LUND UNIVERSITY
School of Economics and Management

Evaluation of Value-at-Risk Models During Volatility Clustering

An Empirical Study on the Financial Crisis of 2008

Master Programme in Finance
Lund University
Spring 2014

Author:
Medjit Yalmaz

Supervisor:
Ph.D. Hans Byström

Abstract

In the light of the financial crisis of 2008, risk management has become one of the most important topics in the financial world. This study applies five different VaR approaches, normal distribution, student's t distribution, historical simulation, age weighted historical simulation and volatility weighted historical simulation under three different sample windows. These parametric, non-parametric and semi-parametric approaches were applied on the historical closing prices of five highly diversified stock indices, OMXS 30, S&P 500, NIKKEI 225, DAX 30 and FTSE 100, where the focus has been on the period of 2007-2012. Performance was evaluated by comparing the expected number of VaR breaks to the actual number of VaR breaks, the so called VaR ratio. The study found that most of the models using a larger sample window failed to cope with sudden changes in volatility, while the age weighted historical simulation seemed to cope well with sudden changes in market conditions in all sample windows. The study also found that forecasting volatility using EWMA in extreme market conditions failed to give accurate VaR estimates.

Acknowledgements

I would like to take this opportunity to thank my supervisor Mr Ph.D. Hans Byström for his valuable guidance and comments throughout this process. I would also like to thank programme coordinator Mr Ph.D. Hossein Asgharian and all other teachers I have had the pleasure to have during my time at Lund University.

Table of Contents

1. Introduction

1.1 Background

The 2008 financial crisis severely impacted financial markets and most importantly economies worldwide. Crashing stock prices with large corporations filing for bankruptcy on a daily basis, and government bailout programmes left the economies crawling on its knees. While the key drivers of the crisis were deep rooted and complex, there is a consensus among scholars and governments that failures in financial risk management was in fact one of the reasons behind the magnitude of the crisis (Sollis, 2009). Goldman Sachs CEO, Lloyd Blankfein, stated, while testifying in front of the Financial Crisis Inquiry Commission (FCIC), that: *"risk models, particularly those predicated on historical data were too often allowed to substitute for judgement"* (CSPAN, 2010). In the aftermath of the 2008 financial crisis, many studies have shown that Value-at-Risk (VaR) models couldn't predict nor forecast the magnitude of the financial collapse. Current market conditions, with inflated stock prices, high volatility and risk appetite, are showing similarities with the pre 2008 crisis conditions. The question is whether or not the financial world has learned from its mistakes or if it will hit again?

Risk modelling and measurement is a unified leg in the operations of a financial institution and an important part of financial regulations. The last decades include several major and minor crises and regulatory changes are constantly appearing to address a better understanding of risk. While many other risk measurements exist, VaR overshadows the rest in usage and compatibility. VaR's practical superiority over its theoretical shortcomings is the main reason for the widely spread usage. Danielsson et al (2013) argues that this assumption is supported in most cases, both theoretically and empirically (Daníelsson et al., 2013).

In its simplest form, VaR is used as an instrument to evaluate the losses arising from a potential decrease in the price of an asset. Simply, VaR quantifies the maximum occurred loss over a given time period and a given probability. In 1996 the Basel Committee incorporated VaR in the Basel I accord, continuing with Basel II and III (Daníelsson et al., 2013). Financial firms are by law forced to engage in risk management stated by the Basel accord. The aim is to advocate financial stability by forcing financial institutions to hold enough capital reserves enabling them to minimize the risk of insolvency and default. Basel rules states that financial institutions must hold 8% of their risk-weighted assets as a capital reserve, also known as the capital ratio (Nilsson, 2013). The recognition of VaR by financial and regulatory commissions in recent times is confirming that the use of VaR as a risk measure is widely spread. The recent Basel III regulation advocates the use of these models (BIS, 2011), and the launch of J.P. Morgan's introduction of the RiskMetrics database for use with third-party VaR softwares explains the growing usage of VaR models by both financial and non-financial firms (Hendricks, 1996).

A vast majority of parametric VaR models relies on the assumption that returns are normally distributed, and theoretically, this isn't something new in financial theory. Both the Black-Scholes option pricing formula and modern portfolio theory are based on the same assumption. Extensive research has shown that returns are seldom normally distributed but rather show signs of kurtosis and/or skewness (Mandelbrot, 1963). This leads to the fact that VaR models consistently miscalculate and in most cases underestimate the probability of high impact events. These VaR models depend excessively on the normal return distribution of the data sample at hand. The underestimation of high impact events was one of the reasons financial institutions were unable to react in a timely manner when the crisis of 2008 struck. Financial institutions generally apply a 99% confidence level when calculating VaR and should anticipate around 2.5[1] VaR breaks (VB) a year if returns are normally distributed. Financial institutions, such as UBS, Credit Suisse and Morgan Stanley, have disclosed that they experienced 50, 24 and 18 VB's respectively during 2008 (Campbell, 2009).

The objective of this study is to investigate different VaR approaches' results during the recent financial crisis i.e. during volatility clustering in five stock indices. The choice of indices is based on three criteria, geographical location,

[1] $1 - (252 * 0{,}99) = 2{,}52$

high market capitalization and stock diversification. Daily closing prices of OMX Stockholm (OMXS 30), Standard & Poors 500 (S&P 500), NIKKEI 225, Deutscher Aktie IndeX (DAX 30) and FTSE 100 are used as the underlying data in this study.

1.2 Defining Value-at-Risk

VaR is the smallest loss ℓ such that the probability of a future portfolio loss L for an investor is larger than the loss ℓ is less than or equal to $1-\alpha$. VaR can be defined with the following mathematical equation:

$$VaR_{\alpha}(L) = min\{\ell : \Pr(L > \ell) \leq 1-\alpha\} \qquad 1.1$$

where α is a confidence level, e.g. 95% or 99%). VaR is, from a probabilistic view, the $(1-\alpha)$ quantile of the return/loss distribution. Typical time periods used when determining VaR is $h = 1$ day or $h = 10$ days. The VaR equation may be rewritten in the following way under the presumption of a continuous loss distribution:

$$\Pr\big(L > VaR_{\alpha}(L)\big) = 1-\alpha. \qquad 1.2$$

This definition of VaR says that the probability of a loss L being larger than VaR is equal to $1-\alpha$.

1.3 Purpose of the thesis

The main objective of the thesis is to investigate the results of different parametric, non-parametric and semi-parametric VaR approaches during times of volatility clustering[2] (Dowd, 2002). The time period investigated will date from beginning of 1994 until end of 2013 and will therefore include the dot-com bubble of 2000-2002 and the latest financial crisis of 2008, while the thesis will focus on the latter. There are no intentions to create a new approach to VaR, but rather to evaluate the behaviour of VaR models during a financial crisis. This will give a deep understanding of how financial institutions assess and how fast they can point out times of volatility clustering by using VaR as a risk measurement. Regulators are pushing for an increased level of control and extensive requirements on risk measurements with Basel III, commencing in 2018 as a

[2] Volatility clustering occurs when financial returns show alternating moments/periods of high and low volatility i.e. if volatility is high (low) one day, it is expected to be high (low) the next day as well.

direct response to the recent financial crisis. It will be out of utter importance to evaluate the accuracy and shortcomings of VaR models, in order to judge whether or not VaR is capable of minimizing the risk of capital losses, and if the governmental support is justified. With this stated, the purpose of the thesis is therefore to evaluate the accuracy of a group of parametric, non-parametric and semi-parametric VaR models over the time period 2007-2012. Different assumptions will be applied, including distributions and forecasting models, during the financial crisis of 2008 using historical closing prices of 5 different indices.

1.4 Delimitation

Due to the limited time of writing the thesis certain delimitations are necessary. Including more than the chosen five stock indices would probably not increase the accuracy of the thesis materially since the chosen indices are well diversified, both in an industrial and geographical point of view. Though, other asset classes like foreign exchange (FX), rates, commodities, and fixed income (FI) has been excluded. This was done since most of the mentioned asset classes are traded over-the-counter (OTC) and each asset class has a numerous amount of different products. E.g. FX products include spot, forward/futures and swaps, where each product have a large amount of currency pairs to choose from. Therefore, due to the magnitude of different combinations of product types and the limited time of writing the thesis, these have been excluded.

There are also some delimitations to which approaches have been selected to be analysed and tested. The thesis only focuses on VaR since it is still the most widely used model to estimate risk of a portfolio or an asset. Expected Shortfall (ES) is a slightly more sophisticated way of estimating risk, but given the lack of recognition this method has had at financial institutions, the objective is to evaluate VaR models.

It is also important to stress that throughout the thesis, only 95% VaR has been taken in to consideration. Since the thesis focuses on relative short sample windows, it would be difficult to draw reliable conclusions tied to the purpose of the thesis using a higher confidence level.

2. Literature Review

In this section of the thesis the advantages and drawbacks of VaR will be presented as a background. Furthermore the different parametric, non-parametric and semi-parametric approaches will be presented before the backtesting methods are discussed.

2.1 Advantages of Value-at-Risk

There are two main advantages of using VaR as a risk measure. The first is that it gives a consistent risk measure across different asset classes and portfolios. In essence, it can measure and compare the risk of a fixed income portfolio with an equity portfolio. VaR gives a common view of the risk since it is measured in monetary value. Secondly, VaR estimates all types of risks and takes the correlation between the risk factors in to account. E.g. if there are two positions that by themselves are risky, but when combined in a portfolio, VaR could estimate a lower risk if they correlate negatively, and vice versa if the risk is highly correlated (Dowd, 2002).

Dowd (2002) points out several practical ways of using VaR. (1) VaR can be used to set overall risk objectives and maintaining the risk appetite. (2) Since VaR gives an absolute figure on how risky a portfolio is, this figure can be used to determine capital allocation. (3) In the last 20 years, measuring and reporting VaR has become an important part for financial institutions in order to maintain and disclose their market risk e.g. in annual and quarterly reports. (4) Investment decisions can be made on the base of how VaR will change when pursuing an investment opportunity or when implementing hedging strategies. (5) It is also a way of managing the risk taking on trading books and is used to supervise traders. This is especially important in the light of Basel III in order for financial institutions to be within their limits and reducing risks.

2.2 Drawbacks of Value-at-Risk

There are several drawbacks of the VaR model, one of them points out that estimates may be too far from the reality and causing imprecise risk, in essence making the risk estimates useless. Another worrying fact is that different VaR approaches give significantly different results, which will be tested later in the thesis (Beder, 1995). These drawbacks clearly show the risk in using VaR as a risk measure, if VaR is inaccurate, and decisions are based fully on the base of VaR, investors may take on far more risk than what was originally expected (Hoppe, 1998).

Another problem that is being stressed by Ju and Pearson (1999) is that if VaR is used to manage and supervise risk taking by traders. Traders will eventually be incentivised to seek positions where risk is over- or understated. Ju and Pearson (1999) show in their empirical results that the magnitude of VaR underestimations that rises from this behaviour is substantial.

Taleb (1997) stresses that the widely spread use of VaR could impact financial markets. Since financial institutions constantly revises their positions and hedges due to changes in market prices, all players in a market might have the same behaviour, since they in the end rely on the same information. The result, or risk as you might call it, being that uncorrelated risk in the end becomes correlated, resulting in higher risk than what the VaR models might have suggested in the beginning.

In the end, (1) VaR is silent about the loss of a VaR breach, i.e. how large the actual loss might be. The fact that (2) VaR is widely used, and in point of a financial crisis, everyone in the market will run for the fire exits at the same time. (3) It is not coherent, i.e. it does not always encourage diversification of the portfolio. The largest drawback may however be that (4) VaR is sensitive to incorrect assumptions of the loss distribution and therefore relying too heavily on the underlying data (Nilsson, 2013).

2.3 Parametric approach

The parametric approach to estimate VaR is done by using probability curves and fitting them to the data, consequently deriving the VaR estimate from the probability curve. The main assumption behind this approach is that market volatility, or in other words price changes, follows a probability distribution curve such as the bell-shaped normal distribution or the student's t-distribution curve. The main drawbacks of using this approach are related to the assumption that the market returns are normally distributed; while empirical evidence has shown that this is not the case. Mandelbrot (1963) also stresses the fact that the parametric models disregard the fact that financial returns are not identically and independently distributed (IID). In other words, high returns are usually followed by high returns, and low returns are usually followed by low returns, a phenomenon called volatility clustering.

2.3.1 Normal Distribution

There has been criticism of VaR since the financial crisis of 2008 where the U.S. housing market and global financial markets collapsed. The criticism is related to the presence of fat tails in financial returns distribution (Olson & Desheng, 2013). As discussed earlier, financial returns are rarely normally distributed leading to VaR estimated under normal distribution could be either under- or overstated (Zhang & Cheng, 2005). This is due to the fact that, under normality, it allows for financial losses larger than the initial capital investment, wiping out more than the initial investment (Dowd, 2002).

Although, the main attraction of using the normal distribution is because there are only two independent factors to consider, the mean, μ, and the standard deviation, σ (Dowd, 2002).

$$VaR_{\alpha} = \mu + \sigma z_{\alpha} \tag{2.1}$$

where z_{α} is the critical value corresponding to the confidence level α.

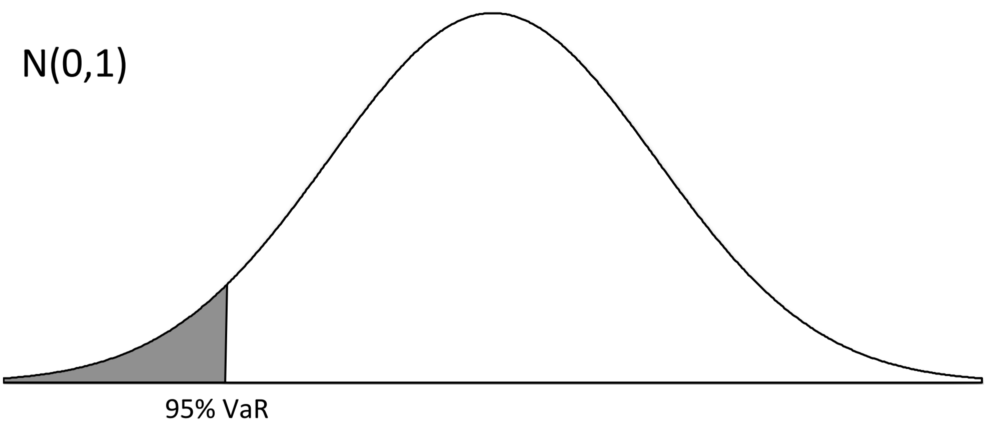

Figure 2.1 VaR under the normal distribution

Under a zero-mean, and standard deviation equal to 1 normal distribution curve VaR is -1,64. In this sample it is the 95th percentile largest loss shown above where the grey area starts. However, it is important to stress that VaR does not estimate how big the loss can be, but rather the smallest loss given the confidence level.

