Stockholms universitet

A Comparison between ARCH and GARCH(1,1) Models Fitted to Nasdaq Nordic Indices

Joohee Kim

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2019:25
Matematisk statistik
Juni 2019

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2019:25**
http://www.math.su.se

A Comparison Between ARCH and GARCH(1,1) Models Fitted to Nasdaq Nordic Indices

Joohee Kim*

June 2019

Abstract

A financial time series is a set of variables observed at different time points. A major concern of financial time series study is to evaluate the changes in the values of an underlining asset and to forecast losses or gains in the future. One distinct characteristic of a financial time series is that it contains a factor of uncertainty. Volatility is a conditional standard variance for an underlying asset. In empirical study, volatility is difficult to measure, and yet it is a key to building a model for a financial time series. In this study, we will study two most representative volatility models, namely the autoregressive conditional heteroscedastic (ARCH) models, and the general autoregressive conditional heteroscedastic (GARCH) models. These two models will be fitted to Nasdaq Nordic indices, the Large Cap and the Small Cap - two market segment indices with different dynamics. This study has two main purposes. First, we want to see if these two models are adequate to describe the two different data sets even though these data sets have different dynamics. If that is not the case, we want to see which model is more adequate to describe a certain data set. Second, we want to see which model has better forecasting power. To measure forecasting power, we will use backtesting based on value at risk, a risk measure that gives a point estimate for a potential loss. The study results show that the differences between the models are not as distinctive as the differences between the data sets. For a same index, both ARCH models and GARCH models show fairly similar results. Backtesting results, however, show that the most distinctive differences in forecasting power come from distribution assumption for innovation of a model. When it comes to forecasting power, differences between distribution assumption for innovation seem more distinctive than differences between ARCH and GARCH models.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: joki1322@student.su.se. Supervisor: Kristofer Lidensjö and Mathias Lindholm.

Contents

1 Introduction

Financial time series analysis is aimed to observe and forecast value evaluation over time. Return series show gains or losses from an underlying asset over time. Volatility is a conditional standard deviation of an underlining asset return series, and it plays an important role when we build models for financial time series. Many models have been developed to capture the dynamics of volatility. In this study two major volatility models will be studied, namely autoregressive conditional heteroscedastic (ARCH) models, and general ARCH (GARCH) models. These two models will be fitted to two Nasdaq indices, the Large Cap and the Small Cap. We will check if fitted models are adequate to describe each index, by studying their standardized residuals. Most importantly, we will compare the power of forecasting of these two models with backtesting based on value at risk measure.

2 Theory

2.1 Prices and asset returns

Let P_t be the price of an asset at time t. A one-period simple gross return is

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

A one-period simple net return or simple return R_t is

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

A log return r_t is the natural logarithm of the simple gross return, and defined as

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}}.$$

2.2 Autocorrelation

2.2.1 Stationarity and covariance

A time series $\{r_t\}$ is strictly stationary if the joint distribution of $(r_{t_1}, \ldots, r_{t_k})$ is identical even when we make a shift in time, where k is an arbitrary positive integer. A time series $\{r_t\}$ is weakly stationary if the mean of r_t and the covariance between r_t and r_{t-l} do not depend on t but depend only on l, where l is an arbitrary positive integer. The covariance $\gamma_l = \text{Cov}(r_t, r_{t-l})$ is called the lag-l autocovariance of r_t. [1]

2.2.2 Autocorrelation function (ACF)

The correlation coefficient between two variables X and Y is defined as

$$\rho_{x,y} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

When we have a return series r_t, we want to study the linear dependence between r_t and its past values r_{t-l}. The correlation coefficient between r_t and r_{t-l} is called the lag-l autocorrelation of r_t and is denoted by ρ_l. We assume that our return series is stationary. For a stationary return series $\text{Var}(r_t) = \text{Var}(r_{t-l})$. Then the lag-$l$ autocorrelation is defined as

$$\rho_l = \frac{\text{Cov}(r_t, r_{t-l})}{\sqrt{\text{Var}(r_t)\text{Var}(r_{t-l})}} = \frac{\text{Cov}(r_t, r_{t-l})}{\text{Var}(r_t)}.$$

For a sample $\{x_t, y_t\}$, the correlation can be estimated by its sample correlation coefficient

$$\hat{\rho}_{x,y} = \frac{\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{T}(x_t - \bar{x})^2 \sum_{t=1}^{T}(y_t - \bar{y})^2}}$$

where $\bar{x} = \sum_{t=1}^{T} x_t/T$ and $\bar{y} = \sum_{t=1}^{T} y_t/T$ respectively.
The lag-l sample autocorrelation of r_t is defined as

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^{T}(r_t - \bar{r})(r_{t-l} - \bar{r})}{\sum_{t=1}^{T}(r_t - \bar{r})^2}, \qquad 0 \leq l < T-1,$$

where $\bar{r}$ denotes the sample mean for a return series $\{r_t\}_{t=1}^{T}$. [2]

2.2.3 Partial Autocorrelation Function (PACF)

The partial correlation function of a stationary time series is a function of its ACF, and it is useful when determining the order of an AR model. We can consider the AR models as follows.

$$r_t = \phi_{0,1} + \phi_{1,1} r_{t-1} + \epsilon_{1t},$$

[1] Tsay, Chapter 2, p 30.
[2] Tsay, Chapter 2, p 30-31.

$$r_t = \phi_{0,2} + \phi_{1,2} r_{t-1} + \phi_{2,2} r_{t-2} + \epsilon_{2t},$$

$$r_t = \phi_{0,3} + \phi_{1,3} r_{t-1} + \phi_{2,3} r_{t-2} + \phi_{3,3} r_{t-3} + \epsilon_{3t},$$

$$r_t = \phi_{0,4} + \phi_{1,4} r_{t-1} + \phi_{2,4} r_{t-2} + \phi_{3,4} r_{t-3} + \phi_{3,4} r_{t-3} + \epsilon_{4t},$$

where $\phi_{i,j}$ are the coefficients of r_{t-i} and $\{\epsilon_{jt}\}$ are the error term of an AR(j) model. The lag-2 sample PACF $\hat{\phi}_{2,2}$ shows the added contribution of r_{t-2} to r_t over the first equation, the AR(1) model. The lag-3 sample PACF shows the added contribution of r_{t-3} to r_t over an AR(2) model. For an AR(p) model, the lag-p sample PACF is significantly different from zero, while $\hat{\phi}_{j,j}$ are close to zero for all $j > p$. In other words, for an AR(p) series, the sample PACF cuts off at lag p. [3]

2.2.4 Ljung-Box test

The Ljung-Box test is to test jointly that several autocorrelations of r_t are zero. The null hypothesis is $H_0 : \rho_l = \cdots = \rho_m = 0$ and the alternative hypothesis is $\rho_i \neq 0$ for at least one $i \in \{1, \ldots, m\}$. For an iid sequence $\{r_t\}$, the test statistic $Q(m)$ is asymptotically a chi-squared random variable with m degrees of freedom.

The Ljung-Box test statistic is

$$Q(m) = T(T+2) \sum_{l=1}^{m} \frac{\hat{\rho}_l^{\,2}}{T-l}.$$

The decision rule is to reject H_0 if p-value is less than a significance level. [4]

2.3 Simple linear time series models

Simple linear time series models are aimed to capture the linear relationship between r_t and its past values. These models will not be used in this study, but they are the basis for more advanced time series models. So it can be worthwhile to understand these simple linear models before we move on to ARCH and GARCH models.

[3] Tsay, Chapter 2, p 46-47.

[4] Tsay, Chapter 2, p 32-33.

2.3.1 Simple AR models

An autoregressive model of order p, or AR(p), says that the past p values jointly determine the conditional expectation of r_t given the past data.[5] If we assume that $\mu = 0$, a general AR(p) model can be written as

$$r_t = \theta_0 + \theta_1 r_{t-1} + \cdots + \theta_p r_{t-p} + a_t,$$

where a_t is called shock or innovation of an asset return at time t.

2.3.2 Simple MA models

If we assumed that $\mu = 0$, we have an AR model with an infinite order as below.

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \cdots + a_t.$$

This model has infinite many parameters. To make the number of parameters finite, we put some constraints on the coefficients θ_i. An example is

$$r_t = \phi_0 - \theta_1 r_{t-1} - \theta_1^2 r_{t-2} - \theta_1^3 r_{t-3} - \cdots + a_t \tag{1}$$

This model can be rewritten as

$$r_t + \theta_1 r_{t-1} + \theta_1^2 r_{t-2} + \theta_1^3 r_{t-3} + \cdots = \phi_0 + a_t \tag{2}$$

We multiply (2) by θ_1 and then subtract the result from (1). We get

$$r_t = \phi_0(1 - \theta_1) + a_t - \theta_1 a_{t-1} \tag{3}$$

The equation (3) can be rewritten as

$$r_t = c_0 + a_t - \theta_1 a_{t-1} \tag{4}$$

The equation (4) is a structure of an MA(1) model. An MA(q) model is

$$r_t = c_0 + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q},$$

where $q > 0$. [6]

[5] Tsay, Chapter 2, 38
[6] Tsay, Chapter 2, 58

2.3.3 ARMA Models

The AR or MA models sometimes need many parameters to fit a data. The autoregressive moving-average models, or ARMA models, combine AR and MA models into a simpler form, and they usually have less numbers of parameters. An ARMA(1,1) model can be written as

$$r_t - \phi_1 r_{t-1} = \phi_0 + a_t - \theta_1 a_{t-1}.$$

The left side of the equation is the AR part and the right side of the equation is the MA part. ϕ_0 is a constant term.

