

iscte

INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Can we improve the accuracy of the Value-at-Risk with asymmetric and long memory GARCH models?

Rafael Manuel Vaz Lima

Master in Finance

Supervisor:

PhD, António Manuel Rodrigues Guerra Barbosa, Assistant Professor, Iscte-lul

July, 2021



BUSINESS
SCHOOL

Department of Finance

Can we improve the accuracy of the Value-at-Risk with asymmetric and long memory GARCH models?

Rafael Manuel Vaz Lima

Master in Finance

Supervisor:

PhD, António Manuel Rodrigues Guerra Barbosa, Assistant Professor, Iscte-lul

July, 2021

Acknowledgment

Firstly, I would like to thank my supervisor, professor António Barbosa, for all the help and availability during my dissertation. Without a doubt this work would fall short without his comments and advice.

I would also like to thank my family that supported me through my education journey and gave me all the help I could ask for.

Moreover, I would like to extend my gratitude to my friends that constantly pressured me to work hard and finish the dissertation in time. A special mention to João Neri and João Mota, that gave me multiple advice and insights, and Alexandre Serrano that accompanied me from the bachelor's to the master's degree.

Lastly, I would like to express my gratitude to all the professors that accompanied me throughout my ISCAL and ISCTE Business School journeys.

Resumo

Estudos empíricos têm demonstrado que muitos retornos de ativos financeiros são caracterizados por um comportamento assimétrico da volatilidade. Para além da assimetria, outra propriedade importante sobre a volatilidade é o seu efeito de memória longa. Dado que o cálculo da volatilidade é uma das peças centrais na estimação do Value-at-Risk (VaR), estas duas propriedades podem ter um impacto na precisão das estimativas do VaR, e modelá-las poderá ser benéfico para melhorar a precisão das estimativas. Nesta dissertação exploramos a importância da assimetria e memória longa da volatilidade para a estimação do VaR para vários índices de mercado. Utilizamos dois dos mais importantes modelos de volatilidade, Exponential Weighted Moving Average (EWMA) e Generalized Auto Regressive Conditional Heteroscedasticity (GARCH), e também modelos assimétricos e de memória longa, Exponential GARCH (EGARCH) e Fractional Integrated GARCH (FIGARCH). Além disso, alargamos a literatura, que é tendencialmente focada em apenas um modelo de VaR, para quatro modelos: Parametric Normal VaR, Parametric Skewed Generalized Student-t VaR, Volatility-Adjusted Historical VaR and Quantile Regression VaR. A performance dos quatros modelos de VaR é avaliada através de um backtest que recorre ao Unconditional Coverage Test e ao Berkowitz, Christoffersen and Pelletier test. Os resultados deste estudo mostram que não existe evidência que as estimativas produzidas por modelos que acomodam a assimetria e a memória longa da volatilidade são mais precisos que os produzidos por modelos mais simples, sugerindo que acomodar estas propriedades não tem um impacto direto na precisão das estimativas do VaR.

Palavras-Chave: Value-at-Risk, Modelo de Volatilidade, Memória Longa, Assimetria, Backtesting.

Classificação JEL: C10, G32

Abstract

Empirical studies have revealed that many financial asset returns are characterized by an asymmetric behavior of the volatility. Aside from asymmetry, another important stylized fact about the volatility is its long memory effect. Given that the computation of the volatility is one of the central pieces for the estimation of the Value-at-Risk (VaR), these two properties could have an impact on the precision of the VaR estimates, and accounting for them might prove beneficial to improve its' accuracy. In this dissertation we'll explore the relevance of the asymmetry and long memory of the volatility on the estimation of the VaR for various market indices. We'll use two of the most well-known volatility models, Exponential Weighted Moving Average (EWMA) and Generalized Auto Regressive Conditional Heteroscedasticity (GARCH), as well as asymmetric and long memory models, Exponential GARCH (EGARCH) and Fractional Integrated GARCH (FIGARCH). Furthermore, we extend the literature, which tends to be focused on a single VaR model, to four models: Parametric Normal VaR, Parametric Skewed Generalized Student-t VaR, Volatility-Adjusted Historical VaR and Quantile Regression VaR. The performance of the four VaR models will be assessed through a backtest recurring to the Unconditional Coverage test and the Berkowitz, Christoffersen and Pelletier test. The results from this study show no evidence that estimates produced by models that account for the asymmetry and long memory of the volatility are more accurate than the ones produced by simpler models, suggesting that accounting for these properties has no direct impact in the precision of the VaR estimates.

Keywords: Value-at-Risk, Volatility Model, Long Memory, Asymmetry, Backtesting.
JEL Classification: C10, G32

Contents

Acknowledgment	i
Resumo	iii
Abstract	v
List of abbreviations	ix
Chapter 1. INTRODUCTION	1
Chapter 2. LITERATURE REVIEW	3
Chapter 3. DATA, DESCRIPTIVE STATISTICS, STATIONARITY AND LONG MEMORY	9
3.1. Data and Descriptive Statistics	9
3.2. Stationarity test	10
3.3. Long Memory test	11
Chapter 4. METHODOLOGY	13
4.1. Volatility Models	13
4.2. Value-at-Risk Models	17
4.3. Evaluating the performance of VaR models	20
Chapter 5. BACKTEST	25
5.1. Backtest – 1% Significance Level	25
5.2. Backtest – 2.5% and 5% Significance Levels	31
5.3. Backtesting Conclusions	31
Chapter 6. CONCLUSION	35
References	37
Appendices	41
Chapter A. Appendix A - EWMA Volatility Estimation	43
Chapter B. Appendix B - UC and BCP test results for 2.5% significance level	45
Chapter C. Appendix C - UC and BCP test results for 5% significance level	51
Chapter D. Appendix D - Source Code	57

List of abbreviations

- ADF** – Augmented Dickey-Fuller
- APARCH** – Asymmetric Power Auto Regressive Conditional Heteroscedasticity
- AR** – Autoregressive
- ARCH** - Autoregressive Conditional Heteroscedasticity
- BCP** – Berkowitz, Christoffersen and Pelletier
- BIS** – Bank of International Settlement
- CHI** – Christoffersen, Hahn, and Inoue
- DAX** – Deutscher Aktienindex
- DF** – Dickey-Fuller
- EGARCH** – Exponential Generalized Auto Regressive Conditional Heteroscedasticity
- EWMA** – Exponential Weighted Moving Average
- FIAPARCH** – Fractionally Integrated Asymmetric Power Auto Regressive Conditional Heteroscedasticity
- FIGARCH** – Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroscedasticity
- FIGARCH** – Fractional Integrated Generalized Auto Regressive Conditional Heteroscedasticity
- FTSE** – Financial Times Stock Exchange
- GARCH** – Generalized Auto Regressive Conditional Heteroscedasticity
- GED** – Generalized Error Distribution
- GJR-GARCH** – Glosten-Jagannathan-Runkle Generalized Auto Regressive Conditional Heteroscedasticity
- HSI** – Hang Seng Index
- KPSS** – Kwiatkowski–Phillips–Schmidt–Shin
- OLS** – Ordinary Least Squares
- R/S** – Range over Standard Deviation
- RM** – RiskMetrics™
- SGSt** – Skewed Generalized Student-t
- S&P500** – Standard & Poor’s 500
- UC** – Unconditional Coverage
- VaR** – Value-at-Risk

CHAPTER 1

INTRODUCTION

Risk has become an everyday topic for any finance practitioner. Managing it, is a must for any financial institution. More than that, it has become a requirement of regulatory entities. While it's impossible to precisely estimate the future value of an asset, it is possible to reasonably estimate how likely, and how much, we are expected to lose. By doing so, we are able to prepare ourselves against bad swings and ease future negative outcomes.

Although many sources of uncertainty exist, this dissertation will be focused on the market risk, which arises from variations in the market prices of financial assets. To measure the market risk, a metric capable of capturing the uncertainty associated with fluctuations of the assets composing a portfolio is needed. The Value-at-Risk (VaR) has become the standard market risk metric used by banks, investors, risk managers and other finance practitioners. It can be defined as the maximum expected lost in the value of a portfolio for a given time horizon and confidence level (Alexander, 2009). It started to gain traction in 1994 when J.P. Morgan publicly released a standardized procedure to compute the VaR, the RiskMetrics (Longerstaey and Spencer, 1996).

To estimate the VaR, the volatility of the returns needs to be modeled. For that, the RiskMetrics model uses an EWMA volatility. While simple, computing the volatility recurring to an EWMA implies a failure in capturing some properties that could potentially improve the accuracy of the VaR estimation. The GARCH model is another popular way to model the volatility. Proposed by Bollerslev (1986), this model has faced various mutations that add the possibility to accommodate more properties into it. Two important properties, and those that will receive attention in this dissertation, are the asymmetric behavior of the volatility and its long memory effect.

The asymmetric behavior of the volatility is observed in the responses that the volatility has to negative and positive shocks: negative shocks impact more the volatility than positive shocks of the same magnitude. To account for the asymmetric behavior of the volatility various models have been presented, but throughout this dissertation we will work with the EGARCH model introduced by Nelson (1991).

The long memory effect of the volatility shows itself when comparing observations distant in time. If the long memory effect is present these observations display high correlation between them (Horta, 2015). What that means is that past occurrences still affect the way volatility reacts to a shock today. We can accommodate the long memory effect through a model proposed by Baillie et al. (1996), the FIGARCH model.

Although much research exists surrounding the ability of these models to improve the accuracy of the VaR estimates, as far as we know, these studies focus mainly on a single VaR model for one index or market. We'll build upon that, studying four VaR models: Parametric Normal, Parametric Skewed Generalized Student-t, Volatility-Adjusted Historical and Quantile Regression, for various indexes of different geographies.

The volatility will be estimated through five models: two EWMA volatility models, GARCH, EGARCH and FIGARCH. Using these, allows us to check both the asymmetry and the long memory of the volatility, while comparing them with three benchmarks, which in this dissertation will be the two EWMA volatility models and the GARCH model. Including the GARCH model is also important since it's one of the most widely used models and the model that serves as the base for the EGARCH, FIGARCH and EWMA.

The VaR will be computed for several indices: S&P500, DAX, FTSE and HSI. Working with these indices allows us to explore various geographies, including an emerging economy. In addition, the advantage of using indices and not stocks or other single asset is that an index provides a good representation of a large and diversified portfolio.

Through a backtest, using the Unconditional Coverage (UC) test (Kupiec, 1995) and the BCP test (Berkowitz et al., 2011), we'll assess the performance of the five volatility models for each of the twenty model combinations and index pairs, finding out which of them best estimates the VaR.

Our goal is to extend the current literature to a broader set of VaR models and geographies. With that we expect to find more robust leads about the importance of modeling the asymmetry and the long memory of the volatility to compute the VaR estimates. If in fact modeling one of these two properties allows to compute more accurate estimates, the choice of which model to select when computing the VaR will be eased. The results of this study are important to researchers, since it provides an analysis over various markets and for a big timeframe, and especially for risk managers, in the sense that, having a better understanding on which model is able to produce the most accurate estimates, allows the risk to be modeled more accurately, leading to a better capital allocation.

The results of this study show that, while having good performances, the models that account for the asymmetry and the long memory of the volatility do not dominate the others in study, entailing that there is no evidence to affirm that accommodating these two properties improves the accuracy of the Value-at-Risk estimation.

This dissertation is organized as follows: Section 2 reviews the most relevant literature; Section 3 goes over the data, investigates the stationarity of the prices and returns and analyzes the long memory property of the returns and it's volatility; Section 4 presents all methodology employed; Section 5 presents the results of this study; Section 6 summarizes our findings and goes over some suggestions for future investigation.

CHAPTER 2

LITERATURE REVIEW

The concept of market risk derives from the uncertainty concerning the future value of an asset or a portfolio. While this uncertainty can resolve into positive results, the main concern revolves around negative ones.

The Value-at-Risk is currently the standard metric of market risk. It is a statistical measure of the potential loss that can be defined as the maximum expected loss in the value of a portfolio, over a time horizon for a given confidence level (Alexander, 2009).

To compute the VaR it becomes imperative to model the volatility of the portfolio, which can be done through complex models or simple and intuitive ones. While various models exist, the EWMA volatility model has become one of the most cherished. One of the main reasons for the popularity of the EWMA volatility model, aside from its simplicity, is the fact that it's the model used in the RiskMetrics model.

Aside from the EWMA volatility model, GARCH models are also very popular volatility models. Initially introduced by Bollerslev (1986), the original GARCH has seen various mutations that aim to accommodate multiple properties and improve the sharpness of the volatility estimates.

One of these properties is the asymmetric behavior of the volatility, also referred has leverage effect. Negative shocks have a larger impact on the volatility than positive shocks, that is, when responding to a negative shock, the volatility increases more than it would for a positive shock of the same magnitude. In that sense, it is typical to say that negative news have a greater impact in volatility than positive news (Campbell and Hentschel, 1992; Christie, 1982).

There are two widely accepted explanations for the asymmetric behavior of the volatility. The first one is explored in Black (1976) and Christie (1982). These authors explained that in a situation where the equity of a firm decreases, the firm's debt-to-equity ratio will obviously rise. Assuming the perspective of an equity holder, having a higher debt-to-equity ratio, means that, in case of a bankruptcy, the investor will receive less than it would before. Knowing this, the equity holder perceives his investment as riskier than before. The second explanation is referred as the volatility feedback effect, it says that if the expected return of an asset increases when that asset's volatility rises, the asset's price will fall in face of a volatility rise (Kayal and Maheswaran, 2018).

A number of asymmetric GARCH volatility models started to be suggested and two rose as the most widely used: the EGARCH, introduced by Nelson (1991), and the Glosten-Jagannathan-Runkle-GARCH (GJR-GARCH), suggested by Glosten et al.

(1993). Throughout this dissertation we will focus on the EGARCH model, which includes an asymmetry coefficient that allows the leverage effect to be accounted for.

To explore the impact of the volatilities' asymmetric behavior on the Parametric Normal 5% VaR, throughout 15 years, Brooks and Persaud (2003) looked at five Southeast Asian economies, Hong Kong, Japan, Singapore, South Korea and Thailand, as well as the S&P500, used as a benchmark. Beyond using the five indices correspondent to each of these economies, the authors also created an equally weighted portfolio. The objective was to find out if modeling the volatility recurring to asymmetric models: EGARCH, Semi-Variance, Multivariate GJR-GARCH and Multivariate EGARCH; lead to more accurate estimates for the VaR than using symmetric models: EWMA, GARCH and Multivariate GARCH. The performance of the various models was evaluated recurring to the Bank of International Settlement (BIS) regulatory backtest. The results pointed to the existence of an unconditional skewness and conditional asymmetric response of the volatility. In terms of the VaR, the Semi-Variance models lead to more stable estimates, which is appreciated in the light of the BIS. Another discovery was that asymmetrical models were able to outperform symmetrical ones, which lead to inappropriate small VaR estimates.

Bucevska (2013) delved into the Macedonian Stock Exchange to test the performance of various models, analyzing how well they were able to gauge the volatility. To do so, the author considered various GARCH models: symmetrical (GARCH and Multivariate GARCH) and asymmetrical (EGARCH, GJR-GARCH and Asymmetric Power Auto Regressive Conditional Heteroscedasticity (APARCH)). The parameters of the GARCH models were computed following the Berndt–Hall–Hall–Hausman algorithm and no re-estimation was employed. The volatility, estimated by the different models, was used to estimate the 1% and 5% VaR for a period of approximately 6 years. The accuracy of these estimates was evaluated following Engle and Ng's test for sign and size bias (Engle and Ng, 1993). Looking at the results of the paper, the authors were able to detect the presence of an ARCH effect in the residuals as well as a volatility clustering effect. Regarding the volatility models, it was found out the most adequate models, that is, the models that lead to more accurate VaR estimates, were the EGARCH, both with the normal and student-t distribution, and the GJR-GARCH, pointing towards the preeminence of asymmetrical models when compared to symmetrical ones.

The same study was done for Nordic Indices, Swedish OMXS30 and Danish OMXC20, as well as Nordic stocks, H&M, Volvo, Carlsberg and Maersk, by Berggren and Folkelid (2015), from 2007 to 2014. The 1% and 5% Parametric Normal VaR were computed recurring to the GARCH, ARCH, EGARCH and GJR-GARCH models, all under three innovation distributions: normal, student-t and generalized error distribution (GED). The parameters needed to compute the models were re-estimated every 500 trading days and the performance of the estimates was evaluated recurring to the unconditional and independence tests. Various results arose from this study, firstly the authors mention that the volatility models studied underestimate the risk for the 1% VaR and overestimate it

for the 5% VaR. For the 5% VaR, leptokurtic distributions had a slight edge, while for the 1% VaR they had a big advantage. Moreover, for the 5% VaR, while no model was found to be clearly superior to the others, asymmetrical models appear to dominate symmetrical ones. For the 1% VaR, the ARCH model was found to be the best (Berggren and Folkelid, 2015).

Angelidis et al. (2004) weren't able to find a volatility model that clearly outperformed the rest. The authors studied a variety of GARCH models applied to various indices, S&P500, NIKKIE, DAX, CAC and FTSE, for a period of 15 years. As mentioned above, no model was able to outperform the rest, but, recurring to the unconditional and conditional coverage tests, the authors found a strong evidence that the EGARCH model, under a student-t distribution, produced the most adequate VaR estimates for the majority of the markets.

Aside from asymmetry, another important stylized fact about the volatility is its long memory effect. A time series exhibits long memory when the present observations are dependent on preceding distant observations. Put another way, successive observations display a slow decay in their dependency and correlation. Looking at the financial markets, the long memory effect means that the market reacts slowly to the data it receives daily. This implies that shocks in the volatility are not momentaneous, they propagate through time and have a long-lasting effect. Interestingly enough, the long memory effect wasn't firstly researched in the scope of economics or finance. It was firstly documented by Hurst (1951) when studying the presence of long-range dependency on the flow of the Nile river.

Ding et al. (1993) took the findings of Hurst to the financial world and showed that many financial returns display long memory on the volatility. Given this discovery, a new model, one capable of accommodating the long memory effect of the volatility, was needed. Baillie et al. (1996) introduced the FIGARCH model, a model that adds a fractional integration to the GARCH model and allows long memory to be accounted for.

Recurring to the FIGARCH and comparing it with a GARCH model for the Parametric VaR, Wu and Shieh (2007) sought to understand the impact of the long memory effect of the volatility on the long-term T-Bond interest rate futures VaR estimation, from 1977 until 2005. The authors concluded that the return series are a $I(0)$ ¹ process, which does not have a long memory characteristic, but the volatilities are a $I(d)$ ² long memory process, where all the estimated degrees of integration are significant and different from 0 and 1. Given this, the volatility can be characterized by a slowly mean-reverting fractionally integrated process, in the case of this study, the FIGARCH model. Resorting to Kupiec's LR test (Kupiec, 1995), the authors understood that the FIGARCH model outperformed the GARCH, for both the in-sample and out-of-sample VaR estimation. Furthermore, there was strong evidence that interest rate futures exhibit long memory in

¹Non-integrated (stationary) process

²Integrated non-stationary process, which can be made stationary through a series of successive differences, d

their volatility, which points out to a potential benefit of accommodating the long memory effect of the volatility to estimate the VaR.

