



Stockholms
universitet

Value at Risk with and without report dates

Oscar Wissén

Kandidatuppsats 2016:17
Matematisk statistik
Juni 2016

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm



Mathematical Statistics
Stockholm University
Bachelor Thesis **2016:17**
<http://www.math.su.se>

Value at Risk with and without report dates

Oscar Wissén*

June 2016

Abstract

Mathematical and/or statistical models are used to help predict future value and variance of the financial market variables. One model used is Value at Risk. It describes potential losses attached to an investment. This thesis will use the joint estimation of an AR(p)-GARCH(1,1) model to develop a model for Value at Risk. This thesis concentrates on the difference between one data-set containing the returns for report days and another set where these are excluded. The results reveal that for the stock Swedbank.A the exclusion of reports seem to give a better fit for our model.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden.
E-mail: oscar.wissen@gmail.com. Supervisor: Joanna Tyrcha, Filip Lindskog and Mathias Lindholm.

Contents

1	Introduction	3
1.1	Aim of thesis	4
2	Data	4
3	Theoretical Background	5
3.1	Characteristics of a time series	5
3.2	Return series	5
3.3	Autocorrelation function (ACF)	6
3.4	Autoregressive models	7
3.5	Heteroskedasticity	7
3.6	ARCH and GARCH models	7
3.6.1	ARCH	7
3.6.2	GARCH	8
3.7	Akaike information criterion (AIC)	9
3.8	Value at Risk (VaR)	9
4	Models and Methods	10
4.1	Descriptive analysis	11
4.2	Test for normality	12
4.3	Test for autocorrelation	14
4.3.1	Applying an AR-model	15
4.4	Test for heteroskedasticity in residuals	17
4.4.1	Applying an AR-GARCH model to data	17
4.5	Backtesting	18
4.6	Software	19
5	Result	19
6	Discussion and conclusion	23
7	Acknowledgement	24
8	Appendix	26
8.1	Distributions	26
8.2	Partial Autocorrelation Function	26
8.3	Maximum likelihood estimation	26
8.4	Central moments	27
8.5	Graphs	29
8.5.1	Graphs for fitted GARCH models	30
8.5.2	Graphs for backtesting	32

1 Introduction

Risk is calculated in order to optimize exposure to it. Having only a vague or no estimate of risks may lead to catastrophic outcomes. Investment banks, insurance companies and government institutions did not have full awareness of their risks in 2008. This contributed to the collapse of banks like Lehman Brothers, the bailout of AIG for 85 billion US dollars and ultimately escalated to a world-wide recession. To avoid a new crisis, the period after the crises was filled with new regulatory rules aimed to minimize the risks in financial markets. In Europe the EBA (European Banking Authority) were formed by the EU in 2012. Since its formation the EBA has made stress tests on all major banks in Europe. The tests were conducted in 2014 and can be found on <http://www.eba.europa.eu/>. A new stress test has been announced to be carried out in 2016. With risk so high up on the agenda we will try and measure the risk for one stock listed on Nasdaq Stockholm. The method used is Value at Risk.

The Value at Risk (VaR) was developed in the late 1980:s by J.P Morgan to estimate the risk of a portfolio or asset. VaR is a point estimate of the worst potential loss that will not be exceeded more than a certain percentage of times. One way to obtain the VaR is to use historical data to estimate the empirical distribution of the underlying financial asset. With this distribution we may calculate the p :th quantile. This quantile can then be used to estimate the VaR. This thesis estimates the quantile and models the empirical distribution.

One of the many problems when using the time series approach to VaR is that financial time series commonly possess autocorrelation and heteroskedasticity. This occurs when the time series are dependent on the previous outcome. If a crisis hits, the volatility in the financial market would be higher for a period of time and not just one day. Given this information, a approach where we let all observed data have the same weight for today's distribution is problematic.

To be able to predict future distributions, methods for autocorrelation and heteroskedasticity has been developed. If the return of an underlying financial asset is observed to have patterns in it, we have autocorrelation. This is adjusted for with an autoregressive model. The model tries to catch these patterns and can be used to help predict future patterns in the return of a asset. Heteroskedasticity is dealt with through a GARCH model. The GARCH model uses today's variance to predict tomorrow's.

The complex behaviour of the financial market has been studied intensively under the last decades. One of the models developed will be used in this paper, namely the joint model of a AR(p)-GARCH(1,1) model. This model will be used to try and a predict future VaR. The estimated AR(p)-GARCH(1,1) model uses the last p days to determine the mean model of the distribution and yesterdays variance to predict todays variance. We will in this paper set up two models were the first one is the time series with all data observed 2007-2015, the other with the days when Swedbank report their quarterly results excluded out of the data. Both models will follow the same framework, that is AR(p)-GARCH(1,1) but because we have different databases the number of parameters and estimates for them will differ. The reasoning to exclude report days is that report days are known to the trader and have a expected high volatility and can be assumed to lack characteristic of heteroskedasticity. The efficient market hypothesis says that if new information reaches the market the price of the security react and price this information directly [12](page. 141). We will in this thesis assume that the market on the day for quarterly reports are efficient and will price this new information accordingly.

1.1 Aim of thesis

The aim of this thesis is to estimate VaR for the traded stock "Swedbank.A". We will do this with the AR(p)-GARCH(1,1) model. An attempt to improve the model will be done by excluding the return for the report dates. We will use the same method but with a different data set the parameter estimates will look different. A comparison between the models will be done to determine what model seem to estimate VaR better.

2 Data

The data for this paper is collected from Yahoo Finance. Throughout we use the adjusted close price, this is to adjust for the dividend and splits out of the data. When using one single asset for calculations of VaR, one has to be aware of quarterly reports. Quarterly reports are normally when new information reaches the market, with the assumption of a efficient market on report dates the reports will directly be visible in the data. Big jumps in the stock price are to be expected. We will also make a subset of our data where we exclude the return for report days. The dates for report days are found on <https://www.avanza.se/aktier/om-aktien.html/5241/swedbank-a> where we manually read the charts for "Swedbank.A".

3 Theoretical Background

In this section we will describe the theory behind the methods used to calculate VaR. The theory for this section can be found in [1], in chapters 1,2,3 and 7 unless otherwise is stated.

3.1 Characteristics of a time series

A univariate time series $\{X(t), t \geq 0\}$ is a statistical process where observations are taken on set times, this could be daily, monthly or yearly observations. Time series could be Markov chains, how wind changes over time in a certain location or as in this paper a financial time series. A time series $\{X(t), t \geq 0\}$ is said to be stationary if the series are independent of time. This means that for all (n,s) $X(t_1), X(t_2), \dots, X(t_{n-1}), X(t_n)$ has the same joint distribution as $X(t_1+s), X(t_2+s), \dots, X(t_{n-1}+s), X(t_n+s)$ [8].

Weakly or a second-order stationary time series requires the two first moments to be finite and time invariant i.e. $E[X(t)] = c$, where c is a constant. The covariance $\text{Cov}[X(t), X(t+s)]$ is also independent of time (t) , this means that $\text{Var}(t_n) = \text{Var}(t_{n-a})$, with a is a constant [8]. Normally one assumes that financial time series are weakly stationary. [1].

The time series that we will be working with in this paper is:

$$r_t = \mu_t + a_t Z_t \tag{1}$$

where Z_t is assumed to be a white noise series. A white noise series refers to a time series that is uncorrelated, stationary, that $E[Z_t] = 0$ and has a finite constant variance [8]. The mean μ_t can be modelled with a autoregressive model and the variance t with a generalized autoregressive conditional heteroskedasticity model. We will return to these subjects later in the text.

3.2 Return series

When measuring profit or loss in the financial market, we talk about return. Return unlike the price is a scale free measurement for a movement over time of the asset. Consider a person telling you "the upside of Apple is 1 dollar", the number has no value if not compared to something. If we know that the Apple stock originally was traded at 1 dollar a share that would be a return of 100%, if the share was valued at 100 dollar it would a return of 1%. The return defined as

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}$$

gives the recipient an idea of profit or loss of an investment. Also in being a scale free measurement, the return has better statistical properties and makes it easier to compare different financial assets. The theories and models for financial time series are built around the log return. It is defined as

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(R_t + 1) \quad (2)$$

Where log is the natural logarithm.