A general ARMA(p, q) model can be written as

$$r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i},$$

where $\{a_t\}$ is a white noise and p and q are nonnegative integers. [7]

2.4 Volatility

2.4.1 Definition and characteristics

Volatility means the conditional standard deviation of the underlying asset return. Stock volatility is not directly observable because there is only one observation per day when we take, for instance, only starting prices or only closing prices each banking day. Although volatility is not directly observable, it displays some common characteristics. First, there are volatility clusters, which means volatility usually is high for a certain period and low for another period. Figure 2 shows the time plots of the log returns of the Large Cap and the Small Cap. We can find that volatility is high for a period, and it is low for another period. Second, volatility evolves in a continuous manner. It is rare to observe volatility jumps. Third, volatility varies within some fixed range.[8]

2.4.2 Volatility models

The log returns of an asset at time t, r_t, is either serially uncorrelated or shows minor lower order serial correlation. It is, however, a dependent series. Volatility models are aimed to capture such dependency between the values of a series. The model structure is

$$\mu_t = E(r_t \mid F_{t-1}), \qquad \sigma_t^2 = \mathrm{Var}(r_t \mid F_{t-1}) = \mathrm{E}[(r_t - \mu_t)^2 \mid F_{t-1}].$$

[7] Tsay, Chapter 2, p 64.

[8] Tsay, Chapter 3, p 111.

The model for μ_t is called the mean equation and the model for σ_t^2 is the volatility equation for r_t. Here F_{t-1} denotes the information available at time $t-1$. If we assume that r_t follows a stationary ARMA(p,q) model, we can build the model

$$r_t = \mu_t + a_t, \qquad \mu_t = \sum_{i=i}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i}$$

We have a volatility equation as

$$\sigma_t^2 = \mathrm{Var}(r_t \mid F_{t-1}) = \mathrm{Var}(a_t \mid F_{t-1}).$$

The evolution of σ_t^2 is the main concern for volatility models. Different volatility models describe different ways under which σ_t^2 evolves over time.[9]

2.4.3 Model building

There are four steps in building a volatility model for an asset series. First, a mean equation is to be built. Second, the residuals of the mean equations are to be tested for potential ARCH effects. Third, if ARCH effects are statistically significant, specify a volatility model. The mean and volatility equations are to be jointly estimated. Fourth, the model is to be verified and, if necessary, refined. [10]

2.4.4 The ARCH effect

We can denote the residuals of the mean equation as $a_t = r_t - \mu_t$. The series of squared residuals are used to test if conditional heteroscedasticity exists. Such conditional heteroscedasticity is also called ARCH effects. The Ljung-box test will be used to see if there are ARCH effects. The null hypothesis is that the firs m lags of ACF of the a_t^2 series are zero.[11]

2.5 ARCH models

In an ARCH model, the shock or innovation a_t of an asset return is a serially uncorrelated but a dependent series. This dependency of the innovation a_t can be written as a quadratic function of its past values. Assuming that an asset return series has mean zero for all t, an ARCH model can be written as

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2,$$

[9] Tsay, Chapter 3, p 111-113.

[10] Tsay, Chapter 3, p 113

[11] Tsay, Chapter 3, p 114.

where $\{\epsilon_t\}$ is a series of independent and identically distributed random variables with mean zero and variance 1. Usually $\{\epsilon_t\}$ is assumed to follow the standard normal or a standardized Student-t distribution. The coefficients $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$. The model structure shows that the large sum of past squared shocks $\{a_{t-i}^2\}_{i=1}^m$ leads to a large conditional variance σ_t^2 for the shock a_t. [12]

2.5.1 The ARCH(1) model

We can study the ARCH(1) model carefully before we move on to the ARCH models in general. Given that we already know a_{t-1}, the ARCH(1) model is defined as

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2,$$

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$. We can get the mean and the variance of a_t. The mean of a_t is zero because

$$E(a_t) = E[E(a_t|F_{t-1})] = E[\sigma_t E(\epsilon_t)] = 0$$

The variance of a_t is

$$\mathrm{Var}(a_t) = \mathrm{E}(a_t^2) = \mathrm{E}[\mathrm{E}(a_t^2|F_{t-1})] = \mathrm{E}(\alpha_0 + \alpha_1 a_{t-1}^2) = \alpha_0 + \alpha_1 \mathrm{E}(a_{t-1}^2).$$

For the variance of a_t to be positive we need to satisfy the requirement $0 \leq \alpha_1 < 1$.

Because $\{a_t\}$ is a stationary series, we have $\mathrm{E}(a_t) = 0$ and $\mathrm{Var}(a_t) = \mathrm{Var}(a_{t-1}) = \mathrm{E}(a_{t-1}^2)$. Thus, we get $\mathrm{Var}(a_t) = \alpha_0 + \alpha_1 \mathrm{Var}(a_t)$ and $\mathrm{Var}(a_t) = \alpha_0/(1 - \alpha_1)$.

2.5.2 Order determination

If an ARCH effect is significant, the PACF of a_t^2 can be used to determine the order of an ARCH model. We have an ARCH(m) model as below.

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2$$

The equation for σ_t^2 has a structure of an autoregressive model of order m for a_t^2. In that sense, a_t^2 is linearly related to its past values in this model, and that is why PACF of a_t^2 can be used for order determination. [13]

[12] Tsay, Chapter 3, p 116.
[13] Tsay, Chapter 3, p 119-120.

2.5.3 Estimation under normality assumption

If ϵ_t is assumed to be normally distributed, the likelihood function of an ARCH(m) is

$$
\begin{aligned}
f(a_1, \ldots, a_T|\boldsymbol{\alpha}) &= f(a_T|F_{T-1})f(a_{T-1}|F_{T-2})\ldots f(a_{m+1}|F_m)f(a_1, \ldots, a_m|\boldsymbol{\alpha}) \\
&= \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{a_t^2}{2\sigma_t^2}\right) \times f(a_1, \ldots, a_m|\boldsymbol{\alpha}),
\end{aligned}
$$

where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_m)'$, and $f(a_1, \ldots, a_m|\boldsymbol{\alpha})$ is the joint probability density function of $a_1, \ldots, a_m$. We can write $A_m = (a_1, \ldots, a_m)$. When the sample is large, we can use the conditional-likelihood function.

$$
f(a_{m+1}, \ldots, a_T|\boldsymbol{\alpha}, A_m) = \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{a_t^2}{2\sigma_t^2}\right)
$$

Then we want to maximize the conditional likelihood function. For this purpose we can maximize its logarithm because logarithm is easier to handle. The conditional log-likelihood function is

$$
l(a_{m+1}, \ldots, a_T|\boldsymbol{\alpha}, A_m) = \sum_{t=m+1}^{T} \left[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma_t^2) - \frac{1}{2}\frac{a_t^2}{\sigma_t^2}\right]
$$

The term $\ln(2\pi)$ does not include any parameters, so the log-likelihood function can be written as

$$
l(a_{m+1}, \ldots, a_T|\boldsymbol{\alpha}, A_m) = -\sum_{t=m+1}^{T} \left[\frac{1}{2}\ln(\sigma_t^2) + \frac{1}{2}\frac{a_t^2}{\sigma_t^2}\right].
$$

[14]

2.5.4 Estimation with t-innovation

Sometimes we may have a good reason to assume that $\{\epsilon_t\}$ follows a heavy-tailed distribution such as a standardized Student-t distribution. Let X_ν denote a random variable which has a Student-t distribution with ν degrees of freedom. Then we have $\mathrm{Var}(X_\nu) = \nu/(\nu-2)$ for $\nu > 2$, and $\boldsymbol{\epsilon_t} = X_\nu/\sqrt{\nu/(\nu-2)}$. The probability density function of $\boldsymbol{\epsilon_t}$ is

$$
f(\epsilon_t|\nu) = \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}} \left(1 + \frac{\epsilon_t^2}{\nu-2}\right)^{-(\nu+1)/2}, \qquad \nu > 2,
$$

[14]Tsay, Chapter 3, p 120-121.

where $\Gamma(x) = \int_0^\infty y^{x-1}e^{-y}dy$. Because $a_t = \sigma_t\epsilon_t$, the conditional-likelihood function of a_t is

$$f(a_{m+1},\ldots,a_T|\boldsymbol{\alpha},A_m) = \prod_{t=m+1}^{T} \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}}\frac{1}{\sigma_t}\left[1+\frac{a_t^2}{(\nu-2)\sigma_t^2}\right]^{-(\nu+1)/2},$$

where $\nu > 2$. The degrees of freedom of the t distribution can be also estimated with the maximum likelihood method. The conditional log-likelihood function is then

$$l(a_{m+1},\ldots,a_T|\boldsymbol{\alpha},A_m) = -\sum_{t=m+1}^{T}\left[\frac{\nu+1}{2}\ln\left(1+\frac{a_t^2}{(\nu-2)\sigma_t^2}\right)+\frac{1}{2}\ln\left(\sigma_t^2\right)\right].$$

[15]

2.5.5 Model validation

The standardized residuals of an ARCH model is defined as

$$\tilde{a}_t = \frac{a_t}{\hat{\sigma}_t}.$$

If an ARCH model is properly built, the standardized residuals follow an independently and identically distributed series. The series $\tilde{a}_t$ can be examined in order to check whether the fitted ARCH model is adequate to describe the data. The Ljung-box statistics of $\tilde{a}_t$ is used to check the adequacy of the mean equation. The Ljung-box statistics of $\tilde{a}_t^2$ is used to check the adequacy of the volatility equation. Histograms and quantile-to-quantile plot of $\tilde{a}_t$ can be used to verify if the distribution assumption is valid. [16]

2.5.6 Forecasting

Given that we know the values of $\{a_h,\ldots,a_{h+1-m}\}$, the 1-step-ahead forecast of σ_{h+1}^2 is

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \cdots + \alpha_{h+1-m}^2,$$

where h is the forecast origin.[17]

[15] Tsay, Chapter 3, p 121.
[16] Tsay, Chapter 3, p 122-123.
[17] Tsay, Chapter 3, p 123.

