LUND UNIVERSITY
School of Economics and Management

An empirical evaluation of Value-at-Risk during the financial crisis

Department of Business Administration
BUSM26 Degree project in Finance I VT10
Author:
Nicklas Norling
Daniel Selling
Advisor:
Göran Andersson

Abstract

In the latest financial crisis, risk management and forecasts of market losses played a crucial role in the area of finance. This thesis evaluates the theory of Value-at-Risk through a quantitative study of two non-parametric approaches and three parametric: Basic- and volatility-weighted Historical Simulation, Normal distribution, Log-normal distribution and Student's t-distribution. The thesis compares 1-day VaR estimates predictive performance of market losses defined by the index of Standard & Poor's 500 on a 99% and 95% confidence level. The study is made by a rolling window forecast between 2005-12-28 and 2008-12-31 which includes 786 observations that accommodate one tranquil and one crisis period. Performance is evaluated by backtesting, using the Kupiec's test. The result shows that VaR as a risk measure is very imprecise in its forecasts; however the predictability is different across the five approaches. Under the tranquil period all approaches were significant on the 99% significance level on the 95% however some approaches were not, due to overestimates of the risk. Another finding among the approaches is that they fail to account for rapid shifts in risk. During the crisis period all VaR models included in the backtest was rejected at both confidence intervals. This means that VaR is proven to serious destabilizing during times of crisis. The failure of non-parametric approaches is mostly due to that they are determined by a historical distribution which reflects the past market climate that not always correspond with the climate tomorrow. We also conclude that the parametric approaches rely too heavily on unrealistic assumptions of future distribution and therefore fail as well.

Keywords: Value-at-Risk, Backtesting, Kupiec's test, Historical Simulation, Normal distribution, Log-normal distribution, Student's t-distribution, GARCH(1,1), Basel.

Table of Contents

List of Tables and Figures

1.0 Introduction

This chapter gives a background to the subject treated. We explain the emergence and the growing need for reliable risk measures. Further on, the purpose of the thesis along with the delimitations is presented.

1.1 Background

The emergence of financial risk management is due to several factors. One factor is the tremendous growth in trading activity since the beginning of 1960's, where we have seen a growth of shares traded per day from $4 million in 1961 to $1.6 trillion in the early 2005 in the New York Stock Exchange. Another is the growth of trades in financial derivatives and the development of new types of instruments e.g. exotic options and CDS contracts. We have also seen that markets occasionally have been extremely volatile with increasing uncertainty e.g. Dow Jones fell 23% on October 19th 1987 and erased over $1 trillion in equity capital; other western stock markets have experienced similar falls. We have furthermore experienced occasional extreme exchange rate events, the crisis of the peso 1994, the ruble crisis of 1998, Argentina in 2001 and Brazil in 1999, just to mention some. Commodity markets have also been volatile, where most commodities could have had a stable price for a long period and suddenly become a victim for a large drop. Not to mention the occasional major fluctuations in the interest rates and the emergences of off balance sheet activities. These events together played its role in increasing the demand for proper risk management tools.[1]

With the background of crises in emerging markets, huge losses from trading activities by institutions such as Long Term Capital Management Fund and Orange County, Value-at-Risk (VaR) was developed as an instrument to understand and manage market risk.[2] VaR has been widely used since 1993 and is one of the most popular methods for estimating market risk. Its popularity is mostly due to its simplicity and easy interpretation.[3] VaR finds its use among many different practitioners such as hedge funds, pension funds, investment banks, commercial banks and other financial institutions. Besides from VaR fundamental use, measuring risk, it also has other usages such as controlling and managing risk. It can be used for investment decisions, allocating capital and satisfying external regulations as well.

---

[1] Dowd (2005) p.2-4
[2] Kuester (2006)
[3] Dowd (2005) p.10

Different regulators have grown interested in VaR. In 1995 the Basel Committee proposed that banks could calculate their own capital requirements by using VaR with certain criteria's stated by the committee.[4] The US Federal Reserve also proposed that banks would be allowed to calculate their capital requirements using VaR with penalties if the losses exceeded the calculated capital requirement. The European Union's Capital Adequacy Directive allows VaR to be used for calculating capital requirements for foreign exchange positions.[5]

Mismanagement of risk was one of the reasons to the recent financial crisis between 2007 and 2008 which tumbled the whole financial system. The internal models that managers relied on in their risk management underestimated the risks which lead to incorrect decisions regarding capital requirements and other safety procedures. This however caused financial distress to many leading financial institutions, whereas some even went bankrupt. The fall of Lehman Brothers, that was the largest default in history, combined with other big defaults of financial institutions such as Bear Sterns, Merrill Lynch and the Icelandic banks Kaupthing Bank, Glitnir banki and Landsbanki Islands proves the severity and the magnitude of the crisis that has it consequences even today (2010).[6] The demand for a risk measure that is precise and reliable in its forecasts has never been bigger. Proper risk management is essential, it may be expensive but it is very important to prevent distress, crisis and excessive risk taking. The financial crisis put the risk measure VaR on the edge. Is it a good risk measure or not? It has been compared to an airbag that works every time except when there's an accident.[7]

One of the problems with VaR is that it can be estimated by several approaches that most of the time give different values. This is making it hard to compare VaR outputs generated from different approaches with each other; it also makes it problematic to decide which estimate we should rely on in our risk management, and the probability for an under or overestimates increases. Financial institutions also have incentives to base their estimates on VaR approaches that forecasts a lower risk, and therefore minimizing their capital requirements i.e. regulatory arbitrage. Therefore it would be interesting to see if there is one VaR approach that is better than the others and therefore should be used. There has also been disagreement among practitioners regarding if VaR is a good risk measure or if it actually works. What we on the other hand know for sure is that Value-at-Risk plays an extremely important role in the

[4] Basel committee on banking supervision (2006)
[5] Linsmeier (2001)
[6] Moody's (2009)
[7] Einhorn (2008)

financial system today; which motivates further studies in the area of VaR and an investigation of the problems regarding it very relevant.

1.2 Purpose

The main purpose of this thesis is to convey better knowledge of the forecasting ability of VaR during times of crisis. During financial crisis risk measures generally tend to fail to incorporate the increasing risk whilst in tranquil period’s risk measures perform better. Especially under crisis the models fails to take account of the shifts from a tranquil to unstable period which make further research into the area interesting. By using a broad coverage of different methodologies we aim to answer if VaR have provided reliable forecasts of the increased risk during the financial crisis.

Our secondary purpose is to focus on which of some estimation approaches that will generate the best VaR forecast according to Kupiec’s test, we will explain if there is one universal VaR approach that should be used by financial institutions in order to provide correct estimates instead of estimates that are in line with their interest. Considering that VaR can be calculated by a range of approaches stretching from parametric to non-parametric approaches it would be interesting to have a closer look at some of the five most common too see if we could find if any approach works better.

Hence the purpose of the thesis is to answer the following questions:

Have VaR worked as a risk measure during the financial crisis based on our chosen methodologies?

Which of our chosen forecast methods will be the most significant according to Kupiec’s test during our time horizon?

1.3 Delimitations

In order to provide quality to this thesis we have to make certain restrictions, understandable we cannot therefore cover all aspects of VaR. There are numerous of different methods to forecast VaR and we will focus on some of the most common. Hence we will limit our thesis to contain the calculation of two non-parametric and three parametric methods. The reason for this is to cover a broad spectrum of methods from the two most common categories of approaches; thereby we will get estimates of VaR that are based on different underlying assumptions regarding the data distribution. We choose to work with Basic Historical Simulated VaR, volatility-weighted Historical Simulated VaR, normal distributed VaR, log-normal distributed VaR and Student's t-distributed VaR. We will also limit the study to stock returns approximated by the index of S&P 500. We believe the financial crisis and the market risk will be reflected well in the S&P 500 since it includes the biggest companies in the United States, which were the core of the financial crisis.

2.0 Theory

In this section we present a brief review of the empirical findings of previous studies. Later on we convey the theoretical framework of VaR along with some major advantages and drawbacks. We give a presentation of the non-parametric and parametric approaches and highlight the most important differences between these two approaches. The backtesting method, Kupiec's test, is described and finally the Basel Capital Accord framework and its influence of VaR are presented.

2.1 Previous studies

There have been numerous of studies about Value-at-Risk in the past. Studies about the best forecast method have also been a popular research area. This thesis could be of interest because of its up-to-date relevance concerning the financial crisis and it could be valuable to know which VaR forecast method is best applied when a crisis is at hand. Below we give an introduction to studies that are related to our thesis.

Beder (1995) applied eight different VaR approaches on three portfolios and stated that VaR relies heavily on the input parameters, data, assumptions and methodology. Her findings surprised managers, since the firms risk report can change considerably under other assumptions. She proved that there is no such thing as one universal VaR estimate; she also clarified that banks could use different VaR methods to match their purposes i.e. regulatory arbitrage. The conclusion was that VaR does not provide any certainty or confidence but rather expectation of outcomes based on certain assumptions.[8]

Another study, Bao et al. (2006), used several of VaR models on the stock market of five Asian economies that were victim of the financial crisis 1997-1998. They showed that risk forecasting under crisis periods are more difficult and generally yields poorer results than during tranquility periods. They showed that most VaR models behaved similarly before and after the crisis but differently during the crisis. Risk Metrics provided good estimates in tranquil periods while models based on the Extreme Value Theory generated the best estimates during the financial crisis.[9]

[8] Beder (1995)
[9] Bao (2006)

Berkowitz and O’Brien (2002) examined the accuracy of commercial banks VaR forecasts in their trading activity by collecting sample returns for six large banks. The article is the first to provide a complete analysis of the performance of models actually used by banks in practice. The banks that are examined are all multinational players and meet the Basel "large trader" criterion i.e. the trading activity is at least 10 percent of total assets or $1 billion. The study compares the banks VaR forecast with a GARCH based forecasting model. For testing their hypothesis they are using a Kupiec’s test with Christoffersens modification, the same as we are supposed to do. The conclusion of the article is that the banks VaR did not outperform standard methods based on the GARCH model.[10]

Hendricks (1996) evaluated twelve different VaR approaches on 1000 randomly chosen foreign exchange portfolios. He concluded that extreme outcomes occur more often than a normal distribution can predict and therefore it cannot produce reliable estimates. He also stated that due to the nature of the market movements, where the size of market movements is not steady over time, that it is hard to find a model that performs well all the time. Therefore he could not suggest any solely VaR approach that is superior towards the others at all market climates.[11]