3.3 Autocorrelation function (ACF)

Time series may have elements that are correlated to previous observed data. Autocorrelation meaning that the covariance between any r_t and $r_{t-\ell}$ is *not* zero. [9], page 24. In financial time series, autocorrelation is often present. "When the linear dependence between r_t and its past values r_{t-i} is of interest, the concept of correlation is generalized to autocorrelation. The correlation coefficient between r_t and $r_{t-\ell}$ is called the lag- ℓ autocorrelation of r_t and is commonly denoted by ρ_ℓ , which under the weak stationarity assumption is a function of ℓ only." [1](page.31). The correlation between two elements are given by:

$$\rho_\ell = \frac{\text{Cov}(r_t, r_{t-\ell})}{\sqrt{\text{Var}(r_t)\text{Var}(r_{t-\ell})}} \quad (3)$$

With the assumption of second order stationarity we have that the denominator becomes $\sqrt{\text{Var}(r_t)\text{Var}(r_{t-\ell})} = \text{Var}(r_t)$. Since $\text{Var}(r_t) > 0$ for all t we can confirm that iff $\text{Cov}(r_t, r_{t-\ell})$ is equal to 0, then $\rho_\ell = 0$. The empirical estimation of Equation (3) is:

$$\hat{\rho}_\ell = \frac{\sum_{t=\ell+1}^T (r_t - \bar{r})(r_{t-\ell} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2} \quad (4)$$

$\hat{\rho}_1, \hat{\rho}_2, \dots$ is called the sample autocorrelation or autocorrelation function (ACF) of our return series.

A Ljung-Box test will be used to check for autocorrelation within the sample, the statistic is given by:

$$Q(m) = T(T+2) \sum_{\ell=1}^m \frac{\hat{\rho}_\ell^2}{T-\ell} \quad (5)$$

T being the endpoint or length of r_t . The test has the null-hypothesis H_0 : "no autocorrelation" and the alternative H_1 : "autocorrelation". $Q(m)$ is under the null-hypothesis χ^2_α distributed with m degrees of freedom. If the p-value obtained from the test is strictly larger than 0.05 we cannot reject the null-hypothesis. The choice of m differs in various literature. In this paper we will use $m=\ln(\# \text{ of observations})$. If we have autocorrelation the time series are not uncorrelated.

3.4 Autoregressive models

If autocorrelation is found in an data material, a autoregressive model (AR-model) can be used to model the mean, μ_t in our time series model, Equation (1). The AR(p) model uses the prior $\{r_{t-p}, r_{t-p+1}, \dots, r_{t-1}\}$ observations to predict r_t . The AR model of order p takes the form of:

$$r_t = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + a_t \quad (6)$$

If the AR model by itself describes our data then $\{a_t\}$ is assumed to e white noise series. To estimate the parameters in an AR-model the maximum likelihood method will be used, see Appendix 8.3.

3.5 Heteroskedasticity

Financial assets do normally not have a constant variance. This variance within the variance is dependent of time, we say that one day of high volatility is often followed by another day of big movements in the market. Shocks can be triggered by financial news, natural disasters or political actions. The phenomenon is called heteroskedasticity. When heteroskedasticity is present the variance today is conditionalized by the past observed variance.

$$Var(X_t | X_{t-1}, X_{t-2}, \dots)$$

If heteroskedasticity is found in a data-material a GARCH- model can be applied to predict $Var(X_{t+1} | X_t, X_{t-1}, \dots)$.

3.6 ARCH and GARCH models

3.6.1 ARCH

The ARCH model was firstly introduced by Engle in 1980 [6] to describe the variance within the British inflation. An ARCH model of order m is given by:

$$a_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 \quad (7)$$

Where Z_t is a white noise series. To estimate the parameters in an ARCH model we use the conditional maximum likelihood estimation (MLE). The maximum likelihood uses the distribution for the innovations (a_t) to determine the parameters. For a more detailed version on the conditional maximum likelihood estimation see Appendix 8.3.

3.6.2 GARCH

The Generalized AutoRegressive Conditional Heteroskedasticity model (GARCH) was introduced in 1986 by T. Bollersky [7]. The model uses the past so called innovations $a_t = r_t - \mu_t = \sigma_t Z_t$ and the past $\sigma_i^2, \{i = 1, \dots, t-1\}$ to determine σ_t^2 , Z_t is assumed to be a white noise series.

$$a_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1} \beta_j \sigma_{t-j}^2. \quad (8)$$

There are different methods described in the literature on how to estimate a GARCH model. This paper will be using a maximum likelihood estimation to determine the coefficients α_i and β_i of the GARCH model. Since a maximum likelihood method will be used the distribution of the innovations will determine how the parameters will be estimated, see Appendix 8.3. A comparison of a GARCH(m,s) and ARCH(m) reveals that Equation (7) is a GARCH(m,0) model. The restrictions of a GARCH model is that $\sum_{i=1}^{max(m,s)} \alpha_i + \beta_i < 1$ and that $\alpha_0 > 0, \alpha_i, \beta_i \geq 0$. The last restraints to make sure that the unconditional variance

$$E(a_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i)} \quad (9)$$

has $E[a_t^2] < \infty$. Equation (9) is also called the long term volatility. It can be proven that when forecasting a GARCH model, the forecast N times ahead of today is Equation (9), $N \rightarrow \infty$.

Once a model is fitted to our data we want to check if it is adequate. To do this we introduce the standardized residuals \tilde{a}_t :

$$\tilde{a}_t = \frac{a_t}{\sigma_t}$$

Where a_t and σ_t comes from Equation (8). If the model catches the autoregressive and heteroskedasticity of the data, $\{\tilde{a}_t\}$ will be an white noise series.

3.7 Akaike information criterion (AIC)

Akaike information criterion or AIC uses the maximum likelihood function to evaluate a model for a time series. The AIC is defined as:

$$AIC = -\ln(L(\hat{\theta})) + 2p \quad (10)$$

$\ln(L(\hat{\theta}))$ is the likelihood equation for the model under consideration and p is the number of parameters used. We will use the AIC when comparing models. The lower the AIC the better the model will fit [4].

3.8 Value at Risk (VaR)

VaR is a measurement for the risk a financial position holds. VaR uses the underlying distribution of the observed data to create an estimate for how big this risk is. [5][Page.106]. We may view VaR as the upper tail (p) behaviour for the distribution of our loss series $-r_t$.

$$VaR_p(r_t) = \inf\{F(r_t) \geq p\} = F_{r_t}^{-1}(q), \quad q = 1 - p \quad (11)$$

A VaR of 95% should not be exceeded in 95% of trading days. That is, $VaR < \text{observed outcome}$ in 95% of cases.

An illustrative example of VaR can be found in Figure 1. Figure 1 shows a normal distribution. We say that a violation of VaR is every observed data point to the left of the red line. It is important to remember that VaR measures the maximum loss that will be exceeded $p\%$ of the time. If we have a VaR of 1000\$ on the 5% level we say that in 5% of trading days the loss will be greater than 1000\$ **but** not with how much. The tail behaviour to the left of red line in Figure 1 can be estimated but is outside the scope of this thesis.

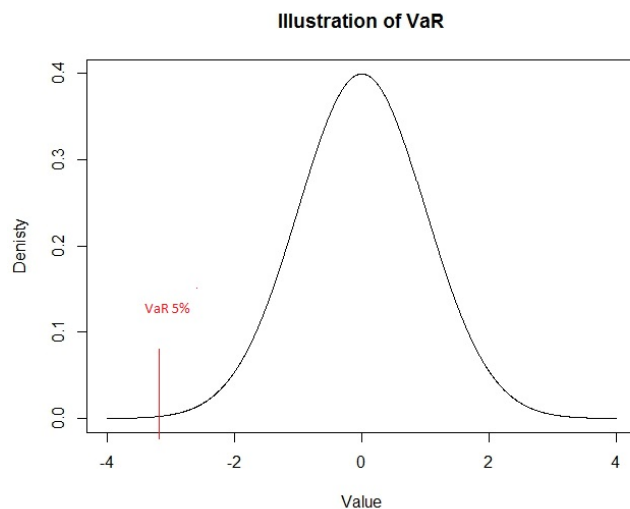


Figure 1: An illustration of VaR. An observation to the left of the red line is a violation of VaR

4 Models and Methods

In this chapter we will analyse our data and apply models to describe it. We will test the assumption of normality, autocorrelation and any ARCH-effects. If data contains only autocorrelation but not heteroskedasticity we will apply a $AR(p)$ model to our data. If the data contains only heteroskedasticity but no autocorrelation we will chose to model data after a $GARCH(1,1)$ model. If it contains both autocorrelation and heteroskedasticity a $AR(p)$ - $GARCH(1,1)$ model will be chosen. To build a $AR(p)$ - $GARCH(1,1)$ model 5 steps will be taken. They are summarized below.