2.6 GARCH models

The ARCH models usually require many parameters to capture the volatility process. To cope with this problem, Bollerslev (1986) designed the generalized ARCH (GARCH) model.[18] For a log return series r_t, we have the innovation at time t, $a_t = r_t - \mu_t$. The GARCH model is

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2,$$

where $\{\epsilon_t\}$ is a series of independent and identical variables with mean zero and variance 1. Here ϵ_t is assumed to follow a standard normal distribution, or standardized Student-t distribution. The α_i and β_j are called ARCH and GARCH parameters, respectively.

2.6.1 The GARCH(1,1) model

Given that a_{t-1} and σ_{t-1} are known to us, the GARCH(1,1) model is

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \qquad 0 \leq \alpha_1, \beta_1 \leq 1, |\alpha_1 + \beta_1| < 1.$$

We can see that a large a_{t-1}^2 gives rise to a large σ_t^2 and a large σ_t^2 gives rise to a large a_t^2. This feature is in agreement with volatility clustering in empirical data.

The constraints are $\alpha_0 > 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$.

2.6.2 Estimation

The modeling procedure of GARCH is analogous to that of ARCH models (see Section 2.7.3 and 2.7.4 for parameter estimation). The conditional maximum-likelihood method is used given that the starting values of the volatility $\{\sigma_t^2\}$ are assumed to be known. The sample variance of a_t can be used as a starting value of σ_1. [19]

2.6.3 Model validation

Model validation for a GARCH(1,1) model is analogous to that for ARCH models (see Section 2.5.5). We check whether the standardized residuals are independently and identically distributed. We can also check whether they follow the distribution assumption that we made for model building, with help of histograms and quantile-to-quantile plots. Ljung-box tests will be run on the standardized residuals to check model adequacy.

[18] Bollerslev, Journal of Econometrics 31

[19] Tsay, Chapter 3, p 134.

2.6.4 Forecasting

We have the forecast origin h. Given that we know a_h and σ_h, 1-step-ahead forecast of a GARCH(1,1) model is

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2.$$

20

2.7 Backtesting

Our models will be validated with a backtesting method. One-step-ahead estimate from a sample set of our data will be compared with the actual value. We continue to roll our sub-sample, or "window" forward and continue this process of comparison. Value at risk will be used for measurement for our backtesting.

2.7.1 Value at risk

Value at risk is a risk measure to estimate potential financial loss. Let $\triangle V(l)$ denote the change in the value of the underlying asset from time t to time $t + l$, and we denote the associated loss function by $L(l)$. We denote the cumulative distribution function of $L(l)$ by $G_l(x)$. Then we can write the value at risk of a financial position over the time period l with tail probability p as

$$p = \Pr[L(l) \geq \text{VaR}] = 1 - \Pr[L(l) < \text{VaR}]$$

The qth quantile of $G_l(x)$ is

$$x_q = \inf\{x | G_l(x) \geq q\},$$

where inf is the smallest real number x for $G_l(x) \leq x_q$. If $L(l)$ is a continuous random variable, $q = \Pr[L(l) \leq x_q]$. If the cumulative distribution function of $G_l(x)$ is known, value at risk is the $(1-p)$th quantile of the cumulative distribution function of the loss function. [21] For a normal distribution, if the tail probability is 5% VaR $= 1.65\sigma_{t+1}$ or VaR $= -1.65\sigma_{t+1}$ for next day, [22] depending on your financial position.[23] For a Student's t distribution, we scale a quantile, by multiplying $\sqrt{(\nu-2)/\nu}$ to a quantile to create variance 1.

[20] Tsay, Chapter 3, p 133.
[21] Tsay, Chapter 7, p 327.
[22] In this paper, -1.64485 is used for a quantile.
[23] Tsay, Chapter 7, p 327-329.

2.7.2 Violation-based tests for value at risk

Let VaR_p^t denote the p-th quantile of the conditional loss distribution $G_{\mathrm{L}_{t+1}|L_t}$ at time t. The event $\{L_{t+1} > \mathrm{VaR}_p^t\}$ is called a VaR violation. The event indicator is defined as $\mathrm{I}_{t+1} = \mathrm{I}_{\{L_{t+1} > \mathrm{VaR}_p^t\}}$. For a continuous loss distribution, we have

$$\mathrm{E}[\mathrm{I}_{t+1}|L_t] = \mathrm{P}[L_{t+1} > \mathrm{VaR}_p^t|L_t] = 1 - p.$$

The indicator variable I_{t+1} is a Bernoullie variable with probability $1 - p$.

The sum of the violation indicators over a certain time period forms binomially distributed random variables as below.

$$\sum_{t=1}^{m} \mathrm{I}_{t+1} \sim \mathrm{B}(m, 1-p),$$

where m is the number of total sequences, or the time points for a time series.

We estimate VaR_p^t based on information available up to $t - 1$, and denote the estimate as $\widehat{\mathrm{VaR}}_p^t$. We compare the actual realized value at time $t + 1$ with our VaR estimate at time t. The violation indicator variable can be written as

$$\hat{I}_{t+1} = I_{\{L_{t+1} > \widehat{\mathrm{VaR}}_p^t\}}$$

Under the null hypothesis that our estimation method is accurate, the sum of violations should form a $\mathrm{B}(m, 1 - p)$ distribution. This hypothesis will be tested with a binomial test. The statistic for a two-sided score test is

$$Z_m = \frac{\sum_{t+1}^{m} \hat{I}_{t+1} - m(1-p)}{\sqrt{mp(1-p)}}$$

The null hypothesis of Bernoulli behavior at the 5% level is rejected if the p-value is less than a significance level.[24]

3 Methodology

3.1 Large Cap and Small Cap indices

All companies listed in the Nasdaq belong to one of the three segments, namely the Large Cap, the Mid Cap and the Small Cap segment. The Large Cap segment has companies with a market value over one billion Euros. The

[24]McNeil, Frey and Embrechts, Chapter 9, p 352-353.

Mid Cap segment includes companies with a market value less than one billion and above 150 million Euros. Companies with a market value below 150 million Euros belong to the Small Cap segment.[25]

As the market values for each segment differ from one another, one can assume that the dynamics of each segment can be different from one another. For that reason, we want to study the Large Cap and the Small Cap indices more closely. We want to study if a same volatility model works differently on these two indices, or it works more or less the same regardless of which segment the index is from.

We take the Large Cap and the Small Cap indices that span from Jan 1, 2003 to Dec 31, 2018. Each segment index has 4081 prices. The time plots of closing prices are shown in Figure 1. Both of the prices seem to follow a similar pattern over the time span.

Figure 2 shows the log returns of both indices. The log returns of the Large Cap move between -0.08 and 0.08, while the log returns of the Small Cap move between a smaller range, between -0.05 and 0.05.

Descriptive statistics for the log returns of the Large Cap and the Small Cap is summarized in Table 1. The means of both log returns are close to zero. The log return series of the Small Cap has a smaller unconditional variance.

	Large Cap	Small Cap
Mean	0.000227	0.000360
Variance	0.000168	0.000054
Standard deviation	0.012944	0.007348

Table 1: Descriptive statistics for the log returns

3.2 ACF and PACF

ACF plots are useful to check serial correlations and dependency of data. The sample PACF of the squared residuals are useful to determine an order for an ARCH model. We plot the ACF of the log returns, of the squared log returns and of the absolute value of the log returns. We also plot the sample PACF of the squared log returns.

Figure 3 shows these plots for the log returns of the Large Cap. The ACF plot for the series shows no significant signs of serial correlations, except for

[25] "Rules for the Construction and Maintenance of the NASDAQ OMX ALL-share, Benchmark and Sector Indexes", version 2.4, March 2018, NASDAQ Copenhagen A/S, NASDAQ Helsinki Ltd, NASDAQ Iceland hf., NASDAQ Stockholm AB

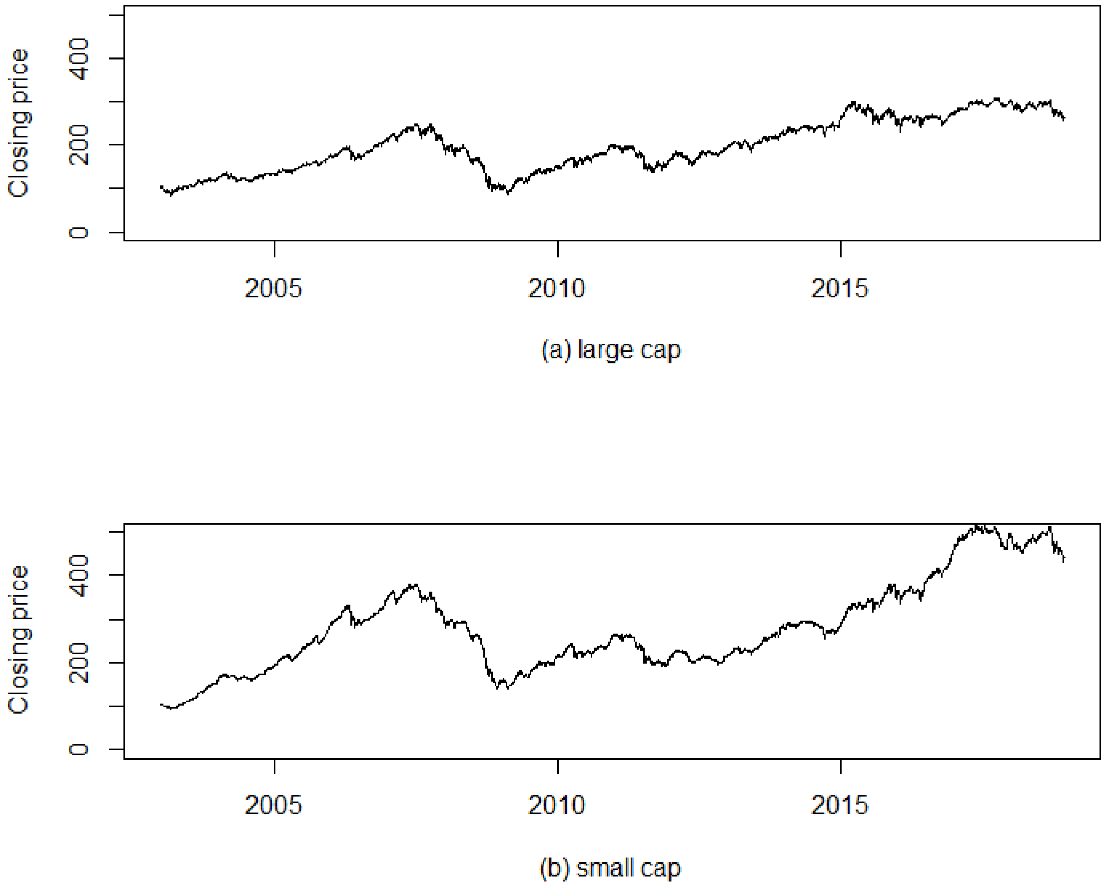

Figure 1: Closing prices of (a) the Large Cap and (b) the Small Cap from 2003 to 2018