Lin et al. (2006) showed that student t-distribution improved and gave more accurate VaR estimates than other methods like the normal VaR especially when the significance level exceeded 98.5%. This is due to the fact that financial data, like stock returns, are often fat tailed distributed and VaR is thus better described by a t-distribution that accommodates this behavior to avoid an underestimation. [12]

2.2 Value-at-Risk

Value-at-risk is defined as the maximum value we risk to lose for a given confidence interval during a certain time horizon. We define the mathematical expression below were R stands for “return” and α is the level of confidence.

$$Pr(\mathrm{R} \leq -VaR) = 1 - \alpha$$

[10] Berkowitz (2002)
[11] Hendricks (1996)
[12] Lin (2006)

The level of confidence will of course affect our VaR estimate. Different levels will fit different firm's needs; the level of the confidence interval could be decided by the management's risk-attitude, whereas a risk averse manager would prefer a higher confidence interval. On the other hand for validation purposes we would often prefer a lower confidence interval; the backside with using a high confidence interval is that they make losses in excess of VaR rare which means that we have to wait a longer time to accumulate data with enough excess losses to be able to draw reliable results in our validation. [13]

The time horizon will also affect VaR, the longer the time horizon the higher VaR we will get assuming that the confidence level remains the same. Also the time horizon is matched to fit the firm's needs, for a bank the 1-day VaR could be of interest since they have highly liquid currencies while for a pension fund a 50-day VaR might be more appropriate. For hedging purposes it is preferable to match the Value-at-Risk horizon with the financial instrument used when hedging. Most often a short time horizon is used because of several reasons, to justify a normal approximation, to accommodate changes in a portfolio and we often prefer a short time horizon for validation purposes as well. [14]

2.2.1 Advantages

There are several reasons to why VaR have become a popular risk measure, below some of these features are listed that make VaR superior to its predecessors.[15]

- VaR has a simple interpretation; it is expressed as "money-at-risk" which makes it very easy to communicate to non-economists as well.
- VaR is probabilistic, which gives the manager useful information about the probability associated with different loss values.
- VaR is holistic i.e. it takes into account all different aspects of risk, many other risk measures looks at single risk factors one at a time.
- VaR can be applied to any type of a portfolio which allows us to compare the risk of investments between each other e.g. a fixed income positions risk can be compared with an equity position.
- VaR allows for risk aggregations and hence decomposing of risk.

[13] Dowd (1998) p.52-53
[14] Dowd (1998) p.50-52
[15] Dowd (2005) p.11

2.2.2 Disadvantages

Ever since the introduction of VaR it has grown in popularity among many practitioners and have become one of the most widely used risk measures, but one should be aware of its drawbacks and the criticism that have been raised towards it. We list some examples on disadvantages below.[16]

- VaR estimates are too imprecise, empirical evidence suggests that different VaR approaches can give diverse estimates. The problem here is oblivious; relying on too inaccurate estimates could lead managers to take on much bigger risk and risk to lose more than they had taken into account.
- VaR is sensitive to incorrect assumptions about the underlying distribution.
- VaR is silent about the size or magnitude if a loss greater than VaR occurs.
- If VaR is used to control risk-taking by for example traders, then there will be an incentive from the trader's part to seek trade positions where risk is over or underestimated. They might therefore take on more risk than suggested by VaR which would lead to a downward biased VaR.
- When using VaR for calculating capital requirements for banks they have incentive to find regulatory arbitrage.
- VaR is not coherent except under the normality assumption.

The latter one is the most crucial disadvantage and need some further explanation. We define a coherent risk measure by four different features where ς is some risk measure and R is the return. The most important quality VaR lack of is sub-additivity, or more intuitive, the diversification effect.[17]

- Monotonicity $R_1 \leq R_2 \rightarrow \varsigma(R_1) \geq \varsigma(R_2)$
- Sub-additivity $\varsigma(R_1 + R_2) \leq \varsigma(R_1) + \varsigma(R_2)$
- Positive homogeneity $\varsigma(hR) = h\varsigma(R)$
- Translation invariance $\varsigma(R + n) = \varsigma(R)$+n

[16] Dowd (2005) p.12
[17] Dowd (2005) p.33

2.2.3 Parametric versus non-parametric methods

As mentioned earlier there are different ways to calculate VaR. The most common way is either by assuming a distribution i.e. using a parametric approach or by using the actual data series in a non-parametric approach. When using a parametric approach we assume that we would be able to draw reliable VaR estimates from a theoretical distribution, like the normal-, Student's t- and log-normal distributions. On the other hand a non-parametric approach uses the historical data returns to forecast future VaR. In almost every case we would expect the VaR estimates to be different depending on which method we choose. It is therefore of interest to make a distinction between parametric and non-parametric approaches and provide some clarification of in what areas the one is preferred to the other.

There have been studies that show that non-parametric approaches have a good track record and are often superior to parametric approaches based on plain assumptions such as the normal distribution. The advantage with this approach is that we let the data speak for itself instead of assuming a theoretical distributions, this is relied heavily on the assumption that the past will reflect the near future and this assumption may not always be valid. Non-parametric approaches are preferable when the climate is fairly stable, the reliability might be jeopardized on the other hand under unstable climates because it relies on the historical data distribution. One drawback is if the period of our data set were unusually quite or unusually volatile, estimates based on this period would probably be misleading. Non-parametric approaches have also gotten criticism of not being able to handling shifts, such as a permanent change in exchange rate risk.[18]

The main advantage of the parametric approaches is that they give us a lot of information based on limited assumptions. Generally if the underlying assumptions of a parametric approach are correct they would expect to generate better estimates than their counterpart. Their main weakness is on the other hand the opposite, if the assumptions are not fulfilled, as with using a normal distribution on bond data. Not fitting the data with a proper distribution is one crucial mistake and would most certainty lead to incorrect estimates. It is therefore important to take account for skewness and kurtosis in our decision of choosing a parametric approach.[19]

[18] Dowd (2005) p.99-101
[19] Dowd (2005) p.182

2.3 Non-parametric approaches

2.3.1 Basic Historical Simulation

Basic Historical Simulation (HS) is a non-parametric approach of estimating VaR. The basic idea behind the method is that we use a historical distribution of returns for an asset to simulate a VaR. If you for example have 1000 historical return observations and want to estimate VaR with a confidence interval of 99% means that we take the eleventh largest loss to be representative for the VaR. The VaR estimate when using the historical simulation is calculated by choosing a quantile in the sample using the formula (2.2) below.

$$VaR = N(1 - \alpha) + 1$$

Some of the advantages of the method is its simplicity, and that much of the necessarily data are often public available, but perhaps the biggest advantage with using the HS is that it does not depend on any assumptions regarding the distribution. Therefore the HS is less restrictive than its counterparts (parametric approaches), and since it does not assume any distribution it is generally better with accommodating fat tails. [20]

Although its advantages the method has some major disadvantages. Historical Simulation assumes that returns are *IID* (Independent and Identically Distributed) and the empirical return distribution is the same as the forecasted distribution. The method gives equal weights to all the observations which means that old observation are treated as relevant as newer observations. For stock returns this is not especially realistic and could be overcome by age-weighting the observations with some decay factor. The HS do not either reflect events that may occur in the future but did not actually occur in our historical data set. For example if no devaluation have occurred during the five year period that we base our VaR estimates on, the HS then accounts for a lower exchange rate risk which not necessarily might be the case for the future. In short, if we base our HS on misleading data our VaR estimates will be biased and therefore not a good proxy for the risk tomorrow. [21]

2.3.2 Volatility-weighted Historical Simulation

Due to the fact that it is counterintuitive and statistically wrong to use volatility measures based on assumptions of constant volatility when the data series, in our case stock returns,

[20] Dowd (1998) p.99-100
[21] Dowd (1998) p.101-102

moves over time we also apply a model for changing volatility by extracting estimates from the return data.[22]

To be able to forecast the volatility-weighted VaR we have to estimate future volatility in the returns conditional on past information, denoted Ω_{t-1}, and we choose to do this by the Generalized Autoregressive Conditional Heteroskedasticity model (GARCH(1,1)) which is a modification of the ARCH model developed by Engle 1982.[23] The definition of a GARCH(1,1) is as follows by these equations. [24]

$$Var[\varepsilon_t|\Omega_{t-1}] = \sigma_t^2$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

$$r_t = \mu + \varepsilon_t \rightarrow \varepsilon_t = r_t - \mu$$

The major advantage of this model is that it incorporates recent volatility and return shocks in a higher grade than older volatility and this is consistent with empirical findings of volatility clustering effects. One disadvantage however is that the history not necessarily has to repeats itself in the future; therefore relying on the historical distribution might lead us to false estimates.[25] When applying a GARCH(1,1) model we have to estimate the parameters from our sample. This process is done by using the method of Maximum Likelihood where the formula to be maximized is shown below.

$$\ln L(\cdot) = \sum_{t=1}^{T}\left[-\frac{1}{2}ln(2\pi\sigma^2) - \frac{\varepsilon^2}{2\sigma^2}\right]$$

The volatility-weighting of the returns are done by forecasting the volatility one day ahead and dividing with the volatility today by the following equation.

$$r_{t,i}^* = \left(\frac{\sigma_{T+1,i}}{\sigma_{t,i}}\right)r_{t,i}$$

To do this we transform the standard form of GARCH(1,1) equation above to forecast the volatility one time step ahead to be able to implement this method. A high β indicates that the

[22] Campbell et al. (1997) p.481
[23] Engle (1982)
[24] Bollerslev (1986)
[25] Campbell et al. (1997) p.482

volatility is persistent and a high α means that volatility is thorny and reacts fast to market changes which are exactly what we want to achieve.[26]

$$\sigma_{T+1}^2 = \omega + \alpha\varepsilon_t^2 + \beta\sigma_t^2$$

The standard GARCH model assumes our return series to be stationary and have a long-run variance to convert towards. This is due to the restrictions below.

$$\omega \geq 0, \alpha, \beta \geq 0, \alpha + \beta < 1$$

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}$$

2.4 Parametric approaches

2.4.1 The Normal distribution

The normal distribution is symmetric around its mean, with small tails. The normal distribution often takes values close to the mean value and very rare takes on values with a big deviation, which is why the distribution has a bell-shaped curve that can be described by the function below. The normal distribution is widely used, and much of its attraction is because of the only two variables we need to calculate VaR is the mean μ and standard deviation σ. However the normal distribution fails to respect the maximum loss constraint which might lead to serious overestimating of the actual risk. Another drawback with the approach is that it is not consistent with extreme value theory.[27] We define the calculation formula of the normal distributed VaR below.