With the same line of thought of Wu and Shieh (2007), Kasman (2009) studied the long memory property of the volatility for the Turkish stock index futures market. The author compared the performance of a FIGARCH and a GARCH model, for a period of 3 years. Recurring to Kupiec's LR test, the author found that the FIGARCH model outperformed the GARCH for the sample in study.

Additional investigation has been made into the importance of the long memory effect on the estimation of the VaR. The conclusions of these studies were harmonious and show that accommodating the long memory of the volatility is beneficial and allows to compute more accurate VaR estimates (So and Philip, 2006; Tang and Shieh, 2006; Bentes, 2015).

Albeit long memory being an important property, asymmetry must not be forgotten. When comparing models that accommodate asymmetry with models that accommodate long memory, the question on which model performs better to estimate the VaR can arise. With this thought in mind, Yao et al. (2006) explored both properties of the volatility on the Shanghai Stock Exchange A Share Index, from 1993 until 2003, using four different volatility models, EWMA, GARCH, EGARCH and FIGARCH. The results were assessed based on two tests, Kupiec's LR test and Christoffersen LR test (Christoffersen, 1998), as well as two Christoffersen, Hahn, and Inoue (CHI) tests (Christoffersen et al., 2001), the CHI specification test and the CHI non-nested test. The authors revealed the both the EGARCH and the FIGARCH models generated better VaR estimates than the EWMA and the GARCH. According to the CHI non-nested test the models could be arranged, from best to worst as follows: FIGARCH, EGARCH, EWMA and GARCH. These results also imply that the asymmetric behavior of the volatility and the long memory effect are important properties for the VaR estimation on the Shanghai Stock Exchange A Share Index, but there was no clear winner when comparing the accommodation of the long memory and the asymmetric behavior of the volatility to estimate the VaR.

Sethapramote et al. (2014) investigated the performance of the FIGARCH model against numerous GARCH models, EGARCH, IGARCH and GJR-GARCH, as well as the standard GARCH, ignoring the re-estimation of the parameters. This study was done for the 1% VaR of the Stock Exchange of Thailand, over a period of 8 years. The authors observed a strong evidence of long memory in the variance of the Stock Exchange of Thailand index returns, while the mean of the stock returns were able to be characterized by a short memory process, the same finding that Wu and Shieh (2007) made for the long-term T-Bond interest rate futures. From Kupiec's LR test and BIS's regulatory backtest, the authors found that the FIGARCH outperformed the simple GARCH model, but was outperformed by the asymmetrical models studied.

As the literature mention until this point shows, most authors focus their research on a single VaR model for a single index or market. In our view, it would be beneficial to study the impact of asymmetry and long memory of the volatility on multiple VaR

models for various markets. As for the volatility models, some authors decide to study only a couple of them, that either don't compare asymmetry and long memory, or leave out important models that could be interesting to compare with the asymmetric and long memory models. Moreover, some literature ignores the re-estimation of the parameters for the volatility and VaR models, which can cause over/under estimations of the VaR.

In order to expand onto the existing literature and overcome the shortcomings discussed above, we'll study the impact of the asymmetry and long memory of the volatility under various VaR models: Parametric (Normal and Skewed Generalized Student-t distributed), Volatility-Adjusted Historical and Quantile Regression, using five volatility models: two EWMA models, GARCH, EGARCH and FIGARCH, for various market indices, re-estimating all the parameters needed every trading month (21 days). This approach will allow us to take stronger and broader conclusions about the importance of modeling the asymmetry and long memory of the volatility to estimate the VaR and perceive if the different volatility models work better under specific VaR models or for any particular market. The tests selected to run the backtest are the UC test and the BCP test. They will let us evaluate the performance of the various models and provide information about the exceedance rate as well as the autocorrelation between the VaR estimates. Ultimately it allows us to answer our main question: Is it possible to improve the accuracy of the VaR estimates modeling the asymmetry and the long memory effect of the volatility?

DATA, DESCRIPTIVE STATISTICS, STATIONARITY AND LONG MEMORY

3.1. Data and Descriptive Statistics

This study examines the impact of modeling the asymmetry and long memory of the volatility on the accuracy of the VaR for multiple indices. Given that indices are composed of various stocks, they provide a good representation of a large and diversified portfolio. To extend our research and be able to take broader and stronger conclusions, we analyze the VaR of four indices of different geographies: S&P500, DAX, FTSE and HSI.

The price of each index was drawn from Yahoo Finance. We use the adjusted daily closing prices and convert them into daily log returns, which allows us to estimate the daily VaR in percentage points:

$$r_t = \ln \left(\frac{S_t}{S_{t-1}} \right), \quad (1)$$

where r_t represents the return at time t and S_t represents the index price at time t .

Timewise, we work for a period of twenty years, from 2000 until 2020. To do so, twenty-four years of data were needed, the first four are required to estimate the first batch of parameters and allow all the study to be done out-of-sample.

Regarding R, the language in which the computations of this dissertation are implemented, several packages are needed:

- Quantmod – contains the functions needed to import data from Yahoo Finance;
- Rugarch – includes the procedures to estimate the GARCH models;
- SGT – to estimate the Skewed Generalized Student-t distributed VaR;
- QuantReg – provides procedures to estimate the Quantile Regression VaR;
- Openxlsx and xlsx – allows to export the results from R to Excel;
- Moments – to compute the skewness and kurtosis;
- Urca – unit root tests for time series data.

Below, in Table 1, are the descriptive statistics for the return series. In Figure 1, we present the plotted returns.

Index returns	Mean	Median	Max	Min	Std	Skewness	Kurtosis
S&P500	0,00029	0,00055	0,1096	-0,0947	0,0110	-0,2674	8,8534
DAX	0,00022	0,00075	0,1080	-0,0963	0,0139	-0,1390	4,7884
FTSE	0,00013	0,00032	0,0938	-0,0926	0,0108	-0,1244	6,2440
HSI	0,00025	0,00056	0,1725	-0,1473	0,0154	-0,1176	9,4883

Table 1. Descriptive statistic. Mean, median, maximum and minimum, standard deviation, skewness and kurtosis of the returns for each index for the period in study, 2000 to 2020.

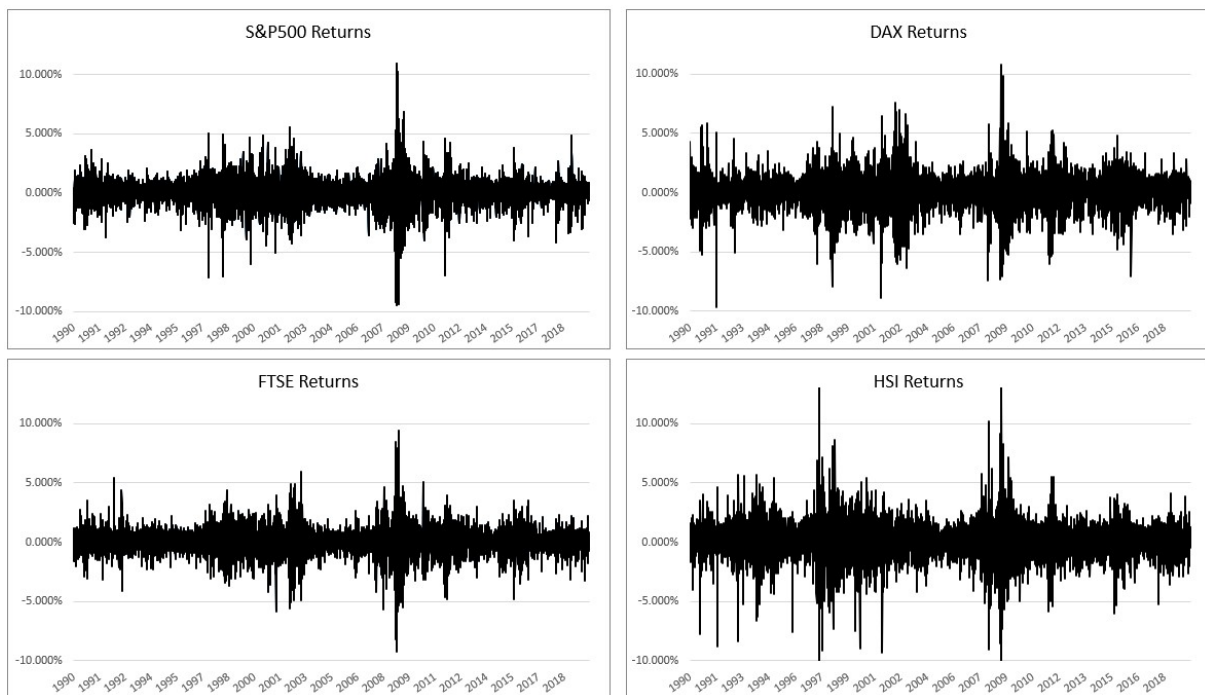


Figure 1. Index returns. Plotted returns of the four indices from 1990 until 2020. Displays stationarity on the returns with volatility clustering.

As expected from financial data, all indices display excess kurtosis ($kurtosis > 3$) on their returns, hence, different from a normal distribution, the unconditional distribution of the returns will reveal fat-tails. Another interesting fact is that all index returns exhibit a negative skewness, meaning that they are asymmetrical, skewed left, with a left tail that is longer than the right tail.

Figure 1 shows that the returns appear to be stationary, but they show volatility clustering. Knowing this, the variance cannot be assumed to be constant over time, and the estimation of the volatility recurring to GARCH models appears to be an appropriate choice.

3.2. Stationarity test

The ADF test is an augmented version of the Dickey-Fuller (DF) test . It tests the null hypothesis that a unit root is present in a given time series, which will imply that the time series is non-stationary, versus the alternative hypothesis that the time series is stationary.

Under the null hypothesis, the t-statistic doesn't follow a student-t distribution, not even asymptotically. The critical values must be taken from an appropriate distribution of t, which will be nonstandard and asymmetric negative, leading to smaller critical values than those for a student-t distribution. Thus, the critical values that will be used to decide about the stationarity of the time series will be the ones provided by Dickey and Fuller (1979).

Different from the ADF test, the KPSS test is used to test the null hypothesis that the given time series is stationary, against the alternative hypothesis of the existence of a unit root, non-stationarity.

Table 2 summarizes the results obtained for both tests. It includes the p-value of each test for the four indices, for both prices and returns.

Index	Data	ADF	KPSS	Results
S&P500	Prices	-0.52	6.59**	Non-stationary
	Returns	-65.90**	0.10	Stationary
DAX	Prices	-2.35	5.37**	Non-stationary
	Returns	-62.77**	0.05	Stationary
FTSE	Prices	-2.72	4.04**	Non-stationary
	Returns	-39.09**	0.05	Stationary
HSI	Prices	-3.88	0.98**	Non-stationary
	Returns	-43.23**	0.05	Stationary

Table 2. ADF and KPSS tests results. Allows to conclude about the stationarity of the prices and the returns. For the ADF test the 5% and 1% critical value are -3.41 and -3.96, respectively. For the KPSS test 0.15 and 0.22, respectively. * and ** denote the rejection of the null hypothesis at 5% and 1% significance levels, respectively.

From the results of the ADF test, the null hypothesis of the existence of a unit root is rejected at a 1% significance level for the returns. For the prices, there is no statistical evidence that allows us to reject the existence of a unit root. That said, the return series can be said to be a stationary process, while the prices are non-stationary. The same conclusion is reached through the KPSS, the null hypothesis, which says that the series is stationary, is only rejected for the prices.

3.3. Long Memory test

Lo's R/S is a standard test to assess the presence of long memory in time series, particularly in stocks and indices. It's a statistic that serves as an extension from the original R/S statistic, proposed by Hurst (1951). Lo (1991) modified the rescale range, constructing a test statistic that is robust to short-range dependence, and derive the limiting distribution under short and long-range dependence.

Table 3 presents the p-values of Lo's R/S test for the returns and the two volatility proxies.

Index	Returns	Returns²	 Returns
S&P500	0,77009	3,4324**	5,0798**
DAX	0,61254	3,4067**	4,3558**
FTSE	0,58373	3,4767**	4,5500**
HSI	0,57195	2,6960**	3,6314**
Results	Short Memory	Long Memory	Long Memory

Table 3. Lo's R/S test results. Allows to conclude about the long memory of the returns and the volatility. The volatility is accessed recurring to two proxies, Returns² and |Returns|. The 5% and 1% critical value for the Lo's R/S test are 1.747 and 2.098, respectively. * and ** denote the rejection of the null hypothesis for these significance levels.

The output, allows us to conclude that the returns do not have a long memory characteristic, given that the results do not reject the null hypothesis. For the squared returns and the absolute-valued returns, the volatility proxies, the results allow us to reject the null for a 1% and 5% significance level. That means that the proxies are described by a long memory process, which indicates that the FIGARCH model is a good candidate to fit the volatility.

METHODOLOGY

4.1. Volatility Models

Volatility, σ , is defined as the standard deviation of the daily returns. In this dissertation we focus on the conditional volatility of the daily returns, σ_t . To model this volatility five models are used: two EWMA volatility models, GARCH, EGARCH and FIGARCH.

Each model has its own particularity. The EWMA is the model used in the Risk-Metrics model and one of the most common and simple models. The GARCH is the base from where various models were developed, furthermore, is one of the most popular volatility models. The EGARCH, based on the GARCH, adds the possibility to model the asymmetric behavior of the volatility. And lastly, the FIGARCH, also based on the GARCH, was developed to allow the long memory of the volatility to be modeled. Using these models, we explore the relevance of modeling the asymmetric behavior of the volatility and its long memory, through the EGARCH and FIGARCH, respectively, while comparing them with more simple and common models, EWMA and GARCH.

Figure 2 illustrates how the estimation of the parameters is done throughout this dissertation. Given that our attention is revolving around the out-of-sample estimation of the conditional variance, one important step in the volatility estimation is the re-estimation of the parameters. Given that the conditions of the financial markets change daily it's important to keep the parameters of the models up to date. To do so, and knowing that the FIGARCH requires at least one thousand lags to truncate its expansion, we use four years of historical data to estimate the parameters. The parameters are re-estimated every trading month, twenty-one trading days, to ensure that the volatility reflects the current state of the market.

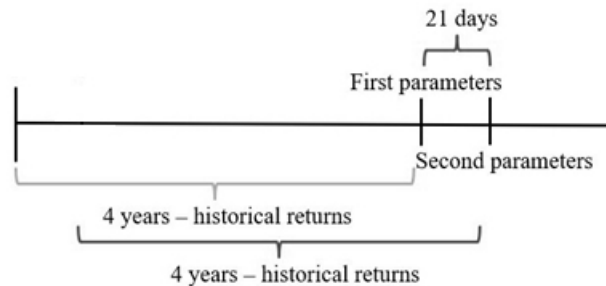


Figure 2. Parameters re-estimation procedure. The parameters of the volatility models are estimated using four years of historical data, and re-estimated every trading month, 21 days, to ensure a high precision on the estimated parameters.

The next sub-chapters will introduce each of these models and present the procedures employed to compute the conditional variance.

4.1.1. GARCH

The GARCH model was introduced by Bollerslev (1986), it works as a generalization of the Autoregressive Conditional Heteroskedasticity (ARCH) model, presented by the Nobel Prize Winner Engle (1982). When compared with the ARCH model, a lower order GARCH can have the same properties of a higher order ARCH without the need to estimate many parameters (Bollerslev, 1986).

The GARCH model is particularly attractive since it captures volatility clustering and exhibits excess kurtosis, properties that might improve the accuracy of the volatility estimation for financial data. Two important features are that the GARCH is a symmetric model, which means that positive shocks have the same impact on the conditional volatility as negative shocks, and that it only accommodates short-term persistence in the volatility.

Under a $GARCH(p, q)$ model the conditional variance can be written as:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^p \beta_j \sigma_{t-1}^2, \quad (2)$$

where $\omega > 0$, $\alpha_i \geq 0$, and $\beta_j \geq 0$. The returns are modeled following:

$$r_t = \mu + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2) \quad (3)$$

where μ is the mean value of the returns.

The p and q define the lag order of the variance of the residuals (squared errors) and the variance of the process, respectively. For instance, for a $p = 0$ the process is reduced to an $ARCH(q)$. While various p 's and q 's can be defined, generally a $GARCH(1, 1)$ is enough to describe the volatility of financial data, thus being the one used in our study Brooks (2008).

The $GARCH(1, 1)$ conditional variance is given by:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (4)$$

All parameters are estimated by maximum likelihood. This implementation is done for the GARCH model, but also for all the other volatility models studied in this dissertation, which are no more than extensions or particular cases of the GARCH model.

4.1.2. EWMA

The EWMA is no more than a special case of the GARCH model with $\omega = 0$ and $\alpha + \beta = 1$. It is one of the simplest and most used volatility models, the one used in the RiskMetrics model, developed by Longerstae and Spencer (1996). Instead of giving the same weight to all observations, the most recent observations have a higher impact on the volatility. This relation is defined by the constant denoted by λ . The larger the value of λ the less weight that is assigned to older observations. While the choice of λ is subjective, it can vary between 0 and 1, for example, the RiskMetrics model uses a λ of 0.94. Another

option to define the value for λ is to fit it to the data using the maximum likelihood function.

An advantage of this model is that it does not suffer from “ghost features”, comparing it with models that assign the same weight to all observations through time. As time passes older returns have a lower impact on the estimated volatility, that is, the impact of an extreme return becomes diluted through time.

Substituting $\omega = 0$, $\beta = \lambda$ and $\alpha = 1 - \lambda$ into Equation 4, we compute the conditional variance recursively as:

$$\sigma_t^2 = (1 - \lambda)r_{t-1}^2 + \lambda\sigma_{t-1}^2, \quad (5)$$

where σ_t^2 is the variance computed at day t and λ is the smoothing constant.

4.1.3. EGARCH

Contrary to the GARCH model, that considers that negative and positive shocks have the same impact on the volatility, the EGARCH model, introduced by Nelson (1991), features an asymmetry coefficient, γ , that allows the asymmetric behavior of the volatility to be accounted for. Various authors have found out that negative returns tend to be followed by periods of higher volatility, in that sense, it's typical to say that bad news impact more the volatility than good news (Black, 1976; Christie, 1982). The γ coefficient allows to model this dynamic. For instance, if $\gamma_i < 0$, the variance increases more when ϵ_{t-1} is negative.

Another important aspect of the EGARCH model is that the conditional variance is computed through a logarithmic transformation, meaning that, no restrictions on the parameters are required to ensure that the volatility is strictly higher than zero.

