June 2021

Application of GARCH Type Models in Forecasting Value at Risk

Katayoon Naimian
naimian@uwindsor.ca

Application of GARCH Type Models in Forecasting Value at Risk

By

Katayoon Naimian

A Major Research Paper
Submitted to the Faculty of Graduate Studies
through the Department of Economics
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts
at the University of Windsor

Windsor, Ontario, Canada

2021

Application of GARCH Type Models in Forecasting Value at Risk

by

Katayoon Naimian

APPROVED BY:

Y. Wang
Department of Economics

D. Li, Advisor
Department of Economics

May 13, 2021

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Dynamic risk management requires the risk measures to adapt to information at different times, such that this dynamic framework takes into account the time consistency of risk measures interrelated at different times. The value-at-risk (VaR) is one of the most well-known downside risk measures due to its intuitive meaning and broad range of applications in practice, however the static version embraces more popularity. This study investigates dynamic VaR modelling using four conditional volatility forecasting models: GARCH, TGARCH, GJRGARCH and IGARCH, and compares the forecasting output of the suggested GARCH-based volatility models. Since the predictive accuracy of Value-at-Risk (VaR) models is crucial for adequate capitalization, we perform backtesting on VaR forecasts and compare our suggested GARCH models, as well as different distributions for their innovations and confidence levels for VaR.

Key words: Value at risk (VaR), GARCH, volatility, dynamic, forecast, backtesting, risk management

ACKNOWLEDGEMENTS

It is a genuine pleasure to express my deep sense of gratitude to my supervisor, Dr. Dingding Li, for her valuable support and insightful feedback that pushed me to sharpen my thinking and brought my work to a higher level. I would also like to acknowledge Dr. Yuntong Wang as the second reader of this paper, and I am gratefully indebted for his very valuable comments. I also appreciate the supporting from everyone who played a role in my academic accomplishments.

TABLE OF CONTENTS

CHAPTER 1

1. Introduction

Uncertainty modelling is an integral part of financial practices, with applications in portfolio allocation, risk management, and financial contract pricing. The basic question of how much volatility we can anticipate for future prices of financial contracts has sparked a wide body of research into the statistical properties of price fluctuations and how we can use them to make better predictions. By taking into account some of the stylized effects of financial data, the ARCH (Autoregressive Conditional Heteroscedasticity) model proposed by Engle (1982) and GARCH (Generalized Autoregressive Conditional Heteroscedasticity) model developed by Bollerslev (1986), represent a landmark breakthrough in the field of financial modelling. Introducing the ARCH model won its author, Robert Engle, the Nobel Prize in 2003, among other honors. Its primary contribution is that it makes uncertainty a dynamic operation. Rather than assuming that future variance is constant, the ARCH and GARCH families of models recognize that it is a time-varying operation.

It can be very costly to ignore the ARCH effect and underplay short-term volatility. Risk management mechanisms in investment portfolios must evaluate the likelihood of a significant loss. An analyst underestimates the inner risk of financial contracts by assuming constant uncertainty and can be shocked by extreme unforeseen losses in the investment portfolio. Similarly, banking regulations such as Basel III require banks to disclose their portfolio risk level systematically and on a regular basis. Given that banks serve as liquidity centers, a miscalculation of uncertainty and risk will endanger the financial system's stability, as an unforeseen financial shock will cause banks to rapidly liquidate financial contracts and increase their cash position.

The value-at-risk (VaR) model, with all its challenges, is still the workhorse of risk management. One of VaR's major flaws is that it fails to account for volatility clustering, resulting in VaR limits being exceeded in serial dependence over time. As a consequence, during a crisis, the risk is undervalued. Combining VaR with GARCH models, which take conditional volatility into account, is a powerful way to solve this problem.

Value-at-Risk is rigorously empirical in that it employs mathematical techniques established in physics and engineering; however, it employs statistical techniques that depend on several assumptions. One of the most important of these assumptions is that the return on financial prices follows a normal distribution.

In mathematics, economics, and finance, the issue of risk measurement is an ancient one. Regulators and financial executives have been concerned with financial risk management for a long time, and this history includes some VaR-like terms. VaR, on the other hand, did not become a distinct term until the late 1980s. The stock market crash of 1987 was the catalyst. This was the first major financial crisis in which a large number of academically trained quants were in high enough positions to be concerned about the firm's long-term survival, as Jorion points out (2007).

Value-at-Risk (VaR) has become an industry standard indicator of business risk. It gives financial institutions details on the estimated worst loss at a given confidence level over a target horizon. Despite its significance and simplicity, there is no widely accepted formula for calculating a portfolio's VaR, and different models will result in substantially different risk measures. One of the most important considerations when using the VaR method to estimate market risk is the selection of the appropriate model; for example, a poorly defined model may be costly to the risk manager and lead to incorrect risk estimates. Furthermore, the massive losses suffered by financial institutions during the recent global financial crisis, in 2007–2008, have posed doubts about the risk models in place. These issues are directly related to the controversy between the financial sector, regulators, and academics about probabilistic market models for VaR forecasting, which can account for extreme events and increased volatility during financial market downturns.

The prediction of market volatility is critical in obtaining accurate VaR measures, particularly given its time-varying existence and some prominent stylized facts of stock returns. Indeed, there is a lot of evidence that small-scale price variations alternate with large-scale price variations; this is known as volatility clustering. To capture the volatility clustering effect, a number of econometric models have been proposed, the most commonly used of which is the GARCH model (Bollerslev, 1986). Former GARCH-based VaR models used the normal distribution resulting from the Brownian motion assumption as a benchmark process for explaining return developments, drawing criticism that such a distributional assumption can not adequately capture the frequency of severe asset price shocks, as well as the amplitude of these shocks, and sometimes leads to risk underestimation. More specifically, we want to see if GARCH-type models can model conditional volatility and VaR for global stock market indices under different error distribution assumptions. We use the standard GARCH model, GJR, IGARCH, and TGARCH among the conditional volatility models.

The aim of this study is to devise a method for accurately estimating Value at Risk in the face of time-varying volatility. Our time series data consists of S&P 500 index prices from 2013-2019 resulting in 1762 observations. We will provide a review of the related literature in the next section. In section 3 we provide information about the data and methodology used for this analysis. We present our empirical results in chapter 4 followed by a conclusion in chapter 5.

CHAPTER 2

2. Literature review

There is a substantial amount of literature on VaR and its forecasting efficiency under various model specifications. After the notorious 2008 financial crisis, more research has been done on strengthening and fixing the inadequacies of the VaR model, as well as its underlying volatility modelling.

Nieto and Ruiz (2016) compared the forecasting potential of various GARCH-based VaR models to their alternatives in an updated report. Surprisingly, the analysis found that forecasting outcomes are affected by the number of out-of-sample observations as well as the time span being studied. They concluded that no single model outperforms another in any situation. Furthermore, only the asymmetric EGARCH-based model with skewed Student's-t distribution can be approved under the various model tests.

When modelling financial market uncertainty, Bentes (2015) and Huang et al (2016) offer new justifications for accounting for long memory characteristics. The former applies various GARCH models to the forecasting of gold return volatility, demonstrating that the long memory FIGARCH outperforms its competitors. As a result, even after the financial crisis, the implementation of long memory models could boost empirical applications such as VaR. Nevertheless, Degiannakis et al (2013) examined returns from 20 established stock market indices and discovered that, despite the evidence of persistence in the volatility process, accounting for long memory does not always boost the resulting VaR forecasts. Updated research on the efficiency of various GARCH-based VaR models can be found in Ardia and Hoogerheide (2014) and Abad et al (2014).

According to So and Yu (2006), different GARCH-based VaR models perform better at different significance levels. These results point to a new line of inquiry, which might look at the dominance and success of these VaR models over time periods with varying market conditions (causing a change in market regime). Given previous findings that the long-memory FIEGARCH was the dominant model for VaR forecasts in the South African industry, it is not immediately clear if it remains the preferred model when evaluated through individual sub-periods with varying market conditions.

Tabasi et al. (2019) used GARCH models to model the volatility-clustering feature and found that using the t-student distribution function instead of the Normal distribution function improved model parameter estimation.

Based on MSCI World Index data from 2006 to 2009, Husng et al. (2015) discovered that ARMA (1,1)-GARCHM (1,1) performs the best in terms of violation measures.

Emenogu et al. (2020) discovered that the persistence of the GARCH models is robust, with the exception of a few cases where IGARCH and EGARCH were unstable. The SGARCH and GJRGARCH models also failed to converge for student t innovation; the mean reverting number of days for returns varied between models.

For all confidence levels, Altun (2018) found that GARCH models listed under the TSLx innovation distribution produce more accurate VaR forecasts than other competing models.

Slim et al. (2017) claimed that in developed markets, the related models show signs of long memory, suggesting that the FIGARCH model is preferable to the GARCH and GJR models. In frontier and emerging markets, the GJR and GARCH are the most important specifications for capturing risk. This means that when analysing frontier markets, risk managers should favour models that account for asymmetry.

Okpara (2015) used the VaR method to conduct a risk analysis of the Nigerian stock market. The study concluded that using the Akaike information criterion (AIC), the EGARCH model with student t innovation distribution could provide a more reliable estimate of VaR, and using the probability ratio tests of proportional failure rates on VaR derived from the EGARCH model.

Apart from the above contradictory observations, there are ongoing questions about the accuracy and validity of GARCH-based VaR models. Hafer and Sheehan (1989), for example, looked at the sensitivity of VaR forecasts to different lag structures in the underlying time series. They conclude that VaR forecasts are responsive to changes in lag structure, and that the relative accuracy of VaR forecasts is strongly influenced by the forecast horizon chosen.

According to Elenjical et al (2016), different VaR models perform better depending on the state or behaviour of the market when examined over periods of different market conditions. Similarly, Ng Cheong Vee et al. (2014) have discovered that if markets are classified, common models could be able to forecast their VaR accurately.

According to Bali and Cakici (2004), stock size, liquidity, and VaR may explain cross-sectional variance in expected returns better than beta and total volatility, and that the relationship between average returns and VaR is robust for various investment horizons and loss probability levels, with VaR having additional explanatory power for stock returns.

CHAPTER 3

3. Data and Methodology

Generally speaking, the development of risk measurement goes through three stages: firstly, the traditional risk measurement stage with variance and risk factors as the main indicators. Secondly, the modern risk measurement stage represented by the VaR, and finally risk measurement stage represented by Conditional VaR (CVaR).

In this study, we focus on VaR method. We use the data on daily S&P 500 closing prices extracted from Yahoo finance database and in order to avoid possible structural breaks, we extract the data from 2013-2019 period resulting in 1762 daily observations during a time span of 7 years. The Standard and Poor's 500, or simply the S&P 500, is a free-float, weighted stock market index that includes 500 of the largest companies listed on US stock exchanges. It is one of the most widely tracked stock market indices. We use Rstudio software for programming and modeling data to provide our results.

In this study, we use log return formula to obtain the return series of S&P 500 closing prices:

$$r_t = \ln p_t - \ln p_{t-1}$$

Where r_t denotes the daily log returns and p_t is the daily closing index price.

3.1. Value at Risk (VaR)

VaR is defined as the predicted loss at a specific confidence level over a given period of time. The VaR concept has emerged as the most prominent measure of downside market risk. It places a lower bound on losses at a given confidence level over a given forecast horizon. Thus, assuming that the VaR model is correct, realized losses will exceed the VaR threshold with only a small target probability α, typically chosen between 1% and 5%.