$$VaR = -\mu_R - z_\alpha\sigma_R \,; z_\alpha < 0$$

$$pdf\colon \phi(R) = \frac{1}{\sqrt{2\pi}\sigma_R} e^{-\frac{1}{2}\left(\frac{R-\mu_R}{\sigma_R}\right)^2}$$

[26] Dowd (2005) p.132

[27] Dowd (2005) p.154-158

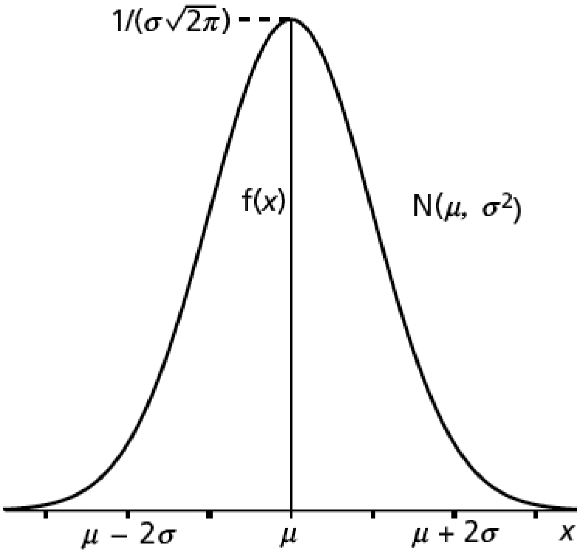


Figure 1: Normal distribution[28]

2.4.2 The Log-normal distribution

In the case of a log-normal distribution, log-returns are used instead of arithmetic returns. Only positive values are possible and the distribution is skewed to the left. The only parameters we need to calculate VaR is the mean μ and the standard deviation σ. The computation of the log-normal VaR is given by:

$$VaR = (1 - Exp\left[h\mu_R - \sqrt{h}\sigma_R Z_\alpha\right])$$

$$pdf: (x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, x > 0$$

One of the greatest advantages of the log-normal distribution is that it takes account of the maximum-loss constraints on long positions. Therefore a portfolio with a positive value can never become a negative-value portfolio as might happen with the other two parametric approaches. Another advantage is that it is consistent with a geometric Brownian motion process.[29]

[28] http://www.answers.com/topic/normal-distribution
[29] Dowd (2005) p.161-163

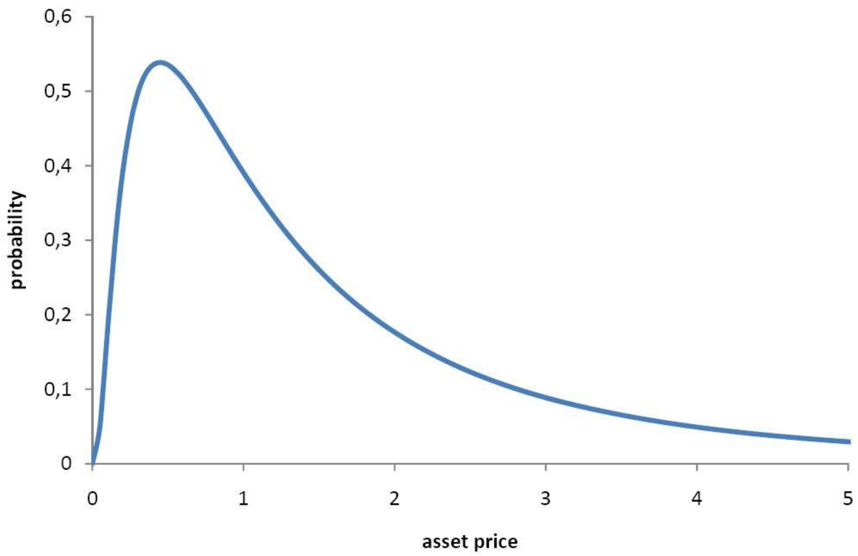


Figure 2: Log-normal distribution[30]

2.4.3 The Student's t-distribution

For the Student's t-distribution, unlike the normal distribution, the probability of a tail event is much higher. An advantage with applying the Student's t-distribution is therefore that it accommodates fat tails and excess kurtosis in the sample. Fat tails is often present in financial data like bond or stock returns, using a normal distribution in these cases will therefore often underestimate the risk.[31] A number of studies have also observed that the t-distribution fit return data better than the normal distribution. These studies have shown that actual losses exceeded normal distributed VaR estimates more often than they exceeded Student's t-distributed VaR estimates. [32] When the number of degrees of freedom goes towards infinity the Student's t-distribution goes towards being normally distributed. We obtain the degrees of freedom υ from the approximation formula below where κ is the kurtosis.

$$\upsilon = \frac{4\kappa - 6}{\kappa - 3}$$

$$VaR = -\mu_R + \sqrt{\frac{V-2}{V}}\, t_{\alpha,V}\sigma_R$$

[30] Dowd (2005) p.51
[31] Lin (2006)
[32] Dowd (1998) p.44

$$pdf: f(R) = \frac{\Gamma\left(\frac{V+1}{2}\right)}{\Gamma\left(\frac{V}{2}\right)\sqrt{V\pi}\sigma_R}\left(1+\frac{1}{V}\left(\frac{R-\mu_R}{\sigma_R}\right)^2\right)^{-\frac{(V+1)}{2}}$$

The Student's t-distribution, like the normal distribution, fails to respect constrains on maximum possible losses which may produce misleadingly high risk estimates. The Student's t-distribution also has the drawback when using very high or very low significance levels of not being consistent with extreme value theory. [33]

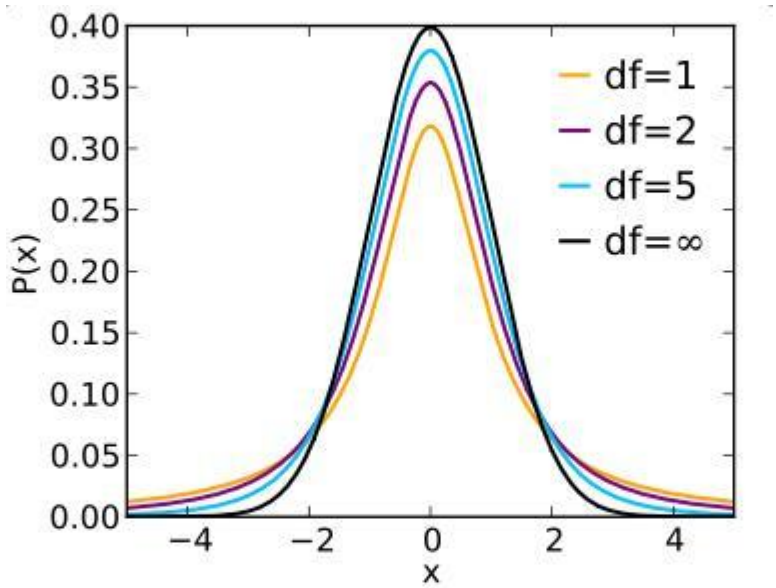


Figure 3: Student's t-distribution[34]

2.5 The Kupiec's test

There are many sources of error when estimating VaR, whereas some of the most common is sampling errors, data problems, using inappropriate assumptions and inappropriate dynamic specifications. Having this in mind our VaR estimates will often be biased in some way, hence being too low or too high. It is therefore of great importance that we have some method to monitor our VaR on a regular basis, to backtest that we actually get reliable estimates of VaR. We therefore turn our attention to the Kupiec's test.

[33] Dowd (2005) p.159-160

[34] http://en.wikipedia.org/wiki/Student's_t-distribution

The Kupiec's test is a statistical test and a standard way of backtesting different models used to forecast VaR. Backtesting means that you test e.g. some forecast to the actual occurred event, in our case stock returns. This is important because we need to determine whether the model is consistent with the assumptions the model is based upon. By calculating the number of actual losses exceeding modeled and forecasted VaR and compare with the expected number of losses that exceeds VaR we are able to test statistically whether to reject or accept a particular model used to forecast VaR. If the model used produces significantly more or significantly less losses exceeding VaR, the Kupiec's test will reject the model. Kupiec assumed that the losses follow a binominal distribution. The equation is shown below where x is the number of exceedances and n is the number of observations. p is the probability of exceedances. [35][36]

$$Pr(x|n,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

By assuming that the VaR model predicts the exceedances to be *IID* we can modify the method above and use a procedure proposed by Christoffersen. To adopt this approach we write the earlier equation in likelihood ratio form and assume it to be $\chi^2 -$ distributed. For simplicity we continue calling the test procedure a Kupiec's test. In the equation below the number of losses exceeding VaR is denoted as$x$; the sample size as n and $q = 1 - \alpha$ where α is the level of confidence. This equation will give the likelihood ratio test-statistic which we assume is χ^2-distributed with one degree of freedom. This makes it possible to compute the *p*-value of the Kupiec's test which can easily be interpreted. Remember that this can never prove a theory completely correct or incorrect, since past results do not necessarily indicate future results. [37]

$$LR = -2ln[q^x(1-q)^{n-x}] + 2ln\left[\left(\frac{x}{n}\right)^x \left(1-\frac{x}{n}\right)^{n-x}\right] \sim \chi^2(1)$$

2.6 The Basel Capital Accord

Financial institutions often play an important role in the economy and their health is reflected on the whole financial system. When a financial institution defaults and goes bankrupt it affects other financial institutions like a domino effect. Therefore banks' defaulting is a threat

[35] Dowd (1998) p.55
[36] Kupiec (1995)
[37] Christoffersen (1998)

to the stability of the entire financial system and have created a need for regulations of banks risk management. Besides from stabilizing the financial market the purpose of the regulation is to protect any parts that are dealing with the bank.

The common framework for treating the risk associated with financial institutions was presented in the Basel Accord of 1988. The accord was primarily designed for handling credit risk and had major drawbacks for its way of treating trading risk e.g. short positions and positions in government holdings were not included. The Basel Accord of 1988 set the capital requirement to be 8% of the private sector assets regardless of the maturity which was a major drawback. Because of this short-maturity private sector bonds were penalized more than corporate bonds with longer maturity. The 1996 amendment to the capital accord incorporated market risk was introduced as a solution to the problems of the accord of 1988. The new accord suggested that capital requirements could be calculated by two methods either by the standardize method or by the banks internal VaR models.