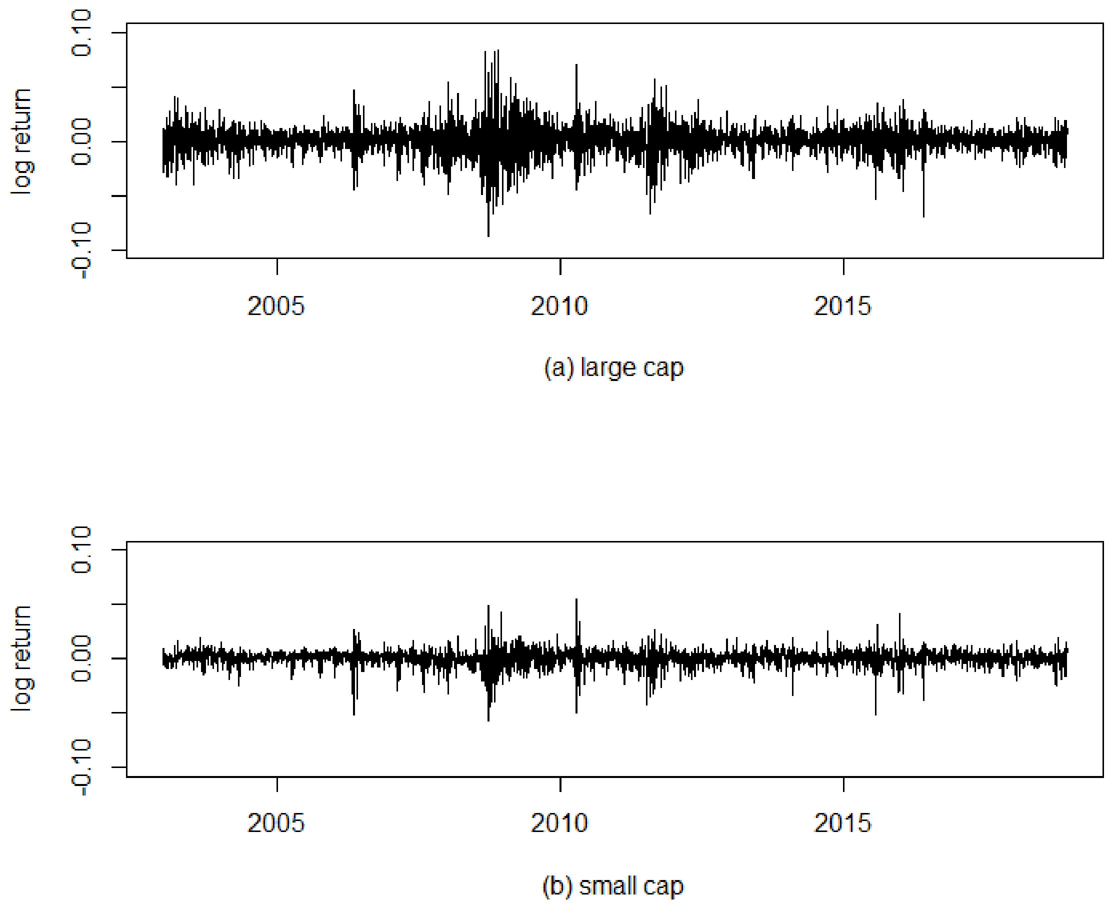

Figure 2: The log returns of (a) the Large cap and (b) the Small cap from 2003 to 2018

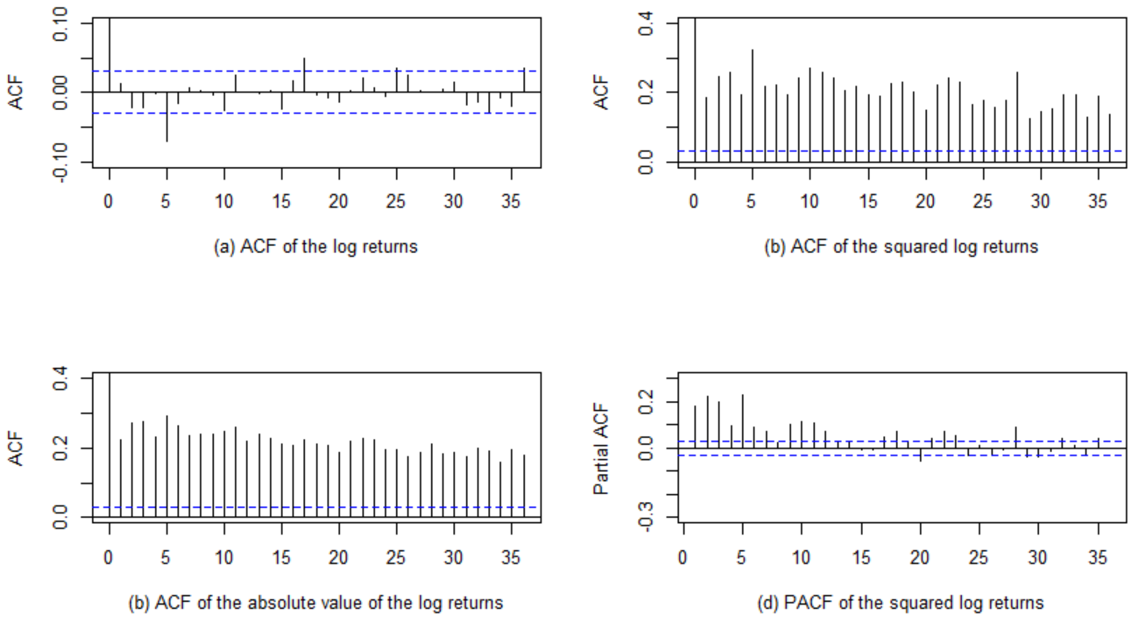

Figure 3: Sample ACF and PACF of the log returns of the Large Cap from 2003 to 2018

lag 5. The ACF plots for the squared and the absolute value of the log returns of the Large Cap show serial dependency. The sample PACF suggests lag 5 for an ARCH model.
Figure 4 shows the sample ACF plots and the PACF plot of the log returns of the Small Cap. Figure 4 suggests that the data has serial correlations up to lag 4, it cuts off, and it appears again, repeating this pattern for a while. The ACF of the squared log returns, and that of the absolute value of the log returns show dependency between the values. The sample PACF of the Small Cap suggests lag 2 for an ARCH model.

In summary, the Large Cap does not show any clear signs of serial correlation, while it shows strong dependency. The Small Cap shows stronger signs of serial correlation and weaker dependency than the Large Cap data. This observation indicates that the Large Cap is more suitable data for building volatility models than the Small Cap data is.

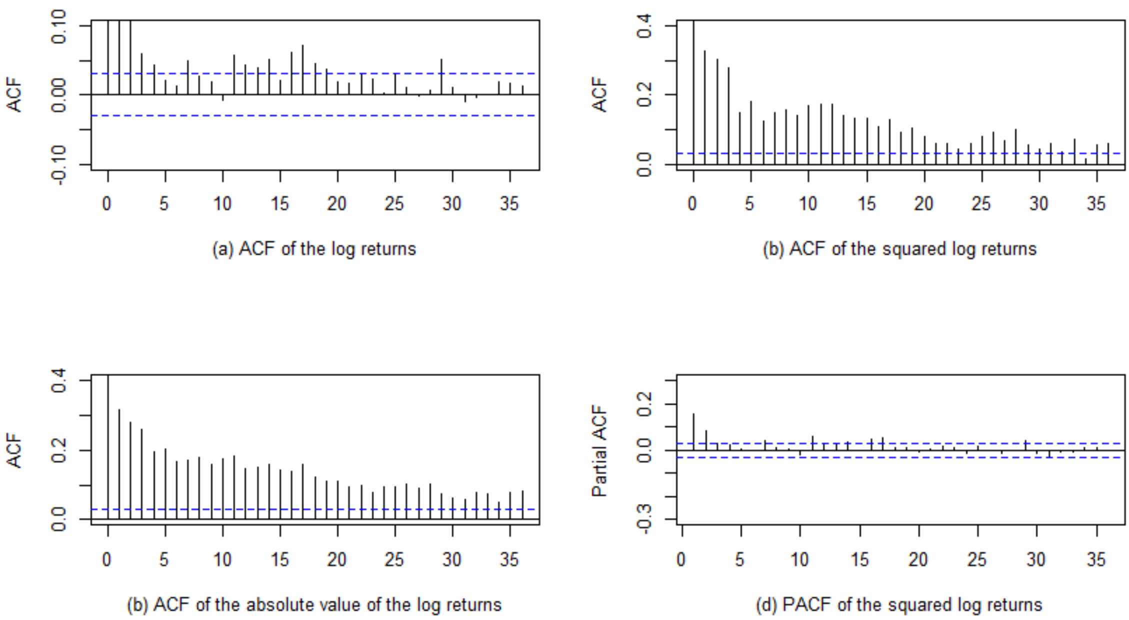

Figure 4: Sample ACF and PACF of the log returns of the Small Cap from 2003 to 2018

Data	p-value	95% confidence interval
Large Cap	0.2625	[-0.0001701865, 0.0006243995]
Small Cap	0.001759	[0.0001345516, 0.0005856323]

Table 2: T-test for mean zero

Data	Residuals, $\{a\}$	Lag	P-value
Large Cap	$\{a\}$	8	0.0012
Large Cap	$\{a^2\}$	8	<2.2e-16
Small Cap	$\{a\}$	8	<2.2e-16
Small Cap	$\{a^2\}$	8	<2.2e-16

Table 3: Ljung-Box tests on the squared residuals

3.3 Mean equation

We want to check the means of both log returns. We run t-tests for the null hypothesis of mean zero. Table 2 shows the test results. For the Large Cap, the p-value is high, and a 95% confidence interval covers zero. For the Small Cap, the p-value is lower than 0.2% and a 95% confidence interval does not cover zero. However, both the lower and the upper limit of the confidence interval are very close to zero. In this paper we will assume mean zero even for the Small Cap data, because the mean is close enough to zero, and we want to focus on volatility modeling.

