VALUE-AT-RISK: MARKET RISK ESTIMATION IN STRESSED MARKET CONDITIONS

Lappeenranta–Lahti University of Technology LUT

Bachelor's Thesis (Business Analytics)

2021

Paavo Kauhanen

Examiners: Mahinda Mailagaha Kumbure, M.Sc.

Christoph Lohrmann, Ph.D.

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT School of Business and Management

Business Administration

Paavo Kauhanen

Value-at-Risk: market risk estimation in stressed market conditions

Bachelor's thesis
2021
45 pages, 4 figures, 7 tables, and 4 appendices
Examiners: Mahinda Mailagaha Kumbure, M.Sc. and Christoph Lohrmann, Ph.D.

Keywords: Value-at-Risk, financial risk management, market risk estimation, stressed market conditions, backtesting

This bachelor's thesis examines the usability and reliability of Value-at-Risk (VaR) modeling in stressed market conditions. The objective of the study is to investigate whether VaR models can produce sufficiently accurate estimates during the increased volatility and unpredictability caused by covid-19 during the year 2020. The study utilizes Historical VaR (HVaR) and Monte Carlo VaR (MCVaR) calculation methods to estimate the market risk of three popular stock indices (S&P 500, STOXX Europe 600, and MSCI Emerging Markets).

The performance of HVaR and MCVaR was analyzed using a statistical backtesting method, which compared the ratio of observed model failures to a prespecified theoretical distribution. The backtesting models were constructed using market data ranging from 31.12.2009 to 31.12.2020. The results obtained from the backtesting indicated that the VaR models were not able to accurately estimate the increased risk and volatility present in stressed market conditions during the year 2020. The performance of both HVaR and MCVaR was overall very similar. However, MCVaR performed, in some cases, slightly better, as it produced a smaller number of total failures. The results also highlighted an interesting relationship between the size of the sampling window and the performance of the models. Using smaller sampling windows resulted in more accurate risk estimates by increasing the reactivity and adaptivity of the VaR models.

Most importantly, it could be concluded, that due to increasing levels of volatility and unpredictability, market risk estimation became increasingly challenging during the stressed market conditions of 2020.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT-kauppakorkeakoulu

Kauppatieteet

Paavo Kauhanen

Value-at-Risk: markkinariskin estimointi stressatuissa markkinaolosuhteissa

Kauppatieteiden kandidaatintutkielma
2021
45 sivua, 4 kuvaa, 7 taulukkoa ja 4 liitettä
Tarkastajat: Mahinda Mailagaha Kumbure, M.Sc. and Christoph Lohrmann, Ph.D.

Avainsanat: Value-at-Risk, rahoitusriskien hallinta, markkinariskin estimointi, stressatut markkinaolosuhteet, toteutumatestaus

Tämä kandidaatintutkielma käsittelee Value-at-Risk (VaR) mallinnusta stressatuissa markkinaolosuhteissa. Tutkimuksen tavoitteena on tarkastella VaR-menetelmien luotettavuutta ja yleistä käytettävyyttä covid-19:n aiheuttamissa korkean volatiliteetin markkinaolosuhteissa. Tutkimuksessa hyödynnetään Historical VaR (HVaR) sekä Monte Carlo VaR (MCVaR) -menetelmiä, joiden avulla estimoidaan kolmen suositun osakeindeksin (S&P 500, STOXX Europe 600, MSCI Emerging Markets) markkinariskiä.

HVaR ja MCVaR -mallien suoriutumista tarkasteltiin tutkimuksessa tilastollisella toteutumatestausmenetelmällä, jossa verrattiin todellisten mallinnusvirheiden suhdetta ennalta määritettyyn teoreettiseen jakaumaan. Nämä toteutumatestausmallit rakennettiin käyttäen markkinadataa väliltä 31.12.2009-31.12.2020. Tutkimuksen tulokset osoittivat, että VaR mallit eivät onnistuneet estimoimaan markkinariskiä riittävällä tarkkuudella vuoden 2020 aikana vallinneissa stressatuissa markkinaolosuhteissa. HVaR ja MCVaR -mallien tuottamien tulosten väliset erot olivat pieniä, mutta MCVaR suoriuitui osassa testejä kuitenkin hieman HVaR:ia paremmin. Huomio perustui mallinnusvirheiden kokonaismäärään. Tämän lisäksi testien tulokset korostivat otoskoon ja -menetelmän vaikutusta mallinnustarkkuuteen; pienemmät otoskoot tuottivat tarkempia riskiestimaatteja, sillä ne lisäsivät mallien joustavuutta ja sopeutumiskykyä uusiin markkinaolosuhteisiin.

