UNIVERSITAT ROVIRA I VIRGILI

AN EVALUATION OF THE PERFORMANCE OF VALUE-AT-RISK MODELS IN FINANCIAL CRISIS

Reem Shayya

UNIVERSITAT
ROVIRA i VIRGILI

An evaluation of the performance of Value-at-Risk Models in Financial Crisis

Reem Shayya

Doctoral Thesis

2024

Reem Shayya

An evaluation of the performance of Value-at-Risk Models in Financial Crisis

Doctoral Thesis

Supervised by:

Dr. Antonio Terceño Gómez

Dra. Maria Teresa Sorrosal Forradellas

UNIVERSITAT ROVIRA i VIRGILI

Department of Business Management

2024

UNIVERSITAT ROVIRA i VIRGILI

FAIG CONSTAR que aquest treball, titulat "An evaluation of the performance of Value-at-Risk Models in Financial Crisis", que presenta Reem Shayya per a l'obtenció del títol de Doctor, ha estat realitzat sota la meva direcció al Departament de Gestió d'Empreses d'aquesta universitat.

HAGO CONSTAR que el presente trabajo, titulado "An evaluation of the performance of Value-at-Risk Models in Financial Crisis", que presenta Reem Shayya para la obtención del título de Doctor, ha sido realizado bajo mi dirección en el Departamento de Gestión de Empresas de esta universidad.

I STATE that the present study, entitled "An evaluation of the performance of Value-at-Risk Models in Financial Crisis", presented by Reem Shayya for the award of the degree of Doctor, has been carried out under my supervision at the Department of Business Management of this university.

Reus 26 de febrer de 2024 / Reus 24 de febrero de 2024 / Reus february 24, 2024

El/s director/s de la tesi doctoral
El/los director/es de la tesis doctoral
Doctoral Thesis Supervisor/s

TERCEÑO GOMEZ, ANTONIO (AUTENTICACIÓN) Firmado digitalmente por TERCEÑO GOMEZ, ANTONIO (AUTENTICACIÓN) Fecha: 2024.02.26 11:36:18 +01'00'

MARIA TERESA SORROSAL FORRADELLAS - DNI 39358079L Firmado digitalmente por MARIA TERESA SORROSAL FORRADELLAS - DNI 39358079L Nombre de reconocimiento (DN): c=ES, sn=SORROSAL FORRADELLAS, givenName=MARIA TERESA, serialNumber=IDCES-39358079L, cn=MARIA TERESA SORROSAL FORRADELLAS - DNI 39358079L Fecha: 2024.02.26 22:11:00 +01'00'

Dr. Antonio Terceño Gómez

Dra. Mª Teresa Sorrosal Forradellas

Acknowledgements

Achieving this endeavor would not have been possible without those who accompanied me throughout my PhD. journey and are so dear to me.

First, I would like to express my deepest gratitude to my thesis supervisors, Antonio and Maria Teresa, who I was very lucky to have known and worked with. I would like to thank you for giving me the opportunity to work with you and pursue my doctoral studies provided with your advice, expertise, and knowledge. Without your support and guidance, it would have been impossible to reach this stage. Thank you for the encouragement and support when I was at my lowest and I needed them most.

Antonio, I am deeply indebted to you, not only for the guidance, advice, and support but more importantly for always showing me the appreciation that we all need to keep going. You have always boosted my confidence with your comments and support. You always directed my steps with patience and wisdom. I have learned a lot from you beyond the academic knowledge, for you are to me the perfect example of the modest, humble, generous, thoughtful, and wise leader. I am extremely grateful for Antonio, the “big brother”, who never missed any chance to make me, and Thalia feel the warmth of family when we were away from home. This is something I can never express my gratitude for. Thank you for making us feel like part of your family with all the care and kindness that you and your family surrounded us with.

Maria Teresa, Teresa, my supervisor, my friend, and my sister, this is who you are to me. Words cannot express my gratitude for all that you have done for me. I have been the luckiest to have you by my side throughout this journey. Thank you for your unwavering support, for believing in me and making me believe in myself, thank you for your patience, support, advice, and guidance throughout this journey. Ever since we met and started working together, I have been grateful for having such a thoughtful, sharp, and critical advisor. Despite the loads of work, repetitions, and changes, you have always been patient and understanding, and you always made sure to encourage me in every opportunity saying, “I see the light at the end of the tunnel”. I always find you the kindest person in the room, spreading your positive vibes with your smile everywhere you go. Thank you for being the amazing person you are, encouraging me at every defeat and celebrating with me every little achievement. Thank you for the care and love that you have surrounded me and Thalia with, always.

I would like to extend my gratitude to the Faculty of Business and Economics and precisely, to the Department of Business Management. In the first place, thank you for giving me the opportunity to join your department as a doctoral student. Thank you for the continuous and prompt support that all the members of the department provided me with throughout the whole period of my doctoral studies. You have always been kind and ready to help.

I would also like to express my gratitude to Dr. Oscar Martinez for his valuable advice and for sharing his skills in R, I really appreciate it.

I would be remiss in not mentioning my colleagues of AS11 office for making the time spent in the office enjoyable and productive. Special thanks to my colleagues who have become friends and more like family to me, Valeria and Ignacio. Together we shared joys and sorrows, and you were always there for me and Thalia and that really meant a lot to me. Thank you for your support and encouragement ever since we have known each other, you are amazing. I would also like to thank my friend, Laura, who always encouraged me especially in my defeats, and kindly offered moral support and shared knowledge and skills.

I would also like to thank the American University of Beirut (AUB), in particular, the Faculty of Arts and Sciences, Department of Economics, for giving me the opportunity to spend my three-month mobility stay at the faculty. My deepest gratitude goes to Dr. Simon Neaime, the Chairperson at AUB, who always showed his encouragement and support throughout my mobility stay at AUB. I would also like to thank Dr. Sumru Altug for her support during my stay at the AUB.

My deepest gratitude also goes to my parents, who always believed in me and encouraged me to keep going and follow my dreams. Thank you for teaching me that hard work always pays off, and that with patience, dedication, and perseverance any goals can be achieved. Thank you for being always there for me and supporting my choices and decisions, I cherish all what you have done for me and thank you is never enough for you both. I would also like to thank my brothers who were always there for me surrounding me with care and emotional support, you are my blessing.

I also thank my husband's family who were always very supportive and understanding, especially my father-in-law, who always encouraged me during my journey, and I am sure he is watching me from above now and happily cheering for me.

Last but foremost, to Nour, my husband, my partner, my friend, and my backbone. Thank you for being the super supportive person you are, for believing in me, for your patience and for keeping up with me. Thank you for lifting me up when I was at my lowest, with a few words you could make me see things in a different way, simpler and easier. You were always there for me when I needed you most despite being sometimes miles away, you never ceased to amaze me with your care and unconditional support.

Finally, I would like to dedicate this doctoral thesis to my daughter, the apple of my eye, the main motive behind this achievement, Thalia. This thesis is for you, for without you in my life, I might have given up. As I became a mother, thanks to you, I realized that it was my duty to be stronger than ever and fight for my dreams as much as I want you to fight for your dreams, now and always.

To Thalia

Abstract

This doctoral thesis presents an extensive study on several Value-at-Risk (VaR) models aiming to compare their performance during different economic situations and in different geographical areas. The VaR models studied are selected after conducting a Systematic Review of the Literature. Accordingly, the top 5 VaR models by number of articles and citation count are the GARCH family of models (applying the GARCH (1,1) model), Extreme Value Theory (EVT), Monte Carlo Simulation (MCS), Historical Simulation and variance-covariance model. Moreover, some other variants and sub-models were also applied: GARCH-EVT, Exponentially Weighted Moving Average, and Filtered Historical Simulation (FHS). These models are applied to the data of the indexes Dow Jones Industrial Average (DJIA), Euro Stoxx 50 (SX5E) and Nikkei 225 (N225), presenting the markets of the Unites States of America, Europe, and Japan, respectively, between 2002 and 2019. The period of the study is divided into three 6-year intervals before, during and after the crisis of 2008, which allows to test the performance of the models in different economic situations. For the model evaluation two methodologies are employed, one using backtesting measures, and the other proposed in this thesis (called "distance measures") which aims to measure the distance between the actual returns and the predicted VaR. The results highlight the strength of the MCS and GARCH (1,1) models in estimating the 99% and 95% VaR of all indexes. Only with DJIA, during the crisis period and under the 99% VaR, GARCH-EVT (90% and 95% thresholds) and FHS models attained good results besides the MCS and GARCH (1,1) models. It is also noted that the models applied in two frameworks of different types of residuals showed better results under Student-t rather than with normally distributed residuals.

Resumen

Esta tesis presenta un estudio sobre modelos de Valor en Riesgo (VaR) con el objetivo de comparar su desempeño en diferentes situaciones económicas y en diferentes áreas geográficas. Los modelos VaR estudiados se seleccionan tras una Revisión Sistemática de la Literatura. Los 5 principales modelos de VaR seleccionados por número de artículos y de citas son: la familia de modelos GARCH (aplicando el modelo GARCH (1,1)), la teoría de los valores extremos (EVT), la simulación de Monte Carlo (MCS), la simulación histórica y el modelo de varianza-covarianza. También se aplicaron algunas variantes y submodelos: GARCH-EVT, media móvil ponderada exponencialmente (EWMA) y simulación histórica filtrada (FHS). Estos modelos se aplican a los índices Dow Jones Industrial Average (DJIA), Euro Stoxx 50 (SX5E) y Nikkei 225 (N225), que representan los mercados de Estados Unidos, Europa y Japón, respectivamente, entre 2002 y 2019. El período de estudio se divide en tres intervalos de 6 años, antes, durante y después de la crisis de 2008, lo que permite comprobar el desempeño de los modelos en diferentes situaciones económicas. Para la contrastación de los modelos se emplean dos metodologías, medidas de backtesting y otra propuesta en esta tesis ("medidas de distancia") que tiene como objetivo medir la distancia entre las rentabilidades reales y el valor estimado del VaR. Los resultados resaltan la fortaleza de los modelos MCS y GARCH (1,1) para estimar el VaR al 99% y 95% de todos los índices. Sólo con el DJIA, durante el período de crisis y al 99%, los modelos GARCH-EVT (con umbrales del 90% y 95%) y el modelo FHS obtuvieron buenos resultados, junto con los modelos MCS y GARCH (1,1). También se observa que los modelos aplicados bajo el supuesto de diferentes tipos de residuos mostraron mejores resultados utilizando la distribución t-Student que con residuos distribuidos normalmente.

Resum

Aquesta tesi presenta un estudi sobre diversos models de Valor en Risc (VaR) amb l'objectiu de comparar el seu comportament en diferents situacions econòmiques i en diferents àrees geogràfiques. Els models VaR estudiats se seleccionen després d'una Revisió Sistemàtica de la Literatura. Els 5 principals models de VaR seleccionats per nombre d'articles i de cites són: la família de models GARCH (apliqcant el model GARCH (1,1)), la teoria dels valors extrems (EVT), la simulació de Monte Carlo (MCS), la simulació històrica i el model de variància-covariància. També es van aplicar algunes variants i submodels: GARCH-EVT, mitjana mòbil ponderada exponencialment (EWMA) i simulació històrica filtrada (FHS). Aquests models s'apliquen als índexs Dow Jones Industrial Average (DJIA), Euro Stoxx 50 (SX5E) i Nikkei 225 (N225), que representen els mercats dels Estats Units, Europa i Japó, respectivament, entre el 2002 i el 2019. El període d'estudi es divideix en tres intervals de 6 anys, abans, durant i després de la crisi de 2008, cosa que permet comprovar els resultats dels models en diferents situacions econòmiques. Per a la contrastació dels models es fan servir dues metodologies, mesures de backtesting i una proposada en aquesta tesi ("mesures de distància") que té com a objectiu mesurar la distància entre les rendibilitats reals i el VaR estimat. Els resultats ressalten la fortalesa dels models MCS i GARCH (1,1) per estimar el VaR al 99% i al 95% de tots els índexs. Només amb el DJIA, durant el període de crisi i al 99%, els models GARCH-EVT (amb llindars del 90% i 95%) i el model FHS van obtenir bons resultats, a més dels models MCS i GARCH (1,1). També s'observa que els models aplicats sota el supòsit de diferents tipus de residus van mostrar millors resultats utilitzant la distribució t-Student que amb residus distribuïts normalment.

Table of Contents

List of Tables

List of Figures

Introduction

Introduction

Over the last few decades, Value-at-Risk (VaR) has been the most commonly used tool in risk management to assess the risk exposures of firms and financial institutions, with the aim to help avoid catastrophic falls due to market fluctuations and sudden events. Its popularity is related with the issuance of the first Basel accord, known as Basel I, by the Basel Committee on Banking Supervision (BCBS), which aimed to provide recommendations on banking regulations to ensure that financial institutions have sufficient capital that would meet any obligations and prevent failures or bankruptcies due to sudden unexpected losses.

Basel I accord (1988) introduced an explicit capital cushion for the market risks that financial institutions are exposed to due to their market trading activities. Later, the BCBS issued the amendment (Basel Committee on Banking Supervision, 1996) on Basel I accord (1988), which allowed firms and banks to use their own internal VaR models to estimate their corresponding market risks. This led to the growth of the academic literature on existing VaR models, comparing alternative modelling approaches and proposing new VaR models as an attempt to improve the already existing models. Furthermore, Basel II (2004) set out the guidelines on the use of VaR models in the banking industry for market risk estimation.

An intuitive definition of VaR can be found in Jorion (2007) stating that "VaR summarizes the worst loss over a target horizon that will not be exceeded with a given level of confidence". According to Holton (2002) the origins of VaR can be traced back as far as 1922 to capital requirements the New York Stock Exchange imposed on member firms, moreover, VaR has roots in portfolio theory and a crude VaR measure was published in 1945. VaR models can be classified into three main approaches: (i) Historical Simulation approach (non-parametric approach), (ii) variance-covariance approach (parametric approach), and (iii) Monte Carlo Simulation approach (semi-parametric approach). These approaches are considered the traditional VaR models which were developed into more sophisticated models incorporating complex statistical, financial, and mathematical notions. Some of the main VaR models that were developed and have become very popular among academics and firms are the Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) family models which are based on the GARCH process proposed by Bollerlsev (1986). Later, it will be seen how these models grew to become the most studied and used models in the literature and in the industry. Another important line of models developed to improve the accuracy of the estimation of VaR is based on the Extreme Value Theory (EVT) which constitutes of two groups of models, the Block Maxima Models (BMM) and the Peaks Over Threshold (POT) models noting that the latter is more extensively studied and used. As per our review of the literature on VaR, the EVT approach was incorporated in the estimation of VaR by Pownall and Koedijk (1999). And later, McNeil and Frey (2000) proposed the GARCH-EVT model which is based on the combination of EVT approach with GARCH models. This line of models was later developed by many authors like Byström (2004), Gençay and Selçuk (2004), and Huang et al. (2011) among others. Other models were developed based on variance-covariance approach like the RiskMetrics developed by Morgan (1996) which is based on the Exponentially Weighted Moving Average (EWMA) model. Many other models on VaR were proposed to improve the quality of the VaR estimates produced to enhance the accuracy of risk measurement, in fact, according to the Systematic Review Literature (SRL) conducted in this doctoral thesis in Chapter 1, 28 models were found to be incorporated in the academic literature on VaR models, of these models we mention Copulas, Moving Average models, Fourier transform based models, Quantile Regression models, Cornish-Fisher based models, models based on the Fuzzy logic and many others.

This doctoral thesis compares the behavior of eight of these VaR models during a crisis, in particular the financial crisis of 2008, aiming to distinguish the VaR models with the best behavior with stock indexes in three different markets, the United States of America presented by Dow Jones Industrial Average index (DJIA), the European market presented by the Euro Stoxx 50 index (SX5E), and the Japanese market presented by the Nikkei 225 index (N225). The financial crisis of 2008 was chosen to test the behavior of the VaR models since it is the most recent crisis with enough data to be used to conduct our research. The period of study lying between 2002 and 2019 is divided into three 6-year intervals presenting the pre-crisis, crisis, and post-crisis periods. This division allows to consider whether a model behaves better than others in crisis situations, and whether a model is better for a particular market.

The choice of the VaR models was not random, the criterion used is to select the models that are most used in the financial literature. In fact, the first stage of this doctoral thesis, presented in Chapter 1, is a SRL on VaR models between 1996 and 2017. This SRL covers almost all VaR models used in the industry and academic literature, highlighting their strengths and weaknesses. Accordingly, the SRL aims to find out the most used VaR models studied and proposed especially after the financial crisis of 2008 to come up with a recommendation regarding the models that showed the highest efficiency. Consequently, the top five VaR models according to the SRL were chosen for the comparative study conducted in this research. These models are: GARCH family models, Extreme Value Theory (EVT), Monte Carlo Simulation (MCS), Historical Simulation (HS), and variance-covariance model (var-cov).

In addition, some other models were also studied in this research considered as variants or sub-models of the top five VaR models. Hence, eight models were applied and analyzed in this study and they are: (i) GARCH (1,1) model (as a GARCH family model), (ii) EVT-Peaks Over Threshold approach (EVT-POT), (iii) GARCH-EVT model considered as a variant of EVT, (iv) MCS model applied under GARCH (1,1) model with normal and Student-t distributed residuals, (v) HS (traditional model), (vi) Filtered Historical Simulation (could be considered a variant of MCS), (vii) Exponentially Weighted Moving Average model (EWMA) (considered a sub-model of the var-cov model), and (viii) the var-cov model (traditional approach). It is also worth noting that some models are applied in more than one framework depending on the methodology followed in the implementation process and whether the model considers residuals which can be modelled with different distributions. MCS and GARCH (1,1) models are applied with normal and Student-t residuals and the implementation of the models under the latter framework involves two methods which distinguish the method of calculation of the degree of freedom of the Student-t residuals. FHS models is also applied under two types of residuals, normal and Student-t residuals. GARCH-EVT model (with 90% and 95% thresholds) is applied in three ways, with normal residuals, Student-t residuals and using the quasi-maximum likelihood estimation (QMLE) method. And EWMA model is also applied with six values of the decay factor (from 0.94 till 0.99, included) and using the root mean squared error (RMSE) only two values (0.94 and 0.95) were included in the analysis of this research.

In Chapter 2, the eight VaR models are defined and explained extensively according to their original definitions and proposals. It is also shown how each of the models can be implemented and the methodology followed in the implementation process. This information is essential to create a better understanding of the ideology behind each model and also know how the model works.

Chapter 3 constitutes the implementation of the VaR models considering the two most usual confidence levels for VaR estimates, the 95% and the 99% levels. The obtained VaR estimates for each of the eight models are analyzed. The results are shown graphically in the chapter and numerical results are provided in the appendix. Moreover, Chapter 3 involves an analysis of the VaR estimates using distance measures which we propose in this study to analyze the performance of VaR models. Considering distances between the VaR estimates and the actual returns (both profits and losses) it is possible to analyze the behavior of the VaR models not only with negative market fluctuations but also when the index is realizing profits. At the end of this chapter, the best performing models are distinguished by index and by time interval. The model implementation was executed using the R Studio software and Microsoft Excel and distance measures were carried out in Excel.

Backtesting measures are typical measures to validate the performance of VaR models. In Chapter 4, backtesting measures are used to further analyze the performance of the eight VaR models applied in Chapter 3. These measures are usually based on analyzing the performance of VaR models considering violations to VaR, i.e., when the actual loss exceeds the VaR predicted by the model to determine whether a model underestimates or overestimates VaR. They evaluate the efficiency of VaR models for example by checking the frequency of these violations, whether it is close to the promised VaR coverage probability. Some measures check the independence among these violations and relate the timing of these violations to market fluctuations, while others, for instance, check the time to the first violation or the duration between two consecutive violations and many other proposals in this sense. Thus, in general, and in this thesis in particular, the backtesting measures used are based on violations and do not consider the magnitude of the losses. In this sense, backtesting measures can be considered complementary to the distance measures proposed in this thesis. The backtesting measures used in this research are Kupiec (Proportion of Failure) test (1995), Dynamic Quantile (DQ) test of Engle and Manganelli (2004), and the Duration-based (D-B) approach of Christoffersen and Pelletier (2004). Kupiec (POF) test was chosen in this research due to its popularity in the literature and in the industry and it is incorporated in the framework for backtesting internal models recommended by the BCBS in the amendment of 1996 (Ziggel et al., 2014). While the DQ test was employed in this study because it is the most popular test among the tests relying on regression models(Dumitrescu et al., 2012). However, the D-B test was involved in this research because it has much better power properties than the two most popular backtesting measures, the unconditional coverage test and conditional cover tests of Christoffersen (1998). These measures were applied using R Studio software.

Chapter 5 constitutes the conclusions of this research and draws the findings clearly distinguishing the models that showed the best performance by combining the analysis of both model validation criteria, distance measures and backtesting measures. The concept of distance measures for evaluating VaR models differs from backtesting measures, which are very popular and widely used in the industry and literature, by the fact that it does not only work with violations to VaR and analyzing their frequency and relating it to the market, but it measures the magnitude of the violations. This is important because it is not important to know when the model fails to predict VaR but also the magnitude of this failure and how much it will affect the market especially if one also considers the frequency of these violations. This was the motivation behind combining the analysis of distance measures and backtesting measures to obtain a better understanding of the models' behavior. Chapter 5 also includes the main limitations of the thesis and some possible future lines research.

In summary, the contribution of this thesis is that it presents a wide comparative analysis between eight of the most popular and important VaR models within a broad period of time between 2002 and 2019 that has been divided into three different intervals presenting different economic situations, pre-crisis, crisis and post-crisis periods. The emphasis of this doctoral thesis is to seek the model(s) that provides the best performance in the different economic situations, i.e., during calm and volatile periods, as well as the model(s) that suits each market (presented by a stock index) according to its geographical area. Moreover, this thesis also highlights that backtesting measures which focus on evaluating VaR models through violations, which are the main concern of institutions, are complemented by distance measures which focus on the magnitude of violations as well as the distance between VaR and positive returns (profits) which is also important for portfolio managers.

References

Bank for International Settlements, & Basel Committee on Banking Supervision. (2004). A revised framework on international convergence of capital measurement and capital standards. *Bank for International Settlements, Basel*, *1*, 285.

Basel Committee on Banking Supervision. (1988). *International Convergence of Capital Measurement and Capital Standards*.

Basel Committee on Banking Supervision. (1996). *Amendment to the capital accord to incorporate market risks*. *January*. https://www.bis.org/publ/bcbs24.pdf

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

Byström, H. N. E. (2004). Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, *13*(2), 133–152. https://doi.org/10.1016/j.irfa.2004.02.003

Christoffersen, P. (1998). EVALUATING INTERVAL FORECASTS. *International Economic Review*, *39*(4), 841–862. https://doi.org/10.2307/2527341

Christoffersen, P., & Pelletier, D. (2004). Backtesting Value-at-Risk: A Duration-Based Approach [Article]. *Journal of Financial Econometrics*, *2*(1), 84–108. https://doi.org/10.1093/jjfinec/nbh004

Dumitrescu, E. I., Hurlin, C., & Pham, V. (2012). Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests [Article]. *Finance (Paris)*, *33*, 79–112.

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, *22*(4), 367–381. https://doi.org/10.1198/073500104000000370

Gençay, R., & Selçuk, F. (2004). Extreme value theory and Value-at-Risk: Relative performance in emerging markets. *International Journal of Forecasting*, *20*(2), 287–303. https://doi.org/10.1016/j.ijforecast.2003.09.005

Holton, G. A. (2002). *History of Value-at-Risk: 1922-1998* (Method and Hist of Econ Thought). University Library of Munich, Germany. https://EconPapers.repec.org/RePEc:wpa:wuwpmh:0207001

Huang, S.-C., Chienb, Y.-H., & Wangc, R.-C. (2011). Applying GARCH-EVT-copula models for portfolio value-at-risk on G7 currency markets. *International Research Journal of Finance and Economics*, *74*, 136–151.

Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*. McGraw-Hill.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, *3*(2), 73–84.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of*

Empirical Finance, *7*(3–4), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8

Morgan, J. P. (1996). *RiskMetrics-Technical Document* (Fourth Edition). J.P. Morgan/Reuters. https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a

Pownall, R. A. J., & Koedijk, K. G. (1999). Capturing downside risk in financial markets: The case of the Asian crisis. *Journal of International Money and Finance*, *18*(6), 853–870. https://doi.org/10.1016/S0261-5606(99)00040-6

Ziggel, D., Berens, T., Weiß, G. N. F., & Wied, D. (2014). A new set of improved Value-at-Risk backtests. *Journal of Banking and Finance*, *48*, 29–41. https://doi.org/10.1016/j.jbankfin.2014.07.005

Chapter 1

Value-at-Risk models: A Systematic Review of the Literature

1.1. Introduction

After witnessing several severe unpredictable catastrophic events, practitioners, academics and regulators from the financial community became more interested in analyzing extreme events and measuring and managing their consequential risk and losses. Measuring different types of risk became a necessity after the issuance of Basel I, also called Basel Accord, which was published in 1988 by the Basel Committee on Banking Supervision (BCBS) after the increase in bankruptcies in the U.S. and throughout the world. This bankruptcy situation was due to extensive lending and simultaneously the countries' external indebtedness that was growing at an unsustainable rate. Basel I was the first publication of the BCBS that included recommendations on

banking regulations regarding capital, market, and operational risks. In fact, the main goal of this accord was to ensure that banks and financial institutions maintain a capital that would be sufficient enough to cover any unexpected losses and meet financial obligations.

Value-at-Risk (VaR) methodology was founded to meet the need of financial institutions to measure market risk. It measures the value of losses likely to be incurred in case the value of a portfolio's assets declines, expressed in terms of a predetermined confidence interval over a specified period of time. In other words, the VaR of the portfolio gives the maximum amount of loss that an investor might face over a certain time horizon with a given probability. VaR has become a basic market risk management tool since financial institutions were required to meet the capital requirements based on VaR estimates. Banks and financial institutions were allowed to use their own internal models for calculating their VaR which was the reason behind the evolution of academic literature and proposals of new models for estimating VaR ever since Basel I was put into action and up till now.

VaR is calculated according to three basic methodologies:

1. Historical Simulation approach (non-parametric method), which consists of re-organizing historical returns and ordering them from worst to best and then taking the assumption that history repeats itself.
2. Variance – Covariance approach (Parametric method), takes the assumption that stocks are normally distributed.
3. Monte Carlo Simulation approach (semi-parametric method), which consists of developing a model for predicting future stock returns and running multiple trials through this model.

This chapter represents a review of the models developed under each methodology within the time period between 1996 and 2017, recognizing the most used ones and analyzing the evolution and trend of growth of the number of the respective articles published on each model, along with citation count, authors and the journals with publication in this field.

It is worth noting that this review is the first of its kind in the domain of VaR estimation and it can be a good reference for academics who are interested in developing new models, and also for practitioners in banks and institutions, who use in-house VaR models developed by their analytics teams. This review covers almost all most popular VaR models used in the industry highlighting their strengths and weaknesses and it concludes with a recommendation of certain models that proved their efficiency especially after the

Financial Crisis of 2008. If firms could accurately measure their VaR, they would be able to avoid huge failures by reserving the required capital sufficient to cover the potential losses.

This chapter is structured as follows:

Section 1.2 discusses the Systematic Review of the Literature (SRL) which is a credible objective reviewing methodology. Section 1.3 contains the search criteria in the selected database. Section 1.4 displays the search results of the review obtained. Section 1.5 is a summary for the whole SRL carried out in the chapter. And finally, the conclusions are drawn in section 1.6.

1.2. SRL Review Methodology

The Systematic Review of the Literature (SRL) initially emerged in the 1990s and it was first developed in the field of medicine (Dickersin et al., 1994; Evans, 2001; Mulrow et al., 1997). Later, this methodology was introduced to other scientific fields like business management and accounting (Crossan & Apaydin, 2010; Mazzi, 2011; Riordan, 2000), social responsibility (Peloza & Shang, 2011) and many other fields. The SRL methodology helps researchers avoid subjectivity in their work contrary to narrative reviews which may be affected by the author's subjective criteria and include studies that support his/her personal opinion (Cook et al., 1997; Rosenthal, 1991). In this sense, the SRL methodology assures the elimination of any possible bias by the author in his/her choice of publications by achieving an unbiased and balanced summary of the literature and identifying all research addressing a specific question and specifically including all negative that might be published in low impact journals (Nightingale, 2009). According to Nightingale (2009), this miantains the balance in the results contrary to ordinary literature reviews which include mainly the easily identified positive studies that are published in high impact journals.

Moreover, another major difference between systematic reviews and traditional narrative reviews is that systematic reviews tend to be more scientific and reliable as they adopt a scientific, replicable, and transparent process which is based on a detailed protocol aiming to minimize bias through detailed search and tracking of published and unpublished studies (Tranfield et al., 2003). As per Cook et al. (1997) and Rosenthal (1991), a major difference between the SRL and the traditional narrative review is that the former incorporates critical analysis for the obtained results and does not generate subjective or personal opinions.

Tranfield et.al. (2003) organize the SRL in three stages that are: (1) planning, (2) execution, and (3) reporting and dissemination.

In the planning stage the researcher defines the importance and the need for the review, prepares a proposal with its general and specific objectives and determines the search criteria and the database being used. The execution stage involves searching in the selected database for the articles related to the subject of the study according to the terms and search questions. In the third stage the search results are analyzed, and a summary is conducted to answer the research questions and objectives.

1.3. SRL: Planning Stage

What are the most popular VaR models? This is an indirect question on the accuracy of the VaR models that were used before the global financial crisis of 2008 and were not accurate enough to avoid the financial failure. Scopus is used with an interval from 1996 until 2017 in order to answer these questions.

The search is based on using the "title, abstract, keyword" to find documents on VaR, using the terms "value-at-risk", "value at risk" and "VaR". Using "VaR" a huge number of documents (105.000 as at 21.11.2018) was obtained, many of which were not relevant to our topic ("VAR" is also used to denote "Autoregressive Vector Regression") and thus were removed from the study. Hence, the search was limited to the following topics: "Economics, Econometrics and Finance", "Business, Management and Accounting" and "Decision Sciences". The document type was limited to Articles, in English, and the source type was limited to Journals.

After obtaining the results, the evolution trend and short reviews of the literature on VaR models with more than 100 articles and more than 30 citations are analysed. The aim of this analysis is to show when each of these models was first introduced into the estimation of VaR, and focus on its evolution trend, the growth in the number of articles published on these models.

1.4. SRL Execution Stage: Search in Scopus

Searching in Scopus for "value-at-risk" or "value at risk", the total number of publications obtained before applying any filters amounted to 5.146 (as at 21.11.2018) (noting that the articles obtained for the keyword "VaR" which are related to value-at-risk are already among those obtained for the keywords "value-at-risk" and "value at risk", so we removed "VaR" keyword to avoid having articles on Autoregressive Vector Regression).

Before limiting the search to any subject areas, the first publication on VaR first appeared in 1982 and it was related to two subject areas: "Earth and Planetary Sciences" and "Environmental Sciences". It is also worth noting that until 1995 the publications issued on VaR were related to the fields of "Earth and Planetary Sciences", "Environmental Sciences", "Immunology and Microbiology" "Medicine", "Engineering", "Materials Science" and "Agricultural and Biological Sciences". However, the total number of publications between 1982 and 1995 amounted to 9, with an average of approximately one publication per year. In 1996, the number of publications increased to 6 (from 1 in 1995) which is considerably higher than the numbers of publications in the previous years. Table 1.1 demonstrates the distribution and the cumulative distribution of number of publications per year.

Year	Number of Articles	Cumulative Distribution of number of Articles	Year	Number of Articles	Cumulative Distribution of number of Articles
1982	1	1	2000	42	119
1983	0	1	2001	69	188
1984	2	3	2002	62	250
1985	1	4	2003	91	341
1986	1	5	2004	98	439
1987	0	5	2005	158	597
1988	0	5	2006	187	784
1989	1	6	2007	229	1,013
1990	0	6	2008	261	1,274
1991	0	6	2009	347	1,621
1992	0	6	2010	378	1,999
1993	2	8	2011	360	2,359
1994	0	8	2012	396	2,755
1995	1	9	2013	456	3,211
1996	6	15	2014	478	3,689
1997	15	30	2015	447	4,136
1998	11	41	2016	476	4,613
1999	36	77	2017	534	5,147
Total Number of Articles				**5,146**	

Table 1.1. Distribution of all publications on Value-at-Risk per year till 2017 as at 21.11.2018.

After applying all the search filters mentioned in the planning stage 1.505 articles were obtained and after removing duplicates the final number of articles amounted to 1.502 and they are presented in Table 1.2 classified by year of publication.

The first publication on VaR, as per the field selections of this study, appeared in 1996 and it was classified in "Economics, Econometrics and Finance". For this reason, our research was limited between 1996 (as a start date) and 2017 (chosen as the end date).

As it is clearly shown in Table 1.2, the number of publications per year follows an increasing trend, from 1 publication in 1996 to an average of 22 between 1997 and 2004, 56 between 2005 and 2006, 71 between 2007 and 2008, and finally to an average of 119 thereafter. The considerable increase in the average number of publications per year from the interval [1997 – 2004] to the interval [2004 – 2006] can be explained by the issuance of Basel II in June 2004. However, the increase from the latter interval of time to the interval [2007 – 2008] can be referred to the global financial crisis of 2008.

Year	Number of Articles	Cumulative Distribution of Number of Articles	Year	Number of Articles	Cumulative Distribution of Number of Articles
1996	1	1	2007	71	360
1997	10	11	2008	71	431
1998	6	17	2009	86	517
1999	16	33	2010	103	620
2000	21	54	2011	109	729
2001	30	84	2012	116	845
2002	37	121	2013	118	963
2003	27	148	2014	134	1,097
2004	30	178	2015	136	1,233
2005	42	220	2016	116	1,349
2006	69	289	2017	153	1,502
Total Number of Articles				**1,502**	

Table 1.2. Distribution of Filtered Articles on Value-at-Risk per year between 1996 and 2017 (as at 21.11.2018).

The next step consisted of searching for the models of VaR also in the "Abstract, Title or Keywords" where the search gives articles which are really focused on the models and not just mentioning them.

The work on VaR model started after data cleaning. The outcome is shown in Table 1.3 by a list of models listed in a descending order as per number of articles. A total of 27 models of VaR was recognized and those with more than 100 articles and having a citation count above 30 were analyzed.

Some models/risk measures that were found to be alternatives to VaR and not precisely modelling VaR were excluded from the results. Some of these models are: Conditional VaR (CVaR), Conditional Autoregressive Value-at-Risk (CAViaR), the Liquidity Adjusted Value-at-Risk, the Modified Value-at-Risk, the Stressed Value-at-Risk, the Entropic Value-at-Risk, and Multidimensional Value-at-Risk, among others.

Model	Sub Model	Number of Articles	Total	Citation Count
ARCH/GARCH Family	GARCH	307	335	4,825
	ARCH	32		
	EGARCH	24		
	FIGARCH	22		
	FIAPARCH	12		
	APARCH	11		
	Hyperbolic GARCH	4		
	AVGARCHM model	1		
Extreme Value Theory	Multivariate Extreme Value at risk	1	178	3,213

	Extreme Value-at-risk	3		
	Extreme Value Theory	178		
Monte Carlo Simulation	Monte Carlo Simulation Method	160	160	2,302
	Quasi Monte Carlo	4		
Historical Simulation	Historical Simulation Method	154	142	2,279
	Filtered Historical Simulation	19		
	Derivative-based VaR/Diversified VaR	1		
Variance Covariance Method	Variance-Covariance Method	134	134	1,733
Copulas	Asymptotic Value-at-Risk	1	108	1,008
	Copulas	108		
Generalized Pareto Distribution	Generalized Pareto Distribution	62	62	882
Moving Average Models	Moving Average	7	59	499
	ARMA	26		
	Exponentially Weighted Moving Average	20		
	ARFIMA	5		
	ARIMA	4		
	Equally Weighted Moving Average	4		
Quantile Regression	Quantile Regression	58	58	732
RiskMetrics	RiskMetrics	44	47	643
	Component Value at Risk	3		
Fourier	Fourier	16	21	89
	Entropic Value-at-Risk	6		
G-and-H Distribution	Generalized Hyperbolic Distribution	17	17	172
	Tukey's G-and-H	2		
Cornish-Fisher Value-at-Risk	Cornish-Fisher	13	13	116
Heterogeneous Autoregressive Model	Heterogeneous Autoregressive	12	12	78
Wavelet Value-at-Risk	Wavelet Value-at-Risk	12	12	234
Levy Model	Levy Model	8	8	41
Quadratic Value at Risk	Quadratic Value at Risk	1	8	168
	Delta-Gamma	8		
Multivariate Value-at-Risk	Multivariate Value at Risk	7	7	67

CreditRisk+	CreditRisk+	7	7	528
CreditMetrics	CreditMetrics	6	6	575
Fuzzy Models	Fuzzy Models	6	6	105
Granularity Adjustment	Granularity Adjustment	6	6	34
Black-Litterman Model	Black-Litterman	4	4	31
KMV	KMV	4	4	551
Artificial Intelligence	Artificial Intelligence	3	3	11
Expectile	Expectile	2	2	60
Threshold stochastic volatility model	Threshold stochastic volatility model	2	2	16
Total before removing duplicates		**1,551**	**1,421**	**20,992**
Total after removing duplicates		**818**	**818** [1]	**11,058**
Excluded Models		**391**	**391**	
Others		**293**	**293**	
Total Number of Articles		**1,502**	**1,502**	

Table 1.3. Distribution of articles per model.

Figure 1.1 presents the number of articles published on each of the five most popular models according to Table 1.3. It is obvious that the ARCH/GARCH family of models is the most widely used in the estimation of VaR with a significant increase in number of publications compared to the other models especially after 2008.

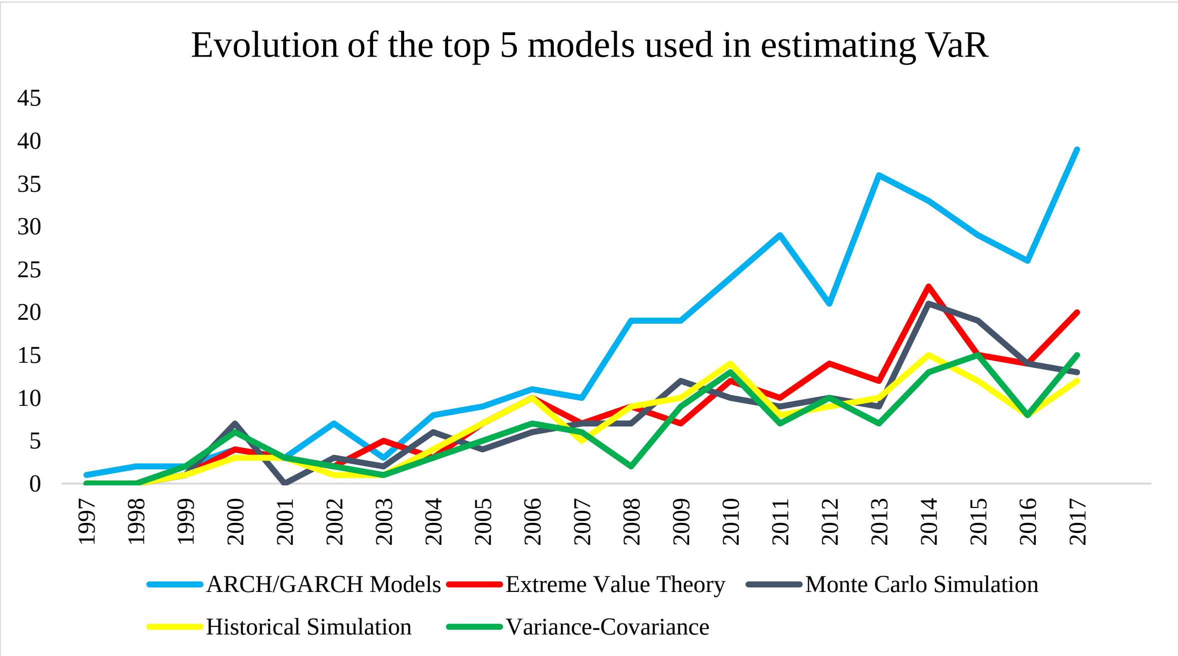

Figure 1.1. Evolution of VaR Models between 1996 and 2017.

[1] It is worth noting that the number of articles decrease from 1,502 to 818 after focusing our analysis on "Models on VaR" and excluding "alternative measures to VaR".

Brief reviews on VaR models that are used in more than 30 articles and have over 100 citation count, are presented in the following subsections, in addition to the evolution of each model, starting with its first published articles along with a glimpse on its development and the way its utility was affected after the crisis of 2008. For detailed results, Appendix A comprises the obtained data classified by models including journals with publications above 10 citations, journals with more than 4 publications, top 10 publications by citation count, top 10 authors by citation count and by number of publications and top 10 journals by number of citations and number of publications.

Regarding the development of VaR models, Table 1.4 gives the percentage of articles of the models published in the same year. The models with more than 30 published articles and more than 100 citations are ARCH/GARCH Models, Extreme Value Theory, Monte Carlo Simulation Method, Historical Simulation Method, Variance Covariance Method and Copulas, Generalized Pareto Distribution, Moving Average, Quantile Regression and RiskMetrics. The number of articles corresponding to these models constitute almost 80% of the total number of articles.

Year	ARCH/ GARCH Models	% Total ARCH/ GARCH Models	Extreme Value Theory	% Total EVT	Monte Carlo	% Total Monte Carlo Simulation	Historical Simulation	% Total Historical	Variance Covariance	% Total Var-Cov	Copulas	% Total Copulas	GPD	% Total GPD	Moving Average	% Total MA	Quantile Regression	% Total QR	RiskMetr.	% Total RiskMetr.	Total
1996	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1997	1	10	0	0	0	0	0	0	0	0	0	0	0	0	1	10	0	0	1	10	10
1998	2	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
1999	2	13	1	6	1	6	1	6	2	13	0	0	0	0	0	0	1	6	2	13	16
2000	4	19	4	19	7	33	3	14	6	29	0	0	0	0	0	0	0	0	2	10	21[2]
2001	3	10	3	10	0	0	3	10	3	10	1	3	0	0	1	3	0	0	0	0	30
2002	7	19	2	5	3	8	1	3	2	5	0	0	0	0	2	5	0	0	0	0	37
2003	3	11	5	19	2	7	1	4	1	4	1	4	3	11	0	0	0	0	1	4	27
2004	8	27	3	10	6	20	4	13	3	10	2	7	1	3	3	10	0	0	6	20	30
2005	9	21	7	17	4	10	7	17	5	12	1	2	1	2	2	5	0	0	1	2	42
2006	11	16	10	14	6	9	10	14	7	10	2	3	4	6	2	3	3	4	1	1	69
2007	10	14	7	10	7	10	5	7	6	8	1	1	4	6	1	1	4	6	3	4	71
2008	19	27	9	13	7	10	9	13	2	3	1	1	4	6	1	1	3	4	1	1	71
2009	19	22	7	8	12	14	10	12	9	10	9	10	0	0	7	8	2	2	3	3	86
2010	24	23	12	12	10	10	14	14	13	13	7	7	6	6	1	1	4	4	6	6	103
2011	29	27	10	9	9	8	8	7	7	6	11	10	2	2	5	5	7	6	3	3	109
2012	21	18	14	12	10	9	9	8	10	9	5	4	9	8	3	3	7	6	4	3	116
2013	36	31	12	10	9	8	10	8	7	6	14	12	3	3	3	3	5	4	3	3	118
2014	33	25	23	17	21	16	15	11	13	10	13	10	3	2	4	3	4	3	2	1	134
2015	29	21	15	11	19	14	12	9	15	11	14	10	10	7	9	7	5	4	6	4	136
2016	26	22	14	12	14	12	8	7	8	7	10	9	5	4	8	7	4	3	2	2	116
2017	39	25	20	13	13	8	12	8	15	10	16	10	7	5	6	4	9	6	0	0	153
Total	335	22	178	12	160	11	142	9	134	9	108	7	62	4	59	4	58	4	47	3	1,502

Table 1.4. Evolution of VaR models by percentage of total articles per year.

1.4.1. ARCH/GARCH Models

The ARCH (AutoRegressive Conditional Heteroskedasticity) process was first introduced in Engle (1982) allowing for conditional variance to vary over time as a function of past errors while keeping the unconditional variance constant.

The Generalized ARCH (GARCH) model was later proposed by the Danish economist Bollerslev (1986) as a generalization of the famous ARCH model of Engle (1982). The main reason for the development of the GARCH model was to facilitate the computation of the polynomial lag operator (that is found in the ARCH model) when it presents a high order. It was mostly used in VaR estimation where the volatility of returns is a central issue.

[2] The sum of articles in a year might exceed the corresponding total number of articles in the same year. This is due to the fact that some articles are classified to belong to more than one model.

According to the search results, the ARCH/GARCH models that have been mostly used in the estimation of VaR are the original ARCH, GARCH models, the Exponential GARCH (EGARCH) model, Fractionally Integrated Asymmetric Power ARCH (FIAPARCH) model, Fractionally Integrated GARCH (FIGARCH) model, Asymmetric Power ARCH (APARCH) model, Hyperbolic GARCH (HYGARCH) model, the Absolute Value GARCH in the Mean (AVGARCHM) model and the Glosten Jagannatan Runkle GARCH (GJR-GARCH) model among others.

The GARCH (p, q) process, where “p” stands for the order of the polynomial which refers to the autoregressive term and “q” is the order of the polynomial which refers to the Moving Average term of the model, improves ARCH (q) because it allows lagged conditional variances contrary to the ARCH (q) process where the conditional variance is specified as a linear function of past sample variances only (Bollerslev, 1986).

The GARCH model has become a part and parcel in building financial models since it has proved its efficiency in predicting and analyzing the size of errors of a model. GARCH models deal with volatility, specifically with the heteroskedasticity, which arises when the variances of error terms in a given data are not equal, i.e., suffer from volatility where the error terms are expected to be larger at some points in the data than others, which is actually the real case rather than the assumption of normality of distribution of returns which was also proved to be some sort of the “ideal case” and cannot be much reliable in the real world.

According to the search criteria followed in this study, the use of GARCH models in the estimation of VaR did not start until 1997 with the first article published on this topic in the “Journal of Derivatives” by Alexander C.O. and Leigh C.T. under the title “On the covariance matrices used in value at risk models”. It can be seen in Table 1.5 that after 1997, the number of articles published yearly witnessed an overall increasing trend, reaching 7 articles in 2002 and then 9 articles in 2005. The average number of articles published before the release of Basel II was 3 articles per year between 1997 and 2003. In 2004, the number of articles increased to 8 then increased to an average of 10 articles per year between 2005 and 2007. Between 2008 and 2017, the average number of publications amounted to approximately 28 articles annually indicating a considerable increase in the trend of evolution of the use of ARCH/GARCH models in estimating VaR. This analysis provides evidence about the usefulness of the family of ARCH/GARCH models in estimating VaR results derived from incorporating GARCH in the estimation of VaR.

Year	Number of Articles	Year	Number of Articles
1997	1	2008	19
1998	2	2009	19
1999	2	2010	24
2000	4	2011	29
2001	3	2012	21
2002	7	2013	36
2003	3	2014	33
2004	8	2015	29
2005	9	2016	26
2006	11	2017	39
2007	10		
	Total	**335**	

Table 1.5. Distribution of articles on VaR estimation using ARCH/GARCH per year of publication.

1.4.2. Extreme Value Theory

The Extreme Value Theory (EVT) is another widely used model in the estimation of VaR. EVT constitutes of two principal groups of models, the Block Maxima Models (BMM) and the Peaks Over Threshold (POT). POT models are considered as the modern group, and they are more popular due to their efficiency in using data on extreme values which is often limited.

As per the search criterion, the incorporation of the EVT in the estimation of VaR started with Pownall & Koedijk (1999). According to Ho et al. (2000), the estimation of VaR using EVT focuses on modeling the tail of the return distribution rather than getting the tail as an outcome of modeling the entire density function which is a primary objective of other approaches for estimating VaR like the RiskMetrics, Historical Simulation and the Monte Carlo simulation methods where the obtained VaR is a threshold value and does not give precise magnitude for the possible losses occurring further in the tail of the distribution.

McNeil and Frey (2000) combined pseudo-maximum-likelihood fitting of GARCH models to estimate the current volatility with EVT for estimating the tail of the innovation distribution of the GARCH model. The EVT-based models are being widely used because they are based on sound statistical theory and provide a parametric form for the tail distribution (McNeil & Frey, 2000).

Thereafter, many works combined EVT with the GARCH model like those of Byström (2004), Huang et al. (2011), and Yi et al. (2014) among others. This combination proved to yield more precise results compared to the regular EVT model sometimes and other VaR estimation methods more often.

According to Table 1.6, the first article issued on the estimation of VaR using EVT was published in 1999 in the "Journal of International Money and Finance" under the title "Capturing downside risk in financial markets: The case of the Asian crisis" by Pownall and Koedijk. Between 1999 and 2004 the average number of articles published was 3 articles per year and increased to an average of 8 articles per year between 2005 and 2008 after the issuance of Basel II in June 2004. Then the average of 10 articles between 2009 and 2010 increased rapidly to 15 articles per year between 2009 and 2017, which can be referred to the issuance of Basel III in July 2010 that was released in the wake of the collapse of Lehman Brothers and the Global financial crisis in 2008.

Year	Number of Articles	Year	Number of Articles
1999	1	2009	7
2000	4	2010	12
2001	3	2011	10
2002	2	2012	14
2003	5	2013	12
2004	3	2014	23
2005	7	2015	15
2006	10	2016	14
2007	7	2017	20
2008	9		
	Total	**178**	

Table 1.6. Distribution of articles on VaR estimation by EVT per year of publication.

1.4.3. Monte Carlo Simulation Method

The Monte Carlo Simulation (MCS) is one of the three basic methodologies for estimating VaR. It is indeed very similar to the HS methodology with the main difference being the construction of the distribution of the potential future portfolio profits and losses. This is explained in detail by Linsmeier and Pearson (2000). The MCS is similar to the HS in some aspects, however, the main difference is that in the former a statistical distribution is assumed to capture or approximate the changes in the market factors which is not applicable in the HS which is non-parametric. Then the hypothetical changes in market factors are generated based on the assumed distribution along with a pseudo-random number generator. Then the hypothetical profit and loss portfolios and their corresponding distribution are constructed based on the market factors hypothetical changes. This distribution is then used to determine the VaR of the portfolio (Linsmeier & Pearson, 2000).

According to these authors, one of the major drawbacks of the MCS in estimating VaR is the flexibility for choosing the statistical distribution to use for market factors where the system designer might make some poor choices and the chosen distribution might not then be reflecting the actual distribution of market factors. McNeil et al. (2015) also mentioned some of the weaknesses of the MCS which mainly constitute the computational cost especially for large portfolios and any results obtained will be only as good as the chosen model used for the distribution of returns. However, MCS proved its dominance as a numerical method for defining market models due to its flexibility that makes it applicable and adaptable to different market conditions. Although MCS can be an expensive method for the estimation of VaR, there are certain methods for obtaining a fast MCS like variance-reduction techniques for evaluating tail probabilities and quantiles (McNeil et al., 2015), or the use of importance sampling, for instance, in Hseih et al. (2014).

Finally, MCS is considered a benchmark to which new models and their performance are compared to determine their accuracy.

As per Table 1.7, the increasing trend of the evolution of publications on MCS, especially after 2008, shows the extensive use of this approach in the estimation of VaR.

According to the search criterion, the first article on MCS with VaR estimation was published in the "Journal of Portfolio Management" under the title "Deterministic simulation for risk management: Quasi-Monte Carlo beats Monte Carlo for value at risk" by Papageorgiou and Paskov (1999). As it can be seen in Table 1.7, the number of publications on the estimation of VaR using MCS was fluctuating between 1999 and 2008 with an average of 5 articles per year. However, after 2008 the number of articles increased from 7 articles in 2008 to 12 articles in 2009. From 2009 and onwards the number of publications witnessed a significant increase with an average of 13 articles per year between 2009 and 2017.

Year	Number of Articles	Year	Number of Articles
1999	1	2009	12
2000	7	2010	10
2001	0	2011	9
2002	3	2012	10
2003	2	2013	9
2004	6	2014	21
2005	4	2015	19
2006	6	2016	14
2007	7	2017	13
2008	7		
Total Number of Articles 160			

Table 1.7. Distribution of articles on VaR estimation using MCS method per year.

1.4.4. Historical Simulation Method

Historical Simulation (HS) is another traditional approach for estimating VaR. It consists of constructing the distribution of the potential future data based on past historical data and then estimating the quantile of the distribution to get the VaR (Linsmeier & Pearson, 2000). The major drawback of the historical simulation lies within the assumption that history repeats itself.

Dockery and Efentakis (2008) compared several models for estimating VaR and found out that the EWMA model outperforms the rest of the compared models which include GARCH and HS. Miletic and Miletic (2015) investigate the performance of the VaR models in the middle of the global financial crisis of 2008 in selected Central and Eastern European emerging capital markets. After the back-testing analysis, they concluded that GARCH-type models outperform HS and RiskMetrics models in measuring VaR (Miletic & Miletic, 2015).

One of the main advantages of HS is that it does not impose any assumptions on the shape of the distribution of the risk factor which affects the value of the portfolio, and this is what makes this approach to a certain point perform better than some models which assume normality of returns (Pritsker, 2006). The HS approach is widely used in practice due to its easy implementation which requires no previous assumptions regarding the distribution of returns. On the other hand, the principal pitfall of HS is the assumption that the risk factors, and consequently, the historically simulated returns, are independently and identically distributed (i.i.d.) through time (Pritsker, 2006).

According to Table 1.8, the first article on HS in the estimation of VaR was published in the "Journal of Derivatives" by Taylor in 1999 under the title "A quantile regression approach to estimating the distribution of multi-period returns". Between 1999 and 2004 the average number of articles published on HS amounted to 2 articles per year and then increased significantly to an average of 8 articles per year between 2005 and 2008. In 2006, the number of articles reached 10 articles for the first time, and this can be referred to the release of Basel II in June 2006. Following a slightly increasing trend, the average number of publications on HS increased to 10 articles per year between 2009 and 2017.

Year	Number of Articles	Year	Number of Articles
1999	1	2009	10
2000	3	2010	14
2001	3	2011	8
2002	1	2012	9
2003	1	2013	10
2004	4	2014	15
2005	7	2015	12
2006	10	2016	8
2007	5	2017	12
2008	9		
Total number of articles 142			

Table 1.8. Distribution of articles on VaR estimation using HS method per year.

1.4.5. Variance Covariance Method

The Variance-Covariance (var-cov) approach for estimating VaR, also known as the delta-normal approach, normal linear or Linear VaR, is a widely used technique for VaR estimation. According to Resti and Sironi (2007), the var-cov approach is one of the easiest and most widespread models for estimating VaR. Linsmeier and Pearson (2000) consider that "the delta-normal approach is based on the assumption that the underlying market factors have a multivariate normal distribution". After that, the distribution of the mark-to-market profits and losses is determined to be also normal. After the distribution of the profits and losses is determined, mathematical properties of the normal distribution are used to determine the quantile of the distribution, that is to say, the VaR.

The major drawback of this approach is the assumption of normality where, in the real world, the distribution of actual data is far from the normal distribution. The effect of non-normality in returns and market capitalization of stock portfolios is studied extensively due to its importance (Sinha & Agnihotri, 2015).

The var-cov method sometimes underestimates the risk due to its assumptions, and another major drawback is that it was not found appropriate for asymmetric distributions (Gençay & Selçuk, 2004).

As per Table 1.9, the first two articles on the estimation of VaR using the var-cov method were published in 1999 in the "European Financial Management" and "Journal of Derivatives" by Peterson and Stapleton, and Taylor, respectively. It is clear that the number of articles published afterwards was fluctuating throughout the period between 1999 and 2008 reaching a maximum of 7 articles in 2006 which could be referred to the issuance of Basel II in June 2006, with an average of 4 articles per year in the same period. However, the number of articles published increased significantly from 2 in 2008 to 9 in 2009 and this can be referred to the global financial crisis of 2008. From 2009 and onwards the number of articles followed an increasing trend with a minimum of 7 articles and a maximum of 15 and an average of 11 articles per year between 2009 and 2017.

Year	Number of Articles	Year	Number of Articles
1999	2	2009	9
2000	6	2010	13
2001	3	2011	7
2002	2	2012	10
2003	1	2013	7
2004	3	2014	13
2005	5	2015	15
2006	7	2016	8
2007	6	2017	15
2008	2		
Total number of articles 134			

Table 1.9. Distribution of articles on VaR Estimation using V-C method per year.

1.4.6. Copulas

According to the search results obtained in this work, the use of Copulas in estimating VaR started in 2001 with Cherubini and Luciano (2001). Cherubini and Luciano (2001) evaluate tail probabilities and market risk trade-offs at a given confidence level using copula functions while dropping the widely used assumption of joint normality of returns. Copulas were introduced into the field of finance and economics in order to calculate the correlations between the returns of assets which are in reality, and contrary to assumptions, not normally distributed making the use of correlation coefficients insufficient.

According to Cherubini and Luciano (2001), although copulas provide a useful tool to represent any joint distribution function when computed at the marginal distribution values, however, contrary to the traditional joint distributions they separate the effect of the marginal distribution from the one of association or dependence between returns.

More specifications for copula models with time-varying dependence structures were brought together by Manner and Reznikova (2012)who compared different copula models and their applicability in different cases. One of their important and useful conclusions is that symmetric copulas that do not allow for tail dependence offer the best fit when considering time-varying dependence parameters contrary to the static case in which asymmetric copulas were found to be more appropriate.

As shown in Table 1.10, the estimation of VaR using copulas started in 2001, with the work of Cherubini and Luciano (2001) that was published in the "Economic Notes" under the title "Value-at-risk trade-off and capital allocation with copulas". However, till 2008, the number of articles published on this topic was considerably low with an average of one article per year between 2001 and 2008. In 2009, the number of publications increased rapidly to 9, reaching its peak in 2017 with 16 articles and an average of 11 articles per year between 2009 and 2017.

Year	Number of Articles	Year	Number of Articles
2001	1	2010	7
2002	0	2011	11
2003	1	2012	5
2004	2	2013	14
2005	1	2014	13
2006	2	2015	14
2007	1	2016	10
2008	1	2017	16
2009	9		
	Total number of articles	**108**	

Table 1.10. Distribution of articles on VaR Estimation using Copulas per year.

1.4.7. Generalized Pareto Distribution

The Generalized Pareto Distribution (GPD) was first used in VaR estimation in 2003. As per the search criterion adopted in this chapter, three articles were published in 2003, with Bali (2003), Vinod (2003) and Bali and Neftci (2003). Bali (2003) was trying to determine the type of asymptotic distribution for the extreme changes in U.S. Treasury yields. Bali (2003) wanted to find the correct limit distributions for maxima and minima and he used the likelihood ratio test between the Fréchet and Gumbel distributions as well as between the Pareto and exponential distributions. Its results indicated the absence of normality in the distribution of the extremes which is why the fat tailed Fréchet and Pareto distributions were more accepted than the thin tailed Gumbel and the exponential distribution. Using the EVT along with the fat tailed distributions – Fréchet and Pareto– Bali (2003) provided a more accurate approach for estimating VaR after proving the superiority of this approach over the traditional approach using normal distribution.

Vinod (2003) discusses the persistence of corruption, and its contribution to Banking distress and rapid transmission across international stock and currency markets. He explains the role of VaR and corruption in discouraging the foreign direct investment (FDI). Since investors often consider worst-case scenarios by VaR methods, Vinod (2003) shows that worst-case corruption costs can be substantial, and since VaR methods cannot cancel losses in one corrupt country with gains in another, due to adverse cross-country relations. This author suggests that both corruption and VaR together discourage direct investment in corrupt developing countries.

Bali and Neftci (2003) were basically studying the extreme interest rate movements in the U.S. federal funds market using daily observations from mid 1950s till the end of 2000. They analyzed the fluctuations of the short-term interest rates and test the significance of the time-varying paths of the mean and volatility of the extremes. Their findings indicated the existence of volatility clustering in the standard deviation of extremes and a positive relationship between the level and volatility of extremes. They proposed a conditional extreme value approach to calculate VaR using a GPD of specified location and scale parameters based on past information. The results of the estimated VaR calculated using tails of GPD were significantly more precise than those obtained by the normal distribution, reflecting the fact that the tails of the empirical distribution were thicker than the tails of the normal distribution.

In general, the GPD model is typically used with the EVT when estimating VaR and depending on the shape and scale parameters, the nature of the distribution is determined according to the value of the shape parameter of the GPD. If it was greater than zero then the distribution corresponds to a heavy tailed distribution like Pareto, Student-*t* among others, and it is also known as the Fréchet case. If the shape parameter of the GPD was equal to zero then the distribution corresponds to distributions like normal, exponential, Gamma and others, this case is known as the Gumbel case. Finally, when the shape parameter of GPD is negative, the distribution corresponds to short tailed distributions with finite endpoint like the uniform distribution, however this case is the least relevant for finance. Brooks et al. (2005) proposed an extreme value semi-nonparametric approach using the GARCH (1,1) model and a bootstrapping approach based on a combination of a GPD and the empirical distribution of returns. They proved that their approach, which separately models the center and the tails, performs better than the competing approaches which constitute a large number of nonparametric tail index estimators.

Different methods for estimating the quantiles of the GPD were introduced by Jocković (2012) and applied to estimate the VaR parameter. These methods include the method of moments (MOM), method of probability weighted moments (PWM), and EPM (elemental percentile method). The mentioned methods were successfully applied on VaR estimates.

Using GPD for estimating VaR for measuring unconditional market risk is more appropriate than the RiskMetrics and the GARCH (1,1) especially when there are fat tails leading to tail-related risk (Lee, 2016).

Using the GPD in estimating VaR started in 2003 as shown in Table 1.11, with three articles published in “Journal of Business”, “Journal of Asian Economics” and “Journal of Empirical Finance”. The average number of articles published on GPD amounted to 3 articles per year between 2003 and 2008. In 2010, a significant increase in the number of publications was witnessed with 6 articles. In general, the trend of evolution of articles on this approach was not stable with ups and downs throughout the whole period of the study. But the average number of articles doubled from 3 between 2003 and 2008 to 6 articles per year between 2010 and 2017.

Year	Number of Articles	Year	Number of Articles
2003	3	2011	2
2004	1	2012	9
2005	1	2013	3
2006	4	2014	3
2007	4	2015	10
2008	4	2016	5
2009	0	2017	7
2010	6		
	Total number of articles	**62**	

Table 1.11. Distribution of articles on VaR Estimation using GPD per year.

1.4.8. Moving Average Models

The family of Moving Average models was introduced to the estimation of VaR in 1997. This family includes a group of models some of which were used in estimating VaR like Autoregressive Fractionally Integrated Moving Average (ARFIMA), Autoregressive Integrated Moving Average (ARIMA), Equally Weighted Moving Average (EQWMA), Exponentially Weighted moving average (EWMA), Autoregressive Moving Average (ARMA) and the simple Moving Average (MA).

The ARMA is most commonly used among the above models with 26 articles followed by the EWMA with 20 articles then the MA with 7 articles followed by ARFIMA model with 5 articles and finally the ARIMA and the EQWMA with 4 articles each.

As per the search criterion followed in this work, the first paper that used ARMA in estimating VaR was published in 2007 in the journal "Annals of Economics and Finance" under the title "Empirical analyses of industry stock index return distributions for the Taiwan stock exchange". In this article, Rachev et al. (2007)study the return distributions of 22 industry stock indexes on the Taiwan Stock Exchange under the conditional homoscedastic independent, identically distributed and conditional heteroskedastic GARCH models, considering two hypotheses, the Gaussian and the Paretian distributions. They proved that the performance of the stable Paretian distribution is much better than that of the Gaussian distribution and they also provided evidence that the stable ARMA-GARCH model outperforms the normal ARMA-GARCH model.

McAleer et al. (2009)developed a constant correlation vector ARMA-asymmetric GARCH model. In general, it can be noticed that ARMA models are often used in combination with GARCH models like in Sun et al. (2009), Zhao and Wu (2016), Uylangco and Li (2016) and Huang et al. (2015) among many others.

In most of the previously mentioned articles, the proposed parametric models combining ARMA and GARCH models proved their superiority to other models.

Apparently, the EWMA model was first used in the estimation of VaR in the late 1980's when RiskMetrics variance model was first established. The EWMA model was formally used in RiskMetrics when J.P. Morgan (1996) launched the RiskMetrics Technical Document. Taking into consideration that its use started way before in the early 1990's. The first article on VaR estimation which included the EWMA model compared to other common volatility forecasting methods was published in 1997 in the "Journal of Derivatives", according to the search criteria results applied in this chapter.

Moosa and Bollen (2002)focused mainly on the volatility measure which lies behind the parametric models used in VaR estimation where any bias in the measure of volatility will directly affect the estimated VaR. In order to measure the bias in the volatility, the authors used the "realized volatility" as a benchmark to which they compared the performance of models based on parametric and historical approaches for estimating VaR.. The results showed that VaR estimates which were based on the EWMA were unbiased and reliable on longer time periods, however, on short time periods, the results remained unbiased in condition that the decay factor is relatively small, whereas VaR estimates that were based on GARCH (1, 1) appeared to be biased.

Guermat and Harris (2002) showed that the EWMA can be obtained as a special case from a more general procedure which allows for time-variation in the higher moments of the returns in addition to the time-variation in the variance of the returns, this procedure is the EWML (exponentially weighted maximum likelihood)..

González-Rivera et al. (2004) compared and analyzed the performance of several volatility models for stock returns using two types of loss functions, the economic loss functions and statistical loss functions: the goodness-of-fit for a VaR calculation and the average predictive likelihood. Briefly, the results of this article show that, for option loss function, the EWMA model performed as well as other sophisticated specifications which is a good result as the simple moving average and the EWMA do not require statistical parameter calculation and thus their implementation is easy. However, when using the goodness of fit function, the EWMA proved to be the worst performer among the compared models which include GARCH (linear and non-linear processes).

It can be seen in Table 1.12 that the first article on the use of Moving Average models in the estimation of VaR was published in the "Journal of Derivatives" by Alexander and Leigh (1997) under the title "On the covariance matrices used in value at risk models". The average number of articles published on MA models in the estimation of VaR was 2 articles per year between 1997 and 2008. Between 2008 and 2017, the average number of articles increased to 5 per year.

Year	Number of Articles	Year	Number of Articles
1997	1	2008	1
1998	0	2009	7
1999	0	2010	1
2000	0	2011	5
2001	1	2012	3
2002	2	2013	3
2003	0	2014	4
2004	3	2015	9
2005	2	2016	8
2006	2	2017	6
2007	1		
Total number of articles		**59**	

Table 1.12. Distribution of articles on Value-at-Risk Estimation using Moving Average models per year.

1.4.9. Quantile Regression

The Quantile Regression is listed under the non-parametric approaches for estimating VaR (Lejeune & Sarda, 1988). According to the adopted search criteria, the use of quantile regression in estimating VaR started with Taylor (1999) who used this approach to estimate the conditional probability of multi-period returns. Taylor (1999) applies quantile regression to the historical returns from a range of different holding periods and produces quantile models that are functions of the one-step-ahead volatility forecast and the length of the holding period, as suggested by theoretically derived variance expressions. The quantile regression method offers good results, and it could be an important alternative for estimating the tail of multi-period returns compared to the traditional methods (including GARCH and EWMA)(Taylor, 1999).

Chernozhukov and Fernández-Val (2011) use the Quantile Regression along with the EVT and apply the extremal Quantile Regression on the conditional VaR. They applied their approach to a bank's portfolio and concluded that unlike unconditional extremal quantiles, conditional extremal quantiles can be useful for stress testing and analyzing the impact of adverse systemic events on the bank's performance.

Taylor (2008) proposes an exponentially weighted quantile regression to estimate the VaR and the Expected Shortfall. The motivation behind the exponentially weighted Quantile Regression was the wide use of the exponential smoothing in volatility forecasting. The non-parametric approach showed to be equivalent to the exponential smoothing of the cumulative distribution function and a comparison with a set of widely used approaches gave encouraging results. (Taylor, 2008).

As shown in Table 1.13, the first article on the estimation of VaR using Quantile Regression was published in 1999 in the "Journal of Derivatives". The number of articles published on this topic is lower than other models where the average number of articles published between 1999 and 2010 amounted to 3 articles noting that between 2000 and 2005 (inclusive) no articles were published on this topic. Between 2011 and 2017 the average number of articles published per year doubled to 6.

Year	Number of Articles	Year	Number of Articles
1999	1	2009	2
2000	0	2010	4
2001	0	2011	7
2002	0	2012	7
2003	0	2013	5
2004	0	2014	4
2005	0	2015	5
2006	3	2016	4
2007	4	2017	9
2008	3		
Total number of articles 58			

Table 1.13. Distribution of articles on VaR Estimation using Quantile Regression models per year.

1.4.10. RiskMetrics

RiskMetrics is a variance model whose roots go back to the late 1980's. Its methodology was launched by J.P. Morgan in 1992, then in RiskMetrics Technical Document was revised and released in 1996. The Technical Document contains techniques and data sets to measure market risk and in particular to calculate the VaR (Morgan, 1996). According to J.P. Morgan (1996) "RiskMetrics is a set of methodologies and data for measuring market risk". The use of RiskMetrics in estimating VaR started in 1997. RiskMetrics was used as a benchmark for testing and comparing the VaR models developed over the past two decades.

Due to its simplicity and ease of implementation, RiskMetrics has been widely used in estimating VaR forecasts and it is one of the most popular models in this area. Oanea and Anghelacheb (2015) emphasized the performance of the RiskMetrics model in forecasting the high volatility during the financial crisis of 2008. The authors concluded that RiskMetrics was sometimes underestimating the decay factor and at other times it was overestimating the decay factor depending on the index under study and the loss function employed.

González-Rivera et al. (2007) investigate the implications of choosing different loss functions in studying VaR, where in VaR modeling in particular, the loss function is well defined. However, in some papers that study VaR, the loss function in estimating and forecasting is still chosen, so this paper investigates the implications of this practice by focusing on RiskMetrics model and the findings show that the RiskMetrics model for the equity markets overestimates the decay factor for a 10-day horizon.

Table 1.14 shows that the first article on estimating VaR using RiskMetrics was published in 1997 in the "Journal of Derivatives" by Marshall and Siegel. It is clear that there is not a huge fluctuation in the number of articles published on this topic between 1997 and 2017, where the average number of articles remained low with a maximum of 6 articles in 2010. However, the average number of articles between 1997 and 2017 published on this topic amounted to only 3 articles per year and the trend of evolution of this model was neither increasing nor decreasing as it was always fluctuating up and down from one year to another.

Year	Number of Articles	Year	Number of Articles
1997	1	2007	3
1998	0	2008	1
1999	2	2009	3
2000	2	2010	6
2001	0	2011	3
2002	0	2012	4
2003	1	2013	2
2004	5	2014	2
2005	1	2015	5
2006	1	2016	2
Total number of articles	**44**		

Table 1.14. Distribution of articles on Value-at-Risk Estimation using RiskMetrics per year.

1.5. Conclusions

This chapter presented a systematic review of the literature on VaR models used in articles published in journals between 1996 and 2017 recognizing the most widely used models and their growth trends and whether it was affected by the crisis of 2008. Some models appeared to be more used than others, this can be referred to their method of implementation or degree of precision, advantages and disadvantages and the need for developing better VaR models from the already existing ones trying to increase the accuracy of VaR estimations.

The ARCH/GARCH models are the most widely used models in estimating VaR and this is supported by statistical evidence as 335 articles were published focusing on ARCH/GARCH models between 1996 and 2017. ARCH/GARCH models treat heteroskedasticity in financial time series as a variance to be modeled. The evolution trend of articles on ARCH/GARCH models in the estimation of VaR witnessed the highest increase among all models throughout the period as it increased from 16% from [1996, 2007] to 24% in [2008, 2017] highlighting the growing interest of engaging these models in estimating VaR

The Extreme Value Theory is another extensively used model in the estimation of VaR. It constitutes of two principal groups of models, the Block Maxima Models (BMM) and the Peaks Over Threshold (POT). McNeil and Frey (2000) combined the POT model with GARCH-type models and obtained relatively accurate VaR estimations. In general, EVT has proved its superiority to other models like RiskMetrics when applied in the Asian market during periods of financial turmoil (Pownall & Koedijk, 1999), and Historical Simulation method (Ho et al., 2000) in terms of accuracy and precision (Shayya et al., 2023). The average percentage of articles published on the estimation of VaR using EVT increased from 9% in [1996, 2007] to 12% in [2008, 2017].

The Monte Carlo Simulation is a traditional semi-parametric approach for estimating VaR and it has gained its popularity among practitioners due to its advantages compared to other modeling methodologies despite being considerably an expensive, time-consuming method. As per the findings in this chapter, the Monte Carlo simulation witnessed a slight growth in the number of articles where the average percentage increased by 2% only reaching 9% of the total number of articles in the interval [2008, 2017].

The Historical Simulation method is another traditional approach for the estimation of VaR which is non-parametric. The major drawback of HS is the assumption that history repeats itself. However, one of its advantages is that, contrary to other methodologies, does not make any assumptions regarding the shape of the distribution of risk factors, since empirically, these distributions are usually fat-tailed, and most models take the assumption of normality which is far from reality. The number of articles published on HS increased from an average percentage of articles of 7% in [1996, 2007] to 10% in [2008, 2017].

