as well as the duration of drawdown and recovery (in the past) and based on that, the investor can assess the riskiness of the asset if the situation were to recur. MDD is furthermore a measure of downside risk, which is usually preferable to a measure that takes upside risk into account as well, like the Volatility. Also, in comparison to e.g. Volatility, the MDD measure refers to a physical reality making it less abstract and more intuitive. Moreover, it can be calculated for any time series regardless of its return distribution.

There are, however, a few reasons for using the MDD measure with some caution (Lhabitant, 2004). First, in order to be able to compare MDD’s between time series of different assets, one needs to make sure that the reporting intervals, i.e. the frequency of the measurement interval, are the same for all time series or make an appropriate correction. This is because of the fact that MDD’s are greater if the frequency of measurement interval is smaller, all other things being equal.

Second, since MDD’s get greater as the time series get longer, one needs to ascertain that the periods used to calculate the MDD’s are equal for all assets that one is interested in comparing. To avoid this in practice, it has become popular to have a three year period as a ground when calculating present MDD’s in an attempt to form an industry standard. On the whole, it is of great importance to choose an appropriate investment horizon and calculate the historical Maximum Drawdown based on that.

Finally, this measure cannot say anything about the risk in a portfolio as the magnitude of losses until after they occur. Moreover, only the worst drawdown is considered – the measure says nothing about the second largest drawdown and so on.

2.1.4.2 Value at Risk (VaR)

Definition:

$$VaR_\alpha = \inf\{t \in \mathbb{R}: P(R > t) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R}: F_R(l) \geq \alpha\} \qquad (14)$$

where R is the random variable (return) with cdf F_R and α is the chosen level of significance.

As an example, the Value at Risk of an asset at the 95% significance level at a one week period of $20 million says that a loss of more than $20 million within one week will not occur in more than 5% of the cases; i.e. there is a 95% certainty that the outcome of the return or profit of the portfolio will incur losses less than $20 million over a given week. This can be seen in figure 6.

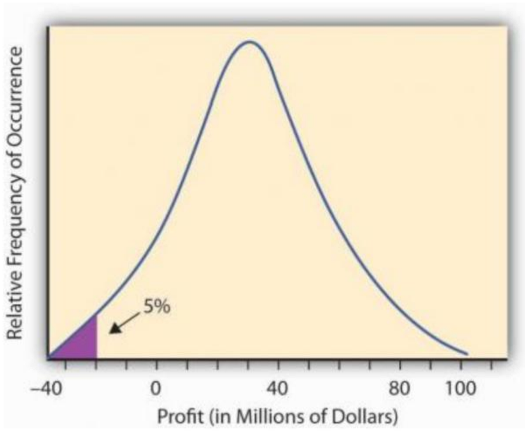

Figure 3 Probability density function of the future profit of a portfolio with values worse than or equal to the 5%-quantile marked in the left end of the figure.

This measure has, since it was first introduced in the early 1990's, become very popular for risk management purposes, and has been embraced by the regulatory entities within the financial industry as well. Its importance and popularity have increased even more due to the current financial crisis, as an alternative to risk adjusted value and probabilistic approaches since it borrows from both (Damodaran, 2005).

There are several reasons for the popularity of the measure. First of all, it gives a direct figure in cash on how big a loss will be less than with a certain high probability. This is something that anyone can understand the meaning of, unlike a Volatility of 20% for example, which might not say all that much for a non-specialist. Also, it is relatively easy to calculate – at least in some of the most straightforward cases – and hence easy to create and compare among assets. This makes it straightforward to come up with rules regulating how much capital for instance banks need to keep available in comparison to the VaR figure for a specific period. Furthermore, the figure also indicates a ceiling for losses under normal circumstances. This is something that is hard to accomplish for other risk measures.

There exist many drawbacks with the value at risk measure too, though. First of all, it says nothing about how big losses actually can get. One never looks beyond the VaR loss, and does consequently not take the rest of the probability distribution into account at all – even though one sometimes calculates Value at Risk for several levels of confidence. How fat the tails of the distribution are is not considered, meaning that two portfolios can have the same VaR value but can be quite different at the very ends of their distributions. This implies that, for example two banks which present the same VaR values should not necessarily need to have equally large amounts of money to back up the risks, namely if the distributions beyond the VaR behave differently. Similarly, VaR does not say anything about the right hand distribution and other measures are necessary as well to make a good evaluation. Finally, VaR is not coherent since the subadditivity feature is not present. This means that the total risk of a portfolio of many assets can be larger than the individual risks added up, which for one thing is quite contrary to the diversification principle commonly accepted among investors.

Several extensions of the VaR have been proposed, one of the main ones being the CVaR (see 2.1.4.3), which considers the tail information beyond the VaR point as well.

The VaR can be calculated in three substantially different manners; the parametric approach (or Variance-Covariance-VaR), historical simulation approach (or non-parametric approach) and the Monte Carlo approach. All three methods have specific advantages and disadvantages. These are relevant for CVaR as well, since one calculates the VaR as a part of the CVaR calculation. Since this thesis will only involve the historical simulation approach, the other methods are not presented, apart from in the following summary.

Historical Simulation

The historical simulation approach uses historical data for the assets' returns to create a return distribution which is then used to calculate the VaR. In general, to create the 99%-day-VaR for instance, one simply goes through all the daily returns over a chosen time period such as five years, and ranks them according to value from best to worst, and locates the 1% quantile in the series. If there are 100 observations, the 99% VaR is equal to the second worst value.

There are several advantages of this approach. The method is easy to implement and to use, given that the data for an appropriate time period is easily available. Also, it is an approach that is easy to explain to most people; non-specialists, regular investors or senior management. Further, this approach can be used regardless of the distribution of the portfolio return, i.e. regardless of what asset classes are among the constituents – options, callable bonds etc. can be part of the portfolio without any problems.

Several shortcomings of the historical approach arise from the underlying assumptions. First, even though all three approaches of calculating the VaR involve the use of historical data, this approach relies heavily on it. Using the past as a prologue for the future might not always be a good idea. Second, even though it is commonly regarded as an approach with no assumptions on the distribution, the returns do need to be IID, i.e. identically and independently distributed, which is in fact quite a strong assumption. Another argument against this approach, related to the one discussed above, is based on the fact that there may exist trends in the data, e.g. for Volatility. Since all the past outcomes are given equal weights, any such trends will then not be incorporated in the VaR calculation, resulting in a wrong estimate of the VaR. Proposed model modifications on this note as well as the former include weighting the recent past more than the distant one. Also, methods that combine historical simulation with time series models, fitting a model, such as an autoregressive moving average (ARMA) model, to historical data and using the parameters to forecast the VaR, have also been proposed, along with Volatility updating methods.

Thirdly, new asset or market risks may arise along the way. Again, even though all three approaches have difficulties dealing with this feature, the historical approach is the one that has the most difficulty succeeding, due to the fact that it only uses historical data. This approach cannot, in its general formula, perform any scenario analyses, and this includes changing market risks and new assets being added (Linsmeier, 1996).

In summary, the three main approaches have their separate advantages and drawbacks. The parametric approach is fast but involves unrealistic assumptions; the historical approach is easy to understand but is relies on the past to foresee the future; and the Monte Carlo approach is the more flexible but more time consuming one. The trade-off between the approaches is evident and needs to be evaluated by whoever is interested in calculating the increasingly popular Value at Risk based on certain case specific features from one situation to another.

2.1.4.3 Conditional Value at Risk (CVaR)

The Conditional Value at Risk is an extension of VaR which takes the mean over all the losses greater than or equal to the VaR point. Thus, it takes the whole distribution tail in consideration and looks beyond the VaR. Another name for this measure is Expected Shortfall. The formal definition is as follows:

$$CVaR_{\alpha} = -E[R \mid R < -VaR_{\alpha}] \tag{15}$$

where R is the random variable (return), α is the confidence level and VaR_{α} is assumed to be a positive number, yielding a positive $CVaR_{\alpha}$ under normal circumstances.

The CVaR is a measure of how big losses will get on average given that they get greater than the VaR with a specified level confidence. Thus, rather than answering the question “How bad can things get?”, CVaR deals with the question “If things do get bad, how bad will it get?” (Hull, 2009).