3.4 Ljung-Box tests

The ACF plots already indicate that serial correlations exist, especially for the Small Cap. We run Ljung-box tests on the residuals and the squared residuals to more closely check if there exist serial correlations and ARCH effects. In our case, the means are assumed to be zero so the log returns series are equal to the residuals.

Table 3 shows the results of the Ljung-Box tests on the residuals. The lag for the test is chosen by $m \approx \ln(n)$.[26] This gives us $m = 8$ for our case. The p-value for the residuals of the Large Cap is 0.0012, and that for the squared residuals is close to zero. The p-values for the residuals and the squared residuals of the Small Cap are close to zero. The null hypothesis of no serial correlation is rejected for both data sets. The test results show that serial correlation exists for both data sets, and there are strong ARCH effects in both data sets.

[26]Tsay, Chapter 2, p 33.

Distribution	Normality assumption		t-innovation	
Parameter	Estimate	p-value	Estimate	p-value
α_0	4.464e-05	<2e-16	4.367e-05	<2e-16
α_1	8.233e-02	7.36e-05	9.226e-02	4.58e-05
α_2	1.860e-01	2.22e-16	1.923e-01	2.73e-12
α_3	1.704e-01	1.09e-13	1.773e-01	3.15e-10
α_4	1.460e-01	3.00e-12	1.452e-01	7.94e-09
α_5	1.609e-01	3.11e-15	1.662e-01	1.11e-10

Table 4: Parameters of ARCH(5) fitted to the log returns of the Large Cap

3.5 Model building

3.5.1 Fitting ARCH models

The PACFs suggest lag 5 for the Large Cap and lag 2 for the Small Cap. We will fit an ARCH model with same order to both data sets because we want to see if a same model works differently depending on the data. We thus choose lag 5 for ARCH models. Another reason for choosing lag 5 over lag 2 is that we already fixed the order of GARCH models as (1,1), so we want to build an ARCH model with a substantially higher order so that we have a better chance to see differences between these two models.

The parameters are estimated with the conditional maximum likelihood method as shown in section 2.7.3 and section 2.7.4. Table 4 shows the parameter estimates for the ARCH(5) models fitted to the log returns of the Large Cap, under normality assumption and with t-innovation. All six parameter estimates are statistically significant. Table 5 shows the parameter estimates for the ARCH(5) models fitted to the log returns of the Small Cap under normality assumption and with t-innovation. All six parameters are significantly different from zero.

3.5.2 Fitting GARCH(1,1) models

We continue to assume mean zero for both log returns.. The parameters are estimated with the conditional maximum likelihood method. Table 6 shows parameter estimates for a GARCH(1,1) model fitted to the Large Cap log returns. All parameters are statistically significant under normality assumption and with t-innovation as well. Table 7 shows parameter estimates for a GARCH(1,1) model fitted to the Small Cap log returns. The p-values for all parameters are close to zero. All parameters are statistically significant.

Model	ARCH(5) fitted to the Small Cap			
Distribution	Normal		t-innovation	
Parameter	Estimate	p-value	Estimate	p-value
α_0	1.567e-05	<2e-16	1.702e-05	<2e-16
α_1	2.484e-01	<2e-16	2.181e-01	4.44e-13
α_2	1.610e-01	5.48e-13	1.608e-01	3.20e-19
α_3	1.107e-01	1.00e-08	1.198e-01	2.09e-06
α_4	1.054e-01	8.96e-07	8.027e-02	0.00103
α_5	1.050e-01	1.67e-08	1.033e-01	4.32e-06

Table 5: Parameters of ARCH(5) fitted to the log returns of the the Small Cap

Model	GARCH(1,1) fitted to the Large Cap			
Distribution	Normal		t-innovation	
Parameter	Estimate	p-value	Estimate	p-value
α_0	1.617e-06	1.8e-06	1.413e-06	0.000157
α_1	8.311e-02	<2e-16	8.269e-02	<2e-16
β_1	9.067e-01	<2e-16	9.097e-01	<2e-16

Table 6: Parameters of GARCH(1,1) fitted to the log returns of the Large Cap

Model	GARCH(1,1) fitted to the Small Cap			
Distribution	Normal		t-innovation	
Parameter	Estimate	p-value	Estimate	p-value
α_0	2.773e-06	6.24e-12	3.130e-06	3.78e-09
α_1	1.802e-01	<2e-16	1.708e-01	<2e-16
β_1	7.699e-01	<2e-16	7.677e-01	<2e-16

Table 7: Parameters of GARCH(1,1) fitted to the log returns of the Small Cap

3.6 Model validation

3.6.1 ARCH Models

Figure 5 shows the ACF plots of the standardized residuals from the ARCH(5) models. The ACF plots suggest no serial correlations for the ARCH(5) models fitted to the log returns of the Large Cap. That means both of the ARCH(5) models fitted to the the log returns of the Large Cap, the one under normality assumption and the other one with t-innovation, are adequate. The plots, however, indicate that the ARCH(5) models fitted to the log returns of the Small Cap still have serial correlations.

The Ljung-Box tests were run to check model adequacy. Ljung-box tests on the standardized residuals check adequacy of the mean equation, while Ljung-box tests on the squared standardized residuals check adequacy of the variance equation. Table 8 shows the results of Ljung-Box tests on the standardized residuals from the ARCH(5) models. For the Large Cap log returns, the p-values of the Ljung-Box tests on the standardized residuals are high, but the p-values of the Ljung-box tests on the squared standardized residuals are close to zero. For the Small Cap log returns, the result is the opposite. The p-values of the Ljung-Box tests on the standardized residuals are close to zero, while the p-values of the Ljung-box tests on the squared standardized residuals are high. In summary, the assumption of mean zero is adequate for the Large Cap data, but not adequate for the Small Cap data. The variance equation from our ARCH(5) model is adequate to describe the Small Cap log returns, but not adequate to describe the Large Cap log returns.

The distribution assumption was checked with histograms and quantile-to-quantile plots. Figure 6 shows histograms of the standardized residuals of the ARCH(5) models under normality assumption and with t innovation, fitted to the Large Cap and the Small Cap log returns. The blue line is the normal distribution curve. It is not quite clear to see whether the data follows the normal distribution or not, so we can check QQ plots instead.

Figure 7 show QQ plots of the standardized residuals from the ARCH models. The plots indicate that the standardized residuals of both the Large Cap and the Small Cap log returns do not follow the normal quantile line. They seem to fit the Student t-distribution better. The degrees of freedom are estimated with the maximum likelihood method (see Section 2.5.4) and listed in Table 10.

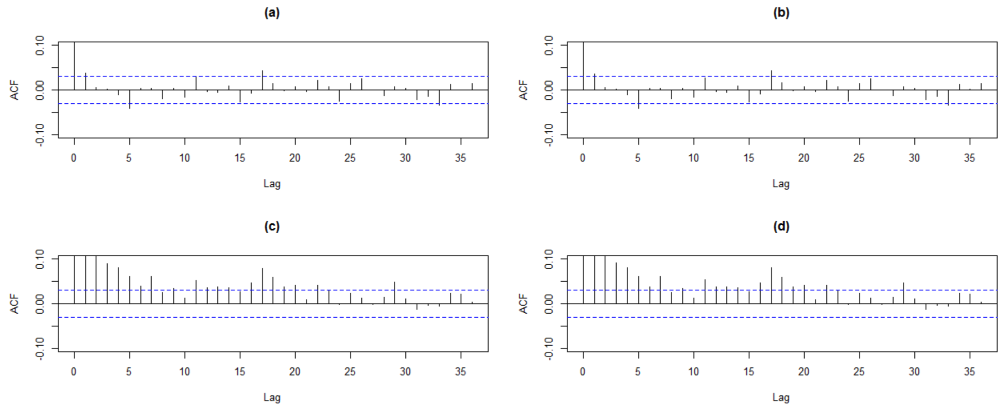

Figure 5: The sample ACF plots of the standardized residuals from the ARCH(5) models fitted to (a) the Large Cap log returns under normality assumption, (b) the Large Cap log returns with t-innovation, (c) the Small Cap log returns under normality assumption, and (d) the Small Cap log returns with t-innovation.

Data	Test	$(\tilde{a})$	Test statistic	p-value for normality	p-value for t-innovation
Large Cap	LB	$\tilde{a}$	Q(10)	0.102019	0.1084049
Large Cap	LB	$\tilde{a}$	Q(15)	0.09553013	0.1005285
Large Cap	LB	$\tilde{a}$	Q(20)	0.05226862	0.05401963
Large Cap	LB	$\tilde{a}^2$	Q(10)	1.619799e-05	1.427092e-05
Large Cap	LB	$\tilde{a}^2$	Q(15)	3.391398e-12	1.236011e-11
Large Cap	LB	$\tilde{a}^2$	Q(20)	2.651213e-12	1.355549e-11
Small Cap	LB	$\tilde{a}$	Q(10)	<2.2e-16	<2.2e-16
Small Cap	LB	$\tilde{a}$	Q(15)	<2.2e-16	<2.2e-16
Small Cap	LB	$\tilde{a}$	Q(20)	<2.2e-16	<2.2e-16
Small Cap	LB	$\tilde{a}^2$	Q(10)	0.3905513	0.3695459
Small Cap	LB	$\tilde{a}^2$	Q(15)	0.07442856	0.0368284
Small Cap	LB	$\tilde{a}^2$	Q(20)	0.006822326	0.001153529

Table 8: Model validation for ARCH(5) models (LB: Ljung-Box test, and $\tilde{a}$: the standardized residuals).