To obtain VaR we need to determine the following three factors: the length of the holding period, the size of the confidence interval and the period of the observation.

i. The length of the holding period is used to decide how long the maximum loss of assets must be calculated. This refers to whether the managers are worried about the assets' value at risk in a day, a week, or a month.

ii. The confidence level, which is the frequency of possible confidence intervals that contain the true value of their corresponding parameter.

iii. The observation period, also known as the historical window, is the overall length of time for the observations. For example, to consider the weekly returns volatility of an asset, we may choose an observation period of the previous 6 months or 1 year. The longer the historical data, the better, in order to avoid the influence of the business cycle. However, the longer the period, the greater the chance of market structural changes resulting in lower accuracy in representing future actual results.

More specifically, according to the definition of VaR, conditional on the information until time t - h, the VaR on time t of one unit of investment is the α quantile of the conditional return distribution, that is:

$$VaR_t = q_\alpha(r_t|\mathcal{F}_{t-h}) = \inf\{x \in \mathbb{R}|P(r_t \leq x|\mathcal{F}_{t-h}) \geq \alpha\}$$

where q_α denotes the quantile function, r_t is the index return in period t, and $\mathcal{F}_{t-h}$ designates the information available at date t-h. When the expected returns, r_t , are assumed to follow a location-scale distribution, they are regarded as a function of an innovation process, ε_t. Therefore, under the specified probability level α, if the return is negative or we have a loss, the probability with which the observed loss exceed estimated loss can be expressed as follows:

$$\Pr\big(r_t \leq VaR_\alpha(r_t)\big) = \alpha$$

According to our results we use an autoregressive model of order 1, AR (1) to model the returns process. The AR (1) model is defined as follows:

$$r_t = K + ar_{t-1} + \varepsilon_t$$

Where ε_t denotes the innovation at time t, K is a constant and a is the AR (1) coefficient. For error terms we have: $E(\varepsilon_t) = 0$

Now let $\mu_t = K + ar_{t-1}$, then we have: $r_t - \mu_t = \varepsilon_t \Longrightarrow \frac{r_t - \mu_t}{\sigma_t} = \frac{\varepsilon_t}{\sigma_t}$

Thus, we have: $r_t - \mu_t = \varepsilon_t = \sigma_t z_t$ where the sequence $z_t = \frac{r_t - \mu_t}{\sigma_t}$ represent the standardized residuals from some probability distribution, D, with mean zero and unit variance. Thus, we have: $E(\varepsilon_t^2) = \sigma_t^2$ and $\varepsilon_t \sim D(0, \sigma_t^2)$.

$$r_t = \mu_t + \sigma_t z_t$$

Now, having the equation for r_t we can obtain the equation for VaR as follows:

$$VaR_t^\alpha(r_{t+1}) = \mu_{t+1} + \sigma_{t+1} VaR_t^\alpha(z)$$

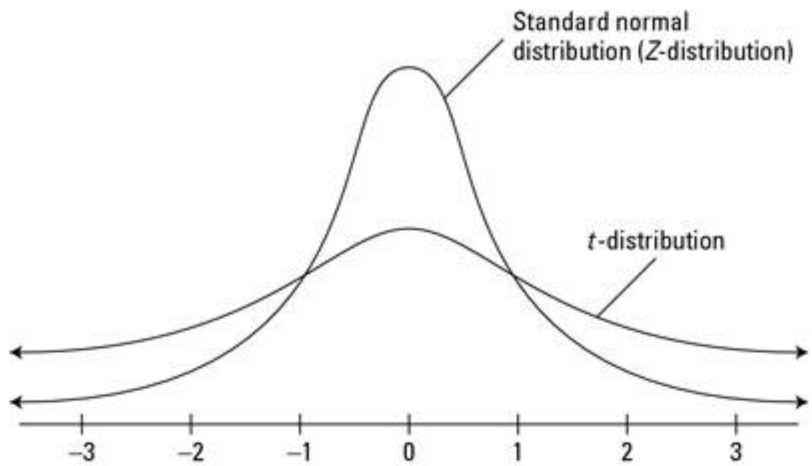
or

$$VaR_t^\alpha = \mu_t + \sigma_t q_z^\alpha$$

where μ_t and σ_t are calculated recursively using the AR(1) and GARCH(1,1) equations and q_z^α is α percent quantile from the fitted distribution to z_t.

As we can see from above equation for VaR, there are three components that result in an estimate or returns. The first component is μ_t, which depends on how we model the mean of returns. However, as we will see in chapter 4, the mean of our return series is about zero, meaning that the accuracy of modeling the mean may not impose a great impact on modeling returns, and this fact is true for most of the other financial data. While the main component that can affect our return series and other financial data is σ_t, since financial data are highly affected by their volatility. Thus, the more we can improve our predictions on volatility, the more we can achieve higher accuracy in predicting the VaR. The last component z_t, is related to the distribution of residuals. Thus, selecting a distribution which is a better representative of our data, will lead to a better forecast for our returns. In this study, we mainly focus on improving the volatility forecasts by comparing the results on four different GARCH models and we consider two distributions for residuals, being normal and student-t distribution, in order to provide related comparisons.

As we can see from graph the below, although normal distribution and student-t distribution look similar in that they are both centered at zero and have a basic bell-shape, but t distribution is shorter and flatter around the center than the normal distribution. Its standard deviation is proportionally larger compared to the normal, which is why we see the fatter tails on each side. Since fat tails are a well-known characteristic of financial data, we expect that considering a student-t distribution may provide better results in our modeling.



3.1.1. Popular approaches of VaR calculation

"One of the most difficult aspects of calculating VaR is selecting among the many types of VaR methodologies and their associated assumptions." (Minnich, 1998)

Although there are many different methods for calculating VaR, there are three main methods that are mentioned in the documented regulations related to financial services mainly the banking industry.

i. Historical simulation Method

The historical method simply re-organizes actual historical returns, putting them in order from worst to best. It then assumes that history will repeat itself, from a risk perspective. It picks an α quantile of the ordered historical series as the α% VaR.

ii. Delta-Normal (Variance-Covariance) Method

This method assumes that stock returns are normally distributed. In other words, it requires that we estimate only two factors—an expected (or average) return and a standard deviation—which allow us to plot a normal distribution curve. The idea behind the variance-covariance is similar to the ideas behind the historical method—except that we use the familiar curve instead of actual data. The advantage of the normal curve is that we automatically know where the worst 5% and 1% lie on the curve. They are a function of our desired confidence and the standard deviation.

iii. Monte Carlo Simulation

The third method involves developing a model for future stock price returns and running multiple hypothetical trials through the model. A Monte Carlo simulation refers to any method that randomly generates trials, but by itself does not tell us anything about the underlying methodology.

In this study, we provide some comparisons on our suggested method and the first two approaches mentioned above.

3.2. Modeling volatility

Financial econometrics and financial time series analysis help us understand how prices behave and how this insight can help us mitigate risk and make better decisions. This is done using time series models for forecasting, option pricing and risk management. Volatility modeling requires two main steps:

- Specify a Mean equation (e.g. ARMA, AR, MA, ARIMA)
- Model a Volatility equation (e.g. GARCH, ARCH)

To determine the mean equation, we use the Box-Jenkins method which consists of three major steps:

- Identification
- Estimation
- Diagnostic Checking

Following the above-mentioned procedure, in order to identify the model we use the Akaike Information Criterion (AIC). AIC estimates the quality of each model relative to each of the other models.

$$AIC = \ln\frac{\sum \hat{\varepsilon}^2}{T} + \frac{2k}{T}$$

Where: $\sum \hat{\varepsilon}^2$ is the sum of squared residuals, T is the number of observations and k is the number of model parameters (p+q+1).

It is obvious that when extra lag parameters are added to the model Sum Squared of Residuals decreases but overfitting problems may occur. AIC deals with both the risk of overfitting and underfitting. The model with the lowest AIC will be selected.

The results in this study approve an ARMA(1,0) model for the mean of the return series. Thus for our volatility modeling an AR(1) model for the mean returns is assumed in all GARCH type models.

The procedure for diagnostics checking includes observing residual plot and its ACF and PACF diagram, and check Ljung-Box test result. If ACF and PACF of the model residuals show no significant lags, the selected model is appropriate.

To further test the hypothesis that the residuals are not correlated, we perform Ljung-Box test.

$$Q_{LB} = T(T+2)\sum_{i=1}^{m}\frac{\hat{\rho}_i^2}{T-i}$$

The Q_{LB} statistic follows asymmetrically a χ^2 distribution with m-p-q degrees of freedom. The null hypothesis refers to $H_0: \rho_1 = \rho_2 = \cdots = \rho_m = 0$

Previously we mentioned our returns equation as: $r_t = \mu_t + \varepsilon_t$. Thus, we assume that the return series is decomposed into two parts, where μ_t is the predictable component and ε_t is the unpredictable part or innovation process.

We defined the unpredictable component as: $\varepsilon_t = \sigma_t z_t$ where z_t is a sequence of independently and identically distributed random variables with zero mean and variance equal to 1. The conditional variance of ε_t is σ_t, a time-varying function of the information set at time t−1. The next step is to define the second part of the error term decomposition, which is the conditional variance, σ_t. For such a task, we can use a GARCH type model with one lag on ARCH and GARCH effects.

GARCH

Developed by Bollerslev (1986), the conditional variance in the GARCH(1,1) specification with AR(1) mean model is represented by:

$$r_t = K + ar_{t-1} + \varepsilon_t$$

And by setting $\mu_t = K + ar_{t-1}$ we have:

$$\varepsilon_t = r_t - \mu_t, \qquad \varepsilon_t \sim D(0, \sigma_t^2)\ iid$$

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Where the parameter α_1 is the ARCH parameter and β_1 is the GARCH parameter, and the conditional variance process is positive and stationary if the following conditions hold:

$$\omega > 0, \alpha_1 > 0, \beta_1 > 0\ and\ \alpha_1 + \beta_1 < 1$$

The restriction on ARCH and GARCH parameters (α_1, β_1) suggests that the volatility is finite. The GARCH(1,1) model can only handle short memory in the volatility process since its autocorrelation function decays rapidly with an exponential rate of $\alpha_1 + \beta_1$.

TGARCH

The threshold GARCH model, developed by Zakoian (1993), is another model used to handle leverage effects, and a TGARCH(1,1) model is given by the following:

$$\sigma_t = \omega + \alpha_1 \sigma_{t-1}(|z_{t-1}| - \eta_{11} z_{t-1}) + \beta_1 \sigma_{t-1}$$

and $\alpha_1\ and\ \beta_1$ are nonnegative parameters and $|\eta_{11}| \leq 1$ satisfying conditions similar to those of GARCH models.

GJRGARCH

In financial markets, it is often the case that downward movements in the market are followed by higher volatility than upward movements of the same magnitude (Engle and

Ng, 1993). This asymmetry can be modeled using the GJR model of Glosten et al. (1993), where the impact of ε_{t-1}^2 depends on the sign of the shock, that is:

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2 + \gamma_1\varepsilon_{t-1}^2 I_{t-1}$$

Where I_{t-1} is equal to unity if $\varepsilon_{t-1} < 0$ and zero otherwise. The conditional volatility is positive when parameters satisfy $\omega > 0,\ \alpha_1 + \gamma_1 > 0$, $\beta_1 > 0$ and γ_1 represents the leverage term. The process is covariance stationary if $\alpha_1 + \beta_1 + \frac{1}{2}\gamma_1 < 1$. The impact of shocks on conditional variance is asymmetric if γ_1 is significantly different from zero. This model allows positive shocks to have a stronger effect on volatility than negative shocks (Rossi 2004).