2.3.2 Student's t-distribution

The student's t-distribution is rather simple and is, in contrast to the normal distribution curve, better suited for explaining financial returns because of it's fat tails and positive excess kurtosis. Because of the shape of the student's t-distribution it has been criticised since it does not capture the asymmetrical distribution of financial returns. Although, empirical results have shown that when using higher confidence levels, the student's t-distribution gives better VaR estimates than when derived from the normal distribution (Lin & Shen, 2006).

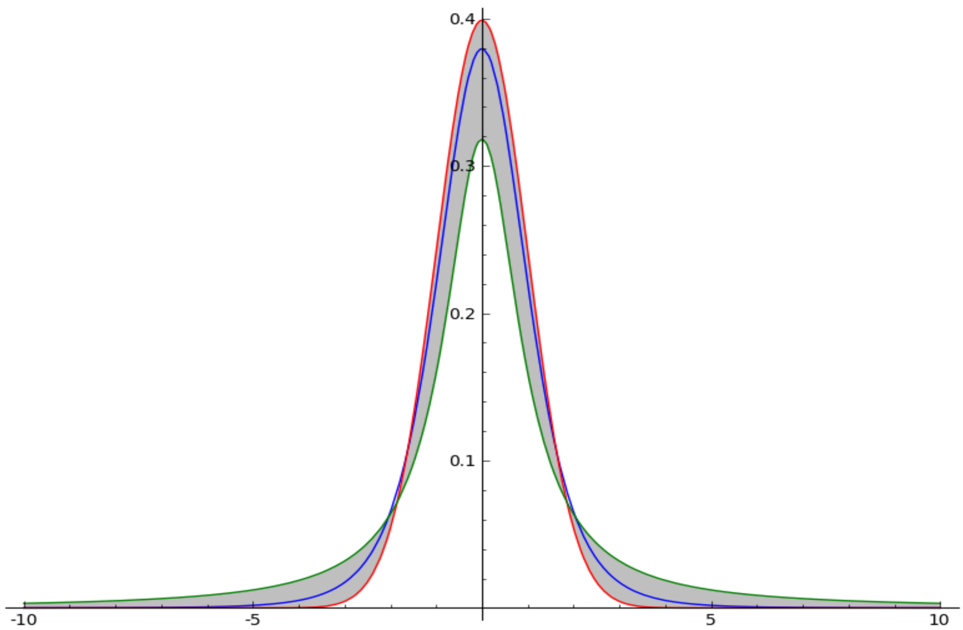

Figure 2.2 Student's t-distribution under different degrees of freedom (Sage, 2012)

In order to estimate VaR using the t-distribution we need to consider the following equation:

$$k = \frac{3(v-2)}{v-4} \Leftrightarrow v = \frac{4k-6}{k-3} \quad 2.2$$

$$VaR_\alpha(L) = \mu + \sqrt{\frac{v-2}{v}}\sigma t_{\alpha,v} \quad 2.3$$

where, v is the degrees of freedom derived from using a maximum likelihood (ML) estimation based on the probability density function. $t_{\alpha,v}$ is the α-quantile for the distribution, while the factor $\sqrt{(v-2)/v}$ is interpreted as a scaling of σ due to excess kurtosis (Dowd, 2002). The degrees of freedom can also be derived by using the Excel function KURT()+3 to get the kurtosis on the sample data. There has been extensive research on the topic on how many degrees of freedom should be used in order to get the most accurate fit of the student's t-distribution to the financial returns. Platen and Sidorowicz (2007) mentions some of the conclusions made by Markowitz and Usmen (1996a, 1996b), Hurst and Platen (1997), and Fergusson and Platen's (2006) ML estimation which showed that the degrees of freedom that gave the best fit was 4,5, 3,0-4,5, and 4,0 respectively. The underlying data in all these research papers were S&P 500 or *"a world stock index, whose constituent weights were determined by market capitalization, and considered different currency denominations of such an index"* (Platen & Sidorowicz, 2007).

2.4 Non-Parametric Approach

The non-parametric approach, in contrast to the parametric approach, does not rely on strong assumptions of the return distribution. The main goal of these approaches is to let the data speak for itself as much as possible, under the assumption that the future will be similar to the past. In other words, the non-parametric distribution relies on the past empirical distribution of returns, rather than using a theoretical distribution curve (Abad et al., 2013).

2.4.1 Historical Simulation

According to Pérignong and Smith's (2010) international survey, Historical Simulation (HS) along with the similar method, Filtered Historical Simulation, are by far the most established methods for estimating VaR at commercial banks.

To forecast the conditional quantiles of financial returns, the HS approach uses unconditional quantiles of financial data (Escanciano & Pei, 2012).

In its simplest form, the HS approach takes the largest loss from the sample window that intersects with the lower 5% of returns and estimates this as VaR on a 95% confidence level. Given a sample of 1000 observations, the VaR will be the 51st largest loss. The reason for taking the 51st and not the 50th largest loss lies in how VaR is defined.

$$VaR_{\alpha}(L) = min\{l: \Pr(L > l) \leq 1 - \alpha\} \qquad 2.4$$

$$\Pr\,(L > VaR_{\alpha}(L) = 1 - \alpha \qquad 2.5$$

In essence, the probability of a loss larger than the VaR is equal to $1 - \alpha$, in this case 95%, therefore, the 51st largest loss is estimated as VaR (Dowd, 2002).

2.4.2 Age Weighted Historical Simulation

A more sophisticated relative to the HS approach is the Age Weighted Historical Simulation (AWHS). One drawback of the HS, which AWHS is trying to overcome, is the fact that all observations in the sample are given the same weight. In other words, an observation will affect the VaR estimate in the same way, no matter if it's the first day or the last day in the sample window. Depending on the sample window, market characteristics may have changed drastically, e.g. during volatility clustering, where an old observation will contribute in a negative way to the VaR estimate. The HS approach depends on the assumption that each observation in a sample period has the same probability to happen again and is independent from other observations over time (IID), which in the end can create ghost effects (Dowd, 2002).

However, the AWHS approach adopted by Boudoukh et al. (1998) overcomes this issue. They took a hybrid approach between the HS and exponential smoothing (EXP) approach, but since the EXP approach is parametrically driven, it assumes normality. As discussed earlier, financial returns are rarely normally distributed, but rather display fat tails, excess kurtosis and unstable correlations, hence the rise of the AWHS approach.

The idea is simple, ranking the largest losses first and assign weights to each of the financial returns, based on their age in the sample window. The next step is to add up the cumulative weights until it breaches the confidence level of choice. The weights are mathematically explained below (Boudoukh et al., 1998):

$$\omega_1 = \frac{(1-\lambda)}{(1-\lambda^N)}$$

$$\omega_2 = \lambda\omega_1$$

$$\omega_3 = \lambda\omega_2 = \lambda^2\omega_1$$

$$\vdots$$

$$\omega_N = \lambda^{N-1}\omega_1$$

2.6

where ω_1 is the probability of the latest observation and ω_N is the probability of the oldest observation in the sample window. λ is the irrelevance factor, i.e. how fast the observations will decrease in probability. A λ close to 0 will make older observations irrelevant quicker than a λ closer to 1 (Boudoukh et al., 1998). In order to adapt the model for volatility clustering, a reasonable λ will be used when conducting the analysis.

2.5 Semi-Parametric Approach

As one could imagine, the semi-parametric approaches combines the parametric approach with the non-parametric approach. The most established semi-parametric methods are volatility weighted historical simulation, CaViaR and the extreme value theory method (Abad et al., 2013), while this thesis will focus on volatility weighted historical simulation.

2.5.1 Volatility Weighted Historical Simulation

In addition to the AWHS approach, there are also other ways of weighting sample data. One such approach is the Volatility Weighted Historical Simulation (VWHS) originally developed by Hull and White (1998). The basic concept of this approach is to weight the observations depending on current market volatility. As an example, if the current market volatility is 2%, while 20 daily observations earlier (i.e. one month ago), the market volatility was 3%. The one month old data overstates the price changes in the market compared to what would be expected in the current state of the market. The same applies to when the volatility was lower but the current state of the market implies a higher volatility, i.e. the older data understates the expected volatility (Hull & White, 1998). The main motivation of using this approach is due to the volatility clustering phenomena mentioned earlier.

The main advantage of using the VWHS instead of the non-parametric approaches HS and AWHS is that the VWHS *"takes account of volatility changes in a natural and direct way".* In contrast, the HS approach completely ignores changes in volatility, while it is not fully incorporated in the AWHS approach. The VWHS approach also enables the VaR estimates to exceed the largest loss in the sample period. This event happens when there are moments of high volatility since the observations in the sample period are scaled upwards (Dowd, 2002). Empirical evidence has also shown that the VWHS approach is producing better VaR estimates than the HS and AWHS approaches (Hull & White, 1998).

The scaling of losses are explained mathematically below:

$$\ell_1^* = \frac{\sigma_{T+1}}{\sigma_1} \ell_1$$

$$\ell_2^* = \frac{\sigma_{T+1}}{\sigma_2} \ell_2$$

$$\vdots$$

$$\ell_T^* = \frac{\sigma_{T+1}}{\sigma_T} \ell_T$$

2.7

where $\sigma_1, \sigma_2, \ldots, \sigma_T$ are the volatilities of each and every observation in the sample period and σ_{T+1} is the forecasted volatility of the next observation, consequently estimated using an exponentially weighted moving average (EWMA) model.

$$\sigma_{T+1}^2 = \frac{1-\lambda}{1-\lambda^T} \sum_{t=1}^{T} \lambda^{T-t} \varepsilon_t^2$$

2.8

λ is a fixed constant equal to 0,94 used by RiskMetrics and the error term, ε is initially set to 0.

2.6 Backtesting VaR

The validity of VaR models is usually measured on their ability to forecast reliable VaR estimates. Since VaR was introduced, several statistical tests have been developed in order to estimate the validity of the VaR estimates, where the two most established ones being the Kupiec's (1995) likelihood ratio test and the independence test of Chrisoffersen (1998).

2.6.1 Kupiec Test

In short, the Kupiec's test is a way of measuring the number of allowed exceptions. The backbone of the test are formal statistical models in order investigate the accuracy of VaR models. Even though Kupiec (1995) found that the main drawback of the test is that it requires large samples to function properly, it is still the most established test and is widely used. The test is formed under Kupiec's null hypothesis where the expected number of violations of the VaR estimate should follow a binominal distribution.

$$H_0: p = \hat{p} = \frac{x}{T} \tag{2.11}$$

Where T is the number of observations in the sample period and x is the number of violations under a certain confidence level. The test's intention is to test if the observed number of violations, $\hat{p}$, is significantly different from the expected number of violations, p. In other words, a VaR break (VB) occurs when a larger loss than the VaR estimate is observed. It has been shown that the hypothesis is best tested under a likelihood ratio (LR) test (Kupiec, 1995).

$$LR = -2\ln\left(\frac{(1-p)^{T-x}p^x}{\left[1-\left(\frac{x}{T}\right)\right]^{T-x}\left(\frac{x}{T}\right)^x}\right) \sim \chi^2(1) \tag{2.12}$$

LR is χ^2 (chi-squared) distributed with one degree of freedom (d.o.f.). In the case where LR is exceeding the critical value of the χ^2-distribution, the null hypothesis will be rejected.

2.6.2 Christoffersen Test

The Kupiec test does not take in to account conditional coverage in order to detect violations of the independence property of an appropriate VaR estimate. Due to this fact a large number of tests have risen, of which the Christoffersen test is one of those. The Christoffersen test is testing whether or not a violation at t is dependent on if there was a violation at t-1 (Campbell, 2005). A conditional coverage test is used to investigate whether or not the underlying model is estimating correct VB frequency and if these VB's are independent from each other. The Christoffersen test also applies a χ^2-distribution using a LR test.

$$LR_{ind} = -2\ln\frac{(1-\pi_2)^{n_{00}+n_{10}}{\pi_2}^{n_{01}+n_{11}}}{(1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}}} \sim \chi^2(1)$$

2.13

where $n_{i,j}$ is the number of days where state i occurred the day before state j;

$$\pi_{01} = \frac{n_{01}}{n_{00}+n_{01}} \qquad \pi_{11} = \frac{n_{11}}{n_{10}+n_{11}} \qquad \pi_2 = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}}$$

However, there are a few shortcomings of this method. The existence of numerous ways where the independence property could be violated is the main drawback. This is due to the shortcoming in the alternative hypothesis where it can't be distinguished in which way the hypothesis is being violated. As for the Kupiec test, the Christoffersen test is being validated using a p-value for the chosen confidence level if the hypothesis can be rejected or not (Campbell, 2005).

3. Methodology

In this section of the thesis the different methodical approaches used will be discussed. The nature of the thesis will be presented firstly, followed by a presentation of each and every index considered in the thesis. After this, a short summary of key statistics will be discussed. Lastly, there will be a presentation of what steps have been taken in order estimate VaR using the different approaches.

3.1 Quantitative Approach

The nature of VaR almost forces the study to take a quantitative approach since the data set used are based on almost 20 years of historical data from five different stock indices. A quantitative approach is a research strategy relying on the quantification in collection and analysis of data. Quantitative researchers <u>measure</u> their result in contrast to qualitative researchers (Bryman & Bell, 2011). These measurements can often be viewed as the truth, but in the case of VaR estimation they should not. It is important to stress that the sample data and sample window may produce biased results, and even if the study is applicable, it may not give true results for other sample data. In order to get to a conclusion, a large amount of calculations have been made on the historical data to estimate VaR. Therefore, for this thesis, a quantitative approach is the best way of conducting a reliable analysis in order to draw relevant and reliable conclusions.