The advantage is that it uses information about the entire tail distribution to create a clear number of the average loss for rare losses, instead of just calculating a loss that will not happen very often which is the case for VaR. Also, CVaR is a measure of downside risk as well as a coherent measure (see above). Hence, CVaR can be regarded as a better way of measuring risk than VaR.

One drawback is that it doesn't take the rest of the distribution into account. Two return distributions with the same CVaR could be quite different above the CVaR point. One needs to include other risk and return measures as well to be able to evaluate investments in a good way using CVaR.

2.2 Evaluation of Risk Measures

Below is a review of pros and cons for the risk measures gone through above, presenting this in a matrix.

Volatility

Pros

- Widely used, including in the MV-model.
- Measures deviations around the expected value in a straightforward way.
- Works well when returns are normal, a good approximation for equity distributions.

Cons

- Hard to interpret.
- Positive and negative deviations are usually not equally bad.
- Other properties of the distribution, like Skewness and Kurtosis, are not accounted for.
- Volatility is not static over time; heteroskedasticity common.
- Not a coherent risk measure.
- Returns not normal in general, especially for non-symmetrical instruments

Beta

Pros

- Well-known measure of covariation with e.g. an index.
- Good when comparing to the market

Cons

- Not an objective measure.
- Not a coherent risk measure.
- Requires constant correlation. Rarely true for longer periods and certainly not true for non-symmetrical instruments.

Higher Order Moments (Skewness, Kurtosis)

Pros

- Contains good information about the distribution.
- Adds to the information included in the lower moments (Mean, Variance).
- Good as complements to the Volatility for non-linear products.

Cons

- Can be hard to understand, given the mathematical definition.
- The robustness is sometimes weak as extreme values can get an over-weighted impact on the value.
- Necessary to try and make a full picture using all moments. Weighing all moments is hard and a subjective choice.

Lower Partial Moments (Semi-Deviation)

Pros

- Measure the downside risk
- Negative deviations are what matter, not positive ones.
- Good when the return distribution is non-symmetrical

Cons

- Not as popular and used as e.g. Volatility.
- Not additive like volatilities are.
- Not a coherent risk measure.
- Not easy to understand and interpret.

Maximum Drawdown

Pros

Cons

- Easy to understand, measures the sustainability and a ceiling for losses
- Increasingly popular, e.g. by hedge-funds
- Measures the downside risk as a graspable value
- Can be calculated regardless of the distribution

- Depends on the time interval
- Could give wrongful expectations about future drawdowns; is just one observation. Could be good to compute the average draw down over several periods
- Time series must be equally long and be the result of equally frequent sampling when comparing several investments
- The second largest draw down is not taken into consideration
- Not a coherent risk measure

VaR

Pros

- Increasingly popular for risk management purposes, especially among regulators
- Relatively easy to calculate and requires no assumptions about the distribution
- Easy to understand, measures the downside risk
- Gives a value on how bad things can get

Cons

- Says nothing about how bad things *can* get, i.e. how things look further out in the tail of the distribution
- For the historical approach; the past distributions do not necessarily say anything about the future
- Not a coherent risk measure - not subadditive, so it contradicts the diversification principle -> Not ideal

CVaR

Pros

- Increasingly popular
- Gives a value of expected losses given that they get bad
- Captures the properties of the tail of the distribution
- Measures the downside risk
- Coherent
- No assumptions about the distributions are needed

Cons

- Looks only at extreme negative out-comes and nothing about the rest of the distribution
- Can be hard to understand
- Two assets with the same CVaR may not be equally risky or desirable; moments and other risk and performance measures as well
- For the historical approach; the past needn't say anything about the future
- The outcome is affected by the time period

3. Methodology

3.1 Data Gathering and Data Selection

The empirical part of the thesis requires data containing prices of the relevant indices. All data necessary are ordered through Handelsbanken Capital Markets and the Reuters 3000Xtra application or www.omxnordicexchange.com.

Three sorts of underlyings are chosen to perform the studies on; one Commodity, one Equity Index and one Bond Index. This is to capture three of the main types of asset classes that investors choose between and their money is allocated, and to see whether or not one will note any differences in predictability of risk.

The three underlyings, denoted Gold, OMXS30 and OMRX respectively later on in this paper, were chosen as the Spot Gold Price (Reuters page XAU=), the OMXS30 (Reuters page .OMXS30) and the OMRX-TOT (GOVT+BOND). They are defined s follows:

OMRX Bond Index
The OMRX Bond Index consists of a mixture of national treasury and mortgage debt. A complete index of the Swedish National Debt Office's and the mortgage institutions' borrowing in Sweden

OMXS30 Equity Index
The OMXS30 Equity Index consists of the 30 largest stocks on the Stockholm Stock Exchange, Nasdaq OMX Stockholm. Can be said to be a good proxy for Equity in the Nordics.

Gold Spot Price
The Gold spot is the price for one troy ounce of gold. One of the most liquid commodities, historically often used as a protection against inflation.

The data from the three underlyings that was found spanned over varying time spans. Gold had data available from 1968-03-22; OMXS30 had data from 1986-09-30 and OMRX from 1990-04-06. The final day for the three time series were 2011-09-23, 2011-09-23 and 2011-09-19, respectively.

Daily returns are used throughout the time series studies. The price type "Last" has been used to denote the price on a day, and the daily return R_t has been defined as:

$$R_t = \frac{(S_t - S_{t-1})}{S_{t-1}} \tag{16}$$

where S_t denotes the price of the index on day t, and t-1 is the banking day preceding day t. Hence, percentage returns are used. These returns, over the defined period, then make up the time series.

3.2 Risk Measures to include in the study

The following risk measures are chosen: Value-at-Risk (VaR); Conditional Value-at-Risk (CVaR) and Maximum Draw-down (MDD). The reason why is that they are all common in recent practice amongst both hedge funds, investment banks and investors for reasons discussed in the theory section above, and the all try to say something about how bad things can get (unlike some other risk measures gone through in the theoretical part above, like Volatility).

Interesting results might appear when comparing the coherent risk measure CVaR and VaR – can the coherence make any difference in terms of how well it works to predict the risk in the future given an estimate created on historical outcomes.

All three risk measures, however, depend on the time period over which the estimations are created. This is also thought to be of interest to keep in mind while looking at the results, since one might want to alter the time period approach etc. and bear it mind while commenting on the results. Given the pros and cons discussed in the theoretical part above, they all depend on whether the period has included “unusually” bad outcomes; the MDD also depends very much on the time periods start and end date.

For Maximum Drawdown, VaR and CVaR, please see Appendix for relevant Python Code to see exactly how they were defined in code. Ultimately, the definitions gone through in the risk measure definition section (equations (14)-(15) for VaR and CVaR) are what the code accomplishes.

3.3 Approach for the study

3.3.1 VaR and CVaR

The general approach is to take historical data from the chosen underlyings (see below), create estimates based on a relative long period ending a few years back, followed by a short period whose outcome is compared to the estimate and look at how well the estimate is able to predict the respective risk measured as percentage difference between the outcome and the estimate.

After the risk have been calculated (and compared) for the first estimation and prediction periods, the estimation “window” is moved one year forward, as is the shorter prediction “window”. The risks are measured and compared and the task is repeated five times, and the final prediction period ends as close to today as possible.

The initial approach is to create an estimation time period as long as possible and with five years available at the end for the out-of-sample outcome tests, of equal length equal to 1 (one) year. Since the final available data point is located at the end of September 2011, September 30th is chosen to be the last data point of each period (the rolling estimation period included) and October 1st is chosen to be the first day of each period (the rolling estimation period included, to create an estimation period consisting of an even number of years).

Given the above and the available data, the Start and End dates for all the periods (the rolling estimation windows as well as the subsequent prediction periods) we chosen as presented in following matrix (even rows being estimation windows / periods and odd rows being the prediction periods):

OMXS30		OMRX	
Start	End	Start	End
1986-10-01	2006-09-30	1990-10-01	2006-09-30
2006-10-01	2007-09-30	2006-10-01	2007-09-30
1987-10-01	2007-09-30	1991-10-01	2007-09-30
2007-10-01	2008-09-30	2007-10-01	2008-09-30
1988-10-01	2008-09-30	1992-10-01	2008-09-30
2008-10-01	2009-09-30	2008-10-01	2009-09-30
1989-10-01	2009-09-30	1993-10-01	2009-09-30
2009-10-01	2010-09-30	2009-10-01	2010-09-30
1990-10-01	2010-09-30	1994-10-01	2010-09-30
2010-10-01	2011-09-23	2010-10-01	2011-09-19

Table 1 Start and End dates for estimate (even) and prediction (odd) calculations for OMXS30 and OMRX respectively for approach 1.