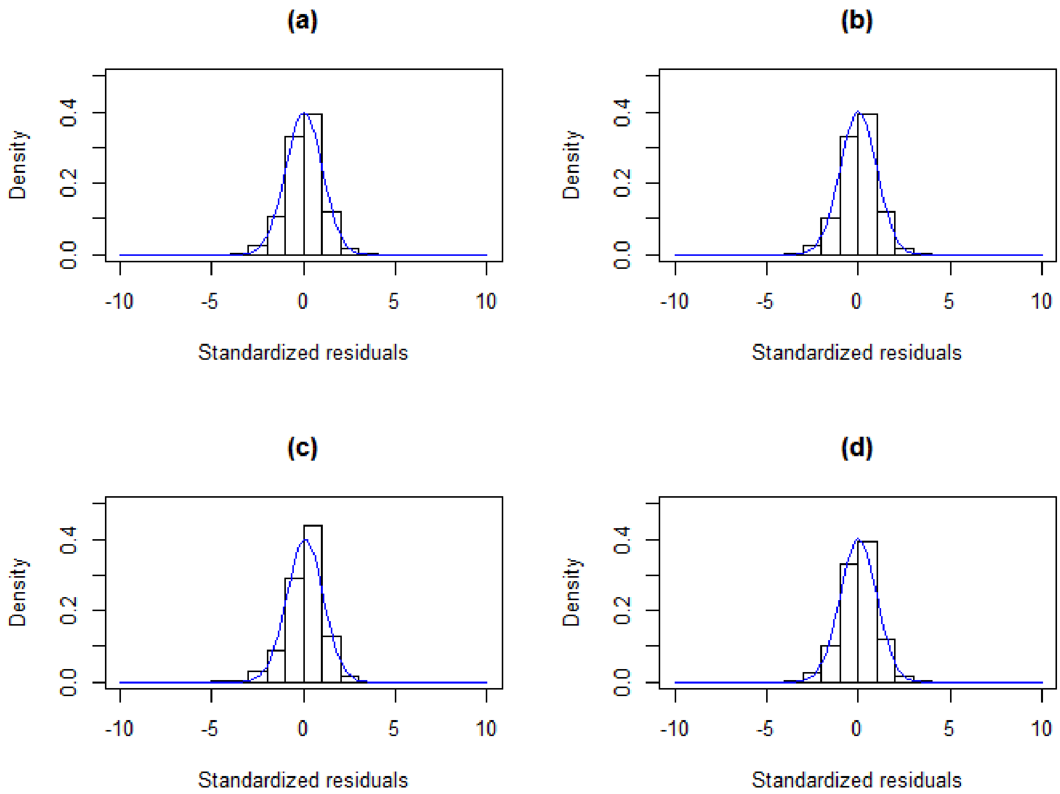

Figure 6: Histograms of the standardized residuals from the ARCH models fitted to (a) the Large Cap log returns under normality assumption, (b) the Large Cap log returns with t-innovation, (c) the Small Cap log returns under normality assumption, and (d) the Small Cap log returns with t-innovation.

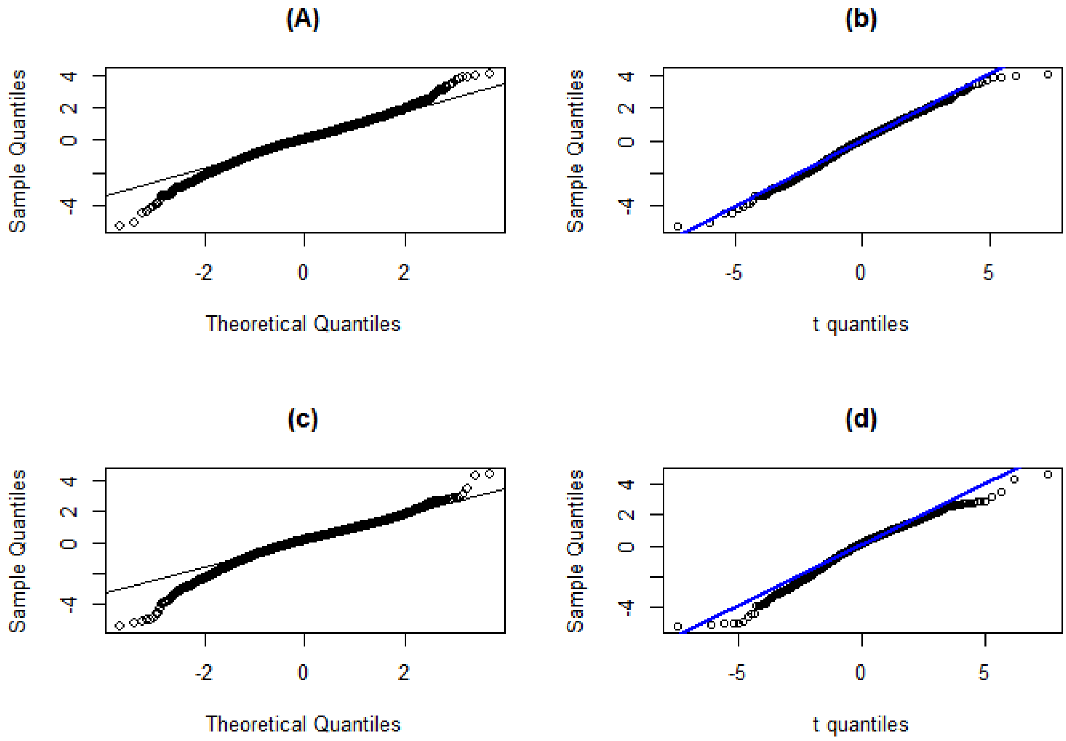

Figure 7: QQ plots for the standardized residuals from the ARCH models fitted to (a) the Large Cap log returns under normality assumption, (b) the Large Cap log returns with t-innovation, (c) the Small Cap log returns under normality assumption, and (d) the Small Cap log returns with t-innovation.

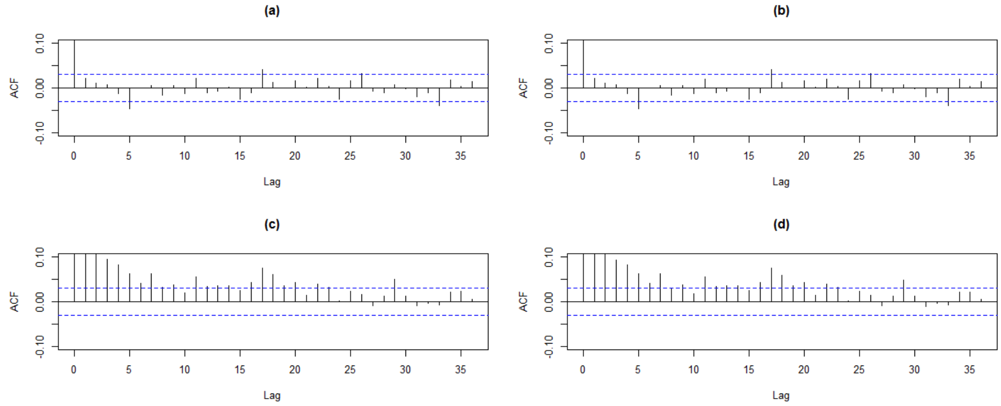

Figure 8: The Sample ACF plots of the standardized residuals from the GARCH models fitted to (a) the Large Cap log returns under normality assumption, (b) the Large Cap log returns with t-innovation, (c) the Small Cap log returns under normality assumption, and (d) the Small Cap log returns with t-innovation.

3.6.2 GARCH Validation

Figure 8 shows ACF plots of the standardized residuals from the GARCH(1,1) models. The plots show that serial correlations are nearly non-existent for the GARCH models fitted to the Large Cap log returns, whether the model was built under normality assumption or with t-innovation. Meanwhile, the ACF plots of the standardized residuals from the GARCH(1,1) fitted to the Small Cap log returns show significant signs of serial correlations.

Table 9 shows the results from Ljung-box tests on the standardized residuals from the GARCH models. According to the test results, the mean equation of the GARCH(1,1) fitted to the Large Cap log returns is adequate and the variance equation of this model is also adequate for lag 10. For the Small Cap data, the mean equation does not seem adequate, while the variance equation seems adequate.

Distribution assumption for innovation is to be checked with histograms and quantile-to-quantile plots. Figure 9 shows the histograms of the standardized residuals from the GARCH(1,1) models. It is, however, not easy to see whether the standardized residuals follow the normal distribution or not. Figure 10 shows quantile-to-quantile plots of the standardized residuals of the GARCH(1,1) models against normal quantiles and t-distribution quantiles. The degrees of freedom for t-distribution are estimated with the maximum likelihood method (see Section 2.5.4) and listed in Table 10. According to the QQ plots, Student t-distribution seems more adequate.