The Variance-Covariance method is the third traditional methodology for estimating VaR. It is a parametric approach also known as the delta-normal approach. It is based on the assumption that the underlying market factors have a multivariate normal distribution, and this is one of its disadvantages. The average percentage of the number of articles on the Variance-Covariance method was constant at 8%.

Copulas aim to evaluate the tail probabilities and market risk trade-offs at a given confidence level. They were mainly used to calculate the correlations between the returns of assets which are not normally distributed contrary to some model assumptions, and this fact in turn makes the use of correlation coefficients insufficient. The average percentage of articles published on copulas in VaR estimation increased from 2% in [1996, 2007] to 8% in [2008, 2017].

The Generalized Pareto Distribution in estimating VaR proved its superiority in measuring the unconditional market risk on other models like the RiskMetrics and the GARCH and in particular when the data constitutes fat tails that lead to tail-related risk. GPD is used along with the Extreme Value Theory, in fact it is a building block of the latter. The average percentage of articles published on GPD in VaR estimation doubled from 2% in [1996, 2007] to 4% in [2008, 2017].

Moving Average models have also been employed in the estimation of VaR and most of the time they are combined with GARCH-type models. The use of these models in VaR estimation increased slightly from 3% in [1996, 2007] to 4% in [2008, 2017].

Quantile Regressions have been used in the estimation of VaR as another non-parametric approach. The average number of articles that use Quantile Regression in estimating VaR increased from an average of 1% in [1999, 2007] to 4% in [2008, 2017] of the total number of articles published in the respective intervals. This shows the positive trend of growth of publications on this model in VaR estimation.

RiskMetrics model is a widely popular model in measuring market risk and estimating VaR. It constitutes using the EWMA model and is also known for its easy implementation process which is why it was highly employed in practice. The average number of articles on RiskMetric in VaR estimation decreased from 5% in [1996, 2007] to 3% in [2008, 2017]. This decrease can be explained by the fact that in the first interval the model was still new and under trials and studies, and most probably, after the crisis of 2008, other models were the center of attention for practitioners and academics like ARCH/GARCH and EVT.

The most popular VaR models recognized in this chapter, and more precisely, the top five models obtained in this review are the ARCH/GARCH family, Extreme Value Theory, Monte Carlo Simulation, Historical Simulation and Variance-Covariance. These models are applied in Chapter 3 along with other models which can be considered sub-models or special cases of the main models like the Filtered Historical Simulation (FHS), which can be considered as a special case of the HS model as it makes the HS consistent with clustering of large returns. GARCH-EVT is a hybrid model that combines a GARCH filter with the EVT and EWMA is another sub-model applied in Chapter 3 which can be considered as a special case of GARCH (1,1), being the basis of the RiskMetrics model which is a pioneer among parametric models. Lengthy definitions and the derivations of these models are explained in Chapter 2.

References

Alexander, C. O., & Leigh, C. T. (1997). On the covariance matrices used in value at risk models. *Journal of Derivatives*, *4*(3), 50–62. https://doi.org/10.3905/jod.1997.407974

Bali, T. G. (2003). An Extreme Value Approach to Estimating Volatility and Value at Risk*. *The Journal of Business*, *76*(1), 83–108. https://doi.org/10.1086/344669

Bali, T. G., & Neftci, S. N. (2003). Disturbing extremal behavior of spot rate dynamics [Article]. *Journal of Empirical Finance*, *10*(4), 455–477. https://doi.org/10.1016/S0927-5398(02)00070-1

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

Brooks, C., Clare, A. D., Dalle Molle, J. W., & Persand, G. (2005). A comparison of extreme value theory approaches for determining value at risk. *Journal of Empirical Finance*, *12*(2), 339–352. https://doi.org/10.1016/j.jempfin.2004.01.004

Byström, H. N. E. (2004). Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, *13*(2), 133–152. https://doi.org/10.1016/j.irfa.2004.02.003

Chernozhukov, V., & Fernández-Val, I. (2011). Inference for extremal conditional quantile models, with an application to market and birthweight risks. *Review of Economic Studies*, *78*(2), 559–589. https://doi.org/10.1093/restud/rdq020

Cherubini, U., & Luciano, E. (2001). Value-at-risk trade-off and capital allocation with copulas. *Economic Notes*, *30*(2), 235–256. https://doi.org/10.1111/j.0391-5026.2001.00055.x

Cook, D. J., Greengold, N. L., Ellrodt, A. G., & Weingarten, S. R. (1997). The relation between systematic reviews and practice guidelines. *Annals of Internal Medicine*, *127*(3), 210–216. https://doi.org/10.7326/0003-4819-127-3-199708010-00006

Crossan, M. M., & Apaydin, M. (2010). A multi-dimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies*, *47*(6), 1154–1191. https://doi.org/10.1111/j.1467-6486.2009.00880.x

Dickersin, K., Scherer, R., & Lefebvre, C. (1994). Systematic Reviews: Identifying relevant studies for systematic reviews. *BMJ*, *309*(6964), 1286. https://doi.org/10.1136/bmj.309.6964.1286

Dockery, E., & Efentakis, M. (2008). An empirical comparison of alternative models in estimating Value-at-Risk: evidence and application from the LSE. *International Journal of Monetary Economics and Finance*, *1*(2), 201–218. https://doi.org/10.1504/IJMEF.2008.019222

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

Evans, D. (2001). Systematic reviews of nursing research. *Intensive and Critical Care Nursing*, *17*(1), 51–57. https://doi.org/10.1054/iccn.2000.1550

Gençay, R., & Selçuk, F. (2004). Extreme value theory and Value-at-Risk: Relative performance in emerging markets. *International Journal of Forecasting*, *20*(2), 287–303. https://doi.org/10.1016/j.ijforecast.2003.09.005

González-Rivera, G., Lee, T.-H., & Mishra, S. (2004). Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting*, *20*(4), 629–645. https://doi.org/10.1016/j.ijforecast.2003.10.003

González-Rivera, G., Lee, T.-H., & Yoldas, E. (2007). Optimality of the RiskMetrics VaR model. *Finance Research Letters*, *4*(3), 137–145. https://doi.org/10.1016/j.frl.2007.06.001

Guermat, C., & Harris, R. D. F. (2002). Forecasting value at risk allowing for time variation in the variance and kurtosis of portfolio returns. *International Journal of Forecasting*, *18*(3), 409–419. https://doi.org/10.1016/S0169-2070(01)00122-4

Ho, L. C., Burridge, P., Cadle, J., & Theobald, M. (2000). Value-at-risk: Applying the extreme value approach to Asian markets in the recent financial turmoil. *Pacific Basin Finance Journal*, *8*(2), 249–275. https://doi.org/10.1016/S0927-538X(00)00008-1

Hsieh, M.-H., Liao, W.-C., & Chen, C.-L. (2014). A Fast Monte Carlo Algorithm for Estimating Value at Risk and Expected Shortfall. *The Journal of Derivatives*, *22*(2), 50–66. https://doi.org/10.3905/jod.2014.22.2.050

Huang, H.-C., Su, Y.-C., & Tsui, J.-T. (2015). Asymmetric GARCH value-at-risk over MSCI in financial crisis. *International Journal of Economics and Financial Issues*, *5*(2), 390–398. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84979824748&partnerID=40&md5=8d5322cff1fccb141df8507ebac2191d

Huang, S.-C., Chienb, Y.-H., & Wangc, R.-C. (2011). Applying GARCH-EVT-copula models for portfolio value-at-risk on G7 currency markets. *International Research Journal of Finance and Economics*, *74*, 136–151.

Jocković, J. (2012). Quantile estimation for the generalized pareto distribution with application to finance. *Yugoslav Journal of Operations Research*, *22*(2), 297–311. https://doi.org/10.2298/YJOR110308013J

Lee, H. (2016). Semi-parametric method for estimating tail related risk measures in the stock market. *Korean Economic Review*, *32*(2), 295–329. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85009431673&partnerID=40&md5=f627f799d6832f28f0cba51c42dfec69

Lejeune, M. G., & Sarda, P. (1988). Quantile regression: a nonparametric approach. *Computational Statistics and Data Analysis*, *6*(3), 229–239. https://doi.org/10.1016/0167-9473(88)90003-5

Linsmeier, T. J., & Pearson, N. D. (2000). Value at Risk. *Financial Analysts Journal*, *56*(2), 47–63. https://doi.org/10.2469/faj.v56.n2.2343

Manner, H., & Reznikova, O. (2012). A Survey on Time-Varying Copulas: Specification, Simulations, and Application. *Econometric Reviews*, *31*(6), 654–687. https://doi.org/10.1080/07474938.2011.608042

Mazzi, C. (2011). Family business and financial performance: Current state of knowledge and future research challenges. *Journal of Family Business Strategy*, *2*(3), 166–181. https://doi.org/10.1016/j.jfbs.2011.07.001

McAleer, M., Hoti, S., & Chan, F. (2009). Structure and asymptotic theory for multivariate asymmetric conditional volatility. *Econometric Reviews*, *28*(5), 422–440. https://doi.org/10.1080/07474930802467217

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3–4), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8

McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative risk management: Concepts, techniques and tools: Revised edition. In *Quantitative Risk Management: Concepts, Techniques and Tools: Revised Edition*. Princeton University Press. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84937787962&partnerID=40&md5=4546a3521a1b5468f3f969c8d2517e01

Miletic, M., & Miletic, S. (2015). Performance of value at risk models in the midst of the global financial crisis in selected cee emerging capital markets. *Economic Research-Ekonomska Istrazivanja* , *28*(1), 132–166. https://doi.org/10.1080/1331677X.2015.1028243

Moosa, I. A., & Bollen, B. (2002). A benchmark for measuring bias in estimated daily value at risk. *International Review of Financial Analysis*, *11*(1), 85–100. https://doi.org/10.1016/S1057-5219(01)00069-2

Morgan, J. P. (1996). *RiskMetrics-Technical Document* (Fourth Edition). J.P. Morgan/Reuters. https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a

Mulrow, C. D., Cook, D. J., & Davidoff, F. (1997). Systematic reviews: Critical links in the great chain of evidence. *Annals of Internal Medicine*, *126*(5), 389–390. https://doi.org/10.7326/0003-4819-126-5-199703010-00008

Nightingale, A. (2009). A guide to systematic literature reviews. *Surgery*, *27*(9), 381–384. https://doi.org/10.1016/j.mpsur.2009.07.005

Oanea, D.-C., & Anghelache, G. (2015). Value at Risk Prediction: The Failure of RiskMetrics in Preventing Financial Crisis. Evidence from Romanian Capital Market. *Procedia Economics and Finance*, *20*, 433–442. https://doi.org/https://doi.org/10.1016/S2212-5671(15)00094-5

Papageorgiou, A., & Paskov, S. (1999). Deterministic simulation for risk management: Quasi-Monte Carlo beats Monte Carlo for value at risk. *Journal of Portfolio Management*, *25*(SUPPL.), 122–127. https://www.scopus.com/inward/record.uri?eid=2-s2.0-20144385376&partnerID=40&md5=a2fa53f0f82988c2c60b5374d6744856

Peloza, J., & Shang, J. (2011). How can corporate social responsibility activities create value for stakeholders? A systematic review. *Journal of the Academy of Marketing Science*, *39*(1), 117–135. https://doi.org/10.1007/s11747-010-0213-6

Pownall, R. A. J., & Koedijk, K. G. (1999). Capturing downside risk in financial markets: The case of the Asian crisis. *Journal of International Money and Finance*, *18*(6), 853–870. https://doi.org/10.1016/S0261-5606(99)00040-6

Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking and Finance*, *30*(2), 561–582. https://doi.org/10.1016/j.jbankfin.2005.04.013

Rachev, S. T., Emeritus, Stoyanov, S. V, Wu, C., & Fabozzi, F. J. (2007). Empirical analyses of industry stock index return distributions for the Taiwan stock exchange. *Annals of Economics and Finance*, *8*(1), 21–31. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84990219474&partnerID=40&md5=b7f5b4ba8a8cd1e6c347c172d34183ac

Resti, A., & Sironi, A. (2007). Risk Management and Shareholders' Value in Banking. From Risk Measurement Models to Capital Allocation Policies. *Risk Management and Shareholders' Value in Banking*.

Riordan, C. M. (2000). Relational demography within groups: Past developments, contradictions, and new directions. In *Research in Personnel and Human Resources Management* (Vol. 19, pp. 131–173). JAI Press. https://doi.org/10.1016/S0742-7301(00)19005-X

Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research*. Sage Publications, Inc.

Shayya, R., Sorrosal-Forradellas, M. T., & Terceño, A. (2023). Value-at-risk models: a systematic review of the literature. *Journal of Risk*. https://doi.org/10.21314/JOR.2022.053

Sinha, P., & Agnihotri, S. (2015). Impact of non-normal return and market capitalization on estimation of VaR. *Journal of Indian Business Research*, *7*(3), 222–242. https://doi.org/10.1108/JIBR-12-2014-0090

Sun, W., Rachev, S., & Fabozzi, F. J. (2009). A new approach for using lévy processes for determining high-frequency value-at-risk predictions. *European Financial Management*, *15*(2), 340–361. https://doi.org/10.1111/j.1468-036X.2008.00467.x

Taylor, J. W. (1999). A quantile regression approach to estimating the distribution of multiperiod returns. *Journal of Derivatives*, *7*(1), 64–78. https://doi.org/10.3905/jod.1999.319106

Taylor, J. W. (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*, *6*(3), 382–406. https://doi.org/10.1093/jjfinec/nbn007

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, *14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

Uylangco, K., & Li, S. (2016). An evaluation of the effectiveness of Value-at-Risk (VaR) models for Australian banks under Basel III. *Australian Journal of Management*, *41*(4), 699–718. https://doi.org/10.1177/0312896214557837

Vinod, H. D. (2003). Open economy and financial burden of corruption: theory and application to Asia [Article]. *Journal of Asian Economics*, *13*(6), 873–890. https://doi.org/10.1016/S1049-0078(02)00188-4

Yi, Y., Feng, X., & Huang, Z. (2014). Estimation of extreme value-at-risk: An EVT approach for quantile GARCH model. *Economics Letters*, *124*(3), 378–381. https://doi.org/10.1016/j.econlet.2014.06.028

Zhao D., & Wu T. (2016). Alarming of exchange rate crisis: A risk management approach. *Risk Governance and Control: Financial Markets and Institutions*, *6*(2), 79–88.

Chapter 2

A review of the most applied models on Value-at-Risk

2.1. Introduction

In 1996, the amendment of Basel I incorporated market risk in the minimum capital that should be attained by banks and financial institutions to avoid failures and control financial activities based on a risk management aspect. Since then, firms as well as academics and professionals have been trying to find the best methods and frameworks to produce reliable forecasts. VaR, the key tool in forecasting the maximum amount of loss that might be incurred in case of sudden shocks, was recommended by the Basel Committee on Banking Supervision to measure the possible risk and account for the capital required to provide a cushion for firms against possible severe losses. This is the main reason behind the very wide and rich literature on VaR and the models that have been used for its estimation.

This chapter presents some of the most widely used models for estimating VaR that have been developed, mentioned, discussed or applied according to the review of the literature in Chapter 1. Based on these results, the most widely used models in estimating VaR, between 1996 and 2017, are the ARCH/GARCH-type models, Extreme Value Theory (EVT), Monte Carlo Simulation (MCS), Historical Simulation (HS) and Variance-Covariance (var-cov) method, listed as per their order of popularity in the literature. Some of these models are among the classical building models of VaR like the Monte Carlo Simulation, Historical Simulation and Variance-Covariance, while the others are among the modern models practically developed and improved attempting to obtain more accurate VaR forecasts than those produced by the classical methodologies. The main issue addressed by the variety of models is their ability to truly reflect the distribution of the data taking into account the stochastic variance of log returns and thus working on the conditional volatility rather than only estimating the unconditional volatility that neglects the time-variation of variance and ignores important aspects when estimating VaR. Moreover, it is noticed that in the literature most models are based on the assumption of normality of log returns. This point can be viewed as a limitation in the estimation of VaR, as empirical evidence shows that log returns cannot be assumed to be normally distributed as they are usually described to follow leptokurtic distributions admitting fatter tails than the normal distribution in addition to some level of skewness.

In the following, the above models are defined and described, including the main idea and approach, as well as a brief review of literature. Moreover, some models can be categorized as sub-models or branches of the main models, like the GARCH (1,1) model, with normally distributed residuals and Student-*t* distributed residuals, listed under the ARCH/GARCH-type models, the GARCH-EVT approach of McNeil and Frey (2000) classified under the EVT approach, the Filtered Historical Simulation (FHS) of Barone-Adesi et al. (1999) related to HS approach, and the Exponentially Weighted Moving Average (EWMA) of the RiskMetrics^{TM} under the var-cov approach.

2.2. ARCH/GARCH Models

Most VaR models are based on the assumption that returns are independent identically distributed (i.i.d), however, if this assumption was true then the mean and variance of returns are to be considered constant over time. On the other hand, it is well known that the variance of returns is not constant, and it varies with time since if the variance should be treated as a constant then yesterday's return should not affect today's return, but this is not the case when it comes to practice and empirical evidence. For instance, it is noted that in general, "large changes tend to be followed by large changes – of either sign – and small changes tend to be followed by small changes…" (Mandelbrot, 1963). Moreover, volatility is time varying and periods of high and low volatility cluster together (Bollerslev, 1986). In fact, as per Kuester at al. (2006), financial returns exhibit "non-standard" statistical properties as they are neither i.i.d nor normally distributed. This is reflected by three widely reported stylized facts about financial returns: (i) volatility clustering which is shown by the high autocorrelation of absolute and squared returns, (ii) substantial kurtosis, i.e., the density of unconditional return distribution is more peaked and exhibits fatter tails than the normal distribution and (iii) mild skewness of returns possibly of a time-varying nature (Kuester et al., 2006). More evidence on the stochastic volatility and clustering in addition to several other stylized facts on financial asset returns are listed and discussed with empirical evidence in Cont (2001). In general, the volatilities of returns are positively correlated which is known as volatility clustering.

The least squares model is a basic tool in applied econometrics, it minimizes the sum of the squared errors (difference between observed and predicted values) of a model. However, the basic assumption of the least squares model is that the expected value of all error terms, when squared, is equal at all points in the data which is known as homoskedasticity. Heteroskedasticity, which is the main focus of ARCH and GARCH models, can be defined as the case when the variances of error terms of the data are not equal, i.e. when error terms may be expected to be larger at some points or ranges of the data than others (Engle, 2001). ARCH and GARCH models consider heteroskedasticity as a variance to be modeled instead of a problem to be corrected, so they correct the deficiencies of least squares and compute predictions of the variance of each error term (Engle, 2001).

The GARCH(p, q) process

The ARCH (Autoregressive Conditional Heteroskedasticity) process was first introduced in Engle (1982) allowing for conditional variance to vary over time as a function of past errors while keeping the unconditional variance constant.

The GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) was later proposed by Bollerslev (1986) as an extension of the ARCH class of models. The GARCH(p, q) process, where "p" stands for the order of the polynomial which refers to the autoregressive term and "q" is the order of the polynomial which refers to the Moving Average term of the model, improves ARCH(q) because it allows lagged conditional variances. GARCH models allow for a longer memory and more flexible lag structure to avoid problems with negative variance parameter estimates, since in practice a long lag is often needed in the conditional variance equation.

A conditionally heteroscedastic series of asset returns r_t can be modelled as follows,

$$r_t = \mu_t + z_t$$

Then, as per the definition of the GARCH(p, q) process given by Bollerslev (1986) it follows,

$$z_t = \sigma_t \varepsilon_t | \psi_{t-1} \sim N(0, \sigma_t^2); \text{ and } \varepsilon_t \sim N(0,1)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i z_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

where r_t denotes the return at time t composed of a deterministic part μ_t (representing the mean) and a random residual term z_t, σ_t is the conditional variance of z_t, ε_t is an i.i.d. random variable, and ψ_{t-1} is the information set through time $t-1$. Moreover, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for all $i = 1, \ldots, q$ and $j = 1, \ldots, p$ are the parameters of the model.

The GARCH(1, 1) process

Normal residuals

The GARCH (1,1) process is the simplest yet the most widely used GARCH process in econometrics and financial applications in particular.

According to Bollerslev (1986), the GARCH(1,1) process has the following form:

$z_t = \sigma_t \varepsilon_t | \psi_{t-1}$, $\varepsilon_t \sim N(0,1)$ and $z_t \sim N(0, \sigma_t^2)$. Then,

$$\sigma_{t+1}^2 = \alpha_0 + \alpha {z_t}^2 + \beta \sigma_t^2 ,$$

where $\alpha_0 > 0$, $\alpha + \beta < 1$ and $\alpha > 0, \beta > 0$.

Assuming residuals are normally distributed, the $(1-q)\%\ VaR$ in the GARCH (1,1) process is

$$VaR_{t,q} = \mu_t + F_\varepsilon^{-1}(q) \sigma_t,$$

where F_ε^{-1} is the inverse of standard normal density function.

Student-*t* residuals

When comparing the actual distribution of ε_t to the normal distribution, the most important thing that is usually noticed is that the former has fatter tails and more pronounced peaks than the normal distribution (Christoffersen, 2011). The Student-*t* distribution is perhaps the most commonly used instead of the normal distribution in order to better reflect the actual distribution of residuals due to its shape, which is closer to the actual distribution, and simplicity as it only depends on one parameter (the degree of freedom). In this study, GARCH (1,1) model is also used with Student-*t* residuals for estimating VaR under the Monte Carlo Simulation which will be detailed in the next section. The Student-*t* distribution has the distribution function

$$f_t(x; d) = \frac{\Gamma((d+1)/2)}{\Gamma(d/2)\sqrt{d\pi}} (1 + x^2/d)^{-(1+d)/2} \text{ for } d > 0,$$

where x is a random variable, $\Gamma(\cdot)$ refers to the gamma function, and d is the degree of freedom. The first two moments, the expected value and variance of x, are given by

$$E(x) = 0 \text{ and } Var(x) = d/(d-2), \text{ respectively.}$$

In particular, the standardized Student-*t* distribution is employed for the tails because of the unified variance modeled under GARCH. So, y, a variable obtained by standardizing x, can be defined as

$$y = \frac{x - E(x)}{\sqrt{Var(x)}} = \frac{x}{\sqrt{d/(d-2)}}$$

Consequently, the density function of the standardized Student-*t* distribution, call it $\tilde{t}(d)$, is given as

$$f_{\tilde{t}}(y, d) = C(d)(1 + z^2/(d-2))^{-(1+d)/2}, \text{ for } d > 2$$

Where $C(d) = \frac{\Gamma((d+1)/2)}{\Gamma(d/2)\sqrt{\pi(d-2)}}$.

It is important to note that the mean and variance of the standardized Student-*t* distribution are equal to zero and 1, respectively, and that the parameter *d* should be greater than two for the distribution to be well defined. It is also worth noting that the significance of the $\tilde{t}(d)$ distribution is in its fat tails feature which is driven by the power function of *d*, producing fatter tails than those obtained from the exponential power of the normal distribution[3]. The first four moments of the $\tilde{t}(d)$ are the mean, variance, skewness (ζ), and excess kurtosis (EK) and they are given by

$$E(y) = 0$$

$$Var(y) = E[(y - E[y])^2] = 1$$

$$\zeta = E[y^3]/\left(\sqrt{Var(y)}\right)^3$$

$$EK = E[y^4]/\left(\sqrt{Var(y)}\right)^4 - 3 = 6/(d-4)$$

For the EK to be well defined, d should be greater than 4. However, for very large values of *d*, the $\tilde{t}(d)$ distribution converges to the standard normal distribution as the EK in this case will have a value of zero ($EK = 0$ refers to the standard normal distribution).

In this case, it is assumed that the series of log returns $\{r_t\}$ is given such that

$$r_t = \mu_t + z_t \text{ and } z_t = \sigma_t \varepsilon_t | \psi_{t-1},$$

however, $\varepsilon_t \sim \tilde{t}(0,1,d)$ with $d > 2$, denotes the standardized Student-*t* distribution, with d degrees of freedom of ε_t. In this case, the VaR is given as follows

[3] If we assume that y is normally distributed, then the density function of y is given as $f(y) = (2\pi)^{-1/2} exp(-z^2/2)$.

$$VaR_{t,q} = \mu_t + \widetilde{t_{d_\varepsilon}}^{-1}(q)\sigma_t$$

where $\widetilde{t_{d_\varepsilon}}^{-1}(q)$ is the q-th quantile of the $\tilde{t}(d)$ distribution. Consequently, it can be written

$$VaR_{t,q} = \mu_t + t_{d_\varepsilon}^{-1}(q)\sqrt{\frac{d-2}{d}}\sigma_t$$

Thus, d is calculated after estimating the EK of log returns as

$$EK = 6/(d-4)$$

$$d = {}^{6}/_{EK} + 4$$

**

The GARCH (1,1) is the most commonly used form of the GARCH(p, q) process especially when it comes to forecasting VaR. Moreover, according to the review of the literature in Shayya et al. (2023) the ARCH/GARCH family of models ranks first among the models used to estimate VaR, being mentioned in more than 300 papers aimed at developing new VaR models or comparing the already existing ones.

Despite its popularity and wide utility in the risk management industry, however, according to Abad et al. (2014) the main drawback of the GARCH (1,1) is the restrictions that should be imposed on the parameters i.e. $\alpha_0 > 0$, $\alpha, \beta > 0$ and $\alpha + \beta < 1$.

Many modifications were developed on the GARCH family of models in order to cope with the necessities required for application on financial series. The IGARCH (Integrated GARCH) model was introduced by Engle and Bollerslev (1986) to deal with the restriction of $\alpha + \beta < 1$, where in empirical applications it was observed that $\alpha + \beta$ is usually very close to one unit. So, the IGARCH was GARCH modified by imposing the restriction $\alpha + \beta = 1$.

Many other GARCH-type models were developed and Abad et al. (2014) list some of them with brief explanations like the FIGARCH (Fractional IGARCH) proposed by Baillie et al. (1996), the PGARCH (Power GARCH) model proposed in Higgins and Bera (1992). In addition, Nelson (1991) proposed the Exponential GARCH (EGARCH) model, which is widely used in empirical applications, it is based on the conditional variance logarithm, and since it is formulated upon logarithm it does not require positiveness condition of the parameters imposed by GARCH (1,1).

The main distinct advantage of GARCH-type models is their ability to capture thick tails in returns and volatility clustering, in addition to their flexibility regarding the choice of the distribution of residuals. This also has a high impact on the performance of the GARCH models since in forecasting VaR, it was noticed that the distribution model of returns has a significant impact on the accuracy of VaR forecasts. In general, it is commonly known and agreed upon that the distribution of financial returns is far from normal due to empirical evidence, however using more fat-tailed distributions of residuals and sometimes also asymmetric distributions (depending on the degree of skewness of the real data) improves the accuracy of VaR estimation. In this work, the GARCH-Normal and GARCH-Student-t models are used upon applying the Monte Carlo Simulation and are also used for fitting the data to obtain VaR in the simple GARCH (1,1) model framework. GARCH (1,1) is also employed in the application of the EWMA approach, which is a special case of GARCH (1,1) summing the parameters α and β to one ($\alpha + \beta = 1$) and $\alpha_0 = 0$. In addition, GARCH (1,1) is also used in this work with the implementation of the GARCH-EVT model proposed by McNeil and Frey (2000).

2.3. Extreme Value Theory

Market and financial institutions failures are mainly caused by sudden huge negative returns, which form the biggest risk to be faced. Knowing the probabilities of these extreme events explicitly is thus the principal target of risk management. According to de Haan and Ferreira (2006), the statistical theory of the Extreme Value theory (EVT) took shape in the 1980's. However, in 1927, the one-dimensional probabilistic EVT was developed by M. Frechet, and in 1928 by R. Fisher and L. Tipppet, then later by R. von Mises in 1936 and further developed by B. Gnedenko in 1943. The so-called first theorem of EVT, the extreme value theorem, is the Fisher-Tippett theorem of 1928 whose first formal proof was provided by Gnedenko (1943) which is why this theorem is sometimes referred to as Fisher-Tippett-Gnedenko. However, the statistical theory was initiated in 1975 by J. Pickands III (de Haan & Ferreira, 2006).

A. J. McNeil (1999) describes EVT as a tool that attempts to provide the best possible estimate of the tail area of the distribution of the empirical data even when there is a shortage in the historical data available. The EVT provides guidance on the kind of tail distributions so that extreme risks in this area are handled conservatively.

Basically, there are two principal groups of models for extreme values, the Block Maxima Models (BMM) and the Peaks-Over-Threshold (POT) models which is considered as the modern group. The BMM are used for large observations collected from large samples of identically distributed observations. However, the POT models are used for all large observations which exceed a high threshold (McNeil, 1999) and they are considered the most useful in practical applications due to their efficiency in using data on extreme values which is often limited. In this study, the focus is on the POT model, and it will be used to obtain VaR.

Peaks-Over-Threshold

The POT approach, also known as threshold exceedances method, is based on the idea that excesses over a certain threshold asymptotically follow a Generalized Pareto Distribution (GPD). There are two styles of analysis within the group of models that deal with the POT: the semi-parametric models and the fully parametric models. The former are built around the Hill estimator and its relatives like in Daníelsson and Vries (1997) among others, and the latter are based on GPD like in Embrechts et al. (1998). In addition, some studies used non-parametric approaches to estimate the tails of the distributions, like McNeil (1999), who used the historical simulation and found out that both approaches are useful and efficient when used correctly.

The POT method is based on modelling the exceedances, y_t, over a certain pre-determined threshold, u, such that $y_t = r_t - u$, for all $r_t > u$. The main result of EVT in this approach is that as u tends to infinity (or to the right end point of the distribution function F of the returns if it is finite), the positive sequence $\{y_t\}$ of exceedances appropriately scaled belongs to the family of GPD with shape parameter ξ. Then one can say that the financial returns whose block maxima follow a Generalized Extreme Value (GEV) distribution with shape parameter ξ_0 as per the EVT, are such that, for a sufficiently high threshold u, the exceedances over u follow a GPD with shape parameter $\xi = \xi_0$ (Rocco, 2014). The parameter ξ can be estimated by the well-known and widely used estimator, the Hill estimator, or using the maximum likelihood estimation and when considering fat-tailed distributions the inverse of ξ, $\eta = 1/\xi$, that is known as the tail index of the distribution. This approach was first introduced by Pickands III (1975) and Balkema and de Haan (1974) and it is known as the Pickands-Balkema-De Haan theorem which is the second theorem in the EVT after the Fisher-Tippett theorem. It states that for a sequence of i.i.d random variables, say here $\{y_t\}$ and y_t as defined above, with a distribution function F_u, then for a sufficiently high threshold u, F_u belongs to the GPD family. It will be seen later how these two theorems together play an important role in showing that as the threshold u is raised, the distributions for which normalized maxima converge to a GEV distribution constitute a set of distributions for which the excess distribution converges to the GPD.

Embrechts et al. (1997) give the GEV representation according to Jenkinson–von Mises and it is defined by the distribution function, H_ξ, which depends on a real parameter, denoted ξ, known as the shape parameter,

$$H_\xi(r) = \begin{cases} \exp\left(-(1+\xi r)^{-1/\xi}\right) & \text{if } \xi \neq 0 \\ \exp\left(-(e^{-r})\right) & \text{if } \xi = 0 \end{cases}$$

where r is such that $1 + \xi r > 0$. The case where $\xi = 0$ should be viewed as the case where $\xi \to 0$ and $\eta = 1/\xi$ is the tail index of the distribution or the extreme value index.

An adequate cut-off between the central part of the distribution and the upper tail is to be determined, which is a problematic aspect of the EVT methods. In the case of POT approach, the cut-off to be determined is the threshold u, which should be high enough to consider the observations in the tail only and at the same time, should not be too high giving too few extreme observations to obtain efficient estimates.

Generalized Pareto Distribution

The GPD is basically a distribution with two parameters having the distribution function, $G_{\xi,\beta}$, defined as follows

$$G_\xi(y) = \begin{cases} 1-(1+\xi y/\beta)^{-1/\xi} & \text{when } \xi \neq 0 \\ 1-\exp(-y/\beta) & \text{when } \xi = 0 \end{cases}$$

where $\beta > 0$, and $y \geq 0$ when $\xi \geq 0$, and $0 \leq y \leq -\beta/\xi$ when $\xi < 0$ where ξ and β are referred to, respectively, as the shape and scale parameters of the distribution. Similar to the GEV and according to the value of ξ, three cases are presented:

- $\xi > 0$, then $G_{\xi,\beta}$ is a reparametrized version of the ordinary Pareto distribution, $Pa(\alpha_P, \kappa_P)$, and it is the most relevant case for risk management because GPD acquires a heavy tail when $\xi > 0$ where $\alpha_P = 1/\xi$ and $\kappa_P = \beta/\xi$. However, unlike the normal distribution, this case does not possess moments of all orders. In this case, $E[r^k]$ is infinite when $k \geq 1/\xi$; for instance, if $\xi = 0.5$ then GPD has infinite variance, and if $\xi = 0.25$ then the GPD has an infinite fourth moment (kurtosis).
- $\xi = 0$, then refers to the exponential distribution.
- $\xi < 0$, refers to a short-tailed Pareto type II distribution.

The GPD distribution function can be scaled by adding a scale parameter σ and a location parameter ψ, such that $\psi \in \mathbb{R}$ and $\beta > 0$, thus the above definition of G_ξ can be written with $(y - \psi)/\sigma$ instead of y.

Estimating Excess Distributions over a threshold "*u*"

The threshold exceedances or excess losses over the pre-determined high threshold, u, are defined as $y_t = r_t - u$ for all $r_t > u$.

Accordingly, denoting the distribution function of underlying log returns by F, the distribution of excess losses over u, denoted F_u can be written as

$$F_u(y) = P\{r - u \leq y \mid r > u\}$$

For $0 \leq y < r_0 - u$ where $r_0 < \infty$ is the right end point of F (McNeil, 1999).

$F_u(y)$ represents the probability that the value of r exceeds the threshold u by at most an amount y given that r exceeds u. Then the conditional probability can be written as

$$F_u(y) = \frac{Pr\{r - u \leq y,\ \ r > u\}}{Pr(r > u)} = \frac{F(y + u) - F(u)}{1 - F(u)}$$

Having $r = y + u$ for $r > u$, then

$$F(r) = [1 - F(u)]\, F_u(y) + F(u), \text{ valid for } r > u.$$

The theorem of Balkema and de Haan (1974) and Pickands III (1975) explains the importance of GPD. It states that, for a sufficiently high threshold u, the distribution function of excess may be approximated by GPD, since as the threshold u tends to r_0, the right end point of F, F_u converges to a GPD. Thus, the GPD can be written as follows.

$$G_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \xi \dfrac{y}{\beta}\right)^{-1/\xi} & \text{when } \xi \neq 0 \\ 1 - \exp(-(y)/\beta) & \text{when } \xi = 0 \end{cases}$$

Where $\beta > 0$, and $y \geq 0$ when $\xi \geq 0$ and $0 \leq y \leq -\beta/\xi$ when $\xi < 0$.

Again, ξ is the shape parameter and β is the scale parameter. The tail index, η, is defined as $\eta = 1/\xi$. Gençay and Selçuk (2004) mention the simple relationship that holds between the GEV, $H_\xi(y)$ defined above, and the GPD, $G_\xi(y)$ where,

$$G_\xi(y) = 1 + \log H_\xi(y) \text{ if } \log H_\xi(y) > -1$$

Estimating the tail index

The GPD model can be approximated using the maximum likelihood method. However, some studies use the Hill estimator for estimating ξ, which was introduced by Hill (1975) and it is a widely used semi-parametric approach that can be used in particular for heavy-tailed data when $\xi > 0$ (Fréchet case) McNeil and Frey (2000). McNeil (1999) used the historical simulation (non-parametric approach) to estimate the tail index.

In order to obtain the tail estimation, McNeil (1999) and Gençay and Selçuk (2004), among others, used the method of maximum likelihood. It is worth noting that the Pickands-Balkema-De Haan- theorem plays a key role in estimating tails since it results in approximating the distribution of exceedances over the threshold by the GPD. Therefore, the above equation

$F(r) = [1 - F(u)]\, F_u(y) + F(u)$, valid for $r > u$ and $r = y + u$,

can be written as,

$F(r) = [1 - F(u)]\, G_{\xi,\beta}(y) + F(u)$ since $F_u(y) \approx G_{\xi,\beta}(y)$.

After choosing a sufficiently high threshold u, $F(u)$ can be determined by $(N - N_u)/N$, i.e., using the Historical Simulation method, where N is the total number of observations and N_u is the number of exceedances over the threshold u. Thus, denoting by $\hat{F}(r)$ the estimator of F, and joining this estimate together with the GPD yields

$$\hat{F}(r) = \left[1 - \frac{(N - N_u)}{N}\right] G_{\hat{\xi},\hat{\beta}}(r - u) + \frac{(N - N_u)}{N}$$

$$= \frac{N_u}{N}\, G_{\hat{\xi},\hat{\beta}}(r - u) + \frac{(N - N_u)}{N}$$

$$\hat{F}(r) = 1 + \frac{N_u}{N}\left(\left(1 + \hat{\xi}\,\frac{(r - u)}{\hat{\beta}}\right)^{-1/\hat{\xi}} - 1\right)$$

$$\hat{F}(r) = 1 - \frac{N_u}{N}\left(1 + \hat{\xi}\,\frac{(r - u)}{\hat{\beta}}\right)^{-1/\hat{\xi}}$$

Where $\hat{\xi}$ and $\hat{\beta}$ denote the maximum likelihood estimators of ξ and β, respectively, such that $r > u$.

Another method for estimating the tail distribution is the Hill estimator, proposed by Hill (1975), for the case when $\xi > 0$. The Hill estimator was based on the intuition of drawing an inference on the behavior of a distribution function in the tails regardless of the parametric form of that distribution. According to Hill (1975), it was interesting to study the behavior of a distribution sample in the tails without assuming a particular parametric form for the distribution hold globally. Examples that gave rise to his article concerned random sample $Z_1, Z_2, \dots, Z_N$ from a distribution F on the unit interval with $F(x) \sim cx^{\alpha}$ as $x \to 0$, where c is a constant to be estimated. The main target desired was to draw an inference about α without making assumptions about the form of F elsewhere. As per Hill (1975), a similar situation occurs in connection with inference about the parameter α of a Pareto Law:

$$1 - F(x) \sim cx^{-\alpha} \text{ as } x \to \infty$$

where c is a real parameter to be estimated.

Hill (1975) proposes a general inference form that can be implemented from either Bayesian or classical approaches. In order to obtain the estimator for ξ, the data is ordered as per the order statistics follows according to their values i.e., $r_1 \geq r_2 \geq \cdots \geq r_N$ (however, it is noticed that the order statistics is used in reverse here which is proposed by Hill to simplify the formulas in applications). The cumulative distribution of the log returns is given as

$$F(r) = 1 - cr^{-1/\xi} \approx 1 - \left(1 + \xi \frac{r}{\beta}\right)^{-1/\xi} = G_{\xi,\beta}(r) \text{ for } r > u \text{ } and \text{ } \xi > 0.$$

By definition, the conditional distribution is given by

$$f(r|r > u) = f(r)/Pr(r > u) = f(r)/\big(1 - F(u)\big), \text{ for } r > u.$$

Note that from the definition of $F(y)$ it can be written,

$$F(u) = 1 - cu^{-1/\xi}$$

$F(r)$ is derived to get $f(r)$ as follows,

$$f(r) = \frac{\partial F(r)}{\partial r} = \frac{1}{\xi} cr^{-1/\xi - 1}$$

Now, the likelihood function for all r_i larger than the threshold u can be constructed as follows,

$L = \prod_{i=1}^{N_u} f(r_i)/\big(1 - F(u)\big) = \prod_{i=1}^{N_u} \frac{1}{\xi} cr_i^{-1/\xi - 1}/\big(cu^{-1/\xi}\big)$, for $r_i > u$ where N_u is the number of observations r larger than the threshold u.

$$\ln L = \sum_{i=1}^{N_u} \left(-\ln(\xi) - (-1/\xi + 1) \ln r_i + \frac{1}{\xi} \ln(u) \right)$$

To get the Hill estimator of the tail index, the derivative of the above equation is obtained with respect to ξ and set it equal to zero, and this implies.

$$\hat{\xi} = \frac{1}{N_u}\sum_{i=1}^{N_u} \ln(r_i/u)$$

Then, the parameter *c* is estimated using the equation,

$$F(u) = 1 - cu^{-1/\xi} = 1 - N_u/N$$

Solving the above equation for c yields,

$$c = \frac{N_u}{N}u^{1/\xi}$$

Therefore, the estimate of the cumulative density function for observations beyond the threshold u, F, is given by,

$$F(r) = 1 - cr^{-1/\xi} = 1 - \frac{N_u}{N}\left(\frac{r}{u}\right)^{-1/\xi}.$$

It is worth mentioning that the above simplified derivation of the Hill estimator follows the display of the estimator given in Christoffersen (2011). De Haan and Ferreira (2006), among others, also provide a detailed display of the Hill estimator. Gençay and Selçuk (2004) discuss the difficulty of the Hill estimator being in the ambiguity of the value of the threshold u, due to the tradeoff faced between the bias and variance. A too high threshold will generate in fewer excesses over the threshold, and thus leads to higher variance of the estimators (Davison & Smith, 1990; MacDonald et al., 2011). And a too low threshold increases the number of observations, and the estimation becomes more precise but, at the same time, introduces some observations from the center of the distribution and the estimation becomes biased.

Benito et al. (2023) categorize the threshold selection methods into two groups: (a) graphical approaches, which are based on visual inspection of plots (Mean Excess plot, stability parameters plot and Hill plot, among others); (b) numerical approaches. The latter can be also categorized into several groups: (i) non-parametrical approaches; (ii) approaches based on goodness-of-fit; (iii) mixture models; (iv) simple naïve methods; (v) computational approaches; (vi) other approaches[4].

Graphical methods like the MEF (mean excess function) and Hill plot are widely used in the literature. However, their main drawback is that they can be subjective and require substantial expertise to interpret the diagnostic graphical methods as a way for choosing the threshold (Benito et al., 2023)

Extensive research has been done on the importance of the optimal threshold selection and its effects on the quantification of market risk. For instance, Benito et al. (2023) found that the choice of the threshold in the POT framework may not be relevant in quantifying

[4] Scarrot and McDonald (2012) and Langousis et al. (2016) present a detailed review on these methods.

market risk when VaR and expected shortfall measures are used for this task. The authors noticed that the high quantile of GPD does not depend on the threshold choice, at least for the wide range of threshold chosen in their paper between the 80th and 90th percentile of the data.

In addition, Harmantzis et al. (2006) experimented in their work with POT many values of the threshold around the vicinity of 10% and noticed no significance change in the results.

In this work, as will be seen later in Chapter 3, two values of the threshold are used, and the respective results are compared. The threshold values used in this thesis are 5% and 10% quantiles of the data. The 10% quantile was proposed by DuMouchel (1983) among the simple methods for threshold selection which is most frequently used in practice as well. In Chavez-Demoulin et al. (2014) and Chavez-Demoulin and Embrechts (2004) among others, 10% of the data sample was also used as threshold exceedances. The 5% quantile threshold is used for comparison purpose.

VaR Estimation with EVT (POT case)

In terms of the loss distribution function F, VaR is the qth quantile of the distribution. The estimator of the loss function F is obtained as follows.

$$\hat{F}(r) = 1 - \frac{N_u}{N}\left(1 + \hat{\xi}\,\frac{(r-u)}{\hat{\beta}}\right)^{-1/\hat{\xi}}$$

Gençay and Selçuk (2004) explain how VaR expression with EVT is obtained. The shape and scale parameters of the GPD, ξ and β respectively, are obtained using the maximum likelihood estimation. Smith (1987) investigates the asymptotic relative error of the above estimator $\hat{F}(r)$ and gets a result of the form$\left(\hat{\xi} - \xi\right)N_u^{1/2}$, that is asymptotically normally distributed with zero mean and ξ^2 variance.

Using maximum likelihood estimation, the density function f of the GPD distribution is given as,

$$f(r) = \frac{1}{\beta}\left(1 + \xi\,\frac{r}{\beta}\right)^{-\frac{1}{\xi}-1}$$

The corresponding log-likelihood function is,

$$\ell(\xi,\beta) = -n\log(\beta) - \left(\frac{1}{\xi} + 1\right)\sum_{i=1}^{N}\log\left(1 + \frac{\xi}{\beta}r_i\right)$$

where N is the sample size. As per evidence from Hosking and Wallis (1987), when $\xi > -0.5$, the maximum likelihood regularity conditions are fulfilled, and the maximum likelihood estimates are asymptotically normally distributed. This implies that the approximate standard errors for the estimators of β and ξ can be also obtained through maximum likelihood estimation.

And again, for a given probability $q > F(u)$, VaR is estimated by inverting the tail distribution and obtained with the same expression as in McNeil (1999),

$$VaR_{\varepsilon}^{t}(q) = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{N}{N_u}(1-q)\right)^{-\hat{\xi}} - 1\right)$$

Where u is the threshold, $\hat{\beta}$ is the estimated scale parameter and $\hat{\xi}$ is the estimated shape parameter, N is the sample size and N_u is the number of exceedances over the threshold u.

One Day Horizon VaR

Now, the VaR of financial return series of a single asset is estimated considering that the returns follow a stochastic volatility model such that.

$$r_t = \mu_t + \sigma_t \varepsilon_t$$

Where σ_t is the volatility of the return on day t and μ_t is the expected return. Assuming also that the noise variables ε_t are i.i.d with an unknown distribution having a mean of zero and unit variance. It is also assumed that the returns are identically distributed with unknown distribution $F(r)$. Models of the ARCH/GARCH family fit into this framework, so it is assumed as per McNeil (1999) that.

$\mu_t = \phi r_{t-1}$ and $\sigma_t^2 = \alpha_0 + \alpha_1(r_{t-1} - \mu_{t-1}) + \beta\sigma_{t-1}^2$

As per the structure of the stochastic volatility model of the returns, and after obtaining the quantile of the noise variable ε_t, VaR can be written as given in McNeil (1999) as follows,

$$VaR_r^t(q) = \mu_{t+1} + \sigma_{t+1} VaR_{\varepsilon}^t(q)$$

Where $VaR_{\varepsilon}^t(q)$ denotes the qth quantile of the noise variable ε_t. The simplest approach for estimating the dynamic VaR in this case is to assume that ε_t with the distribution function $F_{\mathrm{E}}(\varepsilon)$ follows a normal distribution.

***

The main advantage of EVT over traditional VaR methods lies in that it works on estimating the distribution of tails which is the main interest for risk management. However, traditional VaR methods focus on the distribution of central observations (Lewis, 2012). Gençay and Selçuk (2004) mention another advantage of the EVT over other approaches as it provides a convenient framework for the separate treatment of the tail's distribution, unlike other approaches which assume symmetric distributions such as Student-t, normal or even ARCH/GARCH-like distributions. Rocco (2014) lists three main advantages of the EVT approach in estimating VaR. The first is the use of the asymptotic distribution of extreme values which does not depend on the exact cumulative distribution function F of returns. As a result, the precise form of F can thus be ignored and VaR can be estimated using non-parametric or semi-parametric approaches, which is important given that although financial time series usually exhibit skewed or fat-tailed distributions; however, there is no complete agreement about the distribution that would best fit tails. The second advantage is that EVT focuses on the tails, thus contrary to other parametric approaches the estimates of VaR by EVT are not biased by the credit given to the central part of the distribution. And the third advantage is that it allows each of the lower and upper tail of the financial time series to be tackled independently taking into account the skewness of the underlying distribution.

According to the literature, many studies concerning the employment of EVT in the estimation of VaR have showed the accuracy of its results compared to other approaches like the traditional approaches for VaR estimation and also the GARCH-type models (especially the symmetric models). In fact, the EVT approach described above is referred as the standard EVT approach estimating the unconditional distribution of returns, like in Embrechts et al. (1998)., Longin (2000), McNeil (1998) and Daníelsson and Vries (1997) among others. However, none of the mentioned studies used an EVT based approach with VaR estimates that reflect the current volatility background. And since it is commonly agreed that financial time series exhibit conditional heteroskedasticity, thus the above studies among others of same aspect present a major drawback of the standard EVT method.

This was overcome by the work of McNeil and Frey (2000) who introduced the combination of GARCH models with EVT to obtain better estimates of VaR reflecting the time varying volatility of the data and the conditional heteroskedasticity. Thereafter, many studies were carried evolving on this idea and providing good references in this aspect like for instance, Kuester et al. (2006) and Byström (2004) among others. McNeil and Frey (2000) use GARCH modelling and the maximum likelihood estimation to obtain estimates of conditional volatility and then they use the Historical Simulation in the central part of the distribution, and the EVT threshold methods in the tails to estimate the distribution of residuals. The approach of McNeil and Frey (2000) is a bit similar to that of the Filtered Historical Simulation approach of Barone-Adesi et al. (1998) in which they fit a GARCH model to the financial return series and they use historical simulation to obtain the VaR. However, they do not use EVT methods to estimate the tails of the distribution of residuals. The approach of McNeil and Frey (2000) proves to be more accurate than the unconditional EVT approach and GARCH model with normal residuals.

**

2.4. GARCH-EVT approach

According to the literature EVT is very widely used to estimate VaR, especially after the traditional methods raised suspicions regarding their accuracy and credibility mainly after the crisis of 2008. However, it was noticed that efficient EVT methods were usually the ones jointly applied with GARCH type models, which in turn take into account the time varying variance of the data by considering conditional volatility.

In the same context, McNeil and Frey (2000) can be considered as the first formal article proposing the combination of EVT and GARCH models. The authors proposed a method for estimating VaR by describing the tail of the conditional heteroskedastic financial return series. This approach is based on pseudo-maximum likelihood (PML) fitting of GARCH models to estimate the volatility and using EVT for estimating the tail of the residual distribution of the fitted GARCH model. This method showed better 1-day VaR estimates than other methods which ignore the heavy tails of the residuals, or in general those models which ignore the stochastic nature of volatility.

The previous EVT methods, for example Embrechts et al. (1998), McNeil (1998) and Daníelsson and Vries (1997) among other works, do not reflect the current volatility background of the financial return series. McNeil (1998) uses estimation techniques based on limit theorems for block maxima not ignoring the clustering of extremal events caused by stochastic volatility and focuses on showing how to correct this clustering. Whereas, Daníelsson and Vries (1997) use a semi-parametric approach based on the Hill estimator.

This method combines ideas of three approaches to overcome a major drawback of previous EVT methods, which is mainly ignoring the conditional heteroskedasticity of financial return series. It uses the GARCH model fitting with pseudo-maximum-likelihood estimation in order to obtain estimates for the conditional volatility. According to statistical tests and data analysis, the residuals of the GARCH models form an i.i.d series with heavy tails. McNeil and Frey (2000) use historical simulation for the central part of the distribution and threshold methods from EVT to estimate the distribution of residuals. The significance of this approach is that it reflects the stochastic volatility and fat tails exhibited by most financial returns series.

According to McNeil and Frey (2000), a similar work to theirs was proposed by Barone-Adesi et al. (1998). The approach of the latter is known as the Filtered Historical Simulation in which they fit a GARCH model to a financial return series and use historical simulation to infer the distribution of the residuals and estimate VaR, however they do not use EVT to estimate the tails of the distribution of residuals. According to McNeil and Frey (2000), the central idea of their approach, which is the application of EVT to model residuals, has been independently proposed by Diebold at al. (1998).

First, they define $(r_t, t \in \mathbb{Z})$ to be a strictly stationary time series representing daily negative log returns observations on a financial asset price, assuming that the dynamics of r_t are given by

$$r_t = \mu_t + \sigma_t \varepsilon_t$$

Where the residuals (ε_t) are a strict white noise with a marginal distribution function $F_{\mathrm{E}}(\varepsilon)$, and it is assumed that μ_t and σ_t are measurable with respect to Ω_{t-1}, the information set available on the return process up to time $t-1$.

Then, they define $F_R(r)$, the marginal distribution of (r_t), and $F_{r_{t+1}+\cdots+r_{t+h|\Omega_t}}(r)$ denotes the predictive distribution of the returns over the next h days given the information set up to and including day t. The main interest is focused on obtaining the quantiles of the tails of these distributions. The unconditional quantile is a quantile of the marginal distribution denoted by

$$r_q = inf\{r \in \mathbb{R}, F_R(r) \geq q\} \text{ for } 0 < q < 1,$$

And a conditional quantile is a quantile of the predictive distribution for the return over the next h days, denoted by

$$r_q^t(h) = inf\left\{r \in \mathbb{R}: F_{r_{t+1}+\cdots+r_{t+h|\Omega_t}}(r) \geq q\right\}$$

The quantile of the 1-step ($h = 1$ day) predictive distribution is denoted by r_q^t derived as follows.

$$F_{r_{t+1}|\Omega_t}(r) = Pr\{\sigma_{t+1}\varepsilon_{t+1} + \mu_{t+1} \leq r|\Omega_t\} = F_{\mathrm{E}}((r - \mu_{t+1})/\sigma_{t+1})$$

In a simplified form, the above equation is equivalent to

$$VaR_t(q) = r_q^t = \mu_{t+1} + \sigma_{t+1}\varepsilon_q$$

Where ε_q is the upper qth quantile of the marginal distribution of (ε_t) which by assumption does not depend on t.

Among the different models for volatility dynamics in the econometric literature, McNeil and Frey (2000) use the GARCH (1,1) process for volatility and the AR (1) model for the dynamics of the conditional mean noting that their approach can be easily extended to more complex models.

Using the GARCH (1,1) model is often associated with normally distributed residuals and this simply implies that $\varepsilon_q = \Phi^{-1}(q)$ where $\Phi(\varepsilon)$is the standard normal density function, and this model can be fitted with Maximum Likelihood (ML). However, McNeil and Frey (2000) show with empirical evidence that this approach underestimates the conditional quantile for $q > 0.95$ since the distribution of residuals seems to exhibit heavier tails than the normal distribution.

Using Student-*t* residuals (scaled Student-*t* is used with variance 1) is another standard approach when residuals are assumed to have a leptokurtic (heavy-tailed) distribution. GARCH type models with Student-*t* residuals can also be fitted using ML and the degree of freedom parameter of the distribution can be estimated. For instance, consider,

$$\varepsilon = \sqrt{(\nu - 2)/\nu}\, T,$$

where *T* has a Student-*t* distribution and $\nu > 2$ degrees of freedom with density function $F_T(t)$.

Then

$$\varepsilon_q = \sqrt{(\nu - 2)/\nu}\ F_T^{-1}(q)$$

McNeil and Frey (2000) show that this approach is effective if only if the data is symmetric, i.e., the positive and negative tails of the distribution are (roughly) equal. However, they suggest making minimal assumptions regarding the distribution of residuals and for this reason they propose to consider them as a strict white noise with mean zero and unit variance. The method of McNeil and Frey (2000) constitutes two main steps:

1. Fitting a GARCH-type model to the return data using the PML method without making any assumption on the distribution of residuals. This leads to estimating the one-day ahead mean and volatility using the fitted model and calculating the implied residuals.
2. Considering the residuals to be a realization of a strict white noise process (i.i.d with mean zero and unit variance) and using EVT to model their tail and estimate the quantile ε_q for $q > 0.95$.

Estimating σ_{t+1} and μ_{t+1} using PML

For a predictive purpose, a constant memory *n* is fixed so that at the end of day *t*, the data consists of the last *n* negative log returns $(r_{t-n+1}, \dots, r_{t-1}, r_t)$ which are considered to be realizations from an AR (1)-GARCH (1,1) process. Then the conditional variance of the mean-adjusted series $\rho_t = r_t - \mu_t$ is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \rho_{t-1}^2 + \beta \sigma_{t-1}^2$$

Where $\alpha_0 > 0$, $\alpha_1 > 0$ and $\beta > 0$ and $\beta + \alpha_1 < 1$ [5]. Then the conditional mean is given by

$$\mu_t = \phi r_{t-1}$$

The mean-adjusted series is $\rho_t = r_t - \mu_t$.

The above model is fitted using the PML method. The PML provides reasonable estimation of the parameters of the model denoted $\hat{\theta} = \left(\hat{\phi}, \hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}\right)^T$, in fact it can be shown that the PML yields a consistent and asymptotically normal estimator. From the equations of the conditional variance and conditional mean, the conditional standard deviation and conditional mean estimates can be obtained recursively after substituting the sensible starting values. Then the residuals are calculated to check for the GARCH model and to be used in the second stage of this approach where the EVT will be applied on the tail of the residuals above a certain threshold.

The residuals should be i.i.d. if the fitted model is proper and the residuals are calculated as follows,

$$(\varepsilon_{t-n+1}, \dots, \varepsilon_t) = \left(\frac{r_{t-n+1} - \hat{\mu}_{t-n+1}}{\hat{\sigma}_{t-n+1}}, \dots, \frac{r_t - \hat{\mu}_t}{\hat{\sigma}_t}\right)$$

[5] The model employed above by McNeil and Frey (2000) is a special case of the general first order stochastic volatility process considered by Duan (1997) who in turn uses a result by Brandt (1986) to give conditions for strict stationarity. The mean adjusted series (ρ_t) is strictly stationary if $E[\log(\beta + \alpha_1 \varepsilon_{t-1}^2)] < 0$. By using Jensen's inequality and the convexity of $-\log(x)$ it is seen that a sufficient condition for $E[\log(\beta + \alpha_1 \varepsilon_{t-1}^2)] < 0$ is that $\beta + \alpha_1 < 1$, which also ensures that the marginal distribution function $F_r(r)$ has a finite second moment.

The Ljung-Box test can be applied on residuals to check their i.i.d. property although raw data may not acquire this property. If the fitted model provides satisfactory results, the one-day conditional mean and variance forecasts are calculated as follows,

$\hat{\mu}_{t+1} = \hat{\phi} r_t$ and

$\hat{\sigma}_{t+1}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\rho}_t^2 + \hat{\beta} \hat{\sigma}_t^2$ where $\hat{\rho}_t = r_t - \hat{\mu}_t$

Estimating the tail index and $\boldsymbol{\varepsilon_q}$

Now, the second stage is applied which consists of estimating the qth quantile ε_q of the marginal distribution of ε_t.

First, the QQ-plot of residuals against the normal distribution is formed in order to confirm that the assumption of normality of residuals is unrealistic and that the residuals process is leptokurtic. Then, a high threshold u is fixed and the excess residuals over u have a GPD.

The choice of GPD is motivated by a limit result in EVT. Then the cumulative distribution function F and the corresponding excess distribution over the threshold u in the tail of the residuals (ε_t) is given by

$$F_u(y) = P\{\varepsilon - u \leq y \mid r > u\} = \frac{Pr\{\varepsilon - u \leq y,\ \ \varepsilon > u\}}{Pr(\varepsilon > u)} = \frac{F(y+u) - F(u)}{1 - F(u)}$$

For $0 \leq y < \varepsilon_0 - u$, where ε_0 is finite or infinite right endpoint of F.

According to the Fisher-Tippett-Gnedenko theorem, $F \in MDA(H) \Rightarrow H$ is of the type H_ξ for some ξ, i.e., H_ξ is a GEV distribution.

Then Pickands-Balkema-De Haan theorem showed that for a sufficiently large threshold *u,* the distribution function of the excess, $F_u(y)$, converges to the GPD if and only if the cumulative distribution function F belongs to the maximum domain of attraction of H_ξ. Here, it is worth noting that according to the above two theorems, it can be said that the distributions for which normalized maxima (in the BMM method) converge to a GEV distribution constitute a set of distributions for which the distribution of excess converges to the GPD for a sufficiently large threshold *u* where the GEV and the GPD admit the same shape parameter ξ. And since all the commonly used continuous distributions of statistics are in the Maximum Domain of Attraction of H_ξ, MDA(H_ξ), for some ξ, then the Pickands-Balkema-De Haan theorem proves to be a very widely applicable result that firmly states that the GPD is the canonical distribution for modelling excess distributions over high thresholds (McNeil et al., 2015).

As a result, in the class of distributions for which the above result holds are the most common continuous distributions of statistics which according to the value of the parameter ξ can be divided into three groups to limit the GPD approximation to the excess distribution:

- $\xi > 0$ corresponds to heavy tailed distributions like Pareto, Student-*t*, Cauchy, Burr, Log-gamma and Fréchet distributions, whose tails decay like power functions, also known as power-law decaying (Frechet case).
- $\xi = 0$ corresponds to distributions like normal, exponential, gamma and log-normal whose tails decay exponentially (Gumbel case).
- $\xi < 0$ corresponds to short-tailed distributions with finite right endpoint like the uniform and beta distributions, and it is the least relevant case for finance.

Assuming that the tail of the underlying distribution begins at the threshold u, then a number of observations N_u will exceed this threshold such that $N_u > 0$. McNeil and Frey (2000) use the maximum likelihood method to estimate the GPD parameters $\hat{\beta}$ and $\hat{\xi}$ and again they stress on the importance of the choice of the threshold u in the implementation of EVT. Then the following equality for points over the threshold in the tails of F, $\varepsilon > u$, is considered,

$$1 - F(\varepsilon) = \big(1 - F(u)\big)\big(1 - F_u(\varepsilon - u)\big)$$

Then using the random proportion of data in the tail N_u/N where N is the total number of observations (size of the sample) and estimating the second term according to the GPD approximation of excess distribution over the threshold (as a result of Pickands-Balkema-De Haan theorem) fitted by maximum likelihood, the tail estimator becomes as follows,

$\widehat{F(\varepsilon)} = 1 - \frac{N_u}{N}\left(1 + \xi\frac{\varepsilon - u}{\beta}\right)^{-1/\xi}$, for $\varepsilon > u$.

Again, as per Smith (1987), and similar to the above POT method,

$$N_u^{1/2}\left(\frac{1-\widehat{F(\varepsilon)}}{1-F(\varepsilon)} - 1\right) \xrightarrow{d} N(0, \nu^2),$$

as $u = u_N \to \varepsilon_0$ and $N_u \to \infty$.

However, in practice McNeil and Frey (2000) modify the procedure and fix the amount of data in the tail to be $N_u = k$ where $k \ll N$, which gives rise to a random threshold, u, at the $(k+1)$th order statistic (i.e., u is r_{k+1}). Then they let $\varepsilon_1 \geq \varepsilon_2 \geq \cdots \geq \varepsilon_N$ represent the ordered residuals, then the excess over the threshold in the tail to which the GPD, with parameters ξ and β, is fitted is represented as $(\varepsilon_1 - \varepsilon_{k+1} \geq \varepsilon_2 - \varepsilon_{k+1} \geq \cdots \geq \varepsilon_k - \varepsilon_{k+1})$. The form of the tail estimator of $F_{\mathrm{E}}(\varepsilon)$ is then

$$\widehat{F_{\mathrm{E}}(\varepsilon)} = 1 - \frac{k}{N}\left(1 + \hat{\xi}\frac{\varepsilon - \varepsilon_{k+1}}{\hat{\beta}}\right)^{-1/\hat{\xi}}$$

Since the qth quantile ε_q is the inverse of the distribution function of the tail, then for $q > 1 - k/N$, $\widehat{F_{\mathrm{E}}(\varepsilon)}$ can be inverted to obtain the estimate $\hat{\varepsilon}_{q,k}$ (k in the notation is used to emphasize the dependence of the estimator on the choice of k) of ε_q. Then the estimator is written as follows

$$\hat{\varepsilon}_q = \hat{\varepsilon}_{q,k} = \varepsilon_{k+1} + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{1-q}{k/N}\right)^{-\hat{\xi}} - 1\right)$$

McNeil and Frey (2000) illustrate graphically the AR (1)-GARCH (1,1) empirical residuals (without any assumptions on the distribution of the residuals) and also the AR(1)-GARCH (1,1) with standard normal distribution and also with (scaled) Student-*t* residuals. They show in their QQ-plot analysis that the normal distribution underestimates the large losses and large gains, on the other hand, the Student-*t* distribution underestimates losses and overestimates the gains. These results illustrate the drawbacks of using symmetric residual distributions when the data is not symmetric in tails.

However, they note that when the data has a more symmetric behavior, the conditional Student-*t* distribution works quite well and it is considered as a special case of the original method of McNeil and Frey (2000). This would be an example of a heavy-tailed distribution i.e., a distribution whose limiting excess distribution is GPD with $\xi > 0$. Gnedenko (1943), characterized such distributions to be having tails of the form

$$1 - F(r) = r^{-1/\xi} L(r)$$

where $L(r)$ is a slowly varying function and ξ is the positive parameter of the limiting GPD function. The tail index of F is usually referred to as $1/\xi$. For the Student-*t* distribution with ν degrees of freedom, the tail can be written as

$$1 - F(r) \sim \frac{\nu^{(\nu-2)/2}}{\mathrm{B}(1/2, \nu/2)} r^{-\nu}$$

where $\mathrm{B}(a, b)$ denotes the Beta function. According to McNeil and Frey (2000), this provides a simple example of a symmetric distribution when the data is symmetric and $\xi > 0$ and in this case the value of ξ in the limiting GPD will be the reciprocal of the degrees of freedom of the Student-*t* distribution.

McNeil and Frey (2000) perform a simulation study in order to investigate the issue of the choice of the threshold u. This study can be also useful for comparing the GPD tail estimation with the Hill estimator approach and the approach based on the empirical distribution function (historical simulation).

The Hill estimator, as previously described, is designed for data from heavy-tailed distributions admitting the following representation

$1 - F(r) = r^{-1/\xi} L(r)$ for $\xi > 0$, where $L(r) = \prod_{i=1}^{N_u} f(r_i) / \big(1 - F(u)\big) = \prod_{i=1}^{N_u} \frac{1}{\xi} c r_i^{-1/\xi - 1} / \big(c u^{-1/\xi}\big)$, is the likelihood function for all r_i larger than the threshold u and N_u is the number of observations r larger than the threshold u.

The Hill estimator $\hat{\xi}^{(H)}$ is then obtained by deriving the likelihood function with respect to r and then setting it equal to zero to get the estimator for ξ which is based on the k exceedances of the $(k+1)$th order statistic u and it is given as follows,

$$\hat{\xi}^{(H)} = \frac{1}{N_u}\sum_{i=1}^{N_u} \ln(r_i/u)$$

And the associated quantile estimator is obtained as per Daníelsson and Vries (1997) having the following form

$$\hat{\varepsilon}_q{}^{(H)} = \hat{\varepsilon}_{q,k}{}^{(H)} = \varepsilon_{k+1}\left(\frac{1-q}{k/N}\right)^{-\hat{\xi}}$$

McNeil and Frey (2000) generate a simulation study to investigate the issue of threshold choice. For the simulation study, they generate samples of size N from Student-t's distribution, which was observed to provide a rough estimation of the distribution of model residuals. The size of the generated sample corresponds to the window length used in the application of the two-step method. Again, the tail index and the quantiles of the Student-*t* distribution are easily calculated as above

$$1 - F(r) \sim \frac{\nu^{(\nu-2)/2}}{\mathrm{B}(1/2\,, \nu/2)} r^{-\nu}$$

Simulation Study of threshold choice

Then, $\hat{\xi}_k$ and $\widehat{\varepsilon_{q,k}}$ (the maximum likelihood and the GPD based estimators of ξ and ε_q based on the k threshold exceedances) are calculated, and also the tail and quantile estimators, $\hat{\xi}^{(H)}$ and $\hat{\varepsilon}_q{}^{(H)}$ respectively, as per Hill estimator are calculated for various values of k. For the quantile estimates, McNeil and Frey (2000) restrict their attention to the values of k such that $k > N(1-q)$, so that the target quantile is beyond the threshold. Then for each estimator they estimate the mean squared error (MSE) and bias focusing on the dependence of these estimators on k. They use Monte Carlo estimates based on 1000 independent samples. For example, they estimate $\widehat{MSE}\left(\widehat{\varepsilon_{q,k}}\right)$ by

$$\widehat{MSE}\left(\widehat{\varepsilon_{q,k}}\right) = \sum_{j=1}^{N}\left(\widehat{\varepsilon_{q,k}}^{\,j} - \varepsilon_q\right)^2$$

Where $\widehat{\varepsilon_{q,k}}^{\,j}$ represents the quantile estimate obtained from the jth sample.

According to McNeil and Frey (2000), although the Hill estimator is generally the most efficient estimator of the ξ, it might not provide the most stable quantile estimator, and as per their results, the GPD based quantile estimator should be preferred for estimating high quantiles.

These authors used for their study a Student-*t* distribution with 4 degrees of freedom ($\nu = 4$), and they plotted the bias and MSE estimators of the 99th percentile against k. Their results show that for the 99th percentile both Hill estimator and GPD are useful if used correctly.

In the GPD case it is a must to ensure that the variance of the estimator is kept low by setting k sufficiently high, however the issue of choosing an optimal threshold does not seem to be so critical in the GPD case. On the other hand, the threshold choice seems rather important in the Hill estimator case because the efficient range of k is smaller, it is important that the bias is kept under control by choosing a low value of k.

Although these results are referred to a Student-*t* distribution with four degrees of freedom, further simulations suggest the same qualitative conclusions to hold for other values of the degree of freedom and also for other heavy-tailed distributions. Nevertheless, it is important to note that for estimating more distant quantiles the GPD method appears to be more efficient than the Hill method maintaining its relative stability with respect to the choice of k. This can be referred to the greater complexity of the GPD quantile estimator as it involves, besides the tail index estimator $\hat{\xi}^{-1}$, a second estimated scale parameter $\hat{\beta}$ providing better finite sample performance.

In the literature on GARCH-EVT model, the most common value of threshold used is at the 90th percentile of the sample following the empirical example given in McNeil and Frey (2000). In addition, the 90th percentile threshold was found to be useful in finance and insurance (Chavez-Demoulin et al., 2014; Chavez-Demoulin & Embrechts, 2004).

On the other hand, Echaust and Just (2020) carried out a detailed study on the choice of threshold in which four different optimal tail selection algorithms were used including the path stability method, the Eye-ball method, the minimization of the asymptotic mean squared error method and the distance metric method with a mean absolute penalty function. These methods were used in order to estimate the out-of-sample VaR forecasts and compare them to the fixed threshold approach, the 95th quantile of the distribution. The results of this paper showed that the GARCH-EVT model performs well regardless of the choice of the threshold. Surprisingly, the same accuracy of VaR prediction is provided regardless of the choice of the tail. The authors found that the optimal tail selection methods do not improve the accuracy of the VaR forecast relative to the standard method of the fixed 95th quantile threshold.

***

McNeil and Frey (2000) list a number of advantages of the GPD approach which makes it more favorable than other approaches. These advantages can be summarized as follows:

- EVT quantile estimators, whether the maximum likelihood and GPD based or the Hill estimator, provide more efficient results than the historical simulation method in finite samples of order of 1000 points from typical return distributions.
- GPD based quantile estimator is more stable regarding the choice of k (threshold is ε_{k+1}) than the Hill estimator.
- For high quantiles deep in the tail, i.e., $q \geq 0.99$, the GPD method is as efficient as the Hill method.
- The GPD method is applicable for light-tailed ($\xi < 0$) and heavy-tailed data ($\xi > 0$) unlike the Hill estimator which is only applicable for the case of heavy-tailed distributions.

2.5. Monte Carlo Simulation

According to Hammersley and Handscomb (1964), the development of Monte Carlo (MC) simulation goes back to dates around 1944. According to Glasserman (2004), MC methods are based on an analogy between probability and volume in the sense that they use the notion of probability, which is formalized by mathematical measures, in reverse. In other words, the MC simulation defines the volume of an event by sampling randomly from a universe of possible outcomes and then takes the fraction of random draws that fall in a given set as an estimate of the set's volume. Then, by the law of large numbers, this estimate converges to the true value as the number of draws increases. Glasserman (2004), Hammersley and Handscomb (1964), among others, provide extensive description on the employment of MC simulation in VaR estimation and in the Risk Management industry in general.

MC simulation can be computationally demanding and time consuming as it requires highly developed systems, especially when dealing with large portfolios with hundreds of securities. However, its attractive feature is that it does not assume any simplifying distribution for the risk factors (like assuming normality) (Hull, 2018). According to Jorion (2007), MC simulation is by far the most robust technique for the estimation of VaR due to its malleability as it incorporates time variation in volatility or in expected returns, fat tails and extreme scenarios. In addition, it considers a broad range of risks among which are price risk, volatility risk, credit risk, and model risk.

The main idea behind the MC simulation is to find solutions through random number simulations for problems whose solutions are hard to find analytically.