The final study is performed based on "stable" periods, subjectively chosen based on a look at the return graphs of the three time series. The goal in this search for stable periods is to identify periods with roughly the same volatility, but still include fairly recent times. The return graphs are shown below.

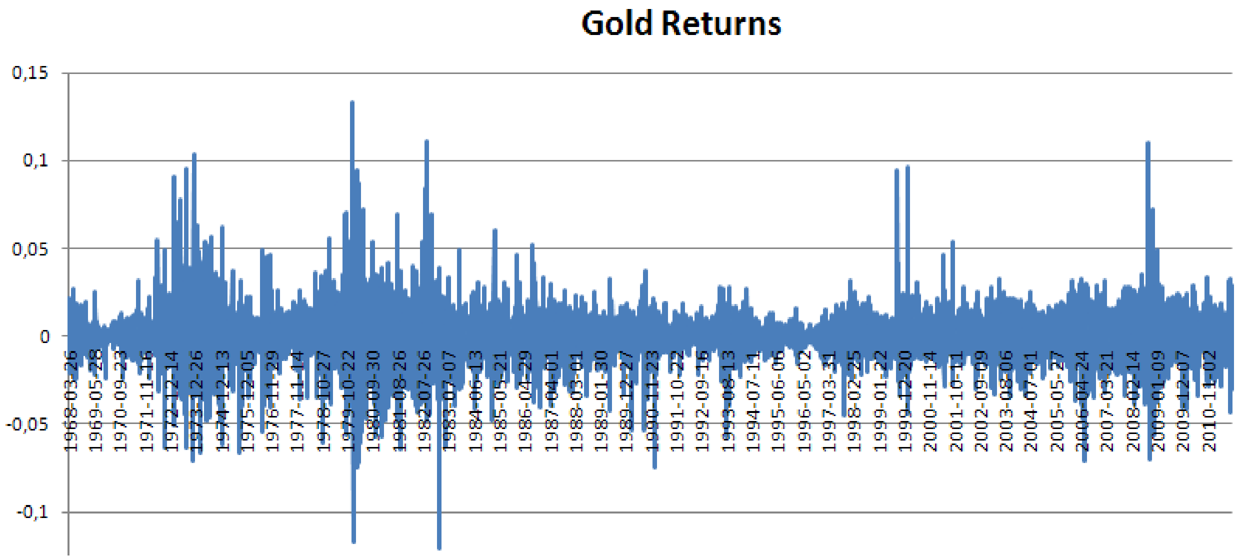

Figure 5 Gold Daily Returns

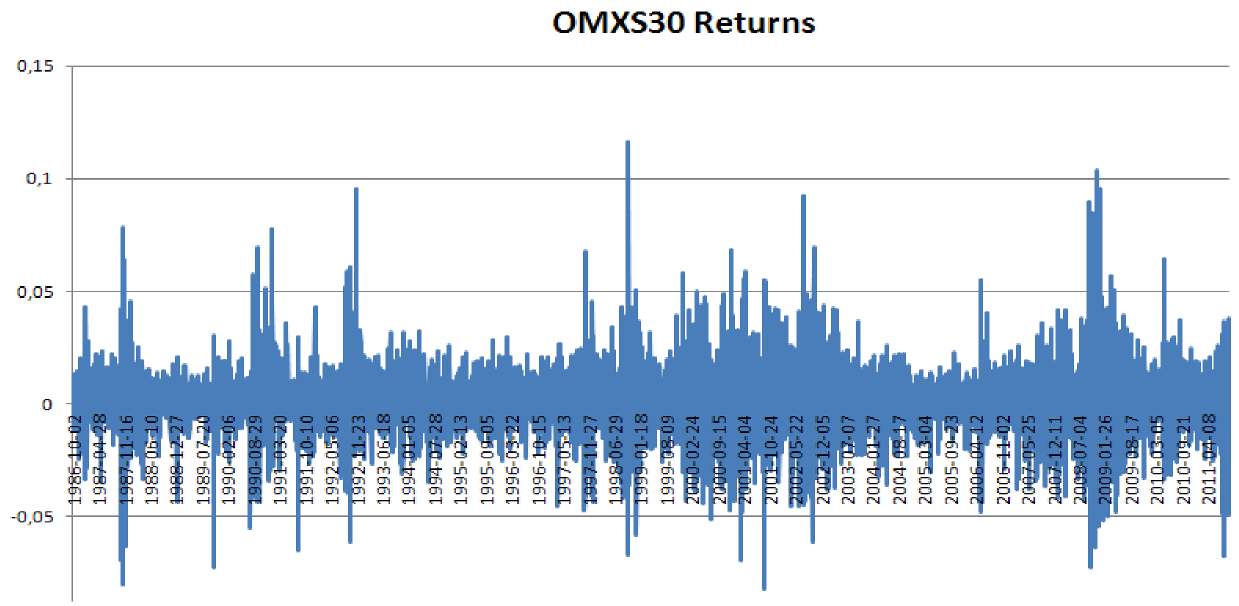

Figure 4 OMXS30 Daily Returns

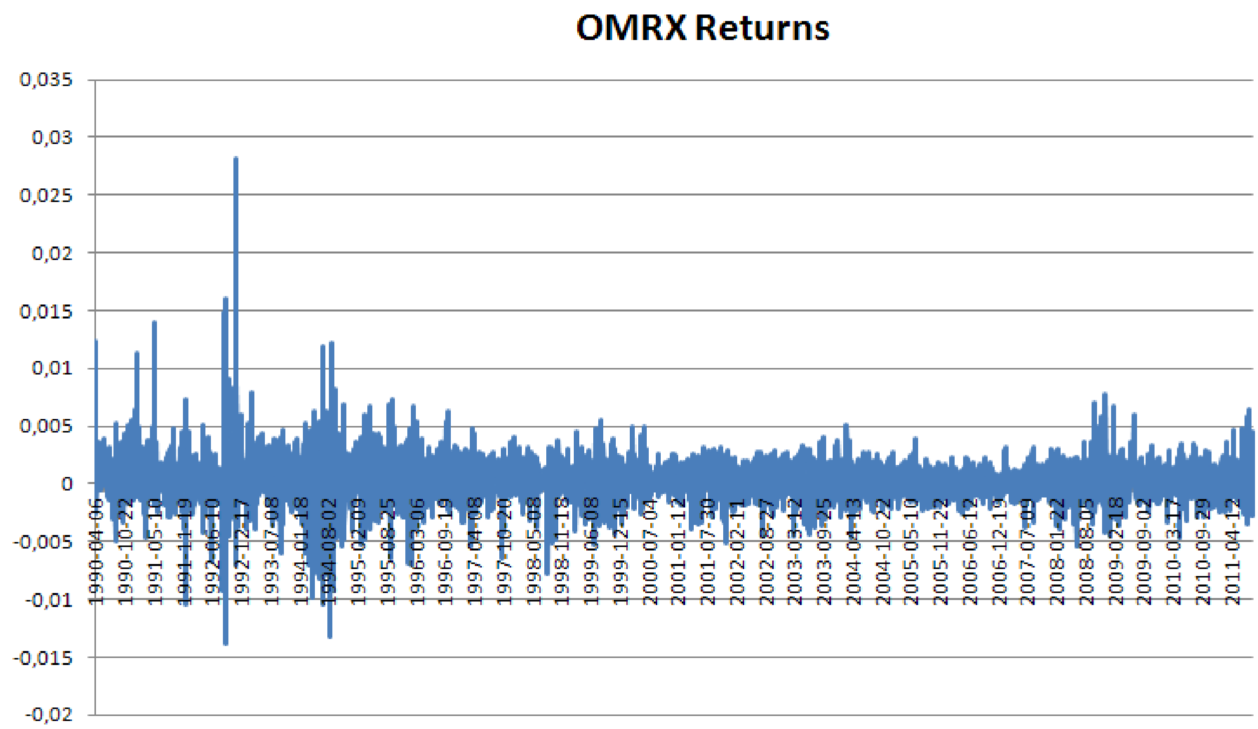

Figure 6 OMRX Daily Returns

The approach with one year prediction periods preceded by an estimation period of sensible length was kept. The instable periods which were to some extent removed from the time series, the latter half of the 00s (for the Equity Index) and the early 90s for the Bond Index – periods that can be ruled as somewhat out of the ordinary in the respective markets. Also, the prediction periods should not be too far back in the future, so that results can be said to be up-to-date once reached. The above and the graph led to the following decisions with regards to the second study for more stable times:

OMXS30		OMRX	
Start	End	Start	End
1992-10-01	2000-09-30	1995-10-01	2002-09-30
2000-10-01	2001-09-30	2002-10-01	2003-09-30
1993-10-01	2001-09-30	1996-10-01	2003-09-30
2001-10-01	2002-09-30	2003-10-01	2004-09-30
1994-10-01	2002-09-30	1997-10-01	2004-09-30
2002-10-01	2003-09-30	2004-10-01	2005-09-30
1995-10-01	2003-09-30	1998-10-01	2005-09-30
2003-10-01	2004-09-30	2005-10-01	2006-09-30
1996-10-01	2004-09-30	1999-10-01	2006-09-30
2004-10-01	2005-09-30	2006-10-01	2007-09-30

Table 2 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for OMXS30 and OMRX respectively for approach 2.

For both of the sets of time period approaches above, the technique is as follows:

1) Create estimation (prediction), defined as the estimate, of the risk using a long period (beginning on October 1^{st} and ending September 30^{th} – or final available day in September). Using the Python code (see Appendix) for all the respective time series types and the Start and End periods, calculate the

VaR and CVaR (for 95% and 99% significance levels). These values are defined as estimates for the respective risk.

2) Calculate the outcome for the risk measures for the year, defined as the prediction period, after the estimation period.
3) Compare the outcome with the estimation and calculate the percentage dispersion of the outcome from the estimate for all risk types to create the Predict Error.

3.3.1 Maximum Draw-down

For the MDD measure, the approach is the same as the first approach for VaR and CVaR gone through above when it comes to the prediction periods. Also, rolling back-testing windows are used in a similar way. However, the estimation procedure is somewhat altered, since it seems implausible that, say, a 15 year period's MDD will correspond to that of a subsequent one year period. The estimations are instead calculated according to:

$$\overline{MDD_T} = \frac{1}{5} \times \sum_{t=T-5}^{T-1} MDD_t \tag{17}$$

where $\overline{MDD_T}$ defines the MDD estimate for period T which should be compared to the outcome in period T (one year in length), and MDD_t is the calculated MDD in period t (one year).

Hence, the MDD estimate for period T is the average MDD of the five (5) precedent periods (years). The procedure is performed – as for VaR and CVaR – for the five last one year periods available in the data. Consequently, T will equal the last 5 one year periods (commencing October 1st and ending September 30th) for all time series. MDD estimations and outcome calculations will be performed in the periods presented in the following matrix:

Start	End
2001-10-01	2006-09-30
2006-10-01	2007-09-30
2002-10-01	2007-09-30
2007-10-01	2008-09-30
2003-10-01	2008-09-30
2008-10-01	2009-09-30
2004-10-01	2009-09-30
2009-10-01	2010-09-30
2005-10-01	2010-09-30
2010-10-01	2011-09-23

Table 3 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for MDD study.

The Predict Error based on the deviation of the calculated MDD outcome from the $\overline{MDD_T}$ is calculated in the same manner as for VaR and CVaR described above.
Since the approach used for the MDD measure includes averaging, and the estimation periods are five year periods during the last decade, no apparent need for additional tests during stable periods is present.

4. Results and Analysis

4.1 VaR and CVaR

The results of these risk measures are presented in two sections: the first approach and the final one.

4.1.1 The first approach

The first results of the empirical study are presented in the matrices below; one per asset class tested. The Risk values (estimations and outcomes) and Predict Errors (PE) are given for each risk and level of significance and period, followed by the average values per risk/level of significance.

4.1.1.1 Gold

			VaR				CVaR			
	Start	**End**	**95%**	**PE**	**99%**	**PE**	**95%**	**PE**	**99%**	**PE**
1	1968-10-01	2006-09-30	1,77%		3,60%		2,92%		4,99%	
	2006-10-01	2007-09-30	1,71%	-3,69%	3,25%	-9,70%	2,21%	-24,26%	3,35%	-32,84%
2	1969-10-01	2007-09-30	1,78%		3,60%		2,92%		4,97%	
	2007-10-01	2008-09-30	2,52%	41,53%	3,86%	7,36%	3,21%	9,73%	3,87%	-22,22%
3	1970-10-01	2008-09-30	1,83%		3,62%		2,96%		4,98%	
	2008-10-01	2009-09-30	2,73%	49,30%	5,63%	55,56%	3,96%	33,44%	6,33%	27,20%
4	1971-10-01	2009-09-30	1,86%		3,68%		3,02%		5,06%	
	2009-10-01	2010-09-30	1,68%	-9,81%	4,08%	10,78%	2,59%	-14,37%	4,12%	-18,53%
5	1972-10-01	2010-09-30	1,86%		3,69%		3,02%		5,04%	
	2010-10-01	2011-09-23	2,24%	20,65%	3,56%	-3,56%	2,80%	-7,28%	3,94%	-21,86%
	Average		2,02%	19,60%	3,89%	12,09%	2,97%	-0,55%	4,63%	-13,65%

Table 4 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for Gold for approach 1. VaR and CVaR estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

The outcomes for VaR and CVaR correspond to their preceding estimations to a varying extent; some are quite close whereas some differ by up to half the size of the risk estimate. On average, these risk measures differ by 10-20%, even though there are some cases with much larger deviations. It is notable that the outcomes are less than the estimates (which is consistent with a negative PE) in 4 out of 10 cases for VaR and in 7 out of 10 cases for CVaR. The outcome being less than the estimate can be regarded as a success, since then, the given percentile has not been as bad as expected. One can also note that the largest positive deviations occur during the instable periods of 2007-10 – 2008-09 (the end of it) and 2008-10 – 2009-09, when the global financial crisis took off. No obvious difference in accuracy is between VaR and CVaR.

4.1.1.2 OMXS30

When it comes to the Equity Index OMXS30, the situation is quite similar to that of the Gold Price. As far as VaR and CVaR are concerned, they differ 10-15% on average, and the instable periods (2007-2009) show larger deviations – especially the 119% for the 95%-VaR of 2008-10 – 2009-09. Negative PE's, outcomes smaller than the estimates, which can be regarded as successful predictions, occur in 3 out of 10 cases for VaR and 5 out of 10 cases for CVaR. This is a little worse than for Gold. VaR and CVaR seem to perform equally well.

	Start	End	VaR 95%	PE	VaR 99%	PE	CVaR 95%	PE	CVaR 99%	PE
1	1986-10-01	2006-09-30	2,23%		4,06%		3,29%		5,08%	
	2006-10-01	2007-09-30	2,32%	4,02%	3,77%	-7,22%	3,02%	-7,98%	3,78%	-25,71%
2	1987-10-01	2007-09-30	2,24%		4,06%		3,30%		5,08%	
	2007-10-01	2008-09-30	3,16%	41,22%	4,26%	4,92%	3,84%	16,14%	4,99%	-1,78%
3	1988-10-01	2008-09-30	2,27%		3,87%		3,24%		4,74%	
	2008-10-01	2009-09-30	4,95%	118,67%	6,33%	63,52%	5,65%	74,43%	6,79%	43,22%
4	1989-10-01	2009-09-30	2,41%		4,16%		3,47%		5,14%	
	2009-10-01	2010-09-30	2,34%	-2,72%	3,27%	-21,42%	2,76%	-20,27%	3,31%	-35,63%
5	1990-10-01	2010-09-30	2,40%		4,08%		3,42%		5,04%	
	2010-10-01	2011-09-23	2,52%	4,89%	4,82%	18,06%	3,68%	7,78%	5,78%	14,59%
	Average		2,73%	33,22%	4,29%	11,57%	3,60%	14,02%	4,96%	-1,06%

Table 5 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for OMXS30 for approach 1. VaR and CVaR estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