Ennen kaikkea tuloksista huomattiin, että korkea volatiliteetti ja kohonnut epävarmuus vaikeuttivat markkinariskin mallinnusta huomattavasti vuoden 2020 aikana vallinneissa stressatuissa markkinaolosuhteissa.

Table of contents

Appendices

1. Introduction

The role of risk management has become one of the most interesting topics in finance during the last couple of decades. Its relevance has risen, as various financial crises have created devastating outcomes for companies and national economies around the world. The increasing complexity and integration of global financial markets have made these crises more damaging and challenging to control (Claessens 2014, xvii). For example, the financial crisis of 2008 started a turmoil in the U.S. banking system, which consequently resulted in the bankruptcies and bailouts of numerous significant financial institutions (FCIC 2011). The initial shock caused by these bankruptcies created a rapidly increasing wave of uncertainty in the financial markets, which developed into a global crisis. This chain of events ultimately started a global economic recession and increased the number of unemployed people by 34 million worldwide (ILO 2010). Most institutions and economies were not prepared for a potential financial shock of such magnitude.

Global crises influence the financial markets by creating adverse market conditions. These conditions are generally referred to as “stressed market conditions”, which are identified by significant short-term changes in market prices and trading volumes (ESMA 2020). During these conditions, market participants face increasing levels of unpredictability making financial risk management more challenging. This increased unpredictability is usually caused by strong market reactions to unprecedented events. Therefore, preparing for potential financial shocks and unfavorable market conditions is crucial.

Financial and economic shocks are challenging to predict by nature. Even the most sophisticated statistical models and measures cannot provide complete certainty about the future. Despite this shortcoming, they may still create highly probable estimates with sufficient accuracies. Technological innovations and the increasing amount of information have created new possibilities to build more advanced models that can estimate potential outcomes more precisely. This emerging opportunity to produce more accurate estimates has created an initiative for extensive research surrounding the different methods of modeling financial risk. Many analytical and statistical tools have been developed to capture various risk factors by examining the mathematical characteristics of financial markets. One of these tools is called Value-at-Risk (VaR); a statistical measure used in estimating the maximum loss that an investment will endure over a given time horizon at a given level of confidence (Jorion 2011, 17).

1.1 Background

VaR analysis was first made popular in 1994 by U.S. investment bank J.P. Morgan (Adamko, Spuchľakova & Valášková 2015). It was included as a part of their risk management product *RiskMetrics™*, which quickly gained the attention of financial professionals and researchers. The launch of *RiskMetrics™* was later accompanied by a technical document that explained the methodology in detail (Longerstaey & Spencer 1996). This was the first time VaR was exposed to a broader audience and not just an exclusive group of quantitative analysts. The initial exposure created a wave of momentum, which set the stage for large-scale adoption in the future.

The popularity of VaR kept increasing towards the end of the 1990s. VaR was usually reported as a single monetary value, which made the measure very easy to understand for the general audience. It was also simple to implement as a risk model. Many banks and investment companies adopted VaR modeling in their risk management during this period. The newly gained popularity also attracted a fair amount of criticism. Taleb (1997) famously debated the general concept of VaR in his fiery answer to Jorion, which led to an intense discussion on the usability of the methodology. Artzner, Delbaen, Eber, and Heath (1999) continued the critical review of VaR by examining the model from a utility-theoretic perspective. Their primary argument was that VaR did not fulfill the requirements of a coherent risk model. Despite the criticisms, VaR took its place as one of the core measures of risk in financial institutions. The importance of VaR was further reinforced in the following decade when the Basel Committee of Banking Supervision (BCBS) released the Basel II accord and announced VaR as the preferred approach to market risk estimation (BCBS 2005). The Basel II accord was the second of the three Basel accords (Basel I, Basel

II & Basel III), which were issued by the BCBS. It provided guidelines and recommendations for banks and financial institutions that revolved around capital requirements, market risk management, and reporting. The members of BCBS consisted of central bank representatives and regulatory authorities of the major global economies. Although the Basel accords were not enforced by law, the financial policies and regulative measures were implemented on a large scale and became the basis for banking regulations. Basel II, therefore, made numerous large banks adopt VaR as their primary tool for calculating minimum capital requirements.