3.2 Deductive Approach

Firstly, the main purpose of this thesis is <u>not</u> to invent a new theory, but solely to test and analyse different established VaR models, therefore the thesis will take a deductive approach. As Bryman and Bell (2011) points out, the process of deductive research starts with the theory, which is then applied to the observations/findings. However, there are some drawbacks when a deductive approach is taken. (1) There could be some new findings or research published by others, before the researcher's findings (thesis) has been published. (2) The data set may become irrelevant for the theory, which is only apparent when the theory is eventually applied. (3) There are no guarantees that the data is suitable for the purpose of the research (Bryman & Bell, 2011). The idea is to apply the theoretical approaches on the empirical evidence and draw conclusions that will either strengthen or weaken the theory.

To justify the thesis taking a deductive approach, even though the approach has its drawbacks, the following comments may be worth taken in to consideration. (1) Due to the time period of which this thesis is conducted, and the amount of research already made on similar topics, there is a limited risk that new findings will make the thesis irrelevant. (2) In order to conduct VaR estimates there is a need of financial data, and this is what the different VaR approaches uses when estimating VaR figures. (3) This is in fact the purpose of the thesis, the data sample contains moments of high volatility, which may be unsuitable for estimating VaR, but in order to make sure which VaR approach gives the best estimates during volatility clustering, this is a necessity. However, this could also lead to inconclusive results if none of the approaches are suitable.

3.3 Reliability and Validity

To ensure that the results and conclusions drawn in this thesis are valid, the historical data that have been used are based on public information and enables reproduction of the conducted tests. There is also extensive research made on the topic of VaR, which ensures the reliability, and validity of the thesis. It is however important to stress that VaR is an estimation of potential future loss and not the absolute truth. The fact that historical data of five stock indices does not tell the entire truth of what has happened in the global financial markets the last decade is also worth considering. The correlation between the stock index and a specific stock is not a perfect match, there is a possibility that a stock price can move in the opposite direction of the market or not even move at all. There are also other asset classes like FX, rates and fixed income etc. that is not considered in this thesis. An investor, or a financial institution that holds large portfolios of different assets may experience different VaR estimates and volatility than for the five chosen indices.

3.4 The Indices

In order to investigate how appropriate the different VaR approaches are, 5 different indices has been chosen, all of which are highly diversified and geographically different. The selection has been made in order to create a robust and reliable analysis of the VaR models. Below there will be a quick introduction to each index and a graph that explains the market conditions from 01.01.1994 to 01.10.2013. For each of the indices, there is significant volatility clustering throughout the observation period. The largest and highest of which happened during the financial crisis of 2008[3], which is also coherent in the logarithmic return distribution showing large positive and negative returns.

3.4.1 OMX Stockholm 30

The OMXS 30 index is a list of the 30 most traded stocks on the Stockholm Stock Exchange and is an index that is priced based on market weights. It was established on the 30th of September 1986 with a base level of 125 and is denoted in Swedish Krona (SEK). The list of stocks on the OMXS 30 is revised twice a year (Bloomberg, 2014).

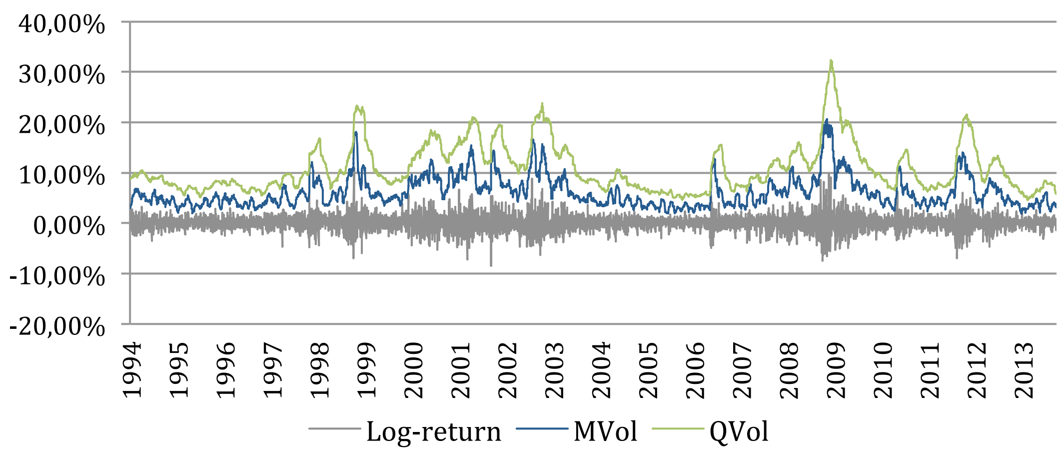

Figure 3.1 The graph above show the Logarithmic return, Monthly volatility and Quarterly volatility (in %) on the OMXS 30 index.

[3] There are many perceptions on when the recent financial crisis actually started. The start date of the recent financial crisis is set to the 15th of September 2008, the day when Lehmann Brothers was filed under bankruptcy.

3.4.2 Standard & Poor's 500

The S&P 500 is one of the most famous stock indices in the world and is a market capitalization weighted index with 500 different stocks. It was established to measure the performance of U.S. economy by looking in to the market value of companies in all the major industries. It was originally established with a base level of 10 for a period 1941-1943 and is denoted in U.S. Dollar (USD) (Bloomberg, 2014).

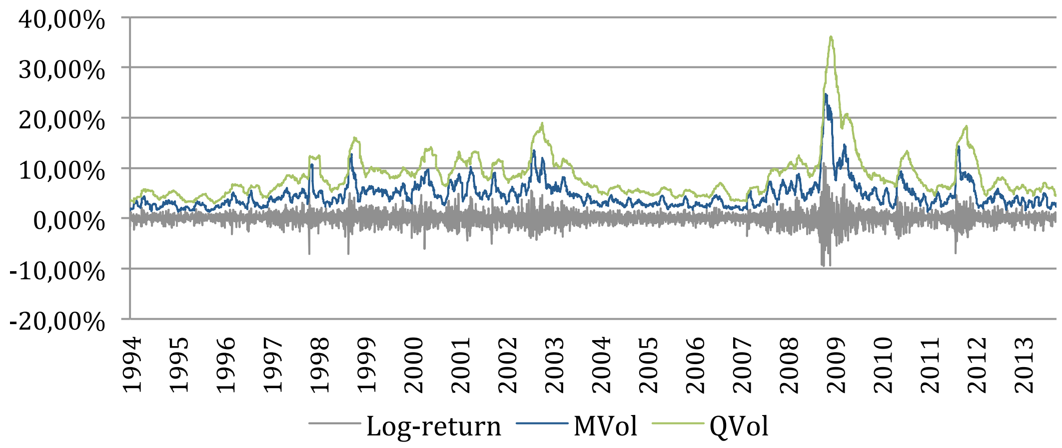

Figure 3.2 The graph above show the Logarithmic return, Monthly volatility and Quarterly volatility (in %) on the S&P 500 index.

3.4.3 NIKKEI 225

225 of the highest rated Japanese companies constitute the NIKKEI 225 index. It is a price-weighted index with companies listed on the First Section of the Tokyo Stock Exchange. NIKKEI 225 was first established on the 16th of May 1949 with a base level of 176,21 Japanese Yen (JPY) (Bloomberg, 2014).

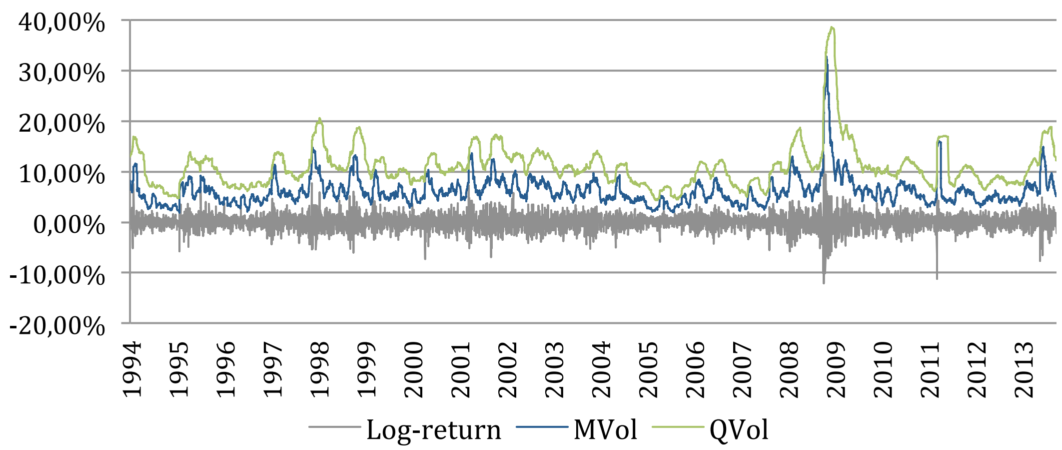

Figure 3.3 The graph above show the Logarithmic return, Monthly volatility and Quarterly volatility (in %) on the NIKKEI 225 index.

3.4.4 Deutscher Aktie IndeX 30

DAX 30 is a total return index consisting of 30 blue chip stocks traded on the Frankfurt Stock Exchange. In order to do the index calculation, DAX 30 is using free float shares. It was established on the 31st of December 1987 with a base value of 1,000 (Bloomberg, 2014).

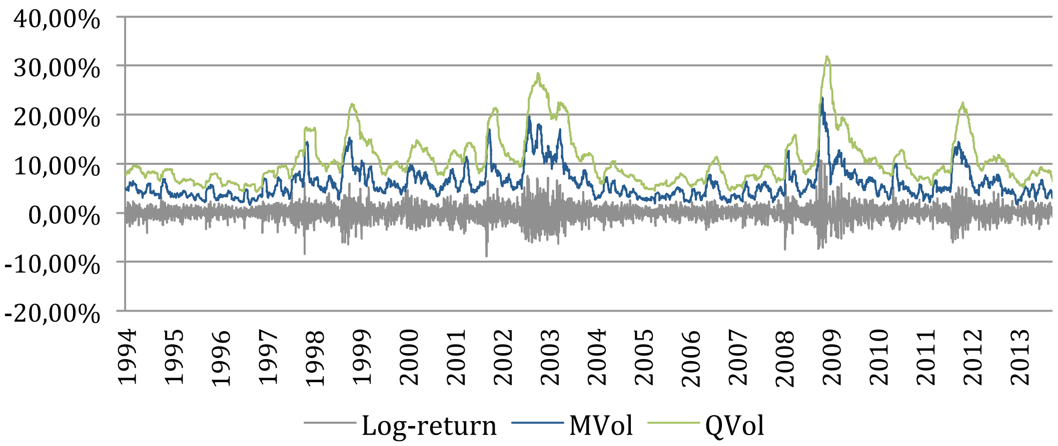

Figure 3.4 The graph above show the Logarithmic return, Monthly volatility and Quarterly volatility (in %) on the DAX 30 index.

3.4.5 FTSE 100

Similar to the S&P 500, the FTSE 100 is also a capitalization weighted index, though, only consisting of 100 of the highest capitalized stocks on the London Stock Exchange. FTSE 100 uses investibility weighting in order to calculate the index level. It was originally established on the 30th of December 1983 with a base value of 1,000 and is denoted in British Pound (GBP) (Bloomberg, 2014).

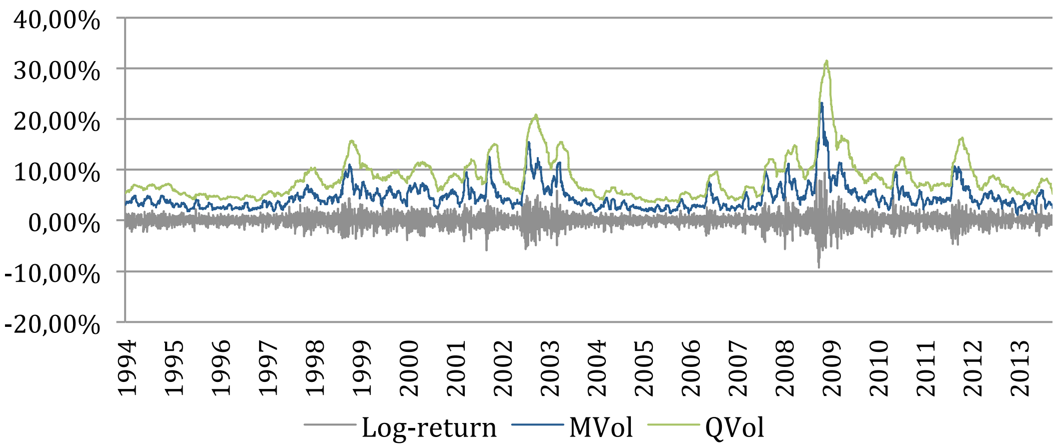

Figure 3.5 The graph above show the Logarithmic return, Monthly volatility and Quarterly volatility (in %) on the FTSE 100 index.

3.5 Data

The data of the five indices described above will be used for the empirical evaluation and has been sourced from Thomson Reuters DataStream. The data contains daily closing prices from the 1st of October 1993 to the 1st of October 2013. It should however be stressed that the time period effectively used will stretch from 1st of October 1994 to 1st of October 2013 due to the computational requirements of the VaR approaches. Effectively, 4,957 VaR observations have been taken in to account when making the empirical evaluation, but due to the nature of this thesis, the focus will lye on the period from 2007 to 2012, during the recent financial crisis. In the event where one or more indices have been closed due to e.g. holidays, the same closing price as the day before has been assigned in order to have consistency between the indices. The logarithmic returns have been calculated by using the formula $\ln(r_t/r_{t-1})$ where r_t is the closing price at time t. It should also be stated that the data does not capture intra-day volatility and this has been disregarded due to the already large magnitude of observations.