1. First we want to check our data for autocorrelation. If autocorrelation is present we implement an adequate $AR(p)$ model for our mean.
2. Test the residuals of our $AR(p)$ model for any ARCH-effects.
3. If heteroskedasticity is found in the residuals we do a joint estimation for an $AR(p)$ - $GARCH(1,1)$ model.
4. With the $AR(p)$ - $GARCH(1,1)$ model we drop the variables not being significant. We will use the normal 5% significance level.
5. With a finished AR - $GARCH$ model we backtest the model to see if it is adequate.

4.1 Descriptive analysis

First we look at our data. Figure 1 show the adjusted closing price for "Swedbank A" from 2007-2016. The crash that occurred 2008-2009 is clearly visible in the data. The ease of monetary policy and the other factors that lead to the recovery 2010-2012 are also visible before the Greek crisis hit 2012. The period 2012-2016 is seen to the right in Figure 2 and is a stable period for the stock market.

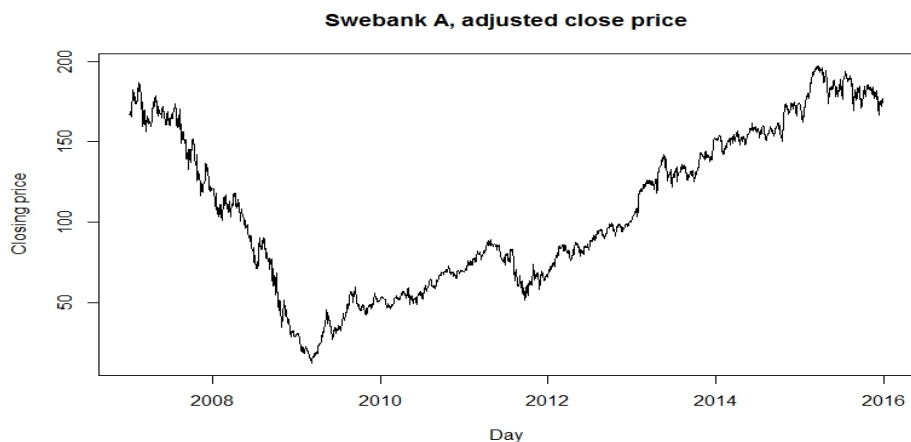


Figure 2: *Adjusted close price for the stock Swedbank A.*

Figure 3 shows the log return for Swedbank.A. The red dots indicating report days. We observe that $|r_t|$ for a number of the report days are large compared to its surroundings. From the plot it seems like the a report does not trigger a longer shock i.e. the characteristic of heteroskedasticity is not apparent. If this is to be the case we would see no difference in number of violation of VaR between a set containing the reports and a set without. The set containing reports would simply treat the reports as shocks and shift its curve, leading to a overestimation of VaR. This overestimation may be costly for companies, it implies extra holdings of capital to counter the risk in the market. With this in mind we will create a subset for our data where we exclude the report days. For this data we will use the same procedure as for the original and see if any differences are to be seen when backtesting VaR.

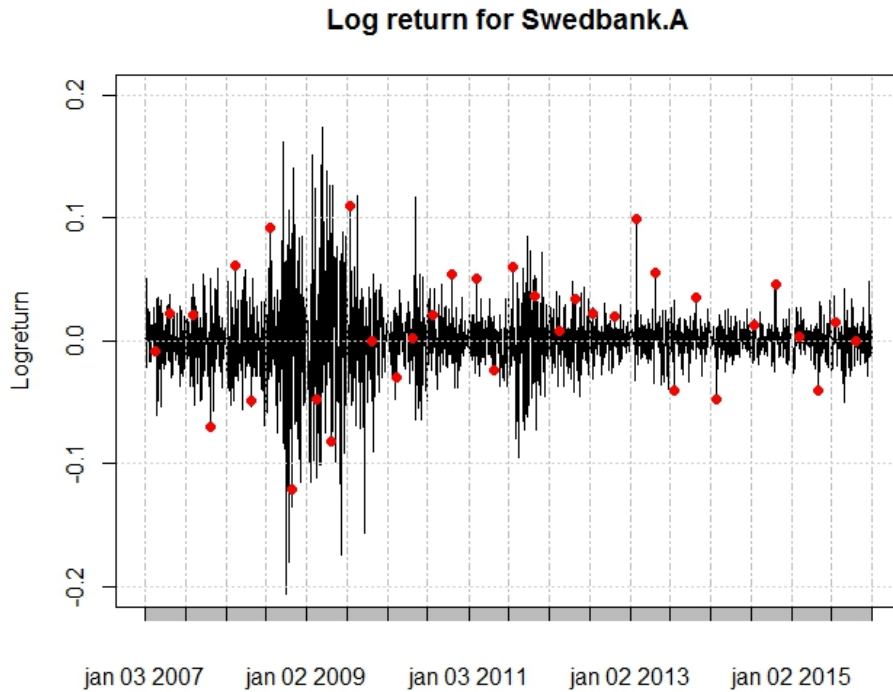


Figure 3: *The log returns for Swedbank A. Each line representing the log return for one day. Red dots indicating report day.*

Next thing we want to check is the assumption of stationary in the loss return series. In figure 3 we see the log return series, we see that the volatility varies with time and does not seem to be constant in the series. With this result we discard the assumption that r_t is a stationary time series. Without further investigation we will assume that the time series are a weakly stationary time series [2].

4.2 Test for normality

The assumption of normality in financial time series is often false. Financial data often inherits fatter tails and has a bigger peak around its mean [1] (p.16). Both these phenomenon can be described with kurtosis. A kurtosis > 3 has a higher peak and heavier tails than the normal distribution. Negative or positive skewness of data can sometimes also be found in data. This means that data are not centred around its mean but rather "tilted" towards either the right or left.

	With report	Without Report
Skewness	-0.218216	-0.2436351
Kurtosis	10.76969	10.70831
P-value skewness-test	0.00002	0.000002
P-value kurtosis test	0	0

Table 1: *Reported point estimates and test if these estimates $\neq 0$*

Table 1 tells us that we have significant skewness and kurtosis for both data sets (that they are similar is to be expected due to the similarity of the series), design of tests may be found in Appendix 8.4. The skewness of -0.21 tells us that data is shifted to the left i.e a heavier tail on the "left" side of the mean. A skewness of -0.21 does not justify the use of a skewed distribution, we would want the value to be larger for that. To further confirm symmetric distribution we look at Figure 4. The two pictures to the right shows two symmetric t-distributions, with apparent skewness we expect alot of points to be on one side of the line. With roughly the same amount of points around the line a symmetric distribution seems to fit the data and will be used.

The kurtosis of the series is high. This means that we have fat tails and high peak for the empiric distribution compared to the normal distribution. A distribution with heavier tails and higher peak than the normal distribution is the t-distribution. In Figure 4 we see how well the empirical distribution for our two data sets fit a normal distribution and a t-distribution. Based on the QQ-plots we can conclude that a t-distribution seem to fit our data better than the normal distribution. There might be another distribution that fits data better. One could consider using the result of extreme value theory. In this thesis we have adopted the t distribution. It seems to fit data good enough based on result found in Figure 4.

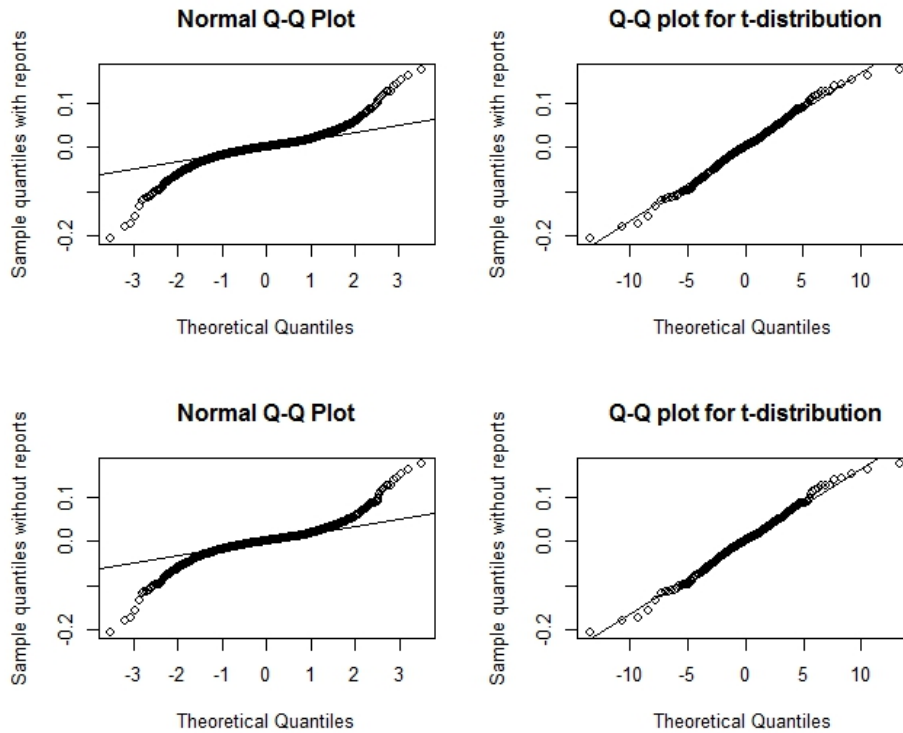


Figure 4: *The first row shows the QQ plots for the data with reports to the left for the normal distribution for a t-distribution with 3 degrees of freedom. The second row QQ plots for data without reports, the left for a normal distribution and the right a t-distribution with 3 degrees of freedom.*