VaR's underlying method of estimating maximum probable losses has since been studied to a great extent. The model has been through numerous empirical studies, and it has been tested comprehensively in real businesses. For example, Berkowitz and O'Brien (2002) examined the statistical accuracy of the VaR models used by six large U.S. banks. The study concluded that on average the VaR forecasts produced overly conservative risk estimates. This overestimation of market risk, while implying higher capital coverage levels, did not accurately model portfolio risk. Multiple scientific publications have then explored improved calculation methods for VaR and suggested new model variations to account for these kinds of inaccuracies. For example, Cvjetković, Stepanov, and Radivojević (2016) developed a new VaR estimation approach that utilized extreme value theory (EVT) in the estimation process. The purpose of this hybrid approach was to improve the model's ability to capture varying volatility and excess kurtosis, while also utilizing a GARCH (p, q) model to capture possible heteroscedasticity in the data. These modified techniques of estimation were implemented in order to better account for extreme returns in volatile markets (Cvjetković, Stepanov & Radivojević 2016). The subject of time-varying volatility was also studied by Gabrielsen, Kirchner, Liu, and Zagaglia (2015) in a paper that explored the application and the role of exponentially weighted moving average (EWMA) framework in VaR modeling. They suggested in the study that more dynamic and accurate VaR models could be constructed by utilizing the higher moments of the data. This observation was meaningful, as the study highlighted the utility of incorporating new statistical concepts into VaR, which improved the modeling of risk in financial markets and various portfolio structures.

Even though multiple advanced VaR models have been developed, the three most popular and widely used methods are Historical VaR, Parametric VaR, and Monte Carlo VaR (Horcher 2005, 217). These methods provide distinct approaches to calculating VaR that lead to slightly different results depending on the data. This study will focus on Historical VaR and Monte Carlo VaR.

1.2 The objective of the study

The objective of this bachelor's thesis is to investigate the usability of Historical VaR (HVaR) and Monte Carlo VaR (MCVaR) in stressed market conditions. Despite the extensive research surrounding these methodologies, it is still unclear how applicable they are under periods of increased unpredictability and extreme volatility. Both models will be studied individually to gain specific insight into their level of performance and reliability. The individual results will also be compared against each other to explore possible differences between the performance of the models. The study will be conducted using a statistical backtesting method, which will be described in depth later.

The VaR models will be backtested using daily market data from the year 2020. This testing period was chosen, as the initial impact of covid-19 provides a suitable period of stressed market conditions for the study. Basuony, Bouaddi, Ali, and EmadEldeen (2021) concluded that the beginning of the global pandemic resulted in an unprecedented increase in conditional volatility across the global stock markets. This effect of extreme volatility was measured in the study by utilizing an asymmetric exponential GARCH model. The increased volatility and unpredictability of the financial markets during the onset of a global pandemic in 2020 offers, therefore, an interesting avenue for examining the performance of VaR models in adverse market conditions. Both HVaR and MCVaR will be used to forecast daily market risk for the same data to capture possible differences in their applicability and reliability. Finally, to ensure an adequate amount of training data, both models will be constructed using ten years of daily market data ranging from 2010 to 2020. This data selection guarantees enough data points for creating robust models and accurate estimates, which will also be discussed in more depth later.

1.3 Research problem and questions

Dramatic market shocks created by global crises are almost impossible to predict. The primary research problem of this study emerges from this unpredictability. Financial markets react strongly to unexpected events, which causes extreme volatility (Schwert 2011). This instability of high-volatility markets is challenging to account for with mathematical models. The study will investigate this problem by exploring the usability of VaR modeling during a period of abnormal market conditions observed in 2020. The differences between the estimates and the actual daily observations are at the core of this investigation. It will be interesting to assess how well VaR models can forecast and measure market risk when facing heightened levels of unpredictability caused by covid-19. These results will be insightful as VaR is still one of the most used risk management tools in the financial industry.

This study aims to answer two central questions related to VaR modeling:

1. *Is VaR a reliable risk measure when stressed market conditions are present?*
2. *Do the Historical VaR and Monte Carlo VaR models perform equally well during the testing?*

These questions were formed based on the objective of the thesis. Their purpose is to guide the focus of this study and help with investigating the fundamental research problem. Both questions will be answered in the empirical section of the study by using the results obtained from the backtesting.