One big drawback with the standardize method is its additive clause, stating that separate capital requirement should be used for each position and summing them into one capital requirement and thereby not taking account for the diversification effect across different asset classes since it assumes a perfect correlation across different assets. The internal approach on the other hand takes into account the diversification effect on risk, and thereby does not systematically overestimate the risk therefore most banks choose to use this method. However there were some general guidelines that have to be met in order to use the method. According to the Basel Committee, banks must have a sound risk management system that has to be integrated into management decisions. The bank also has to have external audits and independent risk control units.

VaR is a standard benchmark today for banks in their capital requirements calculation, but in order to use it some requirements must be fulfilled. The Basel Committee suggests that banks should use a 99% confidence interval over a 10-day horizon. The capital requirement is then given by multiplying the resulting VaR with a safety factor of 3. Without the safety factor banks would once every four years loose worse than the VaR estimate occur. This would be unquestionable and is why a safety factor of three is motivated to provide near absolute insurance against bankruptcy. Presumably the multiplicative factor also accounts for the risk of modeling error e.g. banks may rely their capital requirements calculation on parametric distributions that misfits the data.[38] Furthermore the observation period should be based at minimum one year of historical data and updated on a quarterly basis. The treatment of correlations could be recognized within a category e.g. fixed income or across categories like between fixed income and currencies. A penalty could occur if the backtesting shows that the banks internal models incorrectly forecasted the risk. This gives banks incentives to improve the prediction and accuracy of their models and aims to avoid over optimistic projection of profits and losses. Below are the penalty structure described in a table. Important to undermine is that the number of penalties is based on the amount not the magnitude either individually or cumulatively. In short the colors reflect in which the degree the model works, where green indicates that the model is reliable and red indicates that the model should not be used anymore and that severe punishment may occur if using such a model, where the ultimate penalty is a temporary or permanent suspension of trading activities.[39]

Zone	Number of Violations
Green	0-4
Yellow	5
	6
	7
	8
	9
Red	10+
Note: The number of violations is given for 250 business days.	

Table 1: Basel accord penalty zones

[38] Jorion (2000) p.119-120
[39] McAleer et al. (2009)

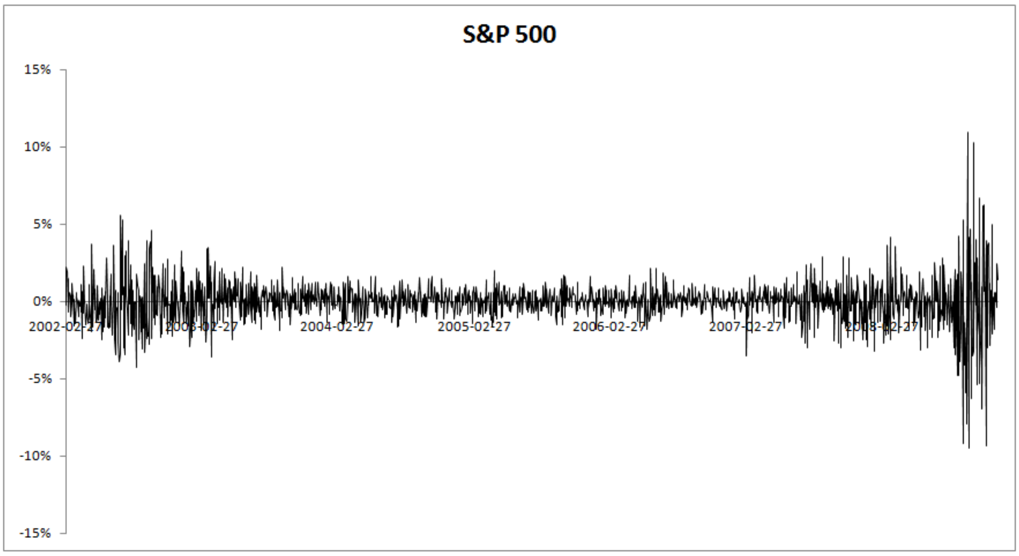
3.0 Methodology

This section begins with specification of our data of choice, where we highlight some important statistical features. The methodology of the working progress is described explicitly in form of application of the theory described earlier. As a final point the validity and reliability is discussed.

3.1 Data

The data represented in the thesis is from the stock index Standard & Poor's 500 and is a daily index ranging from February 27th 2002 to December 31th 2008. S&P 500 consists of the 500 largest publicity traded companies on the US stock market. The index is a composite price index which means it is weighted by market capitalization with reinvested dividends.[40] We use this index as a proxy for stocks as the S&P 500 is one of the most commonly used benchmarks for the overall U.S. stock market and often used in scientific papers. The data is second hand data and was retrieved from Datastream. As we can see in the diagram below we can observe volatility clustering as we expected, especially during the financial crisis. The equation we use to calculate the logarithmic returns are shown below. We use daily returns because we want 1-day VaR estimates.

$$r_t = ln\left(\frac{I_t}{I_{t-1}}\right)$$



[40] http://www.standardandpoors.com/indices/sp-500/en/us/?indexId=spusa-500-usduf--p-us-l--

Figure 4: S&P 500 returns[41]

3.2 Time period definition

We define two periods for our data set, one tranquil and one crisis period. The crisis period is set by looking at the TED-spread. The TED-spread is the difference between the 3-month London Interbank Offered Rate and the 3-month US Treasury bill. It is arbitrary but it is motivated because the TED-spread rose sharply in the beginning of July 2007, an event that can be interpreted as the market, interbank market especially, became very anxious about their trading partners liquidity. We think the end of 2008 will be a fair estimate of the end of the "immediate" financial crisis because the TED-spread fell back to a more modest level. The crisis period is thus set between the 2nd July 2007 to the end of 2008, including 393 observations. Our tranquil period will be defined as the 393 observations before that.

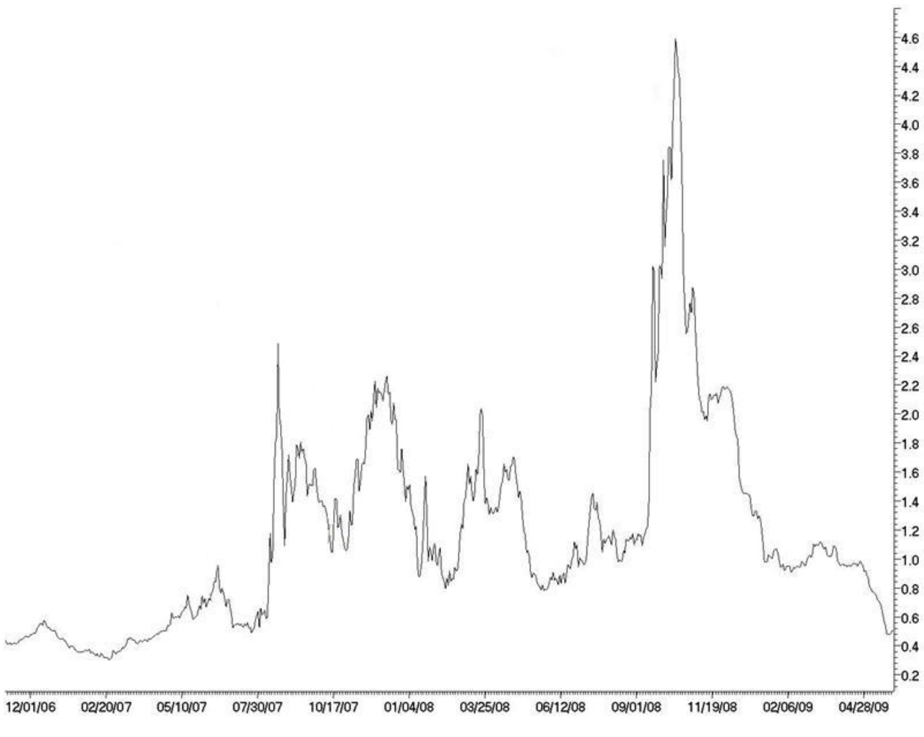


Figure 5: TED-spread[42]

[41] Datastream
[42] Reuters

3.3 Descriptive statistics

We analyze the properties of our data to achieve a better knowledge of the distribution which is relevant when evaluating our result later on. Financial data is often miss-fitted by a normal distribution because of this behavior; therefore it is of interest to determine the distribution by a Jarque-Bera test. We can see that the Jarque-Bera test statistic make us reject the null hypothesis that the data is normally distributed in both cases.[43] The sample data is thus rather leptokurtotic distributed with very high kurtosis and fat tails. A leptokurtotic distribution has higher probabilities for extreme values. A test for non-stationary is also conducted in order to see if the assumption of stationary for the GARCH model is fulfilled. By the result given in Appendix we reject the null hypothesis that we have unit root and thus non-stationary.