4.3 Test for autocorrelation

Knowing that data may hold patterns for its mean we test for autocorrelation within data. To do this, we firstly plot our AFC. Normally we expect around 5% of observations to be outside of the 95% confidence interval, in Figure 5 seen as dotted lines. Given the fact that more than 5% of observations are outside of the confidence interval for both data sets hints that further investigation might be needed, we proceed using a Ljung-Box test. In our Ljung-Box test we will use $\ln(\# \text{ number of observations})$ as our m . The return series with reports has 2254 observations, the series without has 2219 entries, this gives a m of 7.7 for both Ljung-Box tests. The first Ljung-Box test that include reports gave a p-value of 0.04292, this is < 0.05 . We may say that autocorrelation is present on the 5% significance level, even if barely. For the return series without reports the p-value obtained is 0.0037, we can say that we have autocorrelation on the 1% significance level. With this result we proceed to apply a AR-model to our data.

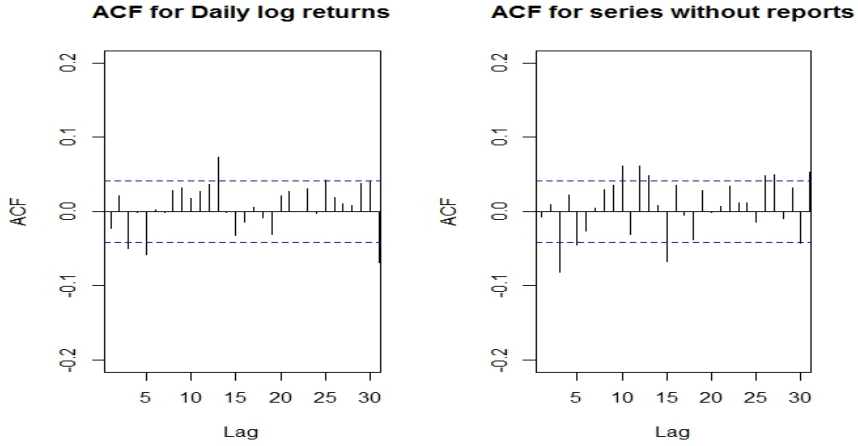


Figure 5: ACF-function for r_t

4.3.1 Applying an AR-model

To determine the order of our AR(p) model we use the AIC. This method uses a predetermined real value integer k and then builds a model for each AR(i), $i=1,,k$. This integer k can be obtained by the PACF, see Appendix 8.2. The AR model that minimizes the AIC will then be used to test for an ARCH-effect.

When a model is fitted with the method described above we want to test if all the ϕ_i in our model are significant. We do this with a t-ratio, the t-ratio is defined as.

$$t\text{-ratio} = \frac{\hat{\rho}_\ell}{\sqrt{1 + 2 \sum_{i=1}^{\ell-1} \hat{\rho}_i^2 / T}} \quad (12)$$

Where T is the number of observations in our data. In the test we have that $H_0 : \rho_\ell = 0$ and $H_1 : \rho_\ell \neq 0$. The decision rule is to reject H_0 if $|t \text{ ratio}| > Z_{\alpha/2}$, Z being the standard normal distribution. [1].

The PACF plot found in Figure 6 shows multiple observations outside of the dotted confidence limit. The choice of k is arbitrary and could here be made for several lags. Our choice fell on $k=13$ for the data with report dates and at lag-15 for the series without report dates. The reasoning behind this is that if the AR model does not hold any ARCH-effect this will be the final model. If the AR model has an ARCH-effect we will proceed by doing a joint estimation of an AR-GARCH model and the order of the AR model will be re-evaluated.

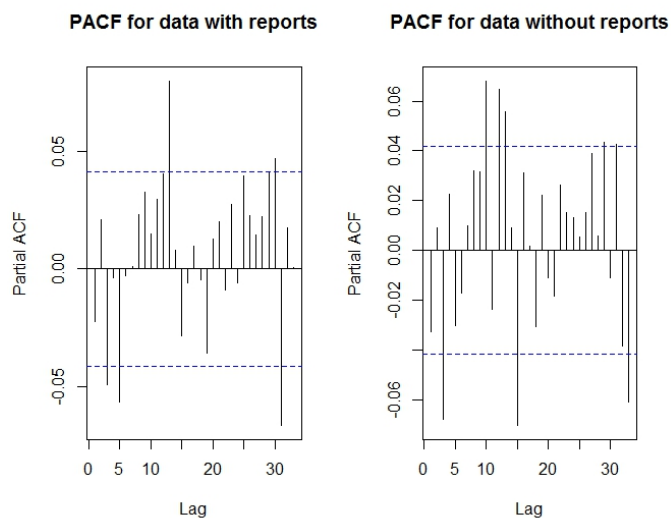


Figure 6: To the left, the PACF for data with reports. To the right the PACF for data with report dates excluded.

	With report	Without report
Order based on AIC	5	12
AIC for full model	-9746.9	-9615.72
σ^2 for full model	0.0007706	0.0007587
Significant parameters	ϕ_3, ϕ_5	$\phi_3, \phi_{10}, \phi_{12}$
AIC*	-9753.15	-9697.12
σ^{2*}	0.0007712	0.0007623
Ljung-Box p-value on residuals	0.9826	0.8553

Table 2: Reported values for the choice of AR model. * indicates the model were insignificant parameters are removed.

With the number k determined the next step is to calculate the AIC for all $\{AR(1), AR(p), \dots, AR(k)\}$. The $AR(i)$, $i=1, \dots, k$ with the lowest AIC will be used as the full model. Calculating the AIC gave the lowest AIC for the $AR(5)$ for the data with reports and $AR(12)$ for the data without. The significant parameters for the $AR(5)$ model were ϕ_3, ϕ_5 and ϕ_3, ϕ_{10} and ϕ_{12} for the series without report dates. The results with reported AIC is found in Table 2.

4.4 Test for heteroskedasticity in residuals

With our selection of AR models in previous Section 4.3.1 we now seek to test the residuals of these models for heteroskedasticity. To do this we take the residuals of the AR model and square them i.e. taking a_t^2 from Equation (6). The ACF plots can be found in the Appendix 9. The plots shows that all observations are outside of the confidence limits. A Ljung-Box test is performed to verify the existence of heteroskedasticity. The p-value for both models are very close to zero (2.2×10^{-16}), we can on any significance level conclude that heteroskedasticity needs to be addressed. With both autocorrelation and heteroskedasticity in the residuals of the AR model we proceed by doing an joint estimation of a AR(p)-GARCH(1,1) model.

4.4.1 Applying an AR-GARCH model to data

With heteroskedasticity found in the residuals of the AR-models we continue to do a joint estimating of an AR(p)-GARCH(1,1) for both our data sets. To estimate the model will start by setting up an AR(5)-GARCH(1,1) and an AR(12)-GARCH(1,1) for the series with and without reports. We use these numbers as a consequence of Section 4.3.1. We will use a backwards elimination on the mean model when estimated in conjunction with the GARCH model. The backwards elimination means that we first estimate all ϕ_i in both models, the parameter that is most insignificant to the model will be removed. When the parameter is removed a new joint estimation of the AR-GARCH model will be done, again removing the parameter with the highest p-value. When all parameters in the mean model are significant on the 5% level we are content with the models. The two models obtained by the backwards elimination can be found in the table below.

Parameter	With report	Without report
μ	0.00774 (0.00307)*	0.000665(0.000292)*
ϕ_1	NA	-0.049466 (0.021206)*
ϕ_5	-0.059(0.020709)**	-0.055905(0.021234)**
α_0	0.000001924 (0.00000091)*	0.000001948 (0.00000082)*
α_1	0.063253 (0.011193)***	0.066982 (0.013136)***
β_1	0.935745 (0.009348)***	0.930854(0.011186) ***
AIC	-4.9931	-5.0568
p-value Ljung-Box test for \tilde{a}_t	0.9711	0.2228
p-value Ljung-Box test for \tilde{a}_t^2	0.2561	0.5256

Table 3: Parameter estimates for both models. The "*" in the tables indicate the significance level, "*" indicating 5% level, "***" 1% and "****" 0.1%.