1.4 The scope and delimitations of the study

Financial risk management and Value-at-Risk modeling are vast fields of research with continuously increasing amounts of information. Both fields contain many topics, which surpass the scope of this thesis. These topics include ARCH and GARCH models, which expand on the traditional risk estimation methods by introducing the concept of conditional variance (Engle 2001). Parametric VaR is also omitted from this study because it relies heavily on various statistical assumptions regarding the underlying characteristics of the data and its distribution (Alexander 2008, 41-42). These assumptions create heavy restrictions that are not satisfied in stressed market conditions. The concept of conditional VaR is also excluded from this study to limit the focus and the scope of the thesis. These delimitations provide a clear focus for the thesis and allow for a more comprehensive analysis of the core concepts of VaR. Finally, the study will primarily examine the market risk observed in the year 2020, which creates additional preliminary limitations for making broad generalizations based on the results.

1.5 The structure of the study

The structure of this study is as follows. Chapter one describes the background of the study and introduces the objective of the thesis. It also defines the research problem, the research questions, and the scope of the study. Chapter two covers the theoretical framework of financial risk management. It also includes a description of the basic statistical and mathematical properties and concepts of financial markets. Chapter three provides an overview of the general VaR methodology and describes the calculation processes of HVaR and MCVaR in detail. Chapter four introduces the data used in this study. It also describes the backtesting method and the construction of the VaR models. Chapter five then presents the backtesting results and discusses possible insights gained from them. It also provides a short outline of the study's limitations. Chapter six ends the study with conclusions and suggestions for future research.

2. Theoretical framework

This chapter presents the main theoretical framework of financial risk management and discusses the mathematical properties of financial markets. This chapter also introduces the various statistical background assumptions that need to be taken into consideration later when constructing the VaR models.

2.1 Financial risk management

Financial risk management can be described as the process of protecting the economic value of a company by identifying and controlling potential risk factors (Christoffersen 2012, 4-6). In other words, the main purpose of financial risk management is to estimate the significance of various risk factors and prepare for negative consequences. This estimation is often done by combining statistical and probabilistic analysis with qualitative assessments of the potential risk factors (Horcher 2005, 206). The process of financial risk management offers valuable information about the different risk factors and potential outcomes that a company can face. This information can be used in strategic decision-making to control the company's exposure to the risk factors that affect it. Furthermore, the management of financial risk can be also done by hedging against the risk factors with different types of investments and financial instruments, such as derivatives (Jorion 2011, 10-11).

The basic theoretical framework of financial risk management can be condensed into a single statement: when everything else is equal, less risk is always preferred over more risk (Bacon 2008, 62). This concept of risk aversion provides a theoretical foundation for managing financial risk. It also highlights the rationale behind controlling the deviation of future outcomes, as unpredictability can be seen as inherently undesirable. Another important concept of financial risk management is the theory of loss aversion. It states that the negative effects of large losses are generally perceived as more significant than the positive effects of equally large gains (Schmidt & Zank 2005). The theory of loss aversion can be seen as a complement to the theory of risk aversion. A high variance in future outcomes implies an increased potential for a significant change without providing information on whether the change is positive or negative. Therefore, higher variance can be seen as an undesirable property, as in the context of financial risk management, negative outcomes have a greater weight than positive outcomes. Even though the concept of loss aversion is fundamentally based on a cognitive bias, its importance must be acknowledged, as it often has a significant role in risk-based decision-making.

Financial risk is usually divided into five distinct categories of risk. These categories are market risk, credit risk, liquidity risk, operational risk, and business risk (Christoffersen 2012, 6-7). This study focuses primarily on market risk, but the other risk categories will also be briefly covered to illustrate the breadth of the subject.

2.1.1 Market risk

The value of a financial asset is defined by its market price, which is determined by the changing levels of supply and demand in the market (Jain 2014). Therefore, market risk can be described as the risk of potential losses that are caused by changes in market prices (Jorion 2011, 22). The changes in market prices are essentially caused by the actions of different market participants and their views on the values of the assets. Hence, the ownership of financial assets creates exposure to market risk, as the value of any given asset can change due to factors that are outside of the investor’s control. Market risk management attempts to evaluate the probabilities of these changes in the market prices and prepare for them. This is done by examining the various market risk factors and constructing statistical models by analyzing the mathematical properties of the market.