IGARCH

Integrated GARCH (IGARCH) models are unit-root GARCH models. The IGARCH(1, 1) model is specified in Tsay (2005) as

$$\sigma_t^2 = \omega + (1-\beta_1)\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2$$

Where 0 <β_1 < 1. The model is also an exponential smoothing model for the ε_t^2 series. To see this, we rewrite the model by repeated substitution as:

$$\sigma_t^2 = (1-\beta_1)(\varepsilon_{t-1}^2 + \beta_1\varepsilon_{t-1}^2 + \beta_1^2\varepsilon_{t-3}^3 + \cdots)$$

which is a well-known exponential smoothing formation in which β_1 is the discounting factor (Tsay 2005).

3.2.1. Evaluating accuracy of the model

Mean Absolute Error (MAE) and Mean squared error (MSE) are two of the most common metrics used to measure accuracy for continuous variables.

Mean Absolute Error (MAE): MAE measures the average magnitude of the errors in a set of predictions, without considering their direction. It’s the average over the test sample of the absolute differences between prediction and actual observation where all individual differences have equal weight. It measures the average of the residuals in the dataset.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\sigma_t^2 - \hat{\sigma}_t^2|$$

Mean squared error (MSE): MSE is a quadratic scoring rule that also measures the average magnitude of the error. It's the square of the average of squared differences between prediction and actual observation. It measures the variance of the residuals.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (\sigma_t^2 - \hat{\sigma}_t^2)^2$$

Both MAE and MSE express average model prediction error in units of the variable of interest. Both metrics can range from 0 to ∞ and are indifferent to the direction of errors. They are negatively oriented scores, which means lower values are better as an indication for more accuracy in the model.

In MSE, since the errors are squared before being averaged, it gives large errors a lot of weight. As a result, when large errors are especially undesirable, the MSE should be more useful. On the other hand, when the total effect is proportionate to the real increase in error, MAE is more useful. For example, if error values increase from 3 to 6, the effect on the result is doubled. It is more common in the financial industry, where a loss of six is twice as bad as a loss of three. In contrast to a non-differentiable function like MAE, MSE is a differentiable function that makes mathematical operations easier. MAE is more robust to data that contains outliers.

3.3. Back-testing procedure

A historical backtest is a good way to check the model's performance. In backtesting a risk model, we compare the estimated VaR with the actual return over the period. A VaR exceedance occurs when the return is more negative than the VaR.

In order to back-test the accuracy for the estimated VaRs, we compute the empirical failure rates. By definition, the failure rate is the number of times returns (in absolute values) exceed the forecasted VaR. If the model is correctly specified, the failure rate should be equal to the specified VaR's level. In this study, the backtesting VaR is based on Kupiec's (1995) and Christoffersen (1998) for unconditional and conditional coverage tests. A variety of backtesting methods have been proposed to gauge the accuracy of VaR estimates.

Backtesting is a formal statistical framework that consists in verifying if actual trading losses are in line with model generated VaR forecasts and relies on testing over VaR violations (also called the hit). A violation is said to occur when the realized trading loss exceeds the VaR forecast. We briefly present the backtesting methods used in our empirical assessment of VaR models regarding the following properties

i. Frequency: The unconditional coverage (UC) test (Kupiec, 1995) is the industry standard, owing to the fact that it is implicitly embedded in the Basel Committee on Banking Supervision's (2006, 2009) "traffic Light" scheme, which is still used by banking regulators as the reference backtest methodology. The test entails determining whether the realized coverage rate (α) of the VaR for a backtesting sample of T non-overlapping observations is equal to the theoretical coverage rate (α). This is the same as determining if the hit variable $I_t(\alpha)$, which takes values of 1 if the loss exceeds the stated VaR measure and 0 otherwise, has a binomial distribution with parameter α. Under the UC hypothesis, the likelihood ratio (LR) test statistic follows a χ^2 distribution with one degree of freedom. That is:

$$LR_{UC}(\alpha) = -2\ln[(1-\alpha)^{T-N}\alpha^N] + 2\ln\left[\left(1-\frac{N}{T}\right)^{T-N}\left(\frac{N}{T}\right)^N\right] \sim \chi^2(1)$$

where N is the number of VaR violations.

ii. Independence: By checking for the independence (IND) of the sequence of VaR violations, the unconditional backtesting framework is improved, resulting in a combined conditional coverage test (CC). The LR test of Christoffersen is used to determine risk models under the joint hypothesis of IND and right UC (1998). The Christoffersen (1998)'s LR test for independence against an explicit first-order Markov alternative is given by:

$$LR_{IND}(\alpha) = -2\ln\left[(1-\frac{N}{T})^{T-N}(\frac{N}{T})^{N}\right] + 2\ln\left[(1-\hat{\pi}_{01})^{n_{00}}\hat{\pi}_{01}^{n_{01}}(1-\hat{\pi}_{11})^{n_{01}}\hat{\pi}_{11}^{n_{11}}\right] \sim \chi^2(1)$$

where n_{ij}; i, j=0,1 is the number of times we have $I_t(\alpha)$=j and $I_{t-1}(\alpha)$=i with $\hat{\pi}_{01} = n_{01}/(n_{00}+n_{01})$ and $\hat{\pi}_{11} = n_{11}/(n_{10}+n_{11})$. The LR statistic for the CC test is then given by:

$$LR_{cc}(\alpha) = LR_{UC}(\alpha) + LR_{IND}(\alpha) \sim \chi^2(2)$$

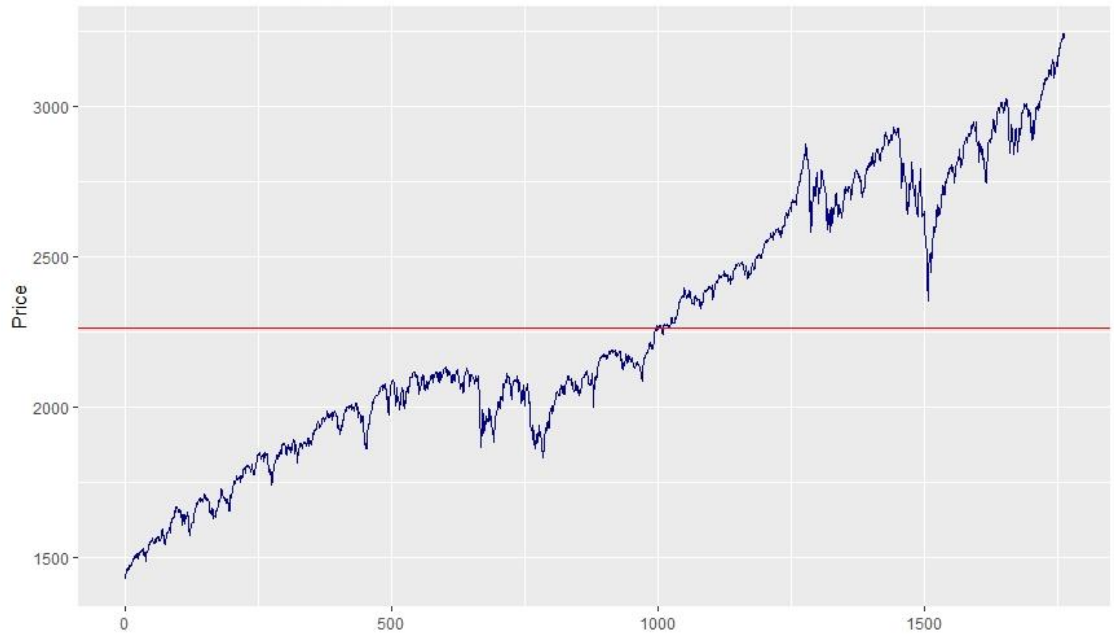
The unconditional coverage test of Kupiec examines whether the sum of expected versus actual exceedances given the tail probability of VaR occur as estimated, while Christoffersen's conditional coverage test examines both the unconditional coverage and the independence of the exceedances. Both the joint and the separate unconditional test are reported since it is always possible that the joint test passes while failing either the independence or unconditional coverage test.

CHAPTER 4

4. Empirical Results

In this section we provide the results. As we can see from figure 1, the time series of daily index price is not stationary. Non-stationary processes have means, variances and covariances that change over time. Using non-stationary time series data leads to unreliable forecasting. A stationary process is mean reverting, i.e., it fluctuates around a constant mean with constant variance. In order to resolve this issue, we mostly use differencing. Thus, in the first step, we obtain the log return series for the S&P 500 index. By calculating the log returns we will employ a log transform and first difference and then we check whether the non-stationarity issue is solved.

Figure1- Daily S&P 500 closing prices – 2013-2019



Red line denotes the average closing price for this particular timeframe. The time series plot appears in clusters, high in certain periods and low in certain periods. It evolves over time in a continuous manner and is thus, volatile.

Table 1- Augmented Dicky-Fuller test results for S&P 500 prices

Model	Dickey-Fuller Critical value
Type 1: No constant and no trend	2.6263
Type 2: With constant but no trend	1.4301
Type 3: With constant and with trend	4.2748

Alternative hypothesis: stationary

Although it can be evident from figure 1 that the time series for index prices is non-stationary, but we also run a Dicky-Fuller test to check for stationarity. Table shows that the null hypothesis of non-stationary is not rejected for all 3 types of Dicky-Fuller test. So the index price series is non-stationary and we need to employ a first difference transformation by using the return of the index prices.

4.1. S&P 500 returns series overview

Based on figure 2, we can see that the return series looks stationary. The excess kurtosis and fat tails are obvious in the histogram, but we can confirm numerically that the kurtosis of the empirical distribution of our sample (3.581421) exceeds that of a normal distribution (which is equal to 3). Table 1 shows the descriptive statistics of the return series.

Table 2- Descriptive statistics of return series

Number of Observations	1761
Minimum	-0.041843
Maximum	0.048403
Mean	0.000463
Median	0.000593
Variance	0.000066
Standard Deviation	0.008102
Skewness	-0.513383
Excess Kurtosis	3.581421
Jarque Bera	1022.6

Jarque-Bera statistic is significant at 0.01 level, rejecting the null hypothesis of normality

From the basic statistics of the log return of the index prices, we observe that the mean is about zero and the distribution of log returns has large kurtosis (fat tails). We observe this further using histogram and Q-Q plot. The negative skewness and the high positive kurtosis indicate that the distribution of the return series has a long left tail and is leptokurtic. Jarque-Bera (JB) statistics also reject the null hypothesis of normal distribution at the 1% level of significance.