In both models the choice fell on a AR(5) model with the difference that the data without reports had a lag-1 factor in in. The Ljung-Box test on the standardized residuals together with the graph found in Appendix 8.5.1 reveals that \tilde{a}_t and \tilde{a}_t^2 for both models behave like white noise series. QQ-plots for \tilde{a}_t can be found in Appendix under Figure 10 and Figure 12 giving more confirmation about the standardized residuals being an iid series. The difference in AIC is around 1 percent, a difference too small to make any statement about what model fits data better. We are content with both models but can not make any statements about which one seems better for estimating VaR.

4.5 Backtesting

Backtesting will be done with a rolling window. A rolling window as the name implies "rolls" over the data. One window size will be set to 504 observations, this is roughly two years if we approximate one year as 252 days (the number of days that the stock exchange normally is open in a year). We chose this size due to the length of the crisis (approximately two years) and the two years of relative stability after the crisis but before the Greek crisis 2012. The rolling window will hence use the last 504 days to predict today's volatility. We will also report the result of the backtesting when setting the window size equal to 700 and 1000. This is done to see if any differences occur. A violation occurs when $VaR < r_t$, were r_t is the loss series.

To test if we have the expected number of violations we will use Kupiec unconditional coverage test, it measures the proportions of failures (POF). The test is given by [11]

$$POF = 2 \log \left(\left(\frac{1 - \hat{p}}{1 - p} \right)^{T - I(p)} \left(\frac{\hat{p}^{I(p)}}{p} \right) \right)$$

With:

$$I(p) = \begin{cases} 1, & \text{if } VaR < r_t \\ 0, & \text{otherwise} \end{cases}$$

$$\hat{p} = \frac{1}{T} I(p)$$

Where r_t is the loss function and T is the total number to days we are backtesting. The test has the null hypothesis H_0 : "Correct exceedence" and the alternative H_1 : "Incorrect number of exceedences".

If $p=0.05$ in Equation (11) and the observed outcome is 5% violations the test will produce a statistic of 1. If on the other hand observed outcome $\neq 5\%$ the statistic is under the null-hypothesis asymptotically χ^2 with one degree of freedom.

4.6 Software

To model our AR-GARCH model we will use the software R, R can be found at <https://www.r-project.org/>. The packages used to make the calculations under "Models and Methods" and "Results" are ("QRM", "rugarch").

5 Result

In this section we will present the results of the backtesting for the two models devolved in the previous section.

The first table of the section reports the backtesting of our two models. The period of testing is the last 2000 days of the sample. The result is presented below.

Window Size=504		
Backtest length=2000	With report	Without Report
Expected exceed, 5% VaR	100	100
Actual exceed, 5% VaR	122 (6.1%)	118 (5.9%)
POF p-value	0.029	0.072
Expected exceed, 1% VaR	20	20
Actual exceed, 1% VaR	16 (0.8%)	20 (1%)
POF p-value	0.35	1

Table 4: *Backtesting with window size=504. For the test p-value, if it is written in **bold text** means that we do not discard the null.*

As seen in the Table 4 the data set with reports performed worse when measuring the 5% VaR and we can reject the null hypothesis of correct exceedance. On the 1% level the AR-GARCH without reports hit the expectation and therefore led to a p-value of one. The AR-GARCH with reports overestimated the risk, but not by that much so that we can reject the null. The graphs for the test can be found in the appendix.

Backtesting was also testing with different sizes of the window. The result is presented in the tables on the next page.

Window size=700		
Backtest length=2000	With report	Without Report
Expected exceed, 5% VaR	100	100
Actual exceed, 5% VaR	118 (6.1%)	114 (5.7%)
POF p-value	0.072	0.16
Expected exceed, 1% VaR	20	20
Actual exceed, 1% VaR	16 (0.8%)	22 (1.1%)
POF p-value	0.35	0.658

Table 5: *Backtesting with window size=700. For the test p-value, if it is written in **bold text** means that we do not discard the null.*

Window size=1000		
Backtest length=2000	With report	Without Report
Expected exceed, 5% VaR	100	100
Actual exceed, 5% VaR	116 (6.1%)	114 (5.7%)
POF p-value	0.109	0.16
Expected exceed, 1% VaR	20	20
Actual exceed, 1% VaR	16 (0.8%)	22 (1.1%)
POF p-value	0.35	0.658

Table 6: *Backtesting with window size=1000. For the test p-value, if it is written in **bold text** means that we do not discard the null.*

We note that letting the window size increase helped increase the accuracy of VaR especially for the AR-GARCH model with reports. We also note that the increase in window size from 700 to 1000 did not change any values for the AR-GARCH without reports.

To visualize differences between the series Figure 7, show the result of backtesting on the 1% for both models. In Figure 8 we have zoomed in on the period late 2011-mid 2014. In Figure 8 one observation is highlighted in blue, this is a report day where the price of the stock "Swedbank.A" gained 10% in one day. The following days we see no extreme observations. The AR-GARCH(1,1) reacts to this report as a shock and the curve shifts outwards.

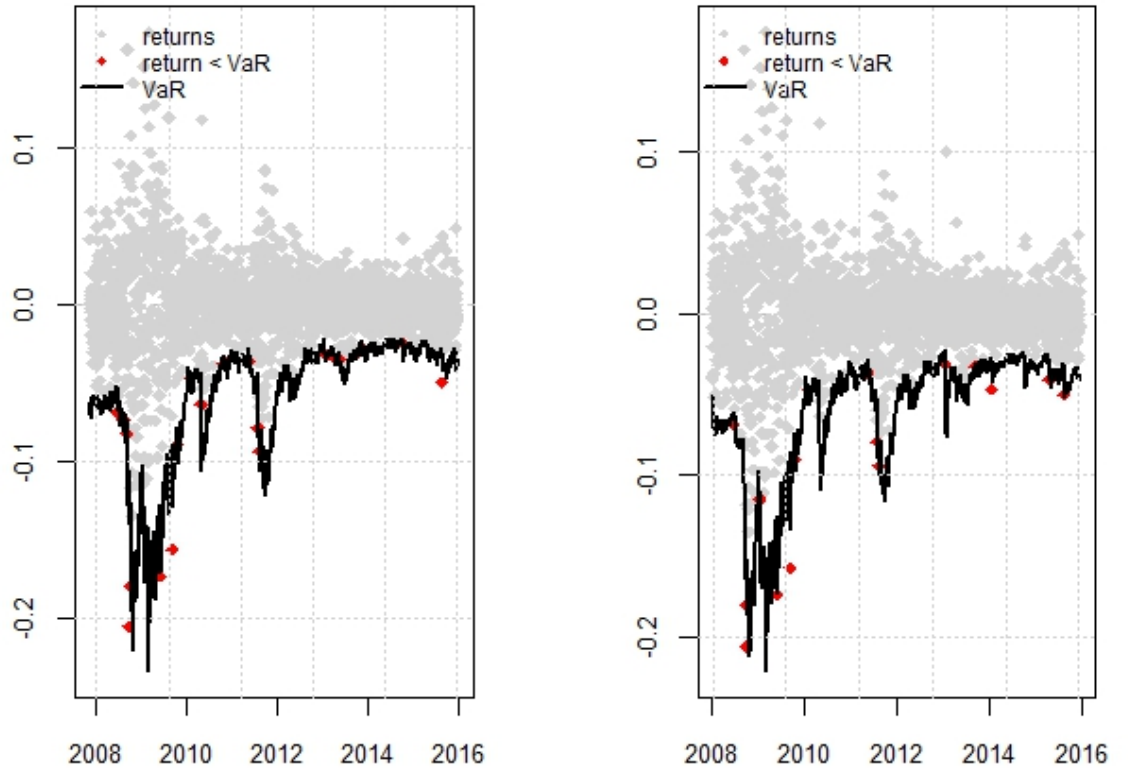


Figure 7: The 1% backtesting, to the left for the model without reports and the right with. Red dots indicating a violation of VaR. The window size is 504.