3.6 Descriptive Statistics

In order to give a view of the data, some preliminary statistics, such as mean, standard deviation, kurtosis, skewness and Jarque Bera[4] (JB) test, are presented in table 3.1. As discussed earlier in the literature review, empirical evidence has shown that financial data are seldom normally distributed. Bearing this in mind, we expect that none of the return distributions for the indices to be normally distributed. In table 3.1 we see that normality is strongly rejected under the JB-test. As well as in the case of non-normality, the data shows significant excess kurtosis, or in other words, the data exhibits leptokurtic features with fat tails. The most different, and also the most interesting, statistic between the five indices is the skewness. OMXS 30 shows a positive skewness meaning that the return distribution is slightly skewed to the right, while the rest of the indices show smaller negative skewness i.e. the return distributions are skewed to the left.

[4] $JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \sim \chi^2(2)$, where skewness, S = 0 and kurtosis, K = 3 under normality.

	OMXS30	S&P500	NIKKEI225	DAX30	FTSE100
Mean	0,0003	0,0003	0,0000	0,0003	0,0001
Std dev	0,0150	0,0120	0,0149	0,0149	0,0117
max	11,02%	10,96%	13,23%	10,80%	9,38%
min	-8,53%	-9,47%	-12,11%	-8,87%	-9,27%
Kurtosis	6,79	11,54	9,12	7,53	9,12
Skewness	0,08	-0,24	-0,29	-0,12	-0,16
Jarque Bera	3 081,95	15 699,04	8 111,22	4 414,96	8 062,01
Number of observations		5 153			

Table 3.1 The table shows descriptive statistics for the 5 indices over the time period 1994.01.01 to 2013.10.04.

3.7 VaR Estimates

The following section will give a detailed description on how the VaR estimates were produced for each VaR approach. This will enable the reader to reproduce and recreate the same approach taken to conduct a similar study. Three different rolling sample windows has been used, 63 days (one quarter), 252 days (one year) and 1000 days (approx. four years). The data, i.e. daily closing prices, has been converted in to daily logarithmic returns, which has been used in order to produce the VaR estimates.

3.7.1 Normal Distribution

In order to produce the VaR estimates under the normal distribution approach the first step is to assign a volatility, σ, to each and every observation. This was made by using Excel's function STDEV.S(). Depending on the sample window, the previous X[5] observations were included in the formula. The same procedure was made in order to calculate the mean, μ, using AVERAGE() in Excel. Since the thesis only considers a 95% confidence level, the z-value, z_α, under normal distribution is equal to -1,64. The calculations were then repeated throughout the period of 2007 to 2012.

3.7.2 Student's t-Distribution

Under the student's t-distribution the first step was to calculate the degrees of freedom. The degrees of freedom were estimated by using the KURT() formula in Excel on the period from 1994.01.01 to 2013.10.04. Under this presumption, the

[5] Sample window - 63, 252 or 1000 observations

degrees of freedom ranged from 4,7 to 5,6. This means that they are significantly different from what previous researchers have found the most appropriate match of around 4,0 degrees of freedom described in the literature review. The student's t-distribution was also tested under degrees of freedom equal to 4,0 (A short analysis can be found in section 5). The critical t-value, $t_{\alpha,v}$, under 5,0 degrees of freedom is equal to -2,015. The mean and volatility was calculated by using the same method as in the normal distribution approach.

3.7.3 Historical Simulation

The historical simulation approach is a relatively simple approach and is fairly straightforward. Depending on the sample window, the Xth largest loss in the sample window is the VaR estimate. X was found by taking the 5th percentile on the number of observations in the sample window, e.g. if the sample window contained 252 observations, the 13th largest loss ($252 * 5\% = 12,6$) was estimated as VaR.

3.7.4 Age Weighted Historical Simulation

In order to estimate VaR in an efficient matter, and including a rolling sample window, a VBA macro was written. The VBA transcript can be viewed in the Appendix. However, the steps the VBA macro takes when estimating VaR using the AWHS approach on a rolling sample window are the following. (1) Copy the X[6] previous observations and assign the weights depending on λ which in this case was set to 0,99 (Formula 2.6). (2) Sort the losses in an ascending order, i.e. from smallest to largest. (3) The VaR estimate is the first loss where the sum of the sum of the weights is larger than 5%. (4) The VBA code then repeated this by taking the next day's X previous observations until reaching the end of the sample period.

3.7.5 Volatility Weighted Historical Simulation

Similarly as for the AWHS VaR estimates, a VBA macro was written and the transcript can be found in the Appendix. The first step was to forecast the volatility, σ^2_{T+1}, using an EWMA model for each day in the observation period (Formula 2.8). When σ^2_{T+1} is estimated for each and every observation in the

[6] Sample window – 63, 252 or 1000 observations

sample period the next step is to scale the losses according to formula 2.7. The scaled losses are then sorted in an ascending order similarly as in the case of AWHS. The VaR estimate is the X^{th} largest loss in the sample window e.g. the 13^{th} largest loss in a sample window of 252 observations.

4. Empirical Evidence

In this chapter the empirical findings will be presented. In order to avoid confusion, each VaR approach and backtest will be presented separately in order to get a better understanding of the results. When presenting the backtesting results, a green number means that the number of VB's is accepted under the Kupiec test, whereas a red underlined number means that it is rejected. In section 4.6, results from the Christoffersen test of independence will be presented, using the same formatting rule as for the Kupiec test.

4.1 VaR under Normal Distribution

Table 4.1 presents the number of VB's under normal distribution with a rolling sample window of 63 observations. Normal distribution, under this approach, constantly underestimates VaR, with an exception in 2009, which shows satisfactory VaR ratios (VR's[7]). Recalling figures 3.1 to 3.5, volatility was constantly at a low decreasing level in 2009 relative to the sample period of 2007-2012, consequently causing fewer VB's.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	16	6,35%	25	9,92%	19	7,54%	20	7,94%	16	6,35%
2008	17	6,75%	22	8,73%	17	6,75%	21	8,33%	20	7,94%
2009	7	2,78%	12	4,76%	13	5,16%	12	4,76%	13	5,16%
2010	13	5,16%	18	7,14%	14	5,56%	14	5,56%	14	5,56%
2011	20	7,94%	17	6,75%	15	5,95%	22	8,73%	19	7,54%
2012	15	5,95%	15	5,95%	13	5,16%	11	4,37%	16	6,35%
2007-2012	**88**	**5,82%**	**109**	**7,21%**	**91**	**6,02%**	**100**	**6,61%**	**98**	**6,48%**

Table 4.1 VaR breaks and VaR ratio under normal distribution with a sample window of 63 observations.

In contrast to the sample window using 63 observations, the 252 observations sample window clearly overestimates VaR in 2009 and 2012 where volatility was lower than previous periods. This clearly displays VaR's attribute of slow adaptability to changes in market conditions. However, in 2010 the VR show satisfactory figures throughout all indices, possibly due to calm volatility changes from 2009 to 2010. Out of all indices, the one index showing satisfactory VR's throughout the whole sample period, with an exception in 2009, is the NIKKEI 225. In figure 3.3, NIKKEI 225 show relatively high and stable volatility with few

[7] VR = 1 – (252 – n) / 252, where n = no. of VB's in a year

exceptions during the sample period, most probably causing an overall satisfactory VR for the sample period.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	22	8,73%	29	11,51%	19	7,54%	22	8,73%	22	8,73%
2008	27	10,71%	33	13,10%	29	11,51%	27	10,71%	26	10,32%
2009	4	1,59%	4	1,59%	3	1,19%	5	1,98%	3	1,19%
2010	13	5,16%	11	4,37%	12	4,76%	10	3,97%	13	5,16%
2011	25	9,92%	23	9,13%	6	2,38%	30	11,90%	23	9,13%
2012	4	1,59%	4	1,59%	7	2,78%	4	1,59%	4	1,59%
2007-2012	**95**	**6,28%**	**104**	**6,88%**	**76**	**5,03%**	**98**	**6,48%**	**91**	**6,02%**

Table 4.2 VaR breaks and VaR ratio under normal distribution with a sample window of 252 observations.

When applying a rolling sample window of 1000 observations the results become somewhat worrying. In table 4.3 there is evidence that slow adaptability can both over- and underestimate VaR. VR's show large spreads in VB's from year to year, with a clear overestimation in 2010 and 2012 and underestimation in 2007 and 2008. Another worrying fact is the lack of VB's in 2012 for NIKKEI 225 and FTSE 100, and only one VB for S&P 500. The VR, ranging from 5.69% to 7.94%, for the overall sample period show that VaR is overestimated through all indices and not displaying a satisfactory number of VB's.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	22	8,73%	34	13,49%	18	7,14%	19	7,54%	28	11,11%
2008	45	17,86%	56	22,22%	49	19,44%	46	18,25%	55	21,83%
2009	14	5,56%	15	5,95%	11	4,37%	17	6,75%	14	5,56%
2010	4	1,59%	7	2,78%	5	1,98%	5	1,98%	6	2,38%
2011	13	5,16%	7	2,78%	3	1,19%	16	6,35%	8	3,17%
2012	4	1,59%	1	0,40%	0	0,00%	4	1,59%	0	0,00%
2007-2012	**102**	**6,75%**	**120**	**7,94%**	**86**	**5,69%**	**107**	**7,08%**	**111**	**7,34%**

Table 4.3 VaR breaks and VaR ratio under normal distribution with a sample window of 1000 observations.

The normal distribution approach only gives satisfactory VaR estimates during moments of low or constant volatility, expectedly, due to the normal distribution's lack of fat tails and leptokurtic features. The results show that the normal distribution approach under a sample window of 252 observations is superior to the other sample windows when estimating VaR over the whole sample period. However, when inspecting the fact that the sample window of 252 observations has a wide spread of VB's from year to year, the sample window of 63 observations clearly gives the most satisfactory VaR estimates during volatility clustering.

4.1.1. Backtest of VaR under Normal Distribution

In table 4.4 the results from the Kupiec's test of unconditional coverage are presented. The worst fit seems to be the sample window of 1000 observations. It has far more VB's than what is acceptable and the difference in VB's from year to year is widely spread. Consequently, rejecting a normal distribution VaR model using a sample window of 1000 observations, where only the NIKKEI 225 index is showing an acceptable number of VB's over the whole sample period. Even though the total VB's for the 225 observations sample window (464 VB's) are less than for the 63 observations sample window (486 VB's), it seems to be a worse model. The only year where the sample window of 252 observations are showing satisfactory results are 2010, while when using a sample window of 63 observations gives satisfactory results from year to year with minor exceptions.

Index				OMXS 30	S&P 500	NIKKEI 225	DAX 30	FTSE 100
Year	**MIN**	**MAX**	**Target**	**63D**	**63D**	**63D**	**63D**	**63D**
2007	6	20	13	16	25	19	20	16
2008	6	20	13	17	22	17	21	20
2009	6	20	13	7	12	13	12	13
2010	6	20	13	13	18	14	14	14
2011	6	20	13	20	17	15	22	19
2012	6	20	13	15	15	13	11	16
2007-2012	**59**	**93**	**76**	**88**	**109**	**91**	**100**	**98**
Year	**MIN**	**MAX**	**Target**	**252D**	**252D**	**252D**	**252D**	**252D**
2007	6	20	13	22	29	19	22	22
2008	6	20	13	27	33	29	27	26
2009	6	20	13	4	4	3	5	3
2010	6	20	13	13	11	12	10	13
2011	6	20	13	25	23	6	30	23
2012	6	20	13	4	4	7	4	4
2007-2012	**59**	**93**	**76**	**95**	**104**	**76**	**98**	**91**
Year	**MIN**	**MAX**	**Target**	**1000D**	**1000D**	**1000D**	**1000D**	**1000D**
2007	6	20	13	22	34	18	19	28
2008	6	20	13	45	56	49	46	55
2009	6	20	13	14	15	11	17	14
2010	6	20	13	4	7	5	5	6
2011	6	20	13	13	7	3	16	8
2012	6	20	13	4	1	0	4	0
2007-2012	**59**	**93**	**76**	**102**	**120**	**86**	**107**	**111**

Table 4.4 Kupiec's Test for 95% VaR under normal distribution.

4.2 VaR under Student's t-Distribution

It is evident that the student's t-distribution under a sample window of 63 observations does produce better VaR estimates than under the same model using the normal distribution approach, even though the results are not significantly different. The student's t-distribution however manages to produce reasonable VR's throughout the sample period, with a few exceptions, OMXS 30 in 2009 being one of them. Overall, the student's t-distribution under a sample window of 63 observations somewhat underestimates VaR throughout the sample period.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	16	6,35%	25	9,92%	18	7,14%	19	7,54%	16	6,35%
2008	16	6,35%	22	8,73%	17	6,75%	20	7,94%	19	7,54%
2009	5	1,98%	12	4,76%	12	4,76%	10	3,97%	13	5,16%
2010	13	5,16%	18	7,14%	12	4,76%	12	4,76%	14	5,56%
2011	16	6,35%	17	6,75%	15	5,95%	21	8,33%	19	7,54%
2012	12	4,76%	15	5,95%	13	5,16%	11	4,37%	16	6,35%
2007-2012	**78**	**5,16%**	**109**	**7,21%**	**87**	**5,75%**	**93**	**6,15%**	**97**	**6,42%**

Table 4.5 VaR breaks and VaR ratio under student's t-distribution with a sample window of 63 observations.

As one could expect, the model becomes less responsive to changes in market conditions when using a larger sample window. In table 4.6 there is evidence that the model fails to adapt to increasing volatility in 2008 and decreasing volatility in 2009, consequently resulting in large number of VB's in 2008 and a small number of VB's in 2009. However, it is important to stress that the overall number of VB's and VR's for the whole period are smaller through all indices except for the outlier OMXS 30. The number of VB's in this case increased from 78 to 87 when using a larger number of observations in the sample window.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	19	7,54%	30	11,90%	19	7,54%	19	7,54%	22	8,73%
2008	25	9,92%	33	13,10%	27	10,71%	25	9,92%	26	10,32%
2009	4	1,59%	4	1,59%	3	1,19%	5	1,98%	3	1,19%
2010	12	4,76%	11	4,37%	12	4,76%	10	3,97%	12	4,76%
2011	24	9,52%	24	9,52%	6	2,38%	29	11,51%	23	9,13%
2012	3	1,19%	4	1,59%	7	2,78%	4	1,59%	4	1,59%
2007-2012	**87**	**5,75%**	**106**	**7,01%**	**74**	**4,89%**	**92**	**6,08%**	**90**	**5,95%**

Table 4.6 VaR breaks and VaR ratio under student's t-distribution with a sample window of 252 observations.