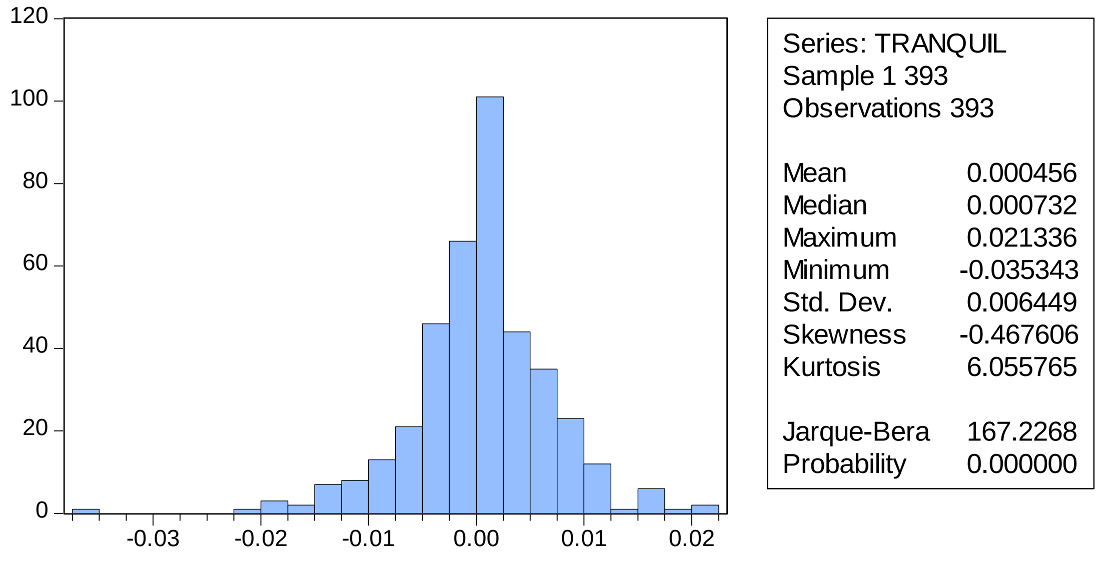


Figure 6: Jarque-Bera test Tranquil Period

[43] Verbeek (2008) p.195

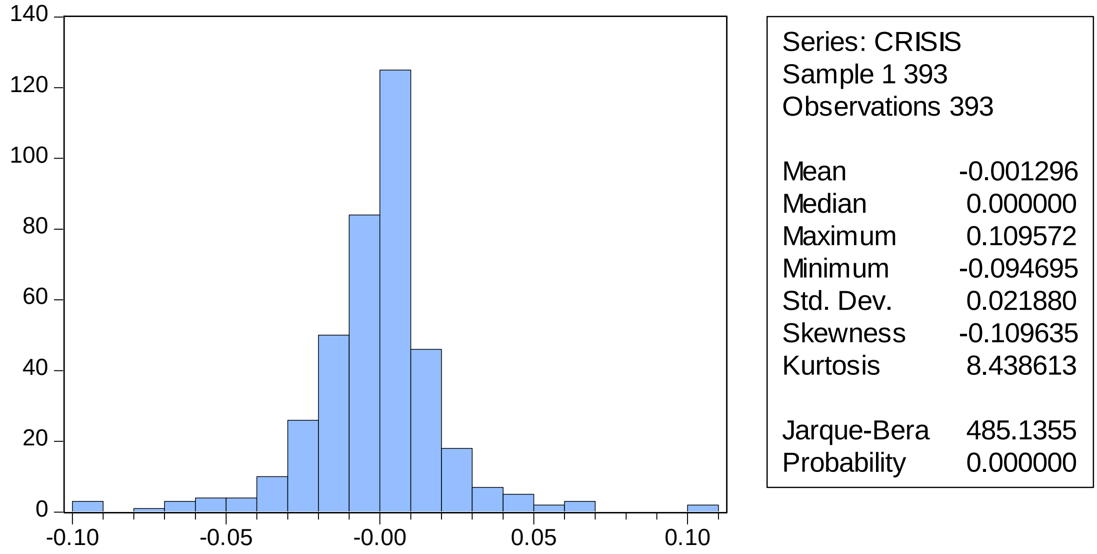


Figure 7: Jarque-Bera test Crisis Period

3.4 Rolling window forecast

We choose to work with 1-day VaR forecasts; this is because horizons longer than 1-day VaR will reduce the number of independent observations and therefore the power of the test. For instance, using a 1-week VaR horizon means that we have only 52 independent observations per year for a 1-day VaR horizon on the other hand we have 252 observations. For this reason a shorter horizon is to prefer to increase the power of the tests. This explains why the Basel Committee performs backtesting over a 1-day horizon, even though the capital requirement is based on a 10-day VaR horizon. This is of importance when it comes to the choice of confidence interval as well because a too high confidence level reduces the number of observations in the tail and hence the power of the test.[44]

The rolling window consists of 1000 observations that start from February 27th 2002. The choice of window size is based on a trade-off between advantages and disadvantages. It cannot be too long and not too short either. We want a sample size that can give us risk estimates of decent precision without the negative effects like the ghost effect. Ghost effects appear when old observations fall out of the sample and create “jumps” in the data which distort the VaR estimate. Experts believe that a couple of year’s daily observations are

[44] Jorion (2000) p.119

needed.[45] On the basis of this we think 1000 observations will be sufficient for our purpose. Especially considering the application of GARCH where the number of observations must be large in order to get reliable parameter estimates which is of importance because GARCH models tend to be unstable when forecasting out of the period.[46]

In the evaluation process we compare the forecast with the outcome and count the number of breaks/exceedances against the level of VaR. We use the Kupiec's test to determine whether the estimation model of VaR is good or not. We are rolling the forecast window to incorporate the tranquil period and the crisis period. By doing so we will determine if the forecasted VaR is destabilizing due to the financial crisis. Hence we will test the null hypothesis, that the model is significant, for each period by calculate the number of exceedances.

$$H_0: VaR\ model\ is\ significant$$

$$H_1: VaR\ model\ is\ not\ significant$$

3.5 Reliability

The aim of the thesis is to provide a high reliability and minimize that the result are given by chance. Our hope is therefore that other researchers, given the same study, data, approaches etc., should come to the same results as us. In order to provide a high reliability we have used statistic gathered from Datastream that we think is a reliable source. It is of highest priority for us to avoid model errors; we therefore have chosen to work with observations at a daily basis and a big amount of observations to minimize the risk for flawed data, which could guide us to misleading estimates. We also believe that S&P 500 is a good proxy for the market risk and therefore applying VaR on it relevant. We are only using one method to secure the reliability of our result, if we instead had used additional backtesting methods or perhaps a stress test we might have gotten a more reliable result. The Kupiec's test only count the number of exceedances but does not account for how much each respective exceedances is, but since Kupiec's test is a popular and well established backtesting method we believe it will sufficient to exclusively use this method.

[45] Dowd (2005) p.100
[46] Dows (1998) p.95-96

3.6 Validity

In our study we will provide a clear picture of how we approach our problem, including all the different parts of the study that yields the result and conclusions. Our result will answer if VaR have produced significant risk estimates based on five of the most used approaches, we cannot however generalize our result to include all VaR approaches and conclude that all of the approaches have worked or have not, due to the result of our research. By choosing representatives from both parametric and non-parametric approaches we will hopefully get a hint of how VaR as a risk measure have worked for a broad coverage of approaches, but again, we cannot deduct that some approaches may behaved differently and deviated from what our study shows. What we with certainty will answer and what is also the aim of this study is if any of our chosen approaches have produced reliable forecasts of the risk during our chosen time horizon. From our result we will not be able to draw conclusions of how these approaches with certainty will perform in the future, we will only have an indication. The reason for this is because we can never rely on that the past will reflect the future. But since our study covers both unstable and tranquil market climate and have sufficient amount of observations we believe that our results will give a good indication of how VaR might behave in the future, especially since we use up-to-date data. The validity of studies based on out-to-date data that only includes one type of market climate might be questioned. Finally, in order to provide a high validity, we will compare our results with other researches on the subject.

4.0 Empirical results

This section presents the empirical results obtained for each VaR models. Our result will firstly be presented by a comparison table and then graphically. We will highlight some of the most important findings and give a short comment on these; the rest of the discussion of our result is left for the next chapter.

4.1 Tranquil period

We display the results from the Kupiec's test for the tranquil period in the table below. For the 99% confidence all VaR approaches performed the same according to the Kupiec's test because they all had the same amount of exceedances, however the log-normal VaR had the most severe exceedances since its magnitude in average percent were the highest. For the VaR at 95% confidence level the result is a bit different, three of the approaches are significant while two are not. This is because they have too few exceedances in contrast to the given significance level, which the LR statistic formula punishes. Hence the overestimation of risk is too high. It is intuitive that the VaR with 95% confidence level has more exceedances because more observations will be incorporated in the left tail.

Tranquil period

99%	Average VaR	Exceedances	Min	Average exceedance	Max	LR	p-value
Basic HS VaR	-2,055%	3	-0,212%	-0,892%	-1,974%	0,242	62,272%
Volatility-weighted HS VaR	-2,152%	3	-0,432%	-0,851%	-1,941%	0,242	62,272%
Normal VaR	-1,916%	3	-0,287%	-0,934%	-2,008%	0,242	62,272%
Log-normal VaR	-1,898%	3	-0,298%	-0,946%	-2,019%	0,242	62,272%
Student's t-VaR	-2,214%	3	-0,054%	-0,695%	-1,768%	0,242	62,272%
95%	Average VaR	Exceedances	Min	Average exceedance	Max	LR	p-value
Basic HS VaR	-1,319%	14	-0,014%	-0,442%	-2,457%	1,893	16,892%
Volatility-weighted HS VaR	-1,337%	11	-0,111%	-0,527%	-2,412%	4,735	2,956%
Normal VaR	-1,344%	14	-0,030%	-0,424%	-2,468%	1,893	16,892%
Log-normal VaR	-1,335%	14	-0,036%	-0,431%	-2,474%	1,893	16,892%
Student's t-VaR	-1,519%	8	-0,016%	-0,545%	-2,327%	9,282	0,231%

Table 2: Backtesting Tranquil period

4.2 Crisis period

During the crisis period the results are stunning, none of the approaches are significant, in fact all of the approaches clearly failed to account for the increasing risk and therefore all are rejected by Kupiec's test. On the 99% confidence level the non-parametric approaches performed the "best" according to the LR-statistic where the Basic Historical Simulation VaR yielded the least average exceedances, i.e. the magnitude and therefore the severity of its underestimates were the least. At the 95% confidence interval the Student's t-distributed VaR had the least exceedances as well as the least average value of exceedances.

Crisis period

99%	Average VaR	Exceedances	Min	Average exceedance	Max	LR	p-value
Basic HS VaR	-2,576%	38	-0,012%	-1,487%	-6,220%	107,374	0,000%
Volatility-weighted HS VaR	-2,581%	38	-0,028%	-1,509%	-6,478%	107,374	0,000%
Normal VaR	-2,005%	50	-0,092%	-1,637%	-6,990%	167,883	0,000%
Log-normal VaR	-1,984%	50	-1,009%	-1,660%	-7,014%	167,883	0,000%
Student's t-VaR	-2,313%	42	-0,034%	-1,583%	-6,653%	126,711	0,000%

95%	Average VaR	Exceedances	Min	Average exceedance	Max	LR	p-value
Basic HS VaR	-1,392%	81	-0,016%	-1,552%	-7,783%	117,436	0,000%
Volatility-weighted HS VaR	-1,408%	81	-0,028%	-1,539%	-7,831%	117,436	0,000%
Normal VaR	-1,413%	81	-0,019%	-1,505%	-7,638%	117,436	0,000%
Log-normal VaR	-1,403%	81	-0,038%	-1,516%	-7,650%	117,436	0,000%
Student's t-VaR	-1,594%	74	-0,009%	-1,454%	-7,440%	95,873	0,000%

Table 3: Backtesting Crisis period

4.3 Non-parametric approaches

4.3.1 Basic Historical Simulation

Our rolling window consist of 1000 observations therefore VaR for this method is decided by the 51st percentile with 95% confidence level and the 11th with 99% confidence level. Based on the graph depicted below, the Basic Historical Simulation reflects the shift from tranquil period to the crisis period very poor and therefore did some major underestimations during this period. Using equally weighted observations is a major drawback of the Basic Historical Simulation since it has a slowness factor and therefore accommodates shifts in the risk poorly. It is arguable that 1000 observation could be a too large window to incorporate new events on the market and could be a reason to the non-parametric approaches weak performance during times of crisis. But as stated earlier the disadvantages with using a larger window is in general undermined by the advantages which motivate our choice.