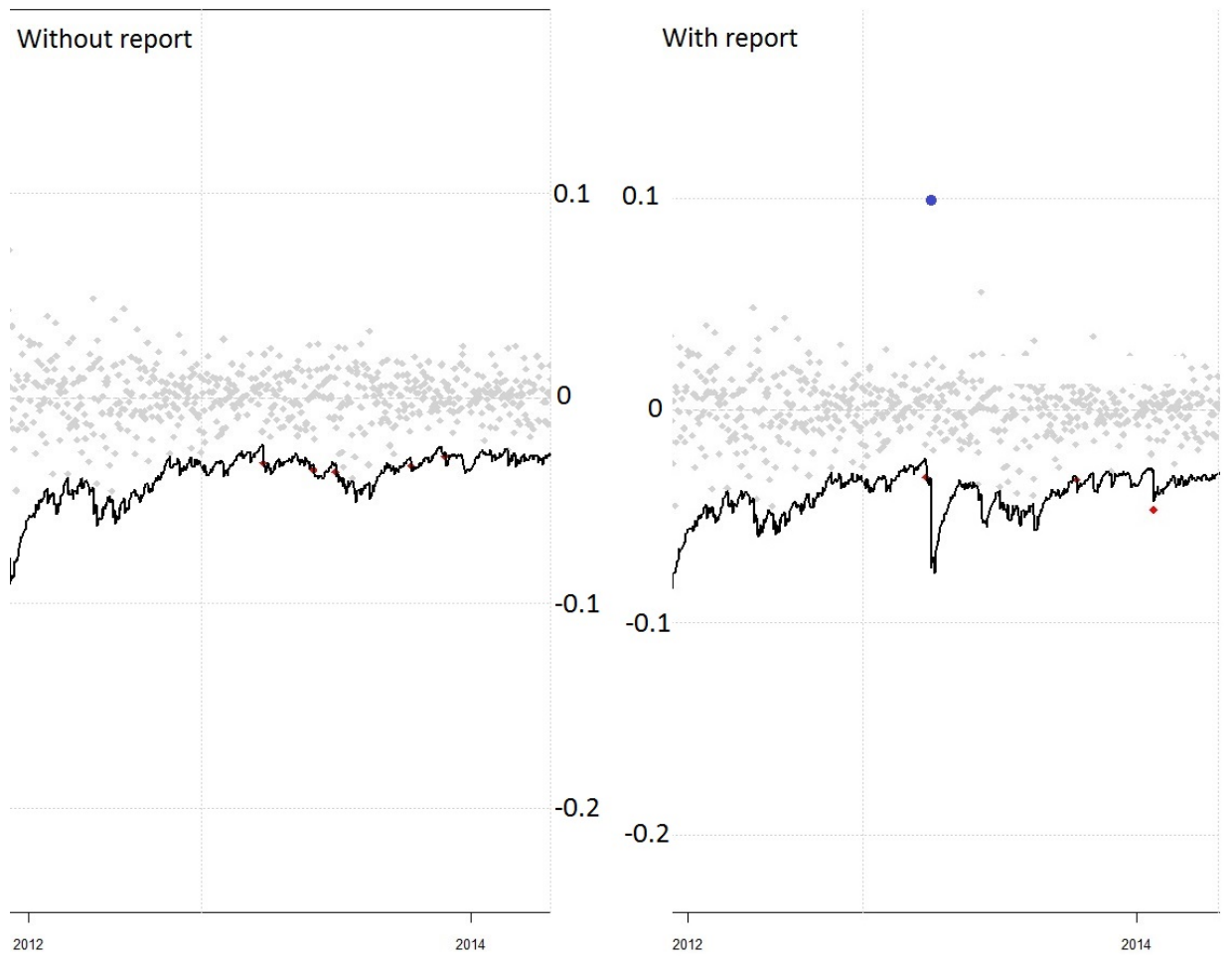


Figure 8: Zoomed in on the period 2012-mid 2014. To the left without reports and the right with reports. The blue dot indicating a highlighted report, red dots being violations of VaR. The window size is 504.

6 Discussion and conclusion

In this thesis we have created two data sets, one set with the all daily returns and one where the dates when Swedbank published a quarterly report excluded out of the return series. The series without reports lay closer to the targeted expected number violations of VaR in the backtesting. The result also reveal that the series with report dates may overestimate the risk after a report. This is a limited study were we have not investigated by how much the data with report overestimated VaR after a report. We have only concluded with that shift outward of the curve happened after a report.

With reports being known for about a year in advance the exclusion of reports make sense. It is easy for a trader to close his or her position if trading with a mathematical program i.e. not trading based on expectations or feeling. It also makes sense out of a cost perspective. With new regulatory treaties for cash requirements a overestimate of VaR forces bigger investors and funds actors to hold more capital, this is costly.

We have not looked at different lengths of the backtesting or at other data than just "Swedbank.A" during 2007-end of 2015. We can **not** conclude that we should generally exclude reports to get a better estimate for VaR. What we can say is that during the period 2007-end of 2015 we should exclude reports for the stock "Swebank.A" if we want to calculate VaR.

Further research could be done for the stock market around report dates, where more data is collected and examined. Alot of studies has been made on post earnings announcement behaviours in finance. Using the knowledge from finance new ways of using the AR-GARCH model when estimating VaR could be developed. One could also measure the distance from the actual returns to the estimated AR-GARCH curve in the periods after a report to see if there is any systematic overestimation of risk after a report.

The main focus of this report has been to model a AR-GARCH model for a rather small data set. The estimates and the backtesting has all been within the same sample. For further research one may consider using a bigger material and dividing it up. With a divided material we can use one part to estimate the models and the other part for backtesting a so called out of sample test.

7 Acknowledgement

I would like to thank my supervisors Joanna Tyrcha, Filip Lindskog and Mathias Lindholm for coming with input, guidance and help for this thesis. A special thanks is given to Joanna Tyrcha who read this thesis multiple times to give direct feedback.

References

- [1] TSAY,R.S. (2010). Analysis of Financial Time Series:Third Edition
- [2] TAYLOR, STEPHEN J. (2007). Modelling Financial Time Series
- [3] ARTZNER,P. DELBAEN,F. EBER,J-M. HEATH,D.(1998). Coherent measure of risk. <https://people.math.ethz.ch/~delbaen/ftp/preprints/CoherentMF.pdf>
- [4] AHO,K. DERRYBERRY,DW. PETERSON,T. (2014). Model selection for ecologists: the worldviews of AIC and BIC. <http://onlinelibrary.wiley.com/doi/10.1890/13-1452.1/full>
- [5] JORION,P (2006). Value at risk
- [6] ENGLE,R.(1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Journal of Econometrics*, 31:307-326, 198 https://www.jstor.org/stable/1912773?seq=1#page_scan_tab_contents
- [7] BOLLERSLEV.T.(1986). Generalized Autoregressive Conditional Heteroskedasticity.*Journal of Econometrics* 31:307-327. http://public.econ.duke.edu/~boller/Published_Papers/joe_86.pdf
- [8] ROSS.S. (2014). Introduction to probability models 11th edition. Chapter 10
- [9] ANDERSSON,P.TYRCHA,J.(2012). Notes in Econometrics.
- [10] REIDER,R. (2009). Volatility Forecasting I: GARCH Models http://cims.nyu.edu/~almgren/timeseries/Vol_Forecast1.pdf
- [11] CAMPBELL, S.D (2005). A Review of Backtesting and Backtesting Procedures. <https://www.federalreserve.gov/pubs/feds/2005/200521/200521pap.pdf>
- [12] MISHKIN, F.S. MATTHEWS, K. GILUIORDORI. M. (2013). The Economics of Money, Banking & Financial Markets.

8 Appendix

8.1 Distributions

Normal distribution:

$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, iff $\mu = 0$ and $\sigma^2 = 1$ we have a standard normal distribution.

T-distribution:

$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}(1+\frac{t^2}{\nu})^{-\frac{\nu+1}{2}}$ where Γ is a gamma function i.e $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx$, and ν is the degrees of freedom. It can be shown that when $\nu \rightarrow \infty$ the t-distribution converges to a standard normal distribution.

χ^2 distribution

$f_r(x) = \frac{x^{(r-1)/2}e^{-x/2}}{\Gamma(\frac{r}{2})2^{r/2}}$ where Γ is the gamma function and r is the degrees of freedom. It can be shown that $\chi_r^2 = \sum_{i=1}^r Y_i^2$ where $\{Y_1, \dots, Y_r\}$ are independent standard normal distributions.

8.2 Partial Autocorrelation Function

The Partial Autocorrelation Function (PACF) can be used to determine the order of an AR model. The PACF can be described by:

$$\begin{aligned}r_t &= \phi_{0,1} + \phi_{1,1}r_{t-1} + a_t \\r_t &= \phi_{0,2} + \phi_{1,2} + \phi_{2,2}r_{t-2} + a_t \\r_t &= \phi_{0,3} + \phi_{1,3} + \phi_{2,3} + \phi_{3,3}r_{t-3} + a_t \\&\dots \\r_t &= \sum_{i=0}^{n-1} \phi_{i,n} + \phi_{n,n}r_{t-n} + a_t\end{aligned}$$

Where n can be any positive integer. The parameter $\phi_{i,i}$ or the lag- i PACF describes the added contribution of r_{t-i} to r_t [1]. With this definition the order p of an AR(p) model can be determined. If plotted the PACF plot shows $\phi_{i,i}$. Any violation of the given confidence interval may be used as our order p .