A simple stepwise description for obtaining the MC VaR is tailored in Christoffersen (2011):

1. Assuming that the returns are independent identically distributed following a GARCH (1, 1)-normal model, at the end of day $t, t = 1 \dots, n$,

   $$r_t = \sigma_t \varepsilon_t, \text{ with } \varepsilon_t \sim i.i.d.N(0,1),$$

   in this case at the end of day t, tomorrow's variance of returns can be calculated according to GARCH (1, 1) variance equation,

   $$\sigma_{t+1}^2 = \alpha_0 \ + \alpha r_t^2 + \beta \sigma_t^2$$

2. Then a large amount of pseudo-random numbers of the error term ε_t is generated, say S, drawn from the standard normal distribution $N(0,1)$ maintaining consistency with the returns model assumption $r_t = \sigma_t \varepsilon_t$. The variable $\varepsilon_{s,t}$ denotes the s-th simulated value of the error on day t (for $s = 1, 2, \dots, S$) which implies that the s-th simulated value of the return can be given as $r_{s,t} = \sigma_t \varepsilon_{s,t}$.
3. In order to predict the volatility of the next period, $\sigma_{s,t+1}$, the returns $r_{s,t}$ are simulated and the volatility is obtained by the assumed GARCH model variance equation,

$$\left(\sigma_{s,t+1}\right)^2 = \alpha_0 + \alpha r_{s,t}^2 + \beta\sigma_t^2$$

then given the predicted variance the simulated returns in day $t+1$ are updated by

$$r_{s,t+1} = \sigma_{t+1}\varepsilon_{s,t+1}$$

where $\varepsilon_{s,t+1}$ is the set of simulated errors in day $t+1$. This iteration is continued till $\sigma_{s,t+k}$ and $r_{s,t+\mathrm{k}}$ are obtained for $k = 1, \dots, K$ where K is the forecast horizon.

4. After calculating the hypothetical K-day returns, the K-day VaR is calculated by estimating the quantile of the simulated returns.

The main advantages of MC simulation are that it is applicable to a variety of distributions of the market factors (returns) and does not rely on relatively small samples to estimate tail volatilities (Pearson, 2002). However, as per Hull (2018), the main drawback of MC simulation is the computational burden especially in the case of large portfolios with a big number of instruments and market factors, the MC simulation will be demanding and time consuming.

According to Abad et al. (2014), some methods provided more accurate VaR estimates than the MC simulation. Pérignon and Smith (2010) found, in 2005, that a 14% of a total of 60 banks worldwide surveyed for VaR disclosure were using MC simulation. However, although the trend for employing the MC simulation was increasing back then, it was reported in Mehta et al. (2012) that according to a survey done in 2011 on 18 financial institutions, only 15% were reported to be using the MC simulation while 75% used the Historical Simulation and the other 10% of the institutions were using hybrid VaR methods combining both parametric and non-parametric features.

The MC simulation has been widely used jointly with other models. It was adopted in McNeil and Frey (2000) for estimating conditional quantiles of returns when conditional volatility was estimated using GARCH model and the tail of the residuals distribution was estimated with extreme value theory (EVT). It was also employed in Yi et al. (2014) with a GARCH model having normal, fat-tailed and skewed residuals and also here the EVT was applied to the GARCH tail observations. It is worth noting that McNeil and Frey (2000) and Yi et al. (2014), among others, showed the efficiency of this approach. Moreover, Berkowitz et al. (2011) employ a comprehensive MC simulation to assess the performance of certain backtesting tests in order to check which one has the best finite-sample size and power properties.

2.6. Historical Simulation

The Historical Simulation (HS) approach is a non-parametric approach for estimating VaR. HS derives VaR forecasts in the simplest way possible without any theoretical assumption regarding the distribution of the asset returns as it only relies on the past historical observations of returns. This method uses the empirical distribution of returns to approximate the cumulative distribution function $F(r)$, and so rather than using the observations to calculate the standard deviation of a portfolio and then estimate VaR, here

$VaR(q)$ is nothing but the actual percentile of the observations over the pre-specified period of the empirical distribution (Hendricks, 1996).

The HS is widely used in estimating VaR and this is due to its simplicity and ease of implementation. As has been previously mentioned, according to Pérignon and Smith (2010), in 2005, 47.4 % of a total of 60 banks worldwide that were surveyed for VaR disclosure used the HS approach to estimate their VaR (73% if only the firms that disclosed their VaR methodology are considered).

Some developments were done on the HS in order to improve its accuracy. The work of McNeil and Frey (2000) combines the GARCH and EVT to obtain estimates for the conditional volatility and to estimate the distribution of residuals they use the HS (for the central part of the distribution) and threshold methods from EVT (for the tails). Moreover, in a previous work, Barone-Adesi at al. (1998) use HS and they model the volatility of their portfolio using an asymmetric GARCH process allowing for leverage effect (positive and negative returns having impact on volatility). They applied this technique on a hypothetical portfolio replicating a stock market index, and they obtained good results.

The HS has its advantages as well as its shortcomings. As per Dowd (2005), the main advantages of HS are its intuitive and conceptually simple nature, and since they do not rely on any assumptions regarding the parametric distribution of returns, it can accommodate fat-tails, skewness or any other features of non-normality which usually causes problems in parametric approaches. Moreover, HS is easy to implement, and its results are easy to report due to the simplicity of the approach. According to Christoffersen (2011), being model-free, HS has an advantage over model-based approaches because sometimes relying on a certain model might be a weakness if the model is poor. Regarding the shortcomings, according to Dowd (2005), HS and other non-parametric approaches might often produce VaR estimates which are too low compared to the actual risk if the data period under study is quiet, and conversely, if the data period considered is turbulent and witnesses unusual volatility then the VaR estimates might be overstated compared to the actual risk. However, being model free and completely dependent on the data set can also be a weakness, for instance this might cause a confusion regarding setting the length of the past data or sample window period, call it m, upon which the VaR will be forecasted; if m is too large, then the most recent observations which are most likely to have a bigger impact on tomorrow's return will carry little weight and on the other hand, if m was too small, then the data sample might not be large enough to include large losses to enable the precise estimation of the 1% VaR (Christoffersen, 2011).

A typical choice in practice for m is between 250 and 1,000 days (i.e., between 1 and 4 years approximately) according to Christoffersen (2011). In this study, $m = 250$ days is considered (250-day rolling window), coinciding approximately with the number of returns in one year.

Although HS has been widely used in the estimation of VaR, many references such as Abad and Benito (2013), Angelidis et al. (2007), Barone-Adesi & Giannopoulos (2001), among others, have reported the inaccuracy of its estimates especially when compared to other approaches. Moreover, Christoffersen (2011) uses evidence from the 2008 financial crisis to demonstrate the inaccuracy of VaR estimates provided by HS after the 1987 crash displayed a dramatic example of the problems embedded in HS approach to estimate VaR. In addition, according to Dowd (2005), non-parametric approaches, like the basic HS, work fairly well if the market conditions remain reasonably stable. However, based on their limitations, they should be supplemented with other approaches to improve their results and he suggested complementing them with stress testing, however, at the end he states that these approaches should never be used alone.

2.7. Filtered Historical Simulation

The Filtered Historical Simulation (FHS) was proposed independently by each of Diebold et al. (1998), Hull and White (1998) and also introduced in Barone-Adesi et al. (1998) and then fully developed in Barone-Adesi et al. (1999). In this section the approach of Barone-Adesi et al. (1999) is illustrated as a modification to the HS approach dealing with non-normality of returns along with heteroskedasticity of conditional variance and it will be applied in the next chapter.

Since HS relies on the data set and does not make any assumptions regarding the distribution of returns, and because it is well known that large returns cluster in time (Mandelbrot, 1963), Barone-Adesi et al. (1998) proposed the FHS in order to make the HS consistent with the clustering of large returns. The FHS is based on modeling the conditional variance of returns using a GARCH process. However, the residuals in the FHS method do not follow a specific distribution like normal or Student-t distribution as in MC simulation where the arbitrary choice of the distribution of residuals destroys any important information about the real distribution of returns. Contrary to that, the residuals in FHS are standardized by the volatility of the GARCH model estimated for the same date, and thus obtaining the “standardized residuals” or “filtered residuals” (hence a FHS) (Barone-Adesi & Giannopoulos, 2001).

Barone-Adesi et al. (1999) recommended using an asymmetric GARCH model, thus, inserting a moving average (MA) term in the conditional mean, when appropriate, to remove any serial dependency. Then the conditional mean of the distribution of returns can be written as

$$r_t = \mu r_{t-1} + \theta z_{t-1} + z_t$$

These authors propose the ARMA-GARCH model as an example of an asymmetric model to explain the application of the FHS where the conditional variance is given as

$$h_t = \alpha_0 + \alpha(z_{t-1} + \gamma)^2 + \beta h_{t-1}$$

In the equations of returns and conditional variance, z_t is a random residual such that $z_t \sim N(0, h_t)$, μ is the AR (1) term, θ is the MA term, and α_0 is a constant. The constant α determines the influence of the last observation and γ determines the effect of asymmetry. As a result, the standardized residuals, s_t, are given by

$$s_t = \frac{z_t}{\sqrt{h_t}}$$

The volatility scaling reduces the volatility clustering, making the returns more uniform and since under the GARCH hypothesis the standardized residuals are independently identically distributed, then they become more suitable for the HS, however empirical observations may deviate from this to some point (Barone-Adesi et al., 1999).

The standardized returns are the basis of the FHS approach. As per Barone-Adesi et al. (1998), the standardized residuals can be randomly drawn (with replacement) and used in the conditional mean and conditional variance equations to obtain the simulated pathways for prices and variances respectively. They explain their methodology in a step-wise manner for generating 10-day return paths as follows:

1. After obtaining the standardized residuals as per the equation of standardized residuals, a vector of standardized residuals s^* is drawn from the data set Θ such that
   $s^* = \{s_1^*, s_2^*, \ldots, s_T^*\} \in \Theta$ where $i = 1, \ldots, 10$ days
2. To get the simulated return at $t+1$, z_{t+1}^*, a random standardized residual return is drawn from the data set Θ and scale it with the volatility of period $t+1$, then
   $$z_{t+1}^* = s_1^* \sqrt{h_{t+1}}$$
3. Now the simulation of the pathway of the asset's price begins starting from the currently known asset price at time t. The simulated price p_{t+1}^* for $t+1$ is given by
   $$p_{t+1}^* = p_t + p_t\left(\hat{\mu} r_t + \hat{\theta} z_t^* + z_{t+1}^*\right)$$
   such that z_t^* and z_{t+1}^* are estimated according to the equation in step (2).
4. After obtaining the simulated prices, the volatility can then be simulated as well from the randomly selected simulated residuals (for $i = 2, 3, \ldots$). The simulated volatility, $\sqrt{h_{t+i}^*}$, for period $t+i$, is obtained according to the equation,
   $\sqrt{h_{t+i}^*} = \sqrt{\hat{\omega} + \hat{\alpha} z_{t+i+1}^* + \hat{\beta} h_{t+i+1}^*}$ for all $i \geq 2$ and z_{t+i+1}^* is obtained according to the equation in step (2).
5. Again, new elements are randomly drawn from Θ to obtain the simulated prices p_{t+i}^* as in the equation of step (3).
6. Finally, the empirical distribution of the simulated prices over the chosen time horizon (here $i = 10$) for a single asset is obtained by replicating the above steps for a large number of times like 5,000 or 10,000.

Christoffersen (2011) describes the FHS in a simple way in which he used the GARCH (1,1) model as a simple example for modeling the conditional variance. Considering the simple GARCH (1,1) model, the return, r_t is given as

$$r_t = \sigma_t \varepsilon_t$$

where the conditional variance equation is given as

$$\sigma_{t+1}^2 = \alpha_0 + \alpha r_t^2 + \beta \sigma_t^2$$

Taking the past sequence of returns $\{r_{t+1-\tau}\}_{\tau=1}^{m}$, the GARCH model is estimated, and the past standardized returns are calculated according to the equation,

$$\varepsilon_{t+1-\tau}^{*} = r_{t+1-\tau}/\sigma_{t+1-\tau}, \text{ for all } \tau = 1, \dots, m.$$

The set of standardized returns will be denoted by $\{\varepsilon_{t+1-\tau}^{*}\}_{\tau=1}^{m}$ and m, which is the number of historical observations, should be as large as possible.

A discrete uniformly distributed random variable is generated from 1 to m, and the draws from the discrete distribution determine which τ to pick and thus which $\varepsilon_{t+1-\tau}^{*}$ from the sequence $\{\varepsilon_{t+1-\tau}^{*}\}_{\tau=1}^{m}$ to choose. In order to adjust the standardized returns to market conditions, at the end of day t, a randomly selected (with replacement) residual $\varepsilon_{t+1-\tau}^{*}$ drawn from the sequence of past standardized residuals $\{\varepsilon_{t+1-\tau}^{*}\}_{\tau=1}^{m}$ is multiplied by GARCH forecast of tomorrow's volatility obtaining the simulated return $r_{t+1-\tau}^{*}$ for $\tau = 1, \dots, m$. Then, $r_{t+1-\tau}^{*}$ is used to estimate the GARCH volatility forecast for the next day, $\sigma_{t+2-\tau}^{*}$. It can be noticed here that the residuals used for the simulation are picked from the standardized returns instead of being drawn from a random generator which relies on a certain distribution (Normal, Student-t, …). Then $\sigma_{t+2-\tau}^{*}$ is multiplied with another randomly selected residual from the sequence $\{\varepsilon_{t+1-\tau}^{*}\}_{\tau=1}^{m}$ to obtain the simulated return which is then used again to update the GARCH volatility forecast for the following day, and so on. This procedure is repeated according to the length of the VaR horizon (1 day, 10 days, etc.…) generating a batch of sample paths of portfolio returns. Finally, $VaR(q)$ is calculated by estimating the q% quantile of the simulated return sample.

Christoffersen (2011) explains the procedure for building the distribution of hypothetical future returns in a map as below,

$$\begin{array}{ccccccccccccccc}
 & \nearrow & \varepsilon_{1,1}^{*} & \rightarrow & r_{1,t+1}^{*} & \rightarrow & \sigma_{1,t+2}^{*}, & \varepsilon_{1,2}^{*} & \rightarrow & r_{1,t+2}^{*} & \rightarrow & \sigma_{1,t+2}^{*} & \cdots & \varepsilon_{1,K}^{*} \rightarrow & \sigma_{1,t+K}^{*} \\
\sigma_{t+1}^{2} & \rightarrow & \varepsilon_{2,1}^{*} & \rightarrow & r_{2,t+1}^{*} & \rightarrow & \sigma_{2,t+2}^{*}, & \varepsilon_{2,2}^{*} & \rightarrow & r_{2,t+2}^{*} & \rightarrow & \sigma_{2,t+2}^{*} & \cdots & \varepsilon_{2,K}^{*} \rightarrow & \sigma_{2,t+K}^{*} \\
 & \searrow & \varepsilon_{FH,1}^{*} & \rightarrow & r_{FH,t+1}^{*} & \rightarrow & \sigma_{FH,t+2}^{*}, & \varepsilon_{FH,2}^{*} & \rightarrow & r_{FH,t+2}^{*} & \rightarrow & \sigma_{FH,t+2}^{*} & \cdots & \varepsilon_{FH,K}^{*} \rightarrow & \sigma_{FH,t+K}^{*}
\end{array}$$

where FH is the number of drawings from the standardized residuals or, in other words, the number of simulations performed, for example 5,000 or 10,000 and K is the time horizon according to which the VaR is measured ($K = 1, 10, \dots$). Thus, an FH sequence of simulated daily returns is obtained from day $t + 1$ through day $t + K$ and from these daily returns the hypothetical K-day returns can be collected as

$$r_{i,t+1:t+K}^{*} = \sum_{k=1}^{K} r_{i,t+K}^{*}$$

Accordingly, the K-day VaR is obtained by calculating the q quantile as follows,

$$VaR_{t+1:t+K}^{q} = -percentile\left(\left\{r_{i,t+1:t+K}^{*}\right\}_{i=1}^{FH}, 100q\right)$$

According to Pritsker (2006), who carried a powerful comparison between the FHS and the HS approach, the FHS is promising since it combines the parametric modelling of conditional volatility with the non-parametric modelling of residuals, moreover, the main advantages of implementing the FHS are that the volatility models are simple and there is no need to estimate the correlation matrix of the standardized residuals. As per Christoffersen (2011), FHS captures the current level of market volatility, and no assumptions are needed for the tails of the distribution.

Empirical evidence shows that FHS performs quite well in VaR estimation (Barone-Adesi & Giannopoulos, 2001). However, it has its limitations which are related to the assumptions on which it is based. These assumptions are that the correlations across asset returns are not related to the scale of returns and that tails of the distribution of asset returns are accurately described by the scaling process.

Alexander (2008) states that volatility adjusted historical VaR (like FHS) make the VaR estimates less sensitive to changes in sample size, and as a result the VaR estimates obtained should be more robust at any confidence level. According to Angelidis et al. (2007), compared to other models, FHS is the best for estimating VaR due to its robustness across sub-samples, stock portfolio and confidence level as the average exception rates obtained were very close to the expected ones.

Moreover, the FHS has been developed and extended to include the use of more complex models for the conditional variance like in Audrino and Barone-Adesi (2005) where they use a multivariate GARCH-type model in connection with functional gradient descent for filtering the residuals. And in a univariate setting, McNeil and Frey (2000) also extended the FHS to incorporate modeling the tails of the distribution of residuals using EVT.

2.8. Variance Covariance

The variance-covariance (var-cov) method for estimating VaR, also known as parametric method and sometimes Delta-normal method, is based on estimating the mean and standard deviation of the portfolio of assets and then calculating VaR accordingly. In other words, it can be viewed as fitting probability curves to data and then inferring VaR from the fitted curve (Abad et al., 2014).

The var-cov approach is perhaps the simplest and most widely used approach among the models for estimating VaR (Glasserrnan, 2004). It gained more popularity because the *RiskMetrics*™, which is employed by a large number of financial institutions in attempt to develop their own in-house models, follows this methodological approach (Resti & Sironi, 2007).

Resti and Sironi (2007) describe the var-cov approach as the easiest and traditionally most widespread approach for estimating VaR, characterized by the following:

- Changes in asset returns follow a normal distribution.
- The variance-covariance matrix summarizes all the possible future values of market factors (here asset returns) and their correlations.
- Consequently, the possible losses of the bank's portfolio depend on this matrix and on the sensitivity (usually approximated by a linear function with constant coefficients) of the individual positions in the portfolio to changes in asset returns.
- VaR is obtained as a multiple of the standard deviation of future losses.

According to Hendricks (1996), the var-cov approach is based on three simplifying assumptions: (i) normality of returns, (ii) serial independence and (iii) the absence of non-linear positions (such as options). The normality assumption along with the serial independence are the main reason behind the simplicity and the ease of use of var-cov methodology. The assumption of normality makes the calculation of VaR easier because all percentiles, on which VaR calculation is based, are assumed to be known multiples of the standard deviation, and the serial independence implies that the size of price change on one day will not affect the price moves on any other day.

In order to estimate VaR, one ought to first obtain the log returns of the assets as follows,

$$r_t = ln\left(\frac{p_t}{p_{t-1}}\right),$$

where p_t is the price value of the asset at day t.

Given Ω_{t-1}, the information set available at time $t-1$, it is assumed that the daily returns $\{r_t\}$ with cumulative distribution function $F(r) = \Pr\,(r_t < r|\Omega_{t-1})$ follow the stochastic process,

$$r_t = \mu_t + z_t \text{ where } z_t = \sigma_t\varepsilon_t \text{ and } \varepsilon_t \sim N(0,1)$$

where $\sigma_t^2 = E(\varepsilon_t^2|\Omega_{t-1})$ and $G(\varepsilon) = Pr(\varepsilon_t < \varepsilon|\Omega_{t-1})$ is the standard normal distribution function of ε_t taking the assumption that ε_t are standard normally distributed.

Thus, VaR at confidence level $(1-q)\%$ is the q quantile of the probability distribution of financial returns,

$$F\big(VaR(q)\big) = Pr\big(r_t < VaR(q)\big) = q$$

Then, VaR can be estimated either by using the inverse of F, F^{-1}, or the inverse of G, G^{-1}, as follows,

$$VaR(q) = F^{-1}(q) = \mu_t + \sigma_t G^{-1}(q)$$

Seen another way, as derived by Alexander (2008), by applying the standard normal transformation it can be written,

$$F(VaR(q)) = Pr(r_t < VaR(q)) = \Pr\left(\frac{r_t - \mu_t}{\sigma_t} < \frac{VaR(q) - \mu_t}{\sigma_t}\right)$$
$$= \Pr\left(\varepsilon_t < \frac{VaR(q) - \mu_t}{\sigma_t}\right) = q = G\left(\frac{VaR(q) - \mu_t}{\sigma_t}\right)$$

since ε_t is the standard normal variable and $G(\cdot)$ is the standard normal distribution function. Then,

$$G\left(\frac{VaR(q) - \mu_t}{\sigma_t}\right) = q$$

And so,

$$\frac{VaR(q) - \mu_t}{\sigma_t} = G^{-1}(q)$$

where $G^{-1}(q)$ is simply the standard normal q quantile value.

Then, in order to calculate VaR, the expression is written as,

$$VaR(q) = \mu_t + \sigma_t G^{-1}(q).$$

If μ_t is assumed significantly small and close to zero then,

$$VaR(q) = \sigma_t G^{-1}(q)$$

Thus, the VaR (as amount of loss) of a single asset can be written as

$VaR = V_t \sigma_t G^{-1}(q)$ where V_t is the value of the asset on day t.

The above model estimates the VaR for one single asset, however, if VaR is to be estimated for a portfolio of *N* assets then,

$$VaR_N = V_N \sigma_N G^{-1}(q) = V_N \sqrt{w' \Sigma\, w}\, G^{-1}(q)$$

where V_N is the value of the portfolio of *N* assets, and σ_N is the volatility of the portfolio returns, w is the weight vector of the assets of the portfolio with an N by 1 dimension and Σ is the variance-covariance matrix with N by N dimension. It is obvious from the simple straight forward derivation of VaR under the var-cov method that its main advantage is its simplicity. According to Resti and Sironi (2007), the main advantages of the var-cov approach that stand behind its widespread are, first its simplicity compared to simulation approaches, and second is that it is the original version of VaR models which was first developed and widely spread among the Anglo-Saxon banks. Finally, with the presence of the RiskMetrics™ approach which was first developed by J.P. Morgan, the use of var-cov methodology became more popular, since RiskMetrics™ is employed by a large number of products developed by the software industry.

Although it is simple and easy to implement however the var-cov approach has its drawbacks which affect the accuracy of the obtained VaR estimates. Its main weaknesses are related to its theoretical hypotheses. A large body of evidence suggest that the tails of distributions of daily percentage changes in financial market prices are fatter than those tails predicted by the normal distribution (Hendricks, 1996). This affects the accuracy of models which are based on the assumption of normality of the distribution of returns like var-cov method. Another disadvantage of the var-cov model is that the portfolio positions are sensitive to changes in market factors (Resti & Sironi, 2007).

2.9. Exponentially Weighted Moving Average

The Exponentially Weighted Moving Average (EWMA) was highly promoted in the RiskMetrics model for VaR estimation introduced by Morgan in 1996. The RiskMetrics from J.P. Morgan (1996) was the first model introduced for estimating VaR under the parametric approach (Abad et al., 2014). The main point of the EWMA is to estimate conditional volatilities and correlations by placing exponentially declining weights on past observations, thus assigning greater weight to recent returns and at the same time not fully neglecting distant past observations, and this represents a reasonable tradeoff between statistical precision and adaptiveness to recent news (Boudoukh et al., 1998).

In this thesis, the EWMA is used to estimate the volatility of returns and estimate VaR using the traditional parametric approach keeping the assumption of normality of returns. Assuming

$$r = \mu_t + z_t \text{ where } E(z_t) = 0 \text{ and } E(z_t^2) = 1 \text{ such that } z_t \sim iid\ N(0,1),$$

then, according to the EWMA, the estimate of today's variance is given by,

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

where λ is the decay factor which determines the weight assigned to recent and older observations. It is worth noting that as the value of λ is maximized, the rate of decay of weights slows down thus giving more weight to distant observations. Then VaR can be written as

$$VaR(q) = \mu_t + \sigma_t G^{-1}(q)$$

where $G^{-1}(q)$ is the standard normal q quantile value, μ_t is the daily mean of returns (most of the time is assumed to be zero for simplification) and σ_t is the standard deviation obtained as per the EWMA approach.

Choosing the decay factor

The choice of the optimal decay factor is discussed by J.P. Morgan (1996) in the RiskMetrics Technical Document, furthermore, it is shown how the optimal decay factors $\lambda = 0.94$ for the daily data set and $\lambda = 0.97$ for the monthly data set are obtained.

It is explained how the root mean squared error (RMSE) is used to obtain the optimal decay factor. By definition, the time $t+1$ forecast of the variance of the return r_{t+1}, made one period earlier, is the expected value of the squared return one period earlier, written as

$$E_t[r_{t+1}^2] = \sigma_{t+1|t}^2.$$

The variance forecast error is then defined as

$$\varepsilon_{t+1|t} = r_{t+1}^2 - \sigma_{t+1|t}^2.$$

Then, the expected value of the forecast error is zero since $E_t\left[\varepsilon_{t+1|t}\right] = E_t[r_{t+1}^2] - \sigma_{t+1|t}^2 = 0$.

Accordingly, the choice of the optimal decay factor, λ, is based on a natural requirement which is to minimize the average squared errors. Applying this to daily forecasts of variance gives the mean square root prediction error which is given as

$$RMSE_v = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(r_{t+1}^2 - \sigma_{t+1|t}^2(\lambda)\right)^2}$$

where the forecast value of the variance is written explicitly as a function of λ.

A similar definition for the $t+1$ covariance forecast between two return series, $r_{1,t+1}$ and $r_{2,t+1}$, made one period earlier is given by J.P. Morgan (1996) as

$$E_t\left[r_{1,t+1}r_{2,t+1}\right] = \sigma_{12,t+1|t}^2$$

This holds true for any forecast made at time $t+j, j \geq 1$.

In practice, the optimal decay factor is found by searching for the smallest value of the RMSE for different values of λ. This means that the optimal decay factor is that with best forecasts, i.e., minimizes the forecast measures.

RiskMetrics derives the same covariance forecast error as $\varepsilon_{12,t+1|t} = r_{1,t+1}r_{2,t+1} - \sigma_{12,t+1|t}^2$ such that $E_t\left[\varepsilon_{12,t+1|t}\right] = E_t\left[r_{1,t+1}r_{2,t+1}\right] - \sigma_{12,t+1|t}^2 = 0$, and then

$$RMSE_c = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(r_{1,t+1}r_{2,t+1} - \hat{\sigma}_{12,t+1|t}^2(\lambda)\right)^2}$$

It is worth noting that the above measures are purely statistical which might not be optimal for risk management purposes because other factors play a role in determining the best forecast. For instance, for some risk managers who do not update their systems on a daily basis, they need optimal decay factors should allow enough stability in the variance and covariance forecasts.

The EWMA gained its popularity mainly because of the RiskMetrics model of J.P. Morgan. Boudoukh et al. (1998) used the EWMA included in the RiskMetrics approach combined with HS and proposed the hybrid approach which estimates VaR by assigning exponentially declining weights to past returns and then finding the quantile of the time-weighted empirical distribution as per the HS. However, Guermat and Harris (2002) showed that the EWMA variance estimator can be considered as a particular case of a more general exponentially weighted maximum likelihood (EWML) procedure which allows for time-variation not only in the variance but also in higher moments like kurtosis of returns. Later, Lin et al. (2006) extend the Guermat and Harris (2002) model known as the Power EWMA in conjunction with the HS to estimate portfolio VaR. These studies among others show the wide utility of EWMA in the estimation of VaR.

According to Resti and Sironi (2007), the two major benefits of using the EWMA for estimating volatility are that the estimate will respond to market factor shocks more rapidly and at the same time, a marked market factor shock will leave the affected volatility estimate gradually avoiding an echo effect. In addition, the logic of EWMA which assigns greater weight on more recent observations gives the volatility estimate obtained high information content and makes it more sensitive to recent shocks. However, the EWMA also has its drawbacks which include the assumption of normality imposed on the returns which is extensively discussed in literature to be inaccurate. Although EWMA captures some non-linear characteristics of volatility like its variation and clustering however, its inability to capture asymmetry (Pagan & Schwert, 1990) and the leverage effect are considered as another drawbacks. Moreover, Abad et al. (2014) note that the EWMA is technically inferior to GARCH family models in modelling persistence of volatility.

In the next chapter, all the aforementioned models are applied to the Dow Jones Industrial Average, Nikkei and EuroStoxx 50 indices, in a 15-year period divided into three intervals that represent the period before, during and after the 2008 financial crisis. The 1-day ahead VaR forecasts at the 99% and 95% confidence levels are estimated using a 250-day rolling window. A preliminary performance analysis is carried out to evaluate the efficiency of the models in estimating VaR using distance measures between the obtained VaR results and the actual incurred returns (losses) More advanced backtesting measures are employed later in Chapter 4 to further validate the performance of the previously implemented models. In a graphical way, the models that has been detailed previously and that will be applied in the next chapter are the following:

GARCH (1,1) model

- GARCH (1,1) with normal residuals
- GARCH (1,1) with Student-t residuals using the fitting method
- GARCH (1,1) with Student-t residuals using the Excess Kurtosis (EK) method

Extreme Value Theory (EVT)

- EVT Peaks Over Threshold model, with 95% threshold
- EVT Peaks Over Threshold model with 90% threshold

GARCH - EVT

- GARCH-EVT QMLE 95% threshold
- GARCH-EVT QMLE 90% threshold
- GARCH-EVT Normal 95% threshold
- GARCH-EVT Normal 90% threshold
- GARCH-EVT Student-*t* 95% threshold
- GARCH-EVT Student-*t* 90% threshold

Monte Carlo Simulation (MCS)

- MCS with normal residuals
- MCS with Student-*t* residuals fitting method
- MCS with Student-*t* residuals EK method

Historical Simulation

Filtered Historical Simulation (FHS)

- FHS with normal residuals
- FHS with Student-*t* residuals

Exponentially Weighted Moving Average (EWMA)

- EWMA with $\lambda = 0.94$
- EWMA with $\lambda = 0.95$
- EWMA with $\lambda = 0.96$
- EWMA with $\lambda = 0.97$
- EWMA with $\lambda = 0.98$
- EWMA with $\lambda = 0.99$

Variance-Covariance Model

References

Abad, P., & Benito, S. (2013). A detailed comparison of value at risk estimates [Article]. *Mathematics and Computers in Simulation*, *94*, 258–276. https://doi.org/10.1016/j.matcom.2012.05.011

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

Alexander, Carol. (2008). *Market risk analysis. Volume IV, Value-at-risk-models* (1st edition) [Book]. Wiley.

Angelidis, T., Benos, A., & Degiannakis, S. (2007). A robust VaR model under different time periods and weighting schemes. *Review of Quantitative Finance and Accounting*, *28*(2), 187–201. https://doi.org/10.1007/s11156-006-0010-y

Audrino, F., & Barone-Adesi, G. (2005). A multivariate FGD technique to improve VaR computation in equity markets. *Computational Management Science*, *2*(2), 87–106. https://doi.org/10.1007/s10287-004-0028-3

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*(1), 3–30. https://doi.org/10.1016/S0304-4076(95)01749-6

Balkema, A. A., & de Haan, L. (1974). Residual Life Time at Great Age [Article]. *The Annals of Probability*, *2*(5), 792–804. https://doi.org/10.1214/aop/1176996548

Barone-Adesi, G., Bourgoin, F., & Giannopoulos, K. (1998). Don't look back. In *Risk* (Vol. 11).

Barone-Adesi, G., & Giannopoulos, K. (2001). Non-parametric VaR techniques. Myths and realities. *Economic Notes*, *30*(2), 167–181. https://doi.org/10.1111/j.0391-5026.2001.00052.x

Barone-Adesi, G., Giannopoulos, K., & Vosper, L. (1999). VaR without correlations for portfolios of derivative securities. *Journal of Futures Markets*, *19*(5), 583–602. https://doi.org/10.1002/(SICI)1096-9934(199908)19:5<583::AID-FUT5>3.0.CO;2-S

Benito, S., López-Martín, C., & Navarro, M. ªÁ. (2023). Assessing the importance of the choice threshold in quantifying market risk under the POT approach (EVT). *Risk Management*, *25*(1). https://doi.org/10.1057/s41283-022-00106-w

Berkowitz, J., Christoffersen, P., & Pelletier, D. (2011). Evaluating Value-at-Risk Models with Desk-Level Data [Article]. *Management Science*, *57*(12), 2213–2227. https://doi.org/10.1287/mnsc.1080.0964

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

Boudoukh, J., Richardson, M. P., & Whitelaw, R. (1998). The best of both worlds: A hybrid approach to calculating value at risk. *SSRN(51420)*, 1–12.

Brandt, A. (1986). The Stochastic Equation $Y_{n+1}=A_nY_n+B_n$ with Stationary Coefficients. *Advances in Applied Probability*, *18*(1), 211–220. https://doi.org/10.2307/1427243

Byström, H. N. E. (2004). Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, *13*(2), 133–152. https://doi.org/10.1016/j.irfa.2004.02.003

Chavez-Demoulin, V., & Embrechts, P. (2004). Smooth extremal models in finance and insurance. *Journal of Risk and Insurance*, *71*(2), 183–199. https://doi.org/10.1111/j.0022-4367.2004.00085.x

Chavez-Demoulin, V., Embrechts, P., & Sardy, S. (2014). Extreme-quantile tracking for financial time series. *Journal of Econometrics*, *181*(1), 44–52. https://doi.org/10.1016/j.jeconom.2014.02.007

Christoffersen, P. (2011). *Elements of Financial Risk Management, 2nd Edition* (2nd ed.) [Book]. Academic Press.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, *1*(2), 223–236. https://doi.org/10.1080/713665670

Daníelsson, J., & Vries, C. G. D. (1997). Beyond the sample: Extreme quantile and probability estimation. *Tinbergen Institute Discussion Papers*.

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *J. R. Statist. Soc. B*, *52*(3), 393–442.

de Haan, L., & Ferreira, A. (2006). Extreme Value Theory: An Introduction. In *Springer Series in Operations Research and Financial Engineering*.

Diebold, F. X., Hickman, A., Inoue, A., & Schuermann, T. (1998). Scale models. *Risk*, *11*(11), 104–107.

Dowd, K. (2005). *Measuring market risk* (2nd ed) [Book]. John Wiley & Sons.

Duan, J.-C. (1997). Augmented GARCH (p, q) process and its diffusion limit [Article]. *Journal of Econometrics*, *79*(1), 97–127. https://doi.org/10.1016/S0304-4076(97)00009-2

DuMouchel, W. H. (1983). Estimating the Stable Index α in Order to Measure Tail Thickness: A Critique. *The Annals of Statistics*, *11*(4), 1019–1031. https://doi.org/10.1214/aos/1176346318

Echaust, K., & Just, M. (2020). Value at risk estimation using the garch-evt approach with optimal tail selection. *Mathematics*, *8*(1). https://doi.org/10.3390/math8010114

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance* (C. Klüppelberg & Thomas. Mikosch, Eds.; 1st ed.) [Book]. Springer. https://doi.org/10.1007/978-3-642-33483-2

Embrechts, P., Resnick, S., & Samorodnitsky, G. (1998). Living on the Edge. *Risk Magazine*, *11*(1), 96–100.

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

Engle, R. F. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, *15*(4), 157–168. https://doi.org/10.1257/jep.15.4.157

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1–50. https://doi.org/10.1080/07474938608800095

Gençay, R., & Selçuk, F. (2004). Extreme value theory and Value-at-Risk: Relative performance in emerging markets. *International Journal of Forecasting*, *20*(2), 287–303. https://doi.org/10.1016/j.ijforecast.2003.09.005

Glasserrnan, P. (2004). *Monte Carlo Menthods in Financial Enngineering*. Springer-Verlag New York, Inc.

Gnedenko, B. (1943). Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire. *Annals of Mathematics*, *44*(3), 423–453. https://doi.org/10.2307/1968974

Guermat, C., & Harris, R. D. F. (2002). Forecasting value at risk allowing for time variation in the variance and kurtosis of portfolio returns. *International Journal of Forecasting*, *18*(3), 409–419. https://doi.org/10.1016/S0169-2070(01)00122-4

Hammersley, J. M. (John M., & Handscomb, D. C. (David C. (1964). *Monte Carlo Methods* (D. C. (David C. Handscomb, Ed.) [Book]. Methuen.

Harmantzis, F. C., Miao, L., & Chien, Y. (2006). Empirical study of value-at-risk and expected shortfall models with heavy tails. *Journal of Risk Finance*, *7*(2), 117–135. https://doi.org/10.1108/15265940610648571

Hendricks, D. (1996). Evaluation of value-at-risk models using historical data. *Federal Reserve Bank of New York Economic Policy Review*, *2*(APRIL), 39–69. https://doi.org/10.2139/ssrn.1028807

Higgins, M. L., & Bera, A. K. (1992). A class of nonlinear ARCH models. *International Economic Review*, *33*(1), 137–158.

Hill, B. M. (1975). A Simple General Approach to Inference About the Tail of a Distribution [Article]. *The Annals of Statistics*, *3*(5), 1163–1174. https://doi.org/10.1214/aos/1176343247

Hosking, J. R. M., & Wallis, J. R. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, *29*(3), 339–349. https://doi.org/10.1080/00401706.1987.10488243

Hull, J. (2018). *Risk management and financial institutions* (Fifth edition.) [Book]. Wiley.

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

Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*. McGraw-Hill.

Kuester, K., Mittnik, S., & Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, *4*(1), 53–89. https://doi.org/10.1093/jjfinec/nbj002

Langousis, A., Mamalakis, A., Puliga, M., & Deidda, R. (2016). Threshold detection for the generalized Pareto distribution: Review of representative methods and application to the NOAA NCDC daily rainfall database. *Water Resources Research*, *52*(4), 2659–2681. https://doi.org/10.1002/2015WR018502

Lewis, N. D. Costa. (2012). *Market risk modelling : applied statistical methods for practitioners*. Risk Books.

Lin, C.-H., Chien, C.-C. C., & Chen, S. W. (2006). Incorporating the time-varying tail-fatness into the historical simulation method for portfolio value-at-risk. *Review of Pacific Basin Financial Markets and Policies*, *9*(2), 257–274. https://doi.org/10.1142/S0219091506000720

Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking and Finance*, *24*(7), 1097–1130. https://doi.org/10.1016/S0378-4266(99)00077-1

MacDonald, A., Scarrott, C. J., Lee, D., Darlow, B., Reale, M., & Russell, G. (2011). A flexible extreme value mixture model. *Computational Statistics and Data Analysis*, *55*(6), 2137–2157. https://doi.org/10.1016/j.csda.2011.01.005

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, *36*(4), 394–419. http://www.jstor.org.sabidi.urv.cat/stable/2350970

McNeil, A. J. (1998). Calculating quantile risk measures for financial return series using extreme value theory. *ETH Working Paper*.

McNeil, A. J. (1999). Extreme Value Theory for Risk Managers. In *Internal modelling and CAD II* (pp. 23–43). London: Risk Books. http://www.sfu.ca/~rjones/econ811/readings/McNeil 1999.pdf

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3–4), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8

McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative risk management: Concepts, techniques and tools: Revised edition. In *Quantitative Risk Management: Concepts, Techniques and Tools: Revised Edition*. Princeton University Press. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84937787962&partnerID=40&md5=4546a3521a1b5468f3f969c8d2517e01

Mehta, A., Neukirchen, M., Pfetsch, S., & Poppensieker, T. (2012). Managing market risk: Today and tomorrow. *McKinsey & Company McKinsey Working Papers on Risk*, *32*(1), 24–36.

https://www.mckinsey.com/~/media/mckinsey/business%20functions/risk/our%20insights/managing%20market%20risk%20today%20and%20tomorrow/managing%20market%20risk.pdf

Morgan, J. P. (1996). *RiskMetrics-Technical Document* (Fourth Edition). J.P. Morgan/Reuters. https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, *59*(2), 347–370.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, *45*(1–2), 267–290. https://doi.org/10.1016/0304-4076(90)90101-X

Pearson, N. D. (2002). *Risk Budgeting: Portfolio Problem Solving with Value-at-Risk*. John Wiley & Sons, Inc.

Pérignon, C., & Smith, D. R. (2010). The level and quality of Value-at-Risk disclosure by commercial banks. *Journal of Banking and Finance*, *34*(2), 362–377. https://doi.org/10.1016/j.jbankfin.2009.08.009

Pickands III, J. (1975). Statistical Inference Using Extreme Order Statistics [Article]. *The Annals of Statistics*, *3*(1), 119–131. https://doi.org/10.1214/aos/1176343003

Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking and Finance*, *30*(2), 561–582. https://doi.org/10.1016/j.jbankfin.2005.04.013

Resti, A., & Sironi, A. (2007). Risk Management and Shareholders' Value in Banking. From Risk Measurement Models to Capital Allocation Policies. *Risk Management and Shareholders' Value in Banking*.

Rocco, M. (2014). EXTREME VALUE THEORY IN FINANCE: A SURVEY [Article]. *Journal of Economic Surveys*, *28*(1), 82–108. https://doi.org/10.1111/j.1467-6419.2012.00744.x

Scarrott, C., & MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *Revstat Statistical Journal*, *10*(1), 33–60.

Shayya, R., Sorrosal-Forradellas, M. T., & Terceño, A. (2023). Value-at-risk models: a systematic review of the literature. *Journal of Risk*, *25*(4), 1–23. https://doi.org/10.21314/JOR.2022.053

Smith, R. L. (1987). Estimating Tails of Probability Distributions [Article]. *The Annals of Statistics*, *15*(3), 1174–1207. https://doi.org/10.1214/aos/1176350499

Yi, Y., Feng, X., & Huang, Z. (2014). Estimation of extreme value-at-risk: An EVT approach for quantile GARCH model. *Economics Letters*, *124*(3), 378–381. https://doi.org/10.1016/j.econlet.2014.06.028

Chapter 3

Estimating VaR

3.1. Data

In this chapter we present the application procedure of the models described in Chapter 2. The data on which the models are applied is described along with the relative statistical description while analyzing some key aspects which affect the model implementation techniques.

The VaR models described previously in Chapter 2 are applied to the data which constitutes the prices of the stock indices Dow Jones Industrial Average (DJIA), Euro Stoxx 50 (SX5E) and Nikkei 225 (N225), with a total of 4,696 observations for each stock index. This data is retrieved from Eikon database for the period between 2002 and 2019 which is later divided into three intervals of equal magnitude (5-year intervals). These intervals represent the time before, during and after the financial crisis of 2008 and are divided as follows: (i) [2002, 2007] with 1,565 observations, (ii) [2008, 2013] with 1,566 observations and (iii) [2014, 2019] with 1,565 observations. The main reason behind splitting the data in such a way is to compare the performance of all the models during calm and volatile periods.

The daily logarithmic returns of the indices are obtained from the daily prices according as

$$r_t = ln(P_t/P_{t-1})$$

where P_t is the price of the index in the day t and r_t its logarithmic return. Table 3.1 illustrates the descriptive statistics of the log returns of the DJIA, SX5E and N225 indices for each subperiod and for the whole period as well.

		[2002, 2007]	[2008, 2013]	[2014, 2019]	[2002, 2019]
Dow Jones Industrial Average	**Maximum**	0.06154722	0.1050835	0.04864331	**0.1050835**
	Minimum	-0.04751468	-0.08200513	-0.04714278	**-0.08200513**
	Mean	0.000179273	0.000142416	0.000347351	**0.000222901**
	Standard Deviation	0.009640121	0.01376961	0.008138655	**0.01077877**
	1st Quartile	-0.004508276	-0.004681962	-0.002876674	**-0.003932641**
	Median	0.000178325	0.00030117	0.000342	**0.00029547**
	3rd Quartile	0.005055668	0.005755084	0.004400189	**0.004999881**
	Semi-Deviation	0.006770231	0.009964219	0.006071484	**0.007793334**
	Skewness	0.217095	-0.035792	-0.523452	**-0.05701**
	Kurtosis	6.854087	11.92227	6.998303	**12.78845**
Euro Stoxx 50	**Maximum**	0.0707832	0.1043765	0.04598324	**0.1043765**
	Minimum	-0.06331463	-0.08207879	-0.0901113	**-0.0901113**
	Mean	9.27E-05	-0.000221878	0.000119029	**-3.43981E-06**
	Standard Deviation	0.01364946	0.01687231	0.01067734	**0.01395969**
	1st Quartile	-0.005979706	-0.00795672800	-0.004891087	**-0.006236857**
	Median	0.000115919	-0.00007145345	0.000296928	**6.35696E-06**
	3rd Quartile	0.006441099	0.00850636000	0.005457825	**0.006464799**
	Semi-Deviation	0.009753988	0.01200560000	0.00786603	**0.01003028**
	Skewness	-0.003484	0.081111	-0.563954	**-0.048536**
	Kurtosis	7.062574	7.906292	8.058382	**8.680707**

Nikkei 225	**Maximum**	0.05735232	0.1323458	0.0742617	**0.1323458**
	Minimum	-0.05569546	-0.1211103	-0.08252932	**-0.1211103**
	Mean	0.00023845	0.000039790	0.000238498	**0.000172144**
	Standard Deviation	0.01236024	0.0175287	0.01195773	**0.01417272**
	1st Quartile	-0.006174119	-0.008020417	-0.0046466850	**-0.006185385**
	Median	0	0	0.0000969551	**0**
	3rd Quartile	0.007185235	0.009272117	0.0058639430	**0.007320683**
	Semi-Deviation	0.008959211	0.01301273	0.008770105	**0.01043807**
	Skewness	-0.217832	-0.582019	-0.293027	**-0.480886**
	Kurtosis	4.290057	11.04966	8.756571	**10.92572**

Table 3.1. Descriptive Statistics of the daily log returns for DJIA, SX5E and N225 between 2002 and 2019 and for each subperiod.

The high kurtosis noticed for all the indices displays a fat-tail distribution of returns, indicating that the assumption of normality adopted in some models might affect the quality of the corresponding VaR estimations. However, on the other hand, the low skewness observed among all indices, close to zero, evidence that the distributions of returns are not far from symmetric, and this justifies the use of the normal and Student-*t* distributions. The skewness is always negative in almost all the intervals; it is only positive in the interval [2002, 2007] for DJIA and in the [2008, 2013] for SX5E. Although the descriptive analysis shows that the data is far from normal, however, the normal distribution is used in the model application since it is highly employed in the literature as well as in practice. Using the normal distribution against another symmetric distribution with heavier tails will help confirm the weakness of the assumption of normality in estimating VaR.

3.2. Models

The models applied vary between: (i) the fully parametric, like the variance-covariance, EWMA and GARCH (1,1) model, which are based on econometric models for volatility dynamics and the assumption of conditional normality; (ii) non-parametric methods like the historical simulation model which is a model free method, (iii) semi-parametric methods which combine non-parametric and parametric approaches like MCS and FHS and the methods based on EVT like the EVT model and GARCH-EVT model.

In this study, the 1-day VaR of returns of the indices is estimated over a 250-day rolling window in each interval. Now the parameters used in the implementation of each model are described, where applicable. Some of the models are implemented using Microsoft Excel like var-cov, EWMA and HS while the remaining models are applied in R software. The R code designed in this thesis is based on the R codes provided in Manzan (2017) and Singh (2017) modified and adjusted to adapt to the thesis data and models' frameworks and parameters.

3.2.1. Variance-Covariance model

There are no specific parameters for this model, it assumes that the returns are normally distributed. It constitutes of estimating VaR by adding the mean to the standard deviation of the returns multiplied by the inverse cumulative normal distribution function of the confidence level chosen (99% or 95%).

3.2.2. Exponentially Weighted Moving Average

The main parameter in the implementation of the EWMA model is the decay factor lambda, λ. RiskMetrics determine in J.P. Morgan (1996) the optimal decay factor by country for equity indices. Accordingly, and following the Root Mean Squared Error (RMSE) criterion used by RiskMetrics to choose the optimal decay factor, the decay factors (among 0.94, 0.95, 0.96, 0.97, 0.98 and 0.99) that yield the minimal RMSE values are used.

	λ	**0.94**	**0.95**	**0.96**	**0.97**	**0.98**	**0.99**
DJIA	2002 - 2007	0.000117987	**0.000117928**	0.000117948	0.000118125	0.00011875	0.000121513
	2008 - 2013	0.000276405	**0.000013441**	0.000014179	0.000015257	0.000281819	0.000020896
	2014 - 2019	**0.000165749**	0.000166236	0.000166892	0.000167772	0.000168941	0.000170587
SX5E	2002 - 2007	**0.000257081**	0.000257624	0.000258443	0.000259699	0.000261936	0.000268484
	2008 - 2013	**0.000476862**	0.000477074	0.000477764	0.000479233	0.000479097	0.000488321
	2014 - 2019	0.000310551	**0.000310524**	0.00031056	0.000310675	0.000310952	0.000312023
N225	2002 - 2007	0.000242381	**0.000242324**	0.000242415	0.000242703	0.000243312	0.000244741
	2008 - 2013	**0.000511688**	0.000512166	0.000512917	0.000514159	0.00051659	0.000523915
	2014 - 2019	**0.000405951**	0.000406099	0.000406384	0.000406828	0.000407556	0.000409314

Table 3.2. RMSE values for DJIA, SX5E and N225 indexes applying EWMA for different values of λ

Consequently, the results in Table 3.2. revealed that the best forecast measures were obtained under the values of $\lambda = 0.94$ and $\lambda = 0.95$. This confirms the optimal decay factor value proposed by RiskMetrics of $\lambda = 0.94$ for a daily forecast horizon. In this study, EWMA is therefore applied with both values of lambda and the obtained results are evaluated and compared.

3.2.3. Historical Simulation

The HS model is the only non-parametric model applied in this study. It is model free and thus has no parameters to be determined. It simply estimates VaR by calculating the quantile of the previous observations of returns by a 250-day rolling window.

3.2.4. Filtered Historical Simulation

The FHS model is a semi-parametric model combining parametric and non-parametric model characteristics. The FHS model in this study is applied in two frameworks: (i) employing the simple but effective GARCH (1,1) model with normally distributed residuals and (ii) also employing the GARCH (1,1) model but with Student-t distributed residuals. Fitting the data to GARCH (1,1)-Student-t model in R is done using the model fitting function from the "rugarch" package. It can be noticed that, unlike the case of MCS, the EK method for fitting the GARCH-Student-t model is not used in the FHS model. In the MCS, where the returns are supposed to follow a GARCH-Student-t model with a certain degree of freedom, EK method is used to compute this degree of freedom "d". Then, a set of random errors is generated from the $t(d)$ distribution to generate the simulated returns which are used to estimate VaR. However, in the

FHS, the obtained standardized residuals are used to generate the simulated returns and there is no need to calculate the degree of freedom of those residuals, thus, using the fitting method to obtain the residuals is enough.

3.2.5. Monte Carlo Simulation

MCS is another semi-parametric model. Like in FHS, it is also applied in two frameworks:

1. GARCH (1,1)-normal model and
2. GARCH (1,1)-Student-t framework, implemented using two methods: (a) fitting method and (b) the Excess Kurtosis (EK) method.

3.2.5.1. MCS with GARCH (1,1)-normal framework

The implementation of MCS consists of fitting the data to the GARCH (1,1)-normal model, by 250-day rolling windows, obtaining the standard deviation of each window and the respective residuals. Then a large amount, S, of pseudo-random numbers is generated, in this case, $S = 10000$, drawn from the $N(0, 1)$ distribution. Simulated returns are then calculated by multiplying the obtained standard deviation of each window with a random number from S generated pseudo-random numbers. The volatility of the next day is calculated using the GARCH (1,1) model variance equation. In this study, the forecast horizon $K = 1$. After calculating the hypothetical set of 250 returns for each window, the 1-day VaR is calculated by estimating the quantile of the simulated returns.

The model implementation in R of the GARCH (1,1)-normal framework is also done using the "rugarch" package.

3.2.5.2. GARCH (1,1)-Student-t framework

The same procedure is followed as in the GARCH (1,1)-normal framework, but instead of generating the $S = 10000$ from the standard normal distribution $N(0, 1)$, the random numbers are generated from the Student-*t* distribution. To do that, the degree of freedom of the residuals obtained from fitting the data to the GARCH (1,1)-Student-*t* distribution is needed. To obtain this degree of freedom, in this study, two methods are adopted.

a) Excess Kurtosis method (EK method)

In the EK method, the degree of freedom, d, per which the pseudo-random Student-*t* distributed random numbers are to be generated is obtained using the following formula:

$$d = ({}^{6}/_{EK}) + 4$$

To obtain EK, the data is fitted to GARCH-Normal distribution. Then, the EK of the standardized residuals is computed and thereafter, d is easily calculated. In R, the main package used in this method is the "e1071" (for calculating the excess kurtosis) and "rugarch" in addition to "metRology" (used for generating random Student-*t* distributed numbers).

b) Fitting method

The fitting method constitutes of fitting the data to the GARCH (1,1)-Student-*t* distribution and obtaining the residuals. After that, these residuals are fitted to a Student-*t* distribution to obtain their respective degree of freedom. In R, the packages used in the implementation of this method are mainly “rugarch” and “metRology”. The “QRM” and “MASS” packages were used to fit the obtained residuals to Student-*t* distribution and obtain the respective degree of freedom for each window of residuals.

3.2.6. GARCH (1,1) model

The GARCH (1,1) model is applied in two frameworks, one with normally distributed residuals and the other with Student-t distributed residuals. However, in the latter framework, it is implemented using two methods, one based on the fitting method to find the degree of freedom “d” of the residuals, and the other uses the Excess Kurtosis method to find the respective degree of freedom.

The methodology is based on fitting the data to the GARCH (1,1) model considering the different types of residuals and accordingly obtaining (by rolling window of 250 days) the corresponding standard deviation and mean for each window to obtain the respective 99% VaR and 95% VaR as per the original definition of VaR.

This model is applied in R using the packages like “rugarch”, “metrology”, “QRM” among others.

3.2.7. Extreme Value Theory (Peaks over Threshold)

The EVT model is applied in this study with the POT approach which requires a pre-specification of a threshold that separates the distribution tails of the returns from its middle part. Two values are selected for the threshold, 90th percentile and 95th percentile of the return distribution. Noting that the threshold value should not be too low and break the asymptotic property, and it should not be too high and have very few data points to estimate the parameters in the excess distribution. Choosing the 90th percentile threshold value is a common practice (Omari et al., 2020) and the 95th percentile threshold value was chosen to maintain consistency with the GARCH-EVT model application. Both choices of the threshold value are supported in the GARCH-EVT model section. Then the results of the VaR prediction on each threshold value are compared.

In R, the packages “fExtremes”, “qrmtools”, “QRM” and “POT” are the main packages used in the estimation of VaR using the EVT method.

3.2.8. GARCH-EVT model

The GARCH-EVT model is applied using the GARCH (1,1) model. The data is fitted to it by 250-day rolling window using: (a) quasi-maximum likelihood, (b) normal residuals and (c) Student-*t* residual. In each framework, the standardized residuals are obtained for each window and the EVT is then applied to the tails of the residuals considering two cases: (i) 95th percentile threshold; (ii) 90th percentile threshold (i.e., 5% and 10% of the window residuals are considered as the excess over the threshold respectively in each case). Researchers are still searching for methods to find the optimal tail selection method to choose the appropriate threshold for EVT based models. However, Echaust and Just (2020) findings contradicted their associated hypothesis that the optimal tail selection methods for appropriate threshold selection improves the accuracy of VaR predictions relative to the standard method approach based on a fixed 95th percentile of distribution as threshold. Thus, they concluded that investors may use the conditional EVT approach taking the 95th percentile of the sample as a threshold in order to obtain accurate VaR estimates. Simultaneously, the choice of the 95th and 90th percentiles of the return distribution as threshold values is based on the recommendations of McNeil and Frey (2000) (Byström, 2004) which was later extensively used in the literature, especially the 90th percentile which became a common practice, like in Raimondo and Tajvidi (2004), Bystrom (2004), Omari et al. (2020), and Huang et al. (2017) among others. These reasons explain the thresholds adopted in this thesis for GARCH-EVT model, and for EVT (POT) as well, to maintain consistency.

3.3. Distance Measures Analysis

In Chapter 4 the backtesting measures will be used to evaluate the performance of the different models. As a preliminary analysis, carry out a distance measures analysis is conducted, and it constitutes of measuring the distance between the obtained VaR estimates and the actual returns of each index, during each of the three intervals.

The measures and their notations used are defined as follows:

N+: Number of positive distances (Actual return – VaR estimate > 0)

N+ (Returns $<$ 0): Number of positive distances while returns are less than zero.

SN+ (Returns $<$ 0): Sum of positive distances for the period when returns are less than zero.

N+ (Returns ≥ 0): Number of positive distances while returns are greater than or equal to zero

SN+ (Returns ≥ 0): Sum of positive distances for the period when returns are greater than or equal to zero.

SN+ Total: Sum of positive distances for the period for all returns, positive and negative

N-: Number of negative distances (Actual return – VaR estimate < 0) which is the number of violations of VaR.

SN–: Sum of negative distances for the period

S Total = Sum of total distances for the period: (SN+) + (SN–)

%N–: % Violations $= (N -)/[(N -) + (N +)]$, being $[(N -) + (N +)]$ the number of observations for the period

CC: Correlation Coefficient between negative returns and VaR estimate

It is worthy to note that the greater the value of SN+ (Returns < 0), the greater the tendency of the VaR model to be overestimating VaR. On the other hand, the greater the value of SN–, the greater the possibility of the employed model to be underestimating VaR.

Throughout this chapter, the distance measures listed above, together with the CC, are presented in tabular forms.

For consistency reasons, the analysis of distance measures is based on classifying their values as low, moderate, or high. This classification is based on taking the minimum and maximum of each measure for each index in each interval and creating three equal intervals in between the corresponding values. The interval boundaries are treated as histogram intervals with all intervals being left closed right open except for the last one. These intervals are also ranked between low, moderate, and high. This way the respective analysis of the distance measures would be unbiased leading to objective analytical comparison between the VaR models.

3.4. Pre-crisis period 2002 -2007

3.4.1. Dow Jones Industrial Average (DJIA) index

3.4.1.1. Market evolution

The DJIA returns witnessed high volatility between 2002 and 2003 due to the Internet bubble burst. The Internet bubble burst (2000 – 2002), also known as the dot-com bubble burst, was a consequence of the speculative current that caused the prices of Internet companies to rise between 1997 and 2000. Between 2004 and 2006, the market variations were relatively calm with regular fluctuations. However, in 2007 when the financial crisis began to appear in the U.S. market, the log returns suffered high volatility. In the following sections, a graphical representation of the obtained VaR estimates using the selected models is given together with a preliminary model behavior analysis using distance measures.

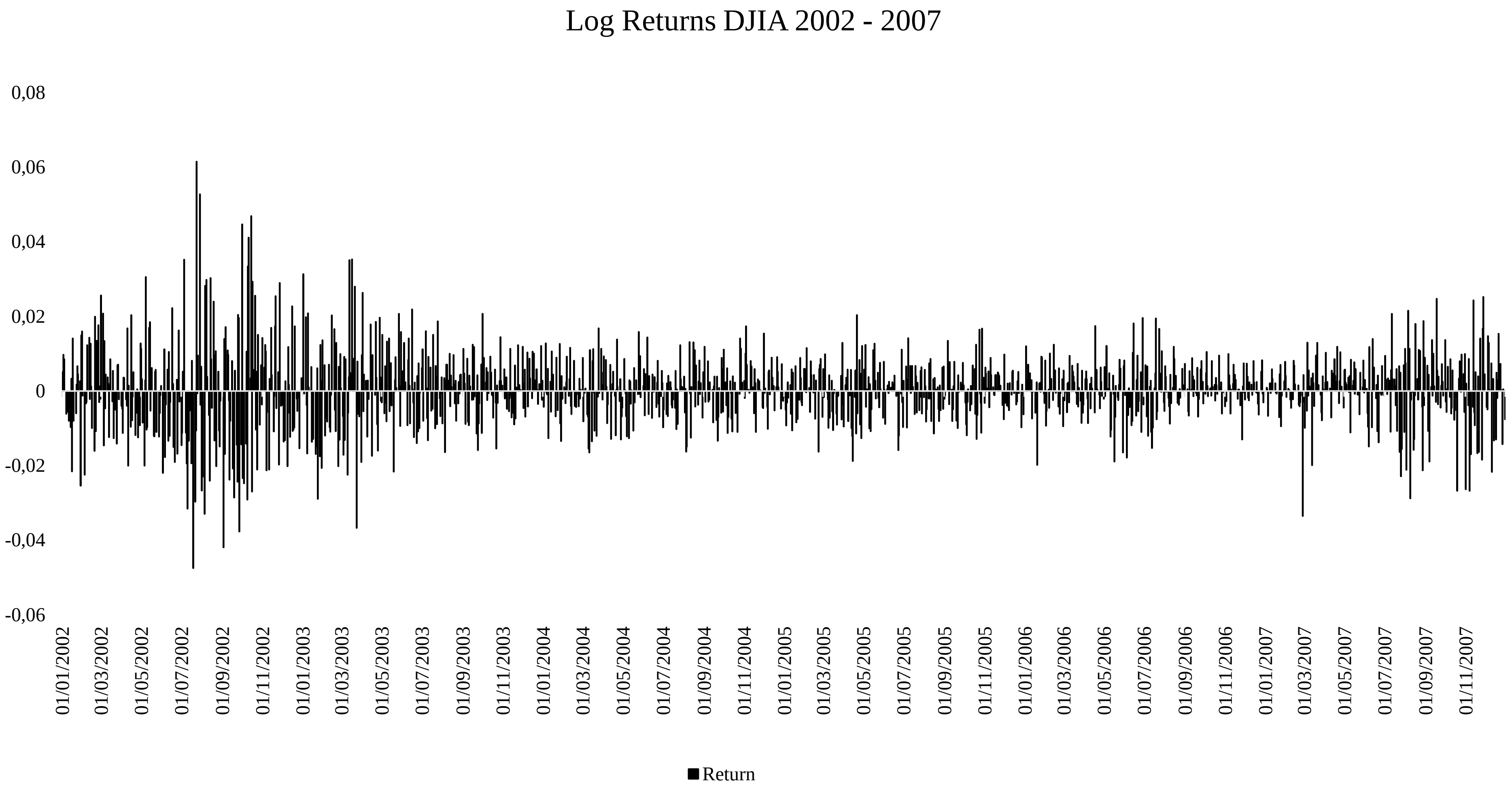

Figure 3.1. DJIA log returns between 2002 and 2007.

3.4.1.2. VaR estimates

Graphical representation

Figure 3.2 constitutes the graphical representation of the obtained 99% VaR estimates of the DJIA index between 2002 and 2007 under all the applied models. It shows that some models produce 99% VaR estimates that do not respond to the market volatility as quickly as others. For instance, the HS and the var-cov models admit nearly flat curves with minimal fluctuations throughout the entire interval. The GARCH-EVT curves also witnesses a constant trend during calm and volatile periods throughout the interval. The GARCH-EVT with Student-*t* residuals at 95% threshold value curve designates an overestimation of VaR in period after 2007, however, the same model at the 90% threshold showed a similar curve to former but with less overestimation after 2007.

The same behavior can be observed in Figure 3.3 for 95% VaR estimates of DJIA in the same period. Since graphically it is difficult to determine the performance of each model at both confidence levels, it is necessary to obtain some numerical measures to address it. Next, the distance analysis is carried out.

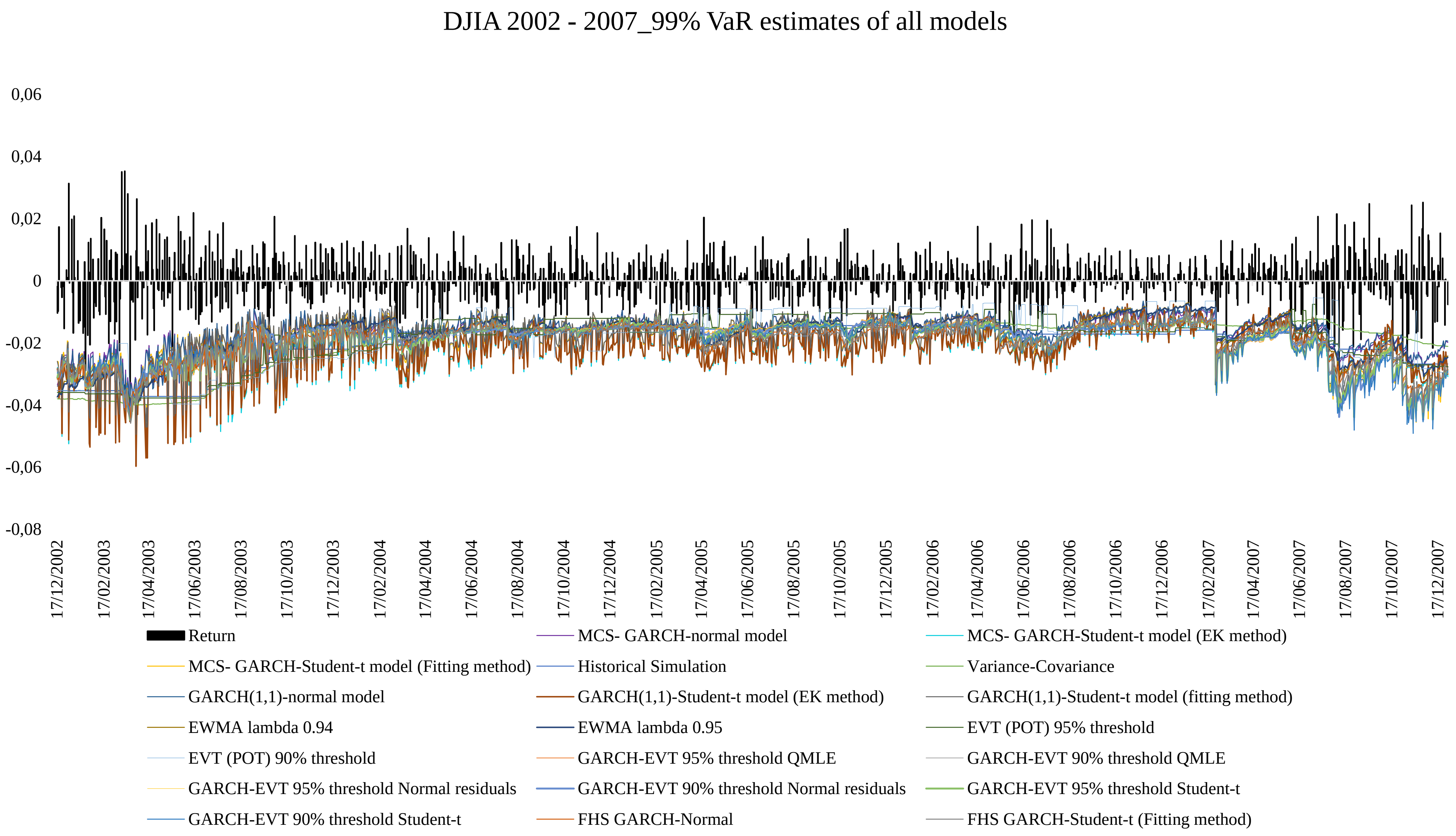

Figure 3.2. DJIA 99% VaR estimates between 2002 and 2007.

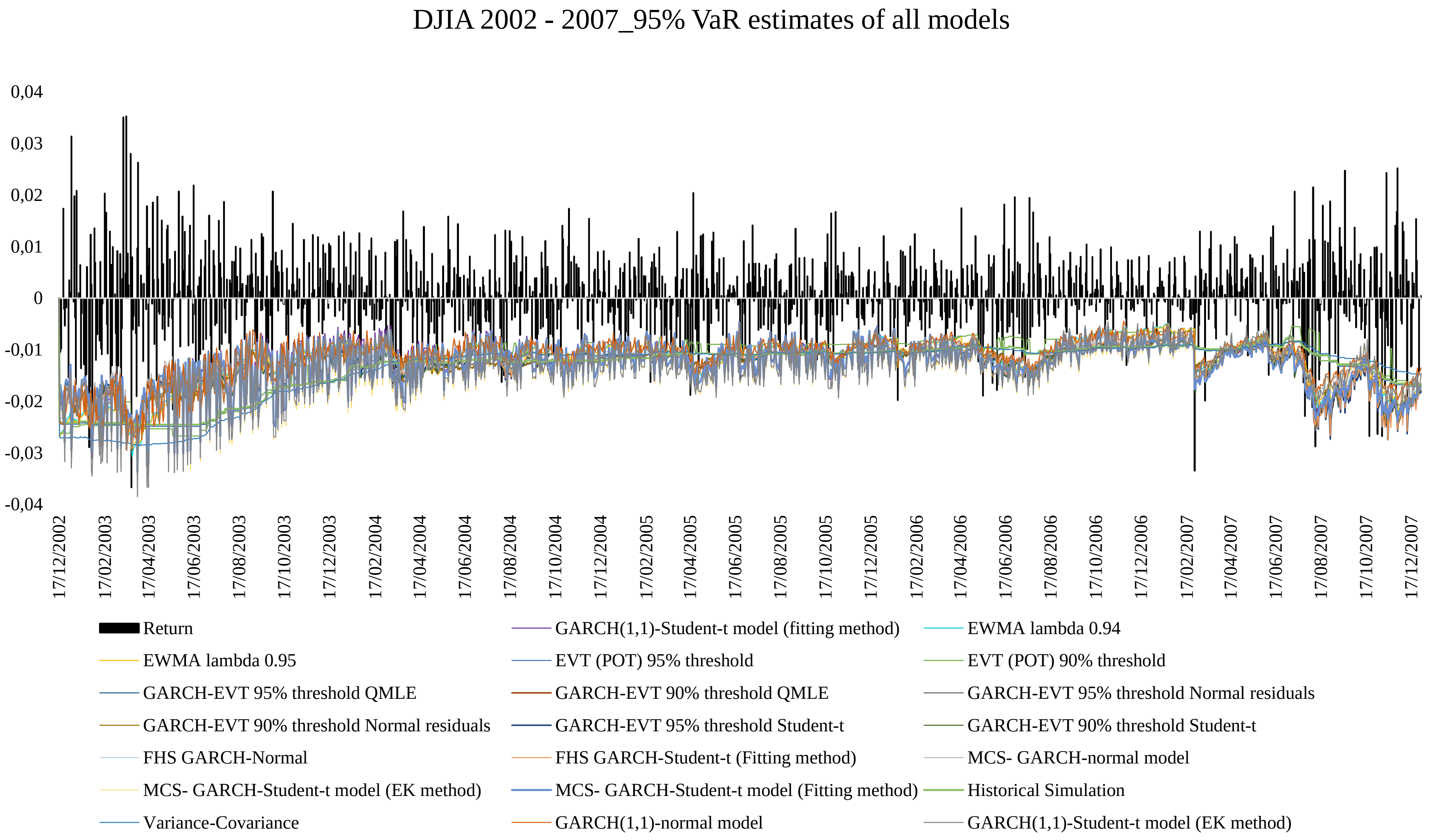

Figure 3.3. DJIA 95% VaR estimates between 2002 and 2007.

Distance Measures

99% VaR

The distance measures calculated for the DJIA in the pre-crisis period for the 99% VaR estimates are represented in Table 3.3 and Table 3.4.

DJIA 2002 – 2007_ 99% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT POT		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student	QMLE	Normal	Student	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1309	1313	1309	1309	1311	1309	1288	1277	1299	1309	1308	1295	1310	1308
N+ (Returns < 0)	582	586	582	582	584	582	561	550	572	582	581	568	583	581
SN+ (Returns < 0)	17.833	18.154	17.833	17.883	18.230	17.883	17.746	17.774	15.838	19.578	18.322	15.653	19.383	18.127
N+ (Returns ≥ 0)	727	727	727	727	727	727	727	727	727	727	727	727	727	727
SN+ (Returns ≥ 0)	7.760	7.976	7.760	7.814	8.019	7.814	7.653	7.927	6.330	9.236	8.149	6.210	9.087	8.000
S N+ Total	25.594	26.131	25.594	25.696	26.249	25.696	25.400	25.700	22.168	28.814	26.472	21.864	28.470	26.127
N- (violations)	6	2	6	6	4	6	27	38	16	6	7	20	5	7
S N- (violations magnitude)	0.016	0.005	0.016	0.014	0.004	0.014	0.088	0.134	0.044	0.018	0.011	0.049	0.019	0.014
S total	25.610	26.135	25.610	25.710	26.252	25.710	25.487	25.835	22.212	28.833	26.483	21.913	28.489	26.141
%N-	0.456%	0.152%	0.456%	0.456%	0.304%	0.456%	2.053%	2.890%	1.217%	0.456%	0.532%	1.521%	0.380%	0.532%
CC	0.434	0.444	0.434	0.434	0.432	0.434	0.271	0.240	0.426	0.336	0.391	0.421	0.341	0.398

Table 3.3. DJIA Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2002 and 2007.

DJIA 2002 - 2007_ 99% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	var-cov	HS
N+	1311	1314	1290	1291	1290	1300
N+ (Returns < 0)	584	587	563	564	563	573
SN+ (Returns < 0)	17.862	18.054	16.434	16.486	17.889	18.330
N+ (Returns ≥ 0)	727	727	727	727	727	727
SN+ (Returns ≥ 0)	7.829	7.899	6.733	6.780	7.928	8.145
S N+ Total	25.691	25.953	23.167	23.266	25.817	26.474
N- (violations)	4	1	25	24	25	15
S N- (violations magnitude)	0.011	0.004	0.102	0.101	0.121	0.046
S total	25.702	25.956	23.268	23.367	25.938	26.520
%N-	0.304%	0.076%	1.901%	1.825%	1.901%	1.141%
CC	0.457	0.454	0.292	0.292	0.232	0.276

Table 3.4. DJIA Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2002 and 2007.