When increasing the number of observations to 1000 in the sample window, similar results as in the case of normal distribution occur. In table 4.7, 2008 show significant underestimation of VaR with a large number of VB's, while the decreasing volatility in 2009 are in line with expectations. However, the lag in adaptability becomes evident in 2010, where VaR is significantly overestimated. In previous sample windows, overestimation occurred in 2009 and therefore showing evidence that longer sample windows are causing lags in VaR estimation. As in the case of the normal distribution approach, the total number of VB's increased and the distribution of VB's from year to year became more widely spread when using a sample window of 1000 observations.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	VB	VR	VB	VR	VB	VR	VB	VR	VB	VR
2007	21	8,33%	34	13,49%	18	7,14%	15	5,95%	28	11,11%
2008	44	17,46%	56	22,22%	49	19,44%	42	16,67%	55	21,83%
2009	13	5,16%	15	5,95%	11	4,37%	16	6,35%	14	5,56%
2010	4	1,59%	7	2,78%	4	1,59%	4	1,59%	6	2,38%
2011	12	4,76%	7	2,78%	3	1,19%	14	5,56%	8	3,17%
2012	4	1,59%	1	0,40%	0	0,00%	4	1,59%	0	0,00%
2007-2012	**98**	**6,48%**	**120**	**7,94%**	**85**	**5,62%**	**95**	**6,28%**	**111**	**7,34%**

Table 4.7 VaR breaks and VaR ratio under student's t-distribution with a sample window of 1000 observations.

The student's t-distribution gives a better estimation of VaR when compared to the normal distribution approach, presumably because of the distribution's fat tails and leptokurtic features. When comparing the three models, clearly, the 63 and 252 observations sample windows are superior, however, as in the case of the normal distribution approach, it's a matter of choosing between a model that gives the best fit from year to year, or for the whole sample period. One important task in risk management is to give unanimous results, regardless of the market conditions. Therefore, the sample window with 63 observations is the most preferred, even though it disregards important historical data.

4.2.1 Backtest of VaR under Student's t-Distribution

Since continuity in VaR estimates is what this study favours, it is evident that the sample window with 1000 observations fails to satisfy this criterion. In table 4.8 there is evidence that the model fails to cope with changes in market condition. Therefore, the model is rejected under the Kupiec test in estimating VaR during moments of volatility clustering. In the case of a sample window of 252 observations the number of VB's over the whole sample period are satisfactory and within the limit of the Kupiec test. However, the purpose of the thesis is to investigate which model in the best possible way copes with volatility clustering. Evidence show that the model's responsiveness is slow and therefore giving large number of VB's during years when volatility was increasing (2008 and 2011). Not surprisingly, due to the similarity of the parametric approaches, the best responsiveness to volatility clustering is made by the model using 63 observations in the sample window. However, as stressed before, this comes with a price. Significant amounts of data are disregarded when taking this approach and there's no guarantee that this model will be the most suitable in "tranquil" market conditions.

Index				OMXS 30	S&P 500	NIKKEI 225	DAX 30	FTSE 100
Year	**MIN**	**MAX**	**Target**	**63D**	**63D**	**63D**	**63D**	**63D**
2007	6	20	13	16	25	18	19	16
2008	6	20	13	16	22	17	20	19
2009	6	20	13	5	12	12	10	13
2010	6	20	13	13	18	12	12	14
2011	6	20	13	16	17	15	21	19
2012	6	20	13	12	15	13	11	16
2007-2012	**59**	**93**	**76**	**78**	**109**	**87**	**93**	**97**
Year	**MIN**	**MAX**	**Target**	**252D**	**252D**	**252D**	**252D**	**252D**
2007	6	20	13	19	30	19	19	22
2008	6	20	13	25	33	27	25	26
2009	6	20	13	4	4	3	5	3
2010	6	20	13	12	11	12	10	12
2011	6	20	13	24	24	6	29	23
2012	6	20	13	3	4	7	4	4
2007-2012	**59**	**93**	**76**	**87**	**106**	**74**	**92**	**90**
Year	**MIN**	**MAX**	**Target**	**1000D**	**1000D**	**1000D**	**1000D**	**1000D**
2007	6	20	13	21	34	18	15	28
2008	6	20	13	44	56	49	42	55
2009	6	20	13	13	15	11	16	14
2010	6	20	13	4	7	4	4	6
2011	6	20	13	12	7	3	14	8
2012	6	20	13	4	1	0	4	0
2007-2012	**59**	**93**	**76**	**98**	**120**	**85**	**95**	**111**

Table 4.8 Kupiec's Test for 95% VaR under student's t-distribution.

4.3 VaR under Historical Simulation

Most surprisingly the simple HS approach seems to cope well with volatile market conditions. The number of VB's and VR's under a sample window of 63 observations are summarised in table 4.9. Even if there is a clear and consistent underestimation of VaR during periods of increasing volatility and overestimation in periods of decreasing volatility, the HS approach is not showing any large discrepancies. The VR's for the whole sample period are showing more than satisfactory figures of around the expected 5,00%. The HS approach is however showing signs of slow adaptability in 2008 and 2009 due to the lack of volatility forecasting features.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	14	5,56%	18	7,14%	13	5,16%	16	6,35%	21	8,33%
2008	20	7,94%	18	7,14%	17	6,75%	19	7,54%	16	6,35%
2009	4	1,59%	6	2,38%	6	2,38%	9	3,57%	9	3,57%
2010	11	4,37%	11	4,37%	9	3,57%	9	3,57%	11	4,37%
2011	15	5,95%	16	6,35%	15	5,95%	17	6,75%	15	5,95%
2012	10	3,97%	14	5,56%	11	4,37%	10	3,97%	14	5,56%
2007-2012	**74**	**4,89%**	**83**	**5,49%**	**71**	**4,70%**	**80**	**5,29%**	**86**	**5,69%**

Table 4.9 VaR breaks and VaR ratio under historical simulation with a sample window of 63 observations.

In table 4.10, the fact that the HS approach actually does not cope well with quick changes in market conditions gets inevitable. Since the VaR estimate moves from the 4th largest loss when using a sample window of 63 observations, to the 13th largest loss when using a sample window of 252 observations, the slow responsiveness can be seen in 2007-2009 and 2011-2012. Although, when looking at the overall sample period, the HS approach shows VR's ranging from 5,09% to 6,28% which should not be seen as unsatisfactory due to the simplicity of the model.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	21	8,33%	29	11,51%	18	7,14%	19	7,54%	22	8,73%
2008	25	9,92%	29	11,51%	28	11,11%	31	12,30%	26	10,32%
2009	3	1,19%	2	0,79%	1	0,40%	5	1,98%	3	1,19%
2010	12	4,76%	9	3,57%	10	3,97%	7	2,78%	12	4,76%
2011	24	9,52%	23	9,13%	14	5,56%	29	11,51%	19	7,54%
2012	3	1,19%	2	0,79%	6	2,38%	4	1,59%	2	0,79%
2007-2012	**88**	**5,82%**	**94**	**6,22%**	**77**	**5,09%**	**95**	**6,28%**	**84**	**5,56%**

Table 4.10 VaR breaks and VaR ratio under historical simulation with a sample window of 252 observations.

However, the drawbacks of the HS approach are fully shown when using a sample window of 1000 observations. Not only does the VR's differ widely from year to year, but also the overall VR's are showing clear underestimation of VaR using the HS approach. Not surprisingly, the HS approach copes well with the volatile market conditions in 2011 due to the fact that large losses were observed in 2008, which are still included in the sample window.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	25	9,92%	34	13,49%	17	6,75%	12	4,76%	30	11,90%
2008	46	18,25%	56	22,22%	46	18,25%	45	17,86%	58	23,02%
2009	13	5,16%	20	7,94%	13	5,16%	20	7,94%	15	5,95%
2010	4	1,59%	6	2,38%	7	2,78%	7	2,78%	7	2,78%
2011	13	5,16%	7	2,78%	4	1,59%	17	6,75%	10	3,97%
2012	4	1,59%	1	0,40%	2	0,79%	6	2,38%	1	0,40%
2007-2012	**105**	**6,94%**	**124**	**8,20%**	**89**	**5,89%**	**107**	**7,08%**	**121**	**8,00%**

Table 4.11 VaR breaks and VaR ratio under historical simulation with a sample window of 1000 observations.

During a financial crisis, market conditions are changing rapidly, and this is the main drawback of the HS approach. It becomes a matter of choosing between historical data or quickly adapting to new conditions. It becomes evident that the HS approach fails to forecast moments of higher/lower volatility without altering the number of observations in a sample window. In this case, clearly, the HS model using a sample window of 63 days is superior due to the fact that responsiveness is key during volatility clustering. Even with a sample window using a small number of observations, the HS approach still underestimates VaR during volatility clustering.

4.3.1 Backtest of VaR under Historical Simulation

Unlike the parametric approaches, the sample window using 63 observations is showing the best fit when taking the whole sample period in to account. In table 4.12 there are only two exceptions, which are the FTSE 100 in 2007 and OMXS 30 in 2009 where the HS approach gets rejected under the Kupiec test. The lags become noticeable also in the HS approach using a sample period of 252 and 1000 observations. NIKKEI 225 however is an outlier since it is accepted under the Kupiec test for the overall sample period. This is most probably due to the fact that NIKKEI 225 has shown relatively high and constant volatility during the "tranquil" periods compared to the other indices.

Index				OMXS 30	S&P 500	NIKKEI 225	DAX 30	FTSE 100
Year	**MIN**	**MAX**	**Target**	**63D**	**63D**	**63D**	**63D**	**63D**
2007	6	20	13	14	18	13	16	21
2008	6	20	13	20	18	17	19	16
2009	6	20	13	4	6	6	9	9
2010	6	20	13	11	11	9	9	11
2011	6	20	13	15	16	15	17	15
2012	6	20	13	10	14	11	10	14
2007-2012	**59**	**93**	**76**	**74**	**83**	**71**	**80**	**86**
Year	**MIN**	**MAX**	**Target**	**252D**	**252D**	**252D**	**252D**	**252D**
2007	6	20	13	21	29	18	19	22
2008	6	20	13	25	29	28	31	26
2009	6	20	13	3	2	1	5	3
2010	6	20	13	12	9	10	7	12
2011	6	20	13	24	23	14	29	19
2012	6	20	13	3	2	6	4	2
2007-2012	**59**	**93**	**76**	**88**	**94**	**77**	**95**	**84**
Year	**MIN**	**MAX**	**Target**	**1000D**	**1000D**	**1000D**	**1000D**	**1000D**
2007	6	20	13	25	34	17	12	30
2008	6	20	13	46	56	46	45	58
2009	6	20	13	13	20	13	20	15
2010	6	20	13	4	6	7	7	7
2011	6	20	13	13	7	4	17	10
2012	6	20	13	4	1	2	6	1
2007-2012	**59**	**93**	**76**	**105**	**124**	**89**	**107**	**121**

Table 4.12 Kupiec's Test for 95% VaR under historical simulation.

4.4 VaR under Age Weighted Historical Simulation

As expected, the AWHS approach shows satisfactory results throughout the sample period except during 2009 where VaR has been overestimated. Table 4.13 summarises the VB's and VR's while using a sample window of 63 days. The AWHS approach models VaR in a sensible way when facing changes in volatility. The fact that the VR's are close to the desired level of 5,00% during the periods of volatility clustering in 2008 and 2011 show that AWHS is superior to the previous models when facing increasing volatility. However, the AWHS approach also shows signs of struggle when facing decreasing volatility as in 2009 with a clear overestimation throughout every index except DAX 30.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	8	3,17%	13	5,16%	12	4,76%	9	3,57%	15	5,95%
2008	17	6,75%	13	5,16%	14	5,56%	14	5,56%	12	4,76%
2009	3	1,19%	3	1,19%	5	1,98%	8	3,17%	5	1,98%
2010	10	3,97%	10	3,97%	7	2,78%	9	3,57%	10	3,97%
2011	10	3,97%	12	4,76%	12	4,76%	14	5,56%	12	4,76%
2012	9	3,57%	11	4,37%	10	3,97%	8	3,17%	9	3,57%
2007-2012	**57**	**3,77%**	**62**	**4,10%**	**60**	**3,97%**	**62**	**4,10%**	**63**	**4,17%**

Table 4.13 VaR breaks and VaR ratio under age weighted historical simulation with a sample window of 63 observations.

Even when the number of observations in the sample window is increased to 252 the AWHS continues to quickly adapt to increasing volatility. As in the previous cases, the AWHS approach has problem to cope with decreasing volatility as shown in the number of VB's in 2009. Using a lower value of λ, which in this case is set to 0,99, may alter the situation and make the AWHS approach more responsive to volatility clustering. For the overall sample period, the AWHS approach in table 4.14 shows that it slightly overestimates VaR, but is still within a satisfactory level.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	13	5,16%	13	5,16%	14	5,56%	11	4,37%	17	6,75%
2008	17	6,75%	21	8,33%	17	6,75%	23	9,13%	16	6,35%
2009	1	0,40%	1	0,40%	1	0,40%	4	1,59%	2	0,79%
2010	10	3,97%	10	3,97%	10	3,97%	9	3,57%	9	3,57%
2011	18	7,14%	17	6,75%	11	4,37%	20	7,94%	13	5,16%
2012	3	1,19%	4	1,59%	7	2,78%	5	1,98%	4	1,59%
2007-2012	**62**	**4,10%**	**66**	**4,37%**	**60**	**3,97%**	**72**	**4,76%**	**61**	**4,03%**

Table 4.14 VaR breaks and VaR ratio under age weighted historical simulation with a sample window of 252 observations.