4.1.1.3 OMRX

	Start	End	VaR 95%	PE	VaR 99%	PE	CVaR 95%	PE	CVaR 99%	PE
1	1990-10-01	2006-09-30	0,25%		0,51%		0,42%		0,75%	
	2006-10-01	2007-09-30	0,13%	-48,36%	0,34%	-33,40%	0,19%	-55,42%	0,36%	-52,12%
2	1991-10-01	2007-09-30	0,25%		0,51%		0,42%		0,75%	
	2007-10-01	2008-09-30	0,21%	-15,47%	0,39%	-23,92%	0,31%	-24,82%	0,46%	-38,07%
3	1992-10-01	2008-09-30	0,24%		0,48%		0,39%		0,68%	
	2008-10-01	2009-09-30	0,21%	-9,60%	0,43%	-10,76%	0,29%	-25,90%	0,44%	-34,91%
4	1993-10-01	2009-09-30	0,23%		0,47%		0,38%		0,67%	
	2009-10-01	2010-09-30	0,20%	-11,77%	0,34%	-28,52%	0,27%	-28,41%	0,40%	-39,30%
5	1994-10-01	2010-09-30	0,21%		0,39%		0,32%		0,49%	
	2010-10-01	2011-09-19	0,21%	0,40%	0,29%	-24,10%	0,26%	-18,65%	0,32%	-34,33%
	Average		0,21%	-16,96%	0,40%	-24,14%	0,32%	-30,64%	0,51%	-39,75%

Table 6 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for OMRX for approach 1. VaR and CVaR estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

OMRX shows better outcomes than the other two for VaR and for CVaR in that the outcomes are more consistent in their PEs, even though their absolute values of the averages are higher than those for Gold and OMXS30. The fact that 9 out of 10 VaR outcomes and all 10 CVaR outcomes fall below the estimates, indicates a certain degree of success on a consistent level. However, if one considers the fact that the early nineties was a very instable period when it comes to interest rates in Sweden – which can be seen in the drop of the estimates for periods towards the end of the study period – it is not surprising to see the outcomes falling short of the estimates on all these occasions. The higher volatility in the beginning of the period that changes the estimates and yields a worse accuracy compared to the outcomes, compared to the less volatile and stable periods. Hence, heteroskedasticity affects the accuracy. In this case, the VaR seems to be performing slightly better than CVaR and the 95% level seems to generate slightly lower PE's than the 99% level.

4.1.1.4 Summing up the approach

To sum all this approach and all matrices above up, the outcomes tend to differ from the estimates – usually quite significantly. However, a majority of the outcomes fort CVaR and VaR (at least for OMRX and Gold) comes in smaller than the estimates, which can be regarded a success – if the worst 1% percentile during a year does not reach the estimates, the investor can be happy.

Except for OMRX, no obvious difference exists between VaR and CVaR. However, one can note that, in general, the outcomes associated with the 95% level of significance have smaller absolute PE's. One possible explanation for this is the fact that a few extreme outcomes affect the 99% values more, especially since the length of the prediction period is only one year, and the 99% level value is based on roughly 2 data points. The 95% level is based on roughly 12 data points, making it less sensitive to extreme outcomes.

After looking at these results – the fact that instable periods give "abnormal" estimates for VaR and CVaR – it seems very intriguing to alter the set-up slightly. It seems apparent at this point that a change in the exchange climate, when an instable period with higher volatility begins, affects the accuracy – as a consequence of the fact that the outcomes become further away from the estimates. A changing volatility is in fact per definition heteroskedasticity. Let's look at the final approach, where subjectively chosen stable (and shorter) periods, with roughly the same volatility, have been chosen to see whether the results get any better.

4.1.2 The final approach

The periods determined by looking at the daily returns and the stableness of them, which can be seen in the methodology section above, resulted in the following estimates and outcomes:

4.1.2.1 Gold

			VaR				CVaR			
	Start	End	95%	PE	99%	PE	95%	PE	99%	PE
1	1986-10-01	2002-09-30	1,17%		2,28%		1,89%		3,34%	
	2002-10-01	2003-09-30	1,55%	31,80%	2,97%	30,22%	2,13%	12,86%	3,14%	-6,02%
2	1987-10-01	2003-09-30	1,20%		2,24%		1,88%		3,26%	
	2003-10-01	2004-09-30	1,77%	46,59%	3,39%	51,19%	2,44%	29,33%	3,41%	4,43%
3	1988-10-01	2004-09-30	1,22%		2,24%		1,90%		3,26%	
	2004-10-01	2005-09-30	1,19%	-2,39%	2,17%	-3,23%	1,56%	-17,86%	2,27%	-30,31%
4	1989-10-01	2005-09-30	1,22%		2,24%		1,89%		3,21%	
	2005-10-01	2006-09-30	2,31%	89,22%	4,72%	110,55%	3,54%	87,78%	5,91%	83,82%
5	1990-10-01	2006-09-30	1,28%		2,36%		2,00%		3,40%	
	2006-10-01	2007-09-30	1,71%	33,59%	3,25%	37,57%	2,21%	10,78%	3,35%	-1,57%
	Average		1,49%	39,76%	2,84%	45,26%	2,17%	24,58%	3,47%	10,07%

Table 7 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for Gold for approach 2. VaR and CVaR estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

If we look at the VaR and CVaR, sadly, the PE's are in general higher than in the first approach – fewer really small PE's are present. They are also more often positive; just 2 out of 10 cases show negative PE's for VaR whereas CVaR have 4 out of 10 negative ones. The period 2005-10 – 2006-09 evidently contains extra negative outcomes, which is seen in the high PE's there. Doing the study over a more stable period is maybe not a way to enhance the results, or the subjective choice of the period wasn't good enough. CVaR seem to be performing slightly better than VaR.

4.1.2.2 OMXS30

	Start	End	VaR 95%	PE	VaR 99%	PE	CVaR 95%	PE	CVaR 99%	PE
1	1992-10-01	2000-09-30	2,05%		3,61%		2,97%		4,48%	
	2000-10-01	2001-09-30	3,53%	72,56%	6,95%	92,36%	4,59%	54,72%	7,56%	68,62%
2	1993-10-01	2001-09-30	2,40%		4,05%		3,35%		4,89%	
	2001-10-01	2002-09-30	3,87%	61,35%	4,55%	12,18%	4,40%	31,07%	5,33%	9,09%
3	1994-10-01	2002-09-30	2,57%		4,29%		3,62%		5,08%	
	2002-10-01	2003-09-30	2,64%	2,78%	3,75%	-12,63%	3,17%	-12,50%	4,07%	-19,92%
4	1995-10-01	2003-09-30	2,70%		4,31%		3,68%		5,08%	
	2003-10-01	2004-09-30	1,90%	-29,91%	3,21%	-25,42%	2,45%	-33,49%	3,39%	-33,28%
5	1996-10-01	2004-09-30	2,72%		4,31%		3,70%		5,08%	
	2004-10-01	2005-09-30	1,27%	-53,17%	2,34%	-45,61%	1,79%	-51,69%	2,66%	-47,61%
	Average		2,62%	10,72%	4,20%	4,18%	3,42%	-2,38%	4,79%	-4,62%

Table 8 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for OMXS30 for approach 1. VaR and CVaR estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

The VaR and CVaR results come out pretty good if you look at the average PE’s. However, they vary and have 4 and 3 PE’s of more than 50% in absolute value, respectively, which is more than in the first study. But, 5 out of 10 and 6 out of 10 of the outcomes for VaR and CVaR respectively, have negative PE’s – a certain degree of success occur more often than in the previous study. There are indications in both directions when it comes to ruling whether the stable period give better results, in this case, even though one can also note that there are fewer really small PE’s here than in the previous approach. No obvious difference in performance is present between VaR and CVaR. Not between levels of significance either.