GARCH-EVT (95% threshold), GARCH-EVT (90% threshold), GARCH (1,1) model under Student-t residuals (fitting method) and FHS admit considerably low values of N-, moderate distances S Total and comparably high values of CC with respect to the other models. It is worth noting that GARCH-EVT and FHS models witness fewer ~~less~~ violations with Student-t residuals than with normally distributed residuals.

Simultaneously, the MCS and GARCH (1,1) models under normally distributed residuals admit moderate values of violations, considerably low values of S Total and high values of CC. However, under the Student-t distributed residuals (fitting method) ~~these~~ both models attain low frequencies of violations, and comparably high CC values, while GARCH (1,1) model shows moderate value of S Total but MCS suffers a bigger value of S Total.

On the other hand, under Student-t distributed residuals (EK method), MCS and GARCH (1,1) models show low values of N-, large distances (S total) and moderate values of CC.

EVT (POT) (95% and 90% threshold) models suffer high frequencies of violations, low CC values and moderate distances, S Total, values.

EWMA (0.94 and 0.95) admit moderate values of N-, low values of CC and low S Total values.

HS admits moderate number of violations but high value of S Total and a low value of CC while var-cov attains moderate values of N- and S Total along with low value of CC.

95% VaR

Tables 3.5 and 3.6 present the distance measures estimated for the 95% VaR of the DJIA obtained under the different models in the pre-crisis period.

DJIA 2002 – 2007_ 95% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT POT		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student	QMLE	Normal	Student	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1272	1270	1272	1274	1271	1274	1252	1231	1245	1268	1257	1237	1261	1249
N+ (Returns < 0)	545	543	545	547	544	547	525	504	518	541	530	510	534	522
SN+ (Returns < 0)	13.315	13.365	13.315	13.408	13.425	13.408	13.663	13.075	12.382	13.959	13.463	12.170	13.743	13.249
N+ (Returns ≥ 0)	727	727	727	727	727	727	727	727	727	727	727	727	727	727
SN+ (Returns ≥ 0)	4.263	4.261	4.263	4.333	4.294	4.333	4.644	4.284	3.654	4.782	4.366	3.534	4.652	4.237
S N+ Total	17.578	17.626	17.578	17.741	17.719	17.741	18.308	17.359	16.035	18.741	17.829	15.704	18.395	17.486
N- (violations)	43	45	43	41	44	41	63	84	70	47	58	78	54	66
S N- (violations magnitude)	0.159	0.147	0.159	0.155	0.146	0.155	0.280	0.364	0.239	0.179	0.183	0.259	0.192	0.200
S Total	17.737	17.773	17.737	17.896	17.866	17.896	18.587	17.723	16.274	18.920	18.012	15.963	18.587	17.686
% N-	3.270%	3.422%	3.270%	3.118%	3.346%	3.118%	4.791%	6.388%	5.323%	3.574%	4.411%	5.932%	4.106%	5.019%
CC	0.443	0.438	0.443	0.442	0.441	0.442	0.251	0.242	0.423	0.339	0.375	0.420	0.343	0.384

Table 3.5. DJIA Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2002 and 2007.

DJIA 2002 – 2007_ 95% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	var-cov	HS
N+	1271	1267	1236	1235	1247	1251
N+ (Returns < 0)	544	540	509	508	520	524
SN+ (Returns < 0)	12.969	13.108	12.722	12.759	13.751	13.590
N+ (Returns ≥ 0)	727	727	727	727	727	727
SN+ (Returns ≥ 0)	4.040	4.091	3.918	3.948	4.752	4.595
S N+ Total	17.009	17.199	16.640	16.707	18.502	18.185
N- (violations)	44	48	79	80	68	64
S N- (violations magnitude)	0.173	0.158	0.311	0.307	0.313	0.286
S Total	17.182	17.357	16.951	17.015	18.815	18.471
% N-	3.346%	3.650%	6.008%	6.084%	5.171%	4.867%
CC	0.454	0.443	0.293	0.293	0.233	0.249

Table 3.6. DJIA Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2002 and 2007

According to Tables 3.5 and 3.6, it can be noticed that GARCH-EVT (95% and 90% threshold) and FHS models admit considerably low values of N-, moderate values of S Total and comparably high values of CC, in a way that the models seem to have similar behavior when estimating the 99% VaR and 95% VaR.

Simultaneously, MCS under normally distributed residuals also shows similar performance to when estimating the 99% VaR, it admits a considerably moderate value of N-, and low values of S Total and high CC.

However, GARCH (1,1) model demonstrates a different behavior when estimating the 95% VaR than with the 99% VaR. As it can be noticed in Table 3.5, the measures show high frequency of violations with normally distributed residuals accompanied by low values of S Total and high values of CC.

On the other hand, both MCS and GARCH (1,1) models with Student-t residuals (EK method) show low values of N- accompanied by high values of S Total and moderate correlation values, whereas with the Student-t residuals (fitting method), MCS and GARCH (1,1) show moderate values of violations and moderate values of CC but the former shows high values of S Total while the latter admits a moderate value of the distance measure S Total.

EWMA (0.94 and 0.95) and EVT (POT) (90% threshold) reveal similar results with high number of violations, moderate S Total, and low correlation values.

EVT (POT) (95% threshold), var-cov and HS admit moderate values of N-. However, they have low correlation values and comparably high values of S Total.

Comments

The distance measures analysis and graphical representation of the models used for the DJIA in the pre-crisis period show the following:

99% VaR:

- GARCH-EVT (95% and 90% threshold), FHS, MCS under normal residuals and GARCH (1,1) under normal and Student-t residuals (fitting method) appear to outperform the remainder models while they admit lower values of violations N-, high CC values and low to moderate values of S Total.

95% VaR:

- GARCH-EVT (95% and 90% threshold), FHS, MCS under normal residuals and GARCH (1,1) under Student-t residuals (fitting method) appear to outperform the remainder models.
- The remainder models admit at least one measure revealing their weak performance.

3.4.2. Euro Stoxx 50 (SX5E) index

3.4.2.1 Market evolution

The SX5E index log returns admitted high clustered volatility between 2002 and 2003, as it can be seen in Figure 3.4, which can be referred to the Internet bubble burst that took place in this period. Between 2004 and almost mid-2006, the market was considerably tranquil with regular market movements to witness again some turbulence around mid-2006. Later in 2007, as the signs of the 2008 financial crisis started to appear, the market was in high frequency turbulent state. In the following sections, the VaR estimates represented graphically, along with preliminary analysis using distance measures on these estimates are demonstrated.

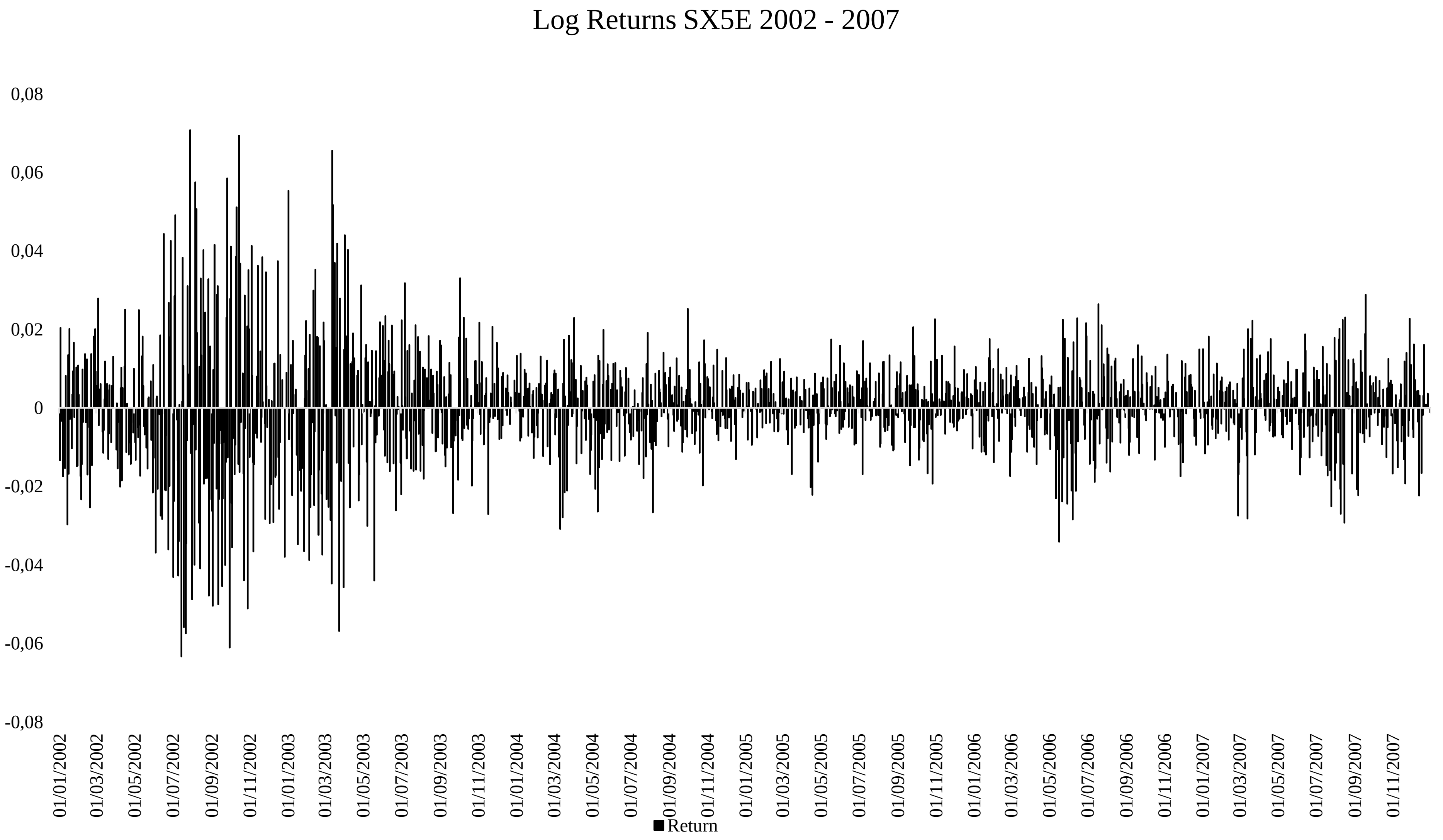

Figure 3.4. SX5E log returns between 2002 and 2007.

3.4.2.2. VaR estimates

Graphical representation

Figure 3.5 and Figure 3.6 illustrate graphically the 99% VaR and 95% VaR of the SX5E index, respectively, estimated in the pre-crisis period under the models used in this thesis. In both figures, the difference between the nature of the VaR curves is obvious, some models admit smoother curves than others acquiring less fluctuations. Examples of such models are the HS and var-cov models, in addition to the EWMA model (λ=0.94 and λ=0.95) and the EVT (POT) model. On the other hand, some models seem to reflect the market movements better like the MCS, GARCH-EVT, FHS and GARCH (1,1) models. At the same time, among the aforementioned models, some models seem to overestimate the VaR with respective curves very far from the plotted returns while others are plotted very close to the returns and thus could be underestimating VaR. The distance measures used in the following sections provide a clearer analysis to the VaR estimates presented graphically.

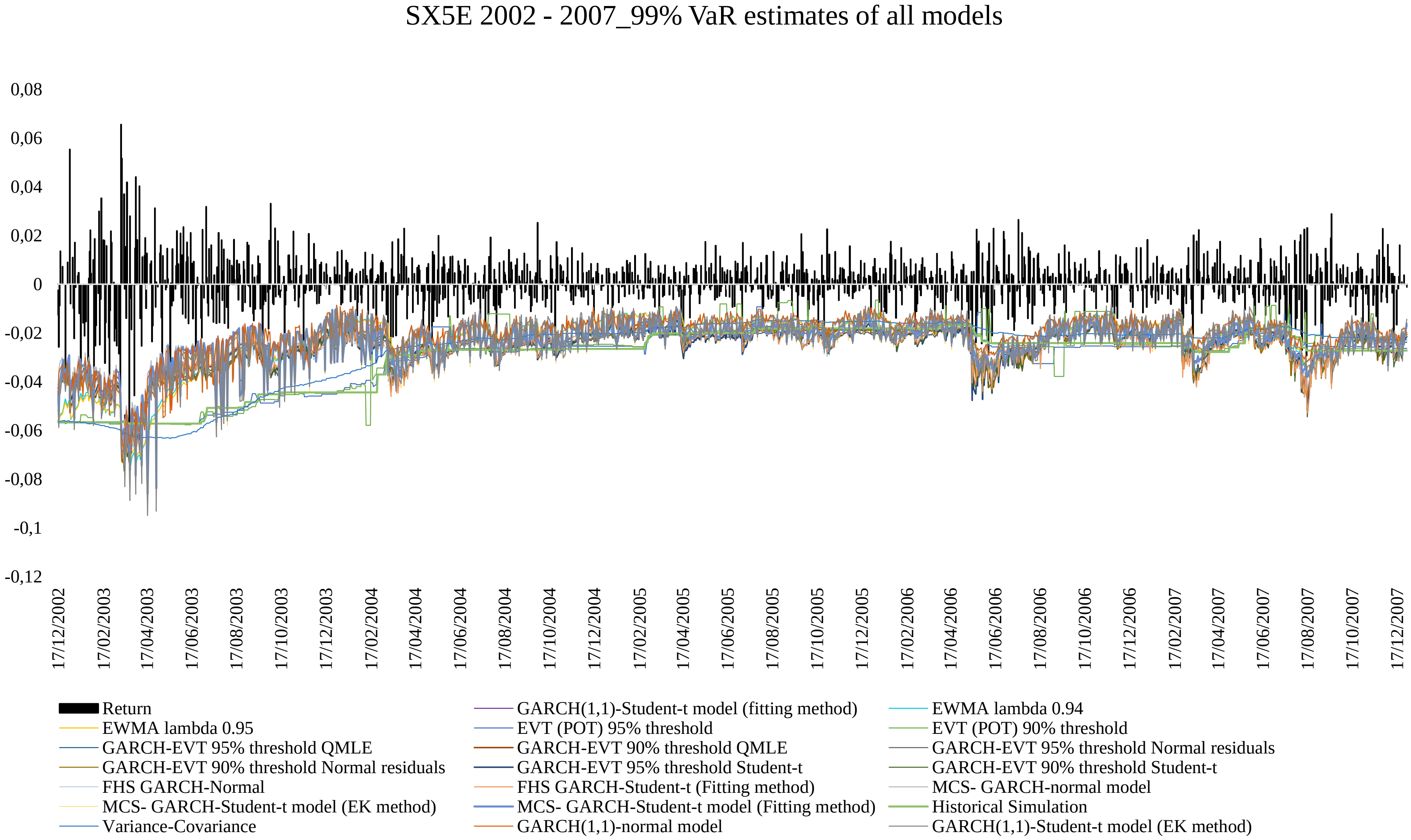

Figure 3.5. SX5E 99% VaR estimates between 2002 and 2007.

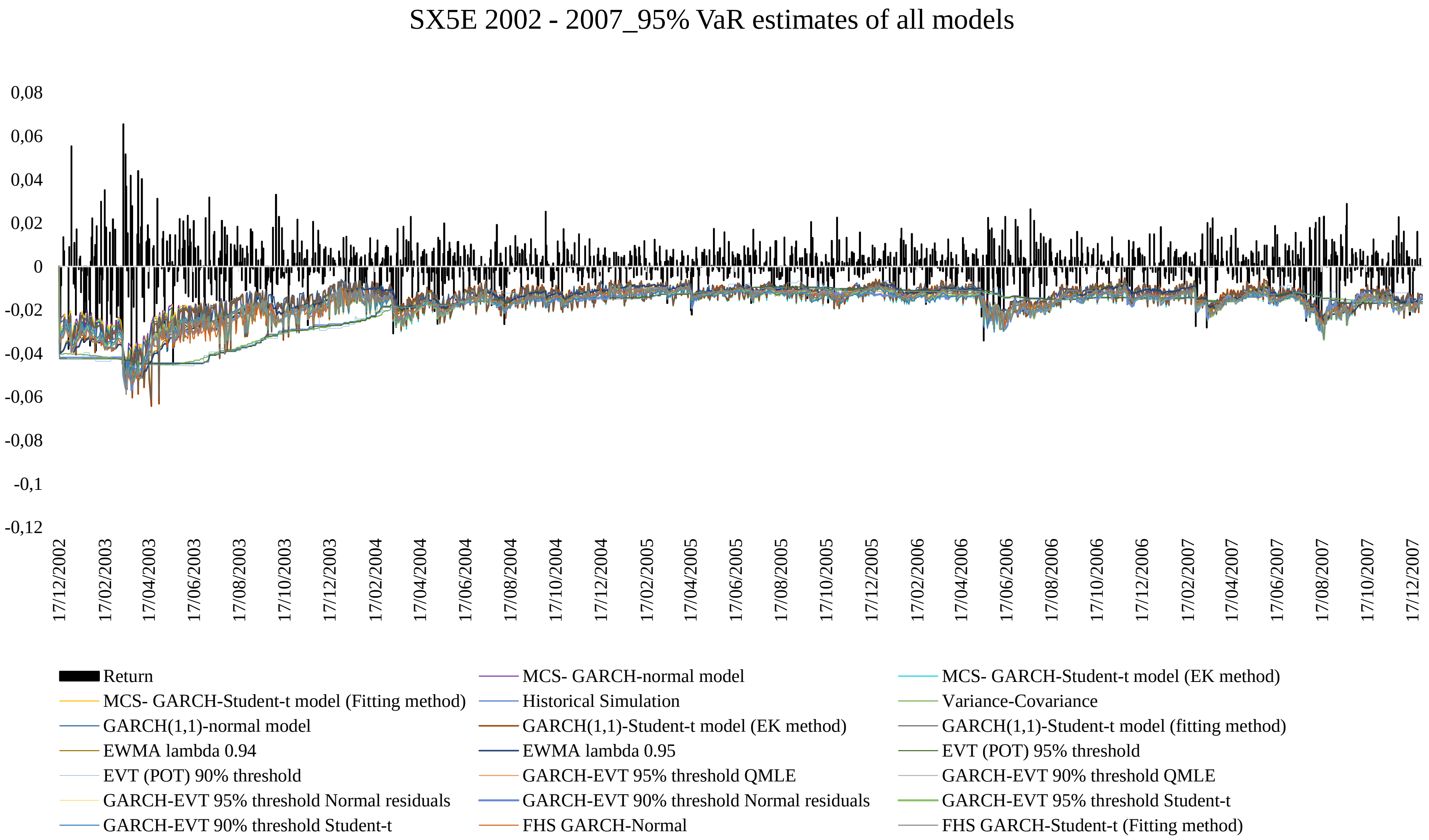

Figure 3.6. SX5E 95% VaR estimates between 2002 and 2007.

Distance Measures

99% VaR

Table 3.7 and Table 3.8 present the distance measures between the 99% VaR estimates and the actual log returns of the SX5E index in the pre-crisis interval between 2002 and 2007.

SX5E 2002 - 2007_99% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student	QMLE	Normal	Student	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1313	1311	1313	1313	1312	1313	1299	1292	1291	1299	1306	1290	1297	1307
N+ (Returns < 0)	605	603	605	605	604	605	591	584	583	591	598	582	589	599
SN+ (Returns < 0)	23.852	24.177	23.852	23.678	24.133	23.678	26.587	25.673	20.264	22.371	22.383	20.083	22.167	22.199
N+ (Returns ≥ 0)	708	708	708	708	708	708	708	708	708	708	708	708	708	708
SN+ (Returns ≥ 0)	11.317	11.479	11.317	11.234	11.487	11.234	13.534	12.957	8.108	9.764	9.827	8.003	9.647	9.692
S N+ Total	35.170	35.656	35.170	34.912	35.619	34.912	40.121	38.630	28.373	32.135	32.210	28.087	31.814	31.892
N- (violations)	2	4	2	2	3	2	16	23	24	16	9	25	18	8
S N- (violations magnitude)	0.003	0.005	0.003	0.003	0.005	0.003	0.055	0.129	0.061	0.033	0.021	0.065	0.039	0.018
S total	35.172	35.661	35.172	34.915	35.624	34.915	40.176	38.759	28.434	32.168	32.230	28.152	31.853	31.910
%N-	0.152%	0.304%	0.152%	0.152%	0.228%	0.152%	1.217%	1.749%	1.825%	1.217%	0.684%	1.901%	1.369%	0.608%
CC	0.553	0.527	0.553	0.554	0.524	0.554	0.317	0.312	0.493	0.460	0.474	0.492	0.464	0.480

Table 3.7. SX5E Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2002 and 2007.

SX5E 2002 - 2007_99% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	Var-cov	HS
N+	1314	1312	1291	1292	1291	1302
N+ (Returns < 0)	606	604	583	584	583	594
SN+ (Returns < 0)	23.749	23.957	21.640	21.756	24.422	26.537
N+ (Returns ≥ 0)	708	708	708	708	708	708
SN+ (Returns ≥ 0)	11.154	11.337	9.422	9.520	11.866	13.483
S N+ Total	34.903	35.295	31.062	31.276	36.288	40.020
N- (violations)	1	3	24	23	24	13
S N- (violations magnitude)	0.001	0.005	0.127	0.124	0.111	0.041
S total	34.904	35.299	31.189	31.400	36.399	40.061
%N-	0.076%	0.228%	1.825%	1.749%	1.825%	0.989%
CC	0.524	0.521	0.408	0.408	0.352	0.330

Table 3.8. SX5E Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2002 and 2007.

GARCH-EVT (95% and 90% threshold) along with FHS and MCS under Student-t residuals (fitting method) attain the best results of all models with the lowest values of N-, moderate values of S Total and the high values of CC.

It is worth noting that GARCH (1,1) under Student-t residuals (fitting method) is the only model that achieves the best results for all distance measures, with low violations, low distance S Total and high correlation coefficient.

MCS under Student-t residuals (EK method) also shows acceptable performance achieving a moderate frequency of violations, and moderate values of S Total and CC.

MCS and GARCH (1,1) models under normal residuals witness high frequencies of violations accompanied by low values of S Total and comparably high values of CC.

GARCH (1,1) model under Student-t residuals (EK method) admits high frequency of violations, low distance S Total and moderate value of CC. This model's behavior is poor in estimating the 99% VaR of SX5E in the pre-crisis interval.

EVT (POT) (90% threshold) and var-cov models admit the worst results of all models with the highest values of S Total and N- and the lowest value of CC.

EVT (POT) (95% threshold) and HS model witness similar results with moderate values of violations, high values of S Total and low values of CC.

EWMA (0.94 and 0.95) attain high frequencies of violations with moderate values of S Total and CC.

95% VaR

Tables 3.9 and 3.10 present the distance measures between the 95% VaR estimates and the actual log returns of the SX5E index between 2002 and 2007.

SX5E 2002 - 2007_95% VaR															
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student	QMLE	Normal	Student	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1279	1270	1279	1282	1269	1282	1252	1245	1229	1238	1239	1226	1232	1236
N+ (Returns < 0)	571	562	571	574	561	574	544	537	521	530	531	518	524	528
SN+ (Returns < 0)	17.891	17.892	17.891	17.976	17.980	17.976	18.891	18.719	15.908	16.791	16.778	15.701	16.580	16.575
N+ (Returns ≥ 0)	708	708	708	708	708	708	708	708	708	708	708	708	708	708
SN+ (Returns ≥ 0)	6.323	6.278	6.323	6.394	6.327	6.394	7.333	7.209	4.713	5.344	5.350	4.593	5.218	5.210
S N+ Total	24.214	24.169	24.214	24.369	24.307	24.369	26.224	25.928	20.621	22.135	22.129	20.294	21.799	21.785
N- (violations)	36	45	36	33	46	33	63	70	86	77	76	89	83	79
S N- (violations magnitude)	0.152	0.185	0.152	0.145	0.180	0.145	0.354	0.387	0.408	0.361	0.345	0.422	0.374	0.354
S total	24.366	24.354	24.366	24.514	24.488	24.514	26.578	26.315	21.029	22.497	22.474	20.716	22.172	22.139
%N-	2.738%	3.422%	2.738%	2.510%	3.498%	2.510%	4.791%	5.323%	6.540%	5.856%	5.779%	6.768%	6.312%	6.008%
CC	0.529	0.511	0.529	0.531	0.515	0.531	0.354	0.351	0.491	0.459	0.468	0.492	0.466	0.475

Table 3.9. SX5E Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2002 and 2007.

SX5E 2002 2007_95% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-Covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	var-cov	HS
N+	1266	1267	1230	1228	1249	1248
N+ (Returns < 0)	558	559	522	520	541	540
SN+ (Returns < 0)	17.650	17.625	16.801	16.884	18.769	18.726
N+ (Returns ≥ 0)	708	708	708	708	708	708
SN+ (Returns ≥ 0)	6.092	6.088	5.547	5.614	7.228	7.194
S N+ Total	23.742	23.713	22.349	22.498	25.996	25.920
N- (violations)	49	48	85	87	66	67
S N- (violations magnitude)	0.197	0.192	0.417	0.412	0.358	0.364
S total	23.939	23.905	22.765	22.909	26.355	26.284
%N-	3.726%	3.650%	6.464%	6.616%	5.019%	5.095%
CC	0.501	0.507	0.410	0.410	0.354	0.354

Table 3.10. SX5E Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2002 and 2007.

GARCH-EVT (95% and 90% threshold) models along with FHS model demonstrate the best performance as per the distance measures as they achieve low values of N- and, in particular, the lowest values are attained under the normal and QMLE fitted residuals for the GARCH-EVT models, while slightly higher values are noticed with Student-t residuals and FHS. Moreover, these models also attain moderate values of S Total and high CC values.

MCS under normally distributed residuals and GARCH (1,1) under normal and Student-t residuals (fitting method) admit low values of S Total and quite high values of CC however this is accompanied with the highest numbers of violations. MCS under Student-t residuals (EK and fitting methods) and GARCH (1,1) under Student-t residuals (Ek method) also suffer high frequencies of violations however they admit low S Total values and moderate CC values.

However, EWMA (0.94 and 0.95) admit high numbers of violations accompanied by moderate values of S Total and low values of CC.

EVT (POT) (95% threshold) along with var-cov and HS models admit moderate values of violations, high values of S Total and considerably low values of CC.

EVT (POT) (90% threshold) shows the weakest performance among all models with high frequency of violations and S Total and low CC value.

Comments

The distance measures analysis of the VaR estimates obtained for the SX5E index in the pre-crisis period along with the graphical representation of the VaR predictions show the differences between the VaR models behavior and performance. As such, the following comments can be made:

99% VaR:

- GARCH-EVT (95% and 90% threshold) and FHS under all residuals along with GARCH (1,1) under Student-t residuals (fitting method), and MCS under Student-t residuals (EK and fitting method) outperform the remainder models.
- The remainder models admit at least one measure that is considered weak compared to the previously mentioned models.

95% VaR:

- GARCH-EVT (95% and 90% threshold) models along with FHS model outperform the remainder models when estimating the 95% VaR of the SX5E in the pre-crisis period.
- EVT (POT) (90% threshold) shows the weakest results among all models.

3.4.3. Nikkei 225 (N225) Index

3.4.3.1. Market evolution

The evolution of the log returns of the N225 index between 2002 and 2007 are presented in Figure 3.7. The first thing to be noticed is that the Internet bubble burst, which took place in

2000 and reached its trough in October 2002, did not have the tremendous effects on the N225 index between 2002 and 2003 as much as it did with the DJIA and SX5E indices. The magnitude of losses in the period between 2002 and 2003 is less than that witnessed in the same period for the other indexes. Moreover, the difference between the N225 log returns and the DJIA and SX5E log returns can be seen as the former admits very high fluctuations almost all the time throughout this interval. The high fluctuations in returns were also present in 2004, 2005 and 2006. These fluctuations can be referred to the asset price bubble that started back in 1986 and burst in 1991 lingering its consequences on the Japanese economy for decades afterwards, which are known as the "lost decades", in which the Japanese economy and in particular the stock market could not recover the consequent losses until sometime around 2012. The crash observed in the N225 index returns in 2007 can be also referred to the financial crisis of 2008 in addition to the bursting of the Japanese asset price bubble. In general, the Japanese market, unlike the U.S. and European markets, suffered a continuously high volatility throughout this interval. This analysis shows the differences between the nature of evolution of the three indexes, and which will be reflected in the results of the VaR estimates.

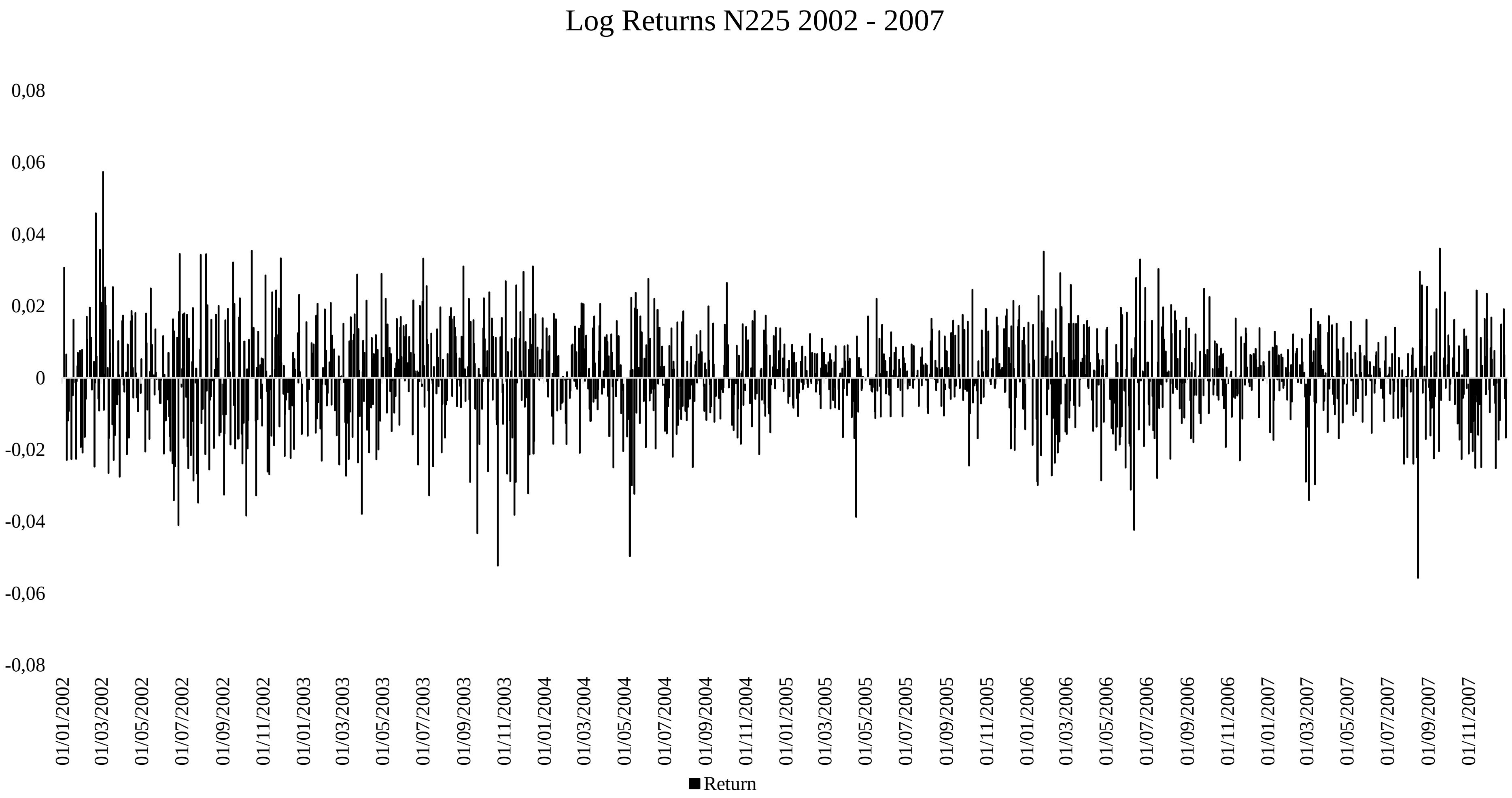

Figure 3.7. N225 log returns between 2002 and 2007.

3.4.3.2. VaR estimates

Graphical representation

Figures 3.8 and 3.9 demonstrate the graphical representation of the 99% and 95% VaR estimates of the N225 index in the pre-crisis period. These graphs show that the behavior of most models differs between the N225, DJIA and SX5E. It can be noticed that, for instance, the EVT (POT) VaR curves in the N225 index lie very close to the plotted returns unlike in the DJIA and SX5E. The GARCH-EVT at 90% threshold with normal residuals shows high overestimation of VaR in the beginning of the interval between 2002 and 2003. Since it is hard to determine which models perform better than the others graphically, the distance measures in the following section clarify some characteristics regarding the performance of each model in estimating VaR for N225 in the pre-crisis period.

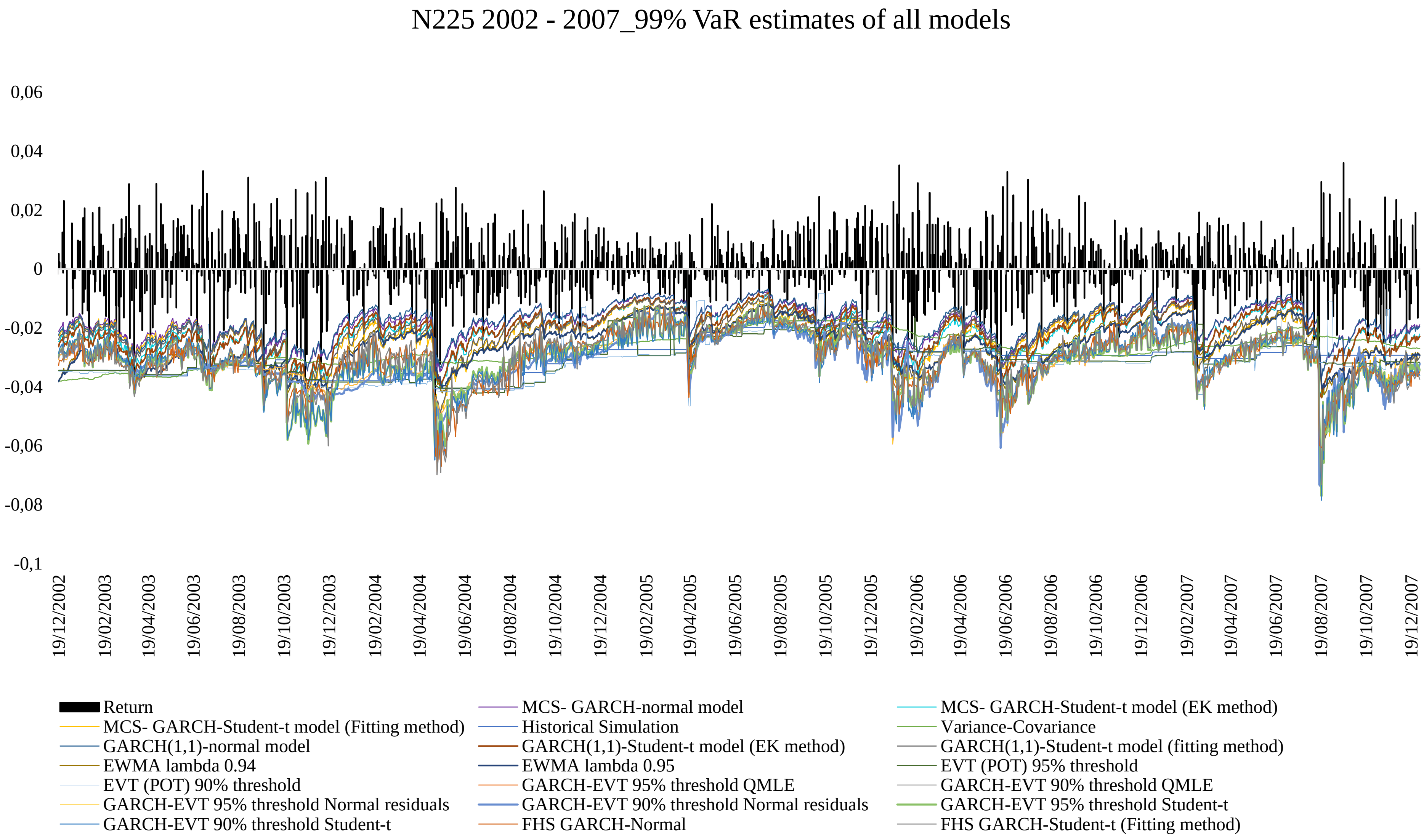

Figure 3.8. N225 99% VaR estimates between 2002 and 2007.

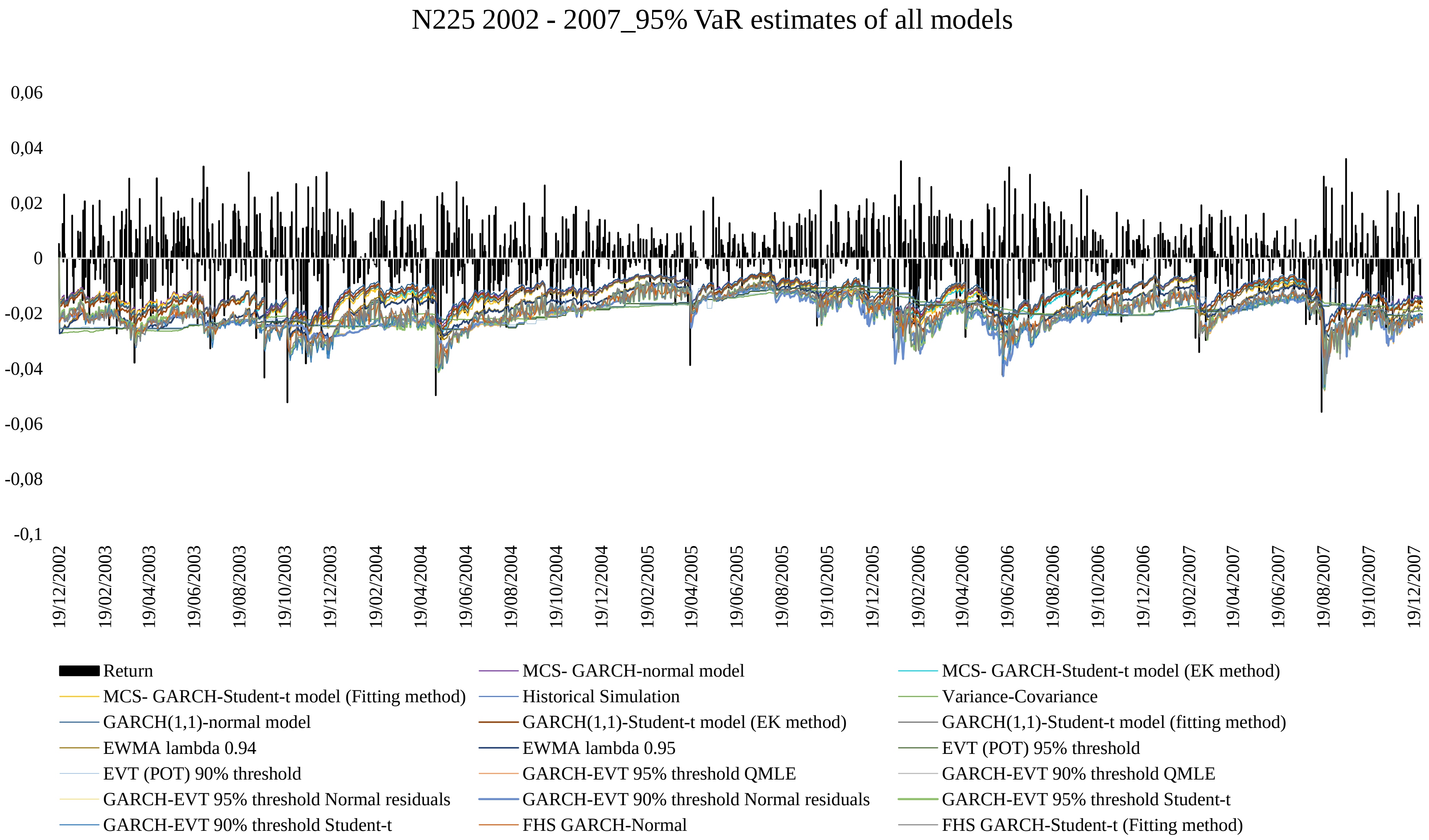

Figure 3.9. N225 95% VaR estimates between 2002 and 2007.

Distance Measures

99% VaR

N225 2002 - 2007_99% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student	QMLE	Normal	Student	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1310	1310	1310	1309	1310	1309	1296	1299	1252	1286	1291	1248	1285	1290
N+ (Returns < 0)	583	583	583	582	583	582	569	572	525	559	564	521	558	563
SN+ (Returns < 0)	28.022	28.122	28.022	28.344	28.264	28.344	27.955	28.329	19.093	21.465	22.575	18.827	21.175	22.278
N+ (Returns ≥ 0)	727	727	727	727	727	727	727	727	727	727	727	727	727	727
SN+ (Returns ≥ 0)	13.009	12.870	13.009	13.291	12.977	13.291	12.931	13.203	5.727	7.481	8.441	5.554	7.273	8.218
S N+ Total	41.030	40.992	41.030	41.635	41.241	41.635	40.886	41.532	24.820	28.946	31.016	24.381	28.448	30.497
N- (violations)	3	3	3	4	3	4	17	14	61	27	22	65	28	23
S N- (violations magnitude)	0.031	0.018	0.031	0.024	0.016	0.024	0.138	0.095	0.262	0.106	0.088	0.288	0.112	0.090
S total	41.061	41.010	41.061	41.659	41.257	41.659	41.024	41.626	25.082	29.052	31.104	24.669	28.559	30.587
%N-	0.228%	0.228%	0.228%	0.305%	0.228%	0.305%	1.295%	1.066%	4.646%	2.056%	1.676%	4.950%	2.133%	1.752%
CC	0.428	0.435	0.428	0.426	0.432	0.426	0.104	0.160	0.497	0.494	0.438	0.497	0.498	0.448

Table 3.11. N225 Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2002 and 2007.

N225 2002 - 2007_99% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-Covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	var-cov	HS
N+	1310	1310	1284	1284	1284	1296
N+ (Returns < 0)	583	583	557	557	557	569
SN+ (Returns < 0)	27.756	27.829	24.466	24.522	25.371	27.531
N+ (Returns ≥ 0)	727	727	727	727	727	727
SN+ (Returns ≥ 0)	12.632	12.631	9.831	9.914	10.843	12.621
S N+ Total	40.389	40.461	34.297	34.436	36.214	40.151
N- (violations)	3	3	29	29	29	17
S N- (violations magnitude)	0.014	0.021	0.206	0.206	0.205	0.106
S total	40.403	40.481	34.503	34.643	36.419	40.257
%N-	0.228%	0.228%	2.209%	2.209%	2.209%	1.295%
CC	0.415	0.426	0.219	0.215	0.096	0.148

Table 3.12. N225 Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2002 and 2007.

The distance measures in Table 3.11 and Table 3.12 correspond to the VaR estimates obtained for the N225 index in the pre-crisis interval.

The results of the distance measures with N225 index are quite different from those generally noticed with the DJIA and SX5E indexes. This can be referred to the different trend of evolution of returns and the volatility continuously embedded in the Japanese market in the pre-crisis period and the different behavior of the VaR models with N225 index is expected.

Consequently, it can be noticed that the GARCH-EVT (95% and 90% threshold) models and FHS model do not outperform the other models as it was the case with DJIA and SX5E indexes during the same interval.

However, with N225 index GARCH-EVT (95% and 90% thresholds) and FHS models admit low violations as it was the case with DJIA and SX5E, they also admit high CC values, however, the distance S Total is quite high compared to other models which designates weak performance.

EVT (POT) (95% and 90% threshold) also admit considerably low frequencies of violations however, they suffer high S Total values and low CC values.

On the other hand, MCS and GARCH (1,1) which witnessed good results with DJIA and SX5E still provide good results with N225, as such, MCS and GARCH (1,1) models under Student-t residuals (EK method) admit moderate values of violations and low distances S Total with high values of CC.

Simultaneously, MCS and GARCH (1,1) under Student-t residuals (fitting method) admits low frequencies of violations, high values of CC and moderate distances S Total.

EWMA (0.94 and 0.95) admit moderate N- and S Total values with comparably low CC values.

Var-cov and HS models present distance measures of similar classification, they admit moderate values of N- along with high values of S Total and low CC values.

95% VaR

The distance measures of the 95% VaR estimates for the N225 index in the pre-crisis interval are shown in Tables 3.13 and 3.14.

N225 2002 -2007_95% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student-t	QMLE	Normal	Student-t	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1278	1270	1278	1277	1273	1277	1252	1246	1162	1182	1195	1154	1179	1187
N+ (Returns < 0)	551	543	551	550	546	550	525	519	435	455	468	427	452	460
SN+ (Returns < 0)	20.910	20.470	20.910	20.919	20.591	20.919	19.956	19.788	15.223	16.140	16.484	14.937	15.839	16.186
N+ (Returns ≥ 0)	727	727	727	727	727	727	727	727	727	727	727	727	727	727
SN+ (Returns ≥ 0)	7.326	6.808	7.326	7.313	6.882	7.313	6.803	6.675	3.222	3.758	4.000	3.058	3.590	3.818
S N+ Total	28.236	27.279	28.236	28.232	27.474	28.232	26.759	26.463	18.445	19.898	20.485	17.995	19.429	20.004
N- (violations)	35	43	35	36	40	36	61	67	151	131	118	159	134	126
S N- (violations magnitude)	0.201	0.221	0.201	0.200	0.214	0.200	0.497	0.533	0.879	0.691	0.617	0.933	0.736	0.661
S total	28.438	27.499	28.438	28.431	27.688	28.431	27.256	26.997	19.324	20.589	21.101	18.928	20.165	20.664
%N-	2.666%	3.275%	2.666%	2.742%	3.046%	2.742%	4.646%	5.103%	11.500%	9.977%	8.987%	12.110%	10.206%	9.596%
CC	0.424	0.450	0.424	0.432	0.448	0.432	0.132	0.111	0.497	0.496	0.481	0.495	0.496	0.485

Table 3.13. N225 Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2002 and 2007.

N225 2002 - 2007_95% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	var-cov	HS
N+	1270	1267	1230	1234	1244	1249
N+ (Returns < 0)	543	540	503	507	517	522
SN+ (Returns < 0)	20.003	20.131	18.924	18.963	19.564	19.783
N+ (Returns ≥ 0)	727	727	727	727	727	727
SN+ (Returns ≥ 0)	6.541	6.579	5.772	5.826	6.484	6.666
S N+ Total	26.543	26.710	24.696	24.789	26.048	26.449
N- (violations)	43	46	83	79	69	64
S N- (violations magnitude)	0.253	0.244	0.565	0.560	0.561	0.511
S total	26.796	26.953	25.261	25.349	26.608	26.960
%N-	3.275%	3.503%	6.321%	6.017%	5.255%	4.874%
CC	0.439	0.450	0.219	0.214	0.095	0.131

Table 3.14. N225 Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2002 and 2007.

The distance measures in Table 3.13 and Table 3.14 show that the var-cov attains the lowest value of CC followed by the EVT (POT) for both thresholds, HS and EWMA.

The GARCH-EVT (95% and 90% threshold) and the FHS models demonstrate results with similar classification of comparatively low frequencies of violations, high CC values and also high S Total values.

GARCH-EVT (95% and 90% threshold) under normal and QMLE distributed residuals witness lower values of violations than under Student-t distributed residuals.

On the other hand, MCS and GARCH (1,1) models also show similar results with comparably high numbers of violations and high values of S Total but high CC values.

EVT (POT) (95% and 90% thresholds), var-cov and HS models admit low values of violations and high distances S Total and they also witness considerably low correlation values.

EWMA (0.94) and EWMA (0.95) both admit moderate values of violations N-, S Total and low CC values, however, as the latter shows high value of S Total, the former model admits a moderate value for this measure.

Comments

The preliminary analysis of the performance of all models for estimating VaR of the N225 index in the pre-crisis interval shows many facts about the difference between the behavior of these models with different indices. The following comments can be drawn according to the distance measures analysis carried out:

99% VaR:

- All models show at least one weak measure except for the MCS and GARCH (1,1) models under Student-t distributed residuals (EK and fitting methods) which show the only acceptable results among all models.

95% VaR:

- The quality of results provided by all models is weak and not one model shows an acceptable performance according to the classification criteria adopted in this analysis.

3.5. Crisis period 2008 -2013

3.5.1. Dow Jones Industrial Average (DJIA) index

3.5.1.1. Market Evolution

The crisis interval between 2008 and 2013 of the DJIA index is characterized by volatility of high magnitude as it can be seen in Figure 3.10. In fact, the crisis of 2008 had already started laying its heavy symptoms and consequences on the market in 2007. The peak of the crisis was reached in 2008 with the collapse of Lehman Brothers. The high clustered volatility in 2008 is obviously reflected. In 2009, the magnitude of fluctuations decreased slightly but the frequency did not witness a considerable change. In 2010, the U.S. market suffered the flash crash on the 6^{th} of May and another flash crash on August 24. Then in 2011, the stock markets in general witnessed another downturn due to fears of contagion of the European sovereign debt crisis in addition to the concerns regarding the slow economic growth of the U.S. and the downgrade of its credit rating, the severe volatility of the stock market indexes remained for the rest of 2011. In 2012 and 2013, the market witnessed a state of relatively lower volatility than in the previous years.

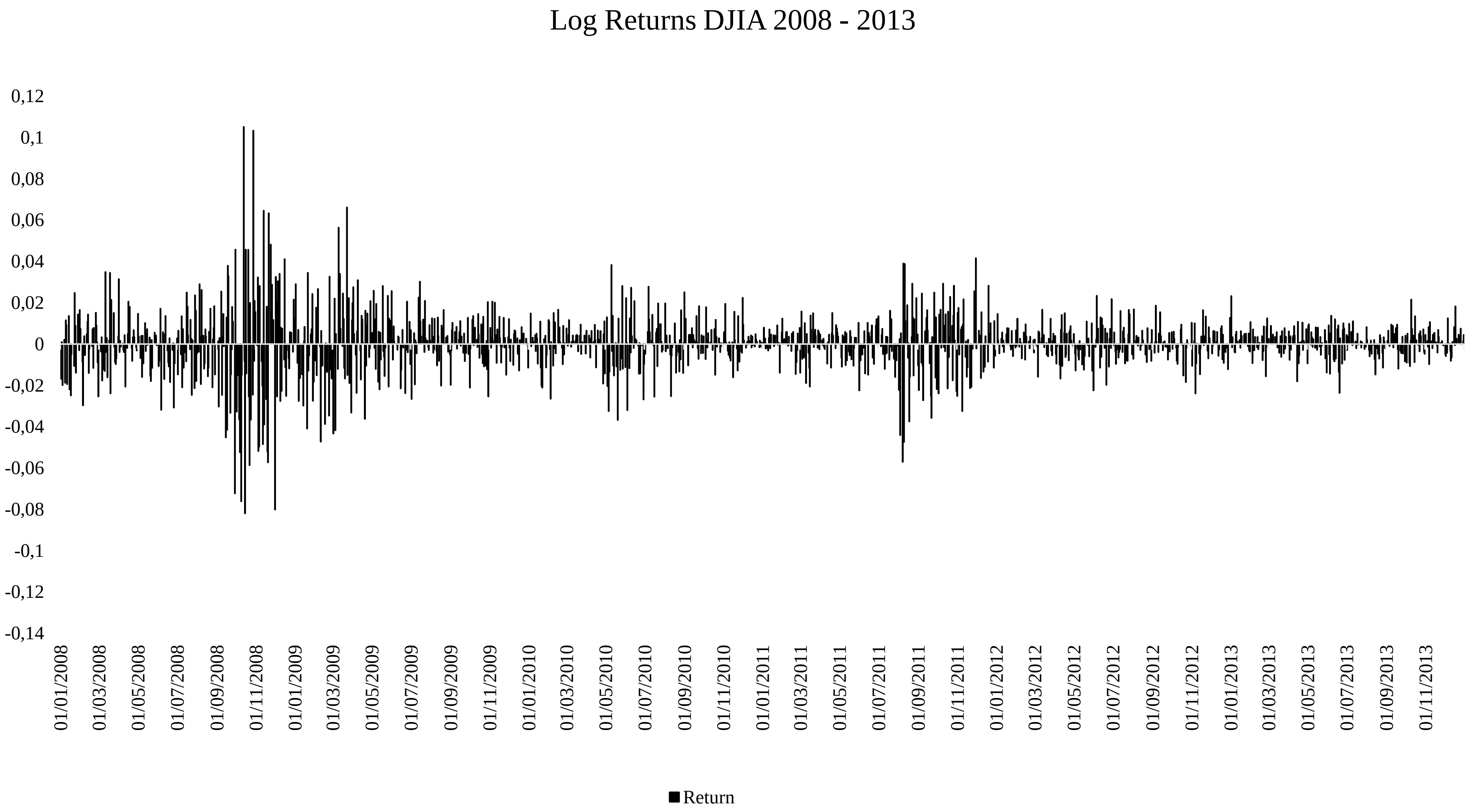

Figure 3.10. DJIA log returns between 2008 and 2013.

3.5.1.2. VaR estimates

Graphical representation

Figure 3.11 and Figure 3.12 show the graphical representation of the 99% and 95% VaR estimates, respectively, of DJIA index between 2008 and 2013. Again, the differences between the nature of the VaR curves is obvious as well as the relative behavior of some models. For instance, the var-cov, HS, and EVT (POT) models show more smooth curves with less fluctuations than the other models like MCS, GARCH-EVT, GARCH (1,1) and FHS. These smooth curves are performing especially poorly in this interval, due to the high volatility experienced as a result of the crisis.

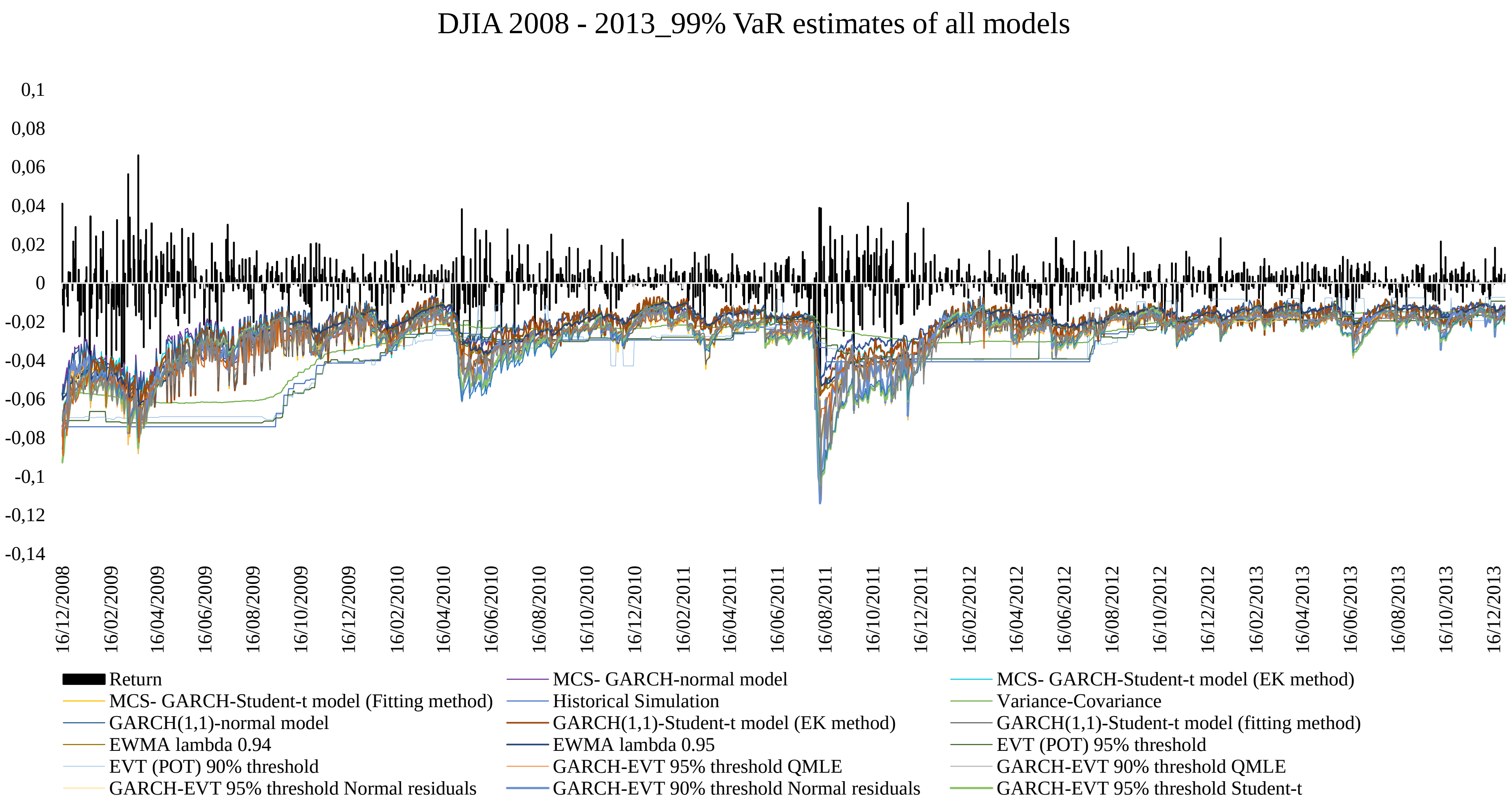

Figure 3.11. DJIA 99% VaR estimates between 2008 and 2013.

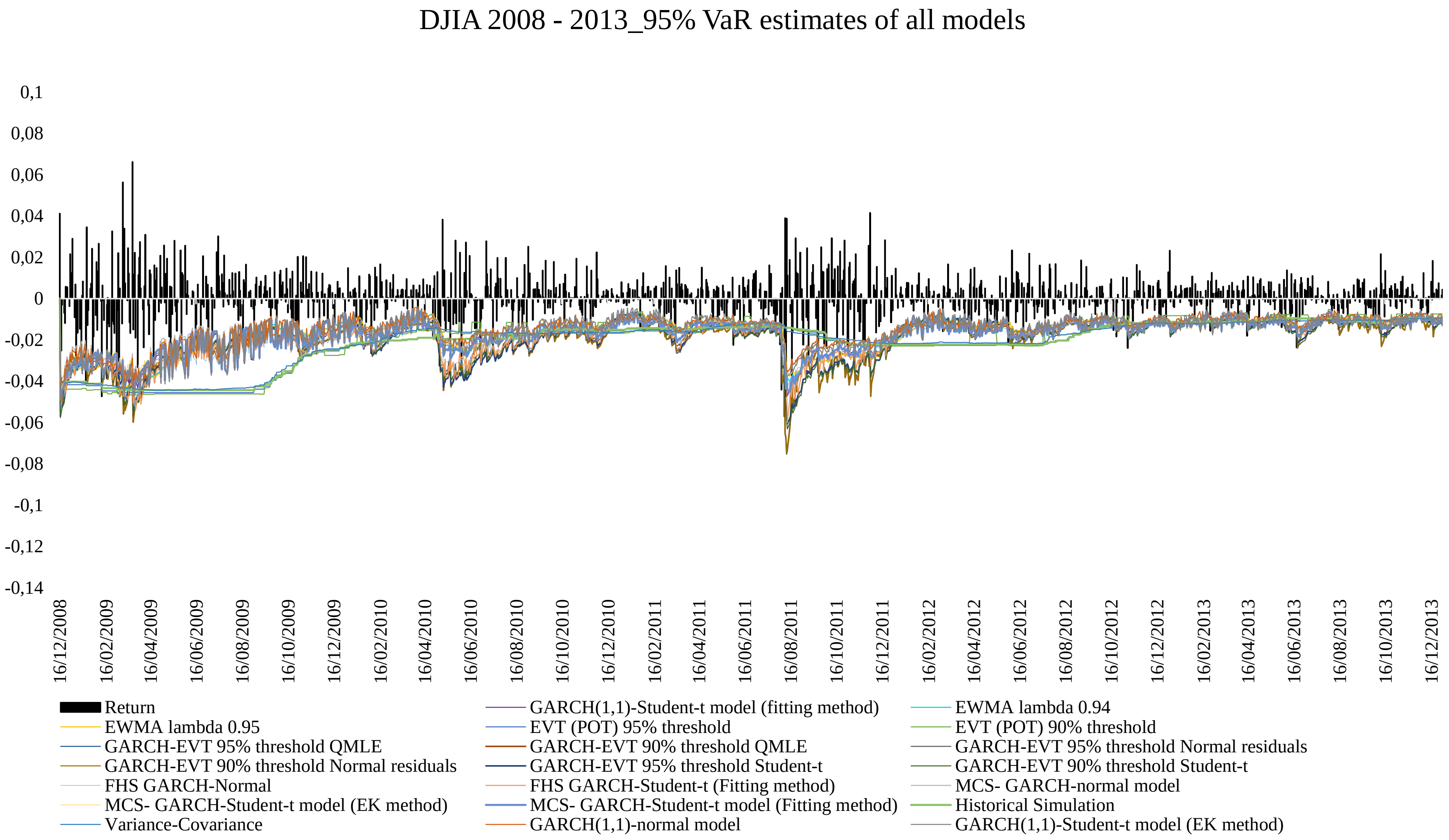

Figure 3.12. DJIA 95% VaR estimates between 2008 and 2013.

Distance Measures

Tables 3.15, 3.16 and 3.17, 3.18 present the distance measures calculated for the estimates 99% VaR and 95% VaR, respectively, of DJIA index during the crisis intervals.

99% VaR

DJIA 2008 - 2013_99% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student-t	QMLE	Normal	Student-t	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1316	1316	1316	1316	1316	1316	1305	1296	1292	1306	1314	1293	1303	1313
N+ (Returns < 0)	577	577	577	577	577	577	566	557	554	567	575	554	564	574
SN+ (Returns < 0)	26.612	26.787	26.612	26.079	26.687	26.079	30.688	29.706	20.780	23.098	25.766	20.636	22.923	25.604
N+ (Returns ≥ 0)	739	739	739	739	739	739	739	739	738	739	739	739	739	739
SN+ (Returns ≥ 0)	12.797	12.878	12.797	12.321	12.799	12.321	16.007	15.276	8.166	9.509	11.714	8.103	9.427	11.617
S N+ Total	39.409	39.665	39.409	38.401	39.485	38.401	46.695	44.982	28.947	32.606	37.480	28.739	32.349	37.221
N- (violations)	0	0	0	0	0	0	11	20	23	10	2	23	13	3
S N- (violations magnitude)	-	-	-	-	-	-	0.109	0.135	0.084	0.029	0.012	0.086	0.034	0.013
S total	39.409	39.665	39.409	38.401	39.485	38.401	46.804	45.117	29.031	32.635	37.493	28.825	32.383	37.234
%N-	0%	0%	0%	0%	0%	0%	0.836%	1.520%	1.749%	0.760%	0.152%	1.748%	0.988%	0.228%
CC	0.620	0.560	0.620	0.625	0.559	0.625	0.251	0.255	0.499	0.474	0.496	0.493	0.472	0.493

Table 3.15. DJIA Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2008 and 2013.

DJIA 2008 - 2013_99% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	var-cov	HS
N+	1315	1314	1282	1282	1298	1306
N+ (Returns < 0)	576	575	543	543	559	568
SN+ (Returns < 0)	25.695	26.111	22.066	22.226	26.437	31.445
N+ (Returns ≥ 0)	739	739	739	739	739	738
SN+ (Returns ≥ 0)	11.914	12.275	9.190	9.323	12.641	16.669
S N+ Total	37.610	38.386	31.256	31.549	39.078	48.114
N- (violations)	1	2	34	34	18	9
S N- (violations magnitude)	0.004	0.020	0.175	0.176	0.182	0.102
S total	37.613	38.406	31.431	31.725	39.261	48.216
%N-	0.076%	0.152%	2.584%	2.584%	1.368%	0.684%
CC	0.515	0.497	0.399	0.393	0.234	0.262

Table 3.16. DJIA Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2008 and 2013.

GARCH-EVT (95% and 90% threshold) models admit no violations to the estimated 99% VaR of DJIA in the crisis period, they also witness the highest values of CC among all models and moderate values of S Total.

In addition, the FHS model with normally and Student-t distributed residuals also admits low number of violations to VaR, high CC values and moderate values of S Total.

EVT (POT) (95% and 90% threshold) admit moderate values of N-, high value of S Total and low values of CC.

MCS models with Student-t distributed residuals (EK and fitting methods) show low numbers of violations, accompanied by low and moderate values of the distance measure S Total, with also moderate and high CC values respectively. Under normally distributed residuals, MCS witnesses high number of violations, low value of S Total and high value of CC.

GARCH (1,1) model shows better performance with Student-t distributed residuals (EK and fitting methods) than with normally distributed residuals showing a similar performance to the MCS model.

EWMA (0.94 and 0.95) models witness high frequencies of violations accompanied by moderate values of CC and low values of S Total.

Var-cov however witnesses a high number of violations, moderate S Total, and low CC value.

HS model shows a low number of violations however, it admits high S Total value with low CC.

95% VaR

DJIA 2008 – 2013_95% VaR														
Distance Measures	**GARCH-EVT 95% threshold**			**GARCH-EVT 90% threshold**			**EVT (POT)**		**Monte Carlo Simulation**			**GARCH (1,1)**		
	Normal	**Student-t**	**QMLE**	**Normal**	**Student-t**	**QMLE**	**95% Threshold**	**90% Threshold**	**GARCH-Normal**	**GARCH-t EK method**	**GARCH-t Fitting method**	**Normal**	**t EK**	**t Fitting**
N+	1293	1286	1293	1299	1284	1299	1269	1262	1236	1234	1260	1233	1233	1252
N+ (Returns < 0)	554	547	554	560	545	560	530	523	498	495	521	494	494	513
SN+ (Returns < 0)	19.370	19.174	19.370	19.667	19.203	19.667	20.641	20.219	16.228	16.971	17.848	16.052	16.789	17.662
N+ (Returns ≥ 0)	739	739	739	739	739	739	739	739	738	739	739	739	739	739
SN+ (Returns ≥ 0)	6.949	6.808	6.949	7.205	6.802	7.205	8.243	7.937	4.829	5.096	5.707	4.749	5.013	5.611
S N+ Total	26.319	25.982	26.319	26.872	26.005	26.872	28.885	28.156	21.057	22.067	23.555	20.801	21.802	23.274
N- (violations)	23	30	23	17	32	17	47	54	79	82	56	83	83	64
S N- (violations magnitude)	0.053	0.096	0.053	0.043	0.085	0.043	0.363	0.412	0.426	0.397	0.283	0.433	0.411	0.287
S total	26.372	26.078	26.372	26.915	26.090	26.915	29.248	28.568	21.483	22.464	23.838	21.234	22.213	23.561
%N-	1.748%	2.280%	1.748%	1.292%	2.432%	1.292%	3.571%	4.103%	6.008%	6.231%	4.255%	6.307%	6.307%	4.863%
CC	0.614	0.548	0.614	0.614	0.556	0.614	0.241	0.247	0.499	0.473	0.489	0.490	0.471	0.489

Table 3.17. DJIA Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2008 and 2013.

DJIA 2008 – 2013_95% VaR						
Distance Measures	**Filtered Historical Simulation**		**EWMA**		**Variance-covariance**	**Historical Simulation**
	GARCH-Normal	**GARCH-t Fitting**	**Lambda 0.94**	**Lambda 0.95**	**var-cov**	**HS**
N+	1275	1280	1233	1235	1265	1263
N+ (Returns < 0)	536	541	494	496	526	525
SN+ (Returns < 0)	18.742	18.860	17.063	17.176	20.153	20.330
N+ (Returns ≥ 0)	739	739	739	739	739	738
SN+ (Returns ≥ 0)	6.554	6.562	5.511	5.597	7.854	8.016
S N+ Total	25.296	25.422	22.573	22.773	28.008	28.345
N- (violations)	41	36	83	81	51	52
S N- (violations magnitude)	0.169	0.172	0.489	0.481	0.398	0.376
S total	25.465	25.594	23.062	23.254	28.406	28.722
%N-	3.116%	2.736%	6.307%	6.155%	3.875%	3.954%
CC	0.492)	0.493	0.399	0.393	0.236	0.240

Table 3.18. DJIA Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2008 and 2013.

According to Table 3.17 and Table 3.18, GARCH-EVT (95% threshold) model admits high CC values accompanied by considerably low numbers of violations to the estimated 95% VaR and comparably moderate values of S Total under normal, Student-t and QMLE distribution of residuals.

However, GARCH-EVT (90% threshold) admits low numbers of violations under the normal, Student-t and QMLE distributed residuals accompanied by high values of S Total under normal and QMLE distributed residuals while with Student-t residuals the value of S Total is considerably moderate, and under all residuals the CC values are considerably high. It is worth noting that GARCH-EVT (95% and 90% thresholds) models admit higher CC values and lower N- values under normal and QMLE fitted residuals than under the Student-t residuals.

FHS under normally distributed residuals witnesses moderate number of violations and a moderate S Total value with a high CC value, however, under Student-t residuals it admits a lower number of violations, with a slightly higher value of correlation and a moderate value of S Total.

MCS model admits high values of N- under normal and Student-t distributed residuals (EK method) while under Student-t (fitting method) it attains a moderate frequency of violations. Under all residuals MCS admits low S Total values however under normal residuals it attains a CC value classified high while those values obtained under the Student-t residuals are classified moderate.

GARCH (1,1) model under all residuals admits high numbers of violations, low values of S Total and moderate values of CC, however under the normally distributed residuals its corresponding CC value is slightly high.

EVT (POT) (95% and 90% thresholds) along with var-cov and HS models attain moderate numbers of violations and high S Total with considerably low values of CC.

EWMA (0.94 and 0.95) witness high numbers of violations and low values of S Total with comparably moderate values of CC.

Comments

According to the distance measures analysis of the VaR estimates of DJIA in the crisis period obtained under the models implemented in this thesis, some comments can be given as follows:

99% VaR:

- GARCH-EVT (95% threshold), GARCH-EVT (90% threshold) under Student-t distributed residual, FHS and MCS and GARCH (1,1) models with Student-t residuals (EK and fitting method) outperform the rest of the models.

95% VaR:

- GARCH-EVT (95% threshold) model, GARCH-EVT (90% threshold) model under Student-t residuals, MCS under Student-t residuals (fitting method), FHS under Student-t residuals and FHS under normally distributed residuals outperform the reminder models.
- The remainder models might show one or more distance measures with good values but also show other values that reveal weak performance.

3.5.2. Euro Stoxx 50 (SX5E) index

3.5.2.1. Market Evolution

The SX5E index returns witnessed high and clustered volatility during the crisis interval between 2008 and 2013. Figure 3.13 shows that the market was in a continuous state of high fluctuations, and more than the volatility witnessed in the same interval by the DJIA index, as the high volatility was not only present during 2008 and 2009 which are the peak crisis years, but also during the entire interval. The high and clustered volatility in 2008 and 2009 can be referred to the financial crisis of 2008, in addition to the sovereign debt crisis of Greece which started in late 2009 which was triggered by the world-wide recession, structural weakness in the Greek economy and lack of monetary policy flexibility as a member of the Eurozone. Due to the stock market downturn in 2011, the returns reached their lowest trough. In 2012 and 2013, the market fluctuations were of less severity than in 2011 and the previous year however the clustered volatility remained but with a slightly lower magnitude.