The $EGARCH(p, q)$ conditional variance is computed as:

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \frac{|\epsilon_{t-i}|}{\sigma_{t-i}} + \sum_{i=1}^q \gamma_i \frac{\epsilon_{t-i}}{\sigma_{t-i}} + \sum_{i=1}^p \beta_i \ln \sigma_{t-i}^2 \quad (6)$$

As in the GARCH model, we work under a $EGARCH(1, 1)$, which is given by:

$$\ln \sigma_t^2 = \omega + \alpha \frac{|\epsilon_{t-1}|}{\sigma_{t-1}} + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \beta \ln \sigma_{t-1}^2 \quad (7)$$

4.1.4. FIGARCH

A problem related with short-memory GARCH models, like the standard GARCH and the EGARCH, is that those models assume that shocks decay at a fast-geometric rate, allowing only short-term persistence. To overcome this problem, Engle and Bollerslev (1986) developed the Integrated GARCH (IGARCH), that contrary to short-memory models, is characterized by having an infinite memory. Having an infinite memory is also not an ideal feature since it means that shocks in the IGARCH volatility will never die out, the initial conditions would impose hard bias on the estimations.

Baillie et al. (1996) introduced the FIGARCH, a long-memory volatility model, that, instead of assuming an infinite memory, implies a slow hyperbolic rate of decay for the

lagged squared innovations in the conditional variance. Starting from Equation 8, which is the equation for a generic $GARCH(p, q)$:

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2, \quad (8)$$

where L is the lag operator, such as $\alpha(L) = \sum_{i=1}^{m-1} \alpha_i(L)$ and $\beta(L) = \sum_{i=1}^{m-1} \beta_i(L)$. Rearranging it and letting $v_t = \varepsilon_t^2 - \sigma_t^2$, the FIGARCH (p,d,q) model is given by:

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1 - \beta(L)]v_t, \quad (9)$$

where $m = \max\{p, q\}$ and $0 \leq d \leq 1$ is the fractional differencing parameter that measures the degree of long memory. The $(1-L)^d$ is obtained as:

$$(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} L^k = 1 - \sum_{k=0}^{\infty} \pi_k L^k, \quad (10)$$

where $\pi_k = \prod_{1 < k \leq i} \frac{k-1-d}{k}$. The expansion needs to be truncated to a large number, most commonly 1000 as Baillie et al. (1996) suggest.

Setting $\phi(L) = (1 - \alpha(L))$ and $(1-L)^d \varepsilon_t^2 = \varepsilon_t^2 + (\sum_{k=0}^{1000} \pi_k L^k) \varepsilon_t^2 = \varepsilon_t^2 + \bar{\varepsilon}_t^2$ (truncating the expansion to 1000 lags), where ε_t^2 is the squared error term and $\bar{\varepsilon}_t^2$ is the squared average error term, we obtain the conditional variance as:

$$\sigma_t^2 = (\omega - \bar{\varepsilon}_t^2) + \sum_{i=1}^q \alpha_i (\varepsilon_{t-i}^2 + \bar{\varepsilon}_{t-i}^2) + \sum_{i=1}^p \beta_i (\sigma_{t-i}^2 + \varepsilon_{t-i}^2) \quad (11)$$

For the $FIGARCH(1, d, 1)$, the specification we use in this dissertation, the conditional variance is computed as:

$$\sigma_t^2 = (\omega - \bar{\varepsilon}_t^2) + \alpha (\varepsilon_{t-1}^2 + \bar{\varepsilon}_{t-1}^2) + \beta (\sigma_{t-1}^2 + \varepsilon_{t-1}^2) \quad (12)$$

4.1.5. Estimating the models

To estimate the volatility, using the models presented above, we use R. With our function *vol_estimation*, Appendix D.2, we are able to estimate the volatility with each model. As mentioned previously, all parameters needed to compute the volatility are estimated by maximizing the log likelihood function.

Using the EWMA we'll compute two sets of volatilities, the first one, following the implementation of the RiskMetrics model, assigning the value of 0.94 to λ (named EWMA $\lambda = 0.94$). For the second one we estimate λ fitting it to four years of past returns and re-estimating it every trading month (referred to as EWMA), ensuring a high precision. The initialization of both implementations is computed as $\sigma^2 = r^2$. A more in-depth analysis on the difference between the two is available at Appendix A.

For the other three models in study, GARCH, EGARCH and FIGARCH, the same process is applied. The parameters are estimated following the same principles and the conditional variance is computed as shown in the previous sub-chapters.

The first volatilities estimated correspond to the beginning of 1996, which gives us four additional years, aside from the twenty in study, to be able to compute the VaR out-of-sample. This is especially important in the EWMA volatility model since it allows to dilute any possible problems with the initialization and arrive at a more proper estimate for the twenty years that are studied.

4.2. Value-at-Risk Models

The Value-at-Risk is a statistical measure of expected loss and it's currently the standard metric to assess market risk. It can be defined as the maximum expected loss of a portfolio, for a given significance level (α) and time horizon (h). Throughout our work we work with $\alpha = 1\%$, 2.5% and 5% , which means that we are 99% , 97.5% and 95% , respectively, sure that a loss in our portfolio will not exceed the estimate VaR. As for the time horizon, we estimate 1-day VaRs ($h = 1$) which allows us to disregard the discounting effect without creating major interferences on the results.

Formally, the $100\alpha\%$ h -day VaR ($VaR_{h,\alpha}$) is minus the α quantile of the h -day discounted return distribution. It is a statistic that is related with the following probability:

$$P(X < x_\alpha) = \alpha \quad (13)$$

The quantile α of the distribution, x_α , can then be obtained as:

$$x_\alpha = F^{-1}(\alpha), \quad (14)$$

where, F^{-1} represents the inverse of the distribution function.

As the VaR represents the maximum loss that won't be exceeded for a certain α , it corresponds to the α quantile of the h -day return distribution:

$$VaR_{h,\alpha} = -F^{-1}(\alpha) \quad (15)$$

Replacing the previous equation onto Equation 14:

$$VaR_{h,\alpha} = -x_\alpha \quad (16)$$

In the previous chapter we explored the choice of a volatility model, but another important step to estimate the VaR is to choose the VaR model. These models tend to be divided by the way they model the return's distribution. In this dissertation four models are studied: Parametric Normal and Parametric Skewed Generalized Student-t VaR, Volatility-Adjusted Historical VaR and Quantile Regression VaR.

4.2.1. Parametric VaR

The Parametric VaR assumes a parametric distribution, in our case Normal or Skewed Generalized Student-t, for the return's distribution of the portfolio. While a big advantage of this model is the analytical tractability, in the sense that we only need to estimate the distribution parameters, this model imposes a parametric distribution that can be unrealistic.

4.2.1.1. Normally distributed

We start under the assumption that the returns of our portfolio are normally distributed, $X_h \sim N(\mu_h, \sigma_h^2)$, where μ_h and σ_h^2 are the estimates for the mean and variance, respectively.

Starting from Equation 16 and applying the equivariance of quantiles to positive transformations³, Koenker and Bassett Jr (1978), the $100\alpha\%$ h -day Parametric Normal VaR, is given as:

$$VaR_{h,\alpha} = -\phi^{-1}(\alpha) \times \sigma_h - \mu_h \quad (17)$$

where $\phi^{-1}(\alpha)$ is the standard normal α quantile.

As Alexander (2009) suggests, we use a μ_h equal to zero. This makes the computation of the VaR easier and only has big impacts in the value when the risk horizon is higher than one month. Furthermore, there are no sufficiently accurate ways to estimate the value of μ_h , so adding another estimation to the process would only increase the computation complexity. With this simplification, the $100\alpha\%$ h -day Parametric Normal VaR is computed recurring to the *normal_var* function, Appendix D.3, as:

$$VaR_{h,\alpha} = -\phi^{-1}(\alpha) \times \sigma_h \quad (18)$$

4.2.1.2. Skewed Generalized Student-t distributed

It is known that financial returns don't typically follow a normal distribution, they have a leptokurtic distribution, with fat-tails. Thus, fitting the returns under a normal distribution, will most likely lead to underestimated VaRs for low significance levels.

The Skewed generalized student-t distribution (Theodossiou, 1998) is a generalization of the Student-t distribution that allows asymmetry and extra flexibility in the shape of the tail and central regions.

Through the equivariance of quantiles to positive transformations, the $100\alpha\%$ h -day Parametric SGSt VaR is :

$$VaR_{h,\alpha} = -T_{0,1,\lambda,p,q}^{-1}(\alpha) \times \sigma_h - \mu_h, \quad (19)$$

where $T_{0,1,\lambda,p,q}^{-1}(\alpha)$ is the quantile of the standard SGSt distribution.

As it's possible to see on Equation 19, to compute the SGSt VaR we need to estimate three parameters: λ that determines the skewness, p that controls the shape of the central region and lastly q that controls the shape of the tail. These are estimated resorting to the *sgt_fit* function, maximizing the likelihood function, Appendix D.4, through the Maximum Likelihood method available in the *sgt* package and based on a sample of four historical trading years, just like in the volatility estimation. Moreover, to ensure that the parameters reflect the current state of the market, this function re-estimates the parameters every trading month.

³ $Q_{g(X)}(\alpha) = g(Q_X(\alpha))$, where $Q_X(\alpha)$ denotes the α quantile of a random X variable which follows a distribution F , and $g(x)$ is an increasing function of x

Finally, with the parameters estimated, the Parametric SGSt VaR is computed using the *sgt_var* function, Appendix D.5.

4.2.2. Volatility-Adjusted Historical VaR

Contrary to the Parametric VaR, the Historical VaR lets the data speak freely. Instead of assuming a parametric distribution, it uses the empirical distribution of the returns.

Formally, the historical $100\alpha\%$ h -day VaR is simply the α quantile of the empirical h -day distribution of the returns. Knowing this, computing the historical VaR is a trivial task: we choose a sample size (n), compute the h -day returns of the portfolio over the sample period, compute the empirical h -day return distribution from those returns (keeping the portfolio weights constant) and find the α quantile that we want to study.

The choice of the sample size is a very important step to compute the Historical VaR. Different sample sizes come with trade-offs, a larger sample will increase the precision of the empirical distribution, however, the more data we use the less the historical distribution reflects the current market conditions.

A problem that arises on the simple historical VaR is that the same weight is assigned to all observations, which will most likely have a negative impact in the VaR estimation. To overcome this problem, Duffie and Pan (1997) and Hull and White (1998) proposed a methodology that adjusts the volatility of the returns to better reflect the current volatility, instead of the historical volatility. The same weight is assigned to every observation, but the series of returns is adjusted with the series of volatility estimates $\hat{\sigma}_t$ and today's volatility $\hat{\sigma}_T$:

$$\hat{r}_t = \frac{\hat{\sigma}_T}{\hat{\sigma}_t} r_t, \quad (20)$$

where $\hat{\sigma}_t$ represents the volatility estimate for day t computed at the end of day $t - 1$. If we are interested in adjusting r_t we have to wait until the end of day t , but by then, we already know σ_{t+1} . Therefore, we can improve the volatility adjustment even more by doing the following:

$$\hat{r}_t = \frac{\hat{\sigma}_{T+1}}{\hat{\sigma}_{t+1}} r_t \quad (21)$$

With the adjusted returns, we can compute the Volatility-Adjusted Historical VaR. At least, to compute the VaR for $\alpha=1\%$ we need 100 daily observations. In our case, we chose to use a sample of two years (504 observations), a large enough sample size to have a good precision of the empirical distribution while allowing the data to be as recent as possible, ensuring that the historical distribution reflects at its best the current market conditions. This is done with the *historical_var* function, Appendix D.6, which deals with both the returns adjustment and the VaR computation.

4.2.3. Quantile Regression VaR

Given that the VaR_α is defined as minus the α -quantile of the return distribution, the VaR estimates can be obtained from a quantile regression of the portfolio returns onto some explanatory variables. Assuming that r represents the returns of the portfolio, the

VaR can be estimated as:

$$VaR_\alpha = -q_{r,\alpha} = -\left(\hat{a}_i + \hat{b}_{xi}\right) \quad (22)$$

In our study, to compute the Quantile Regression VaR, we chose to use one of the simplest specifications, and one of the most attractive, which is the following:

$$r_t = b_i \times \sigma_t + \varepsilon_t, \quad (23)$$

where σ_t is the date t volatility estimate for the portfolio returns computed using one of the four volatility models studied in our work and b is the coefficient related with the explanatory variable.

The parameters are estimated maximizing the likelihood function, recurring to the function *quantile_fit*, Appendix D.7. Like in the past cases, to assure that the coefficients are up to date, they are re-estimated every trading month. While this estimation is based on a historical sample of two trading years and it's based on historical data, the estimation of the VaR is done recurring to the current values of the explanatory variables, the volatility in our case, which guarantees that the VaR reflects the current conditions of the market, much like in the Volatility-Adjusted Historical VaR.

Applying the Equation 22 to our specification, and recurring to the *quantile_var* function, Appendix D.8, the VaR is estimated as:

$$VaR_{\alpha,t} = -\hat{b} \times \sigma_t \quad (24)$$

4.3. Evaluating the performance of VaR models

Our objective is to evaluate the performance of the volatility forecasting techniques, and compare the performance of the various VaR models, estimated using these volatility models. To do so, a backtest composed of two test will be used. The two tests are the Unconditional Coverage (UC) test (Kupiec, 1995) and the BCP test (Berkowitz et al., 2011). The first one is a likelihood test that evaluates the exceedance rate of the estimates, while the second one is a Ljung-Box, which tests for autocorrelation.

From the exceedance rate viewpoint, a well specified VaR model is expected to match the number of exceedances for the chosen significance level. If the observed number of exceedances is much higher, or lower, than the expected value, the model can be said to be badly specified. For instance, for a daily VaR run through a year, and for a significance level of 99%, we expect to have $0.01 \times 252 = 2.52$ exceedances per trading year. Big deviations from this value will implied the rejection that the model is well specified.

Looking at the autocorrelation, it is ideal to avoid exceedance clustering. An exceedance clustering indicates that a model is not fast enough to adapt to the changes in the market conditions, which implies that the VaR is being under or overestimated, depending on the evolution of the market's volatility. Moreover, when VaR exceedances are clustered, some statistical tests, that don't account for autocorrelation, might show good performances, when the risk is being wrongly estimated.

Nevertheless, the most important feature of a model is the exceedance rate, so the preferred source of decision is the UC test. The BCP test is used as a complement to provide more insights on the performance of the models and to compare models that perform similarly according to the UC test.

Both tests are performed for the full period of twenty years, but also for smaller intervals of four years. This analysis allows us to evaluate the performance of the models from a global viewpoint, but also explore if periods of higher stress, for instance the Great Recession, are better estimated by a particular model. In any case, the tests for the full period are more reliable and less volatile, so those should be preferred.

The two tests are computed through the *backtest* function, Appendix D.9, which includes procedures to compute both the exceedances and the autocorrelation for one set of VaR estimates. More details about each test are presented in the next sub-chapters.

4.3.1. Unconditional Coverage test

The UC test is a likelihood ratio test introduced by Kupiec (1995). This test is based on an indicator function that relates with the number of exceedances (N):

$$I_{\alpha,t+1} = \begin{cases} 1, & \text{if } r_{t+1} < VaR_{1,\alpha,t} \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

The null and the alternative hypothesis are defined as follows:

$$\begin{aligned} H_0 : \pi_{obs} &= \pi_{exp} = \alpha \\ H_1 : \pi_{obs} &\neq \pi_{exp}, \end{aligned} \quad (26)$$

where π_{obs} is the observed exceedance rate and π_{exp} is the expected exceedance rate. The test statistic is:

$$LR_{uc} = \frac{\pi_{exp}^{n_1} (1 - \pi_{exp})^{n_0}}{\pi_{obs}^{n_1} (1 - \pi_{obs})^{n_0}}, \quad (27)$$

where n_1 is the observed number of exceedances and n_0 is the number of days without exceedances. Under the null hypothesis the test statistic follows a qui-squared distribution with one degree of freedom:

$$-2 \ln(LR_{uc}) \sim X_1^2 \quad (28)$$

If the null hypothesis is not rejected, the model is assumed to be well specified, that is, the number of expected exceedances is close enough, or equal, to the observed number of exceedances, for the given significance level.

Table 4 displays the acceptance regions for a significance level of 1%, for a sample size of one, four and twenty years. The acceptance regions are presented both in number of exceedances and exceedance rate.

Acceptance Region	1 Year	4 Years	20 Years
Exceedance N^0	$0 \leq N \leq 7$	$2 \leq N \leq 18$	$32 \leq N \leq 69$
Exceedance Rate (%)	[0 ; 2.78]	[0.19 ; 1.79]	[0.64 ; 1.37]

Table 4. UC test acceptance regions. Acceptance region for the 1% significance level backtest, including the interval of exceedances and exceedance rates. If a model presents exceedances outside the presented intervals, the null hypothesis of the UC test is rejected and we conclude that the model is a bad fit to estimate the VaR of the index.

As mentioned before, the backtest will be performed for the full period of twenty years, but also for sub-periods of four years. As Table 2 shows, backtesting the estimates for sub-periods of four years, instead of one year, increases the probability of rejecting the null hypothesis, rising the degree of exigency of the acceptance region.

The *backtest* function includes the procedures to compute the UC test, it provides all the information needed: n , $n1$, $n0$, π_{obs} , test statistic and p-value.

Because we are working with various models and for various indices, there is a lot of information to work with. To keep the information concise and organized, in a way that is easier to manage, we created the function *uctest_table*, Appendix D.10. This function creates a table that contains the p-values for all the models and time intervals in analysis. This way the evaluation of the performance of the various models becomes much easier and simple.

4.3.2. BCP test

The BCP test, (Berkowitz et al., 2011), checks for the first K autocorrelations of exceedances. It's a test that allows to evaluate exceedance clustering. If the VaR model is well specified, the exceedances are independent, meaning that it's impossible to predict when the next exceedance will occur.

The test hypothesis is defined as follows:

$$\begin{aligned}
 H_0 : \hat{\rho}_k &= 0, \forall k \in \{1, \dots, K\} \\
 H_1 : \exists k &\in \{1, \dots, K\} \text{ s.t. } \hat{\rho}_k \neq 0
 \end{aligned}
 \tag{29}$$

And the test statistic is:

$$BCP(K) = T(T + 2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{T - k},
 \tag{30}$$

where $\hat{\rho}_k^2$ is the lag k sample autocorrelation, K is the maximum autocorrelation lag considered and T is the sample size. Under the null hypothesis, the test statistic follows a chi-squared distribution, with K degrees of freedom:

$$BCP(K) \sim X_k^2
 \tag{31}$$

The choice of the parameter K is up to the user. A larger K helps detecting non-independence in higher-order lags, but a smaller K increases the power of the test, given the degrees of freedom are reduced. In this study we start with a K = 1 and go up to a

$K = 5$, this way we are able to explore non-independence up to the lag 5 without losing much power on the test.

Like for the UC test, the *backtest* function is able to compute the BCP test. This function creates the autocorrelation matrix needed to evaluate the autocorrelation of the exceedances and computes the various test statistics and p-values. Since we work with a K up to 5, each model will have five different p-values for each time interval. Resorting to the function *bcptest_table*, Appendix D.11, we organize all p-values for the five K s for all models and time intervals in study, which eases the assessment of the autocorrelation in the exceedances.