8.3 Maximum likelihood estimation

Let $\{x_1, x_2, \dots, x_n\}$ be realizations of the random variables $\{X_1, X_2, \dots, X_n\}$, where X_i are all iid. The maximum likelihood estimation method is the

estimation that maximizes the joint probability function for $\{X_1, \dots, X_N\}$:

$$L = \prod_{i=1}^n f_{X_i}(x_i)$$

To ease the calculation of the L the transformation into log-likelihood is made:

$$l = \log(L) = \sum_{i=1}^n f_{X_i}(x_i)$$

When estimating a GARCH model the returns are not independent of each other so the joint probability function is *not* $\prod_{i=1}^n f_{X_1, \dots, X_n}(x_1, \dots, x_n)$. The density function of X_t is conditional on $X_{t-1}, X_{t-2}, \dots, X_1$ hence we can write the joint density function as.

$$f_{X_t, \dots, X_1}(x_t, \dots, x_1) = f_{X_t|X_{t-1}, \dots, X_1}(x_t | x_{t-1}, \dots, x_1)$$

Again X_{t-1} is conditional on X_{t-2}, \dots, X_1 so that:

$$f_{X_t, \dots, X_1}(x_t, \dots, x_1) = f_{X_t|X_{t-1}, \dots, X_1}(x_t | x_{t-1}, \dots, x_1) f_{X_{t-1}|X_{t-2}, \dots, X_1}(x_{t-1} | x_{t-2}, \dots, x_1)$$

Iterate this t times will yield the joint probability density function for f_{X_t, \dots, X_1} . When modelling for a GARCH model we are interested in the variance. For the normal distribution the estimation would become:

$$L = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(x_t - \mu)^2}{2\sigma_t^2}\right) \frac{1}{\sqrt{2\pi\sigma_{t-1}^2}} \exp\left(-\frac{(x_{t-1} - \mu)^2}{2\sigma_{t-1}^2}\right) \dots \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x_1 - \mu)^2}{2\sigma_1^2}\right)$$

With a rather long but simple calculation the log likelihood becomes:

$$l = \frac{-t}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^t \ln(\sigma_i^2) - \frac{1}{2} \sum_{i=1}^t \left(\frac{(x_i - \mu)^2}{\sigma_i^2}\right)$$

Replacing $\sigma_i^2 = \alpha_0 + \alpha_1 a_{i-1}^2 + \beta_1 \sigma_{i-1}^2$ gives the log likelihood for a GARCH(1,1), given that the innovation are normally distributed. The theory for this section can be found in [10], page 7.

8.4 Central moments

In this text we have been working with the first four central moments for random variables. These are the mean, the variance, the skewness and the kurtosis. When working with central moments, we do not regard data as observations in a time series but rather as realizations of a random variable. This implicates that the order, that is characteristic for a time series are no longer relevant. Let R be a random variable and $\{r_1, \dots, r_T\}$ be its T realizations. To underline the fact that we in this section see the observations as

independent of time we denote the realizations as r_i instead of r_t . The first moment, also known as the sample mean is defined as:

$$\hat{\mu}_r = \frac{1}{T} \sum_{i=1}^T r_i \quad (13)$$

The sample variance is defined as:

$$\hat{\sigma}_r^2 = \frac{1}{1-T} \sum_{i=1}^T (r_i - \hat{\mu}_r)^2 \quad (14)$$

The third central moment or sample skewness:

$$\hat{S}(r) = \frac{1}{(T-1)\hat{\sigma}_r^3} \sum_{i=1}^T (r_i - \hat{\mu}_r)^3 \quad (15)$$

When working with kurtosis there are two different kurtosis. The fourth central moment or the excess kurtosis. Excess kurtosis compares the sample kurtosis with the kurtosis of the standard normal distribution. In this paper we will only work with the central fourth moment and not the excess variant. The kurtosis is then given by:

$$\hat{K}(r) = \frac{1}{(T-1)\hat{\sigma}_r^4} \sum_{i=1}^T (r_i - \hat{\mu}_r)^4 \quad (16)$$

If kurtosis > 3 we say that the distribution is leptokurtic and has a bigger peak around its mean and heavier tails compared to the normal distribution. A test with $H_0: S(r)=0$ $H_1: S(r)\neq 0$ is given by.

$$t = \frac{\hat{S}(r)}{\sqrt{6/T}}$$

The test is under the null a normal standard distribution.

A test for excess kurtosis with the hypotheses $H_0: K(r) - 3 =$, $H_1: K(r) - 3 \neq 0$ is given by:

$$t = \frac{\hat{K}(r) - 3}{\sqrt{24/T}}$$

The statistic is under the null a normal standard distribution. Under the chapter "Models and Methods" a distribution for our data will be fitted.

8.5 Graphs

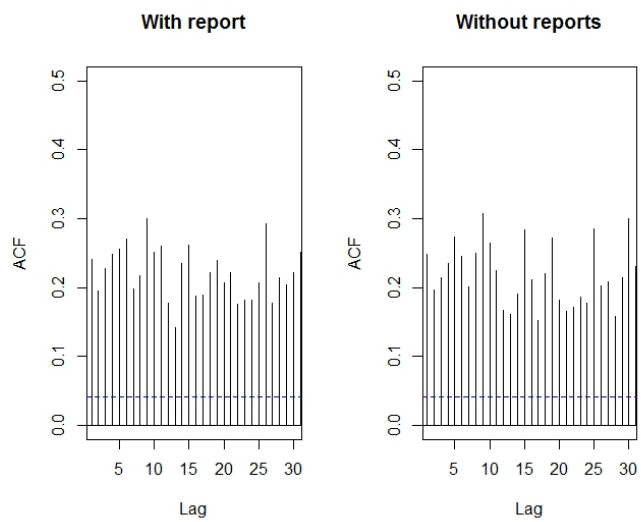


Figure 9: ACF for squared residuals of the AR models in section [4.3.1](#)

8.5.1 Graphs for fitted GARCH models

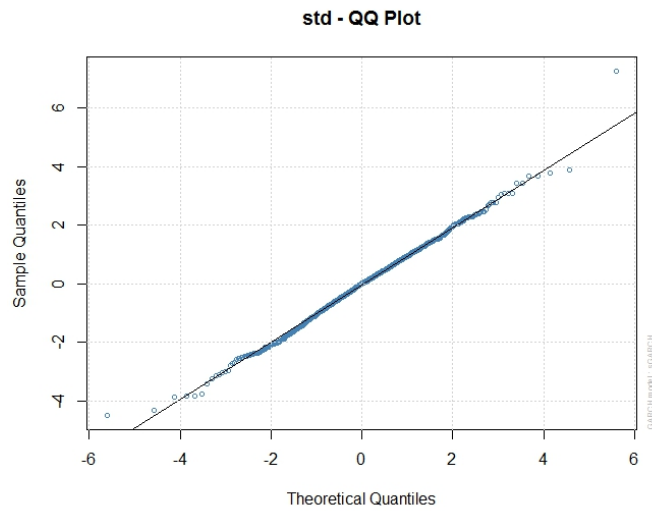


Figure 10: QQ-plot for residuals in the AR-GARCH model for data with reports.

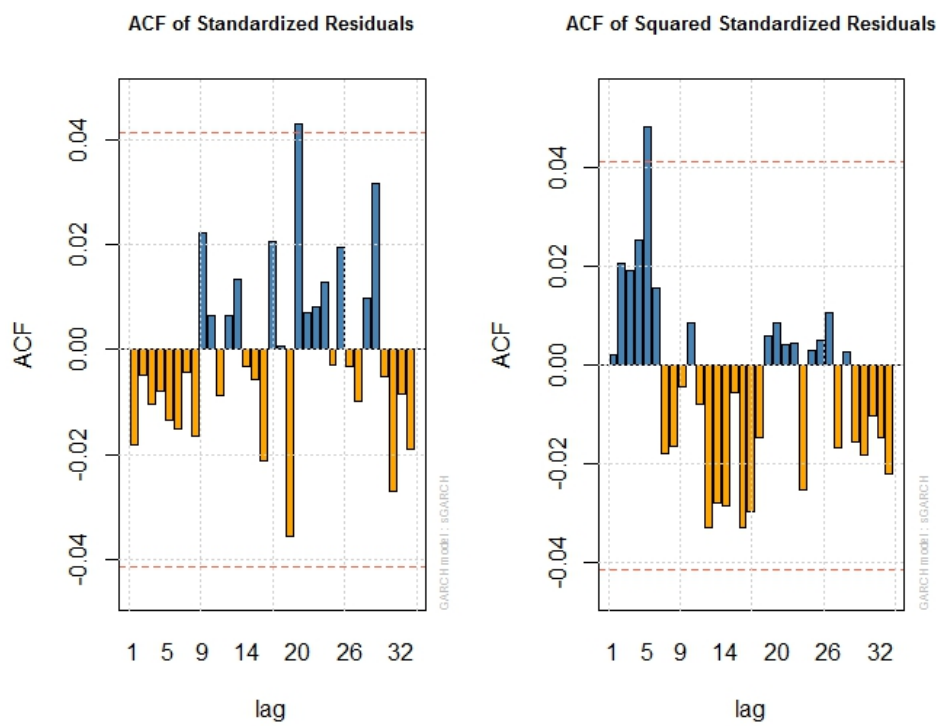


Figure 11: ACF for residuals in AR-GARCH model for data with reports. The picture to the left looking for autocorrelation and the picture to the right looking for a ARCH-effect.