4.1.2.3 OMRX

	Start	End	VaR 95%	PE	VaR 99%	PE	CVaR 95%	PE	CVaR 99%	PE
1	1995-10-01	2002-09-30	0,23%		0,42%		0,35%		0,54%	
	2002-10-01	2003-09-30	0,24%	5,05%	0,44%	6,60%	0,33%	-4,85%	0,45%	-17,15%
2	1996-10-01	2003-09-30	0,23%		0,41%		0,33%		0,50%	
	2003-10-01	2004-09-30	0,22%	-4,57%	0,31%	-23,79%	0,27%	-17,37%	0,39%	-21,47%
3	1997-10-01	2004-09-30	0,22%		0,38%		0,32%		0,48%	
	2004-10-01	2005-09-30	0,16%	-28,19%	0,22%	-41,30%	0,19%	-40,01%	0,22%	-53,08%
4	1998-10-01	2005-09-30	0,21%		0,36%		0,30%		0,42%	
	2005-10-01	2006-09-30	0,17%	-21,81%	0,23%	-37,75%	0,20%	-31,03%	0,25%	-41,84%
5	1999-10-01	2006-09-30	0,19%		0,33%		0,27%		0,40%	
	2006-10-01	2007-09-30	0,13%	-33,44%	0,34%	1,39%	0,19%	-30,25%	0,36%	-10,25%
	Average		0,20%	-16,59%	0,34%	-18,97%	0,27%	-24,70%	0,39%	-28,76%

Table 9 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for OMRX for approach 1. VaR and CVaR estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

In the case of OMRX, the PE’s for VaR and CVaR are at least not worse than in the previous study. 7 and 10 out of 10 of the PE’s for VaR and CVaR are negative; this is slightly worse than the first study – worse from the perspective that a negative PE is a success. The PE’s are increasing towards the end of the periods, which probably has something to with the way the market developed. But, there are more numerous very small

PE's than in the first study. VaR and CVaR perform equally well, but the 95% significance level outcomes perform better in general.

4.2 Maximum Draw-down

The results for the MDD risk measure, with estimations based on five year averages, are presented per asset class.

4.2.1 Gold

The results for gold can be seen below.

	Start	End	$\overline{MDD}_T$	MDD	PE
1	2001-10-01	2006-09-30	13,40%		
	2006-10-01	**2007-09-30**		**7,53%**	**-43,84%**
2	2002-10-01	2007-09-30	13,35%		
	2007-10-01	**2008-09-30**		**25,49%**	**90,98%**
3	2003-10-01	2008-09-30	15,37%		
	2008-10-01	**2009-09-30**		**22,07%**	**43,65%**
4	2004-10-01	2009-09-30	17,35%		
	2009-10-01	**2010-09-30**		**12,47%**	**-28,13%**
5	2005-10-01	2010-09-30	17,96%		
	2010-10-01	**2011-09-23**		**8,91%**	**-50,39%**
					2,45%

Table 10 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for Gold . MDD estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

The PEs vary quite a bit – even though the average of them is close to zero, they are often 40 percent in magnitude; sometimes positive and sometimes negative. Interestingly, the outcomes are smaller than the predictions ($\overline{MDD}_T$'s) in the first period and in the two latest ones. While they increase towards the later estimation periods, the $\overline{MDD}_T$'s themselves are relatively stable – an effect of the average process. Notably, the instable periods from 2007 to 2009 have the largest MDDs.

The fact that the outcomes are small in the beginning and in the end, while the $\overline{MDD}_T$'s increase slightly, shows that in instable markets (where risk is higher) with larger MDD's, trying to predict the risk measured as MDD based on a five year average is a decent but not splendid solution; 3 out of 5 PE's are negative – a case which could be regarded as a success depending on your view.

One can still state that the $\overline{MDD}_T$ is a better risk estimate in a more stable period, since the PE's are of a moderate size in such periods.

4.2.2 OMXS30

The results for OMXS30 are presented below.

The situation for OMXS30 is similar to that of Gold, but the PE's vary even more. This suggests that the fact that the OMXS30 has had a very high volatility during the past decade (especially in the beginning and towards the end, which is also visible in the $\overline{MDD}_T$'s being larger in the beginning and in the end), and hence larger MDD outcomes, yields larger deviations from the five year average estimates. These vary more than for Gold and are larger in the beginning and the end of the period examined, but are fairly stable.

One can note, as for Gold, that the outcomes for the two instable periods between 2007 and 2009 come closer to the large MDD estimate, which is also reflected in the $\overline{MDD_T}$'s becoming larger once these periods are included in them.

	Start	End	$\overline{MDD_T}$	MDD	PE
1	2001-10-01	2006-09-30	22,17%		
	2006-10-01	**2007-09-30**		**12,69%**	**-42,78%**
2	2002-10-01	2007-09-30	14,79%		
	2007-10-01	**2008-09-30**		**39,38%**	**166,31%**
3	2003-10-01	2008-09-30	17,47%		
	2008-10-01	**2009-09-30**		**27,27%**	**56,08%**
4	2004-10-01	2009-09-30	20,76%		
	2009-10-01	**2010-09-30**		**12,77%**	**-38,47%**
5	2005-10-01	2010-09-30	22,11%		
	2010-10-01	**2011-09-23**		**26,86%**	**21,49%**
	Average				32,53%

Table 11 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for OMXS30 . MDD estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

In particular, one can look at the second period check, where the $\overline{MDD_T}$ was just short of 15% but the 2007-10 – 2008-09 period encountered a draw-down of almost 40%. In such an instable environment, it is not surprising that the PE became very large. Except for this period, the PE's are really quite similar to those of Gold, even though 2 out of 5 PE's are negative. Hence, in stable periods, the $\overline{MDD_T}$ gives a fair estimate of the risk to come. This is clearer than in the case of Gold.

4.2.3 OMRX

Below, the results for OMRX are presented.

	Start	End	$\overline{MDD_T}$	MDD	PE
1	2001-10-01	2006-09-30	1,39%		
	2006-10-01	**2007-09-30**		**1,17%**	**-15,88%**
2	2002-10-01	2007-09-30	1,33%		
	2007-10-01	**2008-09-30**		**2,27%**	**70,08%**
3	2003-10-01	2008-09-30	1,41%		
	2008-10-01	**2009-09-30**		**1,51%**	**7,36%**
4	2004-10-01	2009-09-30	1,38%		
	2009-10-01	**2010-09-30**		**1,42%**	**3,08%**
5	2005-10-01	2010-09-30	1,51%		
	2010-10-01	**2011-09-23**		**2,80%**	**84,90%**
	Average				29,91%

Table 12 Start and End dates for estimate (even rows) and prediction (odd rows) calculations for OMRX . MDD estimates and outcomes for 95% and 99% significance levels followed by their Prediction Errors (PE) for the outcomes.

For OMRX, the PE's are quite small in magnitude except for two periods. Only 1 out of 5 PE's is negative. Interestingly, the $\overline{MDD_T}$'s are quite stable, which might reflect the fact that interest rates are more stable; the MDD's are small in magnitude compared to Gold and OMXS30.

Also in this case, instable periods – with a MDD coming out larger than the five year average – yield a poorer $\overline{MDD_T}$ estimate and a large PE. In the periods where the climate is apparently more stable (smaller MDDs), the five year average does quite a good job in predicting the risk. This is clearer than for Gold.

4.3 General observations, explanations and suggestions for future research

In general, the second approach does not seem to improve the VaR and CVaR PE's much – with the exception of OMRX, which might indicate that this works slightly better for interest rates – but even make them higher. There might be several reasons for this. One is that the estimation periods are shorter, which is a bad thing in general for estimations with respect to statistical inference. Another is that the stable period choices, which were conducted based on a subjective look at returns' volatilities, were not accurate or good enough. The volatility could instead have been measured and stable periods found based on this. Plus, heteroskedasticity could be estimated to come to conclusions with regards to this. Nonetheless, it is not apparent that stable periods imply better risk predictability. More research with regards to this is needed.

From a portfolio and investment riskiness follow-up perspective, it might be considered a success if the bad outcome stated as the unlikely bad outcome by the risk measure is consistently never or rarely reached over time. If no or very few really bad outcomes happen – as defined by the risk measure – the investor can be happy. From this point of view, the result of the study is – at least for most of the time series and especially for the interest rate results – really quite satisfying. Also, the fact that the outcome is usually less than 50% away from the estimate when it comes to VaR and CVaR can be interpreted as those risks being quite decent to estimate the risk in the future. The MDD, which includes averaging in the estimation procedure, does a decent job in predicting the risk, at least in stable periods. Here, heteroskedasticity seems to affect the performance.

Interest rate returns seem to work in a way that makes it easier to create good risk estimates, than those of Commodity and Equity. Perhaps, this is due the fact that the volatility is generally greater for commodities and equities, implying that they are harder to make *any* good estimates for.