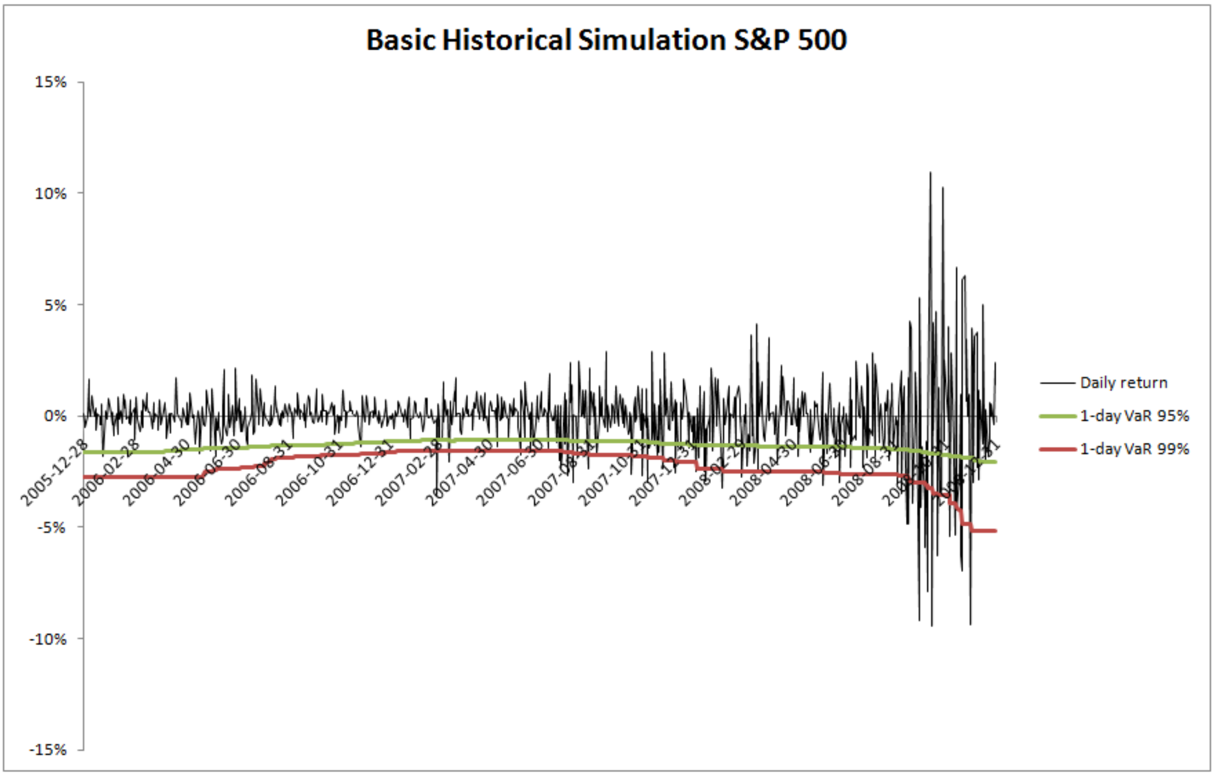


Figure 8: Results Basic Historical Simulation

4.3.2 Volatility-weighted Historical Simulation

In the table below we show the parameters we estimated with the Maximum Likelihood function and are using for applying the volatility-weighted Historical Simulation. From the table we can see that the volatility tomorrow is based on about 90% of the volatility today and roughly 10% depends on the squared shocks. This is consistent with applied theory and proofs the volatility and shocks being sticky and clustered.

We can see in the graph below that during the late 2008 the 99% VaR estimate accommodates changes in the market climate faster than the 95% VaR estimate. This is a common characteristic for both non-parametric approaches.

n=1000, 2002/02/27 - 2005/12/27	
Parameter	**Final values**
μ	0,00038
ω	8,73814E-07
α	0,09264
β	0,90116
Unconditional variance	1,14554E-04
sum ln L	3294,38843

Table 4: GARCH Parameter Values

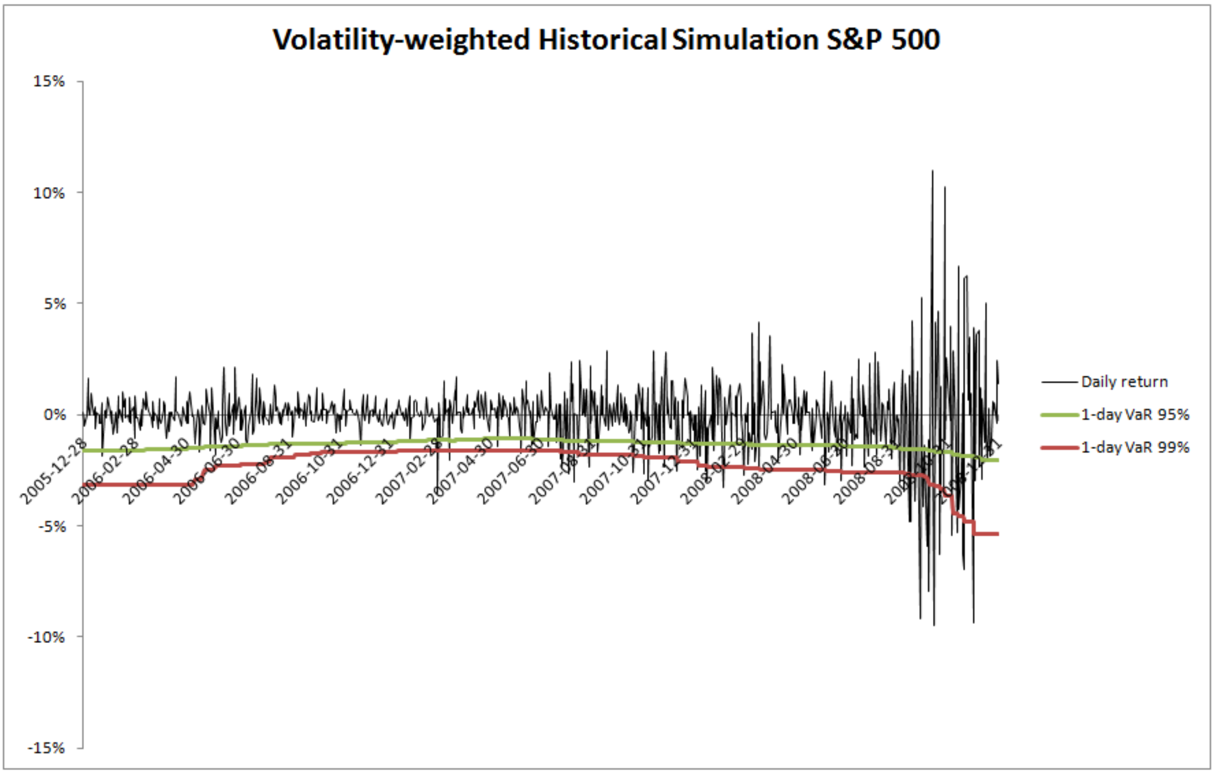


Figure 9: Results volatility-weighted Historical Simulation

4.4 Parametric approaches

4.4.1 Normal distributed Value-at-Risk

The normal distributed VaR is calculated by the formula stated in the theory chapter where we use the average return and sample standard deviation from the 1000 observations incorporated in the window. We can see that the normal distributed VaR does not provide reliable risk estimates especially during the crisis period. Far into the crisis the approach signals higher

VaR estimates but they are still too inaccurate in predicting the risk by looking at the LR statistics. This result however is expected because of the characteristics that the normal distribution possesses, with small tails and much of the distribution centered on the mean. Hence the normal VaR is bad in handling the extreme events that occurred during the financial crisis. Looking at the descriptive statistics of our data series the answer is clear why the normal distribution failed. According to our Jarque-Bera test the null hypothesis of normality is clearly rejected, especially for the crisis period where we observed a very high JB statistic. However the JB statistic were a lot lower during the tranquil period, pointing towards more normally distributed data than the crisis period which explains why the normal distribution performed better during this period. The sample data is rather leptokurtotic distributed with very high kurtosis and fat tails. Financial data is often misfitted by a normal distribution because of this behavior. A leptokurtotic distribution has higher probabilities for extreme values, and as we stated before both normal distribution and the Student's t-distribution is also not consistent with the extreme value theory.