The most basic measure of market risk is volatility σ (sigma). It is the standard deviation of market prices calculated either for a single asset or a portfolio of multiple assets using historical data. Volatility captures the essence of market risk by describing the relative magnitude of changes in the value of an investment during a given period of observation. The measure is very simple to interpret, as higher volatility (standard deviation) directly implies higher levels of risk and variability (Bacon 2008, 63). This positive relationship between volatility and risk exists, even though volatility alone communicates very little about the actual shape of the underlying distribution. Volatility, in the context of market risk, can be used in estimating the standard deviation of unknown future outcomes (Jorion 2011, 76). It is also an integral part of basic statistical stochastic analysis. Therefore, various more sophisticated market risk measures such as VaR incorporate it in their calculation methods.

2.1.2 Other types of financial risks

While market risk is the primary focus of this study, it is essential to introduce the other categories of financial risk. This offers a more comprehensive overview of financial risk management and further outlines the scope of this study. The four other risk categories are credit risk, liquidity risk, operational risk, and business risk. This general classification of risks provided by Christoffersen (2012, 6-7) highlights the great variety of risks that a corporation can encounter. While these categories are theoretically distinct, Christoffersen (2012, 7) also emphasizes the fact that the different types of risks can often overlap and share characteristics.

Credit risk is defined as the risk of the counterparty becoming less likely to properly fulfill their contractual obligations (Christoffersen 2012, 7). This definition includes both the risk of a complete default and the risk of partial fulfillment of the obligation. Credit risk is a significant component of all financial contracts, and it is the result of the counterparty either being unwilling or simply unable to fulfill their obligations (Bacon 2008, 61). Uncertainty about the counterparty's ability to meet their obligations also increases credit risk. Therefore, the risk emerging from this lack of certainty is often compensated with a risk premium (Koulafetis 2017, 21).

Liquidity risk is encountered in financial markets with low liquidity (Christoffersen 2012, 6). Market liquidity can be measured by observing the level of trading volumes and the size of bid-ask spreads, where lower trading volumes and larger bid-ask spreads indicate lower liquidity (Christoffersen 2012, 6; Alexander 2008, 392). Conducting market transactions in these low liquidity market conditions can lead to unexpected price actions (Christoffersen 2012, 6). For example, attempting to buy an asset in a market with low liquidity might push the asset's price higher than in a market with high liquidity. This creates additional risk for market participants, as the reduced availability of buyers and sellers can result in an inability to perform market transactions at the mark-to-market prices (Christoffersen 2012, 6; Alexander 2008, 392).

Business risk is defined as the risk of negative changes in outside factors that affect a company's business plan (Christoffersen 2012, 7). This definition of business risk provided by Christoffersen (2012, 7) covers the various quantifiable risks such as the fluctuations in demand and the impacts of the business cycle. It also includes potential changes in the competitive environment and technological advancements that reduce the competitive advantage of a company. Business risks are strongly connected to the company's core business. Hence, exposure to a specific business risk cannot be avoided without changing the company's strategy (Christoffersen 2012, 7).

Operational risk can be described as the risk of losses caused by physical catastrophes, technical failures, or human errors (Christoffersen 2012, 7). Examples of operational risk include fraud, management failure, and digital system outages. These types of risks are typically more common than, for example, a default of a counterparty. However, their consequences are also usually less severe (Bacon 2008, 61). The monitoring of individual operational risks is crucial, as repeating patterns of small risks may indicate something more serious (Bacon 2008, 61). Operational risk is also often treated as a general catch-all category that includes all financial risks that do not fit within the definitions of any other category. This makes operational risk more difficult to manage, as the risks included in the category are not easily quantifiable and come in great variety. Also, operational risk is very difficult to hedge against in the asset market, which has led to the use of self-insurance and third-party insurances among many firms (Christoffersen 2012, 7).

2.2 Mathematical properties of financial markets

There is more data and computational power available for financial analysis than ever before (Consoli, Recupero & Saisana 2021, 4). This has made the use of various mathematical and statistical models increasingly prevalent in the context of market risk management. The increasing supply of market data offers an opportunity to analyze the various mathematical properties and mechanisms of the financial markets empirically. These properties are useful in estimating potential risk factors, as they offer a framework for analyzing the movements and mechanisms of the market. The mathematical properties are often presented as stylized facts and they include concepts such as the absence of autocorrelation, non-normality of returns, and gain-loss asymmetry (Cont 2001).