Instable periods, or rather: periods which involve extreme market shocks and larger local volatilities, seem to give rise to the VaR and CVaR estimates falling short. The MDD study shows similar behavior, since the five year average $\overline{\mathrm{MDD_T}}$ is closer to the outcome in such periods. Ironically, the risk measures seem to work better for interest rates, which can be interpreted as them working better risky assets; assets where risk management is less important in terms of losses becoming large is less likely.

One can also discuss whether the more fundamental approach of the study should be modified. First, one could argue that the so called prediction period should be longer than one year to see how good an estimate holds, to remove any peculiar effects from ranking quite few outcomes, to turn the test window more similar and to the estimation period in length. Second, perhaps it would add value to the examination to see whether and how many outcomes during the test window come out equal to or worse than the estimates and see this as a means to determine how well the risk measure works. Finally, one could use some sort of averaging scheme for the VaR and CVaR studies – similar to that of this study's MDD study – when creating the estimates. However, this is not in line with common VaR/CVaR historical simulation practice. Nonetheless, it could be of interest.

Interestingly, the 95% level of significance seems to sometimes generate better performance for VaR and CVaR. This could be because of the fact that just 2 data points make up the estimate during the one year prediction period, making few extreme outcomes more influential. This is apparent in both of the two VaR/CVaR approaches.

When it comes to the MDD, even though there is an averaging effect and the fact that five full periods are gone through captures stable and instable years/periods, one could look at periods longer back in time to look at whether stable periods enhances the predictability or not.

Since one can argue that the market has evolved lately (e.g. due to High-Frequency Trading, regulations and more tense and evolved markets in general), the way the risk measures might have worked earlier on need not say anything about how the predictability will be in the future. This is relevant for MDD as well as for VaR/CVaR.

Also, one could change (most likely prolong) the estimation period from five to ten or even more years. However, the averaging effect should take out the impacts of individual years' instability, so this should not alter the $\overline{\mathrm{MDD}_{\mathrm{T}}}$'s significantly. Similarly, one could prolong the prediction period from one to maybe two years.

One could – instead of use averaging over a few precedent years – also simply calculate the estimates based on the MDD over a longer period preceding the prediction periods. However, this would most likely yield estimates that are quite large, since draw-downs are more likely to get larger if one lets the period go over multiple years. This, in turn, will result in the outcomes consistently falling short of the estimates – this is exactly what I found when doing some tests in this manner.

Finally, as far as the MMD measure is concerned, one could try and create a hybrid testing model, where one has a semi-long period of, say, 10 years and creates a 2-year average (taking the average of the five 2-year periods making up the 10 year period) and then compares this to the outcome in the subsequent one year period.

Regardless of the risk measure, a general observation is that the whole thing about choosing a stable period to see whether it works well there is actually quite irrelevant. A risk measure should be usable during any market circumstances. One can argue that it is strange to say that it is fine for a risk measure not to succeed during the financial crisis for instance – it is during such periods that it is the most important for a risk measure to work.

On the other hand, the markets can sometimes be in periods of extreme shocks and somewhat peculiar patterns can be present – like during the financial crisis of 2008-2009 – and it is *understandable* that predicting risk does not work as well as it normally does. One can perhaps not rule out a risk measure completely just because it does not work perfectly during extreme periods; for that reason it can be wise to perform the kind of check that the second approach for VaR and CVaR above and the averaging process in the MDD study aimed at.

When it comes to instable periods, most likely, *no* risk measures can be completely trusted in such environments. What matters then – but also at any given time to be honest – is probably taking wise investment decisions, looking at numerous risk measures and also follow the outcomes closely as time goes by. Also, the risk measure does not tell the whole truth – one should also look at performance over time; a topic not covered by this thesis.

Another thing that could be induced in this kind of study is the frequency of time series measurements. Instead of daily returns, one could imagine looking at weekly or even monthly ones – in which case one needs to maintain long estimation as well as test periods – at least as a complement. A few daily extreme outcomes might affect the picture completely, even though on the whole, if one looks at the development of the portfolio over a little longer period, the picture might look okay. Here, things like skew and partial moments might be helpful.

Furthermore, one could also look at other risk measures to see whether they perform equally well in saying something about how bad things can get for a portfolio. One could also include performance adjusted ones, to see whether the riskiness is compensated for with higher returns in the long run (which is probably true, according to common finance theory and practice, at least in most cases), which is usually of interest even for a risk department. More asset classes than the three included in this study could also be looked at, as well as a larger quantity of different time series in general to create a larger sample.

All of the possible modifications to the study mentioned above could be possible components to future research within this field.

5. Summary

5.1 Conclusion

There exist several risk measures on the market. The most common ones, like beta and volatility, have their disadvantages – like the fact that they are hard to interpret and that they don't say anything about how bad things can get. Higher order moments, like skew and kurtosis, as well as lower partial moments can add more information, especially when it comes to down-side risk, which is what matters. More easily graspable ones, such as VaR and MDD, have become increasingly popular and VaR and CVaR, a coherent measure, have arisen as the modern way of measuring risk from regulators and in the industry as a whole. One should look at more than one risk measure to get the whole picture, remember to revise the models on a regular basis and, most importantly, understand what they say.

The outcome of the empirical part indicates that the risk measures VaR and CVaR can often give a reasonably fair picture of the risk, especially for interest based returns and if you regard outcomes falling short of the estimates as a success. But for some periods, it seems that the outcomes can become significantly different. The 95% level results are usually better than those of 99%, which can be explained by the fact that just 2 data points make up the one year 99% value making extreme outcomes more influential. The Maximum Draw-down measure succeeds better in predicting risk, the results show, even though, in instable markets with the presence of heteroskedasticity, the outcomes become larger than the estimates, not surprisingly. Further research could give more information about exactly how the dynamics behind this work.

5.2 Further Research

Further empirical research, with the theoretical part – which could also be extended with more risk measures etc. – as a starting-point, could include more asset classes and other types of time series (frequency and period length-wise) as well as having more time series or other back-testing methods. For Maximum Draw-down, it would be good to look at testing longer periods than one year, for instance. It could be of interest to include other risk measures and possibly also performance measures to get a wider portfolio- and investment perspective in addition. Also, one can consider evaluating the performance of the risk measures VaR and CVaR by looking at how many times the outcomes become as bad as the estimates say, and add the longer term perspective – not only examining whether a few daily outcomes were bad, but also whether the returns bounced back within a reasonable time period, for instance. Also, averaging schemes in the estimates could be included in VaR and CVaR studies. Finally, more research is necessary to determine whether and how heteroskedasticity affects the accuracy of the predictions.

6. Bibliography

Acerbi, C. T. (2001). *Expected Shortfall: A Natural Coherent Alternative to Value-at-Risk.*

Arrow, K. (1965). *Aspects of the Theory of Risk Bearing .*

Artzner, P. D.-M. (1998). *Coherent Measures of Risk, Pisa Lecture Notes.* ETH Zürich, Switzerland.

Artzner, P. D.-M. (1997). *Thinking Coherently.* Risk Magazine 1997.

Bawa, V. L. (1977). Capital Market Equilibrium in a Mean-Lower Partial Moments Framework. *Journal of Financial Economics 5* , 189-200.

Brachinger, H. (2002). *Measurement of Risk.* University of Fribourg, Switzerland: www.unifr.ch/stat/forschung/publikationen/Braching-Risk02.pdf.

Damodaran, A. (2005). *Value at Risk.* Stern School of Business, New York University.

Diebold, F. (1996). *Converting 1-Day Volatility to h-Day Volatility: Scaling by sqrt(h) is Worse than You Think.*

Hansson, B. (2009). *Preferences towards risk, Method of Choice, Lecture Notes Foundations of Finance.* School of Economics, Lund University Sweden.

Hull, J. (2009). *Options, Futures, and other Derivatives,7th Edition .* Rotman School of Management, University of Toronto, Canada.

Lee, W. R. (1988). Mean Lower Partial Moment Valuation and Lognormally Distributed Returns. *Management Science Vol 34, No 4* .

Lhabitant, F.-S. (2004). *Hedge Funds, Quantitative Insights.*

Linsmeier, T. P. (1996). *Risk measurement: An introduction to Value at Risk.* University of Illinois, Urbana-Champaign .

Magdon-Ismail, M. A. (2004). *An Analysis of the Maximum Drawdown Risk Measure.*

Mewasingh, V. (2006). Downside-Risk Performance Measures and Hedge Funds.