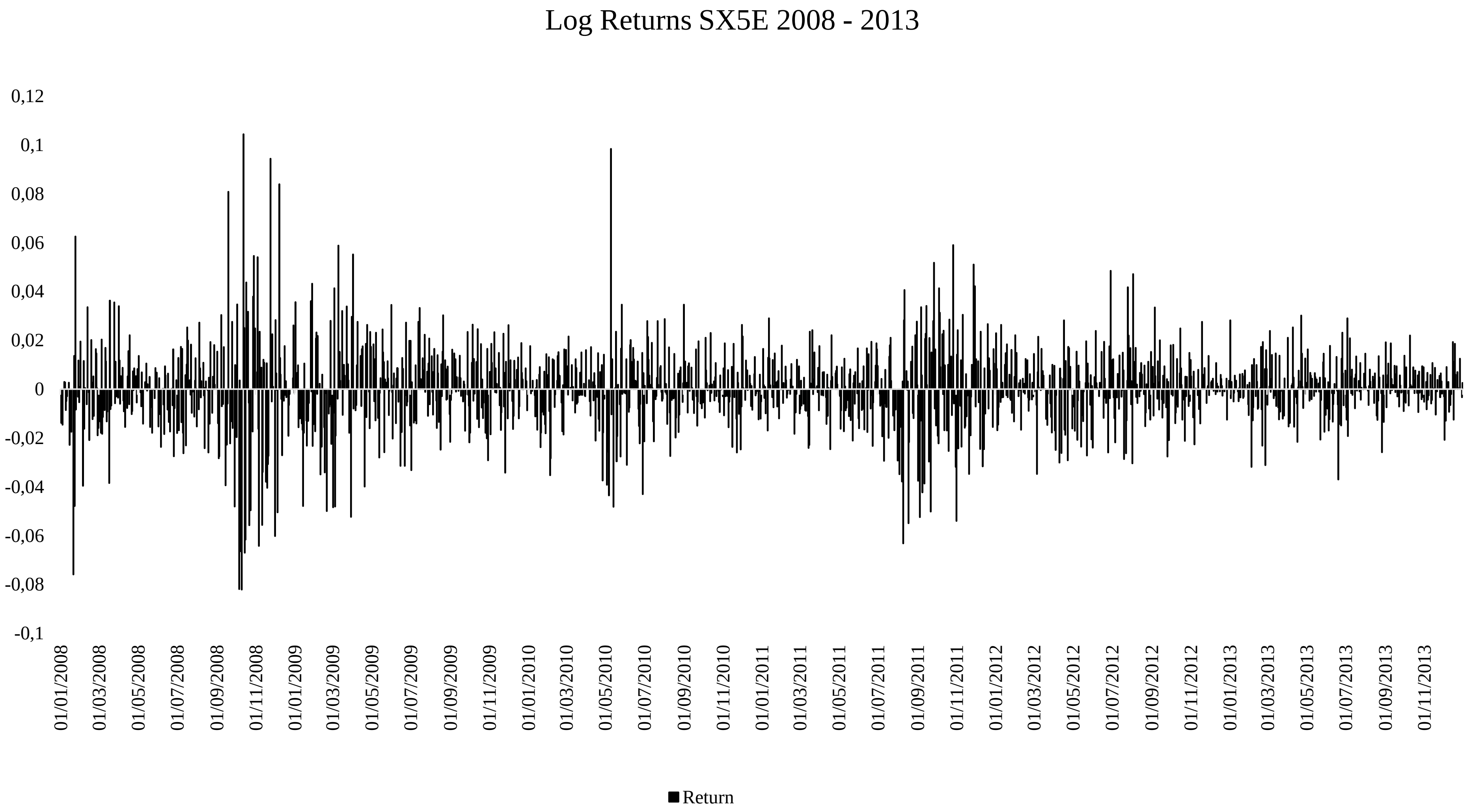

Figure 3.13. SX5E log returns between 2008 and 2013.

3.5.2.2. VaR estimates

Graphical representation

Figure 3.14 and Figure 3.15 demonstrate the 99% and 95% VaR estimates of the SX5E index obtained under all models in the crisis interval. The graphical representation of the VaR estimates shows that the curves of the HS, var-cov and EVT (POT) models acquire shallow curves that don't respond quickly (properly) to market movements. Consequently, these models suffer a higher number of violations than the other models whose fluctuating curves rhyme with the market changes, volatility and tranquility. Here lies the importance of distance measures that will be interpreted and analyzed in the following section to show which models yield VaR estimates that are mostly coherent with the losses mainly which is important for regulators, and also with the gains which is highly important for managers who would not like to save high capitals for overestimated predictions of losses.

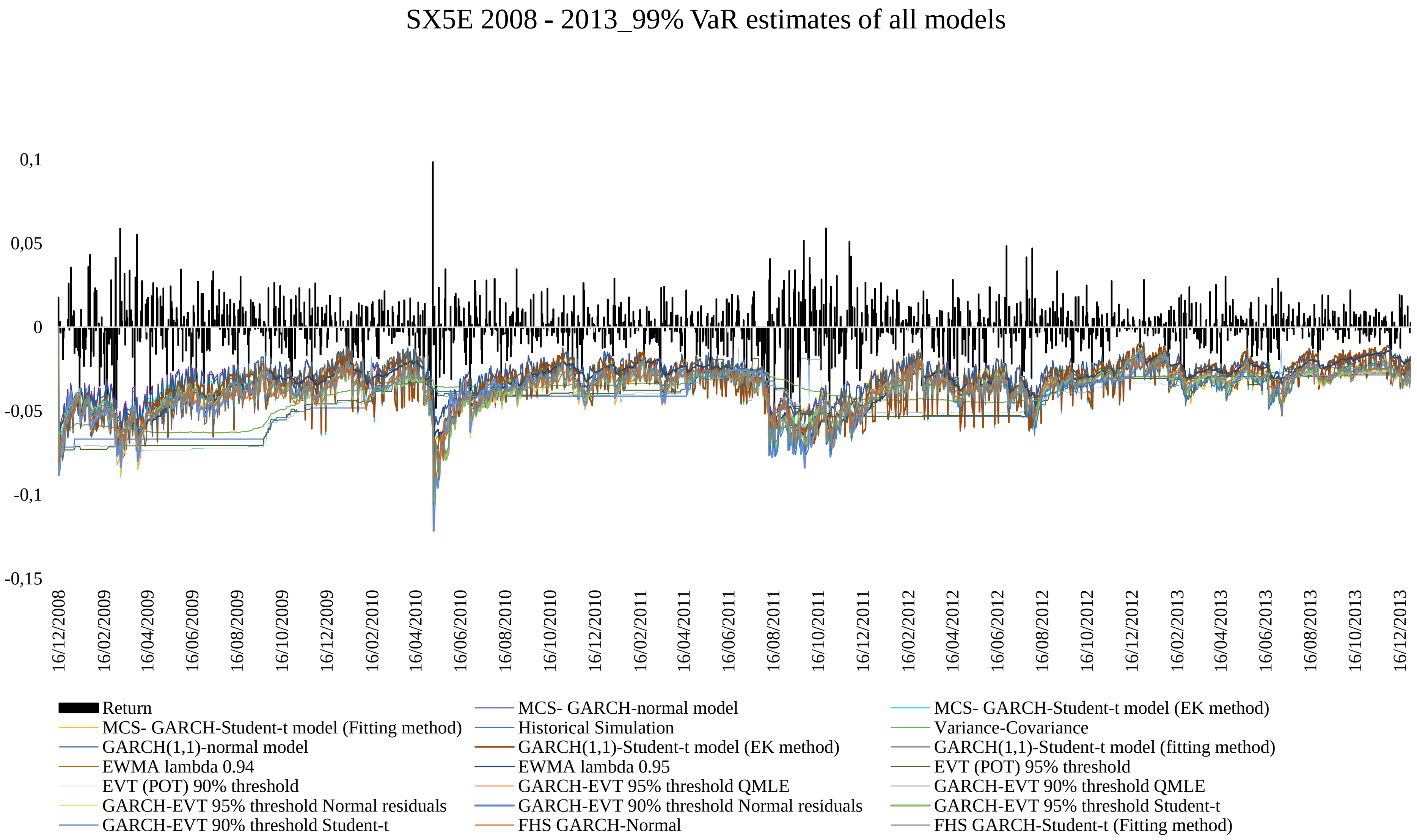

Figure 3.14. SX5E 99% VaR estimates between 2008 and 2013.

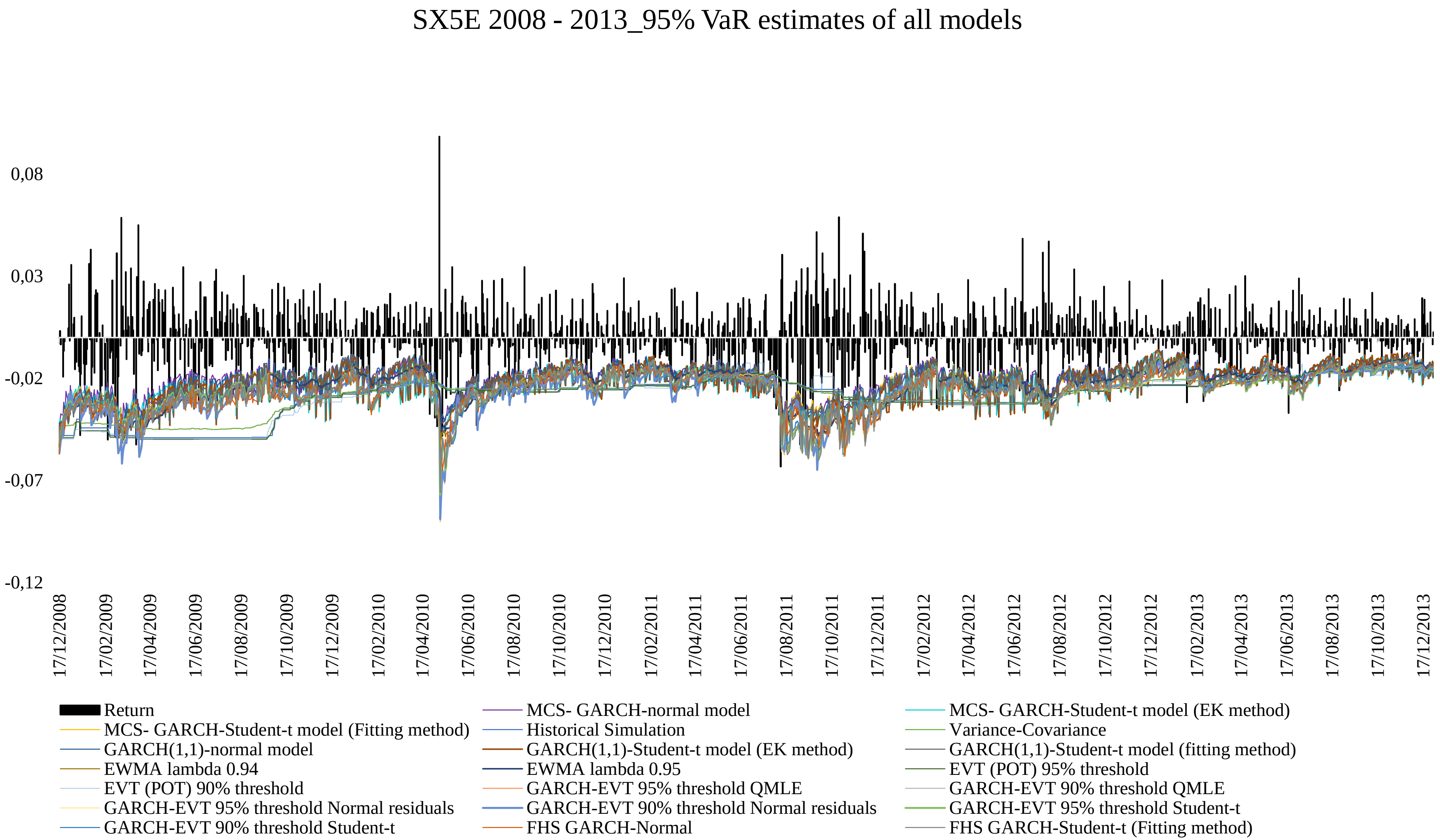

Figure 3.15. SX5E 95% VaR estimates between 2008 and 2013.

Distance Measures

99% VaR

Table 3.19 and Table 3.20 present the distance measures of the 99% VaR estimates obtained under all models for the SX5E index in the crisis period between 2008 and 2013.

SX5E 2008 - 2013_99% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1)		
	Normal	Student-t	QMLE	Normal	Student-t	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1314	1314	1314	1314	1315	1314	1298	1298	1281	1308	1304	1285	1309	1305
N+ (Returns < 0)	646	646	646	646	647	646	630	630	613	640	636	617	641	637
SN+ (Returns < 0)	32.764	32.492	32.764	32.661	32.621	32.661	36.316	36.316	26.766	30.290	30.724	26.871	30.385	30.828
N+ (Returns ≥ 0)	668	668	668	668	668	668	668	668	668	668	668	668	668	668
SN+ (Returns ≥ 0)	17.635	17.280	17.635	17.520	17.322	17.520	21.591	21.591	12.142	15.348	15.666	12.269	15.440	15.780
S N+ Total	50.399	49.772	50.399	50.181	49.943	50.181	57.908	57.908	38.907	45.638	46.391	39.140	45.825	46.608
N- (violations)	2	2	2	2	1	2	18	18	35	8	12	31	7	11
S N- (violations magnitude)	0.005	0	0.005	0.003	0.002	0.003	0.121	0.121	0.152	0.029	0.032	0.136	0.025	0.033
S total	50.403	49.773	50.403	50.184	49.945	50.184	58.028	58.028	39.060	45.667	46.423	39.276	45.850	46.641
%N-	0.152%	0.152%	0.152%	0.152%	0.076%	0.152%	1.368%	1.368%	2.660%	0.608%	0.912%	2.356%	0.532%	0.836%
CC	0.476	0.425	0.476	0.478	0.430	0.478	0.155	0.133	0.392	0.399	0.362	0.393	0.403	0.361

Table 3.19. SX5E Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2008 and 2013.

SX5E 2008 - 2013_ 99% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	Var-cov	HS
N+	1312	1311	1287	1288	1295	1300
N+ (Returns < 0)	644	643	619	620	627	632
SN+ (Returns < 0)	31.874	32.108	29.334	29.468	32.870	36.240
N+ (Returns ≥ 0)	668	668	668	668	668	668
SN+ (Returns ≥ 0)	16.962	16.987	14.604	14.763	18.218	21.485
S N+ Total	48.836	49.094	43.938	44.231	51.088	57.725
N- (violations)	4	5	29	28	21	16
S N- (violations magnitude)	0.009	0.007	0.156	0.152	0.210	0.117
S total	48.845	49.101	44.094	44.383	51.298	57.842
%N-	0.304%	0.380%	2.204%	2.128%	1.596%	1.216%
CC	0.416	0.419	0.287	0.280	0.131	0.148

Table 3.20. SX5E Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2008 and 2013.

GARCH-EVT (95% and 90% threshold), FHS under normal and Student-t residuals, as well as MCS and GARCH (1,1) under Student-t residuals (EK method) models outperform the remainder models as they admit the lowest number of violations to VaR, highest CC values and moderate S Total values.

MCS and GARCH (1,1) models under normal residuals admit similar classifications of their distance measures as they both admit high values of N-, low values of S Total and high CC values. Under Student-t distributed residuals (fitting method) GARCH (1,1) witnesses a low value of N- as well while MCS attains a moderate value of violations and both models admit moderate values of S Total. Regarding the CC values, MCS attains a higher value than GARCH (1,1) whose CC value is considerably moderate.

EVT (POT) (95% and 90% thresholds) and HS admit moderate frequencies of violations to VaR with considerably low values of CC and high values of S Total.

EWMA (0.94 and 0.95) models admit moderate values of CC and low S Total values with high numbers of violations.

However, var-cov admits a moderate value of violations and S Total with low values of CC.

95% VaR

SX5E 2008 - 2013_ 95% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1) model		
	Normal	Student-t	QMLE	Normal	Student-t	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH -t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1284	1268	1284	1286	1266	1286	1269	1269	1229	1237	1238	1231	1238	1236
N+ (Returns < 0)	616	600	616	618	598	618	601	601	561	569	570	563	570	568
SN+ (Returns < 0)	25.054	24.387	25.054	25.062	24.376	25.062	26.287	26.287	21.046	22.090	22.175	21.117	22.156	22.249
N+ (Returns ≥ 0)	668	668	668	668	668	668	668	668	668	668	668	668	668	668
SN+ (Returns ≥ 0)	10.386	9.696	10.386	10.412	9.737	10.412	12.050	12.050	7.148	7.907	7.987	7.235	7.988	8.066
S N+ Total	35.441	34.082	35.441	35.473	34.112	35.473	38.337	38.337	28.194	29.997	30.163	28.353	30.144	30.315
N- (violations)	32	48	32	30	50	30	47	47	87	79	78	85	78	80
S N- (violations magnitude)	0.142	0.237	0.142	0.136	0.240	0.136	0.452	0.452	0.673	0.502	0.514	0.649	0.469	0.487
S total	35.582	34.319	35.582	35.610	34.352	35.610	38.789	38.789	28.867	30.499	30.677	29.002	30.613	30.802
%N-	2.432%	3.647%	2.432%	2.280%	3.799%	2.280%	3.571%	3.571%	6.611%	6.003%	5.927%	6.459%	5.927%	6.079%
CC	0.483	0.436	0.483	0.485	0.431	0.485	0.129	0.120	0.391	0.398	0.388	0.393	0.404	0.389

Table 3.21. SX5E Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2008 and 2013.

SX5E 2008 – 2013_95% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	Var-cov	HS
N+%	1270	1267	1235	1235	1262	1267
N+ (Returns < 0)	602	599	567	567	594	599
SN+ (Returns < 0)	24.297	24.418	22.859	22.954	25.359	25.975
N+ (Returns ≥ 0)	668	668	668	668	668	668
SN+ (Returns ≥ 0)	9.777	9.711	8.817	8.928	11.251	11.760
S N+ Total	34.073	34.129	31.676	31.882	36.610	37.735
N- (violations)	46	49	81	81	54	49
S N- (violations magnitude)	0.228	0.218	0.594	0.590	0.511	0.471
S total	34.301	34.347	32.271	32.471	37.122	38.206
%N-	3.495%	3.723%	6.155%	6.155%	4.103%	3.723%
CC	0.440	0.438	0.288	0.282	0.136	0.127

Table 3.22. SX5E Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2008 and 2013

In Table 3.21 and Table 3.22, the best results are achieved under the GARCH-EVT (95% and 90% thresholds) under Student-t residuals as well as the FHS model with normal and Student-t residuals.

GARCH-EVT (95% threshold) under Student-t fitted residuals along with FHS under normally distributed residuals attain low frequencies of violations with high CC values and considerably moderate S Total values.

On the other hand, GARCH-EVT (90% threshold) and FHS, both under Student-t distributed residuals admit moderate frequencies of violations, moderate S Total values, and high CC values.

However, under normal and QMLE fitted residuals, GARCH-EVT (95% as well as 90% threshold) admit low frequencies of violations, high CC values, however, their corresponding S Total values are considerably high.

MCS and GARCH (1,1) models under all residuals witness high values of violations accompanied by low values of S Total and high CC values.

EVT (POT) (95% threshold and 90% threshold) and HS models admit low values of violations with high distances S Total and considerably low values of CC.

EWMA (0.94 and 0.95) models admit high frequencies of violations and moderate values of CC and S Total.

Var-cov admits moderate values of violations, high values of S total and low valor CC .

Comments:

According to the analysis of the 99% VaR and 95% VaR distance measures of the SX5E index during the crisis period, some comments can be made:

99% VaR:

- The best performing models are the GARCH-EVT (95% and 90% thresholds) under all residuals, FHS with all residuals and MCS and GARCH (1,1) under Student-t distributed residuals (EK and fitting methods).
- The remainder models admit at least one distance measure value that reveals their weakness in estimating the corresponding VaR.

95% VaR:

- The best performing models for estimating the 95% VaR according to distance measures are the GARCH-EVT (95% and 90% thresholds) under Student-t distributed residuals along with the FHS model under normal and Student-t distributed residuals.

3.5.3. Nikkei 225 (N225) index

3.5.3.1. Market evolution

Figure 3.16 displaying the N225 index log returns between 2008 and 2013 shows how its corresponding market responded differently during the period of the crisis. As it can be seen, the high magnitude violations during the crisis in 2008 were followed by fluctuations with a lower magnitude especially when compared to those of SX5E and DJIA indexes. On March 15, 2011, the N225 index witnessed a severe drop due to the massive earthquake in Japan's Fukushima which triggered a tsunami, after which the market showed volatility of lower magnitude. Then, in May 2013, the Japanese market once again suffered another crash which was due to the investors being worried by weak economic data from China and indications that the U.S. Federal Reserve will start slowing down its bond-buying program as early as June 2013.

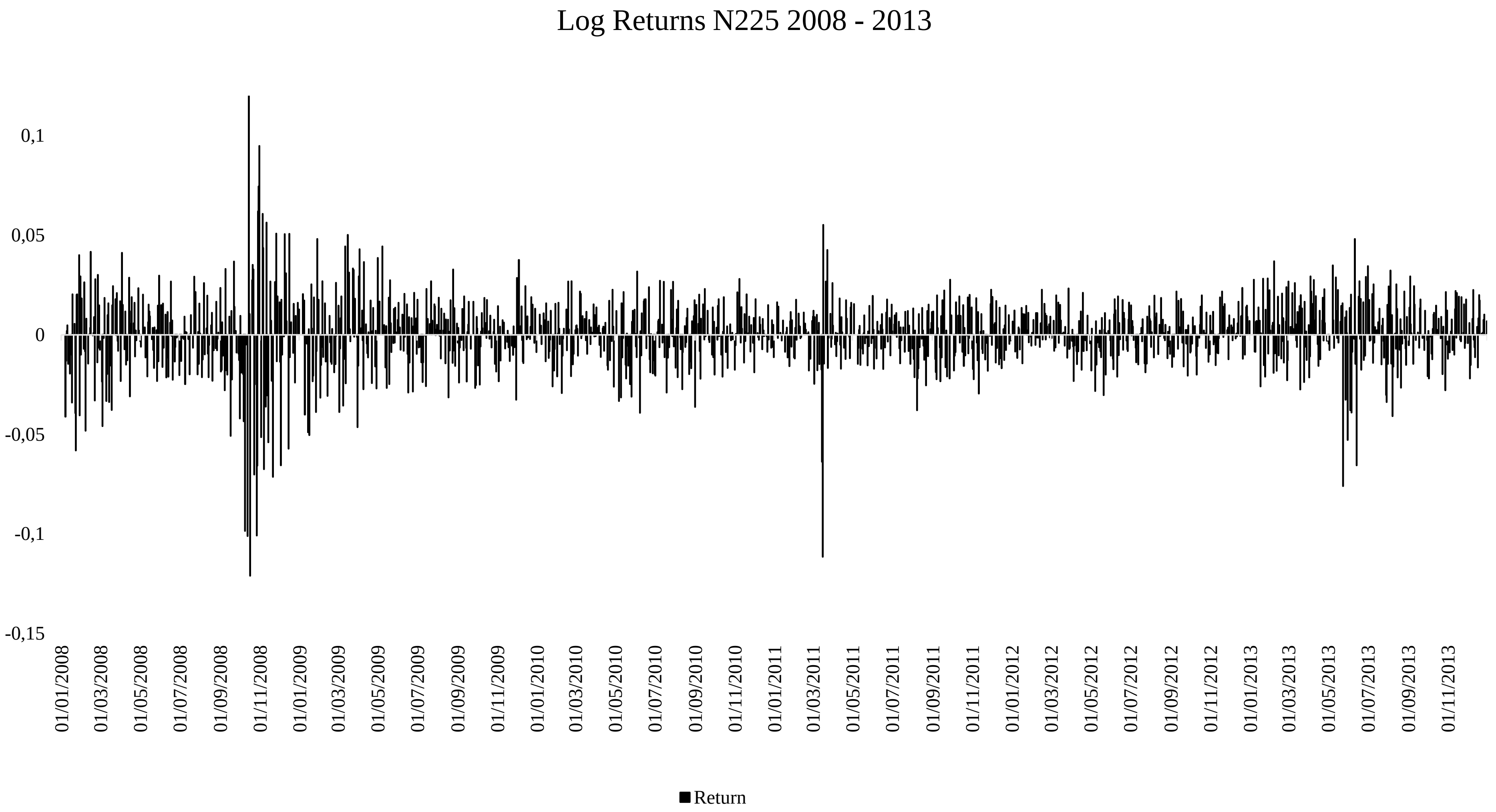

Figure 3.16. N225 log returns between 2008 and 2013.

3.5.3.2. VaR estimates

Graphical representation

Figures 3.17 and 3.18 display the graphical representation of the 99% and 95% VaR estimates of the N225 index in the crisis interval respectively. The features of the plotted VaR curves look similar to those of the DJIA and SX5E VaR curves, for instance, the HS and EVT (POT) and the var-cov models admit curves with slow response to market movements causing their corresponding curves to be almost flat with a constant trend that dominates their flow throughout the entire interval. On the other hand, the other models seem to reflect better the market fluctuations although they lie very close to the plotted returns. The distance measures analysis will reveal the differences between the implemented models in estimating VaR.

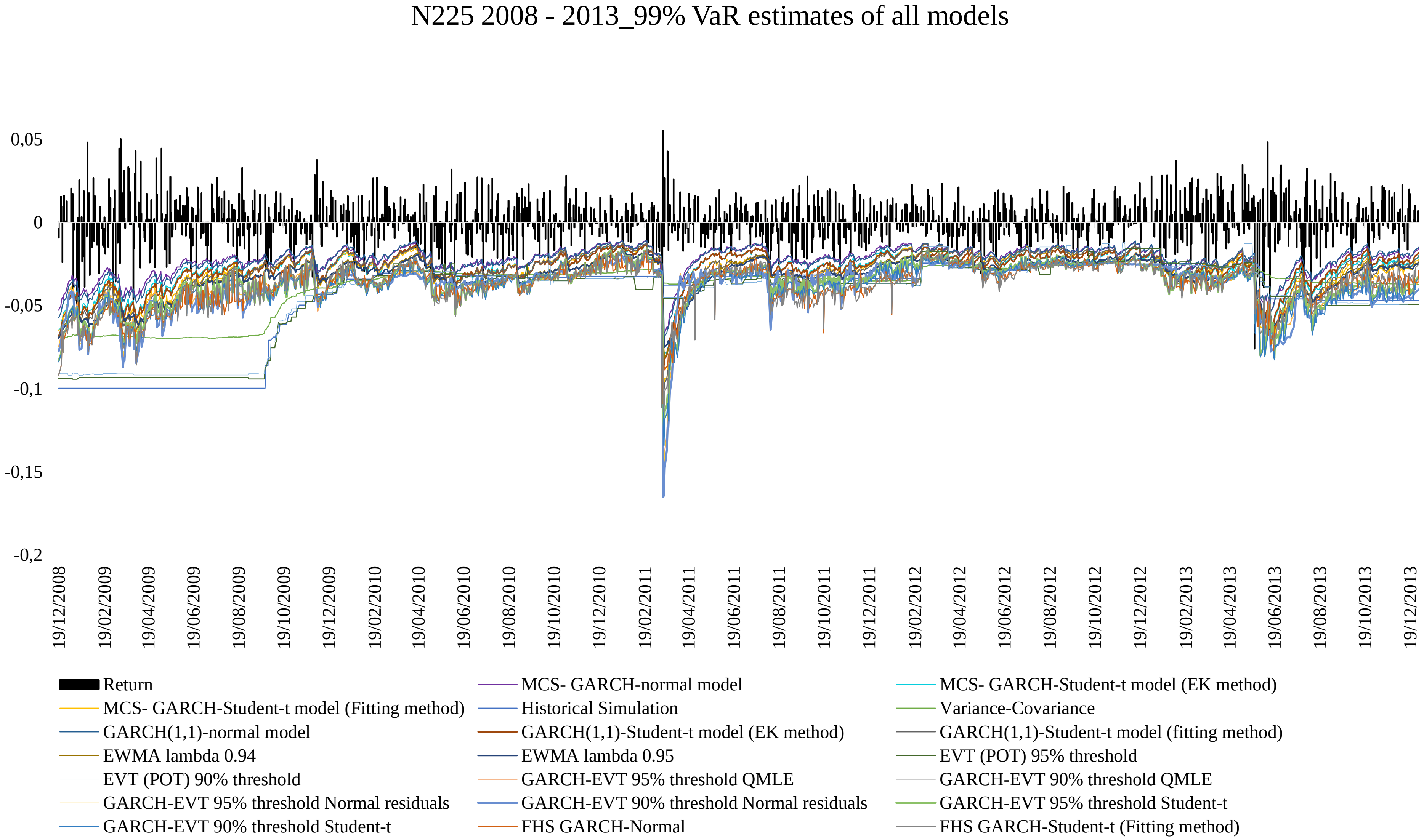

Figure 3.17. N225 95% VaR estimates between 2008 and 2013.

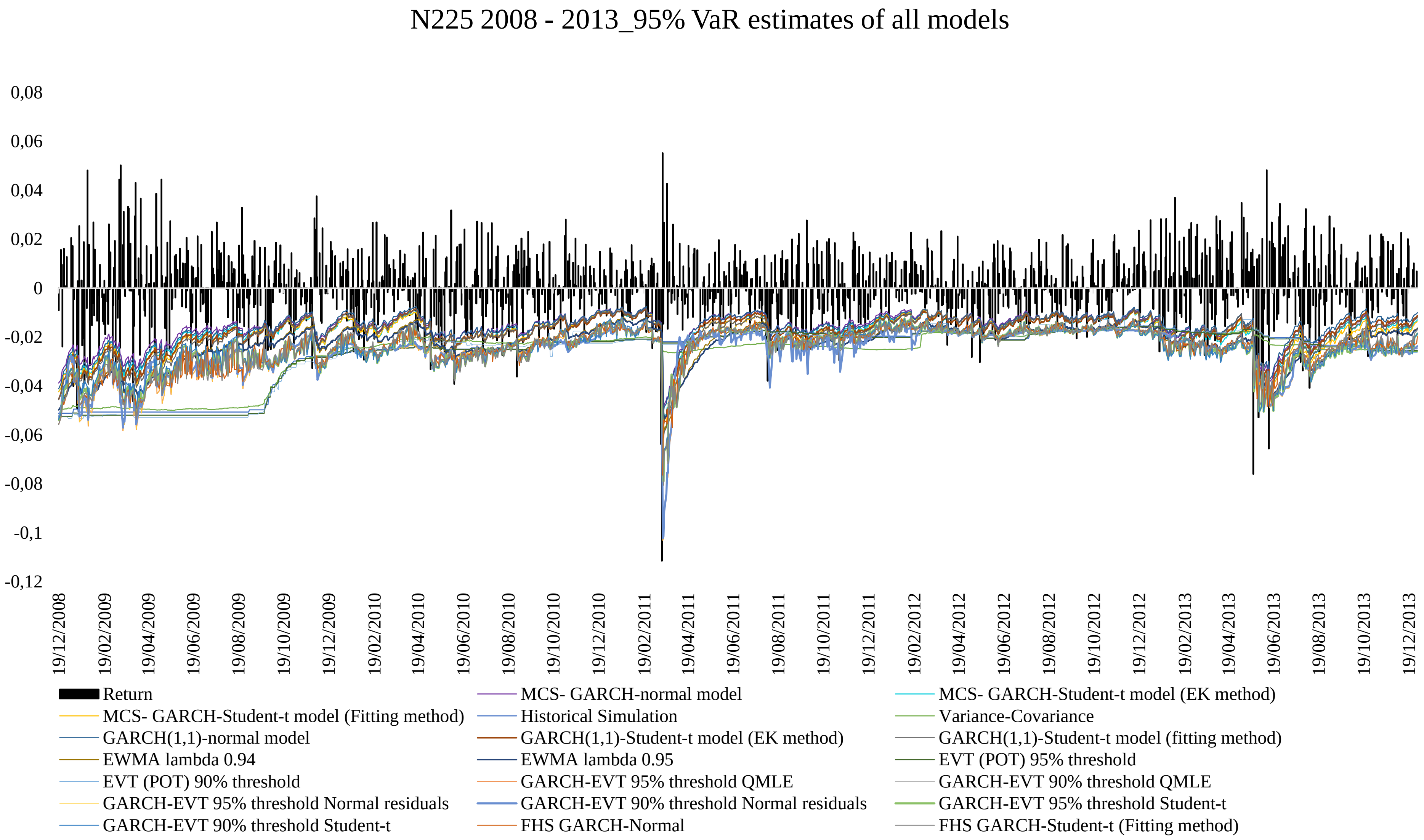

Figure 3.18. N225 95% VaR estimates between 2008 and 2013.

Distance Measures

99% VaR

N225 2008 - 2013_99% VaR																
Distance Measures	**GARCH-EVT 95% threshold**			**GARCH-EVT 90% threshold**			**EVT (POT)**		**Monte Carlo Simulation**			**GARCH (1,1) model**		
	Normal	**Student-t**	**QMLE**	**Normal**	**Student-t**	**QMLE**	**95% Threshold**	**90% Threshold**	**GARCH-Normal**	**GARCH-t EK method**	**GARCH-t Fitting method**	**Normal**	**t EK**	**t Fitting**
N+	1310	1311	1310	1312	1313	1312	1297	1295	1255	1298	1295	1255	1295	1296
N+ (Returns < 0)	581	582	581	583	584	583	568	566	526	569	566	526	566	567
SN+ (Returns < 0)	33.560	33.760	33.560	34.543	34.732	34.543	39.227	39.142	23.640	27.003	28.039	23.573	26.941	27.956
N+ (Returns ≥ 0)	729	729	729	729	729	729	729	729	729	729	729	729	729	729
SN+ (Returns ≥ 0)	15.021	15.249	15.021	15.733	16.023	15.733	18.970	18.859	7.251	9.842	10.685	7.262	9.802	10.655
S N+ Total	48.581	49.009	48.581	50.276	50.755	50.276	58.197	58.000	30.891	36.845	38.723	30.835	36.743	38.611
N- (violations)	4	3	4	2	1	2	17	19	59	16	19	59	19	18
S N- (violations magnitude)	0.030	0.019	0.030	0.024	0.009	0.024	0.189	0.199	0.308	0.104	0.080	0.312	0.108	0.080
S total	48.610	49.028	48.610	50.300	50.764	50.300	58.386	58.200	31.199	36.949	38.804	31.147	36.852	38.691
%N-	0.304%	0.228%	0.304%	0.152%	0.076%	0.152%	1.294%	1.446%	4.490%	1.218%	1.446%	4.490%	1.446%	1.370%
CC	0.464	0.406	0.464	0.452	0.404	0.452	0.172	0.169	0.450	0.450	0.428	0.443	0.444	0.425

Table 3.23. N225 Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2008 and 2013.

N225 2008 - 2013_ 99% VaR						
Distance Measures	**Filtered Historical Simulation**		**EWMA**		**Variance-covariance**	**Historical Simulation**
	GARCH-Normal	**GARCH-t Fitting**	**Lambda 0.94**	**Lambda 0.95**	**Var-cov**	**HS**
N+	1310	1310	1292	1292	1295	1298
N+ (Returns < 0)	581	581	563	563	566	569
SN+ (Returns < 0)	34.224	34.186	30.559	30.737	34.498	39.364
N+ (Returns ≥ 0)	729	729	729	729	729	729
SN+ (Returns ≥ 0)	15.484	15.591	12.690	12.824	15.425	18.990
S N+ Total	49.709	49.777	43.249	43.561	49.923	58.354
N- (violations)	4	4	22	22	19	16
S N- (violations magnitude)	0.037	0.039	0.241	0.244	0.271	0.225
S total	49.746	49.816	43.490	43.806	50.194	58.579
%N-	0.304%	0.304%	1.674%	1.674%	1.446%	1.218%
CC	0.421	0.403	0.189	0.187	0.152	0.158

Table 3.24. N225 Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2008 and 2013.

It can be seen in Table 3.23 and Table 3.24 that MCS and GARCH (1,1) models under Student-t residuals (EK and fitting methods) admit ideal values of distance measures compared to other models, with low numbers of violations, high CC values and low S Total values. While the same models under normally distributed residuals also admit good values for CC and S Total however, they witness high frequencies of violations.

On the other hand, GARCH-EVT (95% threshold) under all residuals witness low values of N-, high values of CC and moderate values of S Total, noting that the CC values are higher under normal and QMLE fitted residuals than under Student-t residuals.

GARCH-EVT (90% threshold) model under all residuals and FHS model (under normal and Student-t residuals) demonstrate distance measures of similar quality with low frequencies of violations, high CC but also high S Total values.

Meanwhile, EVT (POT) (95% and 90% thresholds) along with the HS and var-cov models also witness low frequencies of violations compared to other models however they admit very low values of CC and high values of S Total.

EWMA (0.94 and 0.95) attain moderate vales of N- and S Total with very low values of correlation.

95% VaR

N225 2008 -2013_95% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1) model		
	Normal	Student-t	QMLE	Normal	Student-t	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1284	1273	1284	1285	1273	1285	1266	1258	1165	1189	1192	1163	1188	1193
N+ (Returns < 0)	555	544	555	556	544	556	537	529	436	460	463	434	459	464
SN+ (Returns < 0)	25.657	24.868	25.657	25.534	24.733	25.534	26.371	26.155	18.812	20.100	20.439	18.730	20.014	20.344
N+ (Returns ≥ 0)	729	729	729	729	729	729	729	729	729	729	729	729	729	729
SN+ (Returns ≥ 0)	8.671	8.134	8.671	8.605	8.024	8.605	9.245	9.093	4.062	4.855	5.099	4.044	4.832	5.071
S N+ Total	34.328	33.001	34.328	34.139	32.757	34.139	35.616	35.248	22.875	24.955	25.537	22.774	24.846	25.415
N- (violations)	30	41	30	29	41	29	48	56	149	125	122	151	126	121
S N- (violations magnitude)	0.191	0.252	0.191	0.195	0.267	0.195	0.562	0.587	1.032	0.777	0.730	1.041	0.783	0.734
S total	34.519	33.253	34.519	34.333	33.024	34.333	36.178	35.835	23.907	25.732	26.267	23.815	25.629	26.150
%N-	2.283%	3.120%	2.283%	2.207%	3.120%	2.207%	3.653%	4.262%	11.339%	9.513%	9.285%	11.492%	9.589%	9.209%
CC	0.466	0.408	0.466	0.471	0.413	0.471	0.149	0.153	0.448	0.447	0.433	0.441	0.442	0.434

Table 3.25. N225 Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2008 and 2013.

N225 2008 – 2013_95% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	Var-cov	HS
N+	1269	1275	1237	1235	1267	1264
N+ (Returns < 0)	540	546	508	506	538	535
SN+ (Returns < 0)	24.816	24.778	23.669	23.795	26.454	26.071
N+ (Returns ≥ 0)	729	729	729	729	729	729
SN+ (Returns ≥ 0)	7.960	8.090	7.502	7.579	9.345	9.023
S N+ Total	32.776	32.868	31.172	31.374	35.799	35.094
N- (violations)	45	39	77	79	47	50
S N- (violations magnitude)	0.277	0.256	0.611	0.595	0.541	0.572
S total	33.053	33.123	31.783	31.970	36.340	35.666
%N-	3.425%	2.968%	5.860%	6.012%	3.577%	3.805%
CC	0.432	0.404	0.190	0.188	0.152	0.150

Table 3.26. N225 Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2008 and 2013.

In Table 3.25 and Table 3.26, no model admits considerably good distance measures. While GARCH-EVT (95% and 90% thresholds) and FHS models admit low N- values and high CC values, they also witness high values of S Total.

Simultaneously, MCS and GARCH (1,1) models suffer high frequencies of violations with low values of S Total and high values of CC.

On the other hand, EVT (POT) (95% and 90% thresholds), var-cov and HS all admit considerably low values of N- and CC, while having high values of S Total.

EWMA (0.94 and 0.95) witness moderate values of violations and S Total with comparably low CC value.

Comments

The behavior of the various implemented models with the N225 index in the crisis period was not like that with the DJIA or SX5E indexes. In fact, the Japanese market was affected during this interval by other incidents like the Fukishima tsunami and the crash of 2013. These events and others led that the market trend of N225 turns out to be different from the DJIA and SX5E. This gave rise to difference in the VaR models' behavior in this market. Accordingly, relative to the graphical presentation of VaR estimates and the distance analysis, the following comments can be made:

99% VaR:

- GARCH -EVT (95% thresholds) under normal, QMLE and Student-t fitted residuals, MCS and GARCH (1,1) models under Student-t distributed residuals (EK and fitting methods) outperform the remainder models.

95% VaR:

- According to the distance measures calculated, no model outperforms the remainder models when estimating the 95% VaR estimation of N225 during the crisis period.

3.6. Post Crisis period 2014 -2019

3.6.1. Dow Jones Industrial Average (DJIA) index

3.6.1.1. Market evolution

Figure 3.19 shows the graphical representation of the DJIA log returns between 2014 and 2019. The period starts with slightly high volatility in 2014. Since the beginning of 2015 the DJIA can be seen to witness higher volatility than in the previous year, reaching a huge plunge on Monday August 24, 2015, which is also known as the "Black Monday". This crash, which followed a previous drop on Friday 21st of August, is not referred to one specific news or event, however, it was attributed to economic worries about the Chinese market. The volatility remained considerably high after the crash before the market witnessed another drop in 2016 which was due to the Brexit referendum, with consequences beyond Europe. The magnitude of market fluctuations started to decrease with moderate volatility in 2017, however, in 2018 the DJIA suffered high volatility and severe drops. Since January 2018 and on, the market witnessed several drops which can be referred to the Cryptocurrency crash in January 2018. The market was in high volatility throughout 2018 reaching another steep drop in November 2018 which due to the intensified trade war between U.S. and China, the slowdown in global economic growth and the concerns regarding the Federal Reserve raising interest rates too quickly. In 2019, the market was still in high volatility and witnessed on the 5th and 14th of August plummets due to a series of economic indicators from Germany and China signaling renewed recession fears in the global economy. Accordingly, the VaR estimates will be affected by these intense market movements in this period which will be examined in the following section.

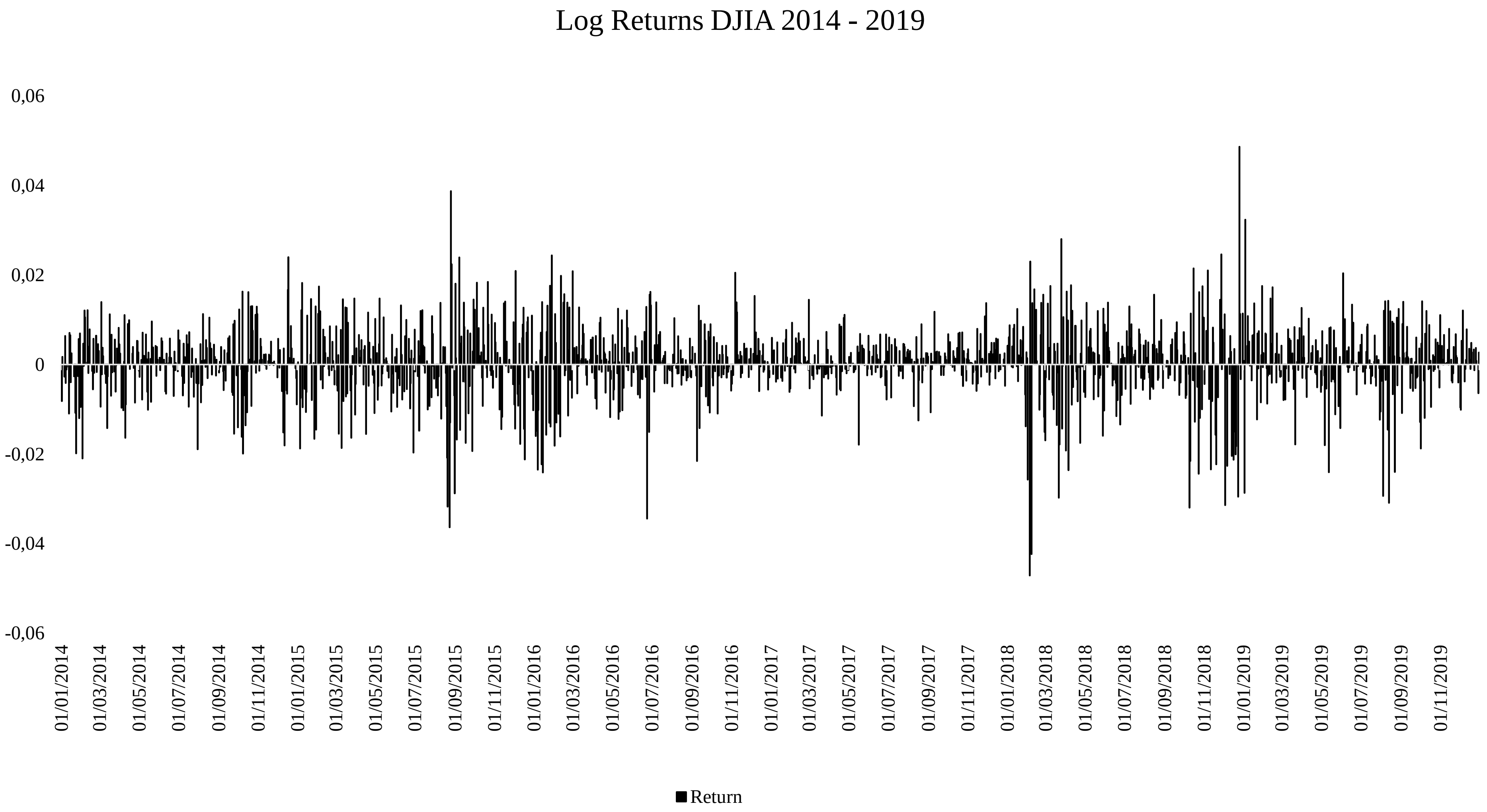

Figure 3.19. DJIA log returns between 2014 and 2019.

3.6.1.2. VaR estimates

Graphical representation

Figure 3.20 and Figure 3.21 show the graphical representation of the 99% VaR and 95% VaR estimates of the DJIA index in the post-crisis period. It can be seen that some models admit nearly smooth curves with minimal fluctuations which do not reflect the market movements, like the HS, var-cov and EVT (POT) models. The other models like GARCH-EVT, MCS, GARCH (1,1) and FHS admit curves with variations that reflect the market movements better than the previously mentioned models. In addition, some models seem to be plotted farthest from the plotted returns throughout the entire interval like the GARCH-EVT (95% threshold) with Student-t residuals. However, this model is sometimes exceeded by the GARCH-EVT (95% threshold) with normally distributed residuals, and this is witnessed more in the case of 95% VaR where it is very obvious in 2015, 2018 and 2019. In the case of 99% VaR, this phenomenon can be noticed during some times of the calm period in 2016 and during the sharp market plunges in 2018 and 2019.

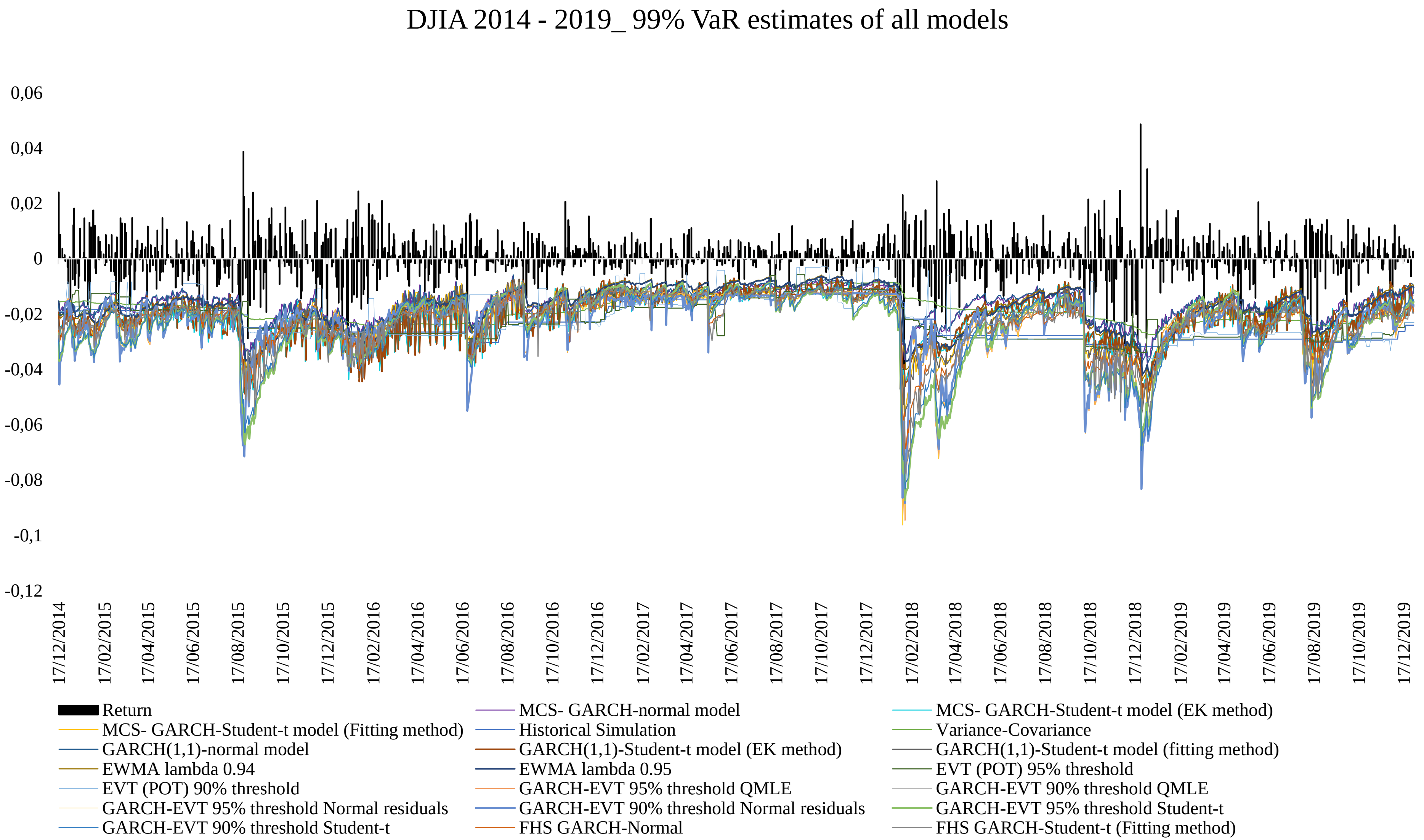

Figure 3.20. DJIA 99% VaR estimates between 2014 and 2019.

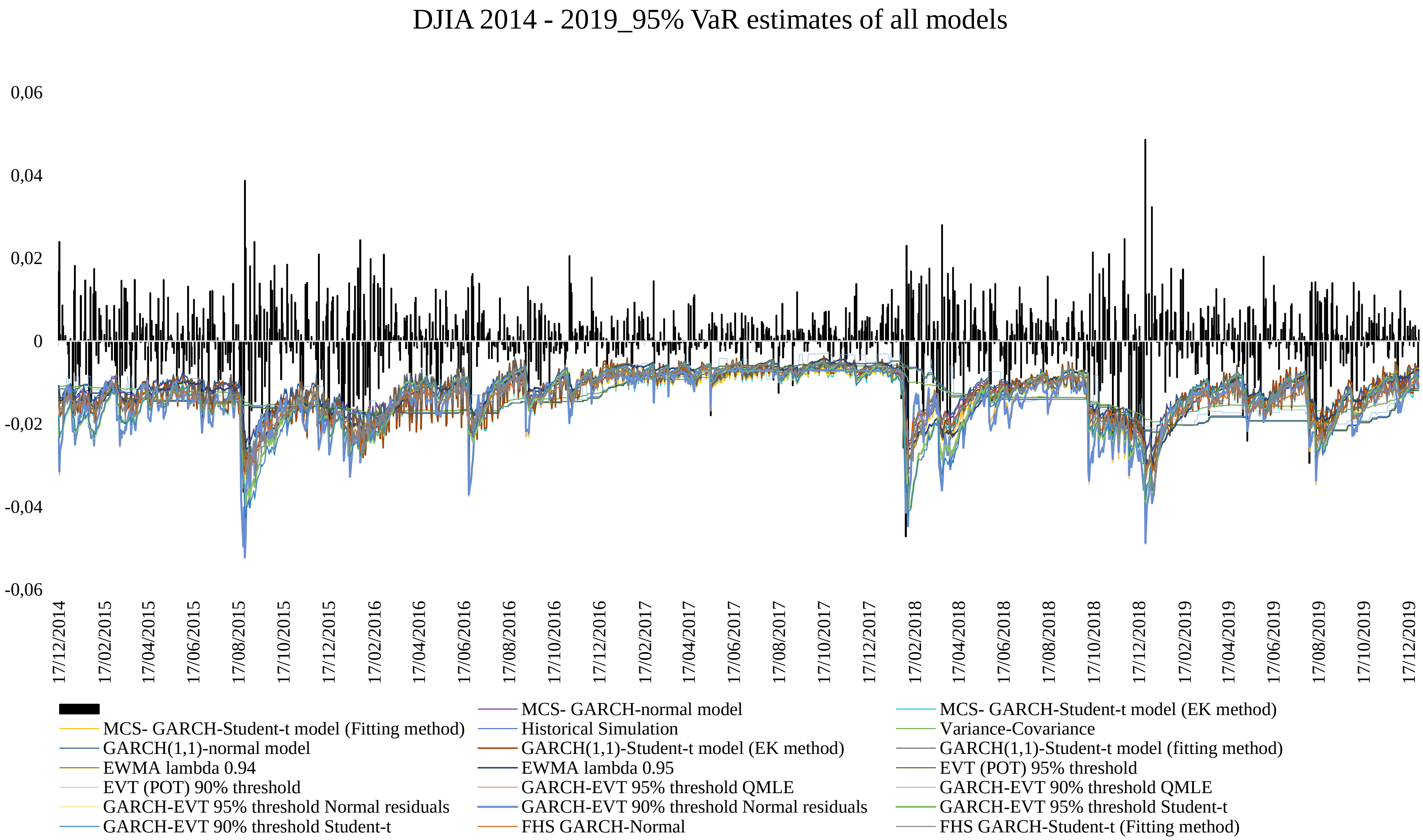

Figure 3.21. DJIA 95% VaR estimates between 2014 and 2019.

Distance Measures

99% VaR

DJIA 2014 - 2019_99% VaR														
Distance Measures	**GARCH-EVT 95% threshold**			**GARCH-EVT 90% threshold**			**EVT (POT)**		**Monte Carlo Simulation**			**GARCH (1,1) model**		
	Normal	**Student-t**	**QMLE**	**Normal**	**Student-t**	**QMLE**	**95% Threshold**	**90% Threshold**	**GARCH-Normal**	**GARCH-t EK method**	**GARCH-t Fitting method**	**Normal**	**t EK**	**t Fitting**
N+	1315	1315	1315	1314	1315	1314	1292	1286	1289	1305	1311	1288	1305	1312
N+ (Returns < 0)	579	579	579	578	579	578	556	550	553	569	575	552	569	576
SN+ (Returns < 0)	20.891	22.167	20.891	20.995	22.111	20.995	20.560	20.247	16.278	19.584	20.128	16.040	19.309	19.868
N+ (Returns ≥ 0)	736	736	736	736	736	736	736	736	736	736	736	736	736	736
SN+ (Returns ≥ 0)	10.137	10.645	10.137	10.233	10.620	10.233	9.785	9.646	6.246	8.742	9.118	6.072	8.583	8.964
S N+ Total	31.028	32.812	31.028	31.228	32.731	31.228	30.345	29.892	22.525	28.325	29.246	22.112	27.891	28.832
N- (violations)	0	0	0	1	0	1	23	29	26	10	4	27	10	3
S N- (violations magnitude)	-	-	-	0.0002	-	0.0002	0.143	0.171	0.115	0.028	0.006	0.121	0.030	0.010
S total	31.028	32.812	31.028	31.228	32.731	31.228	30.488	30.063	22.639	28.354	29.252	22.233	27.921	28.842
%N-	0%	0%	0%	0.076%	0%	0.076%	1.749%	2.205%	1.977%	0.760%	0.304%	2.053%	0.760%	0.228%
CC	0.718	0.577	0.718	0.725	0.585	0.725	0.193	0.196	0.550	0.497	0.568	0.550	0.503	0.571
Euclidean Distance	0.988	1.063	0.988	0.990	1.051	0.990	0.915	0.916	0.717	0.887	0.908	0.709	0.878	0.899

Table 3.27. DJIA Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2014 and 2019.

DJIA 2014 - 2019_ 99% VaR						
Distance Measures	**Filtered Historical Simulation**		**EWMA**		**Variance-covariance**	**Historical Simulation**
	GARCH-Normal	**GARCH-t Fitting**	**Lambda 0.94**	**Lambda 0.95**	**Var-cov**	**HS**
N+	1311	1313	1285	1286	1280	1301
N+ (Returns < 0)	575	577	549	550	544	565
SN+ (Returns < 0)	20.799	21.071	17.032	17.103	17.581	20.865
N+ (Returns ≥ 0)	736	736	736	736	736	736
SN+ (Returns ≥ 0)	9.642	9.870	6.844	6.943	7.572	9.988
S N+ Total	30.441	30.941	23.876	24.046	25.153	30.853
N- (violations)	4	2	30	29	35	14
S N- (violations magnitude)	0.003	0.002	0.222	0.224	0.221	0.110
S total	30.444	30.943	24.097	24.270	25.374	30.963
%N-	0.304%	0.152%	2.281%	2.205%	2.662%	1.065%
CC	0.549	0.560	0.353	0.340	0.187	0.203
Euclidean Distance	0.948	0.975	0.778	0.778	0.771	0.929

Table 3.28. DJIA Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2014 and 2019.

In Table 3.27 and Table 3.28, the distance measures of the estimated 99% VaR of DJIA index in the post-crisis period show that the MCS and GARCH (1,1) models under Student-t (EK and fitting methods) outperform the rest of the models.

On the other hand, it can be seen that the values of measures achieved under Student-t (fitting method) with both models are better than those under the Student-t (EK method), where under the former method, the value of N- was low (and lower than in the EK method), value of CC was high and the S Total measures was moderate while under the Ek method, only the N- value was low while the CC and S Total values were moderate.

It is worth noting also that GARCH-EVT (95% and 90% thresholds) along with FHS witnessed low N- and high CC values with considerably high values of S Total.

HS attains moderate N- value, high S Total, and low CC value.

While EWMA (0.94 and 0.95) witness high N- values accompanied by low S Total measures and low CC values, and var-cov also witnesses measures with similar classification.

EVT (POT) (95% and 90% thresholds) attain the worst classification of measures as they achieve high values of N- and S Total with low value of CC. HS only overcomes EVT (POT) due to a moderate number of violations.

95% VaR

DJIA 2014 – 2019_95% VaR														
Distance Measures	**GARCH-EVT 95% threshold**			**GARCH-EVT 90% threshold**			**EVT (POT)**		**Monte Carlo Simulation**			**GARCH (1,1) model**		
	Normal	**Student-t**	**QMLE**	**Normal**	**Student-t**	**QMLE**	**95% Threshold**	**90% Threshold**	**GARCH-Normal**	**GARCH-t EK method**	**GARCH-t Fitting method**	**Normal**	**t EK**	**t Fitting**
N+	1303	1280	1303	1303	1279	1303	1250	1230	1239	1252	1260	1232	1250	1257
N+ (Returns < 0)	567	544	567	567	543	567	514	494	503	516	524	496	514	521
SN+ (Returns < 0)	14.774	14.400	14.774	14.740	14.501	14.740	14.166	13.499	12.705	13.804	13.952	12.438	13.528	13.677
N+ (Returns ≥ 0)	736	736	736	736	736	736	736	736	736	736	736	736	736	736
SN+ (Returns ≥ 0)	5.264	4.760	5.264	5.239	4.826	5.239	5.110	4.622	3.697	4.489	4.554	3.524	4.316	4.373
S N+ Total	20.037	19.160	20.037	19.979	19.326	19.979	19.276	18.122	16.402	18.293	18.506	15.962	17.843	18.050
N- (violations)	12	35	12	12	36	12	65	85	76	63	55	83	65	58
S N- (violations magnitude)	0.039	0.137	0.039	0.040	0.129	0.040	0.458	0.540	0.385	0.298	0.253	0.397	0.311	0.262
S total	20.076	19.297	20.076	20.018	19.456	20.018	19.734	18.662	16.787	18.591	18.759	16.359	18.154	18.312
%N-	0.913%	2.662%	0.913%	0.913%	2.738%	0.913%	4.943%	6.464%	5.779%	4.791%	4.183%	6.312%	4.943%	4.411%
CC	0.729	0.576	0.729	0.728	0.573	0.728	0.169	0.173	0.550	0.502	0.547	0.549	0.503	0.551

Table 3.29. DJIA Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2014 and 2019.

DJIA 2014 - 2019_95% VaR						
Distance Measures	**Filtered Historical Simulation**		**EWMA**		**Variance-covariance**	**Historical Simulation**
	GARCH-Normal	**GARCH-t Fitting**	**Lambda 0.94**	**Lambda 0.95**	**var-cov**	**HS**
N+	1266	1270	1241	1243	1245	1250
N+ (Returns < 0)	530	534	505	507	509	514
SN+ (Returns < 0)	13.882	13.884	13.139	13.190	13.528	14.034
N+ (Returns ≥ 0)	736	736	736	736	736	736
SN+ (Returns ≥ 0)	4.554	4.485	4.062	4.129	4.623	5.026
S N+ Total	18.436	18.368	17.202	17.319	18.151	19.060
N- (violations)	49	45	74	72	70	65
S N- (violations magnitude)	0.249	0.236	0.461	0.459	0.507	0.472
S total	18.685	18.604	17.663	17.778	18.658	19.532
%N-	3.726%	3.422%	5.627%	5.475%	5.323%	4.943%
CC	0.527	0.538	0.355	0.342	0.192	0.164

Table 3.30. DJIA Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2014 and 2019.

Table 3.29 and 3.30 show that MCS and GARCH (1,1) models under Student-t distributed residuals (fitting method) provide better distance measures of the corresponding 95% VaR estimates, achieving moderate values of N- and S Total and high values of CC.

On the other hand, the same models under normally distributed residuals witness high numbers of violations, low distances S Total and high correlation coefficients. However, under Student-t distributed residuals (EK method) these models also suffer high frequencies of violations but admit moderate values of S Total and CC measures. EWMA (0.94) distance measures admit similar classifications to those of MCS and GARCH (1,1) models under Student-t residuals (EK method).

It can be noticed that FHS models (under normal and Student-t distributed residuals) also show acceptable results, with moderate values of its corresponding distance measures.

Meanwhile, GARCH-EVT (95% and 90% thresholds) models show good results for N- and CC values however, their corresponding S Total values are considerably high.

EWMA (0.95), var-cov and EVT (POT) (90% threshold) witness high values of N- and along with comparably moderate values of S Total, and low values of CC.

EVT (POT) (95% threshold) and HS models admit the worst distance measures being all comparably weak.

Comments

The 99% VaR and 95% VaR distance measures analysis of the DJIA index in the post crisis period shows several indications which can be listed as follows:

99% VaR:

- MCS and GARCH (1,1) models yield the best results under Student-t distributed residuals.
- The remainder models show weak behavior.

95% VaR:

- MCS and GARCH (1,1) under Student-t residuals (fitting method) and FHS models provide the best results among all models.
- The remainder models witness at least one distance measure value revealing weak performance.

3.6.2. Euro Stoxx 50 (SX5E) index

3.6.2.1. Market evolution

The SX5E log returns distribution is represented graphically in Figure 3.22. The market movements witnessed high fluctuations during the entire interval with very short periods of tranquility noticed during 2017. In the years 2014 and 2015, the market witnessed considerably high volatility due to the European debt crisis. In June 2016, the huge plunge of the market occurred when the UK voted for the Brexit. During 2017, the market witnessed some tranquility with volatility of lower magnitude, then in the beginning of 2018 it began to suffer again from considerably high fluctuations due to the Cryptocurrency crash. By the end of 2018 also it witnessed some severe drops which can be referred to the General Election in Italy and the government's decision to conform with the EU budget rules after a period of confrontation. Then in November 2018 when the EU and UK came to an agreement for the UK to leave the EU, the market witnessed again severe drops.

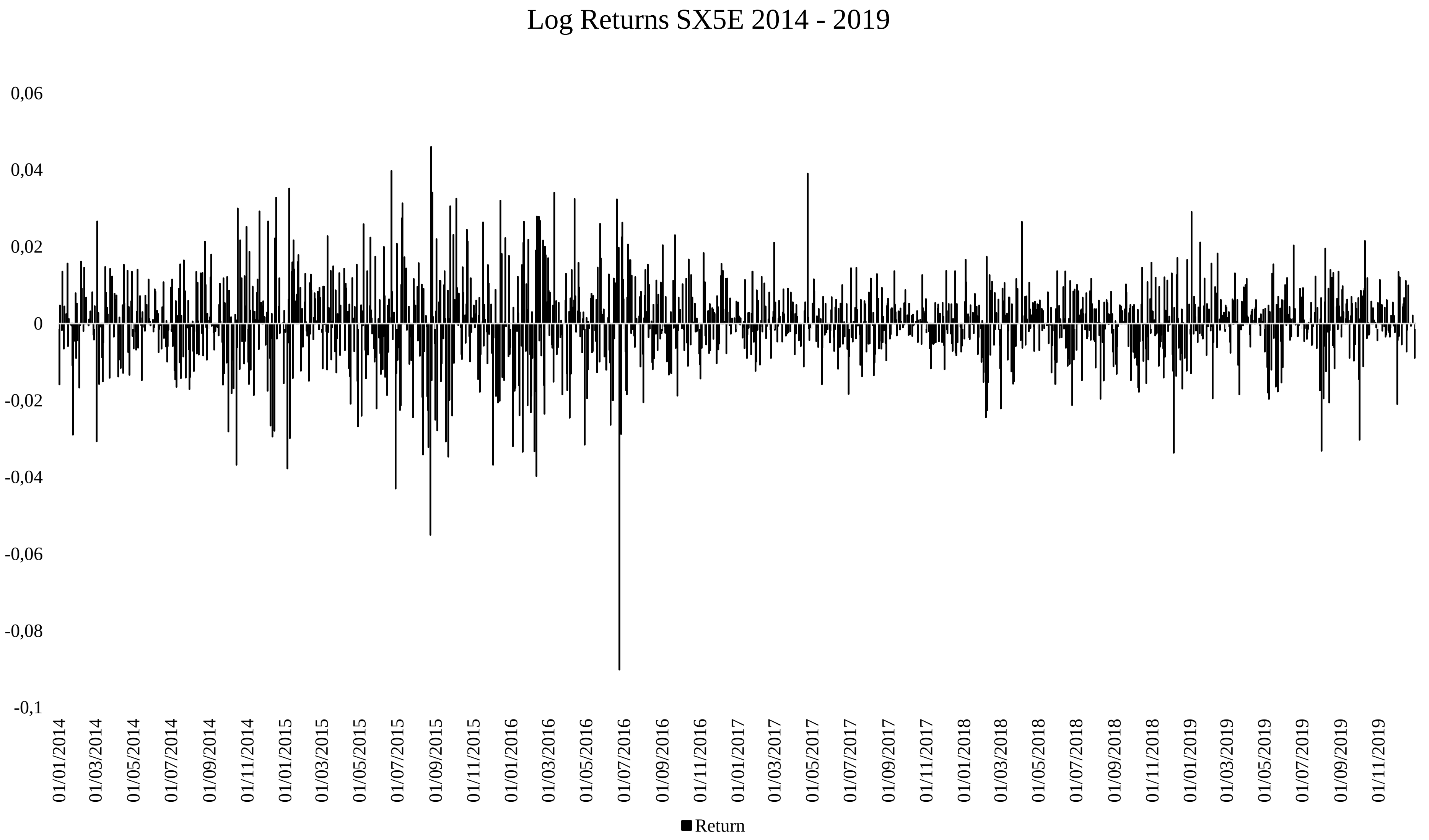

Figure 3.22. SX5E log returns between 2014 and 2019.

3.6.2.2. VaR estimates

Graphical representation

Figure 3.23 and Figure 3.24 represent the 99% VaR and 95% VaR estimates, respectively, of the SX5E index in the post crisis period. As it was previously seen with other indexes and in other intervals, some models admit smoother curves than others like the HS, var-cov and EVT (POT) models. However, most of the time these curves lied slightly farther from the other models and from the plotted returns than in this interval with the SX5E. In Figures 3.23 and 3.24, it can be noticed that the smooth curves of the previously mentioned models lie intertwining or overlapping with the VaR curves of the other models except for in the period between August 2016 and April 2017 where the EVT (POT) 95% threshold model appeared farthest from returns than all the remaining models. On the other hand, the EVT (POT) 95% threshold model appears to be giving the highest estimates of all models especially for the 95% VaR.

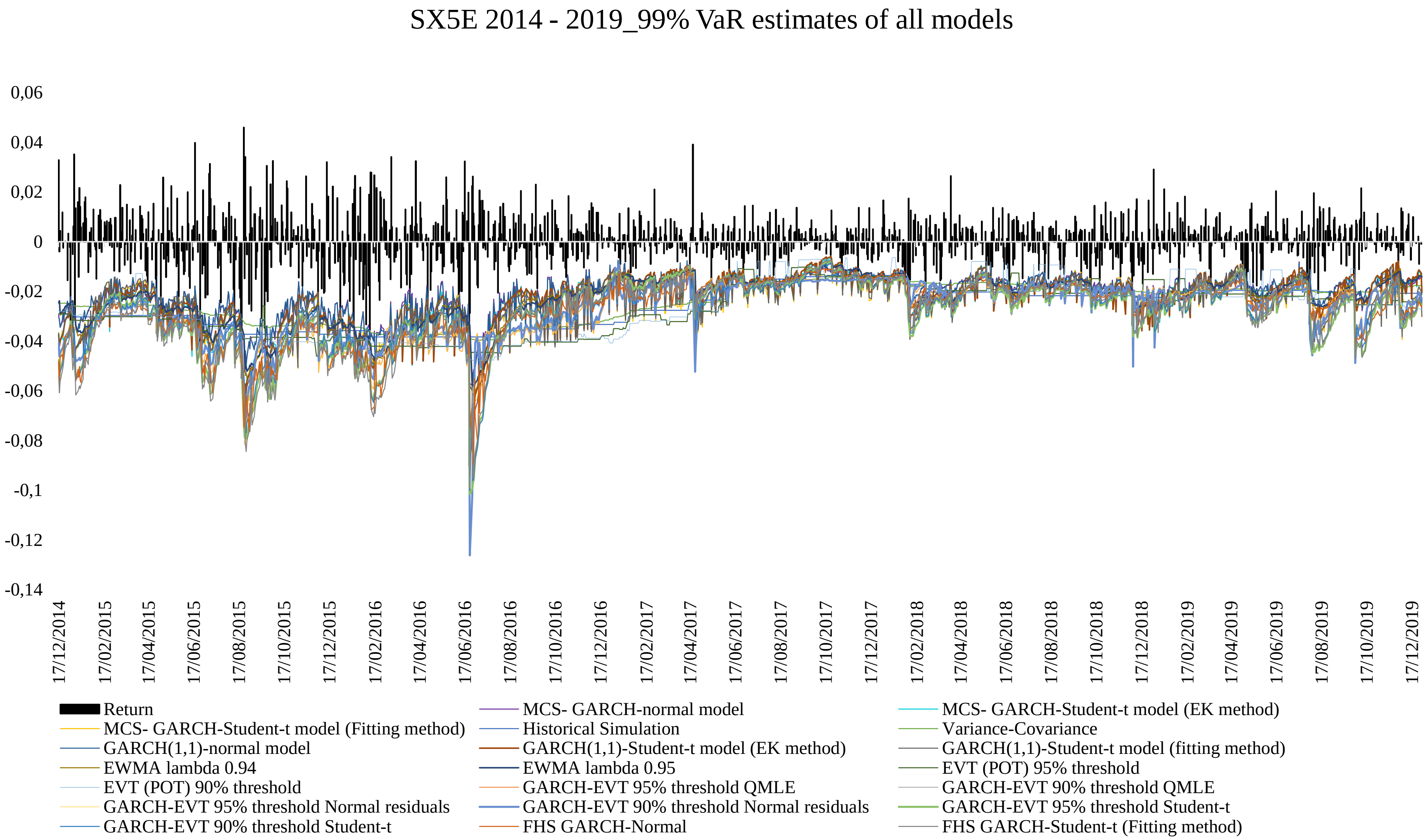

Figure 3.23. SX5E 99% VaR estimates between 2014 and 2019.

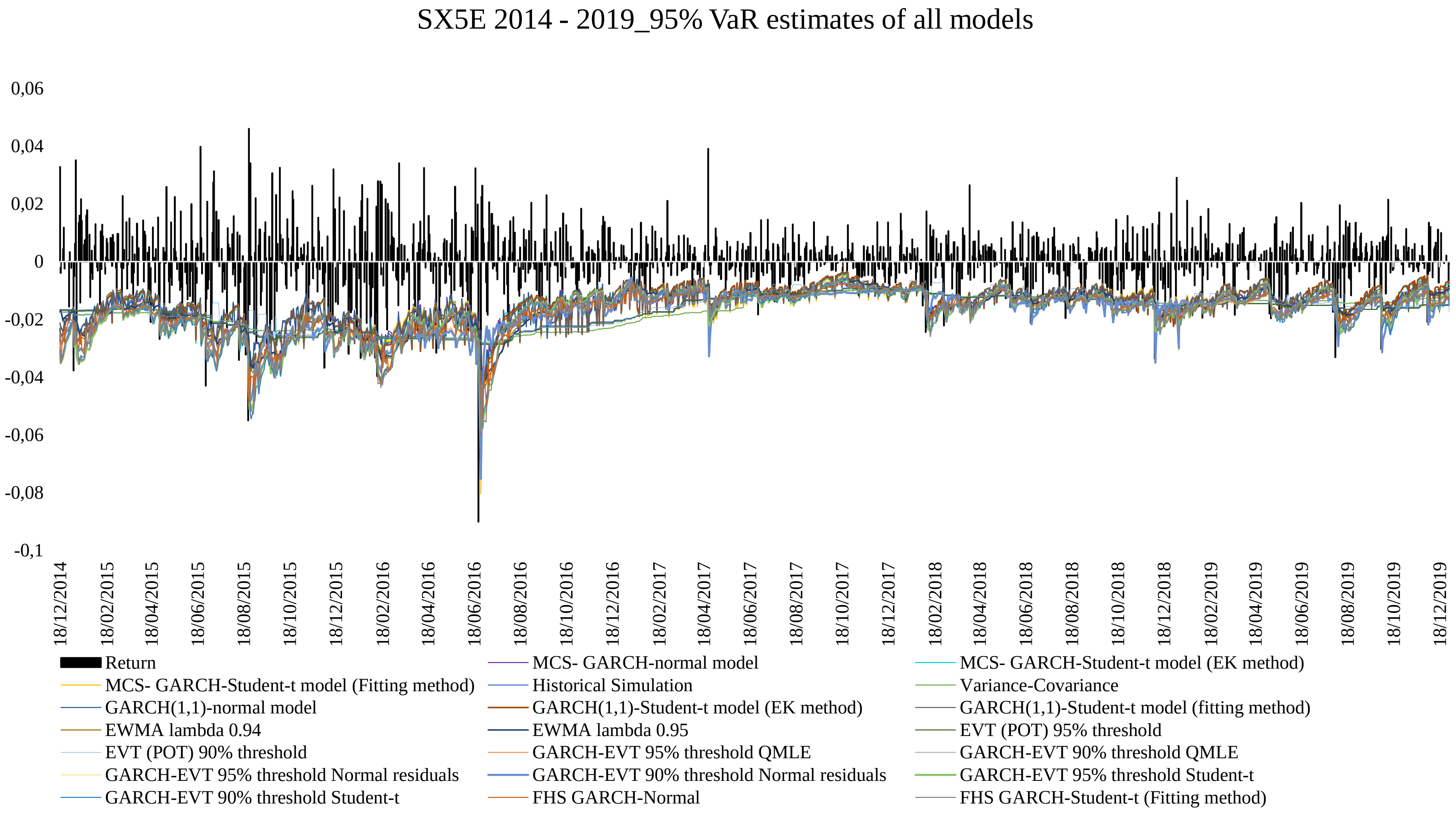

Figure 3.24. SX5E 95% VaR estimates between 2014 and 2019.