CHAPTER 5

BACKTEST

In this section we present the backtest, or validation, of the estimates produced by the various models for each of the four indices. The choice of the models, VaR and volatility, the estimation technique and the sample size are some source of uncertainty that can highly impact the accuracy of the VaR estimates. Therefore, choosing the most adequate combination of models will improve the accuracy of the estimated VaR.

Two tests are used for the backtest, the UC and the BCP test up to five lags. As explained in the methodology, the test that will weigh more in our decision is the UC test, given that the exceedance rate is the most important feature of a model. The BCP test will provide more insight and allow us to compare models that have suitable UC test values. It will only be computed for models that pass the UC test, because it's highly influenced by the number of exceedances, having too many exceedances might cause the model to be accepted by the BCP, when it fails the UC test.

Some model could perform well for the full period in study, 2000 until 2020, but perform rather poorly in specific sub-intervals. To better understand the consistency of the results for each model throughout the years, together with the backtest for the full period, we perform both tests for sub-periods of four years.

A model that yields p-values higher than 5% is accepted by the UC and the BCP test, a p-value smaller than 5% corresponds to a rejection of the null hypothesis. For the UC test, a large p-value implies the model has an exceedance rate that is close to the expected exceedance rate, meaning that it captures well the VaR of the index. For p-values higher or equal than 5% the periods are in the green zone and lower than 5% in the red zone. A model that presents mostly green zone periods can be said to be the most ideal model, since it performs well throughout the years.

5.1. Backtest – 1% Significance Level

In Table 5 below, we present the summary of the backtest for the four indices which includes the UC test p-values as well as the exceedance rates, for the full period and a significance level of 1%. The values in bold denote a model that is within the green zone, a model that is accepted by the UC test.

All models accepted by the UC test, the ones with a p-value higher than or equal to 5%, for the full period, will be further analyzed. These model should provide the most accurate VaR estimates and thus are the most important to investigate.

Model	S&P500		DAX		FTSE		HSI	
	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC
Parametric Normal - GARCH	2.35%	0.00%	1.92%	0.00%	2.17%	0.00%	1.87%	0.00%
Parametric Normal - EGARCH	2.72%	0.00%	2.50%	0.00%	2.45%	0.00%	1.89%	0.00%
Parametric Normal - FIGARCH	2.17%	0.00%	1.78%	0.00%	2.13%	0.00%	1.94%	0.00%
Parametric Normal - EWMA	2.31%	0.00%	1.98%	0.00%	2.39%	0.00%	2.02%	0.00%
Parametric Normal - EWMA $\lambda = 0.94$	2.13%	0.00%	2.06%	0.00%	2.35%	0.00%	1.98%	0.00%
Parametric SGSt - GARCH	1.17%	23,07%	0.83%	21,82%	1.32%	2.88%	0.99%	93,55%
Parametric SGSt - EGARCH	1.47%	0.17%	1.15%	29,67%	1.58%	0.01%	1.03%	83,63%
Parametric SGSt - FIGARCH	1.19%	18,27%	0.71%	3.12%	1.26%	7,28%	1.01%	94,95%
Parametric SGSt - EWMA	1.19%	18,27%	0.73%	4.57%	1.30%	3.98%	1.01%	94,95%
Parametric SGSt - EWMA $\lambda = 0.94$	1.13%	35,33%	0.67%	1.34%	1.16%	26,00%	1.11%	44,09%
Historical - GARCH	1.23%	11,00%	0.91%	52,27%	1.06%	66,03%	1.07%	62,38%
Historical - EGARCH	1.09%	51,27%	1.03%	82,73%	1.10%	47,35%	0.93%	60,32%
Historical - FIGARCH	1.17%	23,07%	0.95%	72,67%	1.14%	32,18%	1.11%	44,09%
Historical - EWMA	0.99%	96,49%	1.11%	44,03%	1.00%	98,64%	1.03%	83,63%
Historical - EWMA $\lambda = 0.94$	0.91%	53,54%	1.11%	44,03%	1.02%	87,38%	1.07%	62,38%
Quantile Regression - GARCH	1.49%	0.11%	1.37%	1.29%	1.40%	0.70%	1.44%	0.37%
Quantile Regression - EGARCH	1.37%	1.22%	1.39%	0.90%	1.44%	0.32%	1.34%	2.42%
Quantile Regression - FIGARCH	1.35%	1.74%	1.37%	1.29%	1.36%	1.45%	1.50%	0.10%
Quantile Regression - EWMA	1.39%	0.84%	1.45%	0.28%	1.40%	0.70%	1.32%	3.37%
Quantile Regression - EWMA $\lambda = 0.94$	1.39%	0.84%	1.55%	0.03%	1.30%	3.98%	1.32%	3.37%

Table 5. Summary of the backtest. For each index we present the exceedance rate and the p-value of the UC test. The values in bold denote the model combinations that were accepted under the UC test, that is, with a p-value higher than 5%. This table includes the results of the backtest done for the global period in study, 2000-2020.

From Table 5 some facts already come in clear for the four indices. Neither the Parametric Normal VaR nor the Quantile Regression VaR work for any index at a significance level of 1%, no matter the volatility model employed. From this fact we already know that the Volatility-Adjusted Historical and the Parametric SGSt are surely the most appropriate when looking at the four indices globally.

5.1.1. S&P500

As mentioned in the summary of the backtest, the individual analysis for each index will only cover the models that have been accepted by the UC test. Table 6 below, presents these models for the S&P500. The values in bold indicate the periods that are within the green zone, a model that is accepted by the test. This connotation will be used in all tables, throughout the analysis of the four indices.

Out of the twenty combinations of models studied, nine passed the test. These nine only include Parametric SGSt and Volatility-Adjusted Historical models. Looking at the p-values for the global period, two Volatility-Adjusted Historical models arise immediately as the most trustworthy, those computed using the EWMA and EWMA $\lambda = 0.94$ volatility, the first with a p-value of 93.49%, showing the largest margin in the UC test, and the

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC
Parametric SGSt - GARCH	1.17%	23.07%	0.90%	73.47%	1.19%	55.30%	1.59%	8.49%	0.60%	16.42%	1.59%	8.30%
Parametric SGSt - FIGARCH	1.19%	18.27%	0.90%	73.47%	0.89%	73.00%	1.49%	14.74%	0.99%	98.48%	1.81%	4.53%
Parametric SGSt - EWMA	1.19%	18.27%	0.90%	30.64%	0.89%	14.55%	1.49%	72.53%	0.99%	37.27%	1.71%	8.30%
Parametric SGSt - EWMA $\lambda = 0.94$	1.19%	35.33%	0.70%	30.64%	1.49%	55.30%	0.89%	72.53%	1.29%	37.27%	1.59%	8.30%
Historical - GARCH	1.23%	11.00%	0.90%	73.47%	1.89%	1.17%	1.29%	37.81%	0.99%	98.48%	1.09%	76.92%
Historical - EGARCH	1.09%	51.27%	0.80%	50.04%	1.99%	0.55%	0.99%	97.72%	0.80%	49.84%	0.89%	73.24%
Historical - FIGARCH	1.17%	23.07%	0.90%	73.47%	1.80%	4.57%	1.59%	8.49%	0.89%	73.24%	0.80%	49.84%
Historical - EWMA	0.99%	96.49%	0.90%	73.47%	1.19%	55.30%	1.19%	55.73%	0.89%	73.24%	0.80%	49.84%
Historical - EWMA $\lambda = 0.94$	0.91%	53.55%	0.70%	30.64%	1.09%	77.16%	0.99%	97.72%	1.09%	76.92%	0.70%	30.50%

Table 6. S&P500 - UC test results. We display the exceedance rate and the p-value of the UC test of all model combinations that were accepted by the UC test for the global period. The values in bold denote the combinations of models that are accepted for each sub-period.

last with 53.55%. Regarding the sub-periods these two combinations also come out as the best models, passing the test for all sub-periods.

Table 7 shows a summary of the BCP test results. It presents the lag that corresponds to the lowest p-value of the BCP test.

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric SGSt - GARCH	2	0.00%	3	0.04%	4	0.00%	2	0.00%	1	0.00%	1	0.00%
Parametric SGSt - FIGARCH	1	0.00%	4	0.09%	4	0.00%	2	0.00%	1	0.00%	1	0.00%
Parametric SGSt - EWMA	1	0.00%	5	63.82%	4	0.01%	3	0.09%	1	0.00%	1	0.00%
Parametric SGSt - EWMA $\lambda = 0.94$	1	0.00%	5	63.82%	5	0.09%	3	0.09%	1	0.00%	1	0.00%
Historical - GARCH	1	0.00%	5	53.88%	5	0.09%	2	0.00%	1	0.00%	1	0.00%
Historical - EGARCH	2	0.00%	5	58.75%	4	0.00%	2	0.00%	2	0.00%	1	0.00%
Historical - FIGARCH	1	0.00%	3	0.04%	5	1.34%	2	0.00%	1	0.00%	1	0.00%
Historical - EWMA	5	0.97%	5	53.89%	5	38.95%	5	1.53%	4	0.09%	1	0.02%
Historical - EWMA $\lambda = 0.94$	2	0.00%	5	63.82%	5	47.07%	5	0.30%	5	0.00%	5	0.00%

Table 7. S&P500 - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

It's possible to understand that all models perform badly, not passing the BCP test for the global period. Like in the UC test, the Volatility-Adjusted Historical EWMA and EWMA $\lambda = 0.94$ have a slight advantage, being the only two that pass the BCP test for two sub-periods.

Given all the analysis above, for the S&P500, during the period in study, the best combination of models to estimate the VaR is the Volatility-Adjusted Historical EWMA, since it presents a higher level of confidence than the Volatility-Adjusted Historical EWMA $\lambda = 0.94$ on the UC test.

5.1.2. DAX

As Table 8 below displays, seven models comply with the acceptance criteria of the UC test. Again, these seven models only include Parametric SGSt and Volatility-Adjusted Historical models. Between these two classes of models, the Volatility-Adjusted Historical models disclose more confidence when compared with the Parametric SGSt.

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC
Parametric SGSt - GARCH	0.83%	21.82%	0.70%	31.94%	0.59%	15.47%	1.47%	15.91%	0.69%	30.22%	0.70%	30.64%
Parametric SGSt - EGARCH	1.15%	29.67%	0.60%	17.33%	0.88%	70.69%	2.06%	0.29%	1.09%	77.41%	1.09%	76.67%
Historical - GARCH	0.91%	52.27%	0.70%	31.94%	0.98%	95.72%	1.08%	80.61%	0.60%	16.25%	1.19%	54.86%
Historical - EGARCH	1.03%	82.73%	0.90%	75.60%	1.08%	79.63%	1.18%	58.38%	0.79%	49.46%	1.19%	54.86%
Historical - FIGARCH	0.95%	72.67%	0.70%	31.94%	1.08%	79.63%	1.08%	80.61%	0.69%	30.22%	1.19%	54.86%
Historical - EWMA	1.11%	44.03%	0.80%	51.81%	1.47%	15.51%	1.18%	58.38%	0.89%	72.77%	1.19%	54.86%
Historial - EWMA $\lambda = 0.94$	1.11%	44.03%	0.90%	75.60%	1.47%	15.51%	0.98%	94.72%	0.99%	97.98%	1.19%	54.86%

Table 8. DAX - UC test results. We display the exceedance rate and the p-value of the UC test of all model combinations that were accepted by the UC test for the global period. The values in bold denote the combinations of models that are accepted for each sub-period.

Looking at the global period, the model that outputs the largest p-value is the Volatility-Adjusted Historical EGARCH, 82.73%, followed by the Volatility-Adjusted Historical FIGARCH and GARCH, 72.67% and 52.27%, respectively. For the sub-periods, all Volatility-Adjusted Historical models and the Parametric SGSt - GARCH perform well, having all sub-periods in the green area.

Table 9 shows the results of the BCP test for the DAX. From it, we understand that all models fail the BCP test for the global period. This is largely due to the sub-periods of 2002 to 2004, 2008 to 2012 and 2016 to 2020, where almost all models fail the BCP test.

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric SGSt - GARCH	2	0.00%	1	0.00%	5	67.04%	2	0.00%	5	61.76%	5	61.71%
Parametric SGSt - EGARCH	2	0.00%	1	0.00%	5	52.20%	2	0.00%	5	0.77%	2	0.00%
Historical - GARCH	2	0.00%	3	0.00%	5	0.00%	2	0.00%	5	66.90%	5	1.11%
Historical - EGARCH	2	0.00%	4	0.00%	5	0.02%	2	0.00%	5	56.79%	5	1.55%
Historical - FIGARCH	2	0.00%	3	0.00%	5	0.02%	2	0.00%	5	61.76%	5	1.11%
Historical - EWMA	5	3.45%	5	56.56%	5	1.17%	5	1.46%	5	52.02%	5	0.09%
Historial - EWMA $\lambda = 0.94$	5	0.67%	5	0.09%	5	1.17%	5	0.28%	5	47.44%	5	0.09%

Table 9. DAX - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

For the DAX, the combination of models that better estimates the VaR is the Volatility-Adjusted Historical EGARCH. It produces good results on the UC test, with more confidence than the other models.

5.1.3. FTSE

Table 10 displays the UC test results for the FTSE.

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC
Parametric SGSt - FIGARCH	1.26%	7.28%	1.42%	21.26%	1.51%	13.27%	1.30%	36.56%	1.30%	36.04%	0.79%	49.07%
Parametric SGSt - EWMA $\lambda = 0.94$	1.16%	26.00%	1.01%	96.43%	1.31%	34.99%	0.80%	50.62%	1.50%	13.81%	1.19%	55.95%
Historical - GARCH	1.06%	66.03%	1.01%	96.43%	1.21%	52.26%	0.50%	7.76%	1.50%	13.81%	1.09%	77.90%
Historical - EGARCH	1.10%	47.35%	1.12%	72.02%	1.41%	22.15%	0.50%	7.76%	1.20%	53.56%	1.29%	37.99%
Historical - FIGARCH	1.14%	32.18%	1.12%	72.02%	1.41%	22.15%	0.70%	31.07%	1.30%	36.04%	1.19%	55.95%
Historical - EWMA	1.00%	98.64%	1.01%	96.43%	1.01%	98.22%	0.90%	74.17%	1.10%	75.20%	0.99%	97.47%
Historial - EWMA $\lambda = 0.94$	1.02%	87.38%	1.01%	96.43%	1.21%	52.26%	0.70%	31.07%	1.00%	99.75%	1.19%	55.95%

Table 10. FTSE - UC test results. We display the exceedance rate and the p-value of the UC test of all model combinations that were accepted by the UC test for the global period. The values in bold denote the combinations of models that are accepted for each sub-period.

For the global period, seven models are in the green zone. Like the two previous indices, these seven are only Parametric SGSt and Volatility-Adjusted Historical VaR models. The models that achieved the best results on the UC test, are the Volatility-Adjusted Historical EWMA and the Volatility-Adjusted Historical EWMA $\lambda = 0.94$, 98.64% and 87.38%, respectively, followed by the Volatility-Adjusted Historical GARCH with 66.03%. Looking at the sub-periods, all models achieve positive results, having all sub-periods in the green zone.

In Table 11 we have the results of FTSE's BCP test. We can see that all models stay in the red zone of the BCP test, for the global period and most of the sub-periods.

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric SGSt - FIGARCH	1	0.00%	4	0.00%	3	0.00%	2	0.00%	5	0.03%	1	0.02%
Parametric SGSt - EWMA $\lambda = 0.94$	1	0.00%	4	0.00%	4	0.00%	5	56.68%	3	0.00%	5	1.53%
Historical - GARCH	1	0.00%	4	0.00%	3	0.00%	2	0.00%	3	0.00%	1	0.00%
Historical - EGARCH	1	0.00%	3	0.00%	3	0.00%	5	72.11%	5	0.01%	1	0.00%
Historical - FIGARCH	1	0.00%	4	0.00%	3	0.00%	5	0.00%	1	0.00%	1	0.00%
Historical - EWMA	5	1.01%	5	0.34%	5	0.33%	5	0.09%	5	0.80%	5	47.49%
Historial - EWMA $\lambda = 0.94$	5	1.25%	5	0.34%	5	1.61%	1	0.00%	1	0.00%	5	39.02%

Table 11. FTSE - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

Due to its' performance on the UC test, and given that all models have similar performances on the BCP test, the Volatility-Adjusted Historical EWMA should be the best combination to estimate the VaR for the FTSE between 2000 and 2020.

5.1.4. HSI

In Table 12 we present the results of the UC test for the HSI index.

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC	Exc. Rate	UC
Parametric SGSt - GARCH	0.99%	93.55%	0.64%	22.78%	1.82%	3.63%	0.62%	19.38%	0.82%	55.47%	1.12%	70.55%
Parametric SGSt - EGARCH	1.03%	83.63%	0.74%	40.29%	1.84%	1.84%	0.72%	35.13%	0.72%	34.66%	1.12%	70.55%
Parametric SGSt - FIGARCH	1.01%	94.95%	0.53%	11.06%	1.73%	3.63%	1.03%	93.61%	0.72%	34.66%	1.02%	94.90%
Parametric SGSt - EWMA	1.01%	94.95%	0.95%	88.47%	1.22%	49.50%	0.62%	19.38%	1.23%	49.08%	1.02%	94.90%
Parametric SGSt - EWMA $\lambda = 0.94$	1.11%	44.09%	1.06%	85.60%	1.22%	49.50%	0.82%	56.10%	1.12%	70.07%	1.33%	32.79%
Historical - GARCH	1.07%	62.38%	0.64%	22.78%	1.84%	1.84%	0.92%	80.68%	0.92%	79.94%	1.02%	94.90%
Historical - EGARCH	0.93%	60.32%	1.06%	85.60%	1.22%	49.50%	0.82%	56.10%	0.82%	55.47%	0.71%	34.35%
Historical - FIGARCH	1.11%	44.09%	0.85%	62.86%	1.63%	6.82%	1.23%	48.46%	0.82%	55.47%	1.02%	94.90%
Historical - EWMA	1.03%	83.63%	0.95%	88.47%	1.02%	94.90%	0.82%	56.10%	1.12%	70.07%	1.22%	49.50%
Historical - EWMA $\lambda = 0.94$	1.07%	62.38%	1.06%	85.60%	0.82%	55.06%	1.44%	19.91%	1.12%	70.07%	0.92%	79.45%

Table 12. HSI - UC test results. We display the exceedance rate and the p-value of the UC test of all model combinations that were accepted by the UC test for the global period. The values in bold denote the combinations of models that are accepted for each sub-period.

The HSI was the index in which more models performed adequately, out of the twenty in study, half of them are in the green area for the global period, some with outstanding p-values on the UC test. From these ten we'd like to highlight the Parametric SGSt GARCH, FIGARCH and EWMA, which achieved p-values of 93.55%, 94.95% and 94.95%.

From all the models that passed the UC test, six did it for all periods, Parametric SGSt EWMA and EWMA $\lambda = 0.94$ and all Volatility-Adjusted Historicals except the GARCH.

Table 13 presents the results for the BCP test.