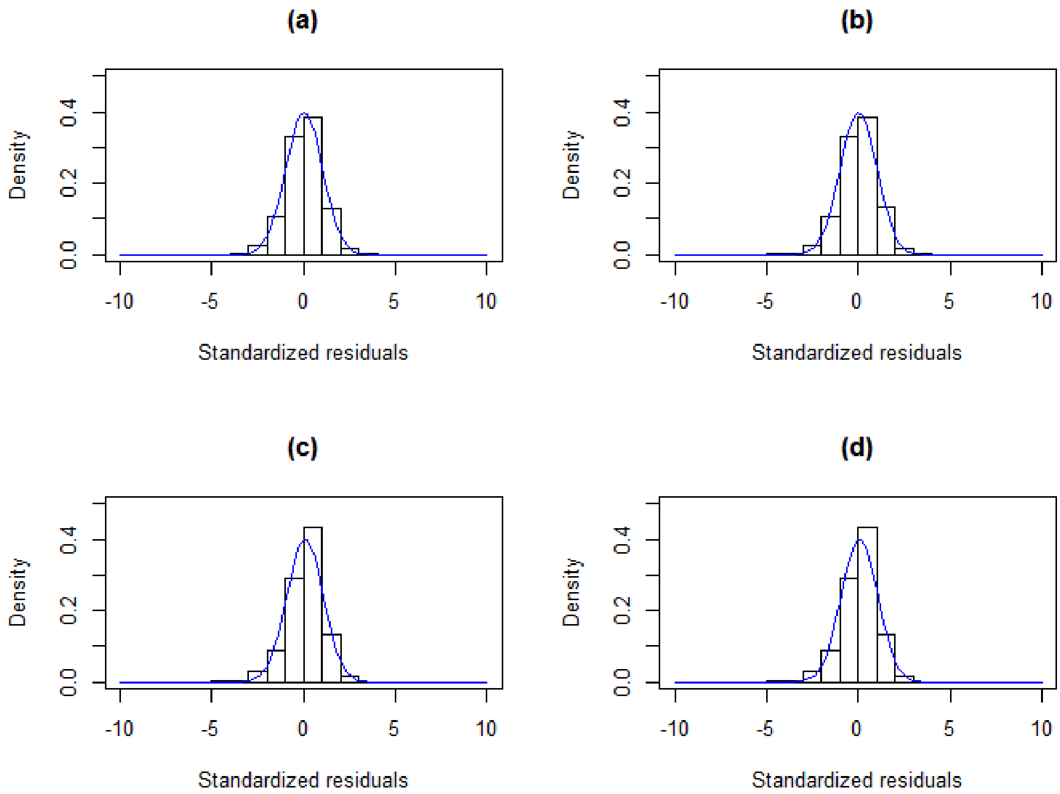

Figure 9: Histograms of the standardized residuals from the GARCH models fitted to (a) the Large Cap log returns under normality assumption, (b) the Large Cap log returns with t-innovation, (c) the Small Cap log returns under normality assumption, and (d) the Small Cap log returns with t-innovation

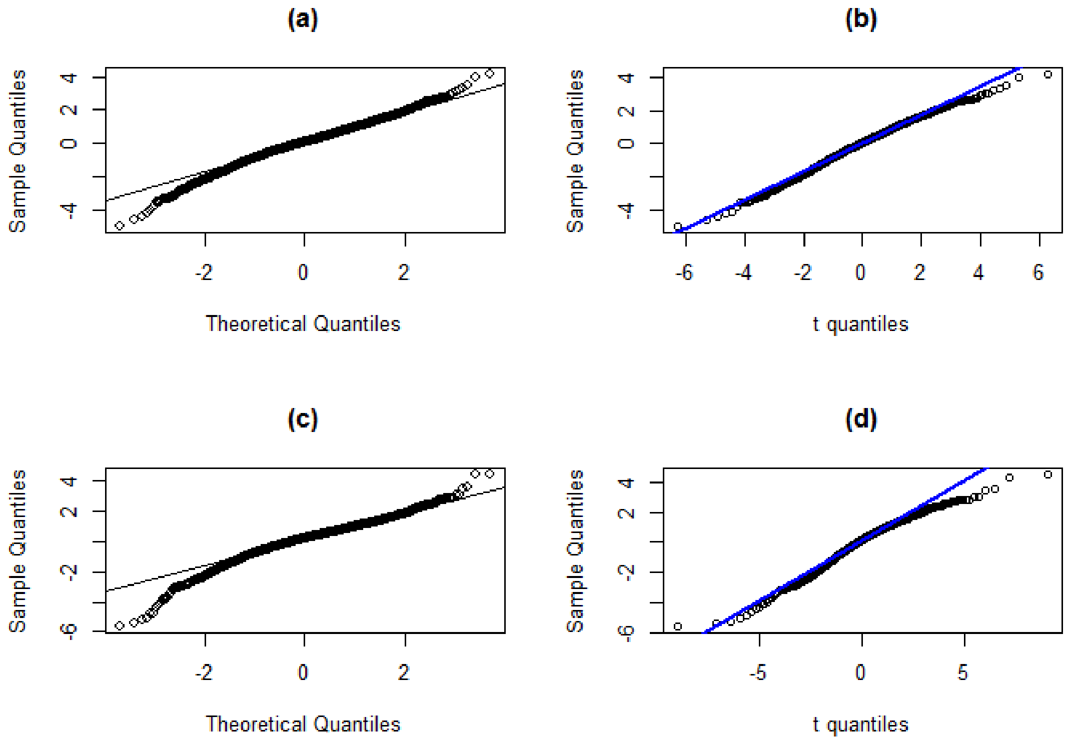

Figure 10: QQ plots of the standardized residuals from the GARCH models fitted to (a) the Large Cap log returns under normality assumption, (b) the Large Cap log returns with t-innovation, (c) the Small Cap log returns under normality assumption, and (d) the Small Cap log returns with t-innovation.

Data	Test	($\tilde{a}$)	Test statistic	p-value for normality	p-value for t-innovation
Large Cap	LB	$\tilde{a}$	Q(10)	0.1854053	0.1887777
Large Cap	LB	$\tilde{a}$	Q(15)	0.2142331	0.2211507
Large Cap	LB	$\tilde{a}$	Q(20)	0.09977573	0.1014339
Large Cap	LB	$\tilde{a}^2$	Q(10)	0.7370318	0.7371919
Large Cap	LB	$\tilde{a}^2$	Q(15)	0.001194696	0.001420884
Large Cap	LB	$\tilde{a}^2$	Q(20)	0.0005675862	0.0005493208
Small Cap	LB	$\tilde{a}$	Q(10)	<2.2e-16	<2.2e-16
Small Cap	LB	$\tilde{a}$	Q(15)	<2.2e-16	<2.2e-16
Small Cap	LB	$\tilde{a}$	Q(20)	<2.2e-16	<2.2e-16
Small Cap	LB	$\tilde{a}^2$	Q(10)	0.5535422	0.4990226
Small Cap	LB	$\tilde{a}^2$	Q(15)	0.4870336	0.3988024
Small Cap	LB	$\tilde{a}^2$	Q(20)	0.3513989	0.2717797

Table 9: Model validation for GARCH(1,1) models (LB: Ljung-Box test, and $\tilde{a}$: the standardized residuals).

3.7 Backtesting

Now we run backtesting to see how the forecasting power of our models work on different data sets. The backtesting procedure is as follows.

1. A window is to be chosen from a data set from $t = 1$ to $t = m$, where m is smaller than the total number of observations, say T.
2. A model is fitted to this window and the parameters are to be estimated.
3. Based on the parameter estimates, 1-step-ahead forecast of volatility is made.
4. The value at risk at time $t = m + 1$ is calculated with this volatility estimate, and a quantile.
5. The estimated value at risk for time $t = m + 1$ is compared with the realized log return at time $t = m + 1$.
6. If the realized log return is lower than the estimated value at risk, the event is recognized as a violation.
7. The steps from 3 to 6 is repeated n times until we reach the point where $m + n = T$.
8. The sum of the violations is supposed to follows a binomial distribution with probability $1 - \alpha$. The binomial test will be run to check if the

Data	Model	d.f	quantile	scaled quantile
Large Cap	ARCH	6.433	-1.92001	-1.59384
Large Cap	GARCH	8.522	-1.88411	-1.64826
Small Cap	ARCH	6.232	-1.93039	-1.59076
Small Cap	GARCH	6.503	-1.91659	-1.59486

Table 10: The 5% quantiles for ARCH and GARCH models with t-innovation (d.f: degrees of freedom).

model estimates risk properly. (In this study, α for the binomial test is set for 95%.)

We use two different lengths for a window, 1 year and 2 years. For value at risk calculation, 5% quantile will be used. For models under the normality assumption, the 5% quantile is -1.64485. For models with t-innovation, scaled 5% quantiles will be used. These scaled quantiles are listed in Table 10.

Table 11 shows the backtesting results for the ARCH(5) models. For the Large Cap log returns, 1-year window with t innovation has the highest p-value, 0.4807. Other models seem to overestimate the risk with the probability ranging from 5.75% to 6.38%. We compare 1-year window tests with 2-year window tests. Tests with a 2-year window test always has a slightly lower probability, except for the model fitted to the Large Cap log returns with t-innovation. Next, we compare if there is any difference between the distribution assumption. We take 2-year window tests for comparison. The models with t-innovation give more conservative results than the models under normality assumption.

Table 12 shows the backtesting results for GARCH(1,1) models. The backtesting with a 2-year window on the GARCH model fitted to the Large Cap with t innovation is statistically significant at 5% level. Probabilities for other models are higher than 5%. We first compare 1-year window test results and 2-year window test results. For the Large Cap log returns, the probabilities from 2-year window tests are slightly lower than the probabilities from 1-year window tests. For the Small Cap log returns, the result is the opposite. The probabilities go slightly up when the window length increases from 1 to 2 years, even though the differences are fairly small. Next, we compare if there is any difference between distribution assumptions. We take 2-year window results for comparison. For the Large Cap log returns, the models under normality assumption show higher probabilities. They tend to give more conservative estimates. For the Small Cap log returns, the models with t innovation give more conservative estimates.