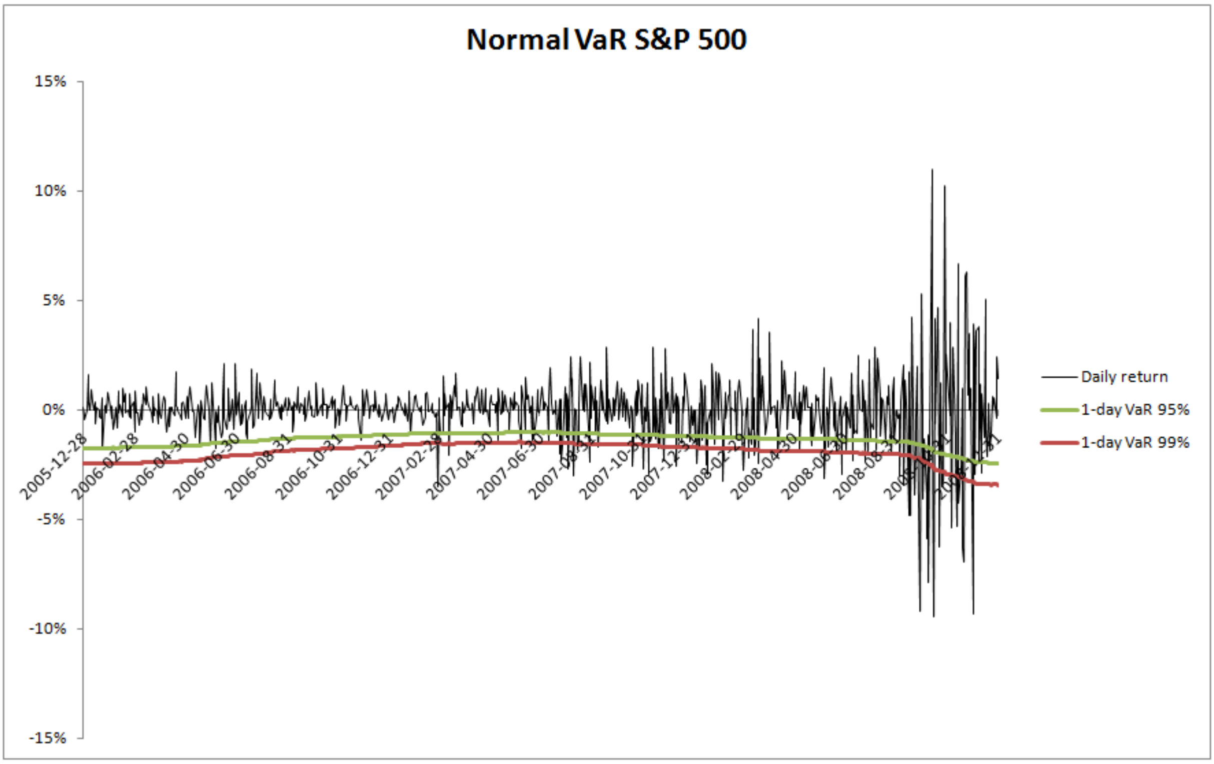


Figure 10: Results Normal distribution

4.4.2 Log-normal distributed Value-at-Risk

The log-normal distribution performed similar to the normal distribution, this behavior however was not that surprising considering that both distributions have similar characteristics and the same inputs needed. However the log-normal distribution slightly estimates VaR lower than the normal distribution.

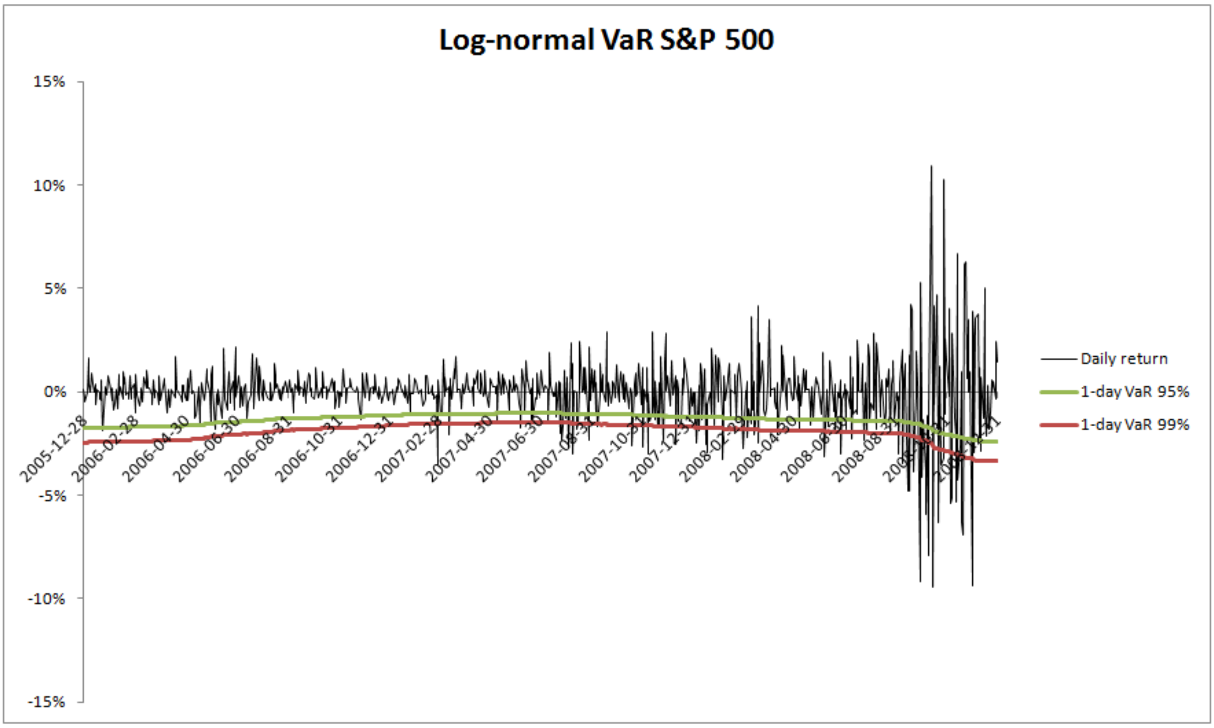


Figure 11: Results Log-normal distribution

4.4.3 Student's t-distributed Value-at-Risk

Besides the same inputs as the other parametric approaches, sample average and standard deviation, the estimation process of Student's t-distribution required the degrees of freedom as well. This was calculated to the number of 5 by the approximation formula defined in the theory chapter above.

The Student's t-distributed VaR for both confidence intervals is slightly higher over the whole period than the normal distributed VaR; this is due to the fact that the Student's t-distribution accommodates fat tails and excess kurtosis. However Student's t-distributed VaR fails as a risk measurement during the financial crisis, since it clearly have a lot of exceedances pointing towards underestimates of the risk.

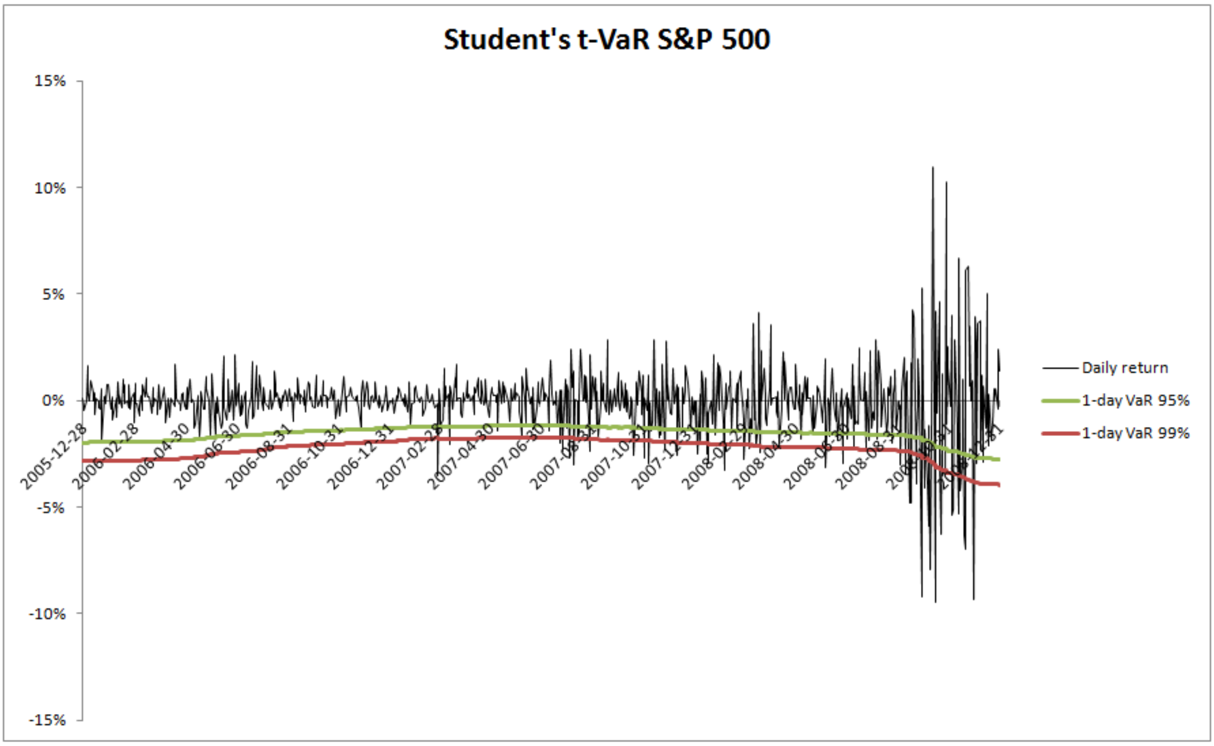


Figure 12: Results Student’s t-distribution

5.0 Analysis

Based on our sample on average the 95% confidence level and the 99% VaR should have had approximately 41 respective 9 amounts of underestimates of risk over the whole period. [47] Clearly none of the approaches were consistent with this which is the most fundamental part of Value-at-Risk, namely based on certain probabilities the losses will only exceed a certain amount a limited number of times. The worrying about these underestimates is not only the amount but also the magnitude, that during the late period of the crisis these underestimates were very severe. A company that rely its risk management solely on a VaR approach during this period therefore would have faced big problems. However it is important to notice that these events during the last part of the crisis is very extreme events and would for any risk measure be hard to deal with. It is times where the volatility is sky high and risk measures as a group tend to perform badly. The evidence for this is the number of defaults during this period that proves that many companies failed in their risk management, especially worrying is that these were large financial institutions that should be the pioneers in risk management. Nevertheless it is under these times we need protection at most.

A finding among our approaches is that they reacted too slowly to the changes in the market climate, they could not incorporate the shift between tranquil and crisis period. The reason for this is very intuitively however, since the non-parametric approaches base their forecast on historical data there is a lag-effect until the new risk is reflected on the forecasts. The reason that the parametric approaches do not account for this is because they are too standardized, they assume a distribution with certain characteristics to last over the whole period. The only change in the risk is if the mean or standard deviation changes but nevertheless a parametric distribution assumes a fairly consistent climate.

Above we discussed the underestimates of risk that occurred during our crisis period, what are also interesting to discuss is the overestimates that occurred in the tranquil period that made some approaches not even significant. The explanation for this is because the models had too few exceedances, than what is expected, for example 1-day VaR on a 95% confidence interval expects 5 days out of 100 to commit underestimations of the risk. Kupiec's test accommodates for this and that is why the t-VaR and volatility-weighted HS were not significant. Overestimates are mainly a problem since it is costly and inefficient for financial institutions to hold too much cash to cover potential losses that is unreasonable. One would

[47] Calculations are based on the following formula: $N(1-\alpha)+1$

also believe that since the models overestimated the risk so much during the tranquil period that they would not commit as many underestimations of the risk in the crisis period. Overestimates of risk are far from as severe as underestimates, but nevertheless it is interesting in our evaluation to take into account all sorts of risk prediction failures in order to determine how accurate our models are.