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric SGSt - GARCH	3	0.00%	1	0.00%	3	0.00%	2	0.00%	2	0.00%	3	0.00%
Parametric SGSt - EGARCH	3	0.00%	1	0.00%	3	0.00%	1	0.00%	5	61.22%	5	0.87%
Parametric SGSt - FIGARCH	3	0.00%	5	73.60%	3	0.00%	2	0.01%	2	0.00%	5	0.00%
Parametric SGSt - EWMA	5	0.83%	5	52.57%	5	0.11%	2	0.00%	5	1.71%	5	42.72%
Parametric SGSt - EWMA $\lambda = 0.94$	5	3.19%	5	47.81%	5	0.11%	4	0.00%	5	42.37%	5	0.19%
Historical - GARCH	4	0.00%	5	68.09%	5	0.00%	4	0.01%	4	0.00%	5	0.35%
Historical - EGARCH	3	0.00%	5	0.44%	3	0.00%	2	0.02%	2	0.02%	3	0.00%
Historical - FIGARCH	5	0.00%	5	57.54%	5	0.00%	5	1.72%	2	0.02%	5	0.35%
Historical - EWMA	5	1.03%	5	52.57%	5	0.35%	2	0.02%	5	0.87%	5	38.28%
Historical - EWMA $\lambda = 0.94$	5	1.58%	5	47.81%	5	56.23%	5	0.73%	5	87.00%	5	10.00%

Table 13. HSI - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

The results of the BCP test were not satisfactory, all models failed the BCP test for the global period. Looking at the sub-periods of the BCP test, all models have similar performances, passing it at maximum for two sub-periods.

Given both tests results, the Volatility-Adjusted Historical EWMA should be the best combination to estimate the VaR of the HSI.

5.2. Backtest – 2.5% and 5% Significance Levels

Following the same analysis done for the 1% significance level, the performance of the various models was assessed for each index for both a 2.5% and 5% significance level. The complete results are available at Appendix B and Appendix C.

Table 14 summarizes the findings about the best model for each index.

Index	Significance Levels	
	2.5%	5%
S&P500	Historical - EWMA	Historical - EWMA $\lambda = 0.94$
DAX	Historical - EGARCH	Historical - EWMA $\lambda = 0.94$
FTSE	Historical - GARCH	Historical - EWMA
HSI	Historical - EWMA	Historical - FIGARCH

Table 14. Backtest results summary for the 2.5% and 5% significance level. Highlighting the best combination of models for each index and significance level.

Between these results and the ones presented in the previous sub-sections, no significant differences were found. While some models that weren't previously accepted were accepted, some Parametric Normals and Quantile Regression VaR models, the confidence levels provided by these was still inferior than the other models. For the BCP test, the results remained unchanged, all models failed the test for the global period.

5.3. Backtesting Conclusions

In Table 15 below, we present the models that output the best estimates, when combining the results from the UC test and the BCP test.

Index	Significance Levels		
	1%	2.5%	5%
S&P500	Historical - EWMA $\lambda = 0.94$	Historical - EWMA	Historical - EWMA $\lambda = 0.94$
DAX	Historical - EGARCH	Historical - EGARCH	Historical - EWMA $\lambda = 0.94$
FTSE	Historical - EWMA	Historical - GARCH	Historical - EWMA
HSI	Historical - EWMA $\lambda = 0.94$	Historical - EWMA	Historical - FIGARCH

Table 15. Backtest results summary for all significance levels studied. Highlighting the best combination of models for each index and significance level.

According to the results of the BCP test, we have some indication of possible problems regarding the independence property of the exceedances, which causes all models to perform poorly under periods of high turbulence, leading to autocorrelated exceedances and clustering.

Contrary to what was expected, there was no dominance of the asymmetric and long memory volatility models, which means that accounting for these two properties is not an obvious decision to achieve better VaR estimates. Moreover, both EWMA volatility

models, appeared as the best models for most of the indices. Out of twelve possibilities, eight times the best combination of models involved an EWMA volatility.

In Table 16 we present the number of rejections that each model received, by the UC test, for the global period in analysis and for the 4 indices.

Model	N ^o of Rejections			
	1%	2.5%	5%	Total
Parametric Normal - GARCH	4	3	1	8
Parametric Normal - EGARCH	4	4	3	11
Parametric Normal - FIGARCH	4	4	0	8
Parametric Normal - EWMA	4	4	4	12
Parametric Normal - EWMA $\lambda = 0.94$	4	4	4	12
Parametric SGSt - GARCH	1	1	2	4
Parametric SGSt - EGARCH	2	3	3	8
Parametric SGSt - FIGARCH	1	0	2	3
Parametric SGSt - EWMA	2	2	4	8
Parametric SGSt - EWMA $\lambda = 0.94$	1	1	4	6
Historical - GARCH	0	0	0	0
Historical - EGARCH	0	0	0	0
Historical - FIGARCH	0	0	0	0
Historical - EWMA	0	0	0	0
Historical - EWMA $\lambda = 0.94$	0	0	0	0
Quantile Regression - GARCH	4	3	1	8
Quantile Regression - EGARCH	4	2	3	9
Quantile Regression - FIGARCH	4	3	3	10
Quantile Regression - EWMA	4	3	1	8
Quantile Regression - EWMA $\lambda = 0.94$	4	3	0	7

Table 16. Number of rejections by model combination. Highlights the number of rejections that each model combination received on the global period of the UC test. The model combinations with the lowest number of total rejections are the ones that better estimate the VaR.

From a global perspective, five combinations of models largely outperform the rest when looking solely at the number of rejections. These were the Volatility-Adjusted Historical models, which were always accepted, for all indices and significance levels. What can be deduced from this is that, for a risk manager that pretends to use a singular implementation for various portfolios, recurring to one of these combinations should increase

the probability of the estimates being accepted by the UC test, hence being a good fit to predict the VaR.

Looking only at the VaR models, the one with the worst performance was the Parametric Normal. Each model could be rejected up to 80 times, 20 times for each significance level, and the Parametric Normal model accumulated 51 rejections. Closely following it, was the Quantile Regression, that racked up 42 rejections. The Parametric SGSt has less rejections, 29 in total. By far, the best model is the Volatility-Adjusted Historical, 0 rejections. Given this, it's obvious that there's a clear advantage of using the Volatility-Adjusted Historical model to estimate the VaR, when paired with good performing volatility models.

Table 17 provides the mean of the VaR estimates for each model, and the global average for each significance level. The values in bold indicate the models that were accepted by the UC test.

Comparing the four indices, we have two pairs with similar results, the first being the S&P500 and the FTSE, and the other the DAX and the HSI. The second pair has higher VaR averages, which implies that those indices are riskier and more volatile.

As we can see, the average that only accounts for accepted models, those that have a p-value higher than 5% on the UC test, is always higher than the global average. Knowing this, we can conclude that the models that are rejected by the UC test tend to underestimate the Value-at-Risk. We can see it clearly for the Parametric Normal models at the 1% backtest, which always have the mean of its' VaR estimates way smaller than the global average.

We can also observe that, amongst the models that have been accepted by the UC test at a 1% significance level, the ones that were chosen as the best to estimate the VaR for each index, have smaller means than the global average. This is an advantage to investors, having a smaller VaR implies that less money is needed to cover for market risk. When looking at the 2.5% and 5% averages, the best models have slightly higher means than the average of all accepted models.

Model	Mean Value-at-Risk											
	S&P500			DAX			FTSE			HSI		
	1%	2.5%	5%	1%	2.5%	5%	1%	2.5%	5%	1%	2.5%	5%
Parametric Normal - GARCH	2.47%	2.08%	1.75%	3.13%	2.64%	2.21%	2.47%	2.08%	1.75%	3.22%	2.71%	2.27%
Parametric Normal - EGARCH	2.27%	1.92%	1.61%	2.91%	2.45%	2.06%	2.31%	1.94%	1.63%	3.10%	2.61%	2.19%
Parametric Normal - FIGARCH	2.47%	2.08%	1.75%	3.22%	2.71%	2.28%	2.50%	2.11%	1.77%	3.18%	2.68%	2.25%
Parametric Normal - EWMA	2.36%	1.99%	1.67%	3.00%	2.53%	2.12%	2.35%	1.98%	1.66%	2.99%	2.52%	2.11%
Parametric Normal - EWMA $\lambda = 0.94$	2.39%	2.01%	1.69%	3.02%	2.55%	2.14%	2.37%	2.00%	1.68%	3.01%	2.53%	2.12%
Parametric SGSt - GARCH	3.02%	2.25%	1.71%	3.81%	2.89%	2.22%	2.93%	2.20%	1.69%	3.87%	2.90%	2.22%
Parametric SGSt - EGARCH	2.77%	2.07%	1.57%	3.54%	2.69%	2.07%	2.73%	2.06%	1.58%	3.72%	2.79%	2.14%
Parametric SGSt - FIGARCH	3.01%	2.25%	1.71%	3.92%	2.97%	2.28%	2.96%	2.23%	1.71%	3.82%	2.87%	2.19%
Parametric SGSt - EWMA	2.88%	2.15%	1.63%	3.65%	2.77%	2.13%	2.78%	2.09%	1.61%	3.59%	2.70%	2.06%
Parametric SGSt - EWMA $\lambda = 0.94$	2.92%	2.18%	1.65%	3.68%	2.79%	2.14%	2.81%	2.12%	1.62%	3.62%	2.72%	2.07%
Historical - GARCH	3.24%	2.50%	1.90%	3.98%	2.95%	2.38%	3.25%	2.53%	1.89%	3.77%	3.04%	2.33%
Historical - EGARCH	3.14%	2.43%	1.87%	3.72%	2.96%	2.31%	3.15%	2.45%	1.86%	3.72%	2.98%	2.31%
Historical - FIGARCH	3.22%	2.46%	1.89%	3.89%	2.98%	2.40%	3.30%	2.51%	1.91%	3.79%	3.03%	2.32%
Historical - EWMA	3.09%	2.41%	1.85%	3.43%	2.89%	2.39%	2.93%	2.41%	1.88%	3.61%	2.89%	2.27%
Historical - EWMA $\lambda = 0.94$	3.11%	2.39%	1.84%	3.45%	2.90%	2.38%	2.94%	2.41%	1.88%	3.63%	2.86%	2.26%
Quantile Regression - GARCH	3.03%	2.33%	1.82%	3.58%	2.81%	2.30%	3.02%	2.36%	1.82%	3.58%	2.85%	2.23%
Quantile Regression - EGARCH	2.99%	2.31%	1.80%	3.53%	2.81%	2.26%	2.96%	2.32%	1.79%	3.57%	2.83%	2.24%
Quantile Regression - FIGARCH	3.01%	2.30%	1.80%	3.59%	2.81%	2.30%	3.08%	2.34%	1.83%	3.60%	2.86%	2.22%
Quantile Regression - EWMA	2.76%	2.23%	1.77%	3.26%	2.81%	2.29%	2.80%	2.25%	1.80%	3.48%	2.75%	2.19%
Quantile Regression - EWMA $\lambda = 0.94$	2.77%	2.21%	1.78%	3.28%	2.82%	2.29%	2.83%	2.26%	1.81%	3.51%	2.71%	2.19%
Global average (only accepted models)	3.07%	2.36%	1.81%	3.69%	2.88%	2.34%	3.05%	2.42%	1.84%	3.71%	2.86%	2.24%
Global Average	2.85%	2.23%	1.75%	3.48%	2.79%	2.25%	2.82%	2.23%	1.76%	3.52%	2.79%	2.21%

Table 17. Mean of the Value-at-Risk estimates. We display the average value of the VaR estimate of each model combination. We also present two global averages. The first one only includes models that were accepted by the UC test for the global period and the second one include all models. The values in bold indicate the models that were accepted by the UC test for the global period in study.

CHAPTER 6

CONCLUSION

In this study, we explored if the Value-at-Risk estimation could be improved by accounting for the asymmetric behavior of the volatility and its' long memory. If these two properties are important to estimate the VaR, using models that accommodate them, would output better estimates than other simpler models.

Before delving into the estimation of the VaR, from a sample of twenty years of four market indices, S&P500, DAX, FTSE and HSI, we analyzed the stationarity of the prices and returns, concluding that the prices were not stationary but the returns were. We also examined if long memory was present in the returns and in the volatility, using the squared returns and absolute-valued returns as proxies, and discovered that, while there was no evidence of long memory on the returns of the indices, for the proxies there was an indication that they were described by a long memory process.

The estimates for the VaR were computed using five volatility models, EGARCH and FIGARCH, that accommodate asymmetry and long memory, respectively, GARCH and two EWMA models. Moreover, to further strengthen our results, we also resorted to four VaR models, Parametric Normal and Skewed Generalized Student-t, Volatility-Adjusted Historical, and Quantile Regression, which gives us a total of twenty batches of estimates for each index. Each of these batches were backtested with the UC and the BCP test.

Looking at the VaR models, there was a striking difference between them. The Parametric Normal model and Quantile Regression VaR weren't able to perform satisfactorily for the backtest of 1% significance level, never being accepted by the UC test. While not performing as bad as the previous two, the Parametric SGSt fell short in comparison with the Volatility-Adjusted Historical, which was the one with the best results, always being found to be the best performing model.

On the volatility we found no dominance of the asymmetric and long memory models. Moreover, from the models that were deemed to be the ones with the best performances, it was found that, eight times out of twelve, an EWMA volatility model performed better than the other models studied.

One important aspect of the results is that, while not dominating the other models, the EGARCH and the FIGARCH, used to compute the Volatility-Adjusted Historical VaR, were accepted by the UC test for all indices and significance levels. This implies that, globally, these two models are very well rounded and have an adequate performance, but, once again, this was not exclusive to the asymmetric and long memory models since the Volatility-Adjusted Historical GARCH, EWMA and EWMA $\lambda = 0.94$ also achieved this feat.

All in all, this study revealed that resorting to more complex volatility models, specifically ones that accommodate the asymmetric behavior of the volatility and its' long memory, is not strictly better than using a simple model like the EWMA volatility. There is no evidence to affirm that modeling these two properties allows to produce better estimates for the Value-at-Risk. While for some indices and significance levels, the EGARCH and FIGARCH have better performances than the others, that could be due to a selection bias, since when looking at global results there is no proof of dominance by neither of these.

For future research on this subject, it would be interesting to extend the analysis to other types of financial assets: bonds, options, commodities, exchange rates, etc. In addition, adding a model that accommodates both asymmetry and long memory together would be an important improvement to the study, some models we recommend are the Fractionally Integrated Asymmetric Power Auto Regressive Conditional Heteroscedasticity (FIAPARCH) and the Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroskedastic (FIEGARCH).

References

- Alexander, C. (2009). *Market risk analysis, value at risk models*, volume 4. John Wiley & Sons.
- Angelidis, T., Benos, A., and Degiannakis, S. (2004). The use of garch models in var estimation. *Statistical methodology*, 1(1-2):105–128.
- Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 74(1):3–30.
- 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.
- Berggren, E. and Folkelid, F. (2015). Which garch model is best for value-at-risk? (Bachelors thesis, Uppsala universitet).
- Berkowitz, J., Christoffersen, P., and Pelletier, D. (2011). Evaluating value-at-risk models with desk-level data. *Management Science*, 57(12):2213–2227.
- Black, F. (1976). Studies of stock market volatility changes. page 177–181.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.
- Brooks, C. (2008). *Introductory econometrics for finance*. Cambridge university press.
- Brooks, C. and Persaud, G. (2003). The effect of asymmetries on stock index return value-at-risk estimates. *The Journal of Risk Finance*.
- Bucevska, V. (2013). An empirical evaluation of garch models in value-at-risk estimation: Evidence from the macedonian stock exchange. *Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy*, 4(1):49–64.
- Campbell, J. Y. and Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of financial Economics*, 31(3):281–318.
- Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of financial Economics*, 10(4):407–432.
- Christoffersen, P., Hahn, J., and Inoue, A. (2001). Testing and comparing value-at-risk measures. *Journal of empirical finance*, 8(3):325–342.
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, pages 841–862.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

- Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106.
- Duffie, D. and Pan, J. (1997). An overview of value at risk. *Journal of derivatives*, 4(3):7–49.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007.
- Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric reviews*, 5(1):1–50.
- Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778.
- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801.
- Horta, P. J. d. B. (2015). The impact of the 2008 and 2010 financial crises on international stock markets: contagion and long memory. (Doctoral dissertation, ISCTE - Instituto Universitário de Lisboa).
- Hull, J. and White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of risk*, 1(1):5–19.
- Hurst, H. (1951). Long term storage capacities of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–779.
- Kasman, A. (2009). Estimating value-at-risk for the turkish stock index futures in the presence of long memory volatility. *Central Bank Review*, 9(1):1.
- Kayal, P. and Maheswaran, S. (2018). Leverage effect and volatility asymmetry. In *International Conference on Economics and Finance*, pages 131–150. Springer.
- Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.
- Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The J. of Derivatives*, 3(2).
- Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica: Journal of the Econometric Society*, pages 1279–1313.
- Longerstaey, J. and Spencer, M. (1996). Riskmetricstm—technical document. *Morgan Guaranty Trust Company of New York: New York*, 51:54.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.
- Sethapramote, Y., Prukumpai, S., and Kanyamee, T. (2014). Evaluation of value-at-risk estimation using long memory volatility models: Evidence from stock exchange of thailand. Available at SSRN 2396531.

- So, M. K. and Philip, L. (2006). Empirical analysis of garch models in value at risk estimation. *Journal of International Financial Markets, Institutions and Money*, 16(2):180–197.
- Tang, T.-L. and Shieh, S.-J. (2006). Long memory in stock index futures markets: A value-at-risk approach. *Physica A: Statistical Mechanics and its Applications*, 366:437–448.
- Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, 44(12-part-1):1650–1661.
- Wu, P.-T. and Shieh, S.-J. (2007). Value-at-risk analysis for long-term interest rate futures: Fat-tail and long memory in return innovations. *Journal of Empirical Finance*, 14(2):248–259.
- Yao, J., Li, Z.-F., and Ng, K. W. (2006). Model risk in var estimation: An empirical study. *International Journal of Information Technology & Decision Making*, 5(03):503–512.

Appendices

Appendix A - EWMA Volatility Estimation

As mentioned in this dissertation, the EWMA volatility is computed recurring to two methodologies. The first, and the most simple, we assume the value for parameter λ . Following the implementation of the RiskMetrics model, we assign the value of 0.94 and keep it constant for all the years in study. The second, we estimate the λ fitting it to four years of past returns, maximizing the likelihood function, and re-estimating it every trading month, ensuring a high precision. Neither of the two methodologies is strictly better than the other, the simpler implementation ignores the data and assumes a fixed parameter, but the second one, where we fit the parameter, introduces estimation errors and more complexity, which can cause deviations from the real value.

The results for the λ estimates are summarized in Table 18 below. As we can see, while the value of the estimated λ varies throughout the years, and for the four indices, the global average is pretty similar for all of them. Another important observation is that the average of the estimated λ remains close to the fixed value suggested by the RiskMetrics model, 0.94. What this indicates is that, while some benefits might appear from fitting the parameter, assuming the fixed value of 0.94 is also a viable option. This is proven by the results of our study, since for some indices the EWMA model with a fixed λ of 0.94 outperformed the one with the estimated λ .