Next we want to check if there is any difference between ARCH and

Data	Distribution	Window	Violations	Trials	Probability	p-value
Large Cap	Normal	1 year	231	3824	0.06040795	0.004242
Large Cap	Normal	2 years	205	3568	0.05745516	0.04565
Large Cap	t-innovation	1 year	181	3824	0.04733264	0.4807
Large Cap	t-innovation	2 years	217	3568	0.06081839	0.003934
Small Cap	Normal	1 year	225	3824	0.05883891	0.01427
Small Cap	Normal	2 years	208	3568	0.05829596	0.0258
Small Cap	t-innovation	1 year	244	3824	0.06380753	0.0001517
Small Cap	t-innovation	2 years	225	3856	0.06306054	0.0005381

Table 11: Backtesting results for the ARCH models

Data	Distribution	Window	Violations	Trials	Probability	p-value
Large Cap	Normal	1 year	229	3824	0.05988494	0.006016
Large Cap	Normal	2 years	210	3568	0.0588565	0.01719
Large Cap	t-innovation	1 year	221	3824	0.05779289	0.02858
Large Cap	t-innovation	2 years	203	3568	0.05689462	0.05984
Small Cap	Normal	1 year	225	3824	0.05883891	0.01427
Small Cap	Normal	2 years	212	3568	0.05941704	0.01121
Small Cap	t-innovation	1 year	235	3824	0.06145397	0.001596
Small Cap	t-innovation	2 years	223	3568	0.0625	0.0009415

Table 12: Backtesting results for the GARCH(1,1) models

GARCH models. We take 1-year window tests on the Large Cap log returns with t innovation, because they have fairly high p values for both models. The ARCH model gives a probability of 4.73%, while the GARCH model gives a probability of 5.78%. The GARCH model gives more conservative results.

Figure 11 shows each data set with value at risk estimates. For the Large Cap log returns, the value at risk estimates from the GARCH model (orange) seem to lie above those estimates from the ARCH model (green) in most points. For the Small Cap log returns, there seem no clear differences between the value at risk estimates from the ARCH model (blue) and the value at risk estimates from the GARCH model (red).

4 Conclusion

4.1 Results of the study

Table 13 summarizes the overall results of this study.

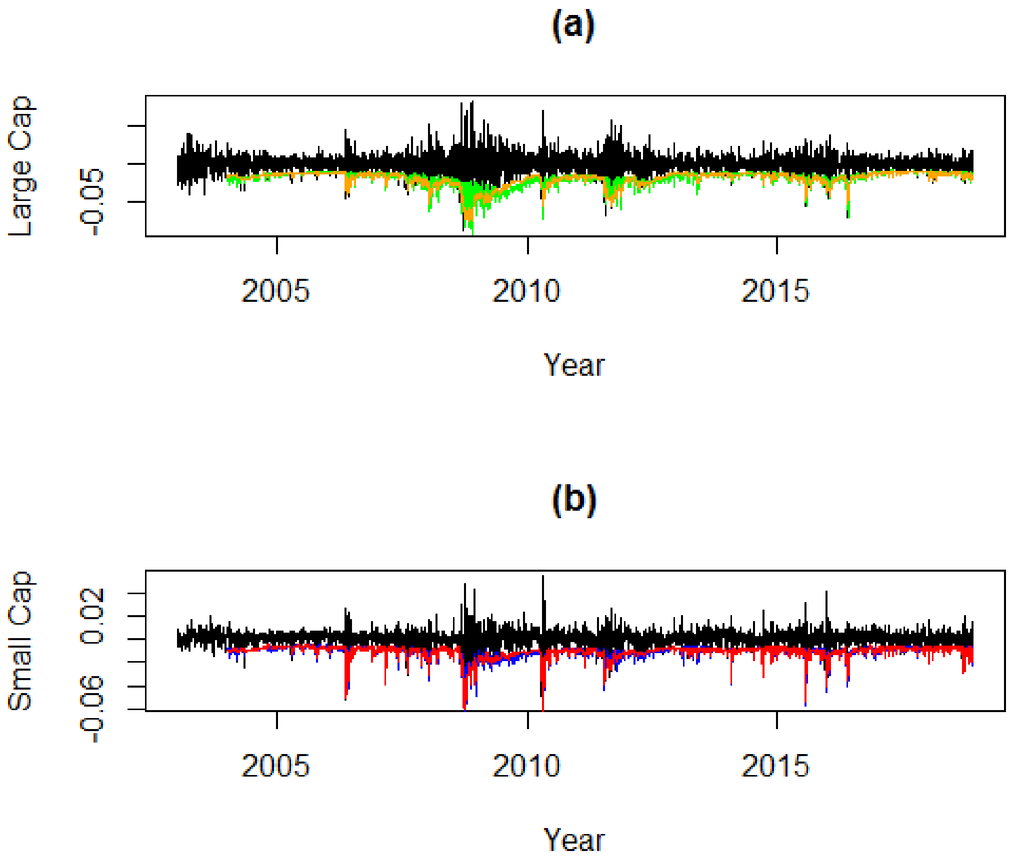

Figure 11: Log returns with VaR estimations. (a) the Large Cap log returns with VaR estimations from ARCH (green) and VaR estimations from GARCH (orange), and (b) the Small Cap log returns with VaR estimations from ARCH and VaR estimations from GARCH (red).

For the Large Cap, the ACF plots show no clear signs of autocorrelations for both ARCH and GARCH models. The mean equation, or mean zero for our case, is adequate for both models. The variance equation, however, seems inadequate to describe the data, according to the Ljung-box test results. For distribution assumption, QQ plots show that Student's t distribution may be more adequate than the normal distribution. Backtesting results show that the models with t-innovation produce more accurate forecasts, both for the ARCH and the GARCH model.

For the Small Cap, the ACF plots show clear signs of autocorrelation, both for the ARCH and the GARCH models. The mean equation seems inadequate for both models, according to the Ljung-Box tests on the standardized residuals. The variance equations seem adequate to describe the data, whether it is the ARCH or the GARCH models. QQ plots suggest that Student's t distribution assumption is more adequate than normality assumption for shocks. Backtesting results show, however, that the models under normality assumption produce more reliable forecasts. The binomial test results for both of the models under normality assumption are statistically significant at 1% significance level. The binomial test results for the models with t innovation are not statistically significant. They give conservative results, 6.38% for the ARCH model and 6.15% for the GARCH model.

In summary, the differences between the data sets are more distinctive than the differences between the models. For a same data set, both the ARCH and the GARCH models show similar test results. For a same model, the adequacy of the model is different depending on the data.

However, that is not the case with backtesting. The distribution assumption makes more difference in forecasting power than the model itself. For the Large Cap data, the models with t innovation produce more accurate forecasts, whether it is an ARCH model or a GARCH model. For the Small Cap, the models under normality assumption produce more accurate forecasts, with their probabilities closer to 5%, regardless of the model.

4.2 Suggestions for further studies

The main differences between the Large Cap and the Small Cap log returns were assumptions about mean zero and no significant autocorrelations. The mean of the Small Cap index was slightly over zero. The ACF plots showed clear signs of autocorrelations for the Small Cap log returns. One can address these issues and try to remove these problems before building any volatility models.

The ACF plots and the Ljung-box tests on the standardized residuals did not seem to agree with each other. The ACF plots show that both of the

		ARCH(5)		GARCH(1,1)	
Data	Test	Normality	t-innovation	Normality	t-innovation
Large Cap	ACF of $\tilde{a}$	Adequate	Adequate	Adequate	Adequate
	LB on $\tilde{a}$	H_0 not rejected	H_0 not rejected	H_0 not rejected	H_0 not rejected
	LB on $\tilde{a}^2$	H_0 rejected	H_0 rejected	H_0 rejected	H_0 rejected
	QQ plot of $\tilde{a}$	Worse fit	Better fit	Worse fit	Better fit
	Backtesting	6.04%	4.73%	5.99%	5.78%
Small Cap	ACF of $\tilde{a}$	Inadequate	Inadequate	Inadequate	Inadequate
	LB on $\tilde{a}$	H_0 rejected	H_0 rejected	H_0 rejected	H_0 rejected
	LB on $\tilde{a}^2$	H_0 not rejected*	H_0 not rejected**	H_0 not rejected	H_0 not rejected
	QQ plot of $\tilde{a}$	Worse fit	Better fit	Worse fit	Better fit
	Backtesting	5.88%	6.38%	5.88%	6.15%

Table 13: Results of the study ($\tilde{a}$: standardized residuals, LB: Ljung-Box test, and JB: Jarque-Bera test. *: H_0 not rejected for Q(10) and Q(15). **: H_0 not rejected for Q(10)).

models were adequate for the Large Cap. The Ljung-box tests, however, raise doubts about the variance equations for the Large Cap data. For the Small Cap, the ACF plots showed signs of autocorrelation, but the Ljung-box tests on the squared standardized residuals suggested that the models are adequate to describe the data nonetheless. It can be worthwhile to study where this discrepancy comes from.

For backtesting, an arbitrary length was chosen for window as 1 year or 2 years. The results show that there is little difference between the window of 1 year and the window of 2 years. For further studies, different lengths of windows can be chosen and tested to see if there are any significant differences in forecasting power, depending on the choice of a window length.

When we measured value at risk violations, we simply counted the numbers of violations. If one is interested in the patterns of these violations, one can study the intervals between violations. If one is interested in degrees of violations, one can measure the distances between VaR estimates and the realized values when violations occur.

5 References

Ruey S. Tsay, *Analysis of Financial Time Series, Third Edition*, John Wiley & Sons, Inc. 2010.

Alexander J. McNeil, Rudiger Frey and Paul Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University

Press, 2005.

Tim Bollerslev, *Generalized Autoregressive Conditional heteroskedasticity*, Journal of Econometerics 31, 1986.

Robert F. Engle, *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United kindom Inflation*, Econometrica, Volume 50, Issue 4, Jul 1982.

Peter J. Brockwell, Richard A. Davis, *Introduction to Time Series and Forecasting, Second Edition*, Springer, 2010.

Sven Erick Alm and Tom Britton, *Stokastik*, Liber, 2008.

John C. Hull, *Options, Futures, and Other Derivatives, Eighth Edition*, Pearson, 2012