We mentioned earlier that the BIS (Bank for International Settlements) suggested VaR as an appropriate risk measure for banks in order to determine the capital requirements; thereby much of the financial stability relies on VaR. Clearly according to our study, VaR have not provided trustworthy forecasts of the risk, therefore capital requirements based on it have been misleading. In fact all our VaR models would be in the red zone on the Table 1: Basel accord penalty zones, meaning that banks using these models under our time horizon would have been subjects to severe punishments such as a temporary or permanent suspension of trading activities.

In order to provide a high validity to our thesis a comparison between our findings with other studies is important. Beder's (1995) result agrees with ours concerning that VaR depends much on the input parameters, data and assumptions. The accuracy of VaR relies heavily on if these assumptions fit the data, therefore VaR can change considerably under other assumptions. Her finding, as ours, could not conclude that there is any VaR approach that is superior to another. Also Hendricks (1996) came to this conclusion, which he explained by the nature of the market movements, where the size of market movements is not steady over time, makes it hard to find a model that fits these sporadic events. Therefore one model that may be good under some circumstances may not be good under changed circumstances meaning that there is no universal VaR approach. Our study agrees with Hendricks (1996) that there is no VaR method that works all the time because of the behavior and the randomness of the market. Our study disagrees with Lin et al. (2006) that showed that the Student's t-distribution gives superior VaR forecasts. Lin et al. (2006) also showed that the t-distributed VaR is significantly better than the normal distributed VaR at significance levels exceeding 98.5%, according to our study both models performed similar at the 99% level; we cannot therefore conclude that one of them is preferred to another. However the t-distributed VaR had lower amount of exceedances pointing towards that the t-distribution has a lower probability of committing underestimations of the risk, which is in line with her study but with that comes also the threat of overestimating the risk. Therefore we cannot conclude that the t-distributed VaR is better than the normal distributed VaR. One possible reason to why our result disagrees with Lin et al. (2006) is because their study is done on a different time horizon and data, and as stated before parametric approaches are very sensitive and behave differently on dissimilar data.

6.0 Inference

This chapter will firstly present our main findings in the conclusion and lastly the thesis is ended by some suggestions concerning future studies.

6.1 Conclusion

This thesis examined the one day predictive power of five of the most common VaR models on the 99% and 95% confidence interval with the purpose to answer if VaR as a risk measure have worked during the financial crisis and which approach that generates the best estimates. We applied our models on S&P 500 with 786 observations between 2005-12-28 and 2008-12-31, divided into an equally sized tranquil and crisis period. We evaluated the models by using the backtesting framework set up by Kupiec's test. Based on our chosen methods we conclude that VaR have not worked during the financial crisis due to the high LR statistic for each approach along with insignificancy stated by the Kupiec's test for both confidence levels. This means that VaR is proven to be serious destabilizing during times of crisis. During the tranquil period all models were significant at the 99% level, on the 95% level however the volatility-weighted HS along with the student t-distribution had overestimates of the risk which made it insignificant. Our result does not show that any VaR methodology is superior towards the others.

Our study confirms much of the criticism raised towards the risk measure by other researchers. The risk measures inaccuracy has produced dangerously misleading risk estimates. This can be explained by VaR's sensitivity to incorrect assumptions about the underlying distribution. The parametric approaches are too standardized to fit the financial data and the non-parametric approaches did not react fast enough to accommodate the changes in the market climate. VaR being silenced about the size or magnitude of losses exceeding VaR have also proven to be hazardous which can be seen by the major underestimates during the late financial crisis of all approaches depicted in the graphs under the empirical results. VaR's advantages, its simplicity, its probabilistic and holistic features to mention some, are clearly undermined by its drawbacks therefore we conclude that VaR as a risk measure have failed during our chosen research period.

6.2 Further studies

Our study has been conducted on S&P 500 with five different approaches to evaluate VaR performance during the time between 2005-12-28 and 2008-12-31. It would be interesting to see if the approach of this thesis applied on a different time period with different data would yield the same result. Also studying VaR on other indices or asset classes would be of interest. Considering the poor performance of parametric approaches according to our study, perhaps they will behave differently for other asset classes or data that have different characteristics.

We limited our thesis to cover different aspects of the traditional VaR measure. However we encourage research on some of the closely related risk measures that are extensions of VaR such as CVaR or CAViaR. As demonstrated, forecasting market risk under different market cycles is a demanding task, both CVaR and CAViaR is promising crisis tools and therefore earns further academic interest. A comparison study between VaR and these models could perhaps be interesting.

7.0 References

7.1 Published sources

Bao, Y. and Lee, T. and Saltoğlu, B. (2006), "Evaluating predictive performance of value-at-risk models in emerging markets: a reality check", *Journal of Forecast,* vol.25, pp.101–128.

Basel Committee on Banking Supervision (2005), *International convergence of capital measurement and capital standards - A revised framework.*

Beder, T. (1995), "VaR: Seductive but Dangerous", *Financial Analysts Journal,* vol.51, pp.12–24.

Berkowitz, J. and O'Brien, J. (2002), "How Accurate Are Value-at-Risk Models at Commercial Banks?", *The Journal of Finance,* vol.62, no.3.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics,* vol.31, pp.307-327.

Campbell, J., Lo, A. and MacKinlay, C. (1997), *The Econometrics of Financial Markets*, Princeton University Press, 2nd edition.

Christoffersen, P.E. (1998), "Evaluating interval forecasts", *International Economic Review*, vol.39, pp.841-862.

Dowd, K. (1998), *Beyond Value at Risk: The new science of risk management,* John Wiley & Sons Ltd.

Dowd, K. (2005), *Measuring Market Risk,* 2nd edition, John Wiley & Sons Ltd.

Einhorn, D. and Brown, A. (2008), "*Private Profits and Socialized Risk*", GARP Risk Review (June/July).

Engle, R. (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation", *Econometrica,* vol.50, pp.987–1008.

Hendricks, D. (1996), "Evaluation of Value-at-Risk Models Using Historical Data", *Economic Policy Review*, vol.2, no.1, April 1996.

Hull, J. C. (2008), *Options, futures and other derivatives,* Prentice Hall, 7th edition.

Jorion, P. (2000), *Value at Risk: the New Benchmark for Managing Financial Risk*, 2nd edition, McGraw-Hill.

Kuester, K. (2006) "Value-at-Risk Prediction: A Comparison of Alternative Strategies Journal of Financial Econometrics", vol.4, pp.53–89.

Kupiec, P. (1995), "Techniques for verifying the accuracy of risk management models", *Journal of Derivatives*, pp.73-84.

Lin, C. and Shen, S. (2006), "Can the student-*t* distribution provide accurate value at risk?", *The Journal of Risk Finance,* vol.7, no.3, pp.292-300.

Linsmeier, T.J. and Pearson, N.D. (2001), "*Risk Measurement: An introduction to Value at Risk*" Econpapers Working paper, pp.2-3.

McAleer, M. and Jiménez-Martín, J.A. (2000), "*The Ten Commandments for Managing Value-at-Risk Under the Basel II Accord*".

Moody's Global Credit Policy (2009), *Corporate Defaults and Recovery Rates, 1920-2008,* Feb 2009.

Verbeek, M. (2008), *A Guide to Modern Econometrics,* 3rd edition, John Wiley & Sons Ltd.

7.2 Electronic sources

Answers 16/5-10

http://www.answers.com/topic/normal-distribution

Datastream 10/4-10

Reuters 5/5-10

Oxford Statistics 16/4-10

http://content.answers.com/main/content/img/oxford/Oxford_Statistics/0199541454.normal-distribution.7.jpg

Standard & Poor’s 19/4-10

http://www.standardandpoors.com/indices/sp-500/en/us/?indexId=spusa-500-usduf--p-us-l--

Wikipedia 16/5-10

http://en.wikipedia.org/wiki/Student's_t-distribution

8.0 Appendix

8.1 Jarque-Bera test

The Jarque-Bera test is a statistical test for normality which depends on the sample skewness and kurtosis. It assumes the JB statistic to be χ^2-distributed with two degrees of freedom. By definition a normal distribution has expected skewness of zero and excess kurtosis of zero. The null hypothesis tested is if the data is normally distributed i.e. the coefficients of excess kurtosis and skewness are jointly zero. Any deviation from this will be punished by the formula and increase the JB statistic.[48]

$$\xi_{LM} = N\left[\frac{1}{6}\left(\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i^3 / \sigma^3\right)^2 + \frac{1}{24}\left(\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i^4 / \sigma^4 - 3\right)^2\right]$$

8.2 Dickey-Fuller test

Dickey-Fuller test is a statistical test which tests whether a unit root is present in a time-series data. The null hypothesis of the test is if the variable has unit root. Unit root is the same as non-stationary where a stationary variable has a mean and a variance that is constant over time and that the covariance between two time periods only depends on the time step and not on the actual time for the observation. The assumptions are defined mathematically below.[49]

$$E(Y_t) = \mu$$

$$E(Y_t - \mu)^2 = \sigma^2$$

$$E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k$$

[48] Verbeek (2008)
[49] Verbeek (2008)

Null Hypothesis: SERIES01 has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic based on SIC, MAXLAG=24)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-34.92318	0.0000
Test critical values:	1% level	-3.433813	
	5% level	-2.862956	
	10% level	-2.567571	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(SERIES01)
Method: Least Squares
Date: 05/11/10 Time: 11:22
Sample (adjusted): 3 1786
Included observations: 1784 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
SERIES01(-1)	-1.230558	0.035236	-34.92318	0.0000
D(SERIES01(-1))	0.104049	0.023600	4.408911	0.0000
C	-0.000144	0.000313	-0.458786	0.6464

R-squared	0.561873	Mean dependent var	9.48E-06
Adjusted R-squared	0.561381	S.D. dependent var	0.019974
S.E. of regression	0.013228	Akaike info criterion	-5.811219
Sum squared resid	0.311659	Schwarz criterion	-5.801992
Log likelihood	5186.607	Hannan-Quinn criter.	-5.807811
F-statistic	1142.015	Durbin-Watson stat	1.986708
Prob(F-statistic)	0.000000		

Table 5: Dickey-Fuller test