In contrast to the other approaches, AWHS, using a sample window of 1000 observations, still copes well with changes in market conditions. Previous models have given extremely high VR's in 2008 when using a sample window of 1000 observations, but the AWHS approach display relatively low VR's compared to them. This means that AWHS still takes important historical data in to account, but due to weighting, are able to give accurate and responsive VaR estimates.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	14	5,56%	18	7,14%	12	4,76%	11	4,37%	17	6,75%
2008	18	7,14%	22	8,73%	17	6,75%	25	9,92%	19	7,54%
2009	1	0,40%	0	0,00%	1	0,40%	4	1,59%	2	0,79%
2010	7	2,78%	8	3,17%	9	3,57%	6	2,38%	8	3,17%
2011	19	7,54%	17	6,75%	11	4,37%	20	7,94%	14	5,56%
2012	3	1,19%	4	1,59%	7	2,78%	4	1,59%	4	1,59%
2007-2012	**62**	**4,10%**	**69**	**4,56%**	**57**	**3,77%**	**70**	**4,63%**	**64**	**4,23%**

Table 4.15 VaR breaks and VaR ratio under age weighted historical simulation with a sample window of 1000 observations.

As have been shown in table 4.13-4.15, the AWHS approach is superior to the regular HS approach since fewer large losses are required in order to alter the VaR estimate. The main attraction of AWHS is that a model that includes a large number of observations in the sample window can still get accurate VaR estimates during volatility clustering. When comparing the different sample windows the worst VR's for the sample period is the sample window using 63 observations, while the 252 and 1000 observations sample window models are giving similar results. However, when looking at consistency, the 252 observations sample window stands out, it has fewer VB's during 2008 and 2011 and consequently displaying more reasonable VaR estimates when faced with volatility clustering.

4.4.1 Backtest of VaR under Age Weighted Historical Simulation

Table 4.16 summarises the Kupiec test under AWHS where it can be concluded that AWHS is superior to the previous approaches when comparing the overall results. Regardless of what number of observations is used in the sample window, AWHS is successful in the Kupiec test for the whole sample period, except for OMXS 30 in the 63- and NIKKEI 225 in the 1000 observations sample window. The only drawback of the AWHS approach is that it shows signs of struggle with decreasing volatility as in the case of 2009. This is most probably an effect of setting λ to 0,99, and lowering λ would make older observations irrelevant quicker, therefore coping better with decreasing (and increasing) volatility.

Index				OMXS 30	S&P 500	NIKKEI 225	DAX 30	FTSE 100
Year	**MIN**	**MAX**	**Target**	**63D**	**63D**	**63D**	**63D**	**63D**
2007	6	20	13	8	13	12	9	15
2008	6	20	13	17	13	14	14	12
2009	6	20	13	3	3	5	8	5
2010	6	20	13	10	10	7	9	10
2011	6	20	13	10	12	12	14	12
2012	6	20	13	9	11	10	8	9
2007-2012	**59**	**93**	**76**	**57**	**62**	**60**	**62**	**63**
Year	**MIN**	**MAX**	**Target**	**252D**	**252D**	**252D**	**252D**	**252D**
2007	6	20	13	13	13	14	11	17
2008	6	20	13	17	21	17	23	16
2009	6	20	13	1	1	1	4	2
2010	6	20	13	10	10	10	9	9
2011	6	20	13	18	17	11	20	13
2012	6	20	13	3	4	7	5	4
2007-2012	**59**	**93**	**76**	**62**	**66**	**60**	**72**	**61**
Year	**MIN**	**MAX**	**Target**	**1000D**	**1000D**	**1000D**	**1000D**	**1000D**
2007	6	20	13	14	18	12	11	17
2008	6	20	13	18	22	17	25	19
2009	6	20	13	1	0	1	4	2
2010	6	20	13	7	8	9	6	8
2011	6	20	13	19	17	11	20	14
2012	6	20	13	3	4	7	4	4
2007-2012	**59**	**93**	**76**	**62**	**69**	**57**	**70**	**64**

Table 4.16 Kupiec's Test for 95% VaR under age weighted historical simulation.

4.5 VaR under Volatility Weighted Historical Simulation

In table 4.17 there is evidence that the VWHS approach using a sample window of 63 observations gives consistent VR's from year to year but with a slight underestimation. As been recalled earlier, even though the NIKKEI 225 index displays consistent high volatility throughout the sample period, disregarding the spike in the end of 2008, VWHS still fails to display reasonable VR's. It was anticipated that the VWHS approach would allow the EWMA volatility forecasting to make sound adjustments to the VaR estimation. Evidently this was not the case, NIKKEI 225 index is the worst performer during the "spiky" period of 2008 and 2011.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	17	6,75%	21	8,33%	19	7,54%	21	8,33%	22	8,73%
2008	21	8,33%	17	6,75%	22	8,73%	22	8,73%	21	8,33%
2009	16	6,35%	16	6,35%	16	6,35%	18	7,14%	19	7,54%
2010	12	4,76%	15	5,95%	16	6,35%	17	6,75%	15	5,95%
2011	17	6,75%	16	6,35%	24	9,52%	17	6,75%	20	7,94%
2012	15	5,95%	18	7,14%	13	5,16%	17	6,75%	18	7,14%
2007-2012	**98**	**6,48%**	**103**	**6,81%**	**110**	**7,28%**	**112**	**7,41%**	**115**	**7,61%**

Table 4.17 VaR breaks and VaR ratio under volatility weighted historical simulation with a sample window of 63 observations.

The VWHS approach using a sample window of 252 observations shows somewhat disperse results. In 2007, before the crisis, VWHS clearly underestimates VaR, but manages to adapt to the increasing volatility in 2008. In 2009 however, the model fails to cope with decreasing volatility, most probably due to λ-factor of 0,94. As expected when volatility decreases after a spike, the model fails to respond to these changes seen in both 2009 and 2012. There is evidence that the VWHS approach using a sample window of 252 observations is giving reasonable VaR estimates. When observing the VR's for the whole sample period, it ranges from 5,49% to 6,02% and is rather consistent throughout all indices.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	20	7,94%	25	9,92%	21	8,33%	17	6,75%	19	7,54%
2008	15	5,95%	16	6,35%	16	6,35%	21	8,33%	12	4,76%
2009	1	0,40%	0	0,00%	2	0,79%	2	0,79%	6	2,38%
2010	17	6,75%	22	8,73%	26	10,32%	19	7,54%	22	8,73%
2011	28	11,11%	21	8,33%	12	4,76%	24	9,52%	20	7,94%
2012	2	0,79%	7	2,78%	14	5,56%	3	1,19%	4	1,59%
2007-2012	**83**	**5,49%**	**91**	**6,02%**	**91**	**6,02%**	**86**	**5,69%**	**83**	**5,49%**

Table 4.18 VaR breaks and VaR ratio under volatility weighted historical simulation with a sample window of 252 observations.

Similar to the HS and parametric approaches, when using a sample window of 1000 observations, the VWHS approach severely over- and underestimates VaR. Old losses with high (low) volatility are still scaled reasonably high (low) when using VWHS, causing an overestimation (underestimation). As mentioned in the previous paragraph, this is probably due to the λ-factor being too large. VWHS fails to give satisfactory VR's throughout all years displaying the drawback of using a large number of observations in the sample window during periods of extreme changes in volatility.

Index	OMXS 30		S&P 500		NIKKEI 225		DAX 30		FTSE 100	
Year	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**	**VB**	**VR**
2007	39	15,48%	44	17,46%	26	10,32%	33	13,10%	52	20,63%
2008	42	16,67%	49	19,44%	41	16,27%	50	19,84%	43	17,06%
2009	6	2,38%	6	2,38%	6	2,38%	12	4,76%	4	1,59%
2010	0	0,00%	0	0,00%	1	0,40%	2	0,79%	0	0,00%
2011	8	3,17%	3	1,19%	3	1,19%	9	3,57%	5	1,98%
2012	3	1,19%	1	0,40%	4	1,59%	4	1,59%	0	0,00%
2007-2012	**98**	**6,48%**	**103**	**6,81%**	**81**	**5,36%**	**110**	**7,28%**	**104**	**6,88%**

Table 4.19 VaR breaks and VaR ratio under volatility weighted historical simulation with a sample window of 1000 observations.

A smaller λ-factor would assign more weight to the returns, rather than variance, when forecasting σ^2_{T+1}. The model will be less responsive to changes in volatility when a larger number of observations are used in the sample window. Therefore, lowering λ would force the model to consider previous day's return, rather than variance, possibly making it more responsive. The only model that gives satisfactory results is the one using a sample window of 63 observations. The forecasted σ^2_{T+1} does not seem to differ enough from the actual variance causing the model to fail. Most probably, the changes in volatility are too quick and this combined with a λ equal to 0,94 makes the VWHS not responsive enough, causing unsatisfactory VR's.

4.5.1 Backtest of VaR under Volatility Weighted Historical Simulation

Even though the model using a sample window of 63 observations gives a consistent and satisfactory yearly number of VB's, it still fails under the Kupiec test when considering the whole sample period. In contrast, the sample window with 252 observations gives inconsistent yearly number of VB's but within the range of the Kupiec test for the period 2007-2012. This phenomenon was also seen in the student's t-distribution approach. Possibly, the assumption that the return distribution of a financial crisis lifecycle can be explained parametrically could be made. The sample window using 1000 observations is however not accepted under Kupiec's test on all levels. NIKKEI 225 is somewhat of an outlier compared to the other indices, where 81 VB's were detected compared to 91 and 110 VB's over the whole sample period.

Index				OMXS 30	S&P 500	NIKKEI 225	DAX 30	FTSE 100
Year	**MIN**	**MAX**	**Target**	**63D**	**63D**	**63D**	**63D**	**63D**
2007	6	20	13	17	21	19	21	22
2008	6	20	13	21	17	22	22	21
2009	6	20	13	16	16	16	18	19
2010	6	20	13	12	15	16	17	15
2011	6	20	13	17	16	24	17	20
2012	6	20	13	15	18	13	17	18
2007-2012	**59**	**93**	**76**	**98**	**103**	**110**	**112**	**115**
Year	**MIN**	**MAX**	**Target**	**252D**	**252D**	**252D**	**252D**	**252D**
2007	6	20	13	20	25	21	17	19
2008	6	20	13	15	16	16	21	12
2009	6	20	13	1	0	2	2	6
2010	6	20	13	17	22	26	19	22
2011	6	20	13	28	21	12	24	20
2012	6	20	13	2	7	14	3	4
2007-2012	**59**	**93**	**76**	**83**	**91**	**91**	**86**	**83**
Year	**MIN**	**MAX**	**Target**	**1000D**	**1000D**	**1000D**	**1000D**	**1000D**
2007	6	20	13	39	44	26	33	52
2008	6	20	13	42	49	41	50	43
2009	6	20	13	6	6	6	12	4
2010	6	20	13	0	0	1	2	0
2011	6	20	13	8	3	3	9	5
2012	6	20	13	3	1	4	4	0
2007-2012	**59**	**93**	**76**	**98**	**103**	**81**	**110**	**104**

Table 4.20 Kupiec's Test for 95% VaR under volatility weighted historical simulation.

4.6 Christoffersen Test of Independence

Since the Christoffersen test of independence is χ^2 distributed with one degree of freedom the critical value is equal to 3,841. Most surprisingly all VaR approaches under all sample windows are showing that the VB's are independently distributed when using S&P 500 as the underlying data. This most probably means that a large loss was eventually followed by another large loss the day after for S&P 500. Consequently, the conclusion that significant signs of volatility clustering, of which the models cannot cope with has been observed, and VB's are thus independently distributed.

However, when observing the other indices, there is a pattern in the models. The normal distribution approach seems to cope well with volatility clustering with a small sample window, but as sample windows grow, it gets less responsive. Student's t-distribution on the other hand shows signs of responding better to volatility clustering using a sample window of 252 observations rather than 63 and 1000 observations.

The non-parametric approaches seem to cope as good as, or even better than, in the case of AWHS, the parametric approaches. Not surprisingly, the non-parametric approaches favour the sample window using 63 observations, making them more responsive than the other models. The AWHS approach seems to prefer the sample window of 1000 observations rather than 252 observations when testing for independence.

The troublesome OMXS 30 together with DAX 30 indices are showing the most significant signs of volatility clustering while the VWHS seems to be the model that is the closest of reaching independence of VB's. The VWHS approach is somewhat showing signs of independence throughout all three models and indices with a few exceptions where higher likelihood ratios were observed. But in the end, compared to the parametric and non-parametric approaches, the VWHS approach seems to cope with volatility clustering, in regards to independence of VB distribution, the best.

Index	OMXS 30	S&P 500	NIKKEI 225	DAX 30	FTSE 100
2007-2012	63D	63D	63D	63D	63D
Norm Dist	3,0834	0,4091	3,8627	3,2138	2,3853
T-Distr	5,6330	0,4091	4,8906	4,9257	1,4955
HS	6,9140	1,7943	0,1943	9,2894	2,1537
AWHS	3,2687	0,0970	1,1321	0,1233	3,8236
VWHS	3,6612	1,4982	2,3945	0,0000	4,9796
2007-2012	252D	252D	252D	252D	252D
Norm Dist	12,2669	0,1902	6,2535	15,6951	2,4843
T-Distr	8,8164	0,1057	3,0731	16,3252	2,6760
HS	15,7914	3,1811	4,0491	23,1222	2,5174
AWHS	11,8200	0,2634	1,1321	5,4098	6,6426
VWHS	10,4433	1,3788	2,4843	7,0700	4,2475
2007-2012	1000D	1000D	1000D	1000D	1000D
Norm Dist	5,6371	0,9288	14,1365	13,5204	9,6591
T-Distr	5,1666	0,9288	14,6537	4,3943	9,6591
HS	8,1492	1,8942	10,2568	7,7916	9,4341
AWHS	6,3068	0,3041	1,5403	8,3618	5,6682
VWHS	3,6612	2,6026	4,7741	6,4534	8,5162

Table 4.21 Christoffersen test of independence using a 95% confidence level.