Date	Estimated Lambda			
	S&P500	DAX	FTSE	HSI
01/01/2000	0,976	0,955	0,978	0,940
22/01/2000	0,976	0,955	0,978	0,934
12/02/2000	0,977	0,954	0,976	0,935
(...)				
01/01/2005	0,942	0,919	0,941	0,958
22/01/2005	0,934	0,923	0,931	0,954
12/02/2005	0,933	0,918	0,931	0,958
(...)				
01/01/2010	0,954	0,933	0,926	0,977
22/01/2010	0,953	0,936	0,925	0,975
12/02/2010	0,954	0,936	0,927	0,973
(...)				
01/01/2015	0,928	0,934	0,925	0,953
22/01/2015	0,941	0,933	0,922	0,959
12/02/2015	0,937	0,934	0,917	0,961
(...)				
01/01/2020	0,923	0,943	0,898	0,959
22/01/2020	0,916	0,937	0,905	0,958
12/02/2020	0,918	0,928	0,903	0,958
(...)				
Total Average	0,937	0,934	0,931	0,946

Table 18. Estimated lambdas. Lambdas estimate for the four indices throughout the years and global average value.

Appendix B - UC and BCP test results for 2.5% significance level

B.1. UC test

In this section we present the results for the UC test. We include all model combinations, accepted and rejected, and the respective p-values. The value in bold highlight accepted model combinations. The model combinations are presented according to a key, following Table 19 below.

Key	Model
1	Parametric Normal - GARCH
2	Parametric Normal - EGARCH
3	Parametric Normal - FIGARCH
4	Parametric Normal - EWMA
5	Parametric Normal - EWMA $\lambda = 0,94$
6	Parametric SGSt - GARCH
7	Parametric SGSt - EGARCH
8	Parametric SGSt - FIGARCH
9	Parametric SGSt - EWMA
10	Parametric SGSt - EWMA $\lambda = 0,94$
11	Historical - GARCH
12	Historical - EGARCH
13	Historical - FIGARCH
14	Historical - EWMA
15	Historical - EWMA $\lambda = 0,94$
16	Quantile Regression - GARCH
17	Quantile Regression - EGARCH
18	Quantile Regression - FIGARCH
19	Quantile Regression - EWMA
20	Quantile Regression - EWMA $\lambda = 0,94$

Table 19. Keys for the UC test. To improve the readability of the UC test tables, we present the models following keys. Each key corresponds to a model studied.

B.1.1. S&P500

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	0.00%	0.00%	0.00%	0.00%	0.00%	12.78%	0.00%	10.76%	2.72%	10.76%	36.20%	77.18%	51.80%	53.71%	53.71%	0.39%	1.08%	0.13%	4.15%	1.38%
01/01/2000	01/01/2004	18.24%	33.90%	97.98%	18.24%	18.24%	86.04%	66.32%	28.32%	56.83%	70.81%	81.89%	97.98%	81.89%	81.89%	81.89%	44.46%	86.04%	44.46%	56.83%	56.83%
01/01/2004	01/01/2008	18.59%	2.55%	9.06%	2.55%	3.99%	97.18%	25.62%	57.53%	18.59%	18.59%	6.09%	3.99%	3.99%	25.62%	25.62%	0.97%	2.55%	0.34%	3.99%	3.99%
01/01/2008	01/01/2012	0.00%	0.00%	0.00%	0.00%	0.00%	1.64%	0.00%	0.60%	13.41%	45.70%	50.63%	27.46%	80.34%	18.97%	27.46%	26.06%	87.64%	9.26%	87.64%	34.96%
01/01/2012	01/01/2016	25.40%	3.94%	3.94%	0.19%	0.33%	86.44%	97.58%	44.77%	18.41%	25.40%	86.44%	65.96%	86.44%	97.58%	51.58%	34.16%	8.96%	44.77%	34.16%	25.40%
01/01/2016	01/01/2020	34.16%	0.01%	25.40%	8.96%	13.01%	57.18%	0.96%	71.19%	86.44%	97.58%	44.77%	97.58%	65.96%	28.11%	51.58%	34.16%	18.41%	34.16%	44.77%	44.77%

Table 20. S&P500 - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 19).

B.1.2. DAX

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	0.00%	0.00%	0.03%	0.00%	0.00%	93.72%	0.02%	26.65%	29.14%	29.14%	86.61%	99.10%	45.90%	84.76%	91.91%	0.57%	5.48%	0.93%	13.61%	13.61%
01/01/2000	01/01/2004	23.28%	0.28%	67.41%	11.71%	5.32%	54.83%	53.72%	30.34%	30.34%	21.17%	41.62%	21.17%	21.17%	54.83%	85.42%	23.28%	69.61%	82.46%	85.42%	85.42%
01/01/2004	01/01/2008	90.84%	2.93%	75.41%	10.12%	6.86%	25.80%	61.07%	25.80%	93.18%	90.84%	93.18%	77.27%	48.15%	90.84%	93.18%	61.07%	27.88%	75.41%	37.14%	48.23%
01/01/2008	01/01/2012	0.00%	0.00%	0.04%	0.00%	0.00%	4.78%	0.00%	91.59%	10.57%	15.14%	38.26%	21.15%	21.15%	38.26%	62.51%	7.19%	10.57%	1.21%	7.19%	3.09%
01/01/2012	01/01/2016	13.28%	4.04%	13.28%	0.59%	0.99%	80.72%	45.39%	80.72%	9.16%	13.28%	50.95%	38.32%	19.13%	27.67%	50.95%	34.69%	87.24%	25.84%	71.96%	71.96%
01/01/2016	01/01/2020	5.94%	3.89%	3.89%	0.94%	1.55%	97.98%	56.83%	97.98%	97.98%	70.81%	33.90%	33.90%	86.04%	86.04%	86.04%	8.86%	8.86%	12.87%	70.81%	86.04%

Table 21. DAX - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 19).

B.1.3. FTSE

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	0.01%	0.00%	0.01%	0.00%	0.00%	4.91%	0.00%	20.27%	0.02%	0.81%	90.63%	73.60%	76.57%	80.59%	73.60%	1.03%	0.48%	0.21%	1.31%	4.00%
01/01/2000	01/01/2004	0.40%	0.40%	1.89%	0.40%	1.15%	4.68%	4.68%	7.10%	29.10%	38.78%	73.36%	32.70%	58.20%	58.20%	44.51%	21.27%	50.35%	21.27%	29.10%	63.68%
01/01/2004	01/01/2008	40.83%	0.77%	5.12%	0.26%	2.09%	66.28%	30.82%	81.26%	16.24%	30.82%	11.34%	40.83%	40.83%	30.82%	30.82%	0.77%	1.29%	0.45%	7.72%	11.34%
01/01/2008	01/01/2012	0.90%	0.00%	2.38%	0.05%	0.53%	1.48%	0.00%	24.54%	0.53%	12.47%	13.34%	67.41%	67.41%	52.87%	83.06%	84.85%	84.85%	69.67%	55.79%	33.11%
01/01/2012	01/01/2016	2.28%	8.27%	1.41%	0.16%	0.50%	42.62%	32.34%	42.62%	3.59%	5.52%	68.54%	99.60%	68.54%	83.65%	99.60%	17.22%	3.59%	8.27%	42.62%	42.62%
01/01/2016	01/01/2020	79.95%	9.37%	64.53%	2.66%	6.31%	12.38%	72.72%	18.80%	35.23%	58.58%	95.98%	79.95%	72.72%	50.32%	37.79%	88.04%	46.01%	58.58%	19.12%	46.01%

Table 22. FTSE - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 19).

B.1.4. HSI

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	6.51%	0.03%	0.52%	0.00%	0.00%	81.18%	54.76%	67.43%	29.36%	29.36%	82.49%	88.33%	61.71%	81.18%	43.44%	18.78%	7.89%	6.51%	2.27%	2.84%
01/01/2000	01/01/2004	5.49%	73.59%	58.08%	37.27%	14.08%	5.49%	32.11%	32.11%	90.00%	62.24%	58.08%	58.08%	58.08%	73.59%	77.25%	93.37%	62.24%	27.68%	20.01%	27.68%
01/01/2004	01/01/2008	2.77%	0.11%	0.62%	0.11%	0.36%	4.33%	1.05%	4.33%	20.12%	37.07%	14.25%	14.25%	20.12%	37.07%	37.07%	6.61%	4.33%	4.33%	6.61%	9.83%
01/01/2008	01/01/2012	59.65%	46.74%	26.51%	9.29%	26.51%	74.15%	62.04%	89.84%	93.85%	77.59%	77.59%	25.17%	93.85%	59.65%	35.67%	59.65%	77.59%	19.19%	19.19%	19.19%
01/01/2012	01/01/2016	13.96%	9.61%	13.96%	13.96%	19.74%	60.98%	92.63%	76.42%	60.74%	75.33%	16.61%	92.63%	6.50%	46.92%	46.92%	92.63%	60.74%	46.92%	76.42%	92.63%
01/01/2016	01/01/2020	14.25%	1.05%	20.12%	1.05%	0.36%	76.12%	91.88%	91.82%	48.37%	37.07%	91.88%	61.47%	91.82%	91.88%	76.12%	61.47%	27.67%	76.12%	37.07%	37.07%

Table 23. HSI - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 19).

B.2. BCP test

In this section we display a summary of the BCP test results for the four indices. It presents the lag that corresponds to the lowest p-value of the BCP test.

B.2.1. S&P500

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric SGSt - GARCH	2	0.00%	2	0.00%	4	0.00%	2	0.00%	5	0.03%	4	0.00%
Parametric SGSt - FIGARCH	2	0.00%	1	0.00%	4	0.00%	2	0.00%	2	0.00%	3	0.00%
Parametric SGSt - EWMA $\lambda = 0.94$	4	0.00%	5	0.18%	5	0.01%	4	0.01%	5	0.48%	4	0.00%
Historical - GARCH	2	0.00%	4	0.00%	4	0.10%	2	0.00%	2	0.00%	4	0.00%
Historical - EGARCH	4	0.00%	5	2.17%	5	0.36%	2	0.00%	2	0.00%	4	0.00%
Historical - FIGARCH	2	0.00%	3	0.00%	5	0.20%	2	0.00%	2	0.00%	4	0.00%
Historical - EWMA	4	0.00%	4	0.03%	5	0.02%	5	0.09%	5	2.48%	4	0.00%
Historical - EWMA $\lambda = 0.94$	4	0.00%	4	0.03%	5	3.38%	4	0.00%	5	0.75%	4	0.00%

Table 24. S&P500 - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

B.2.2. DAX

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric SGSt - GARCH	2	0.00%	2	0.00%	3	0.01%	5	0.00%	5	11.46%	5	0.01%
Parametric SGSt - FIGARCH	1	0.00%	1	0.00%	3	0.00%	2	0.00%	5	1.50%	5	0.06%
Parametric SGSt - EWMA	2	0.00%	5	0.22%	5	0.54%	2	0.00%	5	2.29%	5	0.06%
Parametric SGSt - EWMA $\lambda = 0.94$	2	0.00%	5	0.04%	4	0.04%	2	0.00%	5	9.09%	5	0.24%
Historical - GARCH	2	0.00%	3	0.00%	5	0.43%	5	0.00%	5	10.06%	2	0.00%
Historical - EGARCH	3	0.00%	4	0.00%	5	1.68%	5	0.20%	5	0.03%	3	0.00%
Historical - FIGARCH	2	0.00%	3	0.00%	5	0.06%	2	0.00%	5	6.06%	5	0.05%
Historical - EWMA	2	0.00%	5	0.79%	5	4.36%	4	0.00%	5	7.54%	5	0.22%
Historical - EWMA $\lambda = 0.94$	2	0.00%	5	0.06%	5	0.08%	3	0.00%	5	9.07%	5	1.03%
Quantile Regression - EGARCH	2	0.00%	5	0.02%	5	0.23%	5	0.06%	4	0.00%	2	0.00%
Quantile Regression - EWMA	2	0.00%	5	0.33%	4	4.63%	2	0.00%	5	6.17%	5	5.37%
Quantile Regression - EWMA Lambda 0.94	2	0.00%	5	0.06%	5	0.30%	2	0.00%	5	6.17%	5	1.03%

Table 25. DAX - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

B.2.3. FTSE

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric SGSt - FIGARCH	1	0.00%	2	0.00%	4	0.02%	3	0.00%	1	0.00%	1	0.00%
Historical - GARCH	1	0.00%	3	0.00%	5	0.00%	1	0.00%	3	0.00%	1	0.00%
Historical - EGARCH	1	0.00%	3	0.00%	2	0.00%	2	0.00%	3	0.00%	5	0.00%
Historical - FIGARCH	1	0.00%	2	0.00%	5	0.04%	1	0.00%	3	0.00%	3	0.00%
Historical - EWMA	4	0.00%	5	0.00%	5	5.10%	4	0.01%	3	0.00%	5	0.06%
Historical - EWMA $\lambda = 0.94$	4	0.00%	2	0.00%	5	5.10%	4	0.01%	5	0.02%	5	0.03%

Table 26. FTSE - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

B.2.4. HSI

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric Normal - GARCH	3	0.00%	5	1.18%	3	0.00%	4	0.00%	4	0.78%	5	1.23%
Parametric SGSt - GARCH	3	0.00%	5	1.18%	3	0.00%	4	0.00%	5	0.85%	5	0.16%
Parametric SGSt - EGARCH	3	0.00%	5	6.29%	3	0.00%	5	0.08%	5	11.54%	5	0.08%
Parametric SGSt - FIGARCH	3	0.00%	4	0.00%	3	0.00%	5	0.00%	5	11.12%	5	0.03%
Parametric SGSt - EWMA	3	0.00%	5	1.57%	3	0.00%	5	17.72%	5	4.53%	5	0.50%
Parametric SGSt - EWMA $\lambda = 0.94$	5	1.11%	5	7.68%	5	0.00%	5	13.84%	5	3.64%	5	9.76%
Historical - GARCH	3	0.00%	5	9.27%	3	0.00%	5	11.09%	5	0.08%	5	0.08%
Historical - EGARCH	3	0.00%	5	9.29%	3	0.00%	5	0.18%	5	11.54%	4	0.00%
Historical - FIGARCH	3	0.00%	3	0.00%	3	0.00%	1	0.00%	5	2.34%	5	0.03%
Historical - EWMA	3	0.00%	5	10.06%	3	0.00%	5	6.95%	5	0.55%	5	0.05%
Historical - EWMA $\lambda = 0.94$	5	4.21%	5	30.40%	5	0.00%	5	31.18%	5	0.55%	5	11.93%
Quantile Regression - GARCH	3	0.00%	5	0.30%	3	0.00%	2	0.00%	5	1.66%	5	0.30%
Quantile Regression - EGARCH	3	0.00%	5	1.14%	3	0.00%	1	0.00%	5	2.98%	3	0.00%
Quantile Regression - FIGARCH	3	0.00%	5	0.10%	3	0.00%	3	0.00%	5	0.55%	4	0.04%

Table 27. HSI - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

Appendix C - UC and BCP test results for 5% significance level

C.1. UC Test

In this section we present the results for the UC test. We include all model combinations, accepted and rejected, and the respective p-values. The value in bold highlight accepted model combinations. The model combinations are presented according to a key, following Table 28 below.

Key	Model
1	Parametric Normal - GARCH
2	Parametric Normal - EGARCH
3	Parametric Normal - FIGARCH
4	Parametric Normal - EWMA
5	Parametric Normal - EWMA $\lambda = 0,94$
6	Parametric SGSt - GARCH
7	Parametric SGSt - EGARCH
8	Parametric SGSt - FIGARCH
9	Parametric SGSt - EWMA
10	Parametric SGSt - EWMA $\lambda = 0,94$
11	Historical - GARCH
12	Historical - EGARCH
13	Historical - FIGARCH
14	Historical - EWMA
15	Historical - EWMA $\lambda = 0,94$
16	Quantile Regression - GARCH
17	Quantile Regression - EGARCH
18	Quantile Regression - FIGARCH
19	Quantile Regression - EWMA
20	Quantile Regression - EWMA $\lambda = 0,94$

Table 28. Keys for the UC test. To improve the readability of the UC test tables, we present the models following keys. Each key corresponds to a model studied.

C.1.1. S&P500

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	38.82%	0.02%	19.14%	0.42%	2.11%	10.50%	0.00%	2.48%	0.04%	0.42%	92.01%	54.34%	97.68%	76.79%	92.01%	17.08%	3.39%	2.48%	5.31%	10.50%
01/01/2000	01/01/2004	97.11%	80.11%	43.95%	69.31%	80.11%	80.11%	41.34%	97.11%	33.84%	21.72%	63.46%	53.29%	43.95%	28.25%	35.58%	85.59%	74.29%	59.16%	69.31%	80.11%
01/01/2004	01/01/2008	84.46%	27.96%	42.19%	27.96%	27.96%	95.96%	27.96%	27.96%	34.59%	34.59%	3.04%	1.52%	2.16%	13.55%	10.35%	1.05%	0.48%	0.21%	1.52%	1.05%
01/01/2008	01/01/2012	0.09%	0.00%	0.14%	0.33%	0.50%	0.06%	0.00%	0.06%	0.14%	1.10%	34.13%	15.88%	51.43%	93.68%	93.68%	94.81%	72.13%	94.81%	94.81%	61.20%
01/01/2012	01/01/2016	96.54%	91.95%	50.27%	7.66%	13.36%	69.85%	27.63%	27.63%	4.15%	7.66%	85.02%	96.54%	52.82%	85.02%	85.02%	34.22%	22.00%	17.27%	59.67%	59.67%
01/01/2016	01/01/2020	16.51%	22.00%	21.71%	91.95%	52.82%	52.82%	10.19%	43.54%	50.27%	85.02%	35.22%	41.77%	73.75%	27.93%	35.22%	73.75%	22.00%	96.54%	34.22%	73.75%

Table 29. S&P500 - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 28).

C.1.2. DAX

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	4.42%	0.00%	16.56%	0.00%	0.00%	4.42%	0.02%	28.42%	0.00%	0.00%	71.13%	41.38%	49.03%	59.21%	91.01%	3.80%	0.59%	2.02%	10.15%	16.56%
01/01/2000	01/01/2004	6.48%	2.45%	37.64%	1.73%	2.45%	8.70%	11.52%	45.68%	3.44%	2.45%	90.72%	86.20%	57.59%	47.84%	68.12%	86.20%	54.67%	75.07%	90.72%	75.07%
01/01/2004	01/01/2008	67.90%	15.62%	47.85%	38.50%	46.56%	47.85%	31.39%	39.10%	46.56%	46.56%	46.56%	86.90%	46.56%	98.28%	65.35%	31.39%	15.62%	25.23%	31.39%	38.50%
01/01/2008	01/01/2012	0.30%	0.00%	1.42%	0.30%	0.98%	0.05%	0.00%	0.67%	0.13%	0.30%	48.37%	26.48%	21.06%	78.07%	57.51%	3.96%	5.42%	3.96%	12.76%	16.51%
01/01/2012	01/01/2016	93.10%	81.80%	72.67%	13.75%	13.75%	93.10%	95.38%	61.94%	13.75%	13.75%	51.89%	93.10%	61.94%	72.67%	83.90%	81.80%	42.62%	93.10%	70.93%	95.38%
01/01/2016	01/01/2020	85.59%	80.11%	49.80%	1.45%	2.08%	74.29%	85.59%	91.38%	5.58%	10.03%	91.38%	80.11%	59.16%	63.46%	85.59%	27.31%	17.03%	13.17%	41.34%	59.16%

Table 30. DAX - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 28).