Figure 2 – Time series and histogram of returns

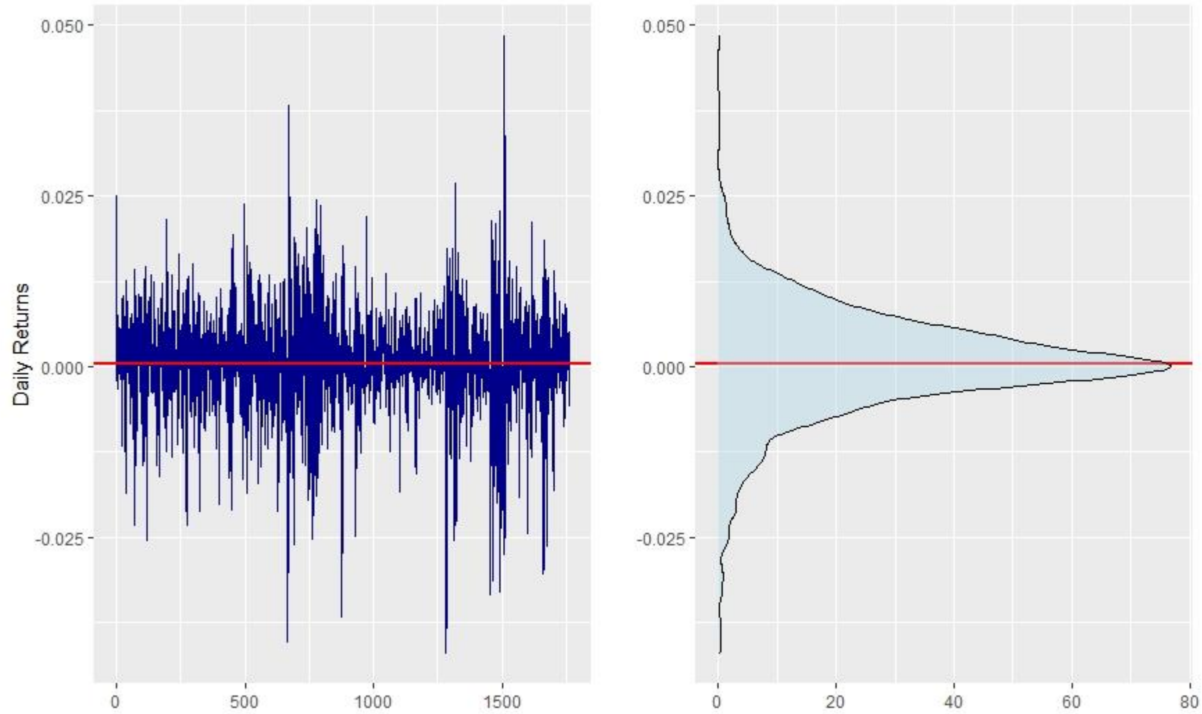
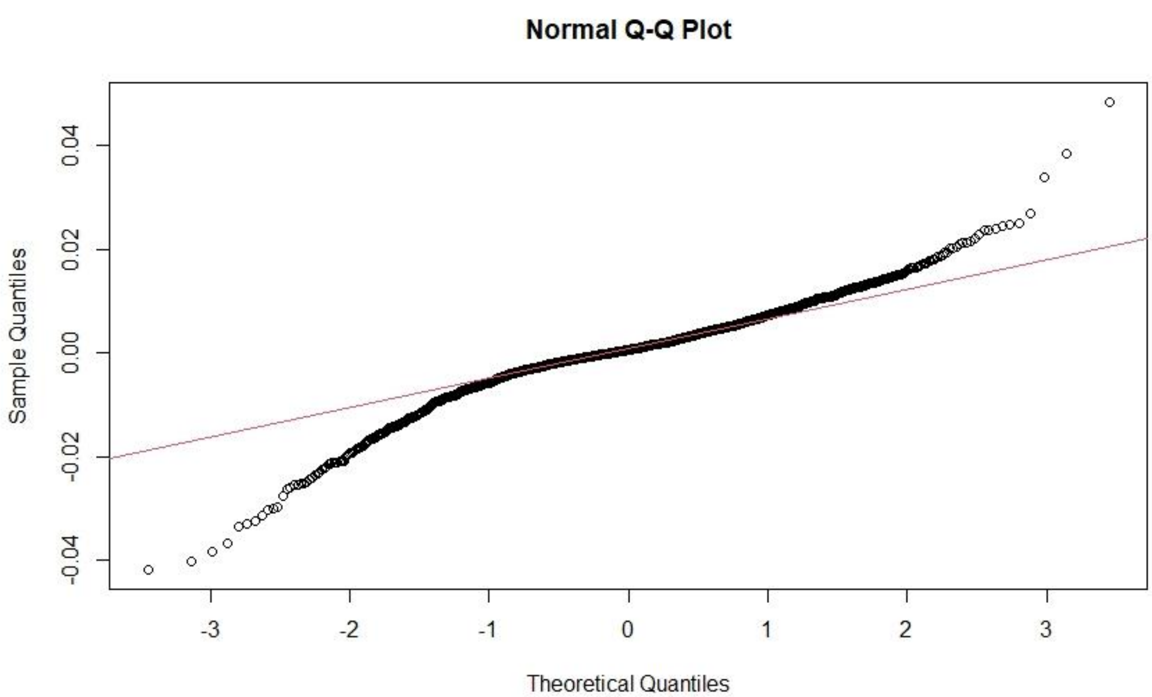
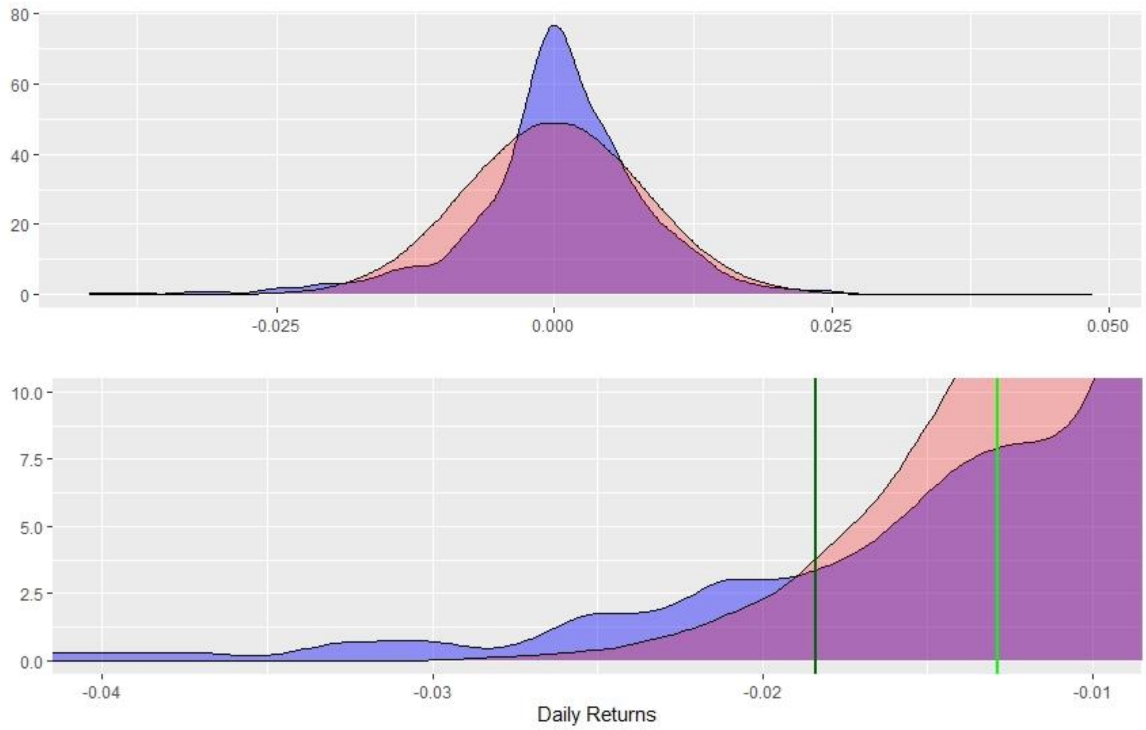


Figure 3 – Q-Q Plot for the returns



The Q-Q plot of the returns also show the same result.

Figure 4 – Comparison with normal distribution



On figure 4, Density plots are shown for stock returns (*blue*) and normally distributed data (*red*). Vertical lines of the lower plot represent the normal corresponding quantile for $\alpha = 0.05$ (*light green*) and $\alpha = 0.01$ (*dark green*). The lower plot indicates that for 95% significance, normal distribution usage may overestimate the value at risk. However, for 99% significance level, a normal distribution would underestimate the risk. In the parametric method to calculate a static VaR, Normal Distribution is adopted to capture the market risk under general market conditions. One key problem of a VaR that resulted from considering a normal distribution and providing a static VaR is that it does not properly account for volatility clustering, which means that VaR limits are breached in serial dependence across time. As a result, risk is underestimated during a crisis. A powerful approach to solve this problem is to combine VaR with GARCH models, which take conditional volatility into account.

4.2. Modelling the mean

4.2.1. Stationarity

To verify the stationarity of the returns, we utilize the Augmented Dickey-Fuller test where the null hypothesis indicates non-stationary time series.

Table 3 - Augmented Dickey-Fuller Test for S&P 500 returns

Model	**Dickey-Fuller Critical value**
Type 1: No constant and no trend	-15.4***
Type 2: With constant but no trend	-15.7***
Type 3: With constant and with trend	-15.6***

Alternative hypothesis: stationary

*** significant at 0.01 level

As we can see the null hypothesis of non-stationarity is rejected at 0.01 level, thus we can consider the return series as stationary.

4.2.2. Identifying the mean model

Table 4: Selecting the ARMA model

ARMA Model	AIC
ARMA(2,2) with non-zero mean	Inf
ARMA(0,0) with non-zero mean	-11960.19
ARMA(1,0) with non-zero mean	-11967.71
ARMA(0,1) with non-zero mean	-11959.53
ARMA(0,0) with zero mean	-11956.45

ARMA(2,0) with non-zero mean	-11966.83
ARMA(1,1) with non-zero mean	-11965.82
ARMA(2,1) with non-zero mean	-11966.60
ARMA(1,0) with zero mean	-11964.01

We can see that ARMA(1,0) with non-zero mean has the lowest AIC: -11967.71. With the process above we computed AIC scores for various ARMA models and we infer that the appropriate model is a **1-order Autoregressive (AR(1)).**

4.2.3. Estimating the mean model

Using AR(1) as the selected model, the results are as follows:

Table 5 - AR(1) estimation results

AR(1)	-0.0268	
	(0.0239)	
Cons	0.0005	**
	(0.0002)	
log likelihood	5982.73	
AIC	-11959.5	
BIC	-11943	

*,**, *** denotes significant at 10%, 5% and 1% respectively
statndard errors are in brakets

Therefore, the mean model can be described as:

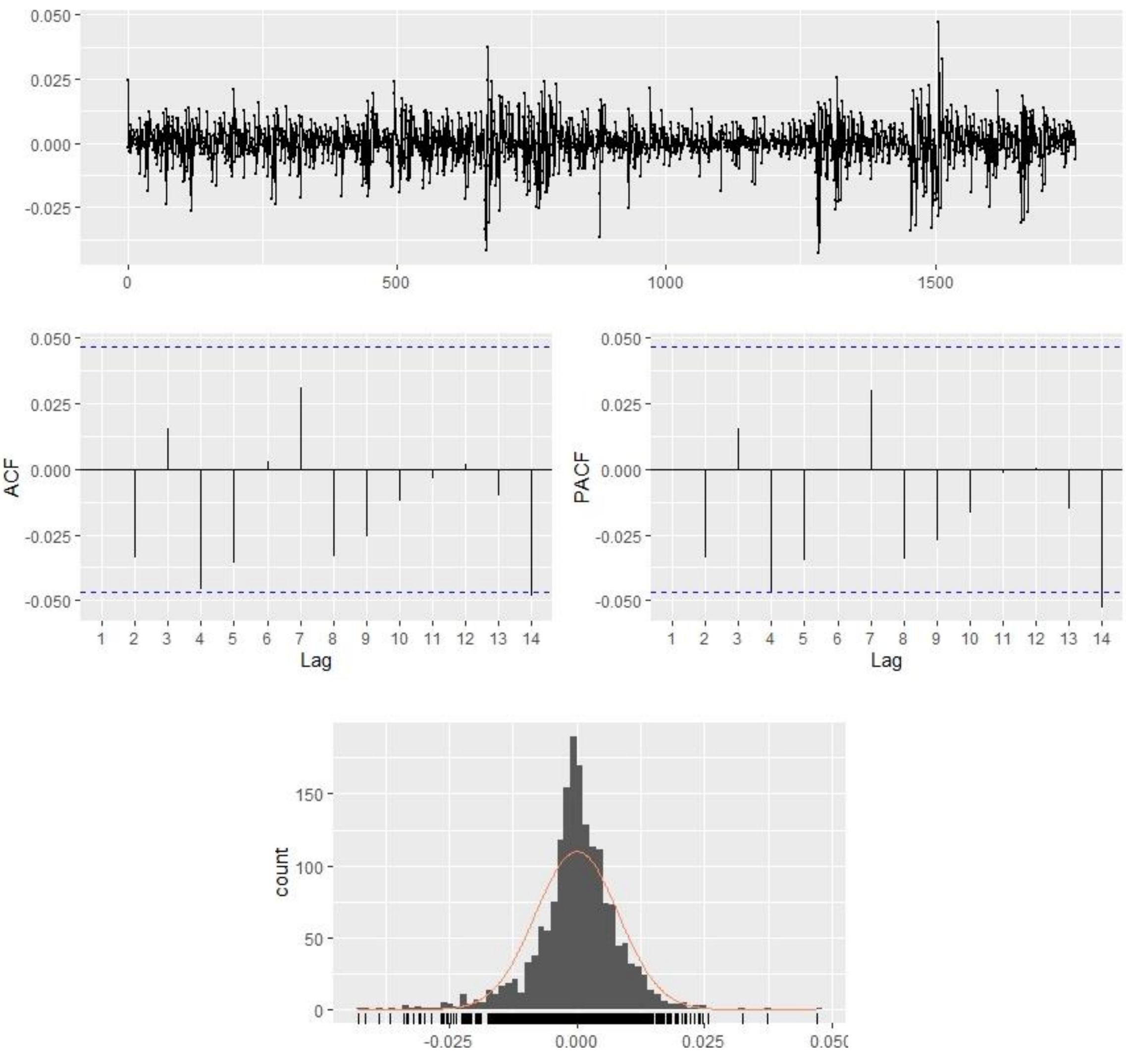
$$\widehat{r_t} = 0.0005 - 0.0268 r_{t-1}$$

Although AR(1) has the lowest information criterion, but based on the results we see that AR(1) coefficient is not significant. However we include it in our modeling since considering an AR(1) for the mean model is suggested in the related literature with this study.

4.2.4. Diagnostic Checking for mean model

We derive the residuals from the fitted AR(1) model and run the diagnostic tests on residuals. Both ACF and PACF plots are similar, and autocorrelations seem to be equal to zero. The lower plot in figure 5 represents the histogram of the residuals compared to a standard normal distribution.

Figure 5 – AR(1) residuals



To check whether the residuals are correlated or not, we perform Ljung-Box test.

Table 6 - Box-Ljung test for AR(1) residuals

X-squared	df	p-value
17.879	12	0.1194

H0: ρ1=ρ2=⋯=ρm=0

From Ljung Box test result, we observe that the residuals are not correlated as the p-value is greater than 0.05 and hence we cannot reject the null hypothesis of no autocorrelation and we conclude that the residuals behave like white noise and there is no indication of pattern that might be modeled. Although ACF & PACF of residuals have no significant lags, the time series plot of residuals shows some cluster volatility meaning that the volatility changes over time and its degree shows a tendency to persist, i.e., there are periods of low volatility and periods where volatility is high, which is a common behavior of GARCH process. Since the model does not represent recent changes or integrate new details, it is important to note that ARMA is a tool for linearly modelling data, and the forecast width remains constant. The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model is used to model volatility.

We perform an Arch test to check for Arch effects in residuals:

Table 7 - ARCH LM-test

Chi-squared	Df	p-value
132.03	1	< 2.2e-16

Null hypothesis: no ARCH effects

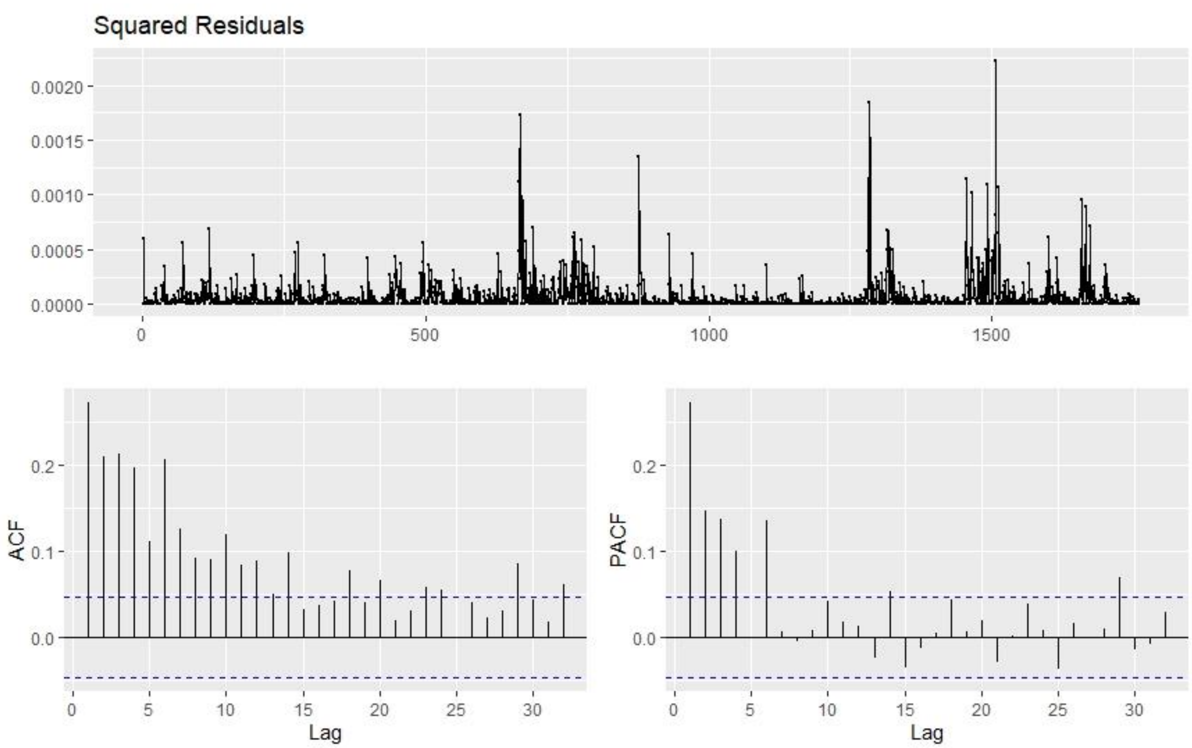
Because the p-value is less than 0.05, we reject the null hypothesis and conclude that ARCH(1) effects exist.

4.3. Volatility modeling

The most commonly used GARCH model, and one that is usually appropriate for financial time series as well, is a GARCH(1,1) model. Thus, for maintaining simplicity and concordant with the related literature we consider four GARCH type models with one lag on ARCH and GARCH effects, to define the second part of the error term decomposition, which is the conditional variance.

Granger and Andersen (1978) discovered that although the ARMA residuals themselves may not appear to be correlated over time, some of the series modelled by Box and Jenkins (1976) have autocorrelated squared residuals, and thus proposed that the ACF of the squared time series could be useful in defining nonlinear time series. The ACF and PACF of the squared residuals, according to Bollerslev (1986), are useful in identifying and checking GARCH behavior. Thus, the GARCH process is valid when the squared residuals from ARMA model are correlated. ACF and PACF plots clearly indicate significant correlation.

Figure 6 – Squared residuals from AR(1) model



We can also perform Ljung-Box test to check for the existence of correlation in squared residuals.

Table 8 - Box-Ljung test for AR(1) squared residuals

X-squared	**df**	**p-value**
585.57	12	2.2e-16

H0: ρ1=ρ2=⋯=ρm=0

We can see that the null hypothesis of no correlation is rejected, so there is correlation in squared residuals.

We can check for the conditional heteroscedasticity of the residuals by running the GARCH model and then check for the significance of α_1 and β_1 parameters.

Table 9 – Fitted GARCH (1,1) with AR (1) mean model

	GARCH		TGARCH		IGARCH		GJRGARCH	
	Normal	Student-t	Normal	Student-t	Normal	Student-t	Normal	Student-t
Panel A: Estiamtion results								
μ	0.0008 ***	0.0009 ***	0.0005 ***	0.0006 ***	0.0009 ***	0.0009 ***	0.0005 ***	0.0006 ***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
AR(1)	-0.0670 ***	-0.0689 ***	-0.0655 **	-0.0579 ***	-0.0704 ***	-0.0693 ***	-0.0763 ***	-0.0594 ***
	(0.0246)	(0.0227)	(0.0302)	(0.0218)	(0.0258)	(0.0239)	(0.0284)	(0.0228)
ω	0.0000 ***	0.0000	0.0005 ***	0.0004 ***	0.0000	0.0000	0.0000 ***	0.0000 ***
	0.0000	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	0.0000	0.0000
α1	0.2066 ***	0.2066 ***	0.1297 ***	0.1333 ***	0.2798 ***	0.2340 ***	0.0000	0.0000
	(0.0214)	(0.0314)	(0.0183)	(0.0148)	(0.0648)	(0.0727)	(0.0058)	(0.0193)
β1	0.7262 ***	0.7630 ***	0.8376 ***	0.8446 ***	0.7202 NA	0.7660 NA	0.7824 ***	0.7851 ***
	(0.0280)	(0.0692)	(0.0205)	(0.0174)	(NA)	(NA)	(0.0183)	(0.0151)
η11			1.0000 ***	1.0000 ***				
			(0.1594)	(0.1203)				
γ1							0.3081 ***	0.3386 ***
							(0.0504)	(0.0480)
Panel B: Diagnostics tests								
Weighted Ljung-Box								
Lag[1]	0.9140	0.8656	0.6808	0.1718	0.7980	0.7724	1.5600	0.2228
	(0.3390)	(0.3522)	(0.4093)	(0.6785)	(0.3717)	(0.3795)	(0.2116)	(0.6369)
Lag[2*(p+q)+(p+q)-1][2]	0.9167	0.8657	0.6813	0.1738	0.7982	0.7725	1.5690	0.2296
	(0.7887)	(0.8171)	(0.9055)	(0.9993)	(0.8523)	(0.8649)	(0.4000)	(0.9980)
Lag[4*(p+q)+(p+q)-1][5]	1.8738	1.7795	1.2291	0.7093	1.6633	1.6477	2.4750	1.1071
	(0.7473)	(0.7737)	(0.9069)	(0.9807)	(0.8051)	(0.8092)	(0.5735)	(0.9295)
Jarque-Bera	1078.7 ***	1080.4 ***	1077.4 ***	1070.5 ***	1081.8 ***	1080.8 ***	1087.2 ***	1071.8 ***
AIC	-7.0377	-7.1096	-7.1168	-7.1700	-7.0297	-7.1096	-7.0883	-7.1520
BIC	-7.0221	-7.0910	-7.0981	-7.1482	-7.0173	-7.0941	-7.0696	-7.1302
SIC	-7.0377	-7.1097	-7.1168	-7.1700	-7.0297	-7.1096	-7.0883	-7.1520
HQIC	-7.0319	-7.1028	-7.1099	-7.1619	-7.0251	-7.1039	-7.0814	-7.1439

Robust statndard errors are in brakets in panel A
Weighted Ljung-Box Test on Standardized Residuals (H0 : No serial correlation) with p-values are in brakets
AIC(Akaike), BIC (Bayes), SIC (Shibata) and HQIC (Hannan-Quinn) are information criterion
Jarque-Bera test on GARCH model residuals (H0: Residuals are normally distributed)
*, **, *** denotes significant at 10%, 5% and 1% respectively

In all models, both α_1 and β_1 are significantly different from zero, therefore it is reasonable to assume time-varying volatility of the residuals.