5. Analysis

In the event of a financial crisis, as we have seen, the phenomenon called volatility clustering occur. Figure 3.1-3.5 clearly displays the market characteristics where a “tranquil” period, with low and steady volatility, suddenly is followed by an extremely volatile market. Throughout the empirical evidence section, the models have reacted differently to these sudden changes. It can be assumed that under “tranquil” periods, all models will display satisfactory number of VB’s and reasonable VR’s. If volatility is constant, regardless of if a parametric, non-parametric or semi-parametric approach is being used to estimate VaR, the different approaches will show somewhat similar VaR estimates. Therefore, it is only in the case of (extreme) changes in market conditions a true and reliable evaluation of VaR models can be made.

Consistently throughout the empirical evidence section, the approaches using a sample window of 63 observations are showing the best responsiveness to changes in volatility. This result is however not surprising, the models favour short sample windows when considering responsiveness. Since the base for all models is volatility, a standard deviation that considers fewer observations will faster adapt to the changes. “Old” observations are in fact not “old”, when using a small number of observations in the sample window. It should be stressed that disregarded observations however contains important information of market conditions and may be important when estimating VaR. A favourable model should be able to cope with “tranquil” periods as well as with volatility clustering. Therefore, a model that considers a larger number of observations will be the most preferable over time.

In order to compare the parametric and non-parametric approaches figure 5.1 and 5.2 were considered. It displays the evolution of VaR estimates over the period 2008-2009 for the student’s t-distribution and AWHS approach[8]. The difference between assuming that returns follow a certain parametric distribution and assuming a return distribution based on historical data becomes

[8] Respective figures for the other approaches can be found in the Appendix

evident. When considering the "tranquil" period before Q3 2008 both approaches show somewhat similar VaR estimates, regardless of what number of observations is used in the sample window. However, when the financial crisis eventually hit in Q3 2008 the differences becomes evident. Where the student's t distribution approach fails to adapt to the volatility clustering when using longer sample windows, the AWHS approach adapts reasonably fast over all sample windows. As soon as the returns behave irrationally and diverge from the assumed return distribution, the parametric approaches fail to give reasonable VaR estimates. Even when using a shorter sample window, the parametric approach is not as responsive as the corresponding AWHS model. The assumption that can be made is that during "tranquil" periods the actual return distribution is similar to a parametric return distribution, but during volatility clustering it certainly is not.

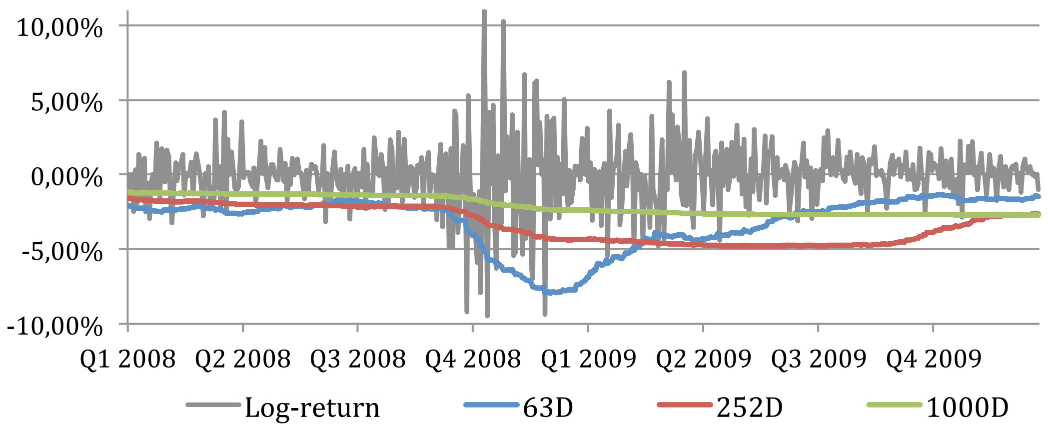

Figure 5.1 Student's t-distribution VaR and logarithmic return for S&P 500 during 2008 and 2009.

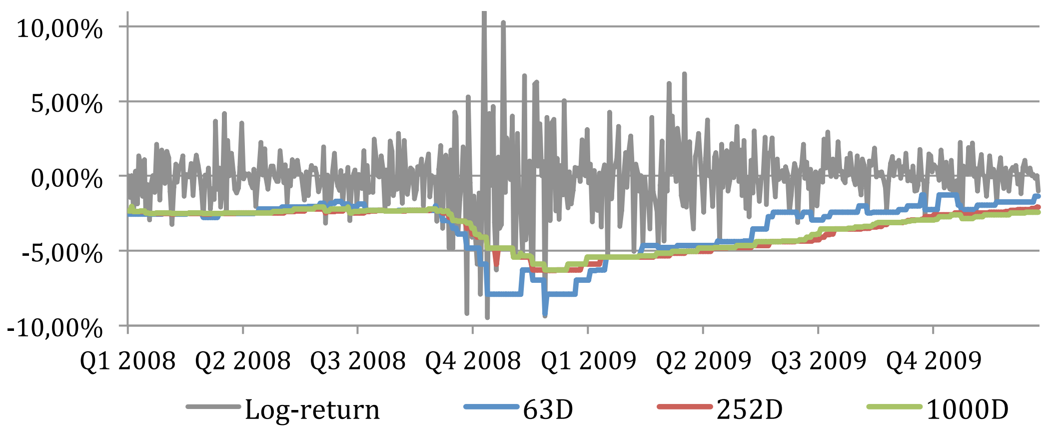

Figure 5.2 AWHS VaR and logarithmic return for S&P 500 during 2008 and 2009.

The student's t-distribution approach is rather complicated in the sense that an estimation of d.o.f. needs to be made. Instead of estimating VaR using a d.o.f. equal to approximately 4,0, proposed by Markowits and Usmen (1996a, 1996b), Hurst and Platen (1997) and Ferguson and Platen (2006), the index specific d.o.f. was used, which ranged from 4,7 to 5,6. This consequently led to higher VaR estimates explained mathematically below. Consider a mean, μ, equal to 0 and standard deviation, σ, equal to 1.

$$\left\{\left(\sqrt{\frac{5-2}{5}}\,t_{5\%,5}\right)\Big/\left(\sqrt{\frac{4-2}{4}}\,t_{5\%,4}\right)\right\}-1 = 3{,}22\% \qquad 5.1$$

While using a d.o.f. equal to 5, the VaR estimate will be 3,22 % larger than using a d.o.f. equal to 4. This eventually leads to the student's t-distribution coping better with high volatility, causing less underestimation, while during moments of low volatility, overestimate VaR. Therefore, it can be assumed that a higher d.o.f. will cause less VB's when volatility is high, and is therefore a better fit when unpredicted increase in volatility occur.

Another aspect to consider when analysing the different approaches is the distribution of VB's. In the empirical evidence section, there have been signs of VB's being evenly distributed from year to year when using a model with fewer observations in the sample window. Similar to the previous discussion, this is linked to the responsiveness of the models. Hull and White (1998) claims that the VWHS approach responds better to moments of volatility clustering than the HS and AWHS approaches since it accounts for changes in market conditions in a direct way. However, the Kupiec test has proven that when using EWMA to forecast volatility, it fails to account for sudden and larger changes. The distribution of VB's in figure 5.3 clearly displays this shortcoming. In contrast to the best performer, in this case the AWHS model (figure 5.4), the VB's only occur when volatility increases. Evidently, the EWMA model fails to forecast volatility in a reliable manner. This is consistent with the study of Mabrouk and Saadi (2012) where they found that the EWMA model proposed by RiskMetrics was the worst performer when comparing different parametric approaches. However, lowering the decay factor proposed by RiskMetrics of 0,94 would make the volatility forecast to consider returns rather than volatility. Since volatility has a significant lag when using larger sample windows, considering previous day's return would make it more responsive, in the case when another large loss is observed the day after. Therefore, the assumption can be made that VWHS using EWMA to forecast volatility is more suitable when smaller changes in volatility are observed. Another suggestion might be to use a GARCH model to forecast volatility since it rescales the weighting factor between variance and the error term, possibly making VWHS more responsive.

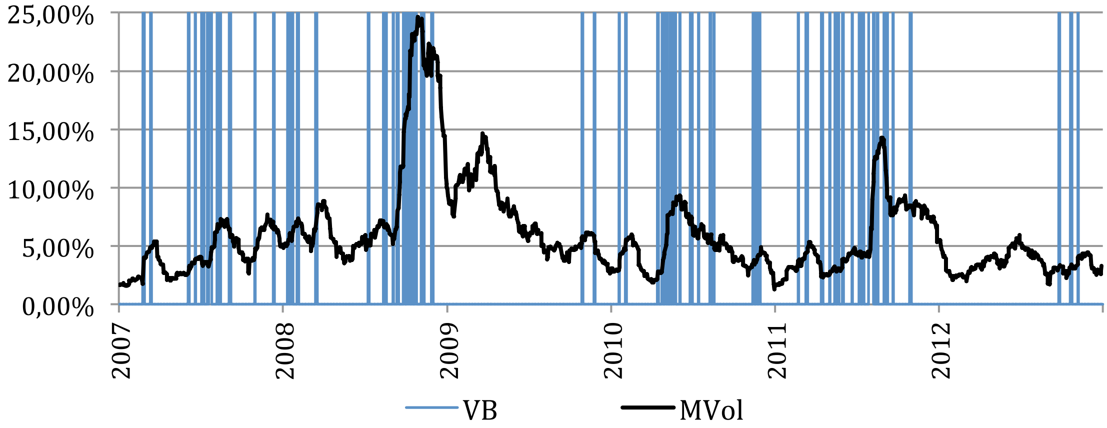

Figure 5.3 DAX 30 monthly volatility and VB distribution using VWHS with a sample window of 252 observations.

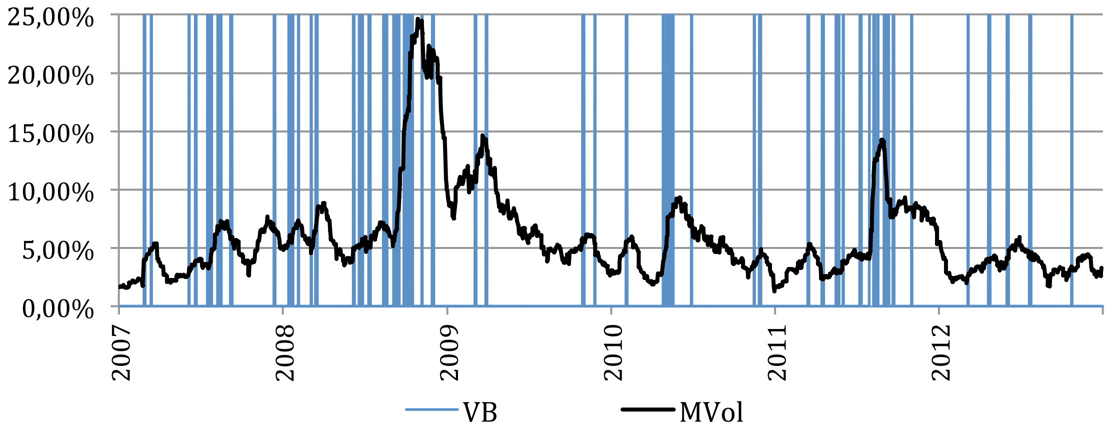

Figure 5.4 DAX 30 monthly volatility and VB distribution using AWHS with a sample window of 252 observations.

When comparing the different approaches side by side it is easy to draw the conclusion that a model favouring responsiveness is the best model to use when a financial crisis strikes. In the previous section it was shown that the superior approach, considering all sample windows, is the AWHS. It has become evident that the parametric approaches fails to cope with irrational return distributions, and that EWMA cannot accurately forecast large changes in volatility. Since both parametric and semi-parametric approaches fails to give reliable VaR estimates, a non-parametric approach that disregards the assumptions made by the other approaches performs on a satisfactory level. The AWHS approach's feature of weighting observations based on their age in the sample window enables it to disregard a parametric return distribution and consider "actual" volatility instead of forecasting. In contrast to regular HS where VaR is estimated as the X^{th} largest loss, VaR under the AWHS can be estimated as almost any of the observed losses, depending on which decay factor is used, therefore responding quicker to volatility clustering.

There is also evidence showing that the results from the Christoffersen test are dependent on the kurtosis of the indices. The index with the highest kurtosis, S&P 500, show that the VB's are independently distributed, while the indices with the lowest kurtosis fails to be accepted under the null hypothesis. This can be explained by higher kurtosis in the return distribution displays fatter tails and consequently gets accepted under the null hypothesis of independently distributed VB's.

6. Conclusion

As Goldman Sachs CEO, Lloyd Blankfein, stated while testifying in front of the FCIC, risk models were too often allowed as a substitute for judgement (CSPAN, 2010). This study has shown that even if some of the models' responsiveness can be questioned, evidence of an arising financial crisis was apparent. A smaller number of observations in the sample window somewhat overcame the issue of slow adaptable VaR models. However, with the arising Basel III accord, some guidelines in which risk models are adequate should be pointed out, possibly minimising a financial crisis of the same magnitude as in 2008.

Since both the responsiveness and a large number of observations in the sample window is important over time when estimating VaR, the conclusion is that the AWHS approach is superior to the other approaches tested. The parametric models might perform on a satisfactory level during low and steady volatility, but as the returns behave irrationally, it fails to cope with these changes. Using the EWMA model to forecast volatility, surprisingly, does not respond well to changes in volatility, most probably due to the magnitude of the volatility clusters. As stated before, there might be more reliable volatility forecasting models such as the GARCH model or altering the λ-factor. The non-parametric approaches, especially the AWHS, are superior to the other models that have been tested. However, the regular HS approach shows significant lags due to the lack of weighting of observations. The AWHS approach has shown features of handling volatility clustering in a more direct and responsive way, but with the drawback of overestimating VaR during decreasing volatility, regardless of which sample window being used.

Finally, it is important to stress that financial institutions and investors should not heavily rely on only one model, but rather analyse results of different VaR approaches using different number of observations in the sample window. Since it has been shown that all models either over- or underestimate VaR during volatility clustering, it will eventually enable them to make sound judgements of their risk.

Bibliography

Abad, P., Benito, S. & López, C., 2013. A comprehensive review of Value at Risk methodologies. *The Spanish Review of Financial Economics*.

Barone-Adesi, G., Bourgoin, F. & Giannopoulos, K., 1998. Don't look back. *Risk*, 11(August), pp.100-03.