C.1.3. FTSE

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	11.54%	0.00%	11.54%	0.00%	0.01%	0.58%	0.00%	0.22%	0.00%	0.00%	97.93%	76.57%	76.57%	91.73%	77.44%	10.16%	1.20%	3.79%	3.24%	5.89%
01/01/2000	01/01/2004	12.99%	1.40%	33.75%	1.40%	3.96%	21.56%	2.84%	41.31%	2.00%	7.39%	43.06%	84.87%	62.59%	59.30%	84.87%	91.87%	69.57%	49.85%	41.31%	69.57%
01/01/2004	01/01/2008	92.44%	6.16%	84.48%	14.39%	44.35%	62.93%	2.31%	44.35%	3.25%	8.29%	14.39%	29.53%	6.16%	62.93%	53.20%	11.00%	0.51%	3.25%	23.58%	29.53%
01/01/2008	01/01/2012	0.28%	0.00%	1.36%	0.43%	0.94%	0.00%	0.00%	0.00%	0.00%	0.00%	87.29%	87.29%	75.93%	89.65%	89.65%	20.90%	48.41%	26.34%	7.17%	12.60%
01/01/2012	01/01/2016	66.10%	31.64%	47.04%	0.59%	0.59%	66.10%	38.85%	25.40%	0.26%	0.40%	47.04%	76.75%	47.04%	99.42%	89.00%	9.13%	15.67%	15.67%	38.85%	25.40%
01/01/2016	01/01/2020	11.63%	94.24%	33.77%	72.02%	61.71%	11.63%	94.24%	33.77%	72.02%	61.71%	15.68%	71.60%	20.68%	33.77%	26.70%	26.70%	82.93%	41.89%	94.25%	94.25%

Table 31. FTSE - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 28).

C.1.4. HSI

Backtest Period		Models																			
Beginning	End	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Global	47.39%	50.46%	85.35%	0.25%	0.46%	78.27%	11.66%	38.88%	0.03%	0.03%	63.70%	85.35%	93.71%	83.35%	78.27%	13.22%	18.91%	5.10%	14.95%	11.66%
01/01/2000	01/01/2004	1.62%	5.67%	20.73%	32.04%	32.04%	8.14%	62.90%	97.62%	2.45%	4.82%	52.45%	34.40%	90.52%	97.62%	78.93%	67.86%	67.86%	39.52%	25.58%	15.58%
01/01/2004	01/01/2008	47.06%	19.93%	19.93%	3.54%	8.97%	25.25%	19.93%	19.93%	4.89%	4.89%	2.52%	6.67%	25.25%	77.08%	77.08%	1.77%	2.52%	4.89%	66.32%	88.38%
01/01/2008	01/01/2012	85.48%	63.64%	74.27%	6.12%	11.00%	44.79%	11.00%	23.72%	1.59%	1.59%	24.22%	10.09%	91.20%	97.07%	63.64%	68.34%	31.01%	63.64%	23.72%	14.42%
01/01/2012	01/01/2016	46.49%	87.22%	37.72%	30.81%	19.40%	66.75%	98.83%	77.91%	37.98%	30.81%	77.91%	77.91%	23.37%	46.49%	46.49%	98.83%	75.96%	77.91%	77.91%	66.75%
01/01/2016	01/01/2020	66.32%	19.93%	100.00%	47.06%	47.06%	47.06%	31.53%	76.79%	66.32%	47.06%	47.06%	66.32%	88.38%	38.80%	77.08%	38.80%	38.80%	19.93%	47.06%	38.80%

Table 32. HSI - UC test results. Displays the p-values for all model combinations, accepted and rejected. The values in bold highlight model combination that are accepted by the UC test. Each key, 1 through 20, corresponds to a model (see Table 28).

C.2. BCP test

In this section we display a summary of the BCP test results for the four indices. It presents the lag that corresponds to the lowest p-value of the BCP test.

C.2.1. S&P500

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric Normal - GARCH	2	0.00%	4	0.01%	2	0.00%	4	0.00%	4	0.01%	4	0.00%
Parametric Normal - FIGARCH	2	0.00%	3	0.00%	3	0.00%	4	0.00%	5	0.00%	1	0.00%
Parametric SGSt - GARCH	2	0.00%	5	0.01%	2	0.00%	4	0.00%	4	0.00%	3	0.00%
Historical - GARCH	2	0.00%	3	0.00%	5	0.83%	4	0.00%	5	0.08%	1	0.00%
Historical - EGARCH	1	0.00%	4	0.00%	5	0.01%	2	0.00%	5	0.03%	1	0.00%
Historical - FIGARCH	2	0.00%	3	0.00%	5	0.97%	5	0.00%	2	0.00%	2	0.00%
Historical - EWMA	4	0.00%	5	11.03%	5	3.52%	5	0.14%	5	1.24%	5	0.00%
Historical - EWMA $\lambda = 0.94$	4	0.00%	5	12.40%	5	2.46%	5	0.14%	5	0.25%	4	0.00%
Quantile Regression - GARCH	1	0.00%	3	0.00%	5	0.01%	4	0.00%	5	0.01%	5	0.00%
Quantile Regression - EWMA	4	0.00%	5	11.28%	5	2.59%	5	0.46%	5	0.32%	4	0.00%
Quantile Regression - EWMA Lambda 0,94	4	0.00%	5	12.46%	5	1.09%	5	0.04%	5	0.38%	4	0.00%

Table 33. S&P500 - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

C.2.2. DAX

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric Normal - FIGARCH	1	0.00%	4	0.00%	5	0.15%	3	0.00%	5	0.03%	5	0.01%
Parametric SGSt - FIGARCH	1	0.00%	4	0.00%	5	3.51%	3	0.00%	5	0.01%	2	0.00%
Historical - GARCH	1	0.00%	4	0.00%	3	0.00%	5	0.01%	3	0.00%	5	0.00%
Historical - EGARCH	1	0.00%	3	0.00%	5	0.01%	2	0.00%	5	0.01%	5	0.02%
Historical - FIGARCH	5	0.00%	5	0.01%	3	0.00%	5	0.02%	5	0.05%	3	0.00%
Historical - EWMA	1	0.00%	3	0.00%	5	0.01%	4	0.00%	5	0.00%	5	0.46%
Historical - EWMA $\lambda = 0.94$	2	0.00%	5	0.68%	5	0.39%	4	0.00%	5	0.33%	5	4.88%
Quantile Regression - EWMA	1	0.00%	5	0.00%	5	0.53%	2	0.00%	5	0.06%	5	2.18%
Quantile Regression - EWMA Lambda 0,94	2	0.00%	4	0.00%	6	0.08%	2	0.00%	5	0.12%	5	0.39%

Table 34. DAX - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

C.2.3. FTSE

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric Normal - GARCH	1	0.00%	2	0.00%	4	0.00%	2	0.00%	3	0.00%	5	0.02%
Parametric Normal - FIGARCH	1	0.00%	1	0.00%	5	0.00%	1	0.00%	3	0.00%	3	0.00%
Historical - GARCH	1	0.00%	3	0.00%	5	0.15%	4	0.00%	3	0.00%	5	0.01%
Historical - EGARCH	1	0.00%	1	0.00%	5	0.90%	4	0.00%	5	0.01%	1	0.00%
Historical - FIGARCH	1	0.00%	4	0.00%	5	0.02%	1	0.00%	3	0.00%	1	0.00%
Historical - EWMA	5	0.20%	5	0.41%	5	13.96%	5	1.36%	5	2.56%	5	0.38%
Historical - EWMA $\lambda = 0.94$	5	0.00%	5	0.10%	5	37.92%	5	0.02%	5	1.88%	3	0.00%
Quantile Regression - GARCH	1	0.00%	2	0.00%	5	0.01%	5	0.01%	3	0.00%	1	0.00%
Quantile Regression - EWMA Lambda 0,94	1	0.00%	5	0.02%	5	20.77%	5	0.02%	5	0.04%	1	0.00%

Table 35. FTSE - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

C.2.4. HSI

Model	Backtest Period											
	Global		2000-2004		2004 - 2008		2008 - 2012		2012 - 2016		2016-2020	
	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP	Lag	BCP
Parametric Normal - GARCH	3	0.00%	5	0.77%	3	0.00%	1	0.00%	5	19.92%	5	9.56%
Parametric Normal - EGARCH	3	0.00%	5	3.98%	3	0.00%	2	0.00%	5	2.82%	5	2.31%
Parametric Normal - FIGARCH	3	0.00%	5	13.56%	3	0.00%	1	0.00%	5	7.04%	5	3.83%
Parametric SGSt - GARCH	3	0.00%	5	2.08%	3	0.00%	3	0.00%	5	11.75%	5	3.94%
Parametric SGSt - EGARCH	3	0.00%	5	0.08%	3	0.00%	1	0.00%	5	2.34%	5	1.24%
Parametric SGSt - FIGARCH	3	0.00%	5	0.61%	3	0.00%	2	0.00%	5	12.45%	5	2.61%
Historical - GARCH	3	0.00%	5	0.79%	4	0.00%	4	0.00%	5	6.74%	5	0.80%
Historical - EGARCH	3	0.00%	5	0.16%	3	0.00%	2	0.00%	5	3.30%	5	0.22%
Historical - FIGARCH	3	0.00%	5	0.33%	3	0.00%	4	0.00%	5	6.68%	5	4.30%
Historical - EWMA	5	0.02%	5	0.01%	3	0.00%	5	21.75%	5	4.90%	5	0.06%
Historical - EWMA $\lambda = 0.94$	5	0.01%	5	0.04%	3	0.00%	5	22.20%	5	4.90%	5	0.01%
Quantile Regression - GARCH	3	0.00%	5	0.46%	3	0.00%	5	0.00%	5	6.79%	5	5.67%
Quantile Regression - EGARCH	3	0.00%	5	1.17%	3	0.00%	5	0.00%	5	2.18%	5	5.58%
Quantile Regression - FIGARCH	3	0.00%	5	0.31%	4	0.00%	4	0.00%	5	10.74%	5	5.41%
Quantile Regression - EWMA	3	0.00%	5	0.47%	3	0.00%	5	0.35%	5	8.04%	5	0.03%
Quantile Regression - EWMA Lambda 0,94	3	0.00%	5	0.50%	3	0.00%	5	0.39%	5	2.78%	5	0.03%

Table 36. S&P500 - BCP test results. For each sub-period we present the lag that corresponds to the lowest p-value of the BCP test, and the respective p-value. Values in bold highlight model combinations that passed the BCP test for the specific sub-period.

Appendix D - Source Code

D.1. Required Packages

```
##### REQUIRED PACKAGES #####
#install.packages('rugarch')
#install.packages('tidyquant')
#install.packages('openxlsx')
#install.packages('sgt')
#install.packages('fGarch')
#install.packages('quantreg')
#install.packages('qrcm')
#install.packages('xlsx')
#install.packages('moments')
#install.packages('urca')
setwd("D:/Faculdade/Thesis/Code")
library(rugarch)
library(quantmod)
library(openxlsx)
library(sgt)
library(quantreg)
library(xlsx)
library(moments)
library(urca)

##### FUNCTIONS #####
```

D.2. Function - Volatility estimation

```
#Function to estimate the volatility estimating the parameters
#Arguments: returns (data.frame), vol_model ("sGarch", "eGarch", "fiGarch" or "ewma"),
reestimation_int (int)
vol_estimation <- function(returns, vol_model, reestimation_int){

one_year <- 252
len <- length(returns)
scope <- c(1:(len-one_year*4))
std <- c(1:(len-one_year*4))

if (vol_model == "ewma"){
vol_model <- "sGARCH"
lambda_ewma <-< matrix(data = -1, nrow = 1, ncol = 1)
spec = ugarchspec(variance.model = list(model = vol_model,
garchOrder = c(1, 1)), distribution.model = "sstd",
fixed.pars = list(omega=0))

counter <- 0
for (x in scope){
if(counter %% reestimation_int == 0){
control <- 1
first <- counter+1
last <- counter+one_year*4-1
fit = ugarchfit(data = returns[first:last], spec = spec, solver = "hybrid")
lambda_ewma <-< rbind(lambda_ewma, c(coef(fit)["beta1"]))
control <- control + 1
}
counter <- counter+1
}
```

```

}
vol_model <- "ewma"
}else{
spec = ugarchspec(variance.model = list(model = vol_model, garchOrder = c(1, 1)),
distribution.model = "sstd")

counter <- 0
for (x in scope){
if(counter %% reestimation_int == 0){
control <- 1
first <- counter+1
last <- counter+one_year*4-1
fit = ugarchfit(data = returns[first:last], spec = spec, solver = "hybrid")
forc1 = ugarchforecast(fit, n.ahead=reestimation_int)
std[x] <- sigma(forc1)[control]
control <- control + 1
}else{
std[x] <- sigma(forc1)[control]
control <- control + 1
}
counter <- counter+1
}
}

if (vol_model == "ewma"){
lambda_ewma <- lambda_ewma[-1,]
lambda_ewma <- data.frame(lambda_ewma)
std <- ewma(returns, lambda_ewma, reestimation_int)
}

df <- data.frame(std)
rownames(df) <- index(returns[(length(returns)-length(std)+1):length(returns)])
#std_df <- subset(x = df, subset = index(df) > which(rownames(df) >= "1996-01-01"))
std_df <- subset(x = df, subset = rownames(df) >= "1996-01-01")
return(std_df)
}

#Internal function to estimate the ewma for a given lambda
#Arguments: returns (data.frame), lambda (int)
ewma <- function(returns, lambda_ewma, reestimation_int){

one_year <- 252
returns <- returns[(one_year*4):length(returns)]

variance <- c(1:(length(returns)-1))
std <- c(1:(length(returns)-1))
counter <- 1
control <- 1
while (counter <= (length(returns)-1)){
if ((counter-1) %% reestimation_int == 0){
lambda <- lambda_ewma[control,]
control <- control+1
}
if (counter == 1){
variance[counter] = returns[counter]^2
}else{
variance[counter] = (1-lambda)*returns[counter]^2+lambda*variance[counter-1]
}
std[counter] = sqrt(variance[counter])
counter <- counter+1
}
return (std)
}

```

D.3. Function - Parametric Normal VaR

```

#Function to compute the Parametric Normal VaR
#Arguments: standard deviation (data.frame), alpha (int)
normal_var <- function(std, alpha){
var <- qnorm(1-alpha)*std

```

```

var_df <- subset(x = var, subset = rownames(var) >= "2000-01-01" &
  rownames(var) < "2020-01-01")
colnames(var_df) <- "var"
return(var_df)
}

```

D.4. Function - Parametric SGSt VaR Fit

```

#Function to estimate the parameters of the Parametric Generalized Skewed Student t VaR
#Arguments: returns (data.frame), reestimation_int (int)
sgt_fit <- function(rt, reestimation_int){

  one_year <- 252
  len <- length(rt)
  scope <- c(1:(len-one_year*4))

  lambda_vec <- c(1:(len-one_year*4))
  p_vec <- c(1:(len-one_year*4))
  q_vec <- c(1:(len-one_year*4))
  par = list(mu = 0, sigma = 2, lambda = 0, p = 2, q = 12)

  counter <- 0
  for (x in scope){
    if (counter %% reestimation_int == 0){
      first <- counter + 1
      last <- counter+one_year*4-1
      rt_to_estimate <- rt[first:last]
      X.f <- X ~ rt_to_estimate
      if (counter == 0){
        par_new <- sgt.mle(X.f, start = par)
        par_old <- par_new
      }else{
        par_old <- par_new
        par_new <- sgt.mle(X.f, start = par)
      }
      if (par_new$estimate[4]*par_new$estimate[5] <= 2){
        lambda_vec[x] <- par_old$estimate[3]
        p_vec[x] <- par_old$estimate[4]
        q_vec[x] <- par_old$estimate[5]
        counter <- counter + 1
      }else{
        lambda_vec[x] <- par_new$estimate[3]
        p_vec[x] <- par_new$estimate[4]
        q_vec[x] <- par_new$estimate[5]
        counter <- counter + 1
      }
    }
  }
  parameters <- data.frame(lambda_vec, p_vec, q_vec)

  rownames(parameters) <- index(rt[(length(rt)-length(parameters[,1])+1):length(rt)])
  #parameters_df <- subset(x = parameters, subset = index(parameters)
  > which(rownames(parameters) >= "1996-01-01"))
  parameters_df <- subset(x = parameters, subset = rownames(parameters) >= "1996-01-01")
  return(parameters_df)
}

```

D.5. Function - Parametric SGSt VaR

```

#Function to compute the Parametric Generalized Skewed Student t VaR
#Arguments: standard deviation (data.frame), parameters (data.frame), alpha (int)
sgt_var <- function(std, par, alpha){

  var <- std*0
  colnames(var) <- "var"
  control <- c(1:length(std[,1]))

  for (x in control){

```



```

var[x,1] <- -qsgt(alpha, mu = 0, sigma = std[x,1], lambda = par[x,1],
  p = par[x,2], q = par[x,3])
}

var_df <- data.frame(var)
var_df <- subset(x = var_df, subset = rownames(var_df) >= "2000-01-01"
  & rownames(var_df) < "2020-01-01")
return (var_df)
}

```

D.6. Function - Historical VaR

```

#Function to compute the Historical VaR
#Arguments: returns (data.frame), standard deviation (data.frame),
  interval (int), alpha (int)
historical_var <- function(rt, std, interval, alpha){

  rt_df <- data.frame(rt)
  std_df <- data.frame(std)

  range_rt <- min(which(rownames(rt_df) >= "2000-01-01")) - interval
  range_std <- min(which(rownames(std_df) >= "2000-01-01")) - interval

  rt_df <- subset(x = rt_df, subset = index(rt_df) >= range_rt)
  std_df <- subset(x = std_df, subset = index(std_df) >= range_std)

  rt_adj <- rt_df/std_df

  var_len <- nrow(rt_df) - min(which(rownames(rt_df) >= "2000-01-01")) + 1
  var <- c(1:(var_len-1))

  std_df <- subset(x = std_df, subset = rownames(std_df) >= "2000-01-01")

  for (x in var){
    var[x] <- -quantile(rt_adj[x:(interval+x-1),1], alpha, type = 3)*std_df[x+1,1]
  }

  var <- data.frame(var)
  rt_df_names <- subset(x = rt_df, subset = rownames(rt_df) >= "2000-01-01"
    & rownames(rt_df) < "2020-01-01")
  rownames(var) <- rownames(rt_df_names)
  return(var)
}