Distance Measures

99% VaR

SX5E 2014 - 2019_99% VaR														
Distance Measures	**GARCH-EVT 95% threshold**			**GARCH-EVT 90% threshold**			**EVT (POT)**		**Monte Carlo Simulation**			**GARCH (1,1) model**		
	Normal	**Student-t**	**QMLE**	**Normal**	**Student-t**	**QMLE**	**95% Threshold**	**90% Threshold**	**GARCH-Normal**	**GARCH-t EK method**	**GARCH-t Fitting method**	**Normal**	**t EK**	**t Fitting**
N+	1313	1314	1313	1312	1313	1312	1297	1290	1291	1304	1307	1291	1303	1307
N+ (Returns < 0)	608	609	608	607	608	607	592	585	586	599	602	586	598	602
SN+ (Returns < 0)	24.326	25.185	24.326	24.066	24.765	24.066	24.297	24.063	19.971	21.528	22.304	19.950	21.512	22.284
N+ (Returns ≥ 0)	705	705	705	705	705	705	705	705	705	705	705	705	705	705
SN+ (Returns ≥ 0)	12.122	12.453	12.122	11.814	12.138	11.814	12.138	12.008	8.256	9.566	10.260	8.233	9.520	10.221
S N+ Total	36.448	37.639	36.448	35.880	36.903	35.880	36.435	36.070	28.228	31.094	32.563	28.183	31.032	32.505
N- (violations)	2	1	2	3	2	3	18	25	24	11	8	24	12	8
S N- (violations magnitude)	0.008	0.004	0.008	0.008	0.002	0.008	0.152	0.173	0.139	0.043	0.024	0.136	0.045	0.023
S total	36.456	37.643	36.456	35.888	36.905	35.888	36.587	36.243	28.367	31.138	32.587	28.319	31.077	32.528
%N-	0.152%	0.076%	0.152%	0.228%	0.152%	0.228%	1.369%	1.901%	1.825%	0.837%	0.608%	1.825%	0.913%	0.608%
CC	0.584	0.533	0.584	0.597	0.513	0.597	0.231	0.231	0.454	0.482	0.471	0.460	0.483	0.477

Table 3.31. SX5E Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2014 and 2019.

SX5E 2014 - 2019_ 99% VaR						
Distance Measures	**Filtered Historical Simulation**		**EWMA**		**Variance-covariance**	**Historical Simulation**
	GARCH-Normal	**GARCH-t Fitting**	**Lambda 0.94**	**Lambda 0.95**	**Var-cov**	**HS**
N+	1310	1312	1285	1285	1291	1298
N+ (Returns < 0)	605	607	580	580	586	593
SN+ (Returns < 0)	24.651	25.313	21.342	21.415	22.180	23.335
N+ (Returns ≥ 0)	705	705	705	705	705	705
SN+ (Returns ≥ 0)	12.114	12.599	9.337	9.418	10.358	11.297
S N+ Total	36.765	37.913	30.679	30.833	32.538	34.632
N- (violations)	5	3	30	30	24	16
S N- (violations magnitude)	0.028	0.003	0.222	0.215	0.188	0.137
S total	36.793	37.916	30.901	31.048	32.726	34.769
%N-	0.380%	0.228%	2.281%	2.281%	1.825%	1.218%
CC	0.472	0.509	0.323	0.319	0.248	0.266

Table 3.32. SX5E Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2014 and 2019.

Table 3.31 and Table 3.32 present the distance measures of the 99% VaR of the SX5E estimated under all models in the post-crisis period.

GARCH-EVT (95% and 90% thresholds) and FHS under Student-t residuals along with MCS and GARCH (1,1) models with Student-t residuals (fitting method) provide acceptable results with low values of N-, high values of CC and considerably also high values of S Total.

In fact, MCS and GARCH (1,1) models under Student-t distributed residuals (EK and fitting method) seem to provide better results during this period for SX5E, as they admit distance measures of good quality compared to other models.

EVT (POT) (95% threshold) and HS models witness similar classifications of distance measures with moderate frequencies of violations and high S Total and low CC values.

EVT (POT) (90% threshold) shows the weakest performance among all models.

Var-cov and EWMA (0.94 and 0.95) models suffer high frequencies of violations and low CC values with S Total values varying between moderate (var-cov model) and low (EWMA).

95% VaR

SX5E 2014 -2019_ 95% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1) model		
	Normal	Student-t	QMLE	Normal	Student-t	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1279	1270	1279	1281	1271	1281	1249	1235	1231	1216	1222	1228	1216	1220
N+ (Returns < 0)	574	565	574	576	566	576	544	530	526	511	517	523	511	515
SN+ (Returns < 0)	17.511	17.612	17.511	17.549	17.648	17.549	17.157	16.649	15.611	15.330	15.560	15.556	15.269	15.498
N+ (Returns ≥ 0)	705	705	705	705	705	705	705	705	705	705	705	705	705	705
SN+ (Returns ≥ 0)	6.231	6.147	6.231	6.285	6.256	6.285	6.228	5.888	4.818	4.582	4.777	4.785	4.549	4.732
S N+ Total	23.742	23.760	23.742	23.833	23.905	23.833	23.385	22.537	20.428	19.912	20.337	20.342	19.818	20.230
N- (violations)	36	45	36	34	44	34	66	80	84	99	93	87	99	95
S N- (violations magnitude)	0.149	0.184	0.149	0.148	0.195	0.148	0.461	0.538	0.483	0.467	0.443	0.478	0.471	0.438
S total	23.890	23.944	23.890	23.981	24.099	23.981	23.846	23.075	20.912	20.379	20.780	20.820	20.289	20.668
%N-	2.738%	3.422%	2.738%	2.586%	3.346%	2.586%	5.019%	6.084%	6.388%	7.529%	7.072%	6.616%	7.529%	7.224%
CC	0.619	0.535	0.619	0.615	0.520	0.615	0.285	0.269	0.456	0.483	0.479	0.461	0.484	0.486

Table 3.33. SX5E Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2014 and 2019.

SX5E 2014 - 2019_95% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	Var-cov	HS
N+	1266	1270	1240	1245	1246	1244
N+ (Returns < 0)	561	565	535	540	541	539
SN+ (Returns < 0)	17.287	17.406	16.541	16.593	17.133	17.019
N+ (Returns ≥ 0)	705	705	705	705	705	705
SN+ (Returns ≥ 0)	5.996	6.044	5.546	5.604	6.270	6.105
S N+ Total	23.283	23.450	22.088	22.197	23.403	23.123
N- (violations)	49	45	75	70	69	70
S N- (violations magnitude)	0.257	0.205	0.520	0.515	0.497	0.474
S total	23.540	23.655	22.608	22.712	23.900	23.597
%N-	3.726%	3.422%	5.703%	5.323%	5.247%	5.327%
CC	0.496	0.529	0.325	0.321	0.252	0.285

Table 3.34. SX5E Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2014 and 2019.

In Table 3.33 and Table 3.34, the distance measures of the 95% VaR estimated for SX5E in the post-crisis period are presented.

It can be noticed that no models achieve acceptable results according to distance. If a model achieves low values of violations, then it simultaneously suffers high values of S Total, and vice-versa.

Although GARCH-EVT (95% and 90% thresholds) and FHS under Student-t residuals models admit considerably low values of N- and high values of CC, however they admit high values of S Total.

EVT (POT) (90% threshold) model is the worst performing model with all three measures, N-, CC and S Total, acquiring weak classifications.

EVT (POT) (95% threshold) along with var-cov and HS models witness moderate frequencies of violations with high values of S Total and low CC values.

MCS and GARCH (1,1) models under normal and Student-t fitted residuals show similar results with high values of N-, low values of S Total and moderate CC values.

EWMA (0.94 and 0.95) attain moderate values of N- and S Total with low CC values.

FHS model under normally distributed residuals shows a low number of violations with high S Total value and moderate CC value.

Comments

The preliminary distance measures analysis provides some insights regarding the performance of the various models in the estimation of VaR for the SX5E index in the post-crisis period. The following comments can be made regarding this matter:

99% VaR:

- MCS and GARCH (1,1) models under Student-t residuals (EK and fitting methods) outperform the remainder models that suffer weak performance when estimating the 99% VaR of SX5E in the post-crisis interval.

95% VaR:

- No VaR model shows acceptable distance measures according to the classification criteria of this analysis.

3.6.3. Nikkei 225 (N225) index

3.4.3.1. Market evolution

Figure 3.25 presents the log returns of the N225 index in the post-crisis period. The market can be seen to have high volatility during the entire period but in some years the magnitude and frequency were higher. In 2014, the market fluctuations were not as high as in 2015 and 2016 which witnessed very high clustered volatility which can be referred to the market selloff as there was a decline in the value of stock prices globally between June 2015 and June 2016. This fall was referred to several reasons of which are the slowing growth of the GDP in China, the Greek debt default in June 2015, a fall in petroleum prices in addition to the referendum of UK membership in the European Union. In 2017, the market was comparably calm with volatility of lower magnitude than in 2015 and 2016. However, in 2018 the N225 index witnessed more turbulence, on the 6^{th} of February 2018 the market losses followed a massive U.S. sell-off on but some also refer this fall to other domestic factors. Then, in December 2018, the market suffered another drop which was also came after developments in the U.S. market.

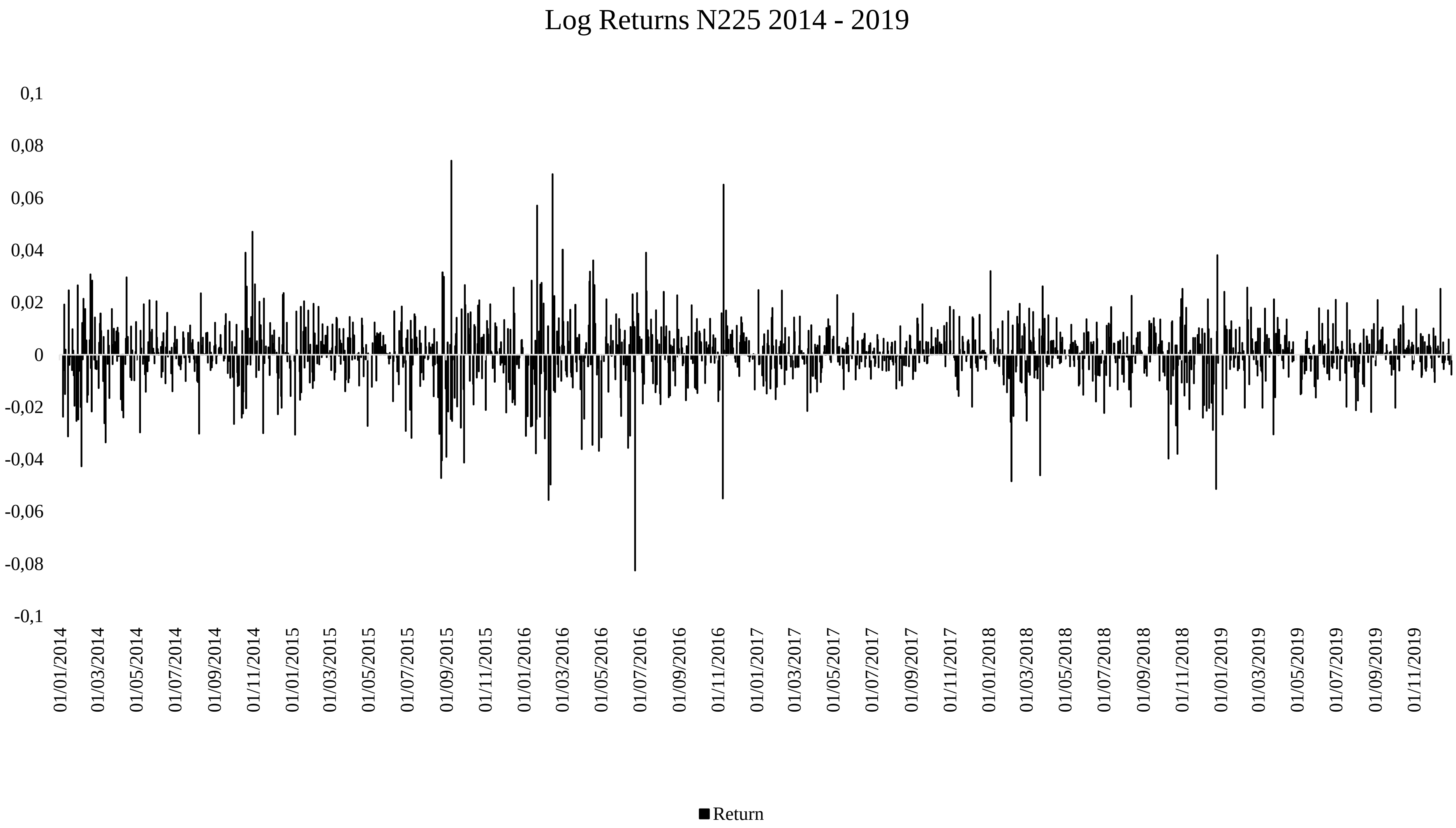

Figure 3.25. N225 log returns between 2014 and 2019.

3.4.3.2. VaR estimates

Graphical representation

The graphical representation of the 99% VaR and 95% VaR of the N225 index obtained under the various models are presented in Figure 3.26 and Figure 3.27. Some models admit smooth curves with constant trends throughout long periods of time during the interval, and this was seen in other intervals and also with other indexes, these models are the EVT (POT), HS, and var-cov. However, these models here look closer to the other models and in many periods during the interval their VaR curves look intertwined with the rest of the models especially when the market is in a considerably tranquil state. Moreover, it can be noticed that the other models run with a similar trend and frequency of fluctuations. GARCH-EVT (95% threshold) seems to be the model plotted the farthest from the plotted returns. Due to the overlapping of VaR curves in the figures, the distance measures analysis will show more about the behavior of the different models when estimating VaR for N225 in the post-crisis period.

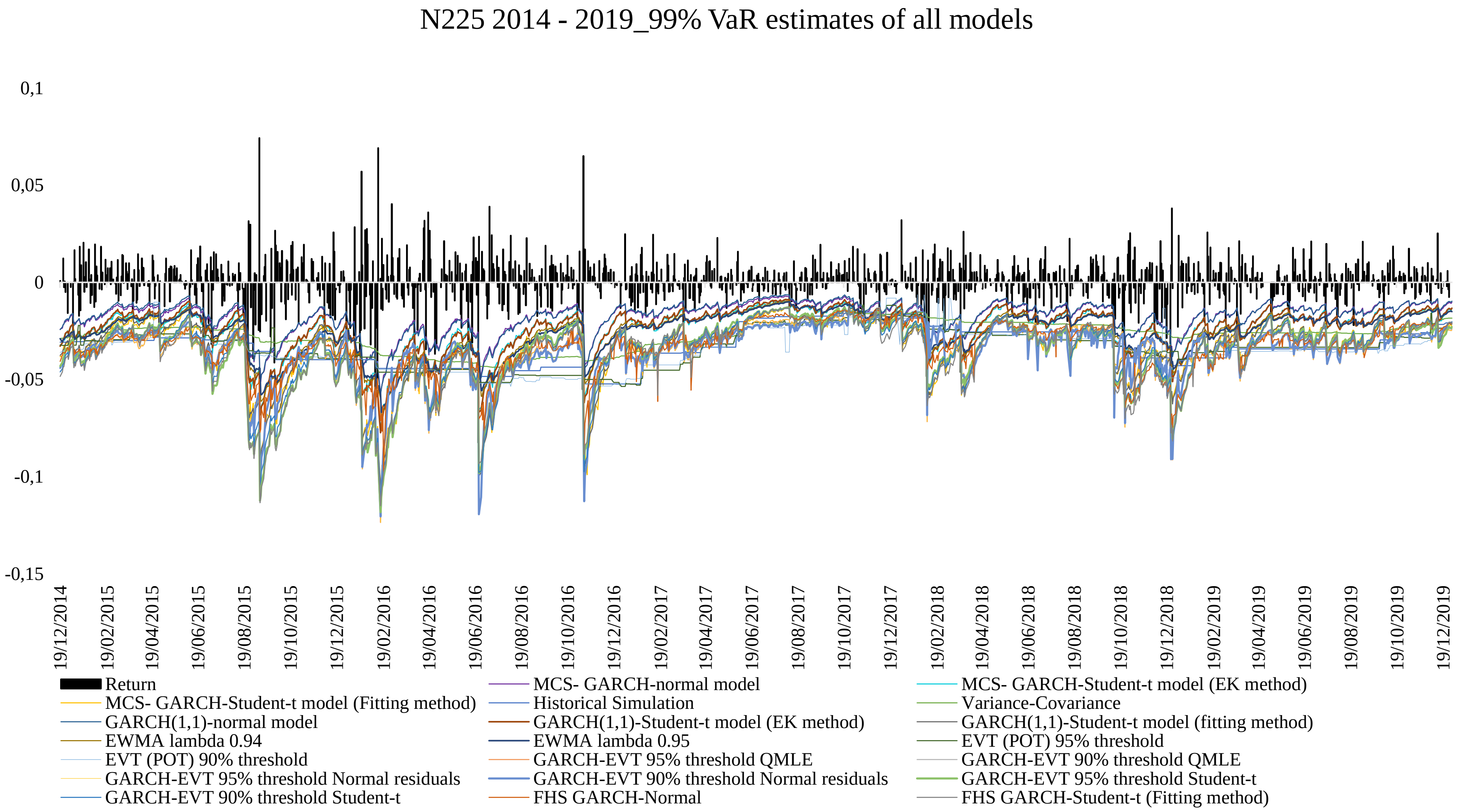

Figure 3.26. N225 99% VaR estimates between 2014 and 2019.

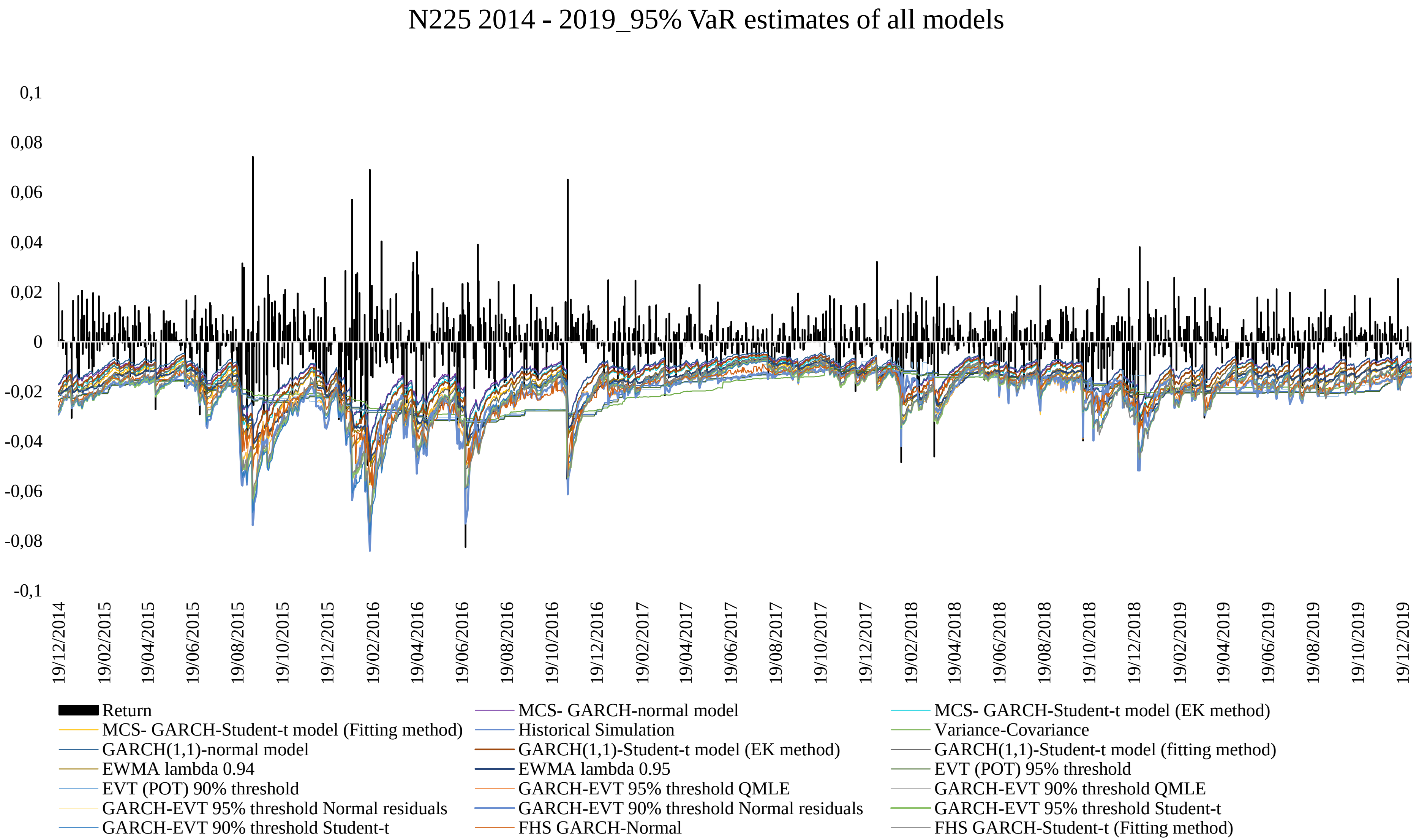

Figure 3.27. N225 95% VaR estimates between 2014 and 2019.

Distance Measures

99% VaR

N225 2014 - 2019_99% VaR														
Distance Measures	GARCH-EVT 95% threshold			GARCH-EVT 90% threshold			EVT (POT)		Monte Carlo Simulation			GARCH (1,1) model		
	Normal	Student-t	QMLE	Normal	Student-t	QMLE	95% Threshold	90% Threshold	GARCH-Normal	GARCH-t EK method	GARCH-t Fitting method	Normal	t EK	t Fitting
N+	1311	1313	1311	1311	1313	1311	1290	1290	1247	1289	1312	1242	1291	1312
N+ (Returns < 0)	564	566	564	564	566	564	543	543	500	542	565	495	544	565
SN+ (Returns < 0)	31.508	31.952	31.508	31.408	31.979	31.408	30.342	31.004	18.711	23.322	30.601	18.524	23.145	30.410
N+ (Returns ≥ 0)	747	747	747	747	747	747	747	747	747	747	747	747	747	747
SN+ (Returns ≥ 0)	15.052	15.113	15.052	14.955	15.086	14.955	14.139	14.511	5.495	8.686	13.996	5.385	8.561	13.911
S N+ Total	46.560	47.065	46.560	46.363	47.065	46.363	44.481	45.515	24.206	32.008	44.597	23.909	31.706	44.321
N- (violations)	2	0	2	2	0	2	23	23	66	24	1	71	22	1
S N- (violations magnitude)	0.002	-	0.002	0.002	-	0.002	0.168	0.160	0.382	0.076	0.00017	0.390	0.080	0.000012
S total	46.562	47.065	46.562	46.365	47.065	46.365	44.649	45.676	24.588	32.084	44.597	24.299	31.786	44.321
%N-	0.152%	0%	0.152%	0.152%	0%	0.152%	1.752%	1.752%	5.027%	1.828%	0.076%	5.407%	1.676%	0.076%
CC	0.637	0.531	0.637	0.639	0.534	0.640	0.179	0.184	0.532	0.533	0.516	0.530	0.531	0.514

Table 3.35. N225 Distance measures analysis of 99% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2014 and 2019.

N225 2014 - 2019_ 99% VaR						
Distance Measures	Filtered Historical Simulation		EWMA		Variance-covariance	Historical Simulation
	GARCH-Normal	GARCH-t Fitting	Lambda 0.94	Lambda 0.95	Var-cov	HS
N+	1309	1313	1279	1281	1285	1293
N+ (Returns < 0)	562	566	532	534	538	546
SN+ (Returns < 0)	30.358	32.728	24.370	24.482	25.703	29.864
N+ (Returns ≥ 0)	747	747	747	747	747	747
SN+ (Returns ≥ 0)	14.003	15.721	9.458	9.577	10.748	13.716
S N+ Total	44.360	48.449	33.827	34.059	36.451	43.580
N- (violations)	4	0	34	32	28	20
S N- (violations magnitude)	0.025	0.000	0.323	0.320	0.321	0.168
S total	44.385	48.449	34.150	34.379	36.772	43.747
%N-	0.305%	0.000%	2.589%	2.437%	2.133%	1.523%
CC	0.503	0.535	0.293	0.288	0.182	0.209

Table 3.36. N225 Distance measures analysis of 99% VaR estimated obtained by FHS, EWMA, VAR-COV and HS models between 2014 and 2019.

Table 3.35 and Table 3.36 present the distance measures of the 99%VaR estimates of the N225 index in the post-crisis period.

MCS and GARCH (1,1) models under Student-t residuals (EK method) outperform the other models with low or moderate N-, high CC and low S Total values. In this case, these models acquire the typical characteristics of a good model with respect to the considered distance measures.

GARCH-EVT (95% and 90% threshold) as well as FHS model, admit low values of N-, high CC values and mainly high S Total values.

EVT (POT) (95% and 90% threshold) and HS model witness similar results classified with low frequencies of violations, high S Total values, and low CC values.

EWMA (0.94 and 0.95) as well as var-cov model show moderate values of N- and S Total with low values of CC.

95% VaR

N225 2014 -2019_ 95% VaR														
Distance Measures	**GARCH-EVT 95% threshold**			**GARCH-EVT 90% threshold**			**EVT (POT)**		**Monte Carlo Simulation**			**GARCH (1,1) model**		
	Normal	**Student-t**	**QMLE**	**Normal**	**Student-t**	**QMLE**	**95% Threshold**	**90% Threshold**	**GARCH-Normal**	**GARCH-t EK method**	**GARCH-t Fitting method**	**Normal**	**t EK**	**t Fitting**
N+	1288	1277	1288	1292	1278	1292	1253	1246	1179	1222	1245	1172	1211	1241
N+ (Returns < 0)	541	530	541	545	531	545	506	499	432	475	498	425	464	494
SN+ (Returns < 0)	20.976	21.339	20.976	21.389	21.263	21.389	20.076	19.725	14.794	16.494	18.636	14.593	16.285	18.425
N+ (Returns >=0)	747	747	747	747	747	747	747	747	747	747	747	747	747	747
SN+ (Returns>=0)	7.134	7.220	7.134	7.447	7.247	7.447	6.706	6.535	3.119	4.103	5.440	3.005	3.981	5.315
S N+ Total	28.110	28.559	28.110	28.836	28.510	28.836	26.782	26.260	17.913	20.596	24.076	17.598	20.266	23.740
N- (violations)	25	36	25	21	35	21	60	67	134	91	68	141	102	72
S N- (violations magnitude)	0.113	0.168	0.113	0.094	0.174	0.094	0.596	0.646	0.955	0.656	0.402	0.977	0.677	0.409
S total	28.224	28.727	28.224	28.930	28.684	28.930	27.378	26.906	18.868	21.252	24.478	18.575	20.943	24.149
%N-	1.904%	2.742%	1.904%	1.599%	2.666%	1.599%	4.570%	5.103%	10.206%	6.931%	5.179%	10.739%	7.768%	5.484%
CC	0.625	0.527	0.625	0.624	0.524	0.624	0.230	0.223	0.531	0.535	0.526	0.529	0.529	0.524

Table 3.37. N225 Distance measures analysis of 95% VaR estimated obtained by GARCH-EVT, EVT (POT), MCS and GARCH (1,1) models between 2014 and 2019.

N225 2014 – 2019_95% VaR						
Distance Measures	**Filtered Historical Simulation**		**EWMA**		**Variance-covariance**	**Historical Simulation**
	GARCH-Normal	**GARCH-t Fitting**	**Lambda 0.94**	**Lambda 0.95**	**Var-cov**	**HS**
N+	1267	1278	1233	1233	1242	1250
N+ (Returns < 0)	520	531	486	486	495	503
SN+ (Returns < 0)	20.235	21.033	18.726	18.805	19.669	19.827
N+ (Returns >=0)	747	747	747	747	747	747
SN+ (Returns>=0)	6.572	7.012	5.664	5.737	6.544	6.547
S N+ Total	26.807	28.045	24.390	24.542	26.213	26.374
N- (violations)	46	35	80	80	71	63
S N- (violations magnitude)	0.324	0.181	0.709	0.695	0.676	0.616
S total	27.132	28.226	25.098	25.237	26.889	26.990
%N-	3.503%	2.666%	6.093%	6.093%	5.407%	4.798%
CC	0.485	0.525	0.294	0.290	0.186	0.229

Table 3.38.N225 Distance measures analysis of 95% VaR estimated obtained by FHS, EWMA, var-cov and HS models between 2014 and 2019.

In Table 3.37 and Table 3.38, the distance measures of the 95% VaR estimates of the N225 index in the post-crisis period are presented.

GARCH-EVT (95% and 90% thresholds) and FHS models witness low frequencies of violations with high values of distance S Total and high CC values.

The MCS under Student-t residuals (EK and fitting methods) along with GARCH (1,1) model under Student-t residuals (fitting method) outperform the remaining models.

MCS and GARCH (1,1) models under Student-t residuals (fitting method) witness moderate frequencies of violations and S Total with high CC values.

MCS under normal residuals and GARCH (1,1) under normal and Student-t (EK method) attain high values of N- as well as high values of CC with low values of S Total.

EVT (POT) (95% threshold) model witnesses a low value of violations with high S Total and low CC value.

EVT (POT) (90% threshold), var-cov and HS acquire moderate values of violations while having high values of S Total and low CC values.

Simultaneously, EWMA (0.94 and 0.95) admit moderate N- and S Total values while suffering low values of CC.

Comments

The preliminary analysis of the distance measures corresponding to the VaR estimates of N225 index in the post-crisis period showed important results regarding the performance of the various models. Some comments to concrete the results of this analysis are:

99% VaR:

- MCS and GARCH (1,1) models under Student-t residuals (EK method) provide the best results regarding the values of N-, S Total and CC among all models.
- The remainder models had at least one measure showing signs of weak performance although some have good measures as to number of violations like GARCH-EVT and FHS which record the lowest values of this measure.

95% VaR:

- MCS under Student-t residuals (EK and fitting methods) and GARCH (1,1) under Student-t residuals (fitting method) outperform the remainder models.
- Some models show good results for some measures like GARCH-EVT and FHS models which admit considerably lower frequencies of violations than the rest of the models however, they also suffer high values of S Total which points out some flaws of these models when estimating the 95% VaR of N225 in the post-crisis period.

3.7. Conclusions

The distance measures analysis reveals important aspects regarding the performance of the various implemented models. Although a low number of violations is important to check the performance of a certain model, however, the distance between the model and the actual returns also counts since a model might have fewer violations than other models not due to its goodness of fit but to the contrary, sometimes it is due to the model overestimating VaR.

In practice and literature, VaR models are usually evaluated according to the number of violations that their corresponding VaR estimates acquire, the duration between these violations, or interdependence between violations. However, the goodness of fit of a model should be examined from various points, starting from the simplest ones like the distance measures and up to more sophisticated backtesting measures which will be examined in Chapter 4. To wrap up this chapter properly, the best performing models according to the preliminary analysis based on distance measures are classified by interval and index in Table 3.39.

		DJIA	SX5E	N225
99% VaR	**2002 - 2007**	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS (normal) • GARCH (1,1) (normal and Student-t (fitting method)) • FHS	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS Student-t (EK and fitting method)) • GARCH (1,1) (Student-t (fitting method)) • FHS	• MCS Student-t (EK and fitting method)) • GARCH (1,1) (Student-t (EK and fitting methods)) •
	2008 - 2013	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS (Student-t EK and fitting methods) • GARCH (1,1) (Student-t EK and fitting methods) • FHS	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS (Student-t (EK and fitting methods)) • GARCH (1,1) (Student-t (EK and fitting methods)) • FHS	• GARCH-EVT (95% threshold) • MCS (Student-t (EK and fitting methods)) • GARCH (1,1) (Student-t (EK and fitting methods))
	2014 - 2019	• MCS (Student-t (EK and fitting method)) • GARCH (1,1) (Student-t (EK and fitting methods))	• MCS (Student-t (EK and fitting method)) • GARCH (1,1) (Student-t (EK and fitting methods))	• MCS (Student-t (EK method)) • GARCH (1,1) (Student-t (EK method))
95% VaR	**2002 - 2007**	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • GARCH (1,1) (Student-t (fitting method)) • FHS	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • FHS	• None
	2008 - 2013	• GARCH-EVT (95% threshold) • MCS (Student-t (fitting method)) • FHS	• GARCH-EVT (95% threshold) Student-t • GARCH-EVT (90% threshold) Student-t • FHS	• None
	2014 - 2019	• MCS (Student-t (fitting method)) • GARCH (1,1) (Student-t (fitting method)) • FHS	• None	• MCS (Student-t (EK and fitting methods)) • GARCH (1,1) (Student-t (fitting method))

Table 3.39. Summary of the chapter; best performing models for estimating 99% VaR and 95% VaR according to distance measures classified by index and interval.

To summarize the findings of the analysis carried out in this chapter, the relative observations are consecutively classified by index and then by interval. This highlights the VaR models that showed the best performance according to the analytical criteria followed in this chapter, at both index and interval levels. It will also simplify the comparison later with the results of the backtesting analysis of Chapter 4.

By index

DJIA

The GARCH-EVT and FHS models provided the best results for the estimation of the 99% VaR and 95% VaR of DJIA index. MCS and GARCH (1,1) models showed good performance with DJIA throughout the three intervals regardless of the type of residuals used which included normally distributed residuals only in the pre-crisis period with the 99% VaR estimates. It is worth noting that the results are very similar for the 99% VaR and 95% VaR with minimal differences.

SX5E

Similar observations are noted for the SX5E index to those of the DJIA index. GARCH-EVT 95% threshold), GARCH-EVT (90% threshold) and FHS provided good results before and during the crisis with the 99% VaR estimates as well as the 95% VaR estimates. However, MCS and GARCH (1,1) models under Student-t distributed residuals achieved good performance with SX5E for estimating the 99% VaR throughout the whole period of the study, before, during and after the crisis. For the 95% VaR estimates in the post-crisis interval, no models provided acceptable distance measures and accordingly it can be said that all models provided poor VaR estimates.

N225

Since the nature of the Japanese market and the distribution of the N225 log returns plotted before in this chapter are quite different from those of the DJIA and SX5E indexes, it is expected that the VaR models' behavior with N225 to be different from the other indexes. The MCS and GARCH (1,1) models provide the best 99% VaR estimates of the Japanese market index before, during and after the crisis noting that during the crisis GARCH-EVT (95% threshold) model also showed good performance. However, on the level of the 95% VaR estimates, only MCS and GARCH (1,1) models revealed acceptable results in the post-crisis period. In the first two intervals all models provided poor results.

By interval

Pre-crisis period

In general, GARCH-EVT, GARCH (1,1) and FHS appear to outperform the remainder models when estimating VaR in the pre-crisis interval. However, for instance, none of the studied models provided acceptable 95% VaR estimates for the N225 index in this interval.

Crisis period

According to the GARCH-EVT, MCS, and FHS also dominate the remainder models in estimating VaR during the crisis interval. Exceptions always apply with the N225 index where no models provide good 95% VaR estimates.

Post-Crisis period

MCS and GARCH (1,1) models and, in a lower level, FHS are the only models that seem to give acceptable VaR estimates in the post-crisis period. No models provide acceptable 95% VaR estimates for the SX5E index in the post crisis interval.

Further analysis and more concrete findings about the performance of VaR models will be investigated using backtesting measures in Chapter 4.

References

Byström, H. N. E. (2004). Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, *13*(2), 133–152. https://doi.org/10.1016/j.irfa.2004.02.003

Echaust, K., & Just, M. (2020). Value at risk estimation using the garch-evt approach with optimal tail selection. *Mathematics*, *8*(1). https://doi.org/10.3390/math8010114

Huang, C.-K., North, D., & Zewotir, T. (2017). Exchangeability, extreme returns and Value-at-Risk forecasts. *Physica A: Statistical Mechanics and Its Applications*, *477*, 204–216. https://doi.org/10.1016/j.physa.2017.02.080

Manzan, S. (2017). Introduction to Financial Econometrics. In *Zicklin School of Business, Baruch College*. http://faculty.baruch.cuny.edu/smanzan/FINMETRICS/_book/_main.pdf

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3–4), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8

Morgan, J. P. (1996). *RiskMetrics-Technical Document* (Fourth Edition). J.P. Morgan/Reuters. https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a

Omari, C., Mundia, S., Ngina, I., Omari, C., Mundia, S., & Ngina, I. (2020). Forecasting Value-at-Risk of Financial Markets under the Global Pandemic of COVID-19 Using Conditional Extreme Value Theory. *Journal of Mathematical Finance*, *10*(04), 569–597. https://doi.org/10.4236/jmf.2020.104034

Raimondo, M., & Tajvidi, N. (2004). A peaks over threshold model for change-point detection by wavelets. *Statistica Sinica*, *14*(2), 395–412.

Singh, A. K. (2017). *R in finance and economics : a beginner's guide* (D. E. Allen, Ed.) [Book]. World Scientific.

Chapter 4

Backtesting VaR

4.1. Introduction

Value-at-Risk (VaR) Backtesting is a method based on a statistical hypothesis which compares the gains and losses to the evaluated VaR to check the accuracy of the applied VaR model. The main idea behind backtesting is to make sure that the forecasted risk exposure is conveniently predicting the realized risk exposure making sure that the predicted VaR forecasts the actual losses incurred considering the aspects of timing and magnitude.

A violation to VaR is an event in which the incurred loss is bigger than the absolute value of the forecasted VaR. Underestimation of VaR is the case when the frequency of losses being above their reported VaR is larger than expected, on the other hand, a very low frequency of violations compared to the expected indicates an overestimation of VaR. The phenomena appear due to problems in the parameter estimation of the VaR model or, in most cases, they are rooted to the estimation/assumption of the returns' distribution.

The literature on backtesting measures is rich, with a variety of tests that were developed in order to help researchers and institutions reduce their exposure to unexpected losses. For instance, in the work of Zhang and Nadarajah (2018) approximately 30 backtesting measures were reviewed which are classified into four main categories, that are: Unconditional test methods, conditional test methods, independence property test methods, and 'other approaches' which includes for example density forecast tests, loss functions, duration-based approach tests and several other approaches.

The unconditional test methods are based on the comparison between the percentage of violations to VaR and the VaR confidence level, of these tests we mention Kupiec's (POF) test (1995), Kupiec's TUFF test, and proportion of failures test of Haas (2001). The independence property test methods are the condition that any two elements of the hit sequence must be independent of each other (Campbell, 2005). This class of tests started with Christofferesen's independence test (1998) (Markov test) which examines whether the likelihood of a VaR violation happening on a certain day is dependent of whether a violation to VaR occurred on the previous day (Campbell, 2005). Of the independence tests we mention independence test (Christoffersen, 1998) and Wald Statistic Test (Engle & Manganelli, 2004) also known as the dynamic quantile test. Conditional test methods are based on considering both properties of unconditional coverage and independence simultaneously, like the joint test or conditional coverage test of Christoffersen (1998) and the time between failures likelihood ratio test of Haas (2001). The fourth category consists of tests that are based on different approaches like the density forecast tests of Berkowitz (2001), Lopez's magnitude loss function (Lopez, 1999), duration-based tests (Berkowitz et al., 2011; Christoffersen & Pelletier, 2004), and multivariate test (Pérignon & Smith, 2008) among many others.

Several backtesting measures are employed in this thesis to evaluate the performance of the VaR models implemented in Chapter 3. The selection of these tests was based on their wide use in the literature. These measures are: (i) Proportion of Failures test (POF) of Kupiec (1995) or, (ii) the Dynamic Quantile (DQ) test of Engle and Manganelli (2004), (iii) the duration-based (D-B) approach of (Christoffersen & Pelletier, 2004) .

Kupiec (1995) was employed since it is the basic backtesting measure of VaR and it is by far the industry standard mostly because it is implicitly incorporated in the framework for backtesting internal models proposed by the BCBS (1996) (Ziggel et al., 2014). The dynamic quantile test of Engle and Manganelli (2004) belongs to the class of tests that rely on regression models, consequently, according to Dumitrescu et al. (2012), this test is the most popular test of this class of models. Although the duration-based test is not as popular as the Kupiec test, but duration-based tests have much better power properties than the unconditional coverage test and the conditional coverage test of Christoffersen (1998) (Christoffersen & Pelletier, 2004).

To implement the above backtesting measures, the "hit sequence" of VaR violations is defined as follows,

$$I_t(q) = \begin{cases} 1 & if \quad r_t < VaR_{t|t-1}(q) \\ 0 & if \quad r_t > VaR_{t|t-1}(q) \end{cases} \text{ for any } t = 1, \dots, n$$

where r_t is the return at time t and $VaR_{t|t-1}(q)$ is the ex-ante VaR for a $q\%$ coverage rate forecast conditional on the information set available at time $t-1$ denoted by $\mathcal{F}_{t-1}$. The sequence $\{I_t\}_{t=1}^{n}$ is constructed for backtesting a risk model which indicates when the past violations occurred. The "hit sequence", $\{I_t\}_{t=1}^{n}$, follows a Bernoulli distribution, Bern (q), i.e., it is unpredictable and independently distributed such that $\Pr(I_t(q) = 1) = q$ and $\Pr(I_t(q) = 0) = (1-q)$ (Christoffersen, 2011).

4.2. Kupiec's POF test

The Kupiec's POF (Proportion of Failures) test is defined as the total number of violations $x = \sum_{t=1}^{n} I_t$ of the VaR model over the total number of observations, n, given as follows,

$$POF = \sum_{t=1}^{n} I_t / n$$

The Kupiec POF test checks the accuracy of a $VaR(q)$ measure by testing if the number of violations, x,follows a Binomial $B(n, q)$ distribution, i.e. acquires an unconditional coverage $\hat{q} = \sum_{t=1}^{n} I_t / n = x/n$ equal to q percent (it has as a null hypothesis $\hat{q} = q$), and the corresponding likelihood ratio, LR_{POF}, defined below[6]

$$POF = LR_{POF} = LR_{UC} = 2ln\left[\frac{(1-\hat{q})^{n-x}\,\hat{q}^{x}}{(1-q)^{n-x} q^{x}}\right]$$

[1]The likelihood of an i.i.d. Bernoulli (q) sequence, say here $\{I_t\}_{t=1}^{n}$, is $L(q) = \prod_{t=1}^{n}(1-q)^{1-I_t} q^{I_t} = (1-q)^{T_0} q^{T_1}$, where T_0 and T_1 are the number of 0s and 1s in the sample, respectively.

follows an asymptotic χ_1^2distribution. The latter equation thus tests whether the empirical frequency $\hat{q}$ is sufficiently close to the predicted frequency q (Dowd, 2005).

The Kupiec POF likelihood ratio tests the null hypothesis that the percentage of violations equals the 99% VaR or 95% VaR significance level, i.e., 1% or 5% respectively. Since $LR_{POF} \sim \chi_1^2$, then the critical values can be obtained from the distribution of the χ_1^2, being $x_{0.01} = 6.6349$ and $x_{0.05} = 3.84146$, corresponding to the 1% and 5% significance levels respectively. The Kupiec test however does not account for the clustering of the zeros and ones of the "hit sequence".

The POF test takes a zero value if the proportion of VaR violations is exactly equal to the promised VaR coverage probability indicating the absence of any evidence of inadequacy in the underlying VaR model, otherwise, the test grows indicating that the tested VaR model either underestimates or overestimates the underlying level of risk (Zhang & Nadarajah, 2018).

The POF test of Kupiec considers the frequency of violations and does not consider the time at which these violations occurred neither the magnitude of these violations and thus it might fail to reject a model with clustered violations.

4.3. Dynamic Quantile Test

The Dynamic Quantile (DQ) test was proposed by Engle and Manganelli (2004) after they built their conditional autoregressive VaR model (CAViaR). The authors mention that this model has been previously independently derived by Chernozhukov and Fernández-Val (2011). There are two versions of the DQ test, the in-sample and the out-of-sample DQ tests. The first one is a specification test applied in particular to test the proposed CAViaR model, however the out-of-sample DQ test is a simpler version of the former and does not depend on the estimation method used, in addition to some good features that make it useful for checking whether the VaR estimates satisfy some of the necessary properties of a good quantile like unbiasedness, independent hits and independence of quantile estimates (Engle & Manganelli, 2004).

The DQ test expands the evaluation of the independence of violations from the basic case of checking the dependence of a violation occurring on a certain day, t, on having a violation on the previous day, $t-1$, to checking the correlation of violations with other variables of the past information set as a whole including explanatory variables and higher order lags (Argyropoulos & Panopoulou, 2019).

The out-of-sample DQ test is a linear regression approach that focuses on the correlation of VaR forecasts with the available information set (Argyropoulos & Panopoulou, 2019). First, the "modified violations sequence" is defined here as follows:

$$V_t(q) = I_t(q) - q = \begin{cases} 1-q & if \quad r_t < VaR_{t|t-1}(q) \\ -q & if \quad r_t > VaR_{t|t-1}(q) \end{cases} \text{for any } t = 1, \dots, n.$$

By definition of VaR and by the Bernoulli (q) distribution of the hit/violation sequence $\{I_t\}_{t=1}^{n}$, $\mathbb{E}[I_t(q)] = q$, then, $\mathbb{E}[V_t(q)|\mathcal{F}_{t-1}] = 0$. The conditional coverage assumption is then tested on the below linear regression model:

$$V_t(q) = \delta + \sum_{k=1}^{n} \beta_k V_{t-k}(q) + \sum_{k=1}^{n} \gamma_k \, \zeta_{t-k} + \varepsilon_t$$

Where β_k and $\gamma_k (k = 1, \dots, n)$ are the coefficients of the model, δ is a constant term and ζ_{t-k} corresponds to a function of past violations and information contained in $\mathcal{F}_{t-k-1}$ from the information set available $\mathcal{F}_{t-k-1}$ and ε_t corresponds to a discrete i.i.d. process.

In other words, the null hypothesis of independence of VaR violations, DQ_{ind}, meaning that present VaR violations are not correlated with past violations, implies that the coefficients $\beta_1 = \beta_2 = \cdots = \beta_k = \gamma_1 = \cdots = \gamma_k = 0$ for all $k = 1, \dots, n$. Moreover, the null hypothesis of unconditional coverage, DQ_{UC}, is satisfied when $\delta = 0$.

This implies that under these joint hypotheses, there will be a joint nullity of coefficients which leads to a correct conditional efficiency test with $\mathbb{E}[V_t(q)] = \mathbb{E}(\varepsilon_t) = 0$ and $\mathbb{E}[I_t(q)] = Pr[I_t(q) = 1] = q$.

This means an equivalence relation between the joint nullity of coefficients and the CC test. The test statistic of the joint nullity, DQ_{CC}, is tested by the likelihood ratio, LR, and it is given in a simple form by Zhang & Nadarajah (2018) and Dumitrescu et al. (2012) as

$$DQ_{CC} = \frac{\hat{\psi}^T Z^T Z \hat{\psi}}{q(1-q)}$$

Where $\psi = (\delta, \beta_1, \dots, \beta_K, \gamma_1, \dots, \gamma_n)$ is a vector of $2n + 1$ parameters of the violation sequence, $\hat{\psi}$ is the estimator of ψ (estimated sequence of VaR values) and Z denotes a matrix of explanatory variables (i.e., the corresponding values of the portfolio). Then, DQ_{CC} follows an asymptotic chi-Square distribution with $2n + 1$ degrees of freedom, $\chi^2_{(2n+1)}$.

Moreover, the test statistic associated with the independence hypothesis, $\beta_k = \gamma_k = 0$, satisfies,

$$DQ_{ind} = \frac{\hat{\psi}^T R^T [R(Z^T Z)^{-1} R^T]^{-1} R \hat{\psi}}{q(1-q)}$$

DQ_{ind} converges to a chi-Square distribution with $2n$ degrees of freedom, $\chi^2_{(2n)}$ and $R = [0\colon I_{2n}]$ is a $(2n + 1, 2n)$ matrix so that $R\psi = \beta$, where $\beta = (\beta_1, \dots, \beta_n, \gamma_1, \dots, \gamma_n)^T$ (Dumitrescu et al., 2012).

Evaluating VaR with DQ test depends on checking the significance of the results obtained based on the corresponding p-value. To check for the statistical significance and judge whether the null hypothesis is accepted or rejected, the obtained p-value is compared to the significance level, if the p-value is smaller than the significance level then the null hypothesis is rejected, and it is accepted otherwise.

4.4. Duration-based approach

The clustering of violations is of particular importance for internal and external risk management since the rapid succession of large failures can lead to disastrous results like bankruptcy. The duration-based (D-B) approach for backtesting proposed by Christoffersen and Pelletier (2004) is based on the duration between the violations of VaR which captures the clustering of these violations.

The same definition is used as above for the "hit sequence", $\{I_t\}_{t=1}^{n}$. The main idea behind the D-B backtesting approach is that clustering of violations leads to relatively short and long no-hit periods, corresponding to market volatility and market calm, respectively.

Based on this intuition, Christoffersen and Pelletier (2004) define the duration of time, in days, between two VaR violations by $D_i = t_i - t_{i-1}$ where t_i denotes the day of violation number i.

As explained in Berkowitz et al. (2011), using the Bernoulli distribution of the "hit sequence", the probability of a hit (violation) in the next period is given by,

$$Pr(D_i = 1) = \Pr(I_{t+1} = 1) = q$$

$$Pr(D_i = 2) = \Pr(I_{t+1} = 0, I_{t+2} = 1\,) = (1-q)q$$

$$Pr(D_i = 3) = \Pr(I_{t+1} = I_{t+2} = 0, I_{t+3} = 1\,) = (1-q)^2 q$$

and the probability of violation in "d" periods can then be given by

$$Pr(D_i = d) = \Pr(I_{t+d} = 1) = \Pr(I_{t+1} = \ldots = I_{t+d-1} = 0, \ldots, I_{t+d} = 1\,) = (1-q)^{d-1} q.$$

The hazard rate of the duration D distribution, defined as probability of having a violation on day d after having $d-1$ days without violations, can be given by

$$\lambda(d) = \frac{Pr(D = d)}{1 - \sum_{j<d} Pr(D = j)} = \frac{(1-q)^{d-1} q}{1 - \sum_{j=0}^{d-2} (1-q)^j q} = q.$$

In order to test the null hypothesis, (Christoffersen & Pelletier, 2004) consider the alternative hypothesis i.e., the violations are dependent and witness clustering. Taking the null hypothesis that the risk model is correctly specified, the no-hit distribution should be memory-free and with mean duration $1/q$.

Thus, Christoffersen and Pelletier (2004) consider that under the null hypothesis, the no-hit durations, D, should follow an exponential distribution since it is the only memory-free continuous random distribution, with probability density function,

$$f_{exp}(D, q) = q\ exp(-qD)$$

In order to test for the independence property, Christoffersen and Pelletier (2004) used the Weibull distribution[7] to test for the exponential distribution as follows,

$$f_W(D,a,b) = a^b b D^{b-1} exp(-(aD)^b)$$

Where D is the random variable following the Weibull distribution with a and b the corresponding scale and shape parameters respectively. This distribution is capable of capturing the clustering of violations (Berkowitz et al., 2011). Its advantage is that it has a closed form for the hazard function given below as

$$\lambda_W(D) = \frac{f_W(D)}{1 - F_W(D)} = a^b b D^{b-1}$$

where the exponential distribution appears with flat hazard function for $b = 1$, and when $b < 1$ the Weibull distribution will have a decreasing hazard function which indicates an excessive number of very short durations between violations (high volatility periods) and also an excessive number of long durations (calm periods), providing evidence that the volatility dynamics of the employed risk model might be misspecified (Christoffersen & Pelletier, 2004). Therefore, the null hypothesis for the independence test in the Weibull case is $H_{0,\ ind}: b = 1$.

Moreover, Christoffersen and Pelletier (2004) also suggest employing the Gamma distribution under the alternative hypothesis where the probability density function is

$$f_\Gamma(D,a,b) = \frac{a^b D^{b-1} \exp(-aD)}{\Gamma(b)},$$

which also tends to an exponential distribution when $b = 1$ and hence in this case the null hypothesis for the independence test is $H_{0,\ ind}: b = 1$.

The Gamma distribution however does not admit a closed form solution for its corresponding hazard function, that's why Christoffersen and Pelletier (2004) use the first two moments, expected value and variance, which are respectively in this case $^b/_a$ and $^b/_{a^2}$, Then excess dispersion, which is used to describe the distribution of data around a central value (average), and it is by definition the variance over the squared expected value becomes $^1/_b$. It is worth noting here that in the case of exponential distribution, where the expected value of durations is $^1/_q$ and the variance is $^1/_{q^2}$, the value of the excess dispersion is 1.

According to Berkowitz et al. (2011), the null hypothesis that the risk model is correctly specified is referred to $b = 1$ and $a = q$.

---

[7] The probability density function of a Weibull random variable is $f(x;\ \lambda, k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} exp^{-\left(^x/_\lambda\right)^k} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$, with $\lambda > 0$ as the scale parameter and $k > 0$ as the shape parameter. By analogy, the above density function of the Weibull distribution $f(W; a, b)$ having the same form of $f(x;\ \lambda, k)$ is a first alternative of the original Weibull density function used usually in econometrics an adopting different parametrization such that the scale parameter $a = \frac{1}{\lambda}$ and the shape parameter $k = b$.

The Weibull and the Gamma distribution duration tests can capture higher-order dependence in the "hit sequence" showing, when they exist, the calm and turbulent market which corresponds to violation clustering. Weibull test is employed in this thesis to check the accuracy of the obtained VaR forecasts since, as per Christoffersen and Pelletier (2004), the Weibull test performs best with samples having more than 750 observations. Testing for the acceptance or rejection of the null hypothesis in this case $H_{0,\ ind}: b = 1$, is done by checking the significance of the obtained p-value as well as the obtained value of b; the null hypothesis is rejected if the p-value obtained is less than the significance level being tested and it is accepted otherwise.

The implementation of the Weibull test is based on the duration between observations, and this is where $D_i = t_i - t_{i-1}$ is actually used. Moreover, Christoffersen and Pelletier (2004), also create a series C_i to indicate whether a duration is censored, i.e., $C_i = 1$, or the duration is uncensored, $C_i = 0$.

Accordingly, if the first observation of the hit sequence starts with a zero, then D_1 is the number of days to the first hit and $C_1 = 1$, because the observed duration is left-censored[8]. However, if the hit sequence starts with a one, then D_1 is simply the number of days until the next hit and $C_1 = 0$. For the last duration, if the last observation of the hit sequence is zero, then $D_{N(T)}$ is the number of days after the last "one" in the hit sequence and $C_{N(T)} = 1$ because the observed duration is right censored[9]. Similarly, if the last observation of the hit sequence is one, then $D_{N(T)} = t_{N(T)} - t_{N(T)-1}$ and $C_{N(T)} = 0$.

According to Christoffersen and Pelletier (2004), the main contribution of an uncensored observation to the likelihood is its corresponding probability density function.

A survival function is a function that gives the probability that a process lasted past a certain time. In this case then, a censored observation shows that the process lasted D_1 or $D_{N(T)}$ days and then its contribution to the likelihood is not a probability density function however it is the aforementioned survival function denoted $S(D_i) = 1 - F(D_i)$, where $F(D_i)$ is the cumulative distribution function. Christoffersen and Pelletier (2004) then combine the censored and uncensored observations and the log-likelihood is then written as follows,

$$\ln L(D;\ \Theta) = C_1 \ln S(D_1) + (1 - C_1) \ln f(D_1) + \sum_{i=2}^{N(T)-1} \ln f(D_i) + C_{N(T)} \ln S\big(D_{N(T)}\big) + \big(1 - C_{N(T)}\big) \ln f\big(D_{N(T)}\big)$$

where Θ is the parameter of the probability distribution function $f(.)$ of D_i. Then according to Christoffersen and Pelletier (2004), once the durations are computed and truncations taken care of, then the likelihood ratio test can be calculated in a straightforward manner. The only complication that arises is that maximum likelihood estimates are no longer available in a

[8] An observed duration is left-censored when a hit sequence starts with zero, which means that the previous event when there was a violation, i.e., hit value is one, occurred before the data is collected, that is only the upper bound of the time is known.

[9] An observed duration is right-censored when the study ends before the event has occurred, in this case, the data sample ends before a next hit occurs because the last observation of the hit sequence is zero.

closed form, and they should be found using numerical optimization. For the unrestricted Weibull likelihood, it should be only numerically maximized over one parameter, since for a given value of b, the first order condition with respect to a, has an explicit solution:

$$\hat{a} = \left(\frac{N(T) - C_1 - C_{N(T)}}{\sum_{i=1}^{N(T)} D_i^b} \right)^{1/b}$$

And since b can take very small values close to zero, it is recommended by Christoffersen and Pelletier (2004) to use a^b instead of a.

And then the null hypothesis according to Berkowitz et al. (2011), the null and alternative hypotheses for the test are:

$H_0: b = 1$ and $a = q;$

$H_a: b \neq 1$ or $a \neq q;$

In the following section, the final backtesting results of the obtained VaR are demonstrated for the DJIA, SX5E and N225 during the pre-crisis, crisis, and post-crisis periods. Several packages were used in the implementation of the backtesting measures in R, like the "GAS" and "rugarch" among others. Detailed numerical values of the test statistics and p-values are found in Appendix B.

4.5. Conclusions

To evaluate and compare the performance of the models between the different intervals and for the three indexes, a summary of the results obtained is presented in Table 4.1 for the 99% VaR and in Table 4.2 for the 95% VaR.

Index	Model		99% VaR								
			2002 - 2007			2008 - 2013			2014 - 2019		
			Kupiec	DQ test	D-B test	Kupiec	DQ test	D-B test	Kupiec	DQ test	D-B test
DJIA	GARCH-EVT (95% threshold)	Normal		X	X		X	X		X	X
		Student-t		X	X		X	X		X	X
		QMLE		X	X		X	X		X	X
	GARCH-EVT (90% threshold)	Normal		X	X		X	X		X	X
		Student-t		X	X		X	X		X	X
		QMLE		X	X		X	X		X	X
	EVT (POT) (95% threshold)		NONE			X			NONE		
	EVT (POT) (90% threshold)		NONE			X			NONE		
	MCS	Normal	X		X			X			X
		Student-t (EK method)		X	X	X		X	X		X
		Student-t (Fitting method)	X	X	X	NONE				X	
	GARCH (1,1)	Normal	X		X			X			X
		Student-t (EK method)		X	X	X		X	X		X
		Student-t (Fitting method)	X	X	X	NONE				X	X
	FHS	Normal		X	X		X	X		X	X
		Student-t		X	X	NONE				X	X

	EWMA (0.94)		NONE					X	NONE		
	EWMA (0.95)		NONE			NONE			NONE		
	Var-cov		NONE			X			NONE		
	HS		X			X			X		
SX5E	GARCH-EVT (95% threshold)	Normal		X			X	X		X	X
		Student-t		X	X		X	X		X	X
		QMLE		X			X	X		X	X
	GARCH-EVT (90% threshold)	Normal		X	X		X	X	NONE		
		Student-t		X	X		X	X		X	
		QMLE		X	X		X	X	NONE		
	EVT (POT) (95% threshold)		X			X			X		X
	EVT (POT) (90% threshold)		NONE			X					X
	MCS	Normal			X			X			X
		Student-t (EK method)	X	X	X	X	X		X		X
		Student-t (Fitting method)	X	X	X	X			X	X	X
	GARCH (1,1)	Normal			X			X			X
		Student-t (EK method)	X	X	X	X	X	X	X	X	X
		Student-t (Fitting method)	X	X		X			X	X	X
	FHS	Normal		X	X		X	X		X	X
		Student-t		X	X			X		X	X
	EWMA (0.94)				X			X			X
	EWMA (0.95)				X			X			X
	Var-cov		NONE			NONE					X
	HS		X			X			X		
N225	GARCH-EVT (95% threshold)	Normal		X	X		X			X	X
		Student-t		X			X	X		X	X
		QMLE		X	X		X			X	X
	GARCH-EVT (90% threshold)	Normal		X	X		X			X	X
		Student-t		X			X	X		X	X
		QMLE		X	X		X			X	X
	EVT (POT) (95% threshold)		X		X	X			NONE		
	EVT (POT) (90% threshold)		X		X	X			NONE		
	MCS	Normal			X			X			X
		Student-t (EK method)			X	X	X	X	NONE		
		Student-t (Fitting method)			X	X	X	X		X	X
	GARCH (1,1)	Normal			X			X			X
		Student-t (EK method)			X	X		X	NONE		
		Student-t (Fitting method)			X	X	X	X		X	X
	FHS	Normal		X	X	NONE				X	X
		Student-t		X				X		X	X
	EWMA (0.94)				X			X			X
	EWMA (0.95)				X			X			X
	Var-cov		NONE			X			NONE		
	HS		X			X			X		

Table 4.1. Summary of backtesting results of the 99% VaR estimates of DJIA, SX5E and N225 before, during and after the crisis. “X” denotes that the model failed to reject the null hypothesis.

As has been said, Table 4.1 shows the backtesting results of the 99% VaR estimates under the various models of the indexes DJIA, SX5E and N225, before, during and after the crisis of 2008.

In the pre-crisis period and for the DJIA index, MCS and GARCH (1,1) models under Student-t distributed residuals (fitting method) witnessed the best backtesting results as their corresponding estimates obtained failed to reject the null hypotheses of the Kupiec, DQ test and D-B test. On the other hand, EVT (POT) models, EWMA and var-cov suffered the weakest results as all backtesting measures were rejected.

In the same period, for the SX5E index, MCS models under Student-t residuals (EK and fitting methods) as well as the GARCH (1,1) model under Student-t residuals (EK method) showed an outstanding performance in which they failed to reject the null hypotheses of all the backtesting measures applied. However, the EVT (POT) (90% threshold) and var-cov models witnessed the most unsatisfactory results among all models.

Regarding the N225 index in the pre-crisis period, GARCH-EVT (95% and 90% thresholds) under normal and QMLE fitted residuals, EVT (POT) models and FHS model under normally distributed residuals showed the best results among all models, and this is quite different from what was noticed with the DJIA and SX5E indexes. Moreover, only the var-cov model provided the worst performance because all backtesting measures were rejected with this model. It is worth noting also that the rate of "fail to reject the null hypothesis" with the N225 index is apparently lower than with the DJIA and SX5E indexes indicating that in general, all VaR models provided 99% VaR estimates of poor quality with this index.

During the crisis period, the general performance of the VaR models estimating the 99% VaR of the DJIA index changed with some noted deterioration. While MCS and GARCH (1,1) provided the best backtesting results in the pre-crisis period, during the crisis, the GARCH-EVT models under normal, Student-t and QMLE fitted residuals failed to the reject the DQ and D-B tests while the MCS and GARCH (1,1) models failed to reject the null hypotheses of Kupiec and D-B tests only under the Student-t distributed residuals (EK method). It is also worth noting that MCS under Student-t residuals (fitting method) which was a superior model in the pre-crisis period, failed all the backtesting measures during the crisis along with FHS under Student-t distributed residuals.

As for the SX5E index during the crisis period, GARCH (1,1) model under Student-t residuals (EK method) maintained similar backtesting results to those provided in the pre-crisis period and passed all the tests. MCS also under Student-t residuals (EK method) passed the Kupiec test and the DQ test. However, under the same residuals but with the fitting method, the models only passed the Kupiec test. Simultaneously, the GARCH-EVT model (95% threshold) showed better results than in the pre-crisis period and failed to reject the DQ and D-B tests under all residuals, behaving equal than GARCH-EVT model with 90% threshold. In addition, EVT (POT) (90% threshold) model's results improved slightly from rejecting all tests in the pre-crisis period to passing the Kupiec test during the crisis. However, the var-cov model maintained the same results as in the pre-crisis period and failed all tests.

While EVT (POT) models showed slightly better performance during the crisis with the DJIA and SX5E indexes and GARCH-EVT models showed better results with SX5E index, however with N225 index, these models witnessed slight deterioration during the crisis. On the other hand, MCS and GARCH (1,1) models showed a deteriorated performance compared to that witnessed in the pre-crisis period with the DJIA and SX5E indexes, to the contrary, with N225 index in the crisis period, the performance of MCS and GARCH (1,1) clearly improved and they outperformed all other models as MCS failed to reject Kupiec, DQ and D-B tests under Student-t residuals (EK and fitting methods) and GARCH (1,1) model also failed to reject the three backtesting measures under Student-t residuals (fitting method).

GARCH-EVT models maintained the same results with DJIA in the three intervals, however, the EVT (POT) models passed only the Kupiec test in the crisis period and failed all tests during the pre-crisis and post-crisis intervals. In general, the backtesting results of the DJIA in the post-crisis interval were similar except for MCS and GARCH (1,1) models which failed more tests in which MCS under Student-t residuals (fitting method) passed only the DQ test and under Student-t EK method passed the Kupiec and D-B tests. Consequently, in the post-crisis period for DJIA index, the models GARCH-EVT, MCS under Student-t (EK method), GARCH (1,1) also under Student-t residuals (EK and fitting methods), and FHS all passed two out of the three measures applied.

For the SX5E index, the GARCH-EVT (90% threshold) model performance deteriorated in the post-crisis period where it failed all tests under normal and QMLE fitted residuals while GARCH-EVT (95% threshold) maintained the same results like during the crisis. It is worth noting that the EVT (POT) performance improved in the post-crisis interval and MCS and GARCH (1,1) models under Student-t residuals (fitting method) showed better results than during the crisis period, failing to reject all backtesting tests. The best performing models for SX5E in the post-crisis period are the MCS under Student-t residuals (fitting method) and GARCH (1,1) model under Student-t residuals (EK and fitting methods) while the worst is GARCH-EVT (90% threshold) under normal and QMLE fitted residuals.

In general, VaR models provided weak 99% VaR estimates for the N225 index in the post-crisis period. While var-cov and FHS under normal distributions failed all tests in the pre-crisis and crisis periods respectively, four models failed all tests in the post crisis period, the EVT (POT) (95% and 90% thresholds), MCS model under Student-t (EK method), GARCH (1,1) model also under Student-t (EK method) and var-cov model. GARCH-EVT models showed the best results with N225 in this period passing the DQ and D-B tests under all residuals, while MCS and GARCH (1,1) models showed a deteriorated performance.

Index	Model		95% VaR								
			2002 - 2007			2008 - 2013			2014 - 2019		
			Kupiec	DQ test	D-B test	Kupiec	DQ test	D-B test	Kupiec	DQ test	D-B test
DJIA	GARCH-EVT (95% threshold)	Normal			X			X		NONE	
		Student-t			X			X			X
		QMLE			X			X		NONE	
	GARCH-EVT (90% threshold)	Normal			X			X		NONE	
		Student-t			X			X			X
		QMLE			X			X		NONE	
	EVT (POT) (95% threshold)		X				NONE		X		
	EVT (POT) (90% threshold)			NONE		X				NONE	

	MCS	Normal	X	X	X	X		X	X		X
		Student-t (EK method)		X	X		X	X	X		
		Student-t (Fitting method)	X	X	X	X	X	X	X		
	GARCH (1,1)	Normal	X	X	X			X			X
		Student-t (EK method)	X	X	X		X	X	X		
		Student-t (Fitting method)	X	X	X	X		X	X		
	FHS	Normal			X			X			X
		Student-t			X			X			X
	EWMA (0.94)		X					X	X	X	X
	EWMA (0.95)		X			X	X	X	X	X	X
	Var-cov		X			X			X		
	HS		X			X			X		
SX5E	GARCH-EVT (95% threshold)	Normal			X			X		NONE	
		Student-t			X		X	X			X
		QMLE			X			X		NONE	
	GARCH-EVT (90% threshold)	Normal			X			X		NONE	
		Student-t			X			X			X
		QMLE			X			X		NONE	
	EVT (POT) (95% threshold)		X				NONE		X		
	EVT (POT) (90% threshold)		X			X			X		
	MCS	Normal			X			X			X
		Student-t (EK method)	X	X	X	X	X	X			X
		Student-t (Fitting method)	X	X	X	X	X	X			X
	GARCH (1,1)	Normal			X			X			X
		Student-t (EK method)		X	X	X	X	X			X
		Student-t (Fitting method)	X	X	X	X	X	X			X
	FHS	Normal		X	X		X	X			X
		Student-t		X	X		X	X		X	X
	EWMA (0.94)				X	X	X	X	X		X
	EWMA (0.95)				X	X		X	X		X
	Var-cov		X			X			X		
	HS		X				NONE		X		
N225	GARCH-EVT (95% threshold)	Normal			X		NONE				X
		Student-t			X			X			X
		QMLE			X		NONE				X
	GARCH-EVT (90% threshold)	Normal			X		NONE				X
		Student-t			X			X			X
		QMLE			X		NONE				X
	EVT (POT) (95% threshold)		X				NONE		X		
	EVT (POT) (90% threshold)		X			X			X		

	MCS	Normal	NONE					X			X
		Student-t (EK method)	NONE					X			X
		Student-t (Fitting method)	NONE					X	X		X
	GARCH (1,1)	Normal	NONE					X			X
		Student-t (EK method)	NONE					X			X
		Student-t (Fitting method)	NONE					X	X		X
	FHS	Normal			X			X			X
		Student-t			X	NONE					X
	EWMA (0.94)				X	X	X	X	X		X
	EWMA (0.95)		X		X	X	X	X	X		X
	Var-cov		X			NONE			X		
	HS		X			NONE			X		

Table 4.2. Summary of backtesting results of the 95% VaR estimates of DJIA, SX5E and N225 before, during and after the crisis. "X" denotes that the model failed to reject the null hypothesis.

Table 4.2 presents the backtesting results of the 95% VaR estimates obtained under the different VaR models for the DJIA, SX5E and N225 indexes before, during and after the crisis of 2008.

For the DJIA index in the pre-crisis interval, the GARCH-EVT and EVT (POT) models seem to have weak performance passing only one test and EVT (POT) (90% threshold) failing all tests. However, MCS under normal and Student-t residuals (fitting method) and all GARCH (1,1) models fail to reject the null hypotheses of Kupiec, DQ and D-B tests and outperforming the remainder of the models.

Similar results were obtained with the SX5E index in the same period as GARCH-EVT and EVT (POT) models only passed one test while MCS under Student-t (EK and fitting methods) and GARCH (1,1) under Student-t (fitting method) failed to reject all the tests. No model failed all tests.

However, the N225 index results in the same interval were, again, apparently different from those of the DJIA and SX5E indexes. The MCS and GARCH (1,1) models failed all tests and under all distributions of residuals while GARCH-EVT and EVT (POT) showed similar results to those of DJIA and SX5E and passed one test under each distribution of residuals and the remainder models witnessed similar results except for the EWMA (0.95) model which failed to reject two tests, the Kupiec and D-B test and thus outperforming all models.

During the crisis, regarding the DJIA index, GARCH-EVT and EVT (POT) models maintained similar results to those in the pre-crisis period. However, the models with best results apparently are no longer the same, only MCS under Student-t (fitting method) and EWMA (0.95) models passed the three tests while the remainder models passed either one or two tests. EVT (POT) (95% threshold) failed all tests.

SX5E index witnessed similar backtesting results during the crisis as in the pre-crisis interval moreover, some models witnessed obvious improvements. Accordingly, the best performing models that passed all tests were MCS and GARCH (1,1) models under Student-t residuals

(EK and fitting methods) and EWMA (0.94). On the other hand, the performance of EVT (POT) (95% threshold) and HS models deteriorated, and they failed all tests.