VaR models are also integrally connected to the mathematical properties of the financial markets, as they provide numerous background assumptions that have to be considered when modeling market risk. Analyzing these properties is, therefore, at the core of the theoretical framework of this study. To conclude, this chapter aims to highlight some of the most basic mathematical properties and stylized facts of financial markets that are connected to the theoretical framework of VaR modeling.

A fundamental principle present in financial market analysis is the concept of observable patterns that can be modeled using various stochastic processes and statistical distributions (Voit, Thirring, Beiglböck, Grosse & Balian 2005, 1-6, 59-60). This assumption implies that there are distinct recurring behaviors and mechanisms in financial markets that can be quantified mathematically. Discovering these patterns can provide tools for creating estimates of future outcomes and modeling risk in financial markets. These patterns are inherently connected to the mathematical properties and stylized facts of the financial markets. Discovering and analyzing these patterns is, therefore, an essential part of financial risk management, as they can be used in estimating future outcomes. Conversely, if such patterns or quantifiable phenomena did not exist, analyzing financial markets and creating estimates based on empirical market data would not produce value for investors. The lack of observable patterns would indicate that the movements of financial markets are completely random and that estimating future outcomes would not be useful.

One of these important mathematical properties of financial markets can be illustrated using an example provided by Christoffersen (2012, 9-10). The example highlights the absence of autocorrelation in financial asset returns, which is also one of the stylized facts introduced by Cont (2001). Autocorrelation is a mathematical measure used to find repeating patterns in data. The presence of autocorrelation means that the observed variables of the data are correlated with delayed versions (lag orders) of the same variables. This example aims to show that the previous returns of financial assets have normally very little correlation with consequent returns. This stylized fact has been established as a mathematical property of the financial markets, and it can be observed and verified by examining the series of returns calculated from the index values of major stock indices. In conclusion, this study will utilize a similar example as Christoffersen (2012, 9-10) to show that the historical returns of the S&P 500 ranging from 1.1.2010 to 31.12.2019 have little to no autocorrelation across the different lag orders:

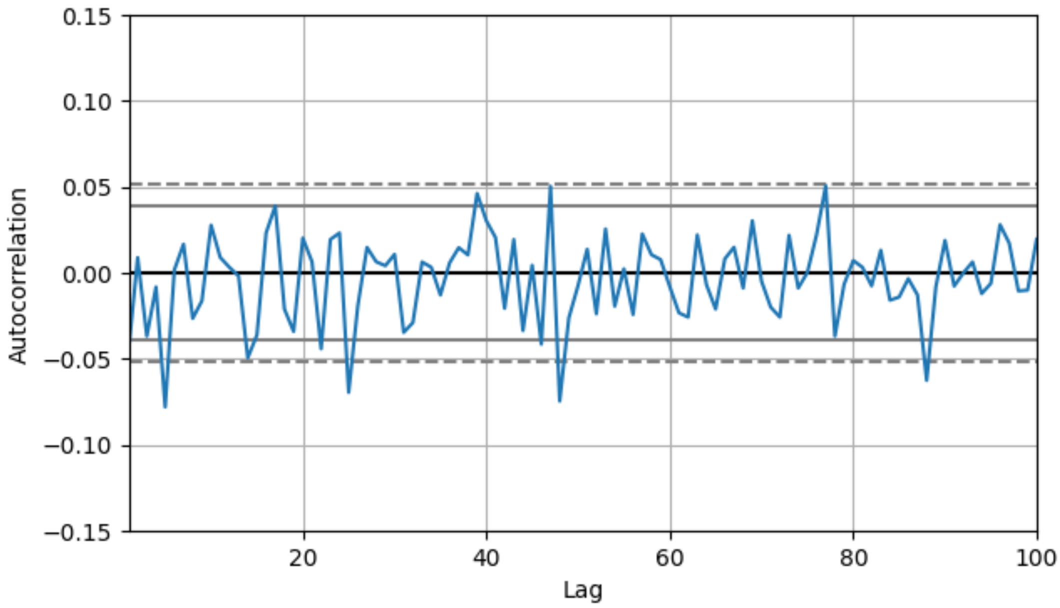

Figure 1: S&P 500 – Autocorrelation of returns (2009-2019)

In Figure 1, the magnitude of delay in the returns is displayed on the x-axis and the level of autocorrelation is displayed on the y-axis. In this example, the amount of correlation between the different lag orders is extremely small. This indicates that the return of a previous day contains very little predictive power when estimating the returns of the consequent days. Even though these types of mathematical properties and historical patterns are based on empirical observations, they can still be used as a theoretical framework for analyzing the behavior of financial markets with high confidence. Furthermore, VaR modeling and general market risk management are also connected to these observable patterns, as they utilize historical information and data to estimate future outcomes. This relationship of VaR and historical observations will be discussed more extensively as the theory of VaR is introduced.