Moors, J. S. (1995). *A New Method For Assessing Judgemental Distributions.* Tilburd University, Netherlands.

Pratt, J. (1964). *Risk-Aversion in the Small and in the Large.*

Prigent, J.-L. (2007). *Portfolio Optimization and Performance Analysis.*

Schwert, G. (1990). Stock Volatility and the Crash of '87. *The Review of Financial Studies, Volume 3, No. 1* .

Sortino, F. S. (2001). *Managing Downside Risk in Financial Markets.*

Sulewski, P. (2008). *On Differently Defined Skewness.* Pomeranian Academy, Arciszewskiego, Poland.

White, H. T.-H. (2008). *Modeling Autoregressive Conditional Skewness and Kurtosis with Multi-Quantile CAViaR.* European Central Bank, Working Paper Series No 957.

Witt, R. H. (2009). *Lower Partial Moments as a Measure of Vulnerability to Poverty in Cameroon.* University of Hannover, Germany: Institute of Development and Agricultural Economics, Discussion Paper No. 434.

7. Appendix

7.1 Python code

The following code has been executed in order to receive the resulting VaR, CVaR and MMD (Maximum Draw-down) measures for the different data series (Gold, OMRX or OMXS30), time periods and levels of significance. Function ael_main is executed, given parameters:

DATEFROM, which defines the start date to be used when creating time series based on entire time series in files (in turn found under paths OMXS30Path, GoldPath and OMRXPath);

DATETO, which defines the end date to be used when creating time series;

FREQUENCY, which defines the frequency under which the time series will be created and risk measured upon;

LEVEL, which is the level of significance (95% or 99%);

DATA, which is Gold, OMRX or OMXS30 and defines on what time series the risks are to be calculated.

NB: Input parameter Frequency has always been "Day" and setup for being able to handle other frequencies have not been fully developed.

```
import os,sys,time

OMXS30Path='H:/Bachelors Thesis/OMXS30-data.txt'
GoldPath='H:/Bachelors Thesis/Guld-data.txt'
OMRXPath='H:/Bachelors Thesis/OMRX-data.txt'


def GetDataDict(Path):

  DataDict={}

  DataFile=open(Path,'r')

  print 'file opened',
  print type(DataFile)

  for m in DataFile:
    Rt=float(m.split('\t')[0].replace(',','.'))
```

```
        Date=m.split('\t')[1].split('\n')[0]
        DataDict[Date]=Rt

    DataFile.close()

    return DataDict


def CompareAndFixDates(OMXS30Dict,goldDict,OMRXDict):

    OMXDates=OMXS30Dict.keys()
    OMXDates.sort()
    GoldDates=OMXS30Dict.keys()
    GoldDates.sort()
    OMRXDates=OMRXDict.keys()
    OMRXDates.sort()

    for date in OMXDates:
        if date not in GoldDates:
            print date,'in OMX but not in Gold'
        if date not in OMRXDates:
            print date,'in OMX but not in Gold'

    for date in GoldDates:
        if date not in OMXDates:
            print date,'in Gold but not in OMX'
        if date not in OMRXDates:
            print date,'in Gold but not in OMRX'

    for date in OMRXDates:
        if date not in OMXDates:
            print date,'in OMRX but not in OMX'
        if date not in GoldDates:
            print date,'in OMRX but not in Gold'


def GetMDD(DataDict,StartDate,EndDate,Freq):

    MDD=0.0
```

```
    dates=DataDict.keys()
    dates.sort()

    Value=1
    ValueList=[]
    DeltaRtList=[]
    DeltaRtPercentList=[]
    n=0
    MDDDict={}
    for date in dates:
        if date >= StartDate and date <= EndDate:
            Value=Value*(1+DataDict[date])
            ValueList.append(Value)
            #print date,Value
            n+=1
    i=0
    while i < n:
        DeltaRt=max(ValueList[0:i+1])-ValueList[i]
        DeltaRtPercent=DeltaRt/max(ValueList[0:i+1])
        DeltaRtList.append(DeltaRt)
        DeltaRtPercentList.append(DeltaRtPercent)
        #print DeltaRt
        i+=1
    MDD=max(DeltaRtPercentList)
    #print max(DeltaRtList),max(DeltaRtPercentList)
    return MDD


def GetCVaR(DataDict,StartDate,EndDate,Freq,Level):

    CVaR=0.0
    dates=DataDict.keys()
    dates.sort()

    n=0
    CVaRDict={}
    #Collecting data
    for date in dates:
        if Freq=='Day':
```

```
            if date>=StartDate and date <= EndDate:
                CVaRDict[DataDict[date]]=date
                n+=1
            else:
                continue
        if Freq=='Week':
            pass
        if Freq=='Month':
            pass

    #print (100-Level)/100.0*n
    BreakPoint=str((100-Level)/100.0*n).split('.')
    print 'BreakPoint:',BreakPoint,'(will be rounded to):',BreakPoint[0]

    WorstList=[]
    m=0
    Rts=CVaRDict.keys()
    Rts.sort()
    #Going through sorted returns
    for Rt in Rts:
        m+=1
        #print CVaRDict[Rt],Rt
        if m<=int(BreakPoint[0]):
            WorstList.append(Rt)

    tot=0.0
    for Rt in WorstList:
        tot+=Rt
        #print Rt
    #print 'Worst:',WorstList[0]
    #print 'VaR / BreakPoint Return:',WorstList[-1]

    CVaR=float(tot/len(WorstList))
    VaR=float(WorstList[-1])
    return VaR,CVaR



ael_variables=[]
```

```
ael_variables.append(['DATEFROM', 'Date From', 'string', None, '1990-04-06',1,0,'Choose Start Date for time series'])

ael_variables.append(['DATETO', 'Date Until', 'string', None, '2006-12-30', 1, 0,'Choose End Date for time series'])

ael_variables.append(['FREQUENCY', 'Frequency', 'string', ['Day','Week','Month'], 'Day', 1, 0,'Choose frequency to sample data at'])

ael_variables.append(['LEVEL', 'Level of confidence', 'string', ['90%','95%','99%'], '95%', 1, 0,'Choose level of confidence for CVaR calculations'])

ael_variables.append(['DATA', 'Data to run for', 'string', ['Gold','OMXS30','OMRX'], 'Gold', 1, 1,'Choose data to run for'])



def ael_main(dict):


  StartDate=dict['DATEFROM']

  EndDate=dict['DATETO']

  Freq=dict['FREQUENCY']

  Level=int(dict['LEVEL'][0:2])

  Data=dict['DATA']

  print '\n\n'

  print 'Level:',Level

  print 'Dates:',StartDate,EndDate

  print 'Frequency:',Freq

  print 'Data:',Data


  #CompareAndFixDates(OMXS30Dict,GoldDict,OMRXDict)


  for timeSeries in Data:
    if timeSeries == 'Gold':
      print '\nGOLD\n'
      GoldDict=GetDataDict(GoldPath)
      GoldMDD=GetMDD(GoldDict,StartDate,EndDate,Freq)
      GoldVaR,GoldCVaR=GetCVaR(GoldDict,StartDate,EndDate,Freq,Level)
      print 'Gold MDD:',GoldMDD
      print 'Gold CVaR:',GoldCVaR
      print 'Gold VaR:',GoldVaR


    if timeSeries == 'OMXS30':
      print '\nOMXS30:\n'
      OMXS30Dict=GetDataDict(OMXS30Path)
      OMXS30MDD=GetMDD(OMXS30Dict,StartDate,EndDate,Freq)
      OMXS30VaR,OMXS30CVaR=GetCVaR(OMXS30Dict,StartDate,EndDate,Freq,Level)
      print 'OMXS30 MDD:',OMXS30MDD
```

```
    print 'OMXS30 CVaR:',OMXS30CVaR
    print 'OMXS30 VaR:',OMXS30VaR

if timeSeries == 'OMRX':
    print '\nOMRX\n'
    OMRXDict=GetDataDict(OMRXPath)
    OMRXMDD=GetMDD(OMRXDict,StartDate,EndDate,Freq)
    OMRXVaR,OMRXCVaR=GetCVaR(OMRXDict,StartDate,EndDate,Freq,Level)
    print 'OMRX MDD:',OMRXMDD
    print 'OMRX CVaR:',OMRXCVaR
    print 'OMRX VaR:',OMRXVaR
```