The backtesting results of the N225 index during the crisis witnessed huge changes. GARCH-EVT models under normal and QMLE residuals failed all tests. Moreover, EVT (POT) (95% threshold), FHS under Student-t residuals, var-cov and HS models also failed all tests. However, the backtesting results of EWMA (0.94 and 0.95) models improved during the crisis and the models passed all tests.

In the post-crisis period, the backtesting results of DJIA index witnessed some changes as the GARCH-EVT (95% and 90% thresholds) models under normal and QMLE fitted residuals and EVT (POT) (90% threshold) failed all tests. On the other hand, EWMA (0.95) model maintained the same result as in the crisis period and EWMA (0.94) model also passed all tests during this period.

SX5E index in the post crisis interval also witnessed deterioration in the results of GARCH-EVT models which failed all tests under normal and QMLE fitted residuals. On the other hand, MCS and GARCH (1,1) models failed the Kupiec and DQ tests and EWMA models also failed the DQ test. The best performing models in this period for the SX5E index were the FHS model with Student-t residuals, and EWMA (0.94 and 0.95) as each of these models passed two tests while the remainder models failed to reject only one test.

As for the N225 index, it can be noticed that it is the only index in this period where all models failed to reject at least one null hypothesis of a backtesting measure. MCS and GARCH (1,1) under Student-t residuals (fitting method) showed an improvement and passed the Kupiec and D-B tests, while EWMA (0.94 and 0.95) models suffered a slight deterioration and rejected the DQ test. It is worth noting also that GARCH-EVT models under normal and QMLE fitted residuals and EVT(POT) (95% threshold) showed better results during this period than in the pre-crisis period and passed one test. However, EWMA models still outperformed the remainder models like in the crisis period along with the MCS and GARCH (1,1), both under Student-t residuals (fitting method).

Table 4.3 shows a summary for the best performing models for estimating the 99% and 95% VaR of the DJIA, SX5E and N225 indexes as per the backtesting results in the pre-crisis, crisis, and post-crisis intervals.

		DJIA	SX5E	N225
99% VaR	2002 - 2007	• MCS (Student-t) (fitting method) • GARCH (1,1) (Student-t) (fitting method)	• MCS (Student-t) (EK and fitting methods) • GARCH (1,1) (Student-t) (EK method)	• GARCH-EVT (95% threshold) (Normal and QMLE) • GARCH-EVT (90% threshold) (Normal and QMLE) • EVT (POT) (95% threshold) • EVT (POT) (90% threshold) • FHS (Normal)
	2008 - 2013	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold)) • MCS (Student-t) (EK method) • GARCH (1,1) (Student-t) (EK method) • FHS (Normal)	• GARCH (1,1) (Student-t) (EK method)	• MCS (Student-t) (EK and fitting methods) • GARCH (1,1) (Student-t) (fitting method)
	2014 - 2019	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS (Student-t) (EK method) • GARCH (1,1) (Student-t) (EK and fitting methods) • FHS (Normal and Student-t)	• MCS (Student-t) (fitting method) • GARCH (1,1) (Student-t) (EK and fitting methods)	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS (Student-t) (fitting method) • GARCH (1,1) (Student-t) (fitting method) • FHS (Normal and Student-t)
95% VaR	2002 - 2007	• MCS (Normal and Student-t fitting method) • GARCH (1,1) (Normal, Student-t EK and fitting methods)	• MCS (Student-t) (EK and fitting methods) • GARCH (1,1) (Student-t) (fitting method)	• EWMA (0.95)
	2008 - 2013	• MCS (Student-t) (fitting method) • EWMA (0.95)	• MCS (Student-t) (EK and fitting methods) • GARCH (1,1) (Student-t) (EK and fitting method) • EWMA (0.94)	• EWMA (0.94) • EWMA (0.95)
	2014 - 2019	• EWMA (0.94) • EWMA (0.95)	• FHS (Student-t) • EWMA (0.94) • EWMA (0.95)	• MCS (Student-t) (fitting method) • GARCH (1,1) (Student-t) (fitting method) • EWMA (0.94) • EWMA (0.95)

Table 4.3. Summary of the chapter; best performing models for estimating 99% VaR and 95% VaR according to backtesting measures classified by index and interval.

According to Table 4.3, as per the backtesting results of the 99% VaR estimates, MCS and GARCH (1,1) models were the models with highest acceptance rate (failing to reject null hypotheses of the backtesting measures) among all models.

To summarize the findings of this chapter, it is a must to analyze the variations observed in Table 4.3 and distinguish which models presented superior performance over the others. This analysis is carried out by index at the 99% and 95% confidence levels, highlighting the models that witnessed the best backtesting results.

DJIA

For the 99% VaR results, MCS and GARCH (1,1) models under Student-t distributed residuals had the highest acceptance rate among all models and they failed to reject all tests in the three intervals. GARCH-EVT and FHS models showed good results during and after the crisis only. As for the 95% VaR estimates backtesting results, MCS showed good performance before and during the crisis periods while GARCH (1,1) model provided good results only in the pre-crisis interval. On the other hand, EWMA models provided good results during and after the crisis periods.

SX5E

As for the SX5E, MCS and GARCH (1,1) models under Student-t residuals provided the best backtesting results for the 99% VaR estimates before, during and after the crisis. Regarding the 95% VaR estimates backtesting results, MCS and GARCH (1,1) acquired the best backtesting results in the pre-crisis and crisis intervals, however, during the post-crisis period, FHS under Student-t residuals and EWMA (0.94 and 0.95) models outperformed the remainder models.

N225

During the pre-crisis interval, the 99% VaR estimates backtesting results of N225 index showed that GARCH-EVT models, EVT (POT) models and FHS model under normal residuals provided the best VaR estimates. However, during the crisis, only MCS and GARCH (1,1) models under Student-t distributed residuals provided acceptable results, while in the post-crisis interval GARCH-EVT, MCS, GARCH (1,1) and FHS models presented similar results outperforming the remainder models. Regarding the 95% VaR estimates backtesting results, the models providing the best VaR estimates were different from those observed at the 99% confidence level. It can be noticed that the EWMA models outperform the remainder models in all intervals. Only in the post-crisis period, MCS and GARCH (1,1) models provide good backtesting results along with the EWMA models.

****

Table 4.4 shows the models that provided the best backtesting results for the 99% and 95% VaR for each index in each interval with details about which backtesting measures they failed to reject. This highlights the performance of the models that witnessed the best backtesting results in each interval with each index and outperformed the remaining models.

99% VaR															
Index	**Model**		**2002 - 2007**			**Model**		**2008 - 2013**			**Model**		**2014 - 2019**		
			Kupiec	**DQ test**	**D-B test**			**Kupiec**	**DQ test**	**D-B test**			**Kupiec**	**DQ test**	**D-B test**
DJIA	MCS	Student-t (Fitting method)	X	X	X	GARCH-EVT (95% threshold)	Normal		X	X	GARCH-EVT (95% threshold)	Normal		X	X
	GARCH (1,1)	Student-t (Fitting method)	X	X	X		Student-t		X	X		Student-t		X	X
							QMLE		X	X		QMLE		X	X
						GARCH-EVT (90% threshold)	Normal		X	X	GARCH-EVT (90% threshold)	Normal		X	X
							Student-t		X	X		Student-t		X	X
							QMLE		X	X		QMLE		X	X
						MCS	Student-t (EK method)	X		X	MCS	Student-t (EK method)	X		X
						GARCH (1,1)	Student-t (EK method)	X		X	GARCH (1,1)	Student-t (EK method)	X		X
						FHS	Normal		X	X		Student-t (Fitting method)		X	X
											FHS	Normal		X	X
												Student-t		X	X
SX5E	MCS	Student-t (EK method)	X	X	X	GARCH (1,1)	Student-t (EK method)	X	X	X	MCS	Student-t (Fitting method)	X	X	X
	MCS	Student-t (Fitting method)	X	X	X						GARCH (1,1)	Student-t (EK method)	X	X	X
	GARCH (1,1)	Student-t (EK method)	X	X	X							Student-t (Fitting method)	X	X	X
N225	GARCH -EVT (95%	Normal		X	X	MCS	Student-t (EK method)	X	X	X	GARCH-EVT (95% threshold)	Normal		X	X

	threshold)														
		QMLE		X	X							Student-t		X	X
	GARCH-EVT (90% threshold)	Normal		X	X		Student-t (Fitting method)	X	X	X		QMLE		X	X
		QMLE		X	X	GARCH (1,1)	Student-t (Fitting method)	X	X	X		Normal		X	X
	EVT (POT) (95% threshold)		X		X						GARCH-EVT (90% threshold)	Student-t		X	X
	EVT (POT) (90% threshold)		X		X							QMLE		X	X
	FHS	Normal		X	X						MCS	Student-t (Fitting method)		X	X
											GARCH (1,1)	Student-t (Fitting method)		X	X
											FHS	Normal		X	X
												Student-t		X	X
95% VaR															
DJIA	MCS	Normal	X	X	X	MCS	Student-t (Fitting method)	X	X	X	EWMA (0.94)		X	X	X
		Student-t (Fitting method)	X	X	X	EWMA (0.95)		X	X	X	EWMA (0.95)		X	X	X
		Normal	X	X	X										

	GARCH (1,1)	Student-t (EK method)	X	X	X										
	GARCH (1,1)	Student-t (Fitting method)	X	X	X										
SX5E	MCS	Student-t (EK method)	X	X	X	MCS	Student-t (EK method)	X	X	X	FHS	Student-t		X	X
	MCS	Student-t (Fitting method)	X	X	X	MCS	Student-t (Fitting method)	X	X	X	EWMA (0.94)		X		X
	GARCH (1,1)	Student-t (Fitting method)	X	X	X	GARCH (1,1)	Student-t (EK method)	X	X	X	EWMA (0.95)		X		X
						GARCH (1,1)	Student-t (Fitting method)	X	X	X					
						EWMA (0.94)		X	X	X					
N225	EWMA (0.95)		X		X	EWMA (0.94)		X	X	X	MCS	Student-t (Fitting method)	X		X
						EWMA (0.95)		X	X	X	GARCH (1,1)	Student-t (Fitting method)	X		X
											EWMA (0.94)		X		X
											EWMA (0.95)		X		X

Table 4.4. The best performing models for 99% and 95% VaR for each index in each interval and the respective backtests that they failed to reject.

99% VaR

In Table 4.4, it can be seen that, for the DJIA and SX5E indexes, MCS under Student-t residuals (Fitting method for DJIA and fitting and EK methods for SX5E)) and GARCH (1,1) model under Student-t residuals (fitting method for DJIA and EK method for SX5E) failed to reject the Kupiec, DQ and D-B tests in the pre-crisis period, indicating that the percentage of violations of the VaR estimates corresponding to these models is close to the promised VaR coverage probability (due to Kupiec test). Moreover, the DQ test hypothesis proposes that the violations of VaR are independent from all previous explanatory variables and higher order lags contained in the past information set. In addition, failing to reject the D-B test also shows that the violations to VaR estimates obtained under MCS and GARCH (1,1) with DJIA and SX5E in the pre-crisis interval are not clustered.

As for the N225 index, MCS and GARCH (1,1) models were outperformed by GARCH-EVT (95% and 90% thresholds), EVT (POT) (95% and 90% thresholds) both under normal and QMLE fitted residuals as well as FHS under normally distributed residuals. GARCH-EVT (95% and 90% thresholds) and FHS fail to reject the DQ and D-B tests showing the absence of violation clustering and the independence of violations from past violations, explanatory variables, and higher order lags. On the other hand, EVT (POT) (90% and 95% thresholds) pass the Kupiec and D-B tests, thus the frequency of violations of VaR is close to the VaR coverage probability and the violations are independent of violations to VaR from past violations and information.

During the crisis period, GARCH (1,1) under Student-t residuals (EK method) with SX5E and N225 indexes fail to reject the null hypotheses of the Kupiec, DQ and D-B tests, while with the DJIA index, MCS and GARCH (1,1) under Student-t residuals (EK method) pass only the Kupiec and D-B tests. Following with the DJIA, the GARCH-EVT (95% and 90% thresholds) models under all residuals along with FHS under normal residuals pass the DQ and D-B tests. It is also worth noting that during the crisis, MCS and GARCH (1,1) models show the best performance with N225 among all models in all periods.

In the post-crisis period, and as for DJIA, GARCH-EVT models (90% and 95% thresholds), MCS under Student-t residuals (EK method), GARCH (1,1) model also with Student-t residuals (EK and fitting methods) and FHS models appear to be the best models according to the backtetsing results. However, as for the SX5E, only MCS under Student-t residuals (fitting method) and GARCH (1,1) model also under Student-t residuals (EK and fitting methods) show the best performance among all models. In the same period, the performance of the best performing models, MCS and GARCH (1,1) of N225 index deteriorated and the 99% VaR estimates of these models failed to reject two tests only, the DQ and D-B tests. Moreover, other models showed a good performance with N225 index in the post-crisis period like GARCH-EVT (95% and 90% thresholds) and the FHS models under normal and Student-t residuals passing also the DQ and D-B tests.

95% VaR

During the pre-crisis period, MCS and GARCH (1,1) models under normal and Student-t residuals appear to be the best performing models for the DJIA index, failing to reject all the backtesting measures hypotheses. The same models were also outperforming the rest of the models also with SX5E index but only under Student-t distributed residuals and they also passed all the backtesting measures. However, only one model showed superior results with N225 index, and it was the EWMA (0.95) model, and it only passed the Kupiec and D-B tests.

During the crisis period, MCS under Student-t residuals and EWMA (0.95) models provide the best backtesting results among all models of DJIA index, noting that they pass all tests. On the other hand, MCS and GARCH (1,1) under Student-t residuals along with EWMA (0.94) model is apparently the best performing models during the crisis period for estimating the 95% VaR of SX5E index. This model here also fails to reject the hypotheses of the three backtesting measures. During the same period, the only models that provided the best backtesting results for the VaR estimates of N225 index were the EWMA (0.94 and 0.95) models which passed all the backtesting measures.

In the post-crisis period, EWMA model seems to be providing good 95% VaR estimates for the three indexes in general. It can be seen that EWMA (0.94 and 0.95) fail to reject the three backtesting measures null hypotheses with the DJIA index. Moreover, FHS under Student-t residuals along with EWMA (0.94 and 0.95) outperform the remainder models with SX5E index, FHS passes the DQ and D-B tests while EWMA models pass the Kupiec and D-B tests. As for the N225 index, MCS and GARCH (1,1) models under Student-t residuals (fitting method) and the EWMA (0.94 and 0.95) fail to reject the Kupiec and D-B tests providing better results than the other models.

To conclude, it can be said that MCS, GARCH (1,1) and GARCH-EVT models are dominant models on the 99% VaR level as they provide the best backtesting results compared to the remainder models. However, it can be argued that percentages of violations to the estimated 99% VaR under MCS and GARCH (1,1) models, with Student-t residuals, almost in all periods and all indexes, were close to the VaR coverage probabilities. Moreover, the violations were independent from the past set of violations and were not clustered. This is contrary to the GARCH-EVT models which always passed the DQ and D-B tests and failed the Kupiec test, which means that the frequency of violations to VaR estimated under these models was not close to the VaR coverage probability.

In Conclusions, an extensive comparative analysis will be conducted on the findings of Chapters 3 and 4 to narrow the options of the best VaR model(s). This analysis will take into account the distance measures along with the backtesting measures and the behavior of VaR models relative to the market evolution of each of the indexes DJIA, SX5E and N225, before during and after the crisis of 2008.

References

Argyropoulos, C., & Panopoulou, E. (2019). Backtesting VaR and ES under the magnifying glass. *International Review of Financial Analysis*, *64*, 22–37. https://doi.org/10.1016/j.irfa.2019.04.005

Basel Committee on Banking Supervision. (1996). *Amendment to the capital accord to incorporate market risks*. *January*. https://www.bis.org/publ/bcbs24.pdf

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics*, *19*(4), 465–474. https://doi.org/10.1198/07350010152596718

Berkowitz, J., Christoffersen, P., & Pelletier, D. (2011). Evaluating Value-at-Risk Models with Desk-Level Data [Article]. *Management Science*, *57*(12), 2213–2227. https://doi.org/10.1287/mnsc.1080.0964

Campbell, S. (2005). A review of backtesting and backtesting procedures. *Journal of Risk*, *9*. https://doi.org/10.21314/JOR.2007.146

Chernozhukov, V., & Fernández-Val, I. (2011). Inference for extremal conditional quantile models, with an application to market and birthweight risks. *Review of Economic Studies*, *78*(2), 559–589. https://doi.org/10.1093/restud/rdq020

Christoffersen, P. (1998). EVALUATING INTERVAL FORECASTS. *International Economic Review*, *39*(4), 841–862. https://doi.org/10.2307/2527341

Christoffersen, P. (2011). *Elements of Financial Risk Management, 2nd Edition* (2nd ed.) [Book]. Academic Press.

Christoffersen, P., & Pelletier, D. (2004). Backtesting Value-at-Risk: A Duration-Based Approach. *Journal of Financial Econometrics*, *2*(1), 84–108. https://doi.org/10.1093/jjfinec/nbh004

Dowd, K. (2005). *Measuring market risk* (2nd ed) [Book]. John Wiley & Sons.

Dumitrescu, E. I., Hurlin, C., & Pham, V. (2012). Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests [Article]. *Finance (Paris)*, *33*, 79–112.

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, *22*(4), 367–381. https://doi.org/10.1198/073500104000000370

Haas, M. (2001). *New methods in backtesting*. https://www.ime.usp.br/~rvicente/risco/haas.pdf

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, *3*(2), 73–84.

Lopez, J. A. (1999). Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of San Francisco Economic Review*, *2*(2), 3–17. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0042141025&partnerID=40&md5=eb6548afcb5e081f266e5983632b678c

Pérignon, C., & Smith, D. R. (2008). A new approach to comparing VaR estimation methods. *Journal of Derivatives*, *16*(2), 54–66. https://doi.org/10.3905/JOD.2008.16.2.054

Zhang, Y., & Nadarajah, S. (2018). A review of backtesting for value at risk. *Communications in Statistics - Theory and Methods*, *47*(15), 3616–3639. https://doi.org/10.1080/03610926.2017.1361984

Ziggel, D., Berens, T., Weiß, G. N. F., & Wied, D. (2014). A new set of improved Value-at-Risk backtests. *Journal of Banking and Finance*, *48*, 29–41. https://doi.org/10.1016/j.jbankfin.2014.07.005

Chapter 5

Conclusions

5.1. Summary and conclusions

This thesis presents an extensive study on the eight main VaR models evaluating their performance in three stock indexes representing the geographical areas of United States (Dow Jones Industrial Average, DJIA), Europe (Euro Stoxx 50, SX5E) and Japan (Nikkei 225, N225), during three different intervals, that is before, during and after the crisis of 2008. These time intervals will allow us to analyze whether there are differences in the behavior of models depending on the characteristics of the period, growth, crisis, or recovery, or if the analyzed period does not affect the validity of the model.

The VaR models implemented and analyzed in this thesis are the variance-covariance (var-cov), historical simulation (HS), Monte Carlo simulation (MCS), GARCH (1,1) model, Filtered Historical simulation (FHS), exponentially weighted moving average model (EWMA), Extreme value theory-peaks over threshold model (EVT-POT) and GARCH-EVT model.

When possible, some models were applied in two frameworks: (i) normally distributed residuals and (ii) Student-t distributed residuals and the latter was implemented in two different ways depending on the calculation method of the degree of freedom “d” of the Student-t distribution. The models applied in two frameworks are MCS, FHS, GARCH (1,1), and GARCH-EVT models. The remainder models were applied according to their original framework.

As seen in Chapter 3, the data of the three stock indexes was found to be leptokurtic due to the high kurtosis, which explains why models with Student-t distributed residuals outperformed those with normally distributed residuals most of the time. This was shown in Chapters 3 and 4.

To evaluate the performance of the implemented models different backtesting measures were employed. The backtesting measures used were Kupiec (1995) test, Dynamic Quantile test of Engle and Manganelli (2004) and Duration-based test of Christoffersen and Pelletier (2004). Moreover, distance measures were also used to recognize the models that provided the nearest VaR estimates to the actual returns. The use of distance measures is a contribution that we consider interesting. They not only take into account violations of the VaR estimates, they also consider which estimates best fit actual returns, with special focus on cases of greater losses or when the actual returns are positive.

To sum up our conclusions, Tables 5.1 and 5.2 show the models that provided the best 99% and 95% VaR estimates, respectively, according to distance measures and backtesting measures for each index in each period. Consequently, as per the criteria of each method of evaluation used here, a model that provides the best results as per distance measures means that its corresponding VaR estimates are the closest to the actual returns than other VaR estimates provided by the competing models. Similarly, the models that provided the best backtesting results means that the violations to the VaR estimates of these models were independent, non-clustered and if they also passed the Kupiec test it means that the percentage of violations met the expected probability coverage of VaR. Then, a model that shows up as the best performing model according to both criteria should be a superior model. The analysis below will be index based since the return distribution of the indexes differs from one to the other and this will also be helpful to see what models work best in each geographical area that the indexes represent.

		DJIA		SX5E		N225	
		Distance Measures	Backtesting measures	Distance Measures	Backtesting measures	Distance Measures	Backtesting measures
99% VaR	2002 - 2007	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS (normal) • **GARCH (1,1)** (normal and **Student-t (fitting method))** • FHS	• MCS (Student-t) (fitting method) • **GARCH (1,1) (Student-t (fitting method))**	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • **MCS (Student-t (**EK and **fitting methods))** • GARCH (1,1) (Student-t (fitting method)) • FHS	• **MCS (Student-t) (EK and fitting methods)** • GARCH (1,1) (Student-t (EK method))	• MCS (Student-t (EK and fitting methods)) • GARCH (1,1) (Student-t (EK and fitting methods))	• GARCH-EVT (95% threshold) (Normal and QMLE) • GARCH-EVT (90% threshold) (Normal and QMLE) • EVT (POT) (95% threshold) • EVT (POT) (90% threshold) • FHS (Normal)
	2008 - 2013	• **GARCH-EVT (95% threshold)** • **GARCH-EVT (90% threshold)** • **MCS (Student-t (EK** and fitting methods)) • **GARCH (1,1) (Student-t (EK** and fitting methods)) • **FHS (Normal** and Student-t)	• **GARCH-EVT (95% threshold)** • **GARCH-EVT (90% threshold)** • **MCS (Student-t (EK method))** • **GARCH (1,1) (Student-t (EK method))** • **FHS (Normal)**	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS **(Student-t (EK** and fitting methods)) • **GARCH (1,1) (Student-t (EK** and fitting methods)) • FHS	• **GARCH (1,1) (Student-t (EK method))**	• GARCH-EVT (95% threshold) • **MCS (Student-t (EK and fitting methods))** • **GARCH (1,1) (Student-t** (EK and **fitting methods))**	• **MCS (Student-t) (EK and fitting methods)** • **GARCH (1,1) (Student-t (fitting method))**
	2014 - 2019	• **MCS (Student-t (EK** and fitting method)) • **GARCH (1,1) (Student-t (EK** and fitting methods))	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • **MCS (Student-t) (EK method)** • **GARCH (1,1) (Student-t (EK** and fitting methods)) • FHS (Normal and Student-t)	• **MCS (Student-t** (EK and **fitting method))** • **GARCH (1,1) (Student-t (EK and fitting methods))**	• **MCS (Student-t (fitting method)** • **GARCH (1,1) (Student-t (EK and fitting methods))**	• MCS (Student-t (EK method)) • GARCH (1,1) (Student-t (EK method))	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • MCS (Student-t) (fitting method) • GARCH (1,1) (Student-t (fitting method)) • FHS (Normal and Student-t)

Table 5.1. Best performing models as per distance measures and backtesting measures on the 99% confidence level. Models in bold font are common between distance measures and backtesting measures.

99% VaR, DJIA index

In the pre-crisis period, GARCH (1,1) under Student-t distributed residuals (fitting method) was the only model that appeared best performing according to both criteria. It has been previously seen in Chapter 4 that this model passed all the backtesting measures, and thus the percentage of violations to its corresponding VaR estimates is consistent with the promised VaR coverage probability, moreover, the violations are independent of past lagged violations and there is no clustering of violations. In addition, since this model also witnesses the best distance measures then its VaR estimates are closest to the actual returns. On the other hand, this does not deny the other models that also witness good backtesting and distance measures their good performance.

During the crisis period, there are several models that fulfill both criteria. GARCH-EVT (95% and 90% thresholds), MCS and GARCH (1,1) both under Student-t (EK method), and FHS under normally distributed residuals. These models outperform the remainder models during the crisis period. While GARCH-EVT (95% and 90% thresholds) along with FHS under normal residuals passed the DQ and D-B tests, the violations to their corresponding VaR estimates were independent of past violations and not clustered. Meanwhile MCS and GARCH (1,1) models under Student-t residuals (EK method) failed to reject the Kupiec and D-B tests implying that the percentages of violations under these models are equal to the promised VaR coverage probability, and they are not clustered. On the other hand, all these models provide VaR estimates closer to the actual returns than the remainder models. These facts make these models superior to the other models when estimating the 99% VaR of DJIA index during the crisis period.

In the post-crisis period, MCS and GARCH (1,1) models under Student-t residuals (EK method) appear as best models under both evaluation criteria and outperform the remainder models. Simultaneously, these two models passed the same backtesting measures as it can be seen in Chapter 4 (Table 4.4). They failed to reject Kupiec and D-B tests, and they also provided the best results for distance measures among models.

99% VaR, SX5E index

In the pre-crisis period, MCS under Student-t residuals (EK and fitting methods) provided the best 99% VaR estimates for the SX5E index. The MCS under Student-t residuals implemented in two different methods achieved the closest 99% VaR estimates to the actual returns and failed to reject all the backtesting measures employed to evaluate the VaR models performance.

During the crisis, the GARCH (1,1) model under Student-t residuals (EK method) fails to reject all backtesting measures and provided the best VaR estimates according to the distance measures.

In the post-crisis period, MCS under Student-t residuals (fitting method) and GARCH (1,1) model under Student-t residuals (EK and fitting methods) outperform the remainder models and achieve good results in distance measures and backtesting measures.

GARCH-EVT models provided good results in the pre-crisis and crisis intervals only according to distance measures.

99% VaR, N225 index

The nature of the N225 index data differs from that of the DJIA and SX5E indexes. Consequently, the VaR models witness different behavior with this index as well. It can be seen in Table 5.1 that in the pre-crisis period, no model acquires a superior behavior according to backtesting and distance measures at the same time. The models that achieve good results according to distance measures do not achieve the same results according to backtesting measures. Thus, no model can be labeled as “superior” in this period with N225 index according to our evaluation criteria.

However, during the crisis period, MCS under Student-t residuals (EK and fitting methods) and GARCH (1,1) model under Student-t residuals (fitting method) achieve good results with the backtesting and distance measures. It is also worth noting that these models during the crisis period passed all the backtesting measures unlike the other models in the pre-crisis and post-crisis period which only failed to reject two out of the three backtesting measures applied. It is important to highlight that GARCH-EVT model (95% threshold) provided good results in this interval only as per distance measures while it performed well in pre-crisis and post-crisis only as per backtesting measures.

The post-crisis period looks slightly like the pre-crisis period as for the best VaR models. Moreover, no models achieve the best results according to both evaluation criteria and thus no model appear to be superior in this period.

In summary, for 99% VaR estimates, there is no model that can be considered to adequately fit all the indexes and in all the periods analyzed. But, surprisingly, in the crisis period it is possible to find one model that behaves better than the others for all indexes: the MCS under Student-t residuals model (EK method). Consequently, the best model to apply will depend on the index analyzed and on the economic situation of the period.

When analyzing distances only, the GARCH (1,1) model under Student-t residuals (fitting method) is the appropriate model for all indexes and periods (except N225 in the post-crisis period). Regarding backtesting measures only, there is no model that outperforms the others.

		DJIA		**SX5E**		**N225**	
		Distance Measures	Backtesting measures	Distance Measures	Backtesting measures	Distance Measures	Backtesting measures
95% VaR	**2002 – 2007**	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • **GARCH (1,1) (Student-t (fitting method))** • FHS	• MCS (Normal and Student-t fitting method) • **GARCH (1,1)** (Normal, Student-t Ek and **fitting methods)**	• GARCH-EVT (95% threshold) • GARCH-EVT (90% threshold) • **GARCH (1,1)** (Normal, Student-t Ek and **fitting methods)** • FHS	• MCS (Student-t) (EK and fitting methods) • **GARCH (1,1) (Student-t) (fitting method)**	• None	• EWMA (0.95)
	2008 – 2013	• GARCH-EVT (95% threshold) • **MCS (Student-t (fitting method))** • FHS	• **MCS (Student-t (fitting method))** • EWMA (0.95)	• GARCH-EVT (95% threshold) (Student-t) • GARCH-EVT (90% threshold) (Student-t) • FHS	• MCS (Student-t) (EK and fitting methods) • GARCH (1,1) (Student-t) (EK and fitting method) • EWMA (0.94)	• None	• EWMA (0.94) • EWMA (0.95)
	2014 – 2019	• MCS (Student-t (fitting method)) • GARCH (1,1) (Student-t (fitting method)) • FHS	• EWMA (0.94) • EWMA (0.95)	• None	• FHS (Student-t) • EWMA (0.94) • EWMA (0.95)	• **MCS (Student-t** (EK and **fitting methods))** • **GARCH (1,1) (Student-t (fitting method))**	• **MCS (Student-t) (fitting method)** • **GARCH (1,1) (Student-t (fitting method))** • EWMA (0.94) • EWMA (0.95)

Table 5.2. Best performing models as per distance measures and backtesting measures on the 95% confidence level. Models in bold font are common between distance measures and backtesting measures.

Table 5.2 presents the best performing VaR models in estimating the 95% VaR of the DJIA, SX5E and N225 indexes according to the distance and backtesting measures.

95% VaR, DJIA index

In the pre-crisis period, GARCH (1,1) model under Student-t residuals (fitting method) was classified among the best performing models with respect to distance measures and backtesting measures. This model failed to reject the null hypotheses of the three backtestting measures applied and provided the neasrest 95% VaR estimates to the actual returns.

However, during the crisis period the MCS model under Student-t residuals (fitting method) was the only model that achieved good results according to both evaluation methods and it also passed the three backtesting measures.

In the post-crisis period, no VaR model provided good results with both evaluation criteria at the same time. While MCS, GARCH (1,1) and FHS provided the best results for distance measures, EWMA (0.94 and 0.95) model provided the best results for backtesting measures.

95% VaR, SX5E index

In the pre-crisis period, GARCH (1,1) model under Student-t residuals (fitting method) provided the best results with backtesting measures and distance measures. This result is similar to that of the DJIA index in the same period and same level of confidence.

However, during the crisis, no models presented good results for both criteria at the same time, while GARCH-EVT models and FHS provided good distance measures, MCS, GARCH (1,1) and EWMA (0.94) models witnessed good results with backtesting measures.

The same can be said about the post-crisis period since none of the models provided good distance measures while FHS and EWMA (0.94 and 0.95) provided good results with backtesting measures.

95% VaR, N225 index

The results of the N225 index were obviously different from those of the DJIA and SX5E indexes.

In the pre-crisis period, none of the models provided good results with distance measures and only EWMA (0.95) provided good outcome with backtesting measures.

During the crisis, the same could be noticed regarding distance measures and EWMA (0.94 and 0.95) were the only models that provided acceptable results with backtesting measures.

However, in the post-crisis period, MCS and GARCH (1,1) models under Student-t distribution (fitting method) were the best performing models according to both evaluation criteria.

It can be also noticed that GARCH-EVT models provided good 95% VaR estimates for DJIA and SX5E in the pre-crisis and crisis intervals only with respect to distance measures. While EWMA models (0.95) showed good backtesting results with the 95% VaR estimates of the N225 index in all intervals and with SX5E index in the post-crisis period and with DJIA index in the crisis and post-crisis periods.

In summary, for 95% VaR estimates, in the pre-crisis period the GARCH (1,1) model under Student-t distribution (fitting method) is the best model for both the DJIA and the SX5E. There is no other coincidence for either periods or indexes.

Taking into account only the backtesting measures, EWMA (0.95) is the best model in the post-crisis period for all indexes, and for the Nikkei in all the analyzed periods. Considering only the distance measures, GARCH-EVT (95%) performs appropriately in the periods before and during the crisis for DJIA and SX5E.

****

To concrete our conclusions, we summarize the analysis in Table 5.3 which shows only the best models for each index estimating the 99% VaR and 95% VaR during each of the three intervals.

	99% VaR			95% VaR		
	DJIA	SX5E	N225	DJIA	SX5E	N225
2002-2007	• **GARCH (1,1) (Student-t (fitting method))**	• **MCS (Student-t (EK and fitting methods))**	• **None (different models under each testing criteria)**	• **GARCH (1,1) (Student-t (fitting method))**	• **GARCH (1,1) (Student-t) (fitting method))**	• **None (different models under each testing criteria)**
2008-2013	• **GARCH-EVT (95% threshold)** • **GARCH-EVT (90% threshold)** • **MCS (Student-t (EK method))** • **GARCH (1,1) (Student-t (EK method))** • **FHS (Normal)**	• **GARCH (1,1) (Student-t (EK method))**	• **MCS (Student-t) (EK and fitting methods)** • **GARCH (1,1) (Student-t (fitting method))**	• **MCS (Student-t (fitting method))**	• **None (different models under each testing criteria)**	• **None (different models under each testing criteria)**
2014-2019	• **MCS (Student-t) (EK method)** • **GARCH (1,1) (Student-t (EK and fitting method))**	• **MCS (Student-t (fitting method)** • **GARCH (1,1) (Student-t (EK and fitting methods))**	• **None (different models under each testing criteria)**	• **None (different models under each testing criteria)**	• **None (different models under each testing criteria)**	• **MCS (Student-t) (fitting method))** • **GARCH (1,1) (Student-t (fitting method))**

Table 5.3. Best performing models as per backtesting measures and distance measures per index per interval.

It can be seen in Table 5.3 that MCS and GARCH (1,1) models, both under Student-t residuals with EK and fitting methods, showed the best performance with all indexes.

Considering the 99% VaR, MCS and GARCH (1,1) models seem to be the only models showing good performance in common with all indexes. While these models seem to be the only models providing good results for SX5E index in all intervals, together with GARCH-EVT (90% and 95% thresholds) and FHS give good results for the DJIA index during the crisis.

As mentioned before, the nature of the N225 index differs from that of the DJIA and SX5E indexes and accordingly the model behavior with it is different. No models satisfy the backtesting and distance measures criteria with N225 index in the pre-crisis and post-crisis intervals, while also MCS and GARCH (1,1) models with Student-t residuals provide the best results during the crisis.

It is worth noting that fewer models seem to produce good VaR estimates at the 95% confidence level. GARCH (1,1) and MCS models with Student-t residuals (fitting method) provide the best VaR estimates for DJIA. While GARCH (1,1) model with Student-t residuals (fitting method) shows the best results for SX5E index. The same models, MCS and GARCH (1,1) with Student-t residuals (fitting method) provide good VaR estimates for the N225 index.

Performing the same analysis by periods, it is possible to affirm that in the 2002-07 interval for the 99% VaR and 95% VaR, the GARCH (1,1) under Student-t residuals with fitting model is the best model for DJIA and SX5E. And due to its nature, no models are dominant with the N225 index in this period.

Focusing on the crisis period, the models that appear to work well in common with all indexes are GARCH (1,1) model under Student-t residuals (EK method) providing the best 99% VaR estimates for the DJIA and SX5E indexes, and with the same residuals (but with the fitting method) with N225 index. Regarding the 95% VaR estimates, only MCS with Student-t residuals (fitting method) presents good behavior with the DJIA index, while there are no dominant models for SX5E or N225.

In the post-crisis period, MCS and GARCH (1,1) models with Student-t residuals provide the best 99% VaR estimates for the DJIA and SX5E indexes, while no models provide distinguished results with N225 index in this period. At the same time, for the 95% VaR estimates, no models provide dominant results for DJIA and SX5E indexes. However, MCS and GARCH (1,1) models under Student-t residuals (fitting method) provide the best results for the N225 index.

Finally, in summary for indexes, the N225 presents a particular behavior, and, in most cases, there is no model that fits well to the VaR value. However, without considering periods, it can be said that the models that best fit the N225 values are MCS and GARCH (1,1), both under Student-t residuals with fitting methods. In the case of the DJIA, it is necessary to distinguish by periods. In the pre-crisis period, the GARCH (1,1) under Student-t residuals with fitting method continues being the best model, but there is no clear model that outperforms the others during and after the crisis. For the SX5E, not even distinguishing by periods, there is no model that can be labeled as the best model.

In conclusion, there are no predominant models, but taking into account the number of times that the models have provided the best VaR estimates, the GARCH (1,1) model under Student-t residuals with fitting method and MCS, also under Student-t residuals with fitting as well as EK method are the models that should be chosen.

5.1. Limitations

The main limitations of this research are basically data related ones. In fact, choosing the crisis of 2008 as the reference crisis in this study was merely due to the fact that it was the most

recent crisis with enough data to establish three intervals upon which the comparison and analysis could be conducted. Working on other recent crises was not possible due to lack of sufficient data to consider after these two crises.
Moreover, in this thesis, we only worked with single indexes, each representing a certain market and consequently a geographical area. However, it is important to also test these models with more complex portfolios, constituting more than one index and with other indexes presenting other geographical areas.

In addition, the behavior of VaR models differs according to the financial instrument it is working with, thus applying the same VaR models to other instruments (equities, options, futures, etc.) might yield different results which is interesting to explore.

5.2. Future lines of research

The main objective of this research was to determine which is the best model among the most used VaR models and to detect if a model is better than the others depending on the economic situation, that is, if there is a crisis or not. This was applied by introducing the period of the 2008 financial crisis. An attempt was also made to detect whether one model can be better than another depending on the analyzed market, which is why three indices from Europe, Asia and America were analyzed. According to the results and to the literature on VaR models, there are many lines of research concerning VaR models accuracy and evaluating their relative performance.

In this sense, we will work on developing new testing approaches that take into account violations as well as distances. Future research will also focus on concereting the proposal of distance measures by formalizing these measures and incoprorating indicators which could be a very efficient tool for evaluating the performance of VaR models.

Future research will also aim to find a new VaR model which might be a combination of one or more models. This VaR model is supposed to be an improvement of an already existing model(s) in a way that it provides more accurate results in different markets and during different periods of time especially during crisis periods.

Thus, future resarch will involve applying the most accurate VaR models obtained in this study and their variations in different crisis, i.e., on different data, which could be different indexes and in different periods of time including other crises like the COVID-19 pandemic and the Ukranian war crisis.

These models include for instance the MCS model under different versions of GARCH family models, and also more sophisticated GARCH family models rather than the simple yet effective GARCH (1,1) model. Moreover, other trials with GARCH-EVT model would be interesting due to the positive results obtained under this model with DJIA index during the crisis period. It should be recalled that this model under both thresholds, the 90% and 95%, showed positive results according to distance measures with DJIA and SX5E at the 99% VaR level, and according to backtesting measures with N225 index with the same VaR confidence level.

A. Appendix A

The information contained in this Appendix A is based on data obtained as at 21.11.2018.

Table A.1. Journals with more than 10 citations_ **ARCH/GARCH Models**

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Journal of Banking and Finance	Alexander C. and Sheedy E., 2008; Alizadeh A.H. and Gabrielsen A., 2013; Allen L. and Bali T.G., 2007; Anand A. et al., 2016; Aramonte S. et al., 2013; Audrino F. and Barone-Adesi G., 2005; Bali T.G. et al., 2007; Bali T.G. et al., 2008; Banulescu G.-D. and Dumitrescu E.-I., 2015; Boubaker H. and Sghaier N., 2013; Chavez-Demoulin V. and McGill J.A., 2012; Consigli G., 2002; Cotter J. and Dowd K., 2006; Crouhy M. et al., 2012; Cui X. et al., 2013; Cui X. et al., 2013; Dias A., 2014; Embrechts P. et al., 2013; Escanciano J.C. and Pei P., 2012; Fermanian J.-D., 2014; Fermanian J.-D. and Scaillet O., 2005; Fortin I. and Hlouskova J., 2011; Frey R. and McNeil A.J., 2002; Giamouridis D. and Vrontos I.D., 2007; Girardi G. and Tolga Ergün A., 2013; Gordy M.B., 2002; Gordy M.B. and Marrone J., 2000; Hua J. and Manzan S., 2013; Huang W. et al., 2012; Inui K. and Kijima M., 2005; Jacobson T. et al., 2006; Jarrow R.A. and Turnbull S.M., 2000; Junker M. et al., 2006; Kadam A. and Lenk P., 2008; Kellner R. and Gatzert N., 2013; Kim Y.S. et al., 2011; Lehar A. et al., 2002; Löffler G., 2003; Longin F.M., 2000; Lönnbark C., 2013; Low R.K.Y. et al., 2013; Martens M. and Poon S.-H., 2001; Mensi W. et al., 2017; Nakazato D., 2000; Ning C. et al., 2015; Novak S.Y. and Beirlant J., 2006; O'Brien J. and Szerszeń P.J., 2017; Okimoto T., 2014; Pais A. and Stork P.A., 2011; Panigirtzoglou N. and Skiadopoulos G., 2004; Paolella M.S. and Taschini L., 2008; Pérignon C. and Smith D.R., 2010; Polanski A. et al., 2013; Pritsker M., 2006; Reboredo J.C. et al., 2016; Renault O. and Scaillet O., 2004; Rossignolo A.F. et al., 2013; Santos A.A.P. et al., 2012; Sbrana G. and Silvestrini A., 2013; Siburg K.F. et al., 2015; Tolikas K., 2014; Trindade A.A. et al., 2007; Vlaar P.J.G., 2000; Weiß G.N.F. and Scheffer M., 2015; Weiß G.N.F. and Supper H., 2013; Wied D. et al., 2016; Wong W.K., 2008; Yamai Y. and Yoshiba T., 2005; Zhu W. et al., 2016; Ziggel D. et al., 2014	2738
Journal of Empirical Finance	Bali T.G. and Neftci S.N., 2003; Becker C. and Schmidt W.M., 2013; Bee M. et al., 2016; Beltratti A. and Morana C., 1999; Berens T. et al., 2013; Billio M. and Pelizzon L., 2000; Brooks C. et al., 2005; Cerrato M. et al., 2017; Chen Y. et al., 2010; Chen Y.-T., 2012; Cheng W.-H. and Hung J.-C., 2011; Choi P. and Nam K., 2008; Chow W.W. and Fung M.K., 2008; Clements M.P. et al., 2008; Diewald L. et al., 2015; Dionne G. et al., 2009; Fries C.P. et al., 2017; Giot P. and Laurent S., 2004; Gouriéroux C. and Monfort A., 2005; Herrera R. and Schipp B., 2013;	1235

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
	Jalal A. and Rockinger M., 2008; Jansen D.W. et al., 2000; Lin C.H. et al., 2014; McNeil A.J. and Frey R., 2000; Opschoor A. et al., 2014; Wong W.K., 2010; Wu P.-T. and Shieh S.-J., 2007	
Journal of Financial Econometrics	Brownlees C.T. and Gallo G.M., 2010; Carlos Escanciano J. and Olmo J., 2011; Chen S.X., 2008; Chen S.X. and Tang C.Y., 2005; Dupuis D.J. et al., 2012; Embrechts P., 2009; Engle R.F., 2011; Ferreira M.A. and Lopez J.A., 2005; Francq C. et al., 2011; Gagliardini P. et al., 2012; Kuester K. et al., 2006; Mancini L. and Trojani F., 2011; Pelletier D. and Wei W., 2016; Santos A.A.P. et al., 2013; Taylor J.W., 2008; Žikeš F. and Baruník J., 2015	656
International Journal of Forecasting	Asai M. and McAleer M., 2008; Bams D. et al., 2017; Chen C.W.S. and So M.K.P., 2006; Chen C.W.S. et al., 2012; Ener E. et al., 2012; Fei F. et al., 2017; Fong Chan K. and Gray P., 2006; Fuertes A.-M. and Olmo J., 2013; Fuertes A.-M. et al., 2009; Gençay R. and Selçuk F., 2004; Gerlach R.H. and Abeywardana S., 2016; González-Rivera G. et al., 2004; Guermat C. and Harris R.D.F., 2002; Herrera R. and González N., 2014; Hoogerheide L.F. and van Dijk H.K., 2010; Leccadito A. et al., 2014; Lucas A. and Zhang X., 2016; Polanski A. and Stoja E., 2012; Polanski A. and Stoja E., 2017; Rubia A. and Sanchis-Marco L., 2013; Sašžiković S. and Aktan B., 2011; Takahashi M. et al., 2016; Taylor J.W. and Buizza R., 2006	589
Journal of Derivatives	Abken P.A., 2000; Alexander C.O. and Leigh C.T., 1997; Burns P. et al., 1998; Chateauneuf A. et al., 2016; Dowd K., 2001; El-Jahel L. et al., 1999; Harris R.D.F. and Shen J., 2004; Hsieh M.-H. et al., 2014; Hsieh M.-H. et al., 2014; Hsieh M.-H. et al., 2014; Hull J. and White A., 1998; Jarrow R.A., 2011; Marshall C. and Siegel M., 1997; Pérignon C. and Smith D.R., 2008; Taylor J.W., 1999; Yueh M.-L. and Wong M.C.W., 2010	388
Quantitative Finance	Albanese C. et al., 2004; Amendola A. and Candila V., 2016; Beckers B. et al., 2017; Bellini F. and Figá-Talamanca G., 2007; Bingham N.H. et al., 2003; Bormetti G. et al., 2012; Chang C.-C. and Tsao C.-Y., 2011; Chen C.W.S. et al., 2011; Chopping T., 2014; Choroś-Tomczyk B. et al., 2014; Chou H.-C. and Wang D.K., 2014; Cochran S.J. et al., 2016; Corlu C.G. and Corlu A., 2015; Dupacová J. and Polívka J., 2007; Elliott R.J. and Miao M., 2009; Ergen I., 2015; Ewald C.-O. and Zhang A., 2006; Fernandez V., 2011; Giacometti R. et al., 2007; Guan L.K. and Xiaoqing L., 2004; Guégan D. and Zhao X., 2014; Hauksson H.A. et al., 2001; Hellmich M. and Kassberger S., 2011; Huber S., 2010; Kawata R. and Kijima M., 2007; Knight J. et al., 2003; Lejeune M.A., 2011; Lönnbark C., 2016; Luo X. and Shevchenko	375

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
	P.V., 2010; Lux T. and Morales-Arias L., 2013; Malevergne Y. and Sornette D., 2004; Masdemont J.J. and Ortiz-Gracia L., 2014; Packham N. et al., 2017; Penzer J. et al., 2012; Rombouts J.V.K. and Verbeek M., 2009; Scheffer M. and Weiß G.N.F., 2017; Siller T., 2013; Silva Filho O.C. et al., 2014; Silvapulle P. and Granger C.W.J., 2001; So M.K.P. and Wong C.-M., 2012; Su J.-B., 2014; Tasche D., 2009; Yao H. et al., 2015; Zumbach G., 2011	
Management Science	Berkowitz J. et al., 2011; Du Z. and Escanciano J.C., 2017; Fu M.C. et al., 2009; Glasserman P. et al., 2000; Goh J. and Hall N.G., 2013; Hong L.J. and Liu G., 2009; Jin X. and Zhang A.X., 2006; Taylor J.W., 2005; Zymler S. et al., 2013	358
International Review of Financial Analysis	Alexander C. et al., 2013; Assaf A., 2009; Berger T. and Missong M., 2014; Bhattacharyya M. and Ritolia G., 2008; Byström H.N.E., 2004; Chou P.-H. et al., 2006; Chrétien S. and Coggins F., 2010; Degiannakis S. and Potamia A., 2017; Degiannakis S. et al., 2013; Diamandis P.F. et al., 2011; Fernandez V., 2005; Fernandez V., 2006; Karmakar M. and Paul S., 2016; Kavussanos M.G. and Dimitrakopoulos D.N., 2011; Liu L., 2014; Masih M. et al., 2010; McMillan D.G. and Kambouroudis D., 2009; Moosa I.A. and Bollen B., 2002; Reber B., 2017; Righi M.B. and Ceretta P.S., 2013; Stavroyiannis S. et al., 2012; Zmeškal Z., 2005	350
ASTIN Bulletin	Afonso L.B. and Corte Real P., 2016; Alink S. et al., 2005; Cai J. and Tan K.S., 2007; Chuliá H. et al., 2015; De Alba E. et al., 2010; Degen M. et al., 2007; Frees E.W. et al., 2009; Hürlimann W., 2014; Kim J.H.T. and Hardy M.R., 2007; Kim J.H.T. and Hardy M.R., 2009; Shi P. and Frees E.W., 2011; Wüthrich M.V., 2003	294
Economic Modelling	Aloui C. and Hkiri B., 2014; Araichi S. et al., 2016; Berger T., 2016; Chang K.-L., 2010; Charfeddine L., 2016; Chen R. and Yu L., 2013; Chen X.B. et al., 2014; Chen Y.-C. et al., 2013; Ghorbel A. and Trabelsi A., 2014; Iglesias E.M., 2015; Li T. et al., 2013; Lourme A. and Maurer F., 2017; Louzis D.P. et al., 2014; Mbairadjim Moussa A. et al., 2014; Mudakkar S.R. et al., 2013; Ormos M. and Timotity D., 2016; Ourir A. and Snoussi W., 2012; Righi M.B. and Ceretta P.S., 2013; Schäfer R. and Koivusalo A.F.R., 2013; Su J.-B., 2015; Su J.-B. and Hung J.-C., 2011; Zhang H.-G. et al., 2017	211
Journal of International Money and Finance	Bams D. et al., 2005; Chang K.-H. and Kim M.-J., 2001; Cotter J., 2007; Martens M., 2001; Polanski A. and Stoja E., 2014; Pownall R.A.J. and Koedijk K.G., 1999; Reboredo J.C. and Ugolini A., 2015; Sirr G. et al., 2011	174
Journal of Risk Finance	Abdul Manap T.A. and Kassim S.H., 2011; Angelidis T. and Degiannakis S., 2005; Dowd K., 2000; Dowd K. et al., 2004; Gubareva M. and Borges M.R., 2017; Guo H. et al., 2007; Harmantzis F.C. et al., 2006; Kinateder H. and Wagner N., 2014; Lechner L.A. and Ovaert T.C.,	159

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
	2010; Lhabitant F.-S., 2001; Lin C.H. and Shen S.S., 2006; Mouatassim Y., 2012; Mozumder S. et al., 2017; Nguyen N.Q.A. and Nguyen T.N.T., 2017; Ozun A. et al., 2010; Sharma B. et al., 2015; Tilman L.M. and Brusilovskiy P., 2001; Uryasev S. and Theiler U.A., 2010; Wang Z., 2002; Yi-Hou Huang A. and Tseng T.-W., 2009; Zheng H. and Shen Y., 2008	
Review of Quantitative Finance and Accounting	Angelidis T. et al., 2007; Baixauli-Soler J.S. and Alvarez S., 2006; Hsu C.-P. et al., 2012; Huang A.Y., 2013; Huang Y.C. and Lin B.-J., 2004; Lee C.-F. and Su J.-B., 2012; Weiß G.N.F., 2013	149
Applied Financial Economics	Aggarwal R. and Qi M., 2009; Asai M. and Brugal I., 2012; Bekiros S.D. and Georgoutsos D.A., 2008; Chen Q. et al., 2012; Chiu C.-L. et al., 2005; Cotter J., 2004; Deacle S. and Elyasiani E., 2014; Dunis C.L. et al., 2010; Fernandes J.L.B. et al., 2008; Gargallo P. et al., 2010; Goncu A. et al., 2012; Huang A.Y., 2009; Hung J.-C. et al., 2006; Iglesias E.M. and Lagoa Varela M.D., 2012; Jaggia S. and Kelly-Hawke A., 2009; Papadamou S. and Stephanides G., 2004; Plunus S. et al., 2012; Reber B., 2013; Ren F. and Giles D.E., 2010; Sadorsky P., 2005; Sjölander P., 2009; Su Y.-C. et al., 2011; Vesper A., 2012; Wong W.K., 2009; Xu Q. and Childs T., 2013; Xu Q. and Li X.-M., 2009; Zhao X. et al., 2010; Zhou J., 2012	143
Journal of International Financial Markets, Institutions and Money	Aloui C. et al., 2015; Angelidis T. and Degiannakis S., 2008; Bekiros S.D. and Georgoutsos D.A., 2005; Riedel C. and Wagner N., 2015; Sarafrazi S. et al., 2014; Slim S. et al., 2017; So M.K.P. and Yu P.L.H., 2006	142
Econometric Reviews	Ferraty F. and Quintela-Del-Río A., 2016; Manner H. and Reznikova O., 2012; Martins-Filho C. et al., 2015; McAleer M. et al., 2009; Nakajima J., 2017; Yu P.L.H. et al., 2010	136
Financial Analysts Journal	Linsmeier T.J. and Pearson N.D., 2000; Xiong J.X. and Idzorek T.M., 2011	135
Journal of Futures Markets	Cao Z. et al., 2010; Conlon T. and Cotter J., 2012; Cotter J. and Dowd K., 2010; Giot P., 2003; Kambouroudis D.S. et al., 2016; Miller D.J. and Liu W.-H., 2006; Park S.Y. and Jei S.Y., 2010; Samuel Y.M.Z.-T., 2008	132
North American Journal of Economics and Finance	Allen D.E. et al., 2013; Aloui C. and Hamida H.B., 2014; de Jesús R. et al., 2013; Araújo Santos P. et al., 2013; Asai M. and Brugal I., 2013; Demiralay S. and Ulusoy V., 2014; Gong X. and Zhuang X., 2017; Haas M. et al., 2013; Hammoudeh S. et al., 2013; Herrera R. and Schipp B., 2014; Li L., 2017; Ñíguez T.-M. and Perote J., 2017; Ortas E. et al., 2015; Reboredo J.C. and Ugolini A., 2015; Rodríguez G., 2017; Su J.-B., 2014; Teply P. and Kvapilikova I., 2017	130

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Annals of Operations Research	AitSahlia F. et al., 2011; Al Janabi M.A.M., 2013; Alonso E. et al., 2015; Bilbao-Terol A. et al., 2016; De Prisco B. et al., 2007; Iaquinta G. et al., 2009; Krzemienowski A. and Szymczyk S., 2016; Lee J. and Prékopa A., 2013; Liu X. et al., 2017; Lu X.F. et al., 2014; Moazeni S. et al., 2016; Prékopa A., 2012; Stoyanov S.V. et al., 2010; Stoyanov S.V. et al., 2013	110
Economic Notes	Alexander C., 2002; Antonelli S. and Iovino M.G., 2002; Barone-Adesi G. and Giannopoulos K., 2001; Cherubini U. and Della Lunga G., 2001; Cherubini U. and Luciano E., 2001; Di Clemente A., 2014; Di Clemente A., 2015; Di Clemente A. and Romano C., 2004	110
European Journal of Finance	Andriosopoulos K. and Nomikos N., 2015; Bernardi M. et al., 2017; Cheng J. et al., 2016; Coroneo L. and Veredas D., 2012; Dias A. and Embrechts P., 2009; Giot P., 2005; Goodworth T.R.J. and Jones C.M., 2007; Liu R. and Lux T., 2015; Ma L. and Pohlman L., 2008; Sermpinis G. et al., 2015; Tolikas K. et al., 2007	106
European Financial Management	Barone-Adesi G. et al., 2002; Hallerbach W.G. and Menkveld A.J., 2004; Liang B. and Park H., 2007; Peterson S. and Stapleton R.C., 1999; Sun W. et al., 2009	105
International Review of Economics and Finance	Asai M. et al., 2015; Chang G.-D. and Chen C.-S., 2014; Chen Y.-H. and Tu A.H., 2013; Hammoudeh S. and McAleer M., 2015; Karmakar M. and Shukla G.K., 2014; Kittiakarasakun J. and Tse Y., 2011; Krehbiel T. and Adkins L.C., 2008; Liu H.C. et al., 2012; Liu Q. and Tse Y., 2017; Mensi W. et al., 2016; Nasr A.B. et al., 2016; Su J.-B. et al., 2014	104
Journal of Econometrics	Cai Z. and Wang X., 2008; Chan N.H. et al., 2007; Francq C. and Zakoïan J.-M., 2015; Macaro C., 2010; Wang C.-S. and Zhao Z., 2016	101
International Journal of Theoretical and Applied Finance	Abad P. and Benito S., 2009; Al-Zoubi H.A. and Maghyereh A.I., 2007; Ané T., 2006; Angelidis T. and Skiadopoulos G., 2008; Ararat C. et al., 2017; Benth F.E. et al., 2006; D'Addona S. and Ciprian M., 2007; Grandits P. et al., 2010; Han C.-H. et al., 2014; Kalyvas L. and Sfetsos A., 2006; Kamdem J.S., 2005; Marinelli C. et al., 2007; Mendes B.V.M. and De Melo E.F.L., 2010; Rutkowski M. and Tarca S., 2015	99
Journal of Business	Bali T.G., 2003	96
Quarterly Review of Economics and Finance	Dimitrakopoulos D.N. et al., 2010; Ergün A.T. and Jun J., 2010; Hammoudeh S. et al., 2011; Karmakar M., 2017; Mabrouk S. and Saadi S., 2012; Sabbaghi O. and Sabbaghi N., 2011	89
Journal of Economic Surveys	McAleer M., 2009; Rocco M., 2014; Ruiz E. and Pascual L., 2002	84
Pacific Basin Finance Journal	Ho L.-C. et al., 2000; Liao Y., 2013; Liu Q. and An Y., 2014; Liu Q. et al., 2014; Mwamba J.M. et al., 2017; Weng H. and Trück S., 2011	84

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Econometrics Journal	Chen X. et al., 2009; Grigoletto M. and Lisi F., 2009; Hoga Y., 2017; Kheifets I.L., 2015; Preminger A. and Storti G., 2017; Wilhelmsson A., 2009	78
Review of Financial Studies	Adrian T. and Shin H.S., 2014	75
Applied Economics Letters	Agiakloglou C. and Bloutsos K., 2011; Allen D.E. et al., 2012; Ardia D. and Hoogerheide L.F., 2013; Chan S. and Nadarajah S., 2015; Choi P. and Min I., 2013; Cotter J., 2005; Diao X. and Tong B., 2015; Hsu Ku Y.-H. and Wang J.J., 2008; Li L. and Lin H.-W.W., 2004; Lin B. and Lin C.N.M. (Lin Chenmiao), 2017; Marzo M. and Zagaglia P., 2010; Park S.-K. et al., 2017; Toque C. and Terraza V., 2014	74
Journal of Financial Economics	Gupta A. and Liang B., 2005	55
Journal of Risk and Insurance	Bali T.G. and Theodossiou P., 2008; Chavez-Demoulin V. et al., 2016	55
Emerging Markets Review	Del Brio E.B. et al., 2014; Kim J.S. and Ryu D., 2015; Mensi W. et al., 2017; Wong A.Y.-T. and Fong T.P.W., 2011	47
Journal of Property Investment and Finance	Liow K.H., 2008; Rong N. and Trück S., 2010; Tajani F. and Morano P., 2017	44
Applied Economics	Blazsek S. and Monteros L.A., 2017; Bollen B., 2015; Chang C.-C. et al., 2015; Chang K.-L., 2011; Chuang C.-C. et al., 2015; Ergen I., 2014; Goode J. et al., 2015; Huang A.Y., 2015; Iglesias E.M., 2012; Liu W.-H., 2014; Lynch M.Á. and Curtis J., 2016; Mi Z. et al., 2017; Shim J. and Lee E.-J., 2016; Sukcharoen K. and Choi H., 2015; Totić S. and Božović M., 2016; Tursunalieva A. and Silvapulle P., 2014; Zhou C. et al., 2016	43
Journal of Operational Risk	Badescu A.L. et al., 2015; Feria-Dominguez J.M. and Jimenez-Rodriguez E., 2017; Figini S. et al., 2015; Guégan D. et al., 2012; Hernández L. et al., 2013; Hess C., 2011; Jiménez J.A. and Arunachalam V., 2016; Li J. et al., 2014; Opdyke J.D., 2017	39
Journal of Risk	Abad P. et al., 2016; Amendola A. and Candila V., 2017; Ardia D. et al., 2017; Auer B.R., 2015; Bellalah M. and Chayeh Z., 2015; Berger T. and Missong M., 2014; Bignozzi V. and Tsanakas A., 2016; Braun V. and Hackethal A., 2013; Buchner A., 2017; Cameron J. et al., 2016; Drapeau S. et al., 2014; Gaisser S. et al., 2011; Hong K., 2017; Izhar H., 2015; Jakobsons E., 2016; Jiménez J.A. and Arunachalam V., 2011; Kellner R. et al., 2016; Kijima M. et al., 2014; Kolman M., 2014; Lau C., 2015; Lee W.-C., 2012; Liu X., 2015; Maciag J. et al., 2016; Martin R.D. and Arora R., 2017; Marumo K. and Wolff R.C., 2016; Meucci A. et al., 2011;	39

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
	Muromachi Y., 2015; Nossman M. and Vilhelmsson A., 2014; Raunig B. and Scheicher M., 2011; Simonato J.-G., 2013; Yu J.-S., 2013	
Journal of Financial Services Research	Jacobson T. et al., 2005; Kupiec P. and Güntay L., 2016	35
Insurance: Mathematics and Economics	Cousin A. and Di Bernardino E., 2014; Ibragimov R. and Walden J., 2008; Torres R. et al., 2015	34
Review of Financial Economics	Huang A.Y., 2010; Karagiannidis I. and Sykes Wilford D., 2015; Karmakar M., 2013; Lu C. et al., 2009; Pearson N.D. and Smithson C., 2002	32
Finance Research Letters	Bao Y. and Ullah A., 2004; Belhajjam A. et al., 2017; González-Rivera G. et al., 2007; Han Y. et al., 2017; Haugom E. et al., 2016; Klein T. and Walther T., 2017; Lehnert T. and Wolff C.C.P., 2004; Siven J.V. et al., 2009; Wang X. et al., 2017	31
Asia-Pacific Financial Markets	Bahlous M., 2013; Ignatieva K. and Platen E., 2010; Ishitani K. and Sato K., 2013; Lin S.-K. et al., 2006; Miura R. and Oue S., 2000; Nagahara Y., 2008; So M.K.P. and Xu R., 2013; Tsuji C., 2003	26
Review of Economic Studies	Chernozhukov V. and Fernández-Val I., 2011	26
Emerging Markets Finance and Trade	Atilgan Y. and Ozgur Demirtas K., 2016; Gencer H.G. and Demiralay S., 2016; Su E. and Knowles T.W., 2006	25
Economics Letters	Ardia D. and Hoogerheide L.F., 2014; Herwartz H., 2009; Herwartz H. and Raters F.H.C., 2015; Ibragimov R. and Prokhorov A., 2016; Krämer W. and Wied D., 2015; Yi Y. et al., 2014	24
Japanese Economic Review	Watanabe T., 2012	24
Journal of Economics and Business	Bauer C., 2000	24
Spanish Review of Financial Economics	Abad P. et al., 2014	24
Journal of Financial Stability	Puzanova N. et al., 2009; Rossignolo A.F. et al., 2012	23
Journal of Money,	Bali T.G., 2007	23

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Credit and Banking		
Managerial Finance	Angelidis T. and Degiannakis S., 2008; Basterfield D. et al., 2010; Degiannakis S. et al., 2012; Liu H.C. and Hung J.-C., 2010; Obi P. and Sil S., 2013; Poon S.-H. and Lin H., 2001; Sabbaghi O., 2011; Uppal J.Y. and Ullah Mangla I., 2013	23
Accounting and Finance	da Veiga B. et al., 2012; Hatherley A. and Alcock J., 2007	22
International Journal of Managerial Finance	de Melo Mendes B.V. and Aíube C., 2011; Maghyereh A.I. and Al-Zoubi H.A., 2006	21
Journal of Credit Risk	Choe G.H. and Kwon S.W., 2014; Ebert S. and Lütkebohmert E., 2011; Fok P.-W. et al., 2014; Hauptmann J. et al., 2014; Huang H. et al., 2015; Huang X. and Oosterlee C.W., 2011; Maciag J. and Löderbusch M., 2017; Owen A.W. et al., 2015; Schmitt T.A. et al., 2015; Schmitt T.A. et al., 2015	21
Applied Financial Economics Letters	Lee M.-C. et al., 2006; Lee M.-C. et al., 2008	20
International Journal of Monetary Economics and Finance	Carlos Escanciano J. and Mayoral S., 2008; Dockery E. and Efentakis M., 2008; Ghorbel A. and Trabelsi A., 2008; Mabrouk S. and Aloui C., 2011; Olmo J., 2008	20
International Economics and Economic Policy	Elsinger H. et al., 2006; Stolbov M., 2017	19
Journal of Business Economics and Management	Atta Mills E.F.E. et al., 2017; Chen D.-H. et al., 2014; Nam D., 2013; Radivojević N. et al., 2017; Suhobokov A., 2007	18
Research in International Business and Finance	Degiannakis S. and Floros C., 2013; Kurita T., 2014; Trabelsi N. and Naifar N., 2017; Walther T. et al., 2017	16
Economic Record	Gronwald M. et al., 2011	14
Global Finance Journal	Di J. and Zhu P., 2015; Shao X.-D. et al., 2009	13
International Journal of Finance and Economics	Gettinby G.D. et al., 2006; Stoyanov S.V. et al., 2017	13

ARCH/GARCH Models		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
International Research Journal of Finance and Economics	Cheng H.-W. and Lin C.-Y., 2011; Deb S.G. and Banerjee A., 2009; Huang S.-C. et al., 2011; Lee M.-C. et al., 2010; Lifschutz S., 2010; Otero L. et al., 2012; Rahman M.M. et al., 2011; Su Y. et al., 2011; Sung C.-H. and Kuo C.-J., 2010; Talebnia G.O. et al., 2011	13
Zbornik Radova Ekonomskog Fakultet au Rijeci	Žiković S., 2011; Žiković S. and Aktan B., 2009	13
Journal of Asian Economics	Vinod H.D., 2003	11
Journal of Multinational Financial Management	Assaf A., 2015; McMillan D.G. et al., 2008; Mokni K. and Mansouri F., 2017	11

Table A.2. Journals with more than 10 citations _ **Extreme Value Theory**

Extreme Value Theory		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Journal of Empirical Finance	Bali T.G. and Neftci S.N., 2003; Bee M. et al., 2016; Brooks C. et al., 2005; Herrera R. and Schipp B., 2013; Jalal A. and Rockinger M., 2008; Jansen D.W. et al., 2000; Lin C.H. et al., 2014; McNeil A.J. and Frey R., 2000	744
Journal of Banking and Finance	Allen L. and Bali T.G., 2007; Consigli G., 2002; Cotter J. and Dowd K., 2006; Dias A., 2014; Huang W. et al., 2012; Kellner R. and Gatzert N., 2013; Longin F.M., 2000; Novak S.Y. and Beirlant J., 2006; Pais A. and Stork P.A., 2011; Rossignolo A.F. et al., 2013; Tolikas K., 2014; Yamai Y. and Yoshiba T., 2005	518
International Journal of Forecasting	Fong Chan K. and Gray P., 2006; Gençay R. and Selçuk F., 2004; Herrera R. and González N., 2014; Sašžiković S. and Aktan B., 2011	259
Journal of Financial Econometrics	Embrechts P., 2009; Kuester K. et al., 2006; Mancini L. and Trojani F., 2011	254
International Review of Financial Analysis	Assaf A., 2009; Berger T. and Missong M., 2014; Bhattacharyya M. and Ritolia G., 2008; Byström H.N.E., 2004; Fernandez V., 2005; Karmakar M. and Paul S., 2016; Liu L., 2014; Reber B., 2017	141
ASTIN Bulletin	Alink S. et al., 2005; De Alba E. et al., 2010; Degen M. et al., 2007; Wüthrich M.V., 2003	106
Journal of Business	Bali T.G., 2003	97
Review of Financial Studies	Adrian T. and Shin H.S., 2014	75

Extreme Value Theory		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Quantitative Finance	Bingham N.H. et al., 2003; Chopping T., 2014; Chou H.-C. and Wang D.K., 2014; Ergen I., 2015; Guégan D. and Zhao X., 2014; Hauksson H.A. et al., 2001; Huber S., 2010; Knight J. et al., 2003; Packham N. et al., 2017	66
Journal of Risk and Insurance	Bali T.G. and Theodossiou P., 2008; Chavez-Demoulin V. et al., 2016	65
Applied Financial Economics	Aggarwal R. and Qi M., 2009; Bekiros S.D. and Georgoutsos D.A., 2008; Chen Q. et al., 2012; Cotter J., 2004; Dunis C.L. et al., 2010; Goncu A. et al., 2012; Iglesias E.M. and Lagoa Varela M.D., 2012; Plunus S. et al., 2012; Ren F. and Giles D.E., 2010; Zhao X. et al., 2010; Zhou J., 2012	61
Journal of International Money and Finance	Cotter J., 2007; Pownall R.A.J. and Koedijk K.G., 1999	56
Journal of Financial Economics	Gupta A. and Liang B., 2005	55
Pacific Basin Finance Journal	Ho L.-C. et al., 2000; Mwamba J.M. et al., 2017	50
Economic Modelling	Araichi S. et al., 2016; Ghorbel A. and Trabelsi A., 2014; Iglesias E.M., 2015; Louzis D.P. et al., 2014; Mudakkar S.R. et al., 2013; Ourir A. and Snoussi W., 2012	49
Journal of Risk Finance	Dowd K., 2001; Harmantzis F.C. et al., 2006; Lechner L.A. and Ovaert T.C., 2010; Lin C.H. and Shen S.S., 2006; Mouatassim Y., 2012; Mozumder S. et al., 2017; Ozun A. et al., 2010	49
Review of Quantitative Finance and Accounting	Angelidis T. et al., 2007; Hsu C.-P. et al., 2012	43
North American Journal of Economics and Finance	Allen D.E. et al., 2013; Araújo Santos P. et al., 2013; de Jesús R. et al., 2013; Herrera R. and Schipp B., 2014	41
International Review of Economics and Finance	Chang G.-D. and Chen C.-S., 2014; Karmakar M. and Shukla G.K., 2014; Kittiakarasakun J. and Tse Y., 2011; Krehbiel T. and Adkins L.C., 2008	36
Journal of International Financial Markets, Institutions and Money	Bekiros S.D. and Georgoutsos D.A., 2005; Sarafrazi S. et al., 2014	36
Quarterly Review of Economics and Finance	Dimitrakopoulos D.N. et al., 2010; Ergün A.T. and Jun J., 2010; Karmakar M., 2017	30
Journal of Economic Surveys	Rocco M., 2014	26
Review of Economic Studies	Chernozhukov V. and Fernández-Val I., 2011	26
Spanish Review of Financial Economics	Abad P. et al., 2014	24

Extreme Value Theory		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Journal of Money, Credit and Banking	Bali T.G., 2007	23
International Journal of Managerial Finance	de Melo Mendes B.V. and Aíube C., 2011; Maghyereh A.I. and Al-Zoubi H.A., 2006	22
Journal of Financial Stability	Rossignolo A.F. et al., 2012	22
Applied Economics	Ergen I., 2014; Iglesias E.M., 2012; Mi Z. et al., 2017; Totić S. and Božović M., 2016	20
International Journal of Theoretical and Applied Finance	Ané T., 2006; Kalyvas L. and Sfetsos A., 2006; Marinelli C. et al., 2007	18
International Journal of Monetary Economics and Finance	Ghorbel A. and Trabelsi A., 2008; Olmo J., 2008	15
Journal of Property Investment and Finance	Liow K.H., 2008	15
Journal of Operational Risk	Figini S. et al., 2015; Guégan D. et al., 2012; Jiménez J.A. and Arunachalam V., 2016	14
International Journal of Finance and Economics	Gettinby G.D. et al., 2006; Stoyanov S.V. et al., 2017	13
Review of Financial Economics	Karmakar M., 2013; Pearson N.D. and Smithson C., 2002	13
Zbornik Radova Ekonomskog Fakultet au Rijeci	Žiković S., 2011; Žiković S. and Aktan B., 2009	13
European Journal of Finance	Cheng J. et al., 2016; Sermpinis G. et al., 2015; Tolikas K. et al., 2007	11

Table A.3. Journals with more than 10 citations _ **Monte Carlo Simulation**

Monte Carlo Simulation		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Journal of Empirical Finance	Chow W.W. and Fung M.K., 2008; Diewald L. et al., 2015; Dionne G. et al., 2009; Jansen D.W. et al., 2000; McNeil A.J. and Frey R., 2000; Wong W.K., 2010	688
Management Science	Berkowitz J. et al., 2011; Du Z. and Escanciano J.C., 2017; Fu M.C. et al., 2009; Glasserman P. et al., 2000; Goh J. and Hall N.G., 2013; Hong L.J. and Liu G., 2009; Jin X. and Zhang A.X., 2006	304