2.2.1 Market prices and returns

The most basic observable variable in the statistical framework of financial markets is the market price. It is the measure of value and the rate of exchange for a given asset determined by the varying levels of supply and demand in the market. Every other mathematical property of financial markets is essentially connected to the market price. Additionally, it is inherently connected to measuring market risk and, therefore, an essential concept in financial risk management.

Another integral concept in market risk management is the return of an investment, which can be calculated using market prices. If P_1 is the initial price of an asset, and P_2 is the price of an asset after the observed period, then the simple rate of return r for that period can be expressed as:

$$r = \frac{P_2 - P_1}{P_1} = \frac{P_2}{P_1} - 1 \tag{1}$$

The simple rate of return is the most widely used method of calculating returns. It is easy to interpret, and the calculation process is straightforward. However, instead of simple returns, this study will use another calculation method called geometric returns (or log-returns). Geometric returns are functionally similar to simple returns, but they offer some properties that make them more attractive in the context of statistical analysis. The formula for the geometric return using the same prices can be written as:

$$r = ln(P_2) - ln(P_1) = ln\left(\frac{P_2}{P_1}\right) \tag{2}$$

This method of calculating returns is advantageous in many ways. Most importantly, using geometric returns improves the reliability and robustness of various statistical risk measures, such as standard deviation, semi-variance, VaR, and expected shortfall (Miskolczi 2017). This is the result of geometric returns removing the positive biases that are included with simple returns (Bacon 2008, 29). Also, they make the calculation of multiperiodic returns more straightforward, as the single period geometric returns are arithmetically additive (Miskolczi 2017). This means that the compounded multiperiodic returns can be calculated as a sum of the single period returns:

$$\sum_{i}^{n} ln(1 + r_i) = \ln(1 + r_1) + \ln(1 + r_2)\ln + \cdots + \ln(1 + r_n) = \ln\left(\frac{P_n}{P_1}\right) \tag{3}$$

The final advantageous property of logarithmic returns compared to simple returns is related to statistical distributions. If the single period geometric returns are normally distributed, then the compounded multi-period geometric returns, calculated by summing the single period returns, are also normally distributed. This is based on probability theory, which states that the sum of normally distributed independent variables is also normally distributed (Lemons 2002, 34). Simple returns do not have this beneficial property, as the compounded multiperiodic simple returns are determined by calculating the product of the single period returns.

In conclusion, market prices and returns provide a substantial amount of statistical information that can be used in financial market analysis and portfolio optimization. Effective utilization of this data and information is at the core of market risk management. Market prices and returns are essential to the theoretical framework of this study, as they are the basic variables of market risk estimation. Historical returns are used as input variables in the VaR models when estimating risk and analyzing the probabilities of various future outcomes. Therefore, understanding the statistical properties and calculation processes of returns is necessary to effectively interpret the estimates produced by the models.

2.2.2 Statistical distributions

Statistical distributions represent the frequency of individual outcomes for a given phenomenon. The distributions describe the probabilities of various quantifiable events, and they can be based on historical observations or theoretical models. In the context of risk management, statistical distributions are an important part of exploring the mathematical properties of financial markets. The most common use of these distributions in financial market analysis is to examine the past returns of investments. The study of return distributions can offer an overview of the historical performance of an asset or portfolio that can be used in various applications, such as estimating expected returns and calculating VaR measures. In the context of VaR estimation, the return distributions are often used in the calculation of the underlying probability density functions that are used in estimating the future portfolio values (Jorion 2011, 107-110). This method of estimation is connected to the framework of repeating statistical mechanisms and quantifiable patterns, which were discussed earlier. While this process of calculating probability distributions using historical observations can help with simplifying the construction of future estimates and statistical models, it relies heavily on various assumptions that can produce significant inaccuracies in the models. As with all statistical analysis, it is crucial to keep in mind that historical data communicates only a limited amount of information about the future and that the true accuracy of the distribution will depend on numerous factors. For example, limited availability or poor quality of data may, in some cases, lead to the construction of statistically erroneous models, as the sampled observations might not correctly represent the underlying phenomena. This concept of exploring statistical distributions in the context of financial market analysis and market risk management can be illustrated with an example.