Based on the Weighted Ljung-Box test on standardized residuals, the null hypothesis of no serial correlation is not rejected thus, we can confirm that there is no serial correlation in GARCH models residuals. TGARCH model provides the lowest information criteria in both cases with student-t or normal distribution. Considering student-t distribution for innovations in all GARCH models result in lower information criterion.

The sum of the two parameters $(\alpha_1 + \beta_1)$ is less than 1, which is good for not resulting in explosive volatility predictions. Since the sum of the parameters is close to one this means that the volatility dies down slowly i.e., it reverts to mean slowly.

Large GARCH lag coefficients, β_1, indicate that shocks to conditional variance take a long time to die out, so volatility is 'persistent'. As for the asymmetric model GJR-GARCH, we see that the γ_1 coefficient is positive and statistically significant, clearly showing how the volatility reacts differently to bad news with respect to good news. Thus, when the bad news hits the market and returns are negative, volatility increases strongly.

Jarque Bera Test Shows that residuals from GARCH models are not normally distributed since the null hypothesis that the data is normally distributed is rejected.

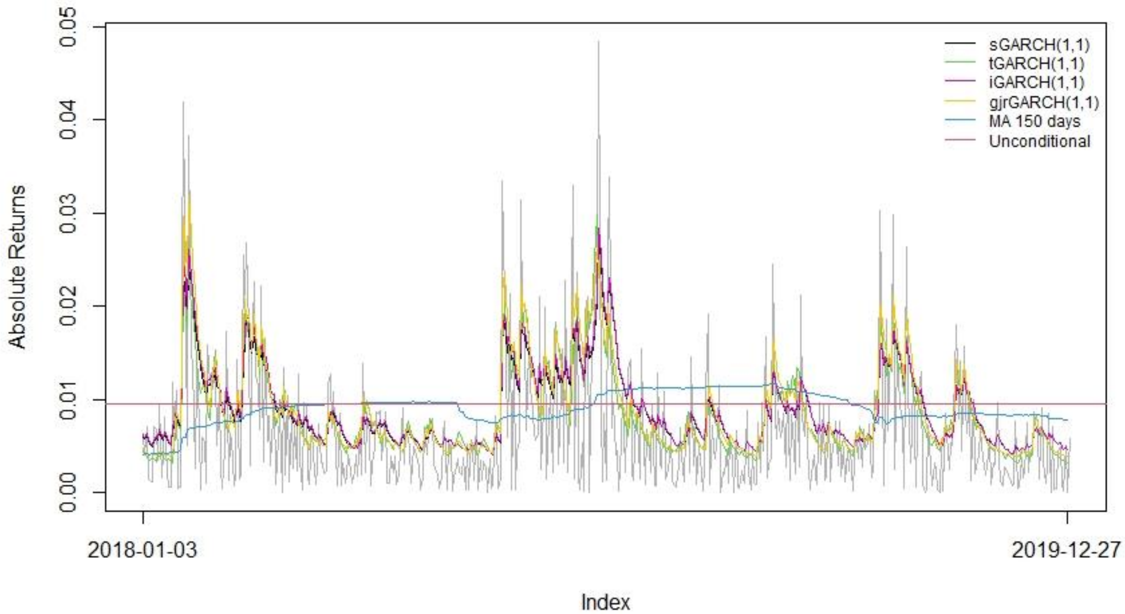
4.4. Forecasting

The returns data has 1761 observations. We use the first 1261 observations to make the initial estimation for the GARCH model. The remaining 500 observations are used for validation and testing.

4.4.1. Forecasting volatility

We use the data from the first 5 years to generate forecasts for the last 2 years based on a rolling estimation. Figure 7 shows how well the GARCH models capture the volatility of returns.

Figure 7 – Volatility forecasting



Based on Figure 7, we can see that GARCH models provide the best results in forecasting volatility comparing to unconditional and moving average volatilities. However, to determine which GARCH model has a better performance we need to employ some other measures.

Here we use 3 forecasting error criteria to compare the forecasting performance of our GARCH models based on 500 observations as our test data.

Table 10 – Comparing forecasting accuracy of GARCH models

	GARCH		TGARCH		IGARCH		GJRGARCH	
	Normal	Student-t	Normal	Student-t	Normal	Student-t	Normal	Student-t
MSE	0.00008963	0.00008973	0.00008956	0.00008947	0.00008973	0.00008973	0.00008961	0.00008953
MAE	0.00652798	0.00653038	0.00654405	0.00653576	0.00653006	0.00653045	0.00655096	0.00653424
DAC	0.55000000	0.55200000	0.51400000	0.54400000	0.54600000	0.55200000	0.51200000	0.55000000

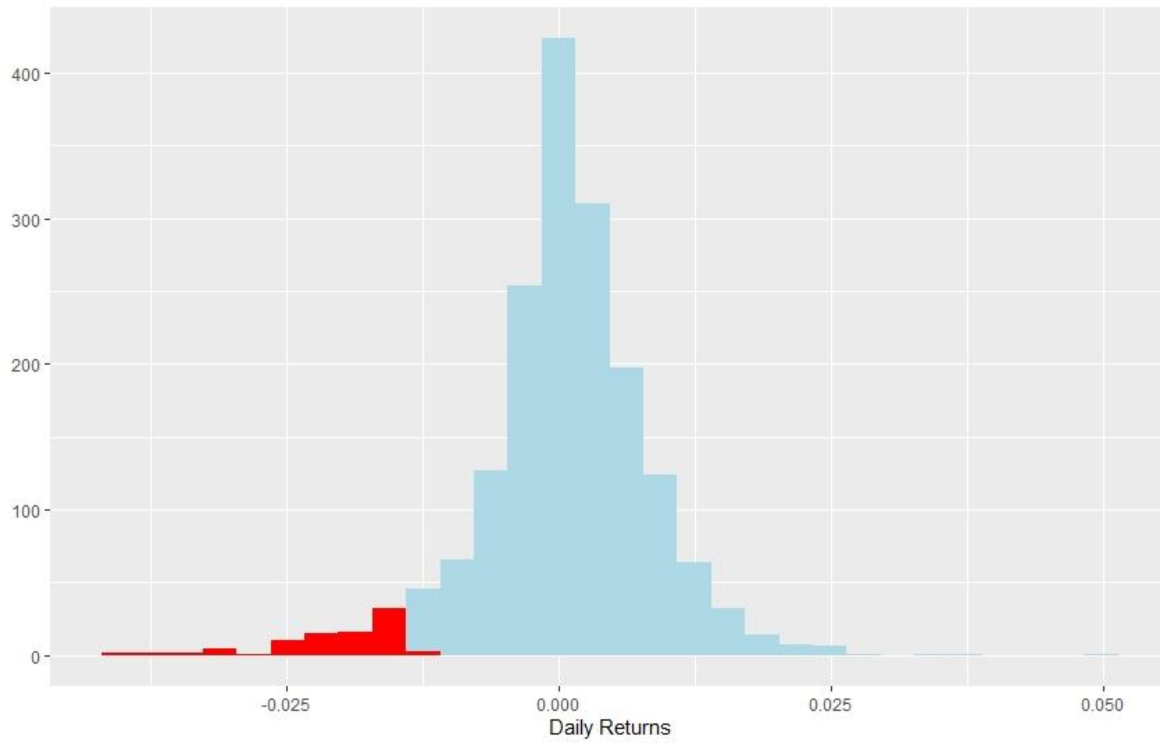
MSE: mean squared error, MAE: mean absolute error and DAC: directional accuracy of the forecast versus realized returns.

We can see that the results on all the measures are very close. TGARCH provides a lower MSE while standard GARCH provides a lower MAE and the lowest DAC is for GJRGARCH. We cannot conclude which GARCH model provides a better performance based on these measures.

4.4.2. Forecasting Value at Risk (VaR)

If we use historical data, we can estimate VaR by taking the 5% quantile value. For our data, this estimation is: -0.01381972 or we can say that for 95% confidence level, the worst daily loss will not exceed 1.38% of the S&P 500 closing prices.

Figure 8 – Historical Value at Risk

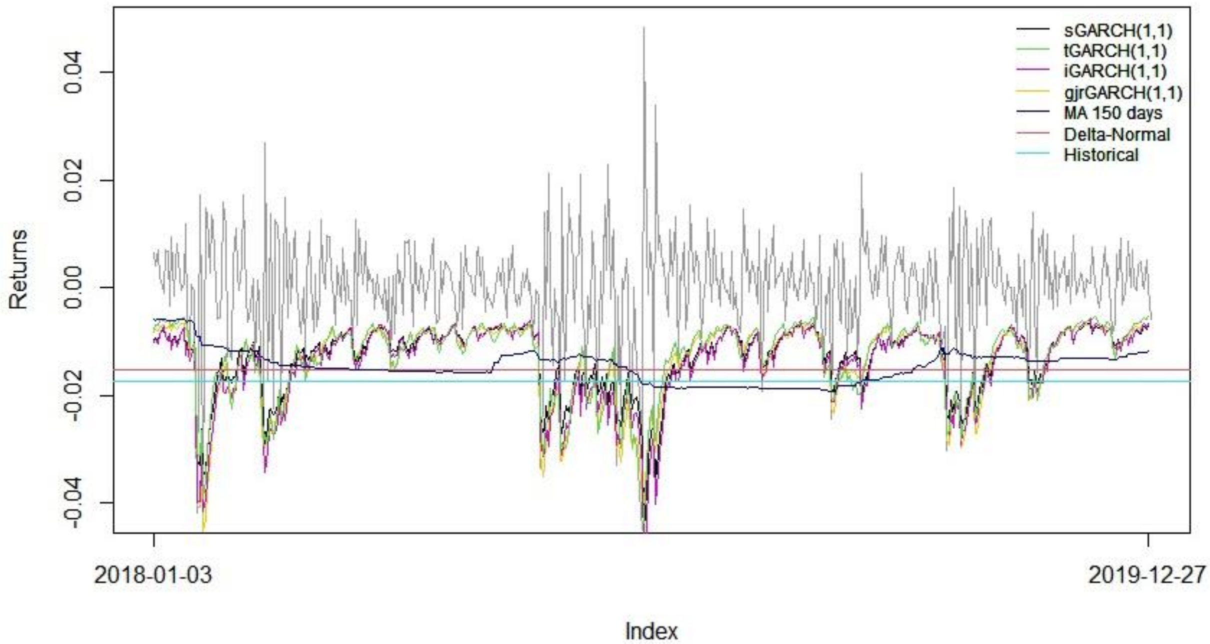


Red bars refer to returns lower than 5% quantile.