```

D.7. Function - Quantile Regression VaR Fit

```

#Function to estimate the parameters of the Quantile Regression VaR
#Arguments: returns (data.frame), standard deviation (data.frame),
  interval (int), reestimation interval (int), alpha (int)
quantile_fit <- function (rt, std, interval, reestimation_int, q_quantile){

  rt_df <- data.frame(rt)
  std_df <- data.frame(std)

  range_rt <- min(which(rownames(rt_df) >= "2000-01-03")) - interval
  range_std <- min(which(rownames(std_df) >= "2000-01-03")) - interval

  rt_df <- subset(x = rt_df, subset = index(rt_df) >= range_rt)
  std_df <- subset(x = std_df, subset = index(std_df) >= range_std)

  rt_df_names <- subset(x = rt_df, subset = rownames(rt_df) >= "2000-01-01")

  a <- c(1:nrow(rt_df_names))

  counter <- 0
  for (x in a){
    if (counter %% reestimation_int == 0){
      temp <- rq(rt_df[x:(interval+counter),1] ~ 0 + std_df[x:(interval+counter),1],
        tau = q_quantile)
    }
  }
}

```

```

a[x] <- temp$coefficients[1]
}else{
a[x] <- temp$coefficients[1]
}
counter <- counter + 1
}
pars <- data.frame(a)
rownames(pars) <- rownames(rt_df_names)
colnames(pars) <- "pars"
return (pars)
}

```

D.8. Function - Quantile Regression VaR

```

#Function to compute the Quantile Regression VaR
#Arguments: standard deviation (data.frame), returns (data.frame),
quantile_Var <- function (std, par){

std_df <- data.frame(std)
std_df <- subset(x = std_df, subset = rownames(std_df) >= "2000-01-01")
var <- -(par[1]*std_df)
var <- subset(x = var, subset = rownames(var) < "2020-01-01")
colnames(var) <- "var"
return (var)
}

```

D.9. Function - Backtest

```

#Function for the UC and BCP test for a single time interval
#Arguments: returns (data.frame), var (data.frame), starting date (int),
ending date (int), test ("uc" or "bcp")
backtest <- function(rt, var, d1, d2, test){

rt_df <- data.frame(rt)
rt_df <- subset(x = rt_df, subset = rownames(rt_df) < "2020-01-01")
var_df <- data.frame(var)

var_df <- subset(x = var_df, subset = rownames(var_df) >= d1)
var_df <- subset(x = var_df, subset = rownames(var_df) <= d2)
rt_df <- subset(x = rt_df, subset = rownames(rt_df) >= d1)
rt_df <- subset(x = rt_df, subset = rownames(rt_df) <= d2)

# Exceedance
exceedance <- c(1:(length(rt_df[,1])))
i <- 1
while (i <= length(rt_df[,1])) {
if (rt_df[,1][i] < (var_df[i,1]*-1)){
exceedance[i] <- 1
}else{
exceedance[i] <- 0
}
i <- i+1
}

# n
n <- nrow(rt_df)

### UC TEST
if (test == "uc"){

# n1
seq <- c(1:n)
n1 <- 0
for (i in seq){
n1 <- n1+exceedance[i]
}

#n0
n0 <- n - n1

```

```

#PI_Obs
pi_obs <- n1/n

#pi_exp
pi_exp <- 0.01

#test-stat
if (n1 > 0){
test_stat = -2*(n1*(log(pi_exp)-log(pi_obs))+n0*(log(1-pi_exp)-log(1-pi_obs)))
}else{
test_stat = -2*n*log(1-pi_exp)
}

p_value <- pchisq(test_stat, df=1, lower.tail=FALSE)

ret <- data.frame(n, n1, n0, pi_obs, test_stat, p_value)

return (ret)

### BCP test
}else{

sig_lvl <- 0.01
exc_sig_lvl <- exceedance-sig_lvl

#Autocorrelation matrix
autocorrelation <- data.frame(matrix(0, nrow = 1, ncol = 5));
colnames(autocorrelation) <- c("lag1", "lag2", "lag3", "lag4", "lag5")
lags <- c(1:5)
for (i in lags){
if (any((exc_sig_lvl[1:(length(exc_sig_lvl)-i)] ==
exc_sig_lvl[(1+i):length(exc_sig_lvl)])) == FALSE){
autocorrelation[i] <- cor(exc_sig_lvl[1:(length(exc_sig_lvl)-i)],
exc_sig_lvl[(1+i):length(exc_sig_lvl)])
if (is.na(autocorrelation[i])){
autocorrelation[i] <- 1
}
}else{
autocorrelation[i] <- 1
}
}

ret <- data.frame(matrix(0, nrow = 2, ncol = 5)); colnames(ret)
<- c("lag1", "lag2", "lag3", "lag4", "lag5"); rownames(ret) <- c("test-stat", "p-value")

#test-stat
for (i in c(1:ncol(ret))){
if (i == 1){
ret[1,i] <- n*(n+2)*autocorrelation[i]^2/(n-i)
}else{
ret[1,i] <- ret[1,i-1] + n*(n+2)*autocorrelation[i]^2/(n-i)
}
}

#p-value
for (i in c(1:ncol(ret))){
ret[2,i] <- pchisq(ret[1,i], df=1, lower.tail=FALSE)
}
}
return(ret)
}

```

D.10. Function - UC test Table

```

#Function that creates a table with the p-values for the global period
and two year intervals,
outputs PI_Obs if Exceedances == TRUE
#Arguments: returns (data.frame), vars (data.frame), exceedances (logical constant)
uctest_table <- function(rt, vars, exceedances){

```

```

if (missing(exceedances)){
  exceedances = FALSE
}

dates <- c("2000-01-03", "2004-01-03", "2008-01-03", "2012-01-03",
  "2016-01-03", "2020-01-03")
m <- data.frame(matrix(0, nrow = (length(dates)), ncol = ncol(vars)))
rownames(m) <- c("Global", "2000-01-03~2004-01-03", "2004-01-03~2008-01-03",
  "2008-01-03~2012-01-03", "2012-01-03~2016-01-03",
  "2016-01-03~2020-01-03")
colnames(m) <- colnames(vars)
first <- 0

for (i in c(1:nrow(m))){

  if (first == 0){
    d1 <- "2000-01-01"
    d2 <- "2020-01-05"
    first <- 1
  }else{
    d1 <- dates[i-1]
    d2 <- dates[i]
  }

  # Table with the decisions UC TEST
  for (x in c(1:ncol(m))){
    temp <- backtest(rt, vars[x], d1, d2, "uc")
    if (exceedances){
      m[i,x] <- temp[4]*100
    }else{
      m[i,x] <- temp[6]
    }
  }
}
return(m)
}

```

D.11. Function - BCP test Table

```

#Function that creates a table with the p-values for the global period
and two year intervals
#Arguments: returns (data.frame), vars (data.frame)
bcptest_table <- function(rt, vars){

  dates <- c("2000-01-03", "2004-01-03", "2008-01-03", "2012-01-03", "2016-01-03",
    "2020-01-03")
  m <- data.frame(matrix(0, nrow = (length(dates)), ncol = 5))
  rownames(m) <- c("Global", "2000-01-03~2004-01-03", "2004-01-03~2008-01-03",
    "2008-01-03~2012-01-03", "2012-01-03~2016-01-03",
    "2016-01-03~2020-01-03")
  colnames(m) <- c("lag1", "lag2", "lag3", "lag4", "lag5")
  ret <- vector("list", 16)

  for (x in c(1:ncol(vars))){
    first <- 0

    for (i in c(1:nrow(m))){
      if (first == 0){
        d1 <- "2000-01-01"
        d2 <- "2020-01-05"
        first <- 1
      }else{
        d1 <- dates[i-1]
        d2 <- dates[i]
      }

      temp <- backtest(rt, vars[x], d1, d2, "bcp")

      for (z in c(1:ncol(m))){

```

```

m[i,z] <- temp[2,z]
}
}
ret[[x]] <- m
}
return(ret)
}

```

D.12. Function - Lo R/S statistic test

```

#Function to compute the Lo R/s statistic test
#Arguments: x (data.frame), q (vector), alpha (optional)
#Significance level: 0.05, 0.1
#Critical value: 1.747, 1.62
#Made by Christoph Helwig - https://stat.ethz.ch/pipermail/r-help/2001-June/013425.html
#References: Lo (1991), Long-term Memory in Stock Market Prices, Econometrica 59, 1279--1313
rs_test <- function(x, q, alpha){
xbar <- mean(x)
N <- length(x)
r <- max(cumsum(x-xbar)) - min(cumsum(x-xbar))
covariance <- NULL
for (i in 1:q){
covariance <- c(covariance, sum((x[1:(N-i)]-xbar)*(x[(1+i):N]-xbar)))
}
if (q > 0)
s <- sum((x-xbar)^2)/N + sum((1-(1:q)/(q+1))*covariance)*2/N
else
s <- sum((x-xbar)^2)/N
rs <- r/(sqrt(s)*sqrt(N))
method <- "R/S_□Test_□for_□Long_□Memory"
names(rs) <- "R/S_□Statistic"
names(q) <- "Bandwidth_□q"
structure(list(statistic = rs, parameter = q, method = method,
data.name=deparse(substitute(x))), class="htest")
}

```

D.13. Runnable Code

```

##### RUN #####

counter <- 1
index_vec = c('GSPC', 'DAX', 'FTSE', 'HSI')
start.time <- Sys.time()

while (counter <= 4){

start <- '1989-12-31'
end <- '2020-1-3'
reestimation_int <- 21

if (index_vec[counter] == "GSPC"){
# S&P500
getSymbols("^GSPC", from = start, to = end)
GSPC <- GSPC$GSPC.Adjusted
index_rt <- na.omit(diff(log(GSPC)))
}else if (index_vec[counter] == "DAX"){
# DAX
getSymbols("^GDAXI", from = start, to = end)
GDAXI <- GDAXI$GDAXI.Adjusted
index_rt <- na.omit(diff(log(GDAXI)))
}else if (index_vec[counter] == "FTSE"){
# FTSE
getSymbols("^FTSE", from = start, to = end)
FTSE <- FTSE$FTSE.Adjusted
index_rt <- na.omit(diff(log(FTSE)))
}else{
# HSI
getSymbols("^HSI", from = start, to = end)
HSI <- HSI$HSI.Adjusted
}
}

```

```

index_rt <- na.omit(diff(log(HSI)))
}

##### ESTIMATES #####

### Volatility
# Volatility Estimation - sGARCH, eGARCH, fiGARCH and EWMA
std_sGARCH <- vol_estimation(index_rt, "sGARCH", reestimation_int)
std_eGARCH <- vol_estimation(index_rt, "eGARCH", reestimation_int)
std_fiGARCH <- vol_estimation(index_rt, "fiGARCH", reestimation_int)
std_ewma <- vol_estimation(index_rt, "ewma", reestimation_int)

### Value-at-Risk
# Parametric Normal VaR
var_parNorm_sGARCH <- normal_var(std_sGARCH, 0.01) ;
  colnames(var_parNorm_sGARCH) <- "var_parNorm_sGARCH"
var_parNorm_eGARCH <- normal_var(std_eGARCH, 0.01) ;
  colnames(var_parNorm_eGARCH) <- "var_parNorm_eGARCH"
var_parNorm_fiGARCH <- normal_var(std_fiGARCH, 0.01) ;
  colnames(var_parNorm_fiGARCH) <- "var_parNorm_fiGARCH"
var_parNorm_ewma <- normal_var(std_ewma, 0.01) ;
  colnames(var_parNorm_ewma) <- "var_parNorm_ewma"

# Parametric SGT VaR
sgt_parameters <- sgt_fit(index_rt, reestimation_int)
var_parSGT_sGARCH <- sgt_var(std_sGARCH, sgt_parameters, 0.01) ;
  colnames(var_parSGT_sGARCH) <- "var_parSGT_sGARCH"
var_parSGT_eGARCH <- sgt_var(std_eGARCH, sgt_parameters, 0.01) ;
  colnames(var_parSGT_eGARCH) <- "var_parSGT_eGARCH"
var_parSGT_fiGARCH <- sgt_var(std_fiGARCH, sgt_parameters, 0.01) ;
  colnames(var_parSGT_fiGARCH) <- "var_parSGT_fiGARCH"
var_parSGT_ewma <- sgt_var(std_ewma, sgt_parameters, 0.01) ;
  colnames(var_parSGT_ewma) <- "var_parSGT_ewma"

# Historical VaR
var_hist_sGARCH <- historical_var(index_rt, std_sGARCH, 500, 0.01) ;
  colnames(var_hist_sGARCH) <- "var_hist_sGARCH"
var_hist_eGARCH <- historical_var(index_rt, std_eGARCH, 500, 0.01) ;
  colnames(var_hist_eGARCH) <- "var_hist_eGARCH"
var_hist_fiGARCH <- historical_var(index_rt, std_fiGARCH, 500, 0.01) ;
  colnames(var_hist_fiGARCH) <- "var_hist_fiGARCH"
var_hist_ewma <- historical_var(index_rt, std_ewma, 500, 0.01) ;
  colnames(var_hist_ewma) <- "var_hist_ewma"

# Quantile Regression VaR
quant_parameters <- quantile_fit(index_rt, std_sGARCH, 500, reestimation_int, 0.01)
var_quant_sGARCH <- quantile_Var(std_sGARCH, quant_parameters) ;
  colnames(var_quant_sGARCH) <- "var_quant_sGARCH"
quant_parameters <- quantile_fit(index_rt, std_eGARCH, 500, reestimation_int, 0.01)
var_quant_eGARCH <- quantile_Var(std_eGARCH, quant_parameters) ;
  colnames(var_quant_eGARCH) <- "var_quant_eGARCH"
quant_parameters <- quantile_fit(index_rt, std_fiGARCH, 500, reestimation_int, 0.01)
var_quant_fiGARCH <- quantile_Var(std_fiGARCH, quant_parameters) ;
  colnames(var_quant_fiGARCH) <- "var_quant_fiGARCH"
quant_parameters <- quantile_fit(index_rt, std_ewma, 500, reestimation_int, 0.01)
var_quant_ewma <- quantile_Var(std_ewma, quant_parameters) ;
  colnames(var_quant_ewma) <- "var_quant_ewma"

save.image(file = paste(index_vec[counter], ".RData", sep = ""))
counter <- counter+1
}
end.time <- Sys.time()
time.taken <- end.time - start.time
time.taken

### Backtesting
# backtest_uc <- backtest(index_rt, var_parNorm_sGARCH, "2000-01-03", "2020-01-03", "uc")
# backtest_bcp <- backtest(index_rt, var_parNorm_sGARCH, "2004-01-03", "2006-01-03", "bcp")

```

```

vars <- data.frame(var_parNorm_sGARCH, var_parNorm_eGARCH, var_parNorm_fiGARCH,
  var_parNorm_ewma, var_parSGT_sGARCH, var_parSGT_eGARCH,
  var_parSGT_fiGARCH, var_parSGT_ewma, var_hist_sGARCH, var_hist_eGARCH,
  var_hist_fiGARCH, var_hist_ewma, var_quant_sGARCH,
  var_quant_eGARCH, var_quant_fiGARCH, var_quant_ewma)

table_uc <- uctest_table(index_rt, vars)
table_bcp <- bcptest_table(index_rt, vars)

exceedances <- uctest_table(index_rt, vars, TRUE)

mean <- data.frame(matrix(NA,1,16)) ; colnames(mean) <- colnames(vars)
counter <-1
while (counter < 17){
mean[,counter] <- mean(vars[,counter])
counter <- counter+1
}

write.xlsx(table_uc, 'D:/Faculdade/Thesis/Code/uc_table.xlsx')
write.xlsx(table_bcp, 'D:/Faculdade/Thesis/Code/bcp_table.xlsx')
write.xlsx(exceedances, 'D:/Faculdade/Thesis/Code/exceedances.xlsx')
write.xlsx(mean, 'D:/Faculdade/Thesis/Code/mean.xlsx')

### Ploting
# Volatility Ploting
plot_std <- data.frame(c(std_ewma, std_sGARCH, std_eGARCH, std_fiGARCH));
  rownames(plot_std) <- rownames(std_ewma)
plot_std <- subset(x = plot_std, subset = rownames(plot_std) >= "2000-01-01")
write.xlsx(plot_std, 'D:/Faculdade/Thesis/Code/plot_std_hsi.xlsx', sheetName="Sheet1",
  col.names = TRUE, row.names = TRUE, append = FALSE, showNA = TRUE)
# Returns Ploting
write.xlsx(data.frame(index_rt), 'D:/Faculdade/Thesis/Code/plot_rt_hsi.xlsx',
  sheetName="Sheet1", col.names = TRUE, row.names = TRUE, append = FALSE, showNA = TRUE)

### Descriptive Statistics
mean <- mean(index_rt)
median <- median(index_rt)
max <- max(index_rt)
min <- min(index_rt)
sd <- sd(index_rt)
skewness <- skewness(index_rt)
kurtosis <- kurtosis(index_rt)
jarque.test(as.vector(index_rt)) ###????

des_stat_table <- data.frame(c(mean, median, max, min, sd, skewness, kurtosis))
rownames(des_stat_table) <- c('mean', 'median', 'max', 'min', 'sd', 'skewness', 'kurtosis');
colnames(des_stat_table) <- colnames(index_rt)

write.xlsx(des_stat_table, 'D:/Faculdade/Thesis/Code/table_descstat_hsi.xlsx',
  sheetName="Sheet1", col.names = TRUE, row.names = TRUE, append = FALSE, showNA = TRUE)

### Stationarity Test
# Prices
summary(ur.df(y=na.omit(HSI), lags=20, type="trend", selectlags="BIC"))
summary(ur.pp(na.omit(HSI), type="Z-tau", model="trend", lags="short"))
summary(ur.kpss(na.omit(HSI), type="tau", lags="short"))

# Returns
summary(ur.df(y=index_rt, lags=20, type="drift", selectlags="BIC"))
summary(ur.pp(index_rt, type="Z-tau", model="drift", lags="short"))
summary(ur.kpss(index_rt, type="tau", lags="short"))

### Long Memory Test
# ACF and PACF
par(mfrow=c(3,2))
# Returns

```

```
acf(index_rt, main = "ACF_□_□HSI_□Returns")
pacf(index_rt, main = "PACF_□_□HSI_□Returns")
# Returns Squared - Variance Proxy 1
acf(index_rt^2, main = "ACF_□_□HSI_□Returns^2")
pacf(index_rt^2, main = "PACF_□_□HSI_□Returns^2")
# |Returns| - Variance Proxy 2
acf(abs(index_rt), main = "ACF_□_□|HSI_□Returns|")
pacf(abs(index_rt), main = "PACF_□_□|HSI_□Returns|")

#LO R/S Statistic Test
rs_test(index_rt, 3)
rs_test(index_rt^2, 3)
rs_test(abs(index_rt), 3)
```