Beder, T.S., 1995. VAR: Seductive but Dangerous. *Financial Analysts Journal*, 51(September/October), pp.12-24.

BIS, B.f.I.S., 2011. *Basel III: A global regulatory framework for more resilient banks and banking systems*. [Online] Available at: http://www.bis.org/publ/bcbs189.pdf [Accessed 20 April 2014].

Bloomberg, 2014. *Bloomberg*. [Online] Available at: www.bloomberg.com [Accessed 03 May 2014].

Boudoukh, J., Richardson, M. & Whitelaw, R., 1998. The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk. *Risk*.

Bryman, A. & Bell, E., 2011. *Business Research Methods*. 3rd ed. Oxford: Oxford University Press.

Campbell, S.D., 2005. A Review of Backtesting and Backtesting Procedures. *Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.*, 21.

Campbell, A., 2009. The risk of value-at-risk. *Risk Magazine*, 01 April. pp.42-46.

CSPAN, 2010. *CSPAN - Financial Crisis Inquiry Commission*. [Online] Available at: http://www.c-spanvideo.org/program/291292-1 [Accessed 16 April 2014].

Escanciano, J.C. & Pei, P., 2012. Pitfalls in backtesting Historical Simulation VaR models. *Journal of Banking & Finance*, 36(8), pp.2233-44.

Daníelsson, J. et al., 2013. Fat tails, VaR and subadditivity. *Journal of Econometrics*, 172(2), pp.283-91.

Dowd, K., 2002. *Measuring Market Risk*. 1st ed. West Sussex: John Wiley & Sons Ltd.

Fergusson, K. & Platen, E., 2006. On the Distributional Characterization of Daily Log-Returns of a World Stock Index. *Applied Mathematical Finance*, 13(1), pp.19-38.

Hull, J. & White, A., 1998. Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, 1(Fall), pp.5-19.

Hurst, S.R. & Platen, E., 1997. The marginal distributions of returns and volatility. In Y. Dodge, ed. *L1-Statistical Procedures and Related Topics, Volume 31 of IMS Lecture Notes - Monograph Series*. Hayward, California: Institute of Mathematical Statistics. pp.301-14.

Hendricks, D., 1996. Evaluation of Value-at-Risk Models Using Historical Data. *Economic Policy Review*, 2(1), pp.39-70.

Hoppe, R., 1998. VaR and the unreal world. *Risk*, 11(July), pp.45-50.

Ju, X. & Pearson, N.D., 1999. Using Value-at-Risk to Control Risk Taking: How Wrong Can You Be? *Journal of Risk*, 1(2), pp.5-36.

Kupiec, P.H., 1995. Techniques for Verifying the Accuracy of Risk Management Models. *Journal of Derivatives*, 3, pp.73-84.

Lin, C.-H. & Shen, S.-S., 2006. Can the student-t distribution provide accurate value at risk? *Journal of Risk Finance*, 7(3), pp.292-300.

Nilsson, B., 2013. *Financial Valuation and Risk Management - NEKN83*. Lecture Notes. Lund: Birger Nilsson Lund University: School of Economics and Management.

Mabrouk, S. & Saadi, S., 2012. Parametric Value-at-Risk analysis: Evidence from stock indices. *The Quarterly Review of Economics and Finance*, 52(2012), pp.305-21.

Mandelbrot, B., 1963. The Variation of Certain Speculative Prices. *The Journal of Business*, 36(4), pp.394-419.

Markowitz, H.M. & Usmen, N., 1996a. The likelihood of various stock market return distributions, Part 1: Principles of inference. *Journal of Risk and Uncertainty*, 13(3), pp.207-19.

Markowitz, H.M. & Usmen, N., 1996b. The likelihood of various stock market return distributions, Part 2: Empirical results. *Journal of Risk and Uncertainty*, 13(3), pp.221-47.

Olson, D.L. & Desheng, W., 2013. The impact of distribution on value-at-risk measures. *Mathematical and Computer Modelling*, 58(9-10), pp.1670-76.

Pérignon, C. & Smith, D.R., 2010. The level and quality of value-at-risk disclosure by commercial banks. *Journal of Banking and Finance*, 34(2), pp.362-77.

Platen, E. & Sidorowicz, R., 2007. *Empirical Evidence on Student-t Log-Returns of Diversified World Stock Indices*. Research Paper Series 194. Sydney: Quantitative Finance Research Centre University of Technology.

Sage, 2012. *t-Distribution Interact*. [Online] Available at: http://interact.sagemath.org/node/79_250414 [Accessed 25 April 2014].

Sollis, R., 2009. Value at risk: a critical overview. *Journal of Financial Regulation and Compliance*, 17(4), pp.398-414.

Taleb, N., 1997. The World According to Nassim Taleb. *Derivatives Strategy*, 2(December/January), pp.37-40.

Zhang, M.-H. & Cheng, Q.-S., 2005. An approach to VaR for capital markets with Gaussian mixture. *Applied Mathematics and Computation*, 168(2), pp.1079-85.

Appendix

2008-2009 VaR

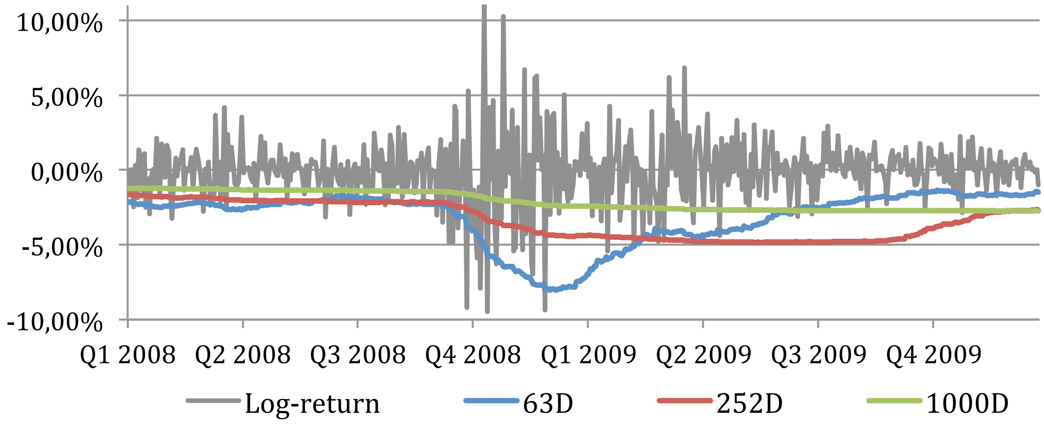

Figure A.1 Normal distribution VaR and logarithmic return for S&P 500 during 2008 and 2009.

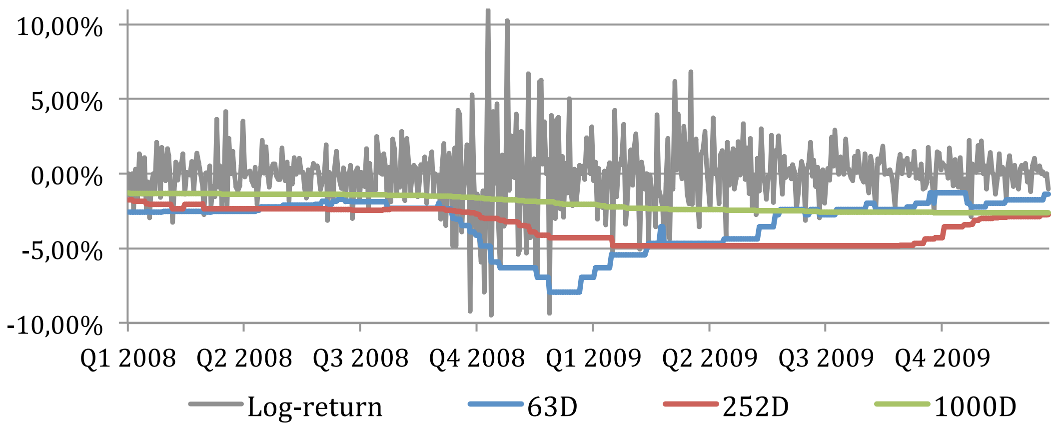

Figure A.2 HS VaR and logarithmic return for S&P 500 during 2008 and 2009.

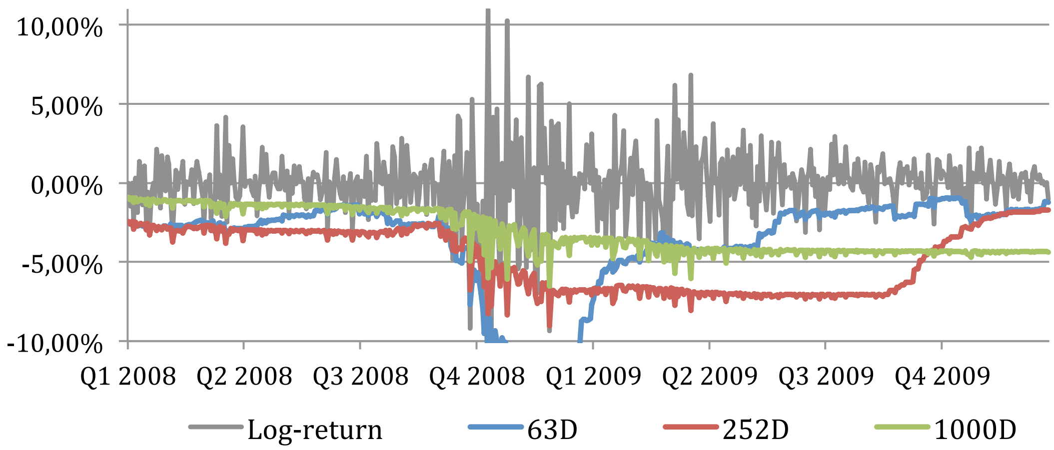

Figure A.3 VWHS VaR and logarithmic return for S&P 500 during 2008 and 2009.

VBA Macro AWHS

```
Sub AWHS()

Dim i As Integer
Dim j As Integer
Dim HistoricalData As Worksheet
Dim Calc As Worksheet
Dim r_input_return As Range
Dim r_weight As Range
Dim r_weight_input As Range

Set data_copy = Worksheets("HistoricalData")
Set data_paste = Worksheets("Calc")

'### 2 = 31st of December 2012 and 1567 = 1st of January
2007
For i = 2 To 1567

'### Column B in sheet HistoricalData contains log-
returns from 1st of January 2006 to 31st of December 2012
    With data_copy
        Set SourceRange = .Range("B" & i & ":B" & i +
252)
    End With

'### "r_input_return" is where the 252 day sample period
is pasted
    With data_paste
        Set destRange = .Range("r_input_return")
    End With

'### Pastes the rolling 252 day sample period in to
"r_input_return"
    destRange.Value = SourceRange.Value

    With data_paste
'### "r_weight" is a prespecified weight distribution
where lambda = 0.99 for the 252 days sample period
        .Range("r_weight").Value =
        .Range("r_weight_input").Value
    End With

'### Sorts return column in ascending order in order to
find what the VaR estimate is – Column A is
"r_input_return" and column B is the weighting
        Columns("A:B").Select

ActiveWorkbook.Worksheets("Calc").Sort.SortFields.Clear
    ActiveWorkbook.Worksheets("Calc").Sort.SortFields.Add
Key:=Range("A1"), _
```

```
        SortOn:=xlSortOnValues, Order:=xlAscending,
DataOption:=xlSortNormal
    With ActiveWorkbook.Worksheets("Calc").Sort
        .SetRange Range("A2:B253")
        .Header = xlNo
        .MatchCase = False
        .Orientation = xlTopToBottom
        .SortMethod = xlPinYin
        .Apply
    End With

'### Pastes the VaR estimate in to a range
    With data_paste
        For j = 2 To 253
            If .Range("C" & j).Value > 0.05 Then
                .Range("H" & i).Value = .Range("A" &
j).Value
                Exit For
            Else
            End If
        Next j
    End With
Next i

End Sub
```

VBA Macro VWHS

```
Sub VWHS()

Dim i As Integer
Dim j As Integer
Dim HistoricalData As Worksheet
Dim Calc As Worksheet
Dim r_input_return As Range
Dim r_EWMA As Range
Dim r_VWHS_loss As Range
Dim r_sigmaT As Range
Dim r_daily_VaR As Range

Set data_copy = Worksheets("HistoricalData")
Set data_paste = Worksheets("Calc")

'### 2 = 31st of December 2012 and 1567 = 1st of January
2007
For i = 2 To 1567

'### Column B in sheet HistoricalData contains log-
returns from 1st of January 2006 to 31st of December 2012
    With data_copy
        Set SourceRange = .Range("B" & i & ":B" & i +
252)
    End With

'### "r_input_return" is where the 252 day sample period
is pasted
    With data_paste
        Set destRange = .Range("r_input_return")
    End With

'### Pastes the rolling 252 day sample period in to
"r_input_return"
    destRange.Value = SourceRange.Value

'### Column C in sheet HistoricalData contains the daily
standard deviations for the sample period
    With data_copy
        Set sigmaTsourceRange = .Range("C" & i & ":C" & I
+ 252)
    End With

'### "r_sigmaT" is where the standard deviations are
pasted
    With data_paste
        Set sigmaTdestRange = .Range("r_sigmaT")
    End With
```

```
'### Pastes the rolling 252 day sample period in to
"r_sigmaT"
    sigmaTdestRange.Value = sigmaTsourceRange.Value

'### Column D in sheet HistoricalData contains the EWMA
forecasted standard deviation for T+1
    With data_copy
        Set EWMAsourceRange = .Range("D" & i)
    End With

'### Single cell in sheet Calc which is used in formulas
on sheet Calc to calculate the scaled losses
    With data_paste
        Set EWMAdestRange = .Range("r_EWMA")
    End With

'### Pastes the EWMA forecasted standard deviation for
T+1
    EWMAdestRange.Value = EWMAsourceRange.Value

'### "r_daily_VaR" is a single cell where the Xth largest
loss in the range "r_VWHS_loss" which contains the
formulas for scaling the losses
    With data_paste
        .Range("H" & i).Value =
.Range("r_daily_VaR").Value
    End With
Next i

End Sub
```