Monte Carlo Simulation		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Journal of Banking and Finance	Audrino F. and Barone-Adesi G., 2005; Dias A., 2014; Escanciano J.C. and Pei P., 2012; Gordy M.B., 2002; Hua J. and Manzan S., 2013; Jacobson T. et al., 2006; Löffler G., 2003; Nakazato D., 2000; Panigirtzoglou N. and Skiadopoulos G., 2004; Renault O. and Scaillet O., 2004; Vlaar P.J.G., 2000; Wong W.K., 2008; Ziggel D. et al., 2014	262
International Journal of Forecasting	Asai M. and McAleer M., 2008; Chen C.W.S. et al., 2012; Chen C.W.S. and So M.K.P., 2006; Ener E. et al., 2012; Leccadito A. et al., 2014; Polanski A. and Stoja E., 2012; Takahashi M. et al., 2016	127
Quantitative Finance	Amendola A. and Candila V., 2016; Bingham N.H. et al., 2003; Bormetti G. et al., 2012; Chang C.-C. and Tsao C.-Y., 2011; Corlu C.G. and Corlu A., 2015; Ewald C.-O. and Zhang A., 2006; Fernandez V., 2011; Guan L.K. and Xiaoqing L., 2004; Lönnbark C., 2016; Lux T. and Morales-Arias L., 2013; Masdemont J.J. and Ortiz-Gracia L., 2014; Rombouts J.V.K. and Verbeek M., 2009; Siller T., 2013; Su J.-B., 2014; Tasche D., 2009	124
Financial Analysts Journal	Linsmeier T.J. and Pearson N.D., 2000	116
Econometric Reviews	Manner H. and Reznikova O., 2012; Martins-Filho C. et al., 2015	61
Journal of Financial Economics	Gupta A. and Liang B., 2005	55
Journal of Financial Econometrics	Carlos Escanciano J. and Olmo J., 2011; Mancini L. and Trojani F., 2011; Pelletier D. and Wei W., 2016; Žikeš F. and Baruník J., 2015	44
Review of Quantitative Finance and Accounting	Angelidis T. et al., 2007; Hsu C.-P. et al., 2012	43
Annals of Operations Research	De Prisco B. et al., 2007; Lu X.F. et al., 2014; Moazeni S. et al., 2016; Stoyanov S.V. et al., 2010	37
Quarterly Review of Economics and Finance	Hammoudeh S. et al., 2011	36
Journal of Financial Services Research	Jacobson T. et al., 2005	31
Journal of Risk Finance	Angelidis T. and Degiannakis S., 2005; Sharma B. et al., 2015; Wang Z., 2002; Zheng H. and Shen Y., 2008	29
Journal of Economic Surveys	McAleer M., 2009	25
Pacific Basin Finance Journal	Liao Y., 2013; Liu Q. et al., 2014	24
Economic Modelling	Chen R. and Yu L., 2013; Chen X.B. et al., 2014; Lourme A. and Maurer F., 2017;	22

Monte Carlo Simulation		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
	Louzis D.P. et al., 2014; Schäfer R. and Koivusalo A.F.R., 2013	
ASTIN Bulletin	De Alba E. et al., 2010; Kim J.H.T. and Hardy M.R., 2007	20
Economic Notes	Antonelli S. and Iovino M.G., 2002; Di Clemente A., 2014; Di Clemente A., 2015; Di Clemente A. and Romano C., 2004	20
International Journal of Theoretical and Applied Finance	Benth F.E. et al., 2006; Marinelli C. et al., 2007	20
Journal of Derivatives	Abken P.A., 2000; Chateauneuf A. et al., 2016; Hsieh M.-H. et al., 2014; Yueh M.-L. and Wong M.C.W., 2010	17
Applied Economics	Blazsek S. and Monteros L.A., 2017; Lynch M.Á. and Curtis J., 2016; Mi Z. et al., 2017; Tursunalieva A. and Silvapulle P., 2014	15
Economics Letters	Herwartz H., 2009; Krämer W. and Wied D., 2015; Yi Y. et al., 2014	13
Journal of Business Economics and Management	Suhobokov A., 2007	12
Journal of Econometrics	Francq C. and Zakoïan J.-M., 2015; Wang C.-S. and Zhao Z., 2016	12
Journal of Futures Markets	Cao Z. et al., 2010	12

Table A.4. Journals with more than 10 citations _ **Historical Simulation**

Historical Simulation		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Journal of Banking and Finance	Aramonte S. et al., 2013; Chavez-Demoulin V. and McGill J.A., 2012; Escanciano J.C. and Pei P., 2012; Fermanian J.-D. and Scaillet O., 2005; Inui K. and Kijima M., 2005; Jacobson T. et al., 2006; Pérignon C. and Smith D.R., 2010; Polanski A. et al., 2013; Pritsker M., 2006; Renault O. and Scaillet O., 2004; Siburg K.F. et al., 2015; Vlaar P.J.G., 2000	454
Journal of Financial Econometrics	Chen S.X., 2008; Chen S.X. and Tang C.Y., 2005; Dupuis D.J. et al., 2012; Kuester K. et al., 2006; Mancini L. and Trojani F., 2011	357
International Journal of Forecasting	Asai M. and McAleer M., 2008; Bams D. et al., 2017; Ener E. et al., 2012; Fong Chan K. and Gray P., 2006; Fuertes A.-M. et al., 2009; Fuertes A.-M. and Olmo J., 2013; Gençay R. and Selçuk F., 2004	322
Financial Analysts Journal	Linsmeier T.J. and Pearson N.D., 2000	116

Historical Simulation		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
European Financial Management	Barone-Adesi G. et al., 2002; Liang B. and Park H., 2007; Sun W. et al., 2009	98
Quantitative Finance	Beckers B. et al., 2017; Bingham N.H. et al., 2003; Kawata R. and Kijima M., 2007; Scheffer M. and Weiß G.N.F., 2017; Silvapulle P. and Granger C.W.J., 2001; Su J.-B., 2014; Yao H. et al., 2015	86
Journal of Empirical Finance	Brooks C. et al., 2005; Cheng W.-H. and Hung J.-C., 2011; Fries C.P. et al., 2017; Gouriéroux C. and Monfort A., 2005	84
Journal of Derivatives	Pérignon C. and Smith D.R., 2008; Taylor J.W., 1999	62
Quarterly Review of Economics and Finance	Dimitrakopoulos D.N. et al., 2010; Hammoudeh S. et al., 2011	61
Economic Modelling	Chen Y.-C. et al., 2013; Louzis D.P. et al., 2014; Ormos M. and Timotity D., 2016; Ourir A. and Snoussi W., 2012; Su J.-B., 2015; Zhang H.-G. et al., 2017	60
Pacific Basin Finance Journal	Ho L.-C. et al., 2000; Weng H. and Trück S., 2011	52
Journal of Econometrics	Cai Z. and Wang X., 2008; Macaro C., 2010; Wang C.-S. and Zhao Z., 2016	47
Journal of Risk Finance	Angelidis T. and Degiannakis S., 2005; Gubareva M. and Borges M.R., 2017; Lhabitant F.-S., 2001; Yi-Hou Huang A. and Tseng T.-W., 2009	47
North American Journal of Economics and Finance	Hammoudeh S. et al., 2013; Ñíguez T.-M. and Perote J., 2017; Su J.-B., 2014	43
Review of Quantitative Finance and Accounting	Angelidis T. et al., 2007; Baixauli-Soler J.S. and Alvarez S., 2006; Huang A.Y., 2013	42
Journal of Financial Services Research	Jacobson T. et al., 2005; Kupiec P. and Güntay L., 2016	35
International Journal of Theoretical and Applied Finance	Angelidis T. and Skiadopoulos G., 2008; Han C.-H. et al., 2014; Kalyvas L. and Sfetsos A., 2006	32
Economic Notes	Barone-Adesi G. and Giannopoulos K., 2001	25
Journal of Economic Surveys	McAleer M., 2009	25
Spanish Review of Financial Economics	Abad P. et al., 2014	25
Applied Financial Economics	Asai M. and Brugal I., 2012; Deacle S. and Elyasiani E., 2014; Huang A.Y., 2009; Papadamou S. and Stephanides G., 2004; Zhou J., 2012	21
Journal of Futures Markets	Cao Z. et al., 2010; Cotter J. and Dowd K., 2010	19

Historical Simulation		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
International Journal of Managerial Finance	Maghyereh A.I. and Al-Zoubi H.A., 2006	13
International Review of Financial Analysis	Chou P.-H. et al., 2006; Chrétien S. and Coggins F., 2010; Kavussanos M.G. and Dimitrakopoulos D.N., 2011	13
Review of Financial Economics	Lu C. et al., 2009	13
Zbornik Radova Ekonomskog Fakultet au Rijeci	Žiković S., 2011; Žiković S. and Aktan B., 2009	13

Table A.5. Journals with more than 10 citations _ **Variance-covariance**

Variance-covariance		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
International Journal of Forecasting	Bams D. et al., 2017; Ener E. et al., 2012; Fong Chan K. and Gray P., 2006; Fuertes A.-M. and Olmo J., 2013; Gençay R. and Selçuk F., 2004; Guermat C. and Harris R.D.F., 2002	301
Journal of Empirical Finance	Billio M. and Pelizzon L., 2000; Brooks C. et al., 2005; Jansen D.W. et al., 2000	165
Quantitative Finance	Beckers B. et al., 2017; Bingham N.H. et al., 2003; Bormetti G. et al., 2012; Chopping T., 2014; Dupacová J. and Polívka J., 2007; Rombouts J.V.K. and Verbeek M., 2009; Scheffer M. and Weiß G.N.F., 2017; Silva Filho O.C. et al., 2014; Silvapulle P. and Granger C.W.J., 2001; Su J.-B., 2014; Yao H., Li Y. and Benson K., 2015	118
Financial Analysts Journal	Linsmeier T.J. and Pearson N.D., 2000	110
Journal of Banking and Finance	Cui X. et al., 2013; Hua J. and Manzan S., 2013; Jacobson T. et al., 2006; Polanski A. et al., 2013; Siburg K.F. et al., 2015; Vlaar P.J.G., 2000; Weiß G.N.F. and Scheffer M., 2015	105
ASTIN Bulletin	Afonso L.B. and Corte Real P., 2016; Degen M. et al., 2007; Shi P. and Frees E.W., 2011	89
International Review of Financial Analysis	Assaf A., 2009; Chou P.-H. et al., 2006; Diamandis P.F. et al., 2011; Moosa I.A. and Bollen B., 2002; Zmeškal Z., 2005	84
Quarterly Review of Economics and Finance	Dimitrakopoulos D.N. et al., 2010; Hammoudeh S. et al., 2011; Mabrouk S. and Saadi S., 2012	77
Journal of Derivatives	Dowd K., 2001; Taylor J.W., 1999	72

Variance-covariance		
Journal	**Author(s), Year of Publication**	**Journal Citation Count**
Econometric Reviews	Manner H. and Reznikova O., 2012	57
Review of Quantitative Finance and Accounting	Angelidis T. et al., 2007; Baixauli-Soler J.S. and Alvarez S., 2006; Weiß G.N.F., 2013	53
Pacific Basin Finance Journal	Ho L.-C. et al., 2000; Weng H. and Trück S., 2011	52
International Journal of Theoretical and Applied Finance	Abad P. and Benito S., 2009; Angelidis T. and Skiadopoulos G., 2008; D'Addona S. and Ciprian M., 2007; Kamdem J.S., 2005; Mendes B.V.M. and De Melo E.F.L., 2010;	43
Journal of Risk Finance	Angelidis T. and Degiannakis S., 2005; Dowd K., 2000; Lechner L.A. and Ovaert T.C., 2010; Ozun A. et al., 2010	40
Applied Financial Economics	Papadamou S. and Stephanides G., 2004; Sadorsky P., 2005; Zhou J., 2012	38
Econometrics Journal	Chen X. et al., 2009; Kheifets I.L., 2015	36
Economic Notes	Barone-Adesi G. and Giannopoulos K., 2001	25
Journal of Economic Surveys	McAleer M., 2009	25
Journal of Financial Econometrics	Mancini L. and Trojani F., 2011; Žikeš F. and Baruník J., 2015	25
Spanish Review of Financial Economics	Abad P. et al., 2014	24
Economic Modelling	Louzis D.P. et al., 2014; Ourir A. and Snoussi W., 2012; Zhang H.-G. et al., 2017	20
Annals of Operations Research	Alonso E. et al., 2015; Moazeni S. et al., 2016; Stoyanov S.V. et al., 2013	19
Emerging Markets Finance and Trade	Su E. and Knowles T.W., 2006	17
Journal of Futures Markets	Cao Z. et al., 2010	15
Review of Financial Economics	Lu C. et al., 2009	13
International Journal of Managerial Finance	Maghyereh A.I. and Al-Zoubi H.A., 2006	12

Table A.6. Journals with more than 4 publications _ **ARCH/GARCH Models**

ARCH/GARCH Models																						
Journal	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**Total**
Journal of Banking and Finance				1	1	1			1		1	3			1	1	6	1	1	1	1	**20**
Journal of Empirical Finance			1	2			1	1	1		1	3	1		1	1	3	2		1		**19**
Quantitative Finance											1		1		2	1	1	2	1	2	3	**14**
International Review of Financial Analysis						1		1	1			1	1	1	2	1	2	1			1	**13**
Applied Financial Economics									1	1			3	2	1	2	1	1				**12**
Applied Economics															1	1			4	2	3	**11**
Economic Modelling														1	1	1	2	2	2	1	1	**11**
International Journal of Forecasting						1		1		3		1	1	1		1					1	**10**
North American Journal of Economics and Finance																	4	3	1		2	**10**
Applied Economics Letters								1				1		1	1	1	1	1	1		1	**9**
Journal of Financial Econometrics									1	1		1		1	3		1					**8**
Journal of Risk																	3	1		2	1	**7**
Managerial Finance												1		2	1	1	2					**7**
International Research Journal of Finance and Economics														1	4	1						**6**
International Review of Economics and Finance																1		2	1	2		**6**
European Journal of Finance									1				1			1			1		1	**5**

ARCH/GARCH Models																						
Journal	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**Total**
Journal of Derivatives	1	2	1									1										**5**
Quarterly Review of Economics and Finance														1	2	1					1	**5**
Review of Quantitative Finance and Accounting								1			1					2	1					**5**

Table A.7. Journals with more than 4 publications _ **Extreme Value Theory**

Extreme Value Theory																				
Journal	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
Journal of Banking and Finance		1		1			1	2	1				1	1	2	2				**12**
Applied Financial Economics						1				1	1	3		5						**11**
Quantitative Finance			1		2							1				3	1		1	**9**
International Review of Financial Analysis						1	1			1	1					2		1	1	**8**
Journal of Empirical Finance		2			1		1			1					1	1		1		**8**
Journal of Risk Finance			1					2				2		1					1	**7**
Economic Modelling														1	1	2	1	1		**6**
Journal of Risk														1			3	1		**5**

Table A.8. Journals with more than 4 publications _ **Monte Carlo Simulation**

Monte Carlo Simulation																			
Journal	**1999**	**2000**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**Total**
Quantitative Finance				1	1		1			2		2	1	2	2	1	2		**15**
Journal of Banking and Finance		2	1	1	2	1	1		1				1	1	2				**13**
International Journal of Forecasting							1		1				3		1		1		**7**
Journal of Risk												1			1	1	1	3	**7**
Management Science		1					1			2		1		1				1	**7**
Journal of Credit Risk															2	3		1	**6**
Journal of Empirical Finance		2							1	1	1					1			**6**
Applied Economics															1		1	3	**5**
Economic Modelling														2	2			1	**5**

Table A.9. Journals with more than 4 publications _ **Historical Simulation**

Historical Simulation																				
Journal	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**Total**
Journal of Banking and Finance		1				1	2	2				1		2	2		1			**12**
International Journal of Forecasting						1		1		1	1			1	1				1	**7**
Quantitative Finance			1		1				1							1	1		2	**7**
Economic Modelling														1	1	1	1	1	1	**6**
Journal of Risk													1			1	2	2		**6**
Applied Financial Economics						1					1			2		1				**5**
Journal of Financial Econometrics							1	1		1			1	1						**5**

Table A.10. Journals with more than 4 publications _ **Variance-covariance**

Variance-covariance																				
Journal	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**Total**
Quantitative Finance			1		1				1		1			1		3	1		2	**11**
Journal of Banking and Finance		1						1							3		2			**7**
International Journal of Forecasting				1		1		1						1	1				1	**6**
International Journal of Theoretical and Applied Finance							1		1	1	1	1								**5**
International Review of Financial Analysis				1			1	1			1		1							**5**
Journal of Risk														1			1	1	2	**5**

Table A.11. Top 10 publications by citation count _ **ARCH/GARCH Models**

ARCH/GARCH Models				
Article Title	**Author(s)**	**Year**	**Journal**	**Article Citation Count**
Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach	McNeil A.J., Frey R.	2000	Journal of Empirical Finance	583
Value-at-risk prediction: A comparison of alternative strategies	Kuester K., Mittnik S., Paolella M.S.	2006	Journal of Financial Econometrics	224
Modelling daily Value-at-Risk using realized volatility and ARCH type models	Giot P., Laurent S.	2004	Journal of Empirical Finance	172
An econometric analysis of emission allowance prices	Paolella M.S., Taschini L.	2008	Journal of Banking and Finance	132
Returns synchronization and daily correlation dynamics between international stock markets	Martens M., Poon S.-H.	2001	Journal of Banking and Finance	113
Value at risk when daily changes in market variables are not normally distributed	Hull J., White A.	1998	Journal of Derivatives	111
Systemic risk measurement: Multivariate GARCH estimation of CoVaR	Girardi G., Tolga Ergün A.	2013	Journal of Banking and Finance	91

ARCH/GARCH Models				
Article Title	**Author(s)**	**Year**	**Journal**	**Article Citation Count**
Using extreme value theory to measure value-at-risk for daily electricity spot prices	Fong Chan K., Gray P.	2006	International Journal of Forecasting	75
On a threshold heteroscedastic model	Chen C.W.S., So M.K.P.	2006	International Journal of Forecasting	74
Value-at-Risk: A multivariate switching regime approach	Billio M., Pelizzon L.	2000	Journal of Empirical Finance	72

Table A.12. Top 10 publications by citation count _ **Extreme Value Theory**

Extreme Value Theory				
Article Title	**Author(s)**	**Year**	**Journal**	**Article Citation Count**
Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach	McNeil A.J., Frey R.	2000	Journal of Empirical Finance	583
Value-at-risk prediction: A comparison of alternative strategies	Kuester K., Mittnik S., Paolella M.S.	2006	Journal of Financial Econometrics	224
From value at risk to stress testing: The extreme value approach	Longin F.M.	2000	Journal of Banking and Finance	210
Extreme value theory and Value-at-Risk: Relative performance in emerging markets	Gençay R., Selçuk F.	2004	International Journal of Forecasting	170
Value-at-risk versus expected shortfall: A practical perspective	Yamai Y., Yoshiba T.	2005	Journal of Banking and Finance	110
An Extreme Value Approach to Estimating Volatility and Value at Risk	Bali T.G.	2003	Journal of Business	97
Using extreme value theory to measure value-at-risk for daily electricity spot prices	Fong Chan K., Gray P.	2006	International Journal of Forecasting	79
Procyclical leverage and value-at-risk	Adrian T., Shin H.S.	2014	Review of Financial Studies	75
The quantitative modeling of operational risk: Between G-and-H and EVT	Degen M., Embrechts P., Lambrigger D.D.	2007	ASTIN Bulletin	61
Do hedge funds have enough capital? A value-at-risk approach	Gupta A., Liang B.	2005	Journal of Financial Economics	55

Table A.13. Top 10 publications by citation count _ **Monte Carlo Simulation**

Monte Carlo Simulation				
Article Title	**Author(s)**	**Year**	**Journal**	**Article Citation Count**
Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach	McNeil A.J., Frey R.	2000	Journal of Empirical Finance	583
Value at Risk	Linsmeier T.J., Pearson N.D.	2000	Financial Analysts Journal	116
Evaluating value-at-risk models with desk-level data	Berkowitz J., Christoffersen P., Pelletier D.	2011	Management Science	107
Variance reduction techniques for estimating value-at-risk	Glasserman P., Heidelberger P., Shahabuddin P.	2000	Management Science	93
On a threshold heteroscedastic model	Chen C.W.S., So M.K.P.	2006	International Journal of Forecasting	74
A Survey on Time-Varying Copulas: Specification, Simulations, and Application	Manner H., Reznikova O.	2012	Econometric Reviews	57
Do hedge funds have enough capital? A value-at-risk approach	Gupta A., Liang B.	2005	Journal of Financial Economics	55
Simulating sensitivities of Conditional value at risk	Hong L.J., Liu G.	2009	Management Science	48
Portfolio selection with limited downside risk	Jansen D.W., Koedijk K.G., De Vries C.G.	2000	Journal of Empirical Finance	46
On the way to recovery: A nonparametric bias free estimation of recovery rate densities	Renault O., Scaillet O.	2004	Journal of Banking and Finance	40

Table A.14. Top 10 publications by citation count _ **Historical Simulation**

Historical Simulation				
Article Title	**Author(s)**	**Year**	**Journal**	**Article Citation Count**
Value-at-risk prediction: A comparison of alternative strategies	Kuester K., Mittnik S., Paolella M.S.	2006	Journal of Financial Econometrics	229
Extreme value theory and Value-at-Risk: Relative performance in emerging markets	Gençay R., Selçuk F.	2004	International Journal of Forecasting	171
The level and quality of Value-at-Risk disclosure by commercial banks	Pérignon C., Smith D.R.	2010	Journal of Banking and Finance	129
Value at Risk	Linsmeier T.J., Pearson N.D.	2000	Financial Analysts Journal	116

Using extreme value theory to measure value-at-risk for daily electricity spot prices	Fong Chan K., Gray P.	2006	International Journal of Forecasting	80
The hidden dangers of historical simulation	Pritsker M.	2006	Journal of Banking and Finance	75
Nonparametric inference of value-at-risk for dependent financial returns	Chen S.X., Tang C.Y.	2005	Journal of Financial Econometrics	61
On the significance of expected shortfall as a coherent risk measure	Inui K., Kijima M.	2005	Journal of Banking and Finance	52
A comparison of extreme value theory approaches for determining value at risk	Brooks C., Clare A.D., Dalle Molle J.W., Persand G.	2005	Journal of Empirical Finance	49
A quantile regression approach to estimating the distribution of multiperiod returns	Taylor J.W.	1999	Journal of Derivatives	48

Table A.15. Top 10 publications by citation count _ **Variance-covariance**

Variance-covariance				
Article Title	**Author(s)**	**Year**	**Journal**	**Article Citation Count**
Extreme value theory and Value-at-Risk: Relative performance in emerging markets	Gençay R., Selçuk F.	2004	International Journal of Forecasting	162
Value at Risk	Linsmeier T.J., Pearson N.D.	2000	Financial Analysts Journal	110
Using extreme value theory to measure value-at-risk for daily electricity spot prices	Fong Chan K., Gray P.	2006	International Journal of Forecasting	79
Value-at-Risk: A multivariate switching regime approach	Billio M., Pelizzon L.	2000	Journal of Empirical Finance	68
The quantitative modeling of operational risk: Between G-and-H and EVT	Degen M., Embrechts P., Lambrigger D.D.	2007	ASTIN Bulletin	61
A Survey on Time-Varying Copulas: Specification, Simulations, and Application	Manner H., Reznikova O.	2012	Econometric Reviews	57
A comparison of extreme value theory approaches for determining value at risk	Brooks C., Clare A.D., Dalle Molle J.W., Persand G.	2005	Journal of Empirical Finance	49
A quantile regression approach to estimating the distribution of multiperiod returns	Taylor J.W.	1999	Journal of Derivatives	49

Variance-covariance				
Article Title	**Author(s)**	**Year**	**Journal**	**Article Citation Count**
Portfolio selection with limited downside risk	Jansen D.W., Koedijk K.G., De Vries C.G.	2000	Journal of Empirical Finance	48
Value-at-risk: Applying the extreme value approach to Asian markets in the recent financial turmoil	Ho L.-C., Burridge P., Cadle J., Theobald M.	2000	Pacific Basin Finance Journal	45

Table A.16. Top 10 authors by citation count _ **ARCH/GARCH models**

Authors	Citation Count
Frey R.	583
McNeil A.J.	583
Paolella M.S.	365
Giot P.	254
Kuester K.	224
Mittnik S.	224
Laurent S.	172
Taylor J.W.	150
Martens M.	149
So M.K.P.	146

Table A.17. Top 10 authors by citation count _ **Extreme Value Theory**

Author	Citation Count
McNeil A.J.	583
Frey R.	583
Bali T.G.	230
Paolella M.S.	224
Mittnik S.	224
Kuester K.	224
Longin F.M.	210
Selçuk F.	170
Gençay R.	170
Yoshiba T.	110
Yamai Y.	110

Table A.18. Top 10 authors by citation count _ **Monte Carlo Simulation**

Authors	Citation Count
Frey R.	583
McNeil A.J.	583
Linsmeier T.J.	116
Pearson N.D.	116
Pelletier D.	109
Berkowitz J.	107
Christoffersen P.	107
Chen C.W.S.	94
Glasserman P.	93
Heidelberger P.	93
Shahabuddin P.	93

Table A.19. Top 10 authors by citation count _ **Historical Simulation**

Authors	Citation Count
Kuester K.	229
Mittnik S.	229
Paolella M.S.	229
Gençay R.	171
Selçuk F.	171
Pérignon C.	143
Smith D.R.	143
Linsmeier T.J.	116
Pearson N.D.	116
Chen S.X.	108

Table A.20. Top 10 authors by citation count _ **Variance-covariance**

Authors	Citation Count
Gençay R.	162
Selçuk F.	162
Linsmeier T.J.	110
Pearson N.D.	110
Fong Chan K.	79
Gray P.	79
Billio M.	68
Pelizzon L.	68
Angelidis T.	67
McAleer M.	65

Table A.21. Top 10 authors by number of publications _ **ARCH/GARCH models**

Author	Number of Articles
Degiannakis S.	7
Lee M.-C.	7
Su J.-B.	7
Liu H.C.	6
Su Y.-C.	6
Aloui C.	5
Fabozzi F.J.	5
Hung J.-C.	5
McAleer M.	5
Weiß G.N.F.	5

Table A.22. Top 10 authors by number of publications _ **Extreme Value Theory**

Author	Number of Articles
Bali T.G.	5
Žiković S.	5
Cotter J.	4
Iglesias E.M.	4
Karmakar M.	4
Embrechts P.	3
Ghorbel A.	3
Hammoudeh S.	3
Herrera R.	3
Trabelsi A.	3

Table A.23. Top 10 authors by number of publications _ **Monte Carlo Simulation**

Author	Number of Articles
Fabozzi F.J.	4
Kim Y.S.	4
Rachev S.T.	3
Bollen B.	2
Chen P.	2
Degiannakis S.	2
Dunis C.L.	2
Floros C.	2
Handika R.	2
Harris R.D.F.	2
Kurosaki T.	2
Laws J.	2
Sermpinis G.	2
Su Y.-C.	2

Table A.24. Top authors by number of publications _ **Historical Simulation**

Author	**Number of articles**
Angelidis T.	3
Barone-Adesi G.	3
McAleer M.	3
Su J.-B.	3
Žiković S.	3
Asai M.; Chen S.X.; Degiannakis S.; Dimitrakopoulos D.N.; Fuertes A.-M.; Giannopoulos K.; Hammoudeh S.; Ho L.-C.; Huang A.Y.; Jacobson T.; Jiang J.; Kavussanos M.G.; Kijima M.; Lindé J.; Liu W.-H.; Olmo J.; Pérignon C.; Perote J.; Polanski A.; Radivojević N.; Roszbach K.; Scaillet O.; Schmidt R.; Smith D.R.; Stoja E.; Trück S.; Wang X.; Weiß G.N.F.	2

Table A.25. Top authors by number of publications _ **Variance-covariance**

Author	**Number of articles**
Benito S.	4
Weiß G.N.F.	4
Abad P.	3
Angelidis T.	3
Trück S.	3
Alonso E.; Bollen B.; Degiannakis S.; Dowd K.; Fabozzi F.J.; Harris R.D.F.; Ho L.-C.; Knowles T.W.; Lee H.; Li Y.; López-Martín C.; McAleer M.; Olmo J.; Polanski A.; Rodríguez G.; Rong N.; Scheffer M.; Silvapulle P.; Stoja E.; Su E.; Su J.-B.; Tejada J.	2

Table A.26. Top 10 journals by citation count _ **ARCH/GARCH models**

Journal	**Citation Count**	**Scopus Coverage**
Journal of Empirical Finance	1096	1993
Journal of Banking and Finance	693	1977
Journal of Financial Econometrics	439	2005
International Journal of Forecasting	345	1985
Journal of Derivatives	274	1996
International Review of Financial Analysis	207	1992
Review of Quantitative Finance and Accounting	131	1991
Journal of International Money and Finance	96	1982
North American Journal of Economics and Finance	96	1992
Economic Modelling	92	1984

Table A.27. Top 10 journals by citation count _ **Extreme Value Theory**

Journal	Citation Count	Scopus Coverage
Journal of Empirical Finance	744	1993
Journal of Banking and Finance	518	1977
International Journal of Forecasting	259	1985
Journal of Financial Econometrics	254	2005
International Review of Financial Analysis	141	1992
ASTIN Bulletin	106	1958 - 1969; 1971 - 1975; 1977 - 1982; 1984
Journal of Business	97	1978; 1996
Review of Financial Studies	75	1996
Quantitative Finance	66	2001
Journal of Risk and Insurance	65	1978 - 1979; 1996

Table A.28. Top 10 journals by citation count _ **Monte Carlo Simulation**

Journal	Citation Count	Scopus Coverage
Journal of Empirical Finance	688	1993
Management Science	304	1969
Journal of Banking and Finance	262	1977
International Journal of Forecasting	127	1985
Quantitative Finance	124	2001
Financial Analysts Journal	116	1996
Econometric Reviews	61	1982 - 1987; 1989 - 1995; 2003
Journal of Financial Economics	55	1974
Journal of Financial Econometrics	44	2005
Review of Quantitative Finance and Accounting	43	1991

Table A.29. Top 10 journals by citation count _ **Historical Simulation**

Journal	Citation Count	Scopus Coverage
Journal of Banking and Finance	454	1977
Journal of Financial Econometrics	357	2005
International Journal of Forecasting	322	1985
Financial Analysts Journal	116	1996
European Financial Management	98	1995
Quantitative Finance	86	2001
Journal of Empirical Finance	84	1993
Journal of Derivatives	62	1996
Quarterly Review of Economics and Finance	61	1992
Economic Modelling	60	1984

Table A.30. Top 10 journals by citation count _ **Variance-covariance**

Journal	Citation Count	Scopus Coverage
International Journal of Forecasting	301	1985
Journal of Empirical Finance	165	1993
Quantitative Finance	118	2001
Financial Analysts Journal	110	1996
Journal of Banking and Finance	105	1977
ASTIN Bulletin	89	1958 - 1969; 1971 - 1975; 1977 - 1982; 1984
International Review of Financial Analysis	84	1992
Quarterly Review of Economics and Finance	77	1992
Journal of Derivatives	72	1996
Econometric Reviews	57	1982 - 1987; 1989 - 1995; 2003

Table A.31. Top 10 journals by number of publications _ **ARCH/GARCH models**

Journal	Number of Articles	Scopus Coverage
Journal of Banking and Finance	20	1977
Journal of Empirical Finance	19	1993
Quantitative Finance	14	2001
International Review of Financial Analysis	13	1992
Applied Financial Economics	12	1991 - 2014
Applied Economics	11	1969
Economic Modelling	11	1984
International Journal of Forecasting	10	1985
North American Journal of Economics and Finance	10	1992
Applied Economics Letters	9	1994

Table A.32. Top 10 journals by number of publications _ **Extreme Value Theory**

Journal	Number of Articles	Scopus Coverage
Journal of Banking and Finance	12	1977
Applied Financial Economics	11	1991 - 2014
Quantitative Finance	9	2001
International Review of Financial Analysis	8	1992
Journal of Empirical Finance	8	1993
Journal of Risk Finance	7	1999
Economic Modelling	6	1984
Journal of Risk	5	2011
Applied Economics	4	1969
ASTIN Bulletin	4	1958 - 1969; 1971 - 1975; 1977 - 1982; 1984
International Journal of Forecasting	4	1985
International Review of Economics and Finance	4	2008
North American Journal of Economics and Finance	4	1992

Table A.33. Top 10 journals by number of publications _ **Monte Carlo Simulation**

Journal	Number of articles	Scopus Coverage
Quantitative Finance	15	2001
Journal of Banking and Finance	13	1977
International Journal of Forecasting	7	1985
Journal of Risk	7	2011
Management Science	7	1969
Journal of Credit Risk	6	2011
Journal of Empirical Finance	6	1993
Applied Economics	5	1969
Economic Modelling	5	1984
Journal of Derivatives	4	1996
Annals of Operations Research	4	1984
Economic Notes	4	2001
Journal of Financial Econometrics	4	2005
Journal of Risk Finance	4	1999

Table A.34. Top 10 journals by number of publications _ **Historical Simulation**

Journal	**Number of Articles**	**Scopus Coverage**
Journal of Banking and Finance	12	1977
International Journal of Forecasting	7	1985
Quantitative Finance	7	2001
Economic Modelling	6	1984
Journal of Risk	6	2011
Applied Financial Economics	5	1991 - 2014
Journal of Financial Econometrics	5	2005
Journal of Empirical Finance	4	1993
Journal of Risk Finance	4	1999
European Financial Management	3	1995
International Journal of Theoretical and Applied Finance	3	2003
International Review of Financial Analysis	3	1992
Journal of Econometrics	3	1973
North American Journal of Economics and Finance	3	1992
Review of Quantitative Finance and Accounting	3	1991

Table A.35. Top 10 journals by number of publications _ **Variance-covariance**

Journal	**Number of Articles**	**Scopus Coverage**
Quantitative Finance	11	2001
Journal of Banking and Finance	7	1977
International Journal of Forecasting	6	1985
International Journal of Theoretical and Applied Finance	5	2003
International Review of Financial Analysis	5	1992
Journal of Risk	5	2011
Journal of Risk Finance	4	1999
Annals of Operations Research	3	1984
Applied Financial Economics	3	1991 - 2014
ASTIN Bulletin	3	1958 - 1969; 1971 - 1975; 1977 - 1982; 1984
Economic Modelling	3	1984
Journal of Empirical Finance	3	1993
Quarterly Review of Economics and Finance	3	1992
Review of Quantitative Finance and Accounting	3	1991

Table A.36. Journals with more than 4 publications_ **All models**

All Models																						
Journals	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
Journal of Banking and Finance				5	1	4	1	2	4	5	4	5		2	3	5	13	5	4	4	2	69
Quantitative Finance					2		2	3		1	4		3	2	6	3	2	7	3	3	3	44
Journal of Risk															4	1	3	5	6	7	5	31
Applied Financial Economics								2	2	1		2	6	4	1	7	2	1				28
Journal of Empirical Finance			1	3			1	1	2		1	4	1	2	1	1	3	2	1	1	2	27
Economic Modelling														1	1	2	6	5	2	4	2	23
International Review of Financial Analysis						1		1	2	2		1	2	2	2	1	3	2		1	2	22
Journal of Risk Finance				1	3	1		1	1	2	1	1	1	3	1	1		1	1		3	22
International Journal of Forecasting						1		2		3		1	1	1	1	2	2	2		3	2	21
Applied Economics															1	1		3	6	4	3	18
North American Journal of Economics and Finance																	6	4	3		5	18
Journal of Financial Econometrics									2	1		3	1	1	4	2	1		1	1		17
Journal of Derivatives	2	2	2	1	1			1						1	2			2		1		15
Annals of Operations Research											1		1	1	1	1	3	1	1	3	1	14
International Journal of Theoretical and Applied Finance									1	3	3	1	1	2				1	1		1	14
Applied Economics Letters								1	1			1		1	1	1	2	1	2		2	13
ASTIN Bulletin							1		1		3		2	1	1			1	1	1		12
International Review of Economics and Finance												1			1	1	1	3	2	2	1	12
European Journal of Finance									1		2	1	1			1			3	1	1	11
International Research Journal of Finance and Economics													1	3	5	1						10
Finance Research Letters								2			1		1							1	4	9
Journal of Credit Risk															2			3	3		1	9
Journal of Futures Markets							1			1		1		4		1				1		9
Journal of Operational Risk															1	1	1	1	2	1	2	9
Management Science				1					1	1			2		1		2				1	9
Asia-Pacific Financial Markets				1			1			1		1		1			3					8
Economic Notes					3	2		1										1	1			8
Managerial Finance					1							1		2	1	1	2					8

All Models																						
Journals	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
Journal of International Financial Markets, Institutions and Money									1	1		1						1	2		1	7
Journal of International Money and Finance			1		2				1		1				1				1			7
Review of Quantitative Finance and Accounting								1		1	1					2	2					7
Econometric Reviews													1	1		1			1	1	1	6
Econometrics Journal													3						1		2	6
Economics Bulletin															2	1			1	1	1	6
Economics Letters													1					2	2	1		6
International Journal of Economics and Financial Issues																1		3	1	1		6
Investment Management and Financial Innovations								1					2	1		1	1					6
Pacific Basin Finance Journal				1											1		1	2			1	6
Quarterly Review of Economics and Finance														2	2	1					1	6
European Financial Management			1			1		1			1		1									5
International Journal of Monetary Economics and Finance												4			1							5
Journal of Business Economics and Management											1						1	1			2	5
Journal of Econometrics											1	1		1					1	1		5
Review of Financial Economics						1							1	1			1		1			5
Studies in Economics and Finance										1										1	3	5

Table A.37. Top 10 authors by number of publications **_ All models**

Author	**Number of Articles**
Hammoudeh S.	10
Degiannakis S.	9
Fabozzi F.J.	8
McAleer M.	8
Weiß G.N.F.	8
Aloui C.	7
Bali T.G.	7
Lee M.-C.	7
Rachev S.V.	7
Su J.-B.	7

Table A.38. Top 10 authors by citation count _ **All models**

Author	Citation Count
Frey R.	662
McNeil A.J.	662
Gordy M.B.	376
Paolella M.S.	365
Marrone J.	332
Bali T.G.	316
Giot P.	249
Kuester K.	224
Mittnik S.	224
Longin F.M.	210
Taylor J.W.	210
Embrechts P.	198
McAleer M.	174

Table A.39. Top 10 journals by citation count _ **All models**

Journal	Citation Count
Journal of Banking and Finance	2737
Journal of Empirical Finance	1235
Journal of Financial Econometrics	656
International Journal of Forecasting	587
Journal of Derivatives	404
Quantitative Finance	375
Management Science	358
International Review of Financial Analysis	350
ASTIN Bulletin	294
Economic Modelling	213

B. Appendix B

1. Dow Jones Industrial Average (2002 – 2007)

Backtesting 99% VaR

	DJIA 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	4.923	0.026	14.862	1.16E-04	4.923	0.026	4.923	0.026	8.843	2.94E-03	4.923	0.026	11.296	7.77E-04	31.426	2.07E-08
D.Q. Test	4.062	**0.773**	9.643	**0.210**	4.062	**0.773**	4.001	**0.780**	6.470	**0.486**	4.001	**0.780**	110.658	0	138.376	0
D-B Test	0.925	**0.826**	5.772	**0.120**	0.925	**0.826**	0.756	**0.429**	1.024	**0.961**	0.756	**0.429**	0.689	5.66E-03	0.730	4.47E-03

Table B.1. Backtesting 99% VaR results of DJIA between 2002 and 2007 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level																	
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	0.584	**0.445**	4.923	0.026	3.502	**0.061**	3.109	**0.078**	6.681	0.010	3.502	**0.061**	8.843	2.94E-03	19.260	1.14E-05
D.Q. Test	23.593	1.34E-03	4.130	**0.765**	3.878	**0.794**	23.346	0.001	5.151	**0.642**	3.894	**0.792**	6.590	**0.473**	11.521	**0.117**
D-B Test	0.849	**0.410**	0.979	**0.955**	0.900	**0.763**	0.813	**0.230**	0.683	**0.292**	0.900	**0.763**	1.734	**0.276**	2	**1**

Table B.2. Backtesting 99% VaR results of DJIA between 2002 and 2007 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	8.531	3.49E-03	7.269	7.02E-03	8.531	3.49E-03	0.251	**0.616**
D.Q. Test	49.958	1.47E-08	59.127	2.25E-10	55.282	1.31E-09	27.437	2.78E-04
D-B Test	1.603	0	1.603	0	0.646	1.14E-03	1.603	0

Table B.3. Backtesting 99% VaR results of DJIA between 2002 and 2007 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

	DJIA 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	9.391	2.18E-03	7.715	5.48E-03	9.391	2.18E-03	11.260	7.92E-04	8.530	3.49E-03	11.260	7.92E-04	0.123	**0.726**	4.921	0.027
DQ Test	21.303	3.35E-03	24.960	7.71E-04	21.303	3.35E-03	22.986	1.71E-03	30.496	7.70E-05	22.986	1.71E-03	44.617	1.62E-07	47.713	4.05E-08
D-B Test	0.892	**0.342**	0.870	**0.228**	0.892	**0.342**	0.905	**0.418**	0.871	**0.237**	0.905	**0.418**	0.702	2.28E-05	0.767	1.88E-04

Table B.4. Backtesting 95% VaR results of DJIA between 2002 and 2007 (GARCH-EVT and EVT (POT) models). Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	0.283	**0.594**	6.223	0.013	1.000	**0.317**	2.273	**0.132**	2.348	**0.125**	9.99E-04	**0.975**	8.530	3.49E-03	5.544	0.019
DQ Test	9.650	**0.209**	8.526	**0.289**	4.351	**0.739**	8.938	**0.257**	7.295	**0.399**	6.827	**0.447**	21.734	2.82E-03	20.104	0.005
D-B Test	1.049	**0.616**	1.049	**0.685**	1.160	**0.172**	0.997	**0.975**	1.031	**0.779**	1.010	**0.921**	0.841	**0.131**	0.954	**0.681**

Table B.5. Backtesting 95% VaR results of DJIA between 2002 and 2007 (MCS, GARCH (1,1) and FHS models). Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	2.648	**0.104**	3.050	**0.081**	0.080	**0.777**	0.049	**0.824**
DQ test	18.545	0.010	18.714	9.13E-03	42.507	4.15E-07	18.695	9.02E-04
D-B test	1.619	0	1.618	0	0.719	4.13E-05	1.607	0

Table B.6. Backtesting 95% VaR results of DJIA between 2002 and 2007 (EWMA, var-cov and HS models). Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

2. Euro Stoxx 50 (2002 – 2007)

Backtesting 99% VaR

	SX5E 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	14.862	1.16E-04	8.843	2.94E-03	14.862	1.16E-04	14.862	1.16E-04	11.512	6.91E-04	14.862	1.16E-04	0.584	**0.445**	6.092	0.014
DQ Test	9.894	**0.195**	8.223	**0.313**	9.894	**0.195**	9.654	**0.209**	8.208	**0.315**	9.654	**0.209**	68.186	3.43E-12	65.564	1.16E-11
D-B Test	10	0.017	1.179	**0.710**	10	0.017	1.781	**0.567**	2.077	**0.207**	1.781	**0.567**	0.539	6.08E-04	0.617	8.12E-04

Table B.7. Backtesting 99% VaR results of SX5E between 2002 and 2007 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2002 - 2007 Backtesting 99% VaR Results 5% Significance Level																			
	MCS						GARCH (1,1) model						FHS							
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value				
Kupiec	7.269	7.02E-03	0.584	**0.445**	1.488	**0.223**	8.531	3.49E-03	1.620	**0.203**	2.369	**0.124**	19.260	1.14E-05	11.512	6.91E-04				
DQ Test	17.177	0.016	7.793	**0.351**	3.720	**0.811**	18.315	0.011	9.610	**0.212**	3.029	**0.882**	11.319	**0.125**	8.250	**0.311**				
D-B Test	0.991	**0.955**	1.278	**0.242**	1.153	**0.616**	0.924	**0.609**	1.011	**0.954**	2.188	0.035	2	**1**	2.077	**0.207**				

Table B.8. Backtesting 99% VaR results of SX5E between 2002 and 2007 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	7.269	7.02E-03	6.092	0.014	7.269	7.02E-03	1.73E-03	**0.967**
DQ Test	21.112	3.61E-03	22.597	2.00E-03	60.985	9.59E-11	36.825	5.06E-06
D-B Test	1.200	**0.322**	1.053	**0.772**	0.587	2.44E-04	1.650	0

Table B.9. Backtesting 99% VaR results of SX5E between 2002 and 2007 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

	SX5E 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	16.834	4.08E-05	7.715	5.48E-03	16.834	4.08E-05	20.854	4.96E-06	6.947	8.40E-03	20.854	4.96E-06	0.123	**0.726**	0.283	**0.594**
DQ Test	24.260	1.03E-03	20.958	3.83E-03	24.260	1.03E-03	25.929	5.19E-04	21.810	2.74E-03	25.929	5.19E-04	33.278	2.35E-05	40.934	8.34E-07
D-B Test	0.955	**0.718**	1.010	**0.930**	0.955	**0.718**	1.041	**0.781**	0.973	**0.815**	1.041	**0.781**	0.714	6.15E-05	0.754	3.51E-04

Table B.10. Backtesting 95% VaR results of SX5E between 2002 and 2007 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	6.010	0.014	1.925	**0.165**	1.605	**0.205**	7.830	5.14E-03	4.414	0.036	2.648	**0.104**	4.908	0.027	5.544	0.019
DQ Test	28.536	1.76E-04	5.915	**0.550**	13.328	**0.065**	25.244	6.87E-04	9.541	**0.216**	12.783	**0.078**	9.844	**0.198**	12.228	**0.093**
D-B Test	1.032	**0.705**	0.984	**0.855**	0.985	**0.869**	1.094	**0.283**	1.044	**0.616**	0.993	**0.937**	1.006	**0.959**	1.048	**0.685**

Table B.11. Backtesting 95% VaR results of SX5E between 2002 and 2007 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	5.453	0.020	6.592	0.010	0.001	**0.975**	0.025	**0.875**
DQ Test	21.135	0.004	24.268	0.001	55.003	1.49E-09	55.950	9.66E-10
D-B Test	0.966	**0.668**	0.927	**0.336**	0.710	2.61E-05	1.650	0

Table B.12. Backtesting 95% VaR results of SX5E between 2002 and 2007 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

3. Nikkei 225 (2002 – 2007)

Backtesting 99% VaR

	N225 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level																
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	11.481	7.03E-04	11.481	7.03E-04	11.481	7.03E-04	8.815	2.99E-03	11.481	7.03E-04	8.815	2.99E-03	1.054	**0.305**	0.057	**0.811**
DQ Test	8.574	**0.285**	9.305	**0.232**	8.574	**0.285**	7.318	**0.397**	9.347	**0.229**	7.318	**0.397**	16.363	0.022	15.932	0.026
D-B Test	0.541	**0.238**	5.370	0.019	0.541	**0.238**	0.961	**0.931**	5.370	0.019	0.961	**0.931**	0.879	**0.525**	0.761	**0.180**

Table B.13. Backtesting 99% VaR results of N225 between 2002 and 2007 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	93.434	0	11.339	7.59E-04	5.031	0.025	106.291	0	12.840	3.39E-04	6.122	0.013	11.481	7.03E-04	11.481	7.03E-04
DQ Test	351.611	0	78.637	2.61E-14	33.735	1.93E-05	355.323	0	75.808	9.83E-14	37.251	4.20E-06	9.883	**0.195**	9.115	**0.244**
D-B Test	1.087	**0.412**	0.903	**0.519**	0.744	**0.055**	1.064	**0.526**	0.887	**0.435**	0.787	**0.127**	0.541	**0.238**	5.370	0.019

Table B.14. Backtesting 99% VaR results of N225 between 2002 and 2007 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2002 - 2007_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	14.414	1.47E-04	14.414	1.47E-04	14.414	1.47E-04	1.054	**0.305**
D-Q test	63.521	2.98E-11	81.384	7.22E-15	140.984	0	27.551	2.65E-04
D-B test	0.856	**0.299**	0.787	**0.105**	0.649	1.07E-03	1.574	0

Table B.15. Backtesting 99% VaR results of N225 between 2002 and 2007 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

N225 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level																
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	18.018	2.19E-05	9.319	2.27E-03	18.018	2.19E-05	16.740	4.29E-05	12.187	4.81E-04	16.740	4.29E-05	0.355	**0.551**	0.029	**0.865**
DQ Test	35.445	9.22E-06	26.572	3.98E-04	35.445	9.22E-06	32.574	3.18E-05	28.039	2.16E-04	32.574	3.18E-05	34.790	1.22E-05	50.235	1.30E-08
D-B Test	0.792	**0.068**	0.888	**0.306**	0.792	**0.068**	0.815	**0.108**	0.872	**0.258**	0.815	**0.108**	0.730	3.17E-04	0.731	1.96E-04

Table B.16. Backtesting 95% VaR results of N225 between 2002 and 2007 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

N225 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level																
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	86.827	0	53.790	2.23E-13	35.906	2.07E-09	101.759	0	58.334	2.21E-14	46.558	8.89E-12	9.319	2.27E-03	6.884	8.70E-03
DQ Test	209.349	0	137.163	0	90.666	1.11E-16	244.792	0	150.960	0	113.054	0	32.691	3.02E-05	27.599	2.60E-04
D-B Test	1.198	5.81E-03	1.234	3.06E-03	1.172	0.031	1.152	0.025	1.211	6.30E-03	1.152	0.047	0.922	**0.498**	0.927	**0.503**

Table B.17. Backtesting 95% VaR results of N225 between 2002 and 2007 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

N225 2002 - 2007_Backtesting 95% VaR Results_5% Significance Level								
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	4.470	0.034	2.691	**0.101**	0.177	**0.674**	0.044	**0.834**
DQ Test	27.379	2.85E-04	27.988	2.21E-04	45.022	1.35E-07	42.568	4.04E-07
D-B Test	0.960	**0.622**	0.929	**0.379**	0.735	1.94E-04	1.574	0

Table B.18. Backtesting 95% VaR results of N225 between 2002 and 2007 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

4. Dow Jones Industrial Average (2008 – 2013)

Backtesting 99% VaR

	DJIA 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	26.452	2.70E-07	26.452	2.70E-07	26.452	2.70E-07	26.452	2.70E-07	26.452	2.70E-07	26.452	2.70E-07	0.379	**0.538**	3.098	**0.078**
DQ Test	13.253	**0.066**	13.253	**0.066**	13.253	**0.066**	13.253	**0.066**	13.253	**0.066**	13.253	**0.066**	92.990	0	133.119	0
D-B Test	2	**1**	2	**1**	2	**1**	2	**1**	2	**1**	2	**1**	0.499	2.07E-03	0.588	8.29E-04

Table B.19. Backtesting 99% VaR results of DJIA between 2008 and 2013 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	6.077	0.014	0.836	**0.361**	14.879	1.15E-04	6.077	0.014	0.002	**0.965**	11.528	6.86E-04	19.279	1.13E-05	14.879	1.15E-04
D.Q. Test	31.131	5.88E-05	42.739	3.75E-07	77.262	4.97E-14	24.904	0.001	32.091	3.91E-05	45.956	8.92E-08	11.476	**0.119**	76.978	5.68E-14
D-B Test	1.029	**0.870**	0.736	**0.262**	0.200	0.016	1.003	**0.988**	0.721	**0.143**	0.344	0.042	2	**1**	0.200	0.016

Table B.20. Backtesting 99% VaR results of DJIA between 2008 and 2013 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	23.199	1.46E-06	23.199	1.46E-06	1.613	**0.204**	1.494	**0.222**
D-Q test	84.019	2.11E-15	109.577	0	61.016	9.46E-11	112.016	0
D-B test	0.822	**0.140**	0.772	0.047	1.541	0	1.541	0

Table B.21. Backtesting 99% VaR results of DJIA between 2008 and 2013 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

	DJIA 2008 - 2013_Backtesting 95% VaR Results_5% Significance Level																
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	38.697	4.95E-10	25.490	4.45E-07	38.697	4.95E-10	53.465	2.63E-13	22.369	2.25E-06	53.465	2.63E-13	6.253	0.012	2.366	**0.124**
DQ Test	34.975	1.13E-05	27.419	2.80E-04	34.975	1.13E-05	43.392	2.80E-07	26.454	4.18E-04	43.392	2.80E-07	24.995	7.60E-04	23.029	1.69E-03
D-B Test	0.754	**0.068**	0.889	**0.399**	0.754	**0.068**	0.828	**0.310**	0.980	**0.885**	0.828	**0.310**	0.683	7.35E-05	0.706	1.31E-04

Table B.22. Backtesting 95% VaR results of DJIA between 2008 and 2013 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2008 - 2013_Backtesting 95% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	2.627	**0.105**	3.907	0.048	1.615	**0.204**	4.386	0.036	4.386	0.036	0.052	**0.819**	11.299	7.76E-04	16.881	3.98E-05
DQ Test	20.704	4.23E-03	9.383	**0.226**	13.524	**0.060**	20.507	4.57E-03	10.981	**0.139**	16.783	0.019	22.292	2.26E-03	22.425	2.15E-03
D-B Test	1.002	**0.985**	1.010	**0.910**	0.985	**0.884**	0.996	**0.966**	1.035	**0.685**	0.962	**0.700**	0.847	**0.161**	1.021	**0.879**

Table B.23. Backtesting 95% VaR results of DJIA between 2008 and 2013 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2008 - 2013_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	4.386	0.036	3.454	**0.063**	3.786	**0.052**	3.273	**0.070**
DQ Test	17.051	0.017	13.423	**0.062**	17.943	0.012	24.236	1.04E-03
D-B Test	0.967	**0.686**	1.002	**0.977**	1.556	0	1.553	0

Table B.24. Backtesting 95% VaR results of DJIA between 2008 and 2013 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

5. Euro Stoxx 50 (2008 – 2013)

Backtesting 99% VaR

	SX5E 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level																
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	14.879	1.15E-04	14.879	1.15E-04	14.879	1.15E-04	14.879	1.15E-04	19.279	1.13E-05	14.879	1.15E-04	1.613	**0.204**	1.036	**0.309**
DQ Test	9.758	**0.203**	9.573	**0.214**	9.758	**0.203**	9.722	**0.205**	11.340	**0.124**	9.722	**0.205**	50.308	1.26E-08	51.291	8.05E-09
D-B Test	2.820	**0.293**	2.494	**0.344**	2.820	**0.293**	2.820	**0.293**	2	**1**	2.820	**0.293**	0.541	1.27E-04	0.495	9.59E-06

Table B.25. Backtesting 99% VaR results of SX5E between 2008 and 2013 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	25.160	5.28E-07	2.377	**0.123**	0.106	**0.744**	17.687	2.60E-05	3.511	**0.061**	0.379	**0.538**	8.857	2.92E-03	6.694	0.010
DQ Test	52.887	3.90E-09	3.504	**0.835**	14.906	0.037	40.341	1.08E-06	3.386	**0.847**	18.154	0.011	6.857	**0.444**	24.617	0.001
D-B Test	1.043	**0.758**	2.004	0.046	0.525	1.75E-03	1.050	**0.740**	1.108	**0.765**	0.496	1.21E-03	1.113	**0.816**	0.539	**0.105**

Table B.26. Backtesting 99% VaR results of SX5E between 2008 and 2013 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	14.340	1.53E-04	12.771	3.52E-04	3.996	0.046	0.579	**0.447**
DQ Test	41.158	7.55E-07	43.846	2.29E-07	56.752	6.70E-10	52.267	5.17E-09
D-B Test	0.973	**0.852**	0.852	**0.257**	0.516	3.83E-06	1.596	0

Table B.27. Backtesting 99% VaR results of SX5E between 2008 and 2013 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

	SX5E 2008 - 2013 Backtesting 95% VaR Results 5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	22.369	2.25E-06	5.572	0.018	22.369	2.25E-06	25.490	4.45E-07	4.339	0.037	25.490	4.45E-07	6.253	0.012	1.012	**0.314**
DQ Test	26.149	4.74E-04	12.535	**0.084**	26.149	4.74E-04	29.632	1.11E-04	17.996	0.012	29.632	1.11E-04	20.842	4.01E-03	26.053	4.93E-04
D-B Test	0.813	**0.108**	1.160	**0.225**	0.813	**0.108**	0.799	**0.088**	0.986	**0.898**	0.799	**0.088**	0.752	7.19E-03	0.746	1.62E-03

Table B.28. Backtesting 95% VaR results of SX5E between 2008 and 2013 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2008 - 2013_Backtesting 95% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	6.558	0.010	2.627	**0.105**	2.253	**0.133**	5.422	0.020	2.253	**0.133**	3.027	**0.082**	6.978	8.25E-03	4.934	0.026
DQ Test	20.929	3.88E-03	9.691	**0.207**	13.069	**0.070**	15.588	0.029	8.550	**0.287**	11.124	**0.133**	13.820	**0.054**	12.479	**0.086**
D-B Test	0.961	**0.625**	1.010	**0.908**	0.958	**0.628**	0.968	**0.696**	1.037	**0.685**	0.980	**0.821**	1.058	**0.642**	1.068	**0.581**

Table B.29. Backtesting 95% VaR results of SX5E between 2008 and 2013 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2008 - 2013_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	3.454	**0.063**	3.454	**0.063**	2.366	**0.124**	4.934	0.026
DQ Test	13.969	**0.052**	15.005	0.036	21.168	3.53E-03	18.232	1.10E-02
D-B Test	0.975	**0.765**	0.957	**0.605**	0.754	3.73E-03	1.590	0

Table B.30. Backtesting 95% VaR results of SX5E between 2008 and 2013 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

6. Nikkei 225 (2008 – 2013)

Backtesting 99% VaR

	N225 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level																
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	8.829	2.96E-03	11.497	6.97E-04	8.829	2.96E-03	14.845	1.17E-04	19.242	1.15E-05	14.845	1.17E-04	1.048	**0.306**	2.320	**0.128**
DQ Test	6.907	**0.439**	8.754	**0.271**	6.907	**0.439**	9.566	**0.215**	11.689	**0.111**	9.566	**0.215**	65.525	1.18E-11	61.429	7.82E-11
D-B Test	10	2.02E-04	2.183	**0.232**	10	2.02E-04	10	0.021	2	**1**	10	0.021	0.553	1.15E-03	0.565	6.80E-04

Table B.31. Backtesting 99% VaR results of N225 between 2008 and 2013 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level																
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	87.137	0	0.588	**0.443**	2.320	**0.128**	87.137	0	2.320	**0.128**	1.628	**0.202**	8.829	2.96E-03	8.829	2.96E-03
DQ Test	243.103	0	11.092	**0.135**	10.166	**0.179**	231.276	0	19.835	5.94E-03	9.877	**0.196**	40.905	8.44E-07	36.969	4.75E-06
D-B Test	1.060	**0.579**	0.991	**0.967**	0.723	**0.063**	1.067	**0.539**	0.847	**0.386**	0.724	**0.075**	0.348	0.010	0.512	**0.158**

Table B.32. Backtesting 99% VaR results of N225 between 2008 and 2013 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2008 - 2013_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	5.017	0.025	5.017	0.025	2.320	**0.128**	0.588	**0.443**
DQ Test	24.478	9.39E-04	25.507	6.17E-04	60.763	1.06E-10	59.284	2.10E-10
D-B Test	0.774	**0.130**	0.774	**0.130**	0.555	3.93E-04	1.523	0

Table B.33. Backtesting 99% VaR results of N225 between 2008 and 2013 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

	N225 2008 - 2013 Backtesting 95% VaR Results 5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	25.377	4.71E-07	11.220	8.09E-04	25.377	4.71E-07	27.036	2.00E-07	11.220	8.09E-04	27.036	2.00E-07	5.516	0.019	1.583	**0.208**
DQ Test	27.977	2.22E-04	25.500	6.18E-04	27.977	2.22E-04	29.505	1.17E-04	26.529	4.05E-04	29.505	1.17E-04	20.899	3.92E-03	18.974	8.27E-03
D-B Test	0.763	0.034	0.902	**0.398**	0.763	0.034	0.740	0.020	0.853	**0.191**	0.740	0.020	0.687	1.24E-04	0.741	9.34E-04

Table B.34. Backtesting 95% VaR results of N225 between 2008 and 2013 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2008 - 2013_Backtesting 95% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	83.103	0	45.066	1.90E-11	40.995	1.53E-10	86.685	0	46.459	9.36E-12	39.674	3.00E-10	7.682	5.58E-03	13.287	2.67E-04
DQ Test	139.099	0	72.474	4.67E-13	60.863	1.02E-10	145.250	0	75.885	9.48E-14	57.812	4.12E-10	15.915	2.59E-02	32.478	3.31E-05
D-B Test	1.096	**0.147**	1.124	**0.095**	1.063	**0.378**	1.077	**0.239**	1.088	**0.222**	1.061	**0.392**	0.906	**0.397**	0.773	0.031

Table B.35. Backtesting 95% VaR results of N225 between 2008 and 2013 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2008 - 2013_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	1.943	**0.163**	2.669	**0.102**	6.193	0.013	4.289	0.038
DQ Test	8.726	**0.273**	12.425	**0.087**	16.535	0.021	19.049	8.04E-03
D-B Test	1.009	**0.922**	0.922	**0.336**	0.733	0.002	1.523	0

Table B.36. Backtesting 95% VaR results of N225 between 2008 and 2013 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

7. Dow Jones Industrial Average (2014 – 2019)

Backtesting 99% VaR

	DJIA 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	26.432	2.73E-07	26.432	2.73E-07	26.432	2.73E-07	19.260	1.14E-05	26.432	2.73E-07	19.260	1.14E-05	6.092	0.014	14.364	1.51E-04
DQ Test	13.242	**0.066**	13.242	**0.066**	13.242	**0.066**	11.372	**0.123**	13.242	**0.066**	11.372	**0.123**	68.160	3.48E-12	150.917	0
D-B Test	2	**1**	2	**1**	2	**1**	2	**1**	0.8741	**1**	2	**1**	0.690	0.019	0.693	7.02E-03

Table B.37. Backtesting 99% VaR results of DJIA between 2014 and 2019 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	9.874	1.68E-03	0.831	**0.362**	8.843	2.94E-03	11.296	7.77E-04	0.831	**0.362**	11.512	6.91E-04	8.843	2.94E-03	14.862	1.16E-04
DQ Test	87.466	4.44E-16	42.753	3.72E-07	11.422	**0.121**	85.881	8.88E-16	42.831	3.60E-07	10.090	**0.184**	6.970	**0.432**	9.565	**0.215**
D-B Test	0.797	**0.154**	0.834	**0.545**	3.145	0.043	0.835	**0.254**	0.834	**0.545**	2.992	**0.151**	1.351	**0.565**	0.476	**0.344**

Table B.38. Backtesting 99% VaR results of DJIA between 2014 and 2019 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	16.006	6.32E-05	14.364	1.51E-04	25.194	5.19E-07	0.054	**0.816**
DQ Test	97.551	0	99.018	0	127.989	0	97.891	0
D-B Test	0.724	0.020	0.745	0.039	0.6950	3.50E-03	1.476	0

Table B.39. Backtesting 99% VaR results of DJIA between 2014 and 2019 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

	DJIA 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	68.957	1.11E-16	18.115	2.08E-05	68.957	1.11E-16	68.957	1.11E-16	16.834	4.08E-05	68.957	1.11E-16	0.009	**0.924**	5.453	0.020
DQ Test	55.598	1.13E-09	27.977	2.22E-04	55.598	1.13E-09	55.538	1.17E-09	28.167	2.05E-04	55.538	1.17E-09	27.085	3.22E-04	49.083	2.19E-08
D-B Test	0.533	1.94E-03	0.935	**0.630**	0.533	1.94E-03	0.550	3.93E-03	0.874	**0.316**	0.550	3.93E-03	0.724	1.79E-04	0.775	5.86E-04

Table B.40. Backtesting 95% VaR results of DJIA between 2014 and 2019 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level																	
	MCS						GARCH (1,1) model						FHS					
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t			
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value		
Kupiec	1.605	**0.205**	0.123	**0.726**	1.954	**0.162**	4.414	0.036	9.04E-03	**0.924**	1.000	**0.317**	4.908	0.027	7.715	5.48E-03		
DQ Test	47.434	4.59E-08	19.686	6.29E-03	23.076	1.65E-03	48.380	3.00E-08	15.901	0.026	21.825	2.72E-03	20.809	4.06E-03	16.868	0.018		
D-B Test	0.855	**0.058**	0.797	0.012	0.805	0.028	0.879	**0.105**	0.827	0.034	0.813	0.030	0.854	**0.146**	0.874	**0.239**		

Table B.41. Backtesting 95% VaR results of DJIA between 2014 and 2019 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	DJIA 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	1.049	**0.306**	0.607	**0.436**	0.283	**0.594**	0.009	**0.924**
DQ Test	11.191	**0.130**	9.095	**0.246**	39.517	1.56E-06	27.239	3.02E-04
D-B Test	0.920	**0.347**	0.922	**0.371**	0.7224	8.18E-05	1.460	0

Table B.42. Backtesting 95% VaR results of DJIA between 2014 and 2019 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

8. Euro Stoxx 50 (2014 – 2019)

Backtesting 99% VaR

	SX5E 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	14.862	1.16E-04	19.260	1.14E-05	14.862	1.16E-04	11.512	6.91E-04	14.862	1.16E-04	11.512	6.91E-04	1.620	**0.203**	8.531	3.49E-03
DQ Test	9.783	**0.201**	11.305	**0.126**	9.783	**0.201**	45.903	9.13E-08	10.444	**0.165**	45.903	9.13E-08	29.629	1.11E-04	81.518	6.77E-15
D-B Test	0.887	**0.894**	2	**1**	0.887	**0.894**	0.345	0.043	10	0.019	0.345	0.043	0.968	**0.863**	0.754	**0.056**

Table B.43. Backtesting 99% VaR results of SX5E between 2014 and 2019 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.x`

	SX5E 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	7.269	7.02E-03	0.376	**0.540**	2.369	**0.124**	7.269	7.02E-03	0.105	**0.746**	2.369	**0.124**	6.681	0.010	11.512	6.91E-04
DQ Test	31.614	4.79E-05	16.095	0.024	7.532	**0.376**	32.949	2.71E-05	14.780	**0.039**	9.267	**0.234**	5.657	**0.580**	7.898	**0.342**
D-B Test	1.008	**0.959**	0.979	**0.924**	0.770	**0.350**	0.968	**0.837**	1.145	**0.557**	0.929	**0.796**	2.461	**0.055**	1.555	**0.474**

Table B.44. Backtesting 99% VaR results of SX5E between 2014 and 2019 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	16.006	6.32E-05	16.006	6.32E-05	7.269	7.02E-03	0.584	**0.445**
DQ Test	43.620	2.53E-07	53.604	2.82E-09	34.718	1.26E-05	26.184	4.67E-04
D-B Test	1.153	**0.330**	1.064	**0.671**	0.910	**0.556**	1.550	0

Table B.45. Backtesting 99% VaR results of SX5E between 2014 and 2019 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Backtesting 95% VaR

	SX5E 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	16.834	4.08E-05	7.715	5.48E-03	16.834	4.08E-05	19.454	1.03E-05	8.530	0.003	19.454	1.03E-05	0.001	**0.975**	3.050	**0.081**
DQ Test	38.543	2.39E-06	15.025	0.036	38.543	2.39E-06	40.341	1.08E-06	16.837	0.018	40.341	1.08E-06	36.961	4.77E-06	69.486	1.88E-12
D-B Test	0.775	0.036	1.220	**0.129**	0.775	0.036	0.774	0.043	1.003	**0.975**	0.774	0.043	0.811	0.014	0.762	3.07E-04

Table B.46. Backtesting 95% VaR results of SX5E between 2014 and 2019 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	4.921	0.027	15.426	8.58E-05	10.592	1.14E-03	6.592	0.010	15.426	8.58E-05	12.1135	5.01E-04	4.908	0.027	7.715	0.005
DQ Test	30.929	6.41E-05	57.418	4.94E-10	41.930	5.36E-07	32.769	2.92E-05	55.978	9.54E-10	49.069	2.20E-08	15.704	0.028	12.053	**0.099**
D-B Test	0.942	**0.484**	0.944	**0.468**	0.962	**0.633**	0.938	**0.443**	0.946	**0.485**	0.966	**0.667**	0.941	**0.565**	1.004	**0.974**

Table B.47. Backtesting 95% VaR results of SX5E between 2014 and 2019 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	SX5E 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	1.313	**0.252**	0.283	**0.594**	0.167	**0.683**	0.283	**0.594**
DQ Test	20.691	0.004	17.706	0.013	40.384	1.06E-06	36.929	4.84E-06
D-B Test	0.963	**0.675**	0.949	**0.577**	0.771	1.32E-03	1.552	0

Table B.48. Backtesting 95% VaR results of SX5E between 2014 and 2019 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

9. Nikkei 225 (2014 – 2019)

Backtesting 99% VaR

	N225 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level															
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	14.828	1.18E-04	26.392	2.79E-07	14.828	1.18E-04	14.828	1.18E-04	26.392	2.79E-07	14.828	1.18E-04	6.122	0.013	6.122	0.013
DQ Test	9.766	**0.202**	13.222	**0.067**	9.766	**0.202**	9.779	**0.201**	13.222	**0.067**	9.779	**0.201**	61.422	7.85E-11	63.514	2.99E-11
D-B Test	0.431	**0.280**	2	**1**	0.431	**0.280**	0.431	**0.280**	2	**1**	0.431	**0.280**	0.616	1.38E-03	0.569	1.23E-04

Table B.49. Backtesting 99% VaR results of N225 between 2014 and 2019 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level															
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	109.588	0	7.303	6.89E-03	19.223	1.16E-05	126.540	0	5.031	0.025	19.223	1.16E-05	8.815	2.99E-03	26.392	2.79E-07
DQ Test	352.334	0	39.793	1.38E-06	11.295	**0.126**	380.333	0	35.085	1.08E-05	11.294	**0.126**	9.458	**0.221**	13.222	**0.067**
D-B Test	0.992	**0.934**	1.427	0.043	2	**1**	0.975	**0.788**	1.846	4.00E-03	2	**1**	0.866	**0.772**	2	**1**

Table B.50. Backtesting 99% VaR results of N225 between 2014 and 2019 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2014 - 2019_Backtesting 99% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	23.296	1.39E-06	19.549	9.81E-06	12.840	3.39E-04	3.130	**0.077**
DQ Test	56.474	7.60E-10	57.195	5.47E-10	67.609	4.49E-12	58.418	3.12E-10
D-B Test	1.091	**0.549**	0.993	**0.962**	0.695	8.22E-03	1.434	0

Table B.51. Backtesting 99% VaR results of N225 between 2014 and 2019 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

95% Backtesting

	N225 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level																
	GARCH-EVT 95% threshold						GARCH-EVT 90% threshold						EVT (POT) 95% threshold		EVT (POT) 90% threshold	
	Normal		Student-t		QMLE		Normal		Student-t		QMLE					
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	34.337	4.63E-09	16.740	4.29E-05	34.337	4.63E-09	43.007	5.45E-11	18.018	2.19E-05	43.007	5.45E-11	0.526	**0.468**	0.029	**0.865**
DQ Test	37.311	4.10E-06	21.521	3.07E-03	37.311	4.10E-06	36.859	4.99E-06	21.712	2.85E-03	36.859	4.99E-06	30.015	9.44E-05	36.161	6.76E-06
D-B Test	0.935	**0.684**	1.209	**0.180**	0.935	**0.684**	1.026	**0.893**	1.181	**0.243**	1.026	**0.893**	0.724	2.82E-04	0.731	1.85E-04

Table B.52. Backtesting 95% VaR results of N225 between 2014 and 2019 (GARCH-EVT and EVT (POT) models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level																
	MCS						GARCH (1,1) model						FHS			
	Normal		t-EK method		t-Fitting method		Normal		t-EK method		t-Fitting method		Normal		Student-t	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	58.334	2.21E-14	9.246	2.36E-03	0.088	**0.767**	69.513	1.11E-16	18.259	1.93E-05	0.628	**0.428**	6.884	8.70E-03	18.018	2.19E-05
DQ Test	122.939	0	49.152	2.12E-08	26.362	4.34E-04	137.867	0	49.908	1.51E-08	26.022	4.99E-04	19.654	6.37E-03	23.063	1.66E-03
D-B Test	1.031	**0.647**	0.974	**0.746**	0.910	**0.304**	1.035	**0.596**	0.989	**0.881**	0.939	**0.488**	0.959	**0.721**	1.204	**0.1996**

Table B.53. Backtesting 95% VaR results of N225 between 2014 and 2019 (MCS, GARCH (1,1) and FHS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

	N225 2014 - 2019_Backtesting 95% VaR Results_5% Significance Level							
	EWMA 0.94		EWMA 0.95		Var-cov		HS	
	Test	P-value	Test	P-value	Test	P-value	Test	P-value
Kupiec	3.096	**0.078**	3.096	**0.078**	0.448	**0.503**	0.114	**0.736**
DQ Test	14.173	0.048	19.151	7.73E-03	37.619	3.58E-06	33.203	2.43E-05
D-B Test	0.995	**0.953**	0.945	**0.515**	0.757	6.13E-04	1.434	0

Table B.54. Backtesting 95% VaR results of N225 between 2014 and 2019 (EWMA, var-cov and HS models); Fonts formatted in bold indicate that the model failed to reject the null hypothesis of the corresponding backtesting measure.

Data (including Excel Files and R code) is available upon request. For contact: reem.shayya@gmail.com.