Modelling Value at Risk with GARCH (1,1)

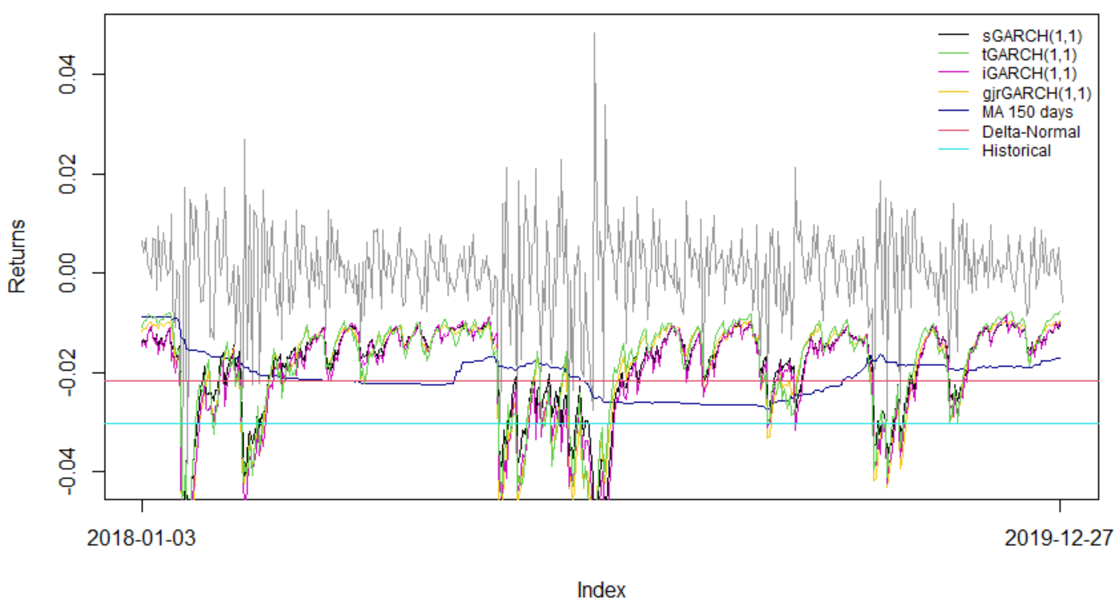
In order to illustrate this method, we apply GARCH (1,1) models with normal and student-t distributions and with confidence level 95% and 99% to provide comparisons.

Figure 9 – 95% VaR forecasting



As we can see from the plot, the VaR-GARCH combination is way more realistic and lowers the VAR limit when volatility clustering occurs, whereas for the static VaR (red line) we observe serial limit breaches.

Figure 10 – 99% VaR forecasting



We can see that a 99% VaR provides a too conservative forecast, and it seems that it's overestimating the risk. Thus, it's required to employ a back testing procedure to understand that how the provided forecasts, differ in reality and which model provides a better performance.

4.5. Back-testing

We use two methods for back-testing in this study. Table 10 shows the results for Kupiec and Christopherson's methods comparing different GARCH models and confidence levels of 95% and 99%.

Table 11 – Back testing results

	GARCH		TGARCH		IGARCH		GJRGARCH	
	Normal	Student-t	Normal	Student-t	Normal	Student-t	Normal	Student-t
Panel A: α = 5%								
Expected Exceed	25	25	25	25	25	25	25	25
Actual VaR Exceed	38	38	35	38	31	38	31	35
Actual Percentage Exceedance	7.60%	7.60%	7%	7.60%	6.20%	7.60%	6.20%	7.00%
Unconditional Coverage (Kupiec)								
LR.uc Statistic	6.181 **	6.181 **	3.765 *	6.181 **	1.413	6.181 **	1.413	3.765 *
Conditional Coverage (Christoffersen)								
LR.cc Statistic	7.706 **	7.706 **	3.936	6.195 **	1.416	7.706 **	1.416	3.936
Panel B: α = 1%								
Expected Exceed:	5	5	5	5	5	5	5	5
Actual VaR Exceed:	15	11	15	9	13	11	14	10
Actual %:	3.00%	2.20%	3.00%	1.80%	2.60%	2.20%	2.80%	2.00%
Unconditional Coverage (Kupiec)								
LR.uc Statistic	13.162 ***	5.419 **	13.162 ***	2.613	8.973 ***	5.419 **	10.994 ***	3.914 **
Conditional Coverage (Christoffersen)								
LR.cc Statistic	13.699 ***	6.848 **	13.699 ***	4.739 *	9.887 ***	6.848 **	11.704 ***	5.665 *

Null-Hypothesis for Kupiec: Correct Exceedances
Null-Hypothesis for Christoffersen: Correct Exceedances and Independence of Failures
Backtest Length: 500 observations
*,**, *** denotes significant at 10%, 5% and 1% respectively

Kupiec's unconditional coverage compares the number of expected versus actual exceedances given the tail probability of VaR, while the Christoffersen test is a joint test of the unconditional coverage and the independence of the exceedances. Based on the results for confidence level 95%, considering a t distribution for innovations results in underestimating the risk while a normal distribution has a better performance. On the other hand, a student-t distribution provides a better performance comparing to normal distribution for a confidence level 99%, however, it seems that a 99% confidence level appears unrealistically conservative which results in rejecting the null hypothesis of "correct exceedances" for most of the models.

Now we compare the back-testing results for α=5%. At this level, both GJRGARCH and IGARCH provide the most accurate exceedances with a normal distribution for innovations in GARCH modeling.

Since a GJRGARCH model assumes that there is asymmetry between negative shocks and positive shocks, which is almost always the case for financial data, we suggest a GJRGARCH model with normal distribution for innovations to provide a 95% VaR forecast for modeling our data.

For a 99% VaR based on the backtesting results a TGARCH-VaR model is selected with student-t distribution for the innovations.

An interesting result that we obtain is that, although we expected that a student-t distribution may perform better for having fat tails, but our results approved a normal distribution for 95% lenel and a t distribution for 99%. We can can find an explanation for this in table 12:

Table 12 – Normal and student-t distribution quantiles

	Shape	α = 0.01	α = 0.05
Student-t	3.038383744	-2.627861	-1.368716
Normal	-	-2.326348	-1.644854

Normal distribution has a bigger quantile at α = 0.05 than t distribution, thus it provides a bigger VaR 95% which results in less exceedances in backtesting. On the other hand, t distribution has a bigger quantile in α = 0.01 resulting in bigger VaR 99% and accordingly less exceedances in backtesting at this level. Thus, we conclude that the chosen distribution for GARCH innovations may provide different performance in different levels of confidence for VaR forecasts. One distribution may not provide good results in all levels of confidence.

4.6. Performance of suggested VaR model

Here we want to see how our selected model performs in reality and how it can be used practically. Figure 11 shows volatility over time plotted with the log returns. Figure 12 depicts graphically the actual exceedances from our selected GJRGARCH-VaR model.

Based on results from table 6, the suggested GJRGARCH(1,1) model can be defined as follows:

$\sigma_t^2 = 0.308062\, I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2 + 0.782385\sigma_{t-1}^2$

Where $I(\varepsilon_{t-1} < 0)$ is an indicator function, which takes the value one if the corresponding lagged conditional standard deviation is less than zero.

Figure 11 – Volatility forecasting with GJRGARCH (1,1)

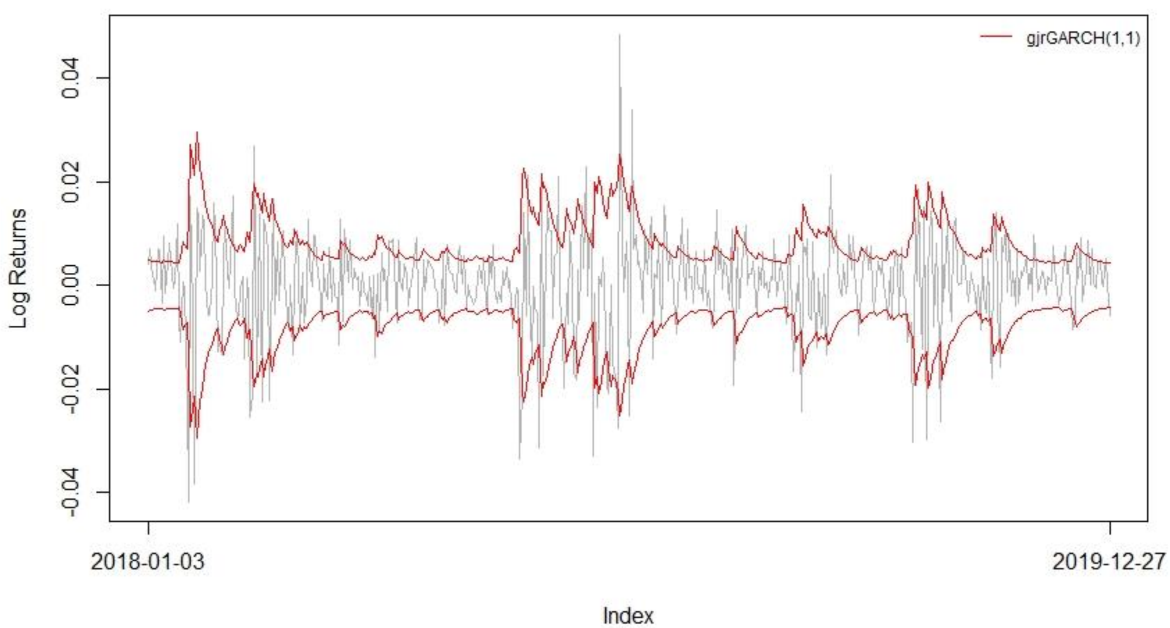
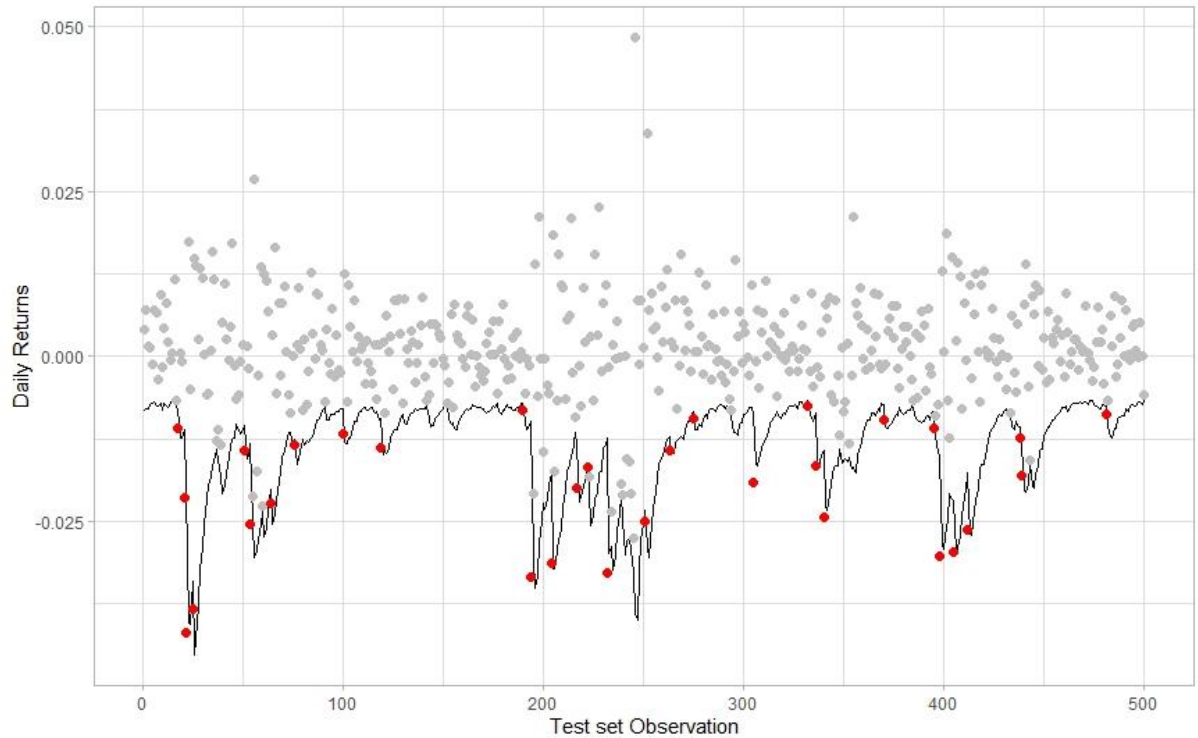


Figure 12-Actual exceedances from GJRGARCH (1,1)-VaR 95% model



The red points illustrate exactly 31 exceedances as calculated in table 10. Based on the graph, we can see that even the exceedances are not so far from the suggested model.