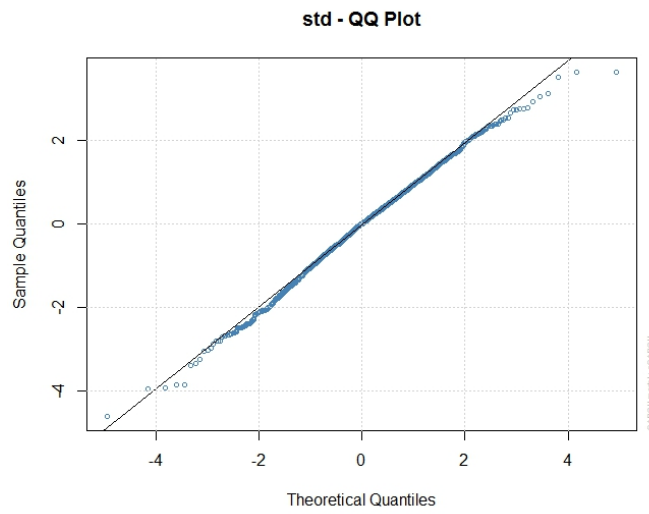


Figure 12: QQ-plot for residuals in the AR-GARCH model for data without reports.

8.5.2 Graphs for backtesting

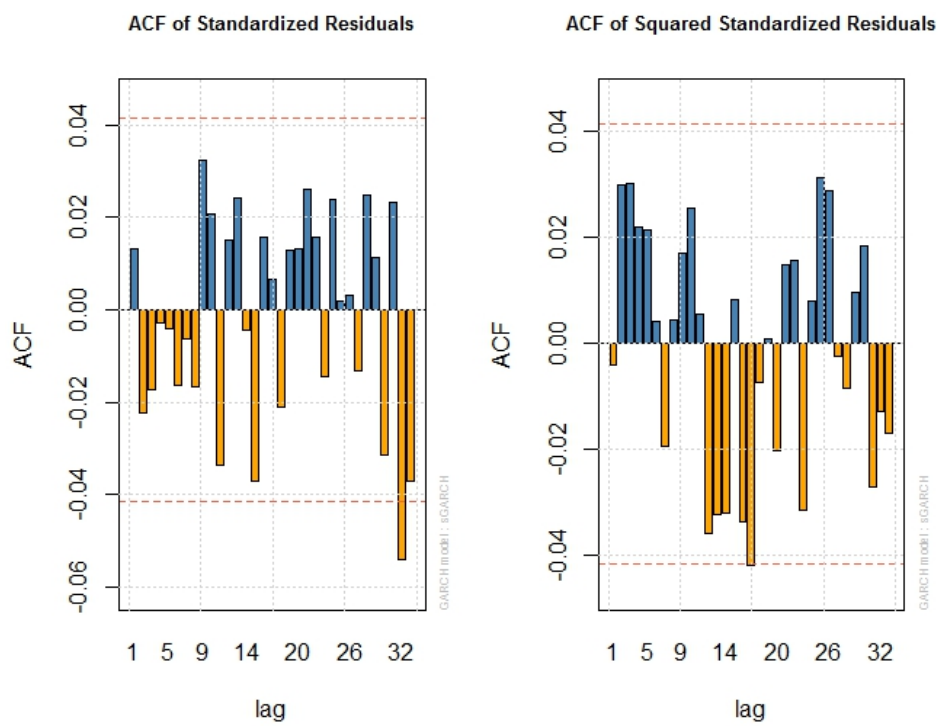


Figure 13: ACF for residuals in AR-GARCH model for data with reports. The picture to the left looking for autocorrelation and the picture to the right looking for a ARCH-effect.

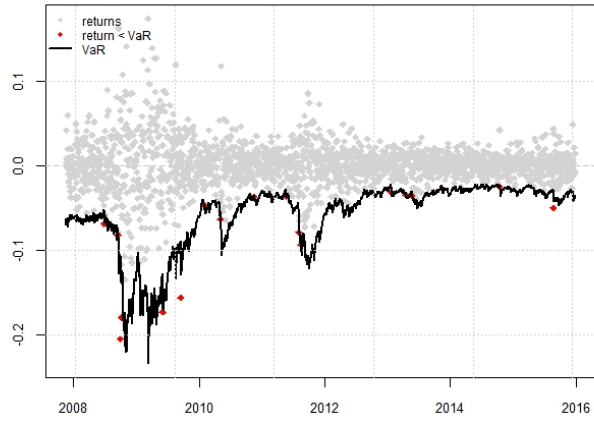


Figure 14: 1% VaR for the AR-GARCH without report dates, red dots indication violation of VaR. The window size is 504.

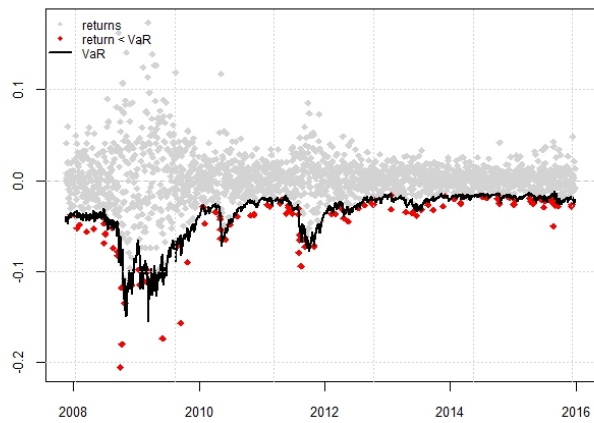


Figure 15: 5% VaR for the AR-GARCH without report dates, red dots indication violation of VaR. The window size is 504.

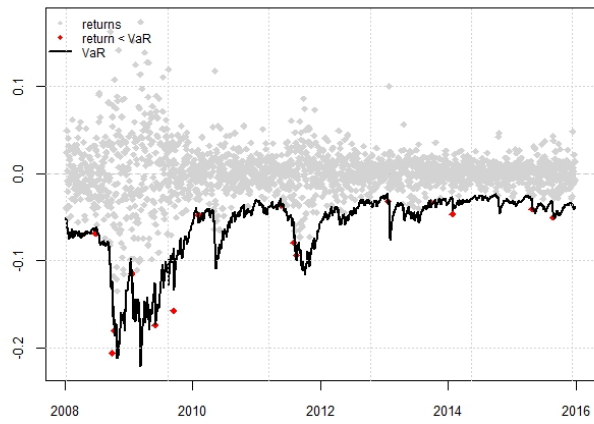


Figure 16: 1% VaR for the AR-GARCH with reports dates, red dots indication violation of VaR. The window size is 504.

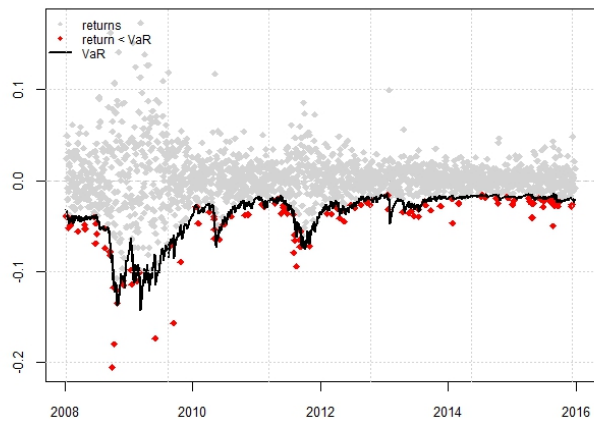


Figure 17: 5% VaR for the AR-GARCH with reports dates, red dots indication violation of VaR. The window size is 504.