For example, assume that today is 30-May-2019 and we have 1000 shares of the S&P 500 index with a market price of 2788.86. We are interested to know how much the potential loss on our portfolio will be tomorrow.

Table 12 – Predicted loss for actual days

	Index price	No. of shares	Portfolio value	Real loss	VaR 95%	Max predicted loss
2-Jan-2019	2,510.03	1,000	2,510,030	-	-	-
3-Jan-2019	2,447.89	1,000	2,447,890	-62,140	-0.0234	-58,734
30-May-2019	2788.86	1,000	2,788,860	-	-	-
31-May-2019	2752.06	1,000	2,752,060	-36,800	-0.0152	-42,391

Based on our forecasted VaR 95% for 31-May-2019 the potential loss will be 42,391 which is providing a good coverage for the real loss of 36,800.

VaR is also useful when we want to compare the riskiness of different portfolios. This is especially important when evaluating how closely a portfolio manager conformed to the stated risk tolerance of his fund. Corporate Treasuries and Banks use VaR for the same purpose. They need to have an idea of how their market exposures behave under normal market conditions. It is a risk management cliché, but you know that you have a bad risk management regime in place if you are surprised by the extent of any gains or losses that you sustain.

CHAPTER 5

5. Conclusion

Having high levels of volatility in financial markets, it's critical to put in place an efficient risk management strategy to protect against market risk. VaR has become the most common risk measurement method for organizations and regulators in this context. Furthermore, employing dynamic risk measures has been successfully implemented in a variety of fields where high volatility is imposing immense impact on the market.

In this study, we perform a dynamic volatility forecasting using four GARCH type models, being GARCH, TGARCH, IGARCH and GJRGARCH models with one lag on ARCH and GARCH effects each. The model suggested for the mean of return time series is AR(1) which is selected through a Box-Jenkins methodology due to having the lowest information criterion (AIC) in comparison with other possible ARMA models with maximum lag of two. The suggested AR-GARCH models are employed to provide forecasts on Value at Risk (VaR) at different confidence levels of 95% and 99%.

We extract the data for this study from Yahoo finance database on S&P 500 daily prices for a time span of 7 years from 2013 to 2019 resulting in 1762 observations on the index closing prices. In order to obtain a stationary time series for our analysis, we compute the log return of the index prices with 1761 observations to provide our results.

Our analysis on the return series shows evidence of non-normality and fat tails, consequently for our analysis we consider both normal and student-t distributions and provide comparisons on their results. Accordingly, we have 8 comparisons on volatility modelling including four GARCH type models with two distributions for their innovations each and we have 16 comparisons on the VaR forecasts considering confidence intervals of 99% and 95% in addition to the 8 volatility results.

Our results on volatility modelling show that, a TGARCH model with student-t distribution for the innovations, provides the lowest information criteria, however, based on MSE and MAE measures we cannot judge the performance of each model.

Our backtesting results on VaR forecasts, show that GJRGARCH and IGARCH with normal distribution for innovations both provide the lowest exceedances in 95% level and at 99% confidence level, TGARCH with student-t distribution provides the best results on backtesting. Since there is always asymmetry between the negative and positive shocks in financial data, and GJRGARCH assumes such asymmetry in modelling, we suggest a GJRGARCH with normal innovations for modelling a 95% VaR. However, based on the risk appetite of the users, one may choose TGARCH with student-t distribution at 99% confidence interval as it provides a much more conservative measure and consequently more costly regarding the required capital charge or other risk cushions based on the risk management strategies. Based on these results, we find that the confidence level considered for forecasting VaR is a decisive factor in selecting a proper distribution for GARCH innovations which can result in better VaR forecasts. Ignoring this fact, may result in lower accuracy of var forecasts in different confidence levels.

Our results are based on an index from the 500 largest companies listed on stock exchanges in the United States during a specific period of time and we might have found different results for different periods of other financial data, especially when market conditions change. As a result, we expect that no single model can be defined as the best performer across all returns data sub-periods. Related literature to this study highlights the possible pitfalls of using VaR as a risk management method to specify the minimum regulatory capital requirement under varying market conditions. Failure to account for such market changes may result in serious model misspecifications and incorrect model selections. Depending on market conditions or the regime of the sub-period being studied, the resulting VaR forecasts can be significantly over- or under-estimated, but they tend to be stable over a longer time horizon. As a result of these misunderstandings, firms and financial intermediaries that manage risk using VaR can find themselves with insufficient capitalizations.

In reality, dynamic volatility has the following consequences: financial returns are more likely than expected to result in significant losses (the "fat tail" effect); uncertain periods appear to cluster, with large price fluctuations within days. As a recent and realistic example, equity markets encountered extreme volatility in 2020, owing primarily to the COVID-19 incident. When considering the historical distribution of price changes, such events are highly unexpected. Thus, in addition to employing accurate VaR modelling we need to have proper scenario analysis and stress testing procedures to foresee the rare events that are not predicted by models based on normal conditions historical data.

Our findings in this study also point to a number of areas in which further research is required. To improve the accuracy of VaR estimations, one potential avenue is to impose model-switching mechanisms rather than parameter switching. The inherent benefit stems from the ability to switch between the best performing models for VaR estimation as market conditions adjust and the market enters a new regime. Alternatively, as suggested by Nieto and Ruiz (2016), further research into the implementation of bias corrections to enhance the forecasting of GARCH models could be conducted in order to reach more conclusive results on the performance of VaR models across various market regimes. Finding more appropriate distributions for GARCH innovations would be another suggestion.

REFERENCES

Abad, P., Benito, S., & López, C. (2014). A comprehensive review of Value at Risk methodologies. The Spanish Review of Financial Economics, 12(1), 15–32. https://doi.org/10.1016/j.srfe.2013.06.001

Altun, E. (2018). A new approach to Value-at-Risk: GARCH-TSLx model with inference. Communications in Statistics - Simulation and Computation, 49(12), 3134–3151. https://doi.org/10.1080/03610918.2018.1535069

Ardia, D., & Hoogerheide, L. F. (2014). GARCH models for daily stock returns: Impact of estimation frequency on Value-at-Risk and Expected Shortfall forecasts. Economics Letters, 123(2), 187–190. https://doi.org/10.1016/j.econlet.2014.02.008

Bali, T. G., & Cakici, N. (2004). Value at Risk and Expected Stock Returns. Financial Analysts Journal, 60(2), 57–73. https://doi.org/10.2469/faj.v60.n2.2610

Bentes, S. R. (2015). Forecasting volatility in gold returns under the GARCH, IGARCH and FIGARCH frameworks: New evidence. Physica A: Statistical Mechanics and Its Applications, 438, 355–364. https://doi.org/10.1016/j.physa.2015.07.011

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Christoffersen, P. F. (1998). Evaluating Interval Forecasts. International Economic Review, 39(4), 841. https://doi.org/10.2307/2527341

Christoffersen, P., Hahn, J., & Inoue, A. (2001). Testing and comparing Value-at-Risk measures. Journal of Empirical Finance, 8(3), 325–342. https://doi.org/10.1016/s0927-5398(01)00025-1

Degiannakis, S., Floros, C., & Dent, P. (2013). Forecasting value-at-risk and expected shortfall using fractionally integrated models of conditional volatility: International evidence. International Review of Financial Analysis, 27, 21–33. https://doi.org/10.1016/j.irfa.2012.06.001

Elenjical, T., Mwangi, P., Panulo, B., & Huang, C. S. (2016). A comparative cross-regime analysis on the performance of GARCH-based value-at-risk models: Evidence from the Johannesburg stock exchange. Risk Management, 1–1000. https://doi.org/10.1057/rm.2016.4

Emenogu, N. G., Adenomon, M. O., & Nweze, N. O. (2020). On the volatility of daily stock returns of Total Nigeria Plc: evidence from GARCH models, value-at-risk and backtesting. Financial Innovation, 6(1), 1–900. https://doi.org/10.1186/s40854-020-00178-1

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4), 987. https://doi.org/10.2307/1912773

Engle, R. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. Journal of Economic Perspectives, 15(4), 157–168. https://doi.org/10.1257/jep.15.4.157

Hafer, R., & Sheehan, R. G. (1989). The sensitivity of VAR forecasts to alternative lag structures. International Journal of Forecasting, 5(3), 399–408. https://doi.org/10.1016/0169-2070(89)90043-5

Huang, Z., Liu, H., & Wang, T. (2016). Modeling long memory volatility using realized measures of volatility: A realized HAR GARCH model. Economic Modelling, 52, 812–821. https://doi.org/10.1016/j.econmod.2015.10.018

Kupiec, P. H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. The Journal of Derivatives, 3(2), 73–84. https://doi.org/10.3905/jod.1995.407942

Ng Cheong Vee, D. A. N., Nunkoo Gonpot, P., & Ul Hacq Sookia, N. (2014). Forecasting value-at-risk for frontier stock market indexes using GARCH-type models and extreme value theory: model validation for dynamic models. The Journal of Risk Model Validation, 8(4), 47–67. https://doi.org/10.21314/jrmv.2014.128

Nieto, M. R., & Ruiz, E. (2016). Frontiers in VaR forecasting and backtesting. International Journal of Forecasting, 32(2), 475–501. https://doi.org/10.1016/j.ijforecast.2015.08.003

Okpara GC. (2015). Downside risk analysis of the Nigerian stock market: a value at risk approach. In: A Conference Paper presented at International Science Index, LOS Angeles USA Part XIX

Slim, S., Koubaa, Y., & BenSaïda, A. (2017). Value-at-Risk under Lévy GARCH models: Evidence from global stock markets. Journal of International Financial Markets, Institutions and Money, 46, 30–53. https://doi.org/10.1016/j.intfin.2016.08.008

So, M. K., & Yu, P. L. (2006). Empirical analysis of GARCH models in value at risk estimation. Journal of International Financial Markets, Institutions and Money, 16(2), 180–197. https://doi.org/10.1016/j.intfin.2005.02.001

Spierdijk, L. (2016). Confidence intervals for ARMA–GARCH Value-at-Risk: The case of heavy tails and skewness. Computational Statistics & Data Analysis, 100, 545–559. https://doi.org/10.1016/j.csda.2014.08.011

Tabasi, H., Yousefi, V., Tamošaitienė, J., & Ghasemi, F. (2019). Estimating Conditional Value at Risk in the Tehran Stock Exchange Based on the Extreme Value Theory Using GARCH Models. Administrative Sciences, 9(2), 40. https://doi.org/10.3390/admsci9020040

Wipplinger, E. (2007). Philippe Jorion: Value at Risk – The New Benchmark for Managing Financial Risk. Financial Markets and Portfolio Management, 21(3), 397–398. https://doi.org/10.1007/s11408-007-0057-3

Rabemananjara, R., & Zakoian, J. M. (1993). Threshold arch models and asymmetries in volatility. Journal of Applied Econometrics, 8(1), 31–49. https://doi.org/10.1002/jae.

VITA AUCTORIS

NAME: Katayoon Naimian

EDUCATION: University of Allameh Tabatabaei, B.Sc., Tehran, Iran, 2003

University of Leeds, M.Sc., Leeds, United Kingdom, 2018

University of Windsor, MA, Windsor, ON, 2021