To provide a more concrete understanding of the role of distributions in financial markets, this example will explore the daily returns of the S&P 500 stock index ranging from the start of 2010 to the end of 2019. These returns are presented as a histogram along with a graph that represents the normal (gaussian) distribution in Figure 2.

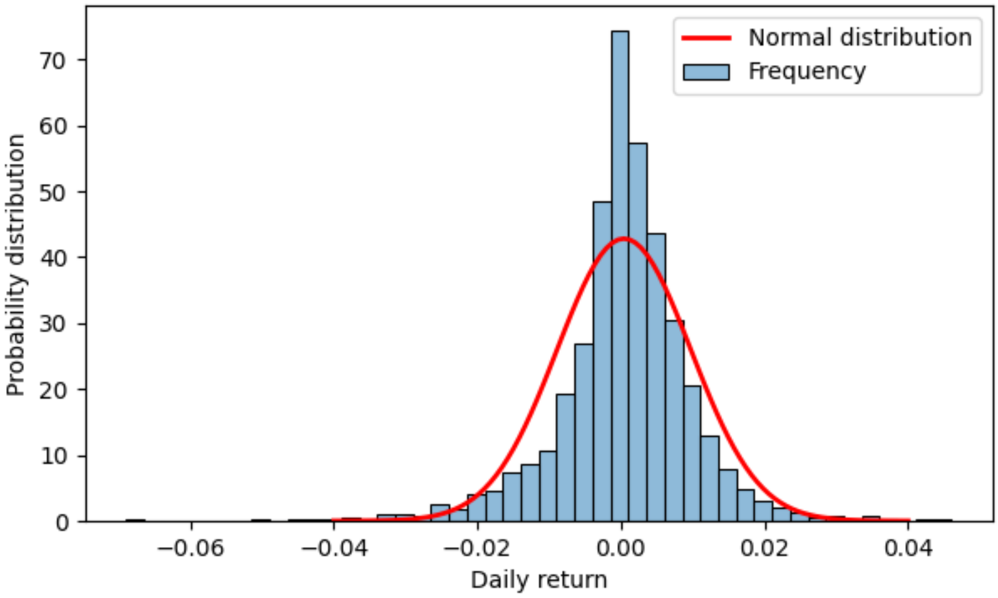

Figure 2: S&P 500 – Histogram of returns (2010-2019)

Multiple mathematical and statistical observations can be made by examining this distribution of the historical returns of the S&P 500 shown in Figure 2. Most importantly, it can be concluded that the distribution of historical returns is different from the normal distribution. The historical distribution of the S&P 500's returns exhibits significantly higher kurtosis than the normal distribution. This observation is established by analyzing the peak of the distribution presented in Figure 2, which indicates that the returns of the S&P 500 are more heavily concentrated around the mean than in the normal distribution. Furthermore, the actual value of kurtosis calculated for the historical returns of the S&P 500 is 4.60, while the kurtosis of the normal distribution is 3.00. This implies that while most of the returns are heavily concentrated around the average return, extreme returns are more likely to occur because of the "fat tails" caused by the excess kurtosis. The presence of the additional tail-risk can also be observed in Figure 2 by looking at the most extreme historical returns. For example, there is a single return with a value of -6.90 %, and given the mean and standard deviation of the empirical returns, the probability of observing such a value in normally distributed data is $3.9224 * 10^{-11}$. This type of analysis of the return distributions is an important part of the theoretical framework of this study, as it is the basis of analyzing the probability of future outcomes. The statistical distributions are also essentially connected to the theory of VaR analysis. The relationship between the historical returns and the risk estimates will be discussed more in the next chapter.

3. Value-at-Risk models

This chapter introduces the general theory of VaR modeling, which will be used as a foundation for the empirical section. It also includes an overview of the statistical properties and the mathematical notation of the VaR measures. The methodology behind both HVaR and MCVaR will be discussed individually.

VaR models are statistical risk measures that estimate the worst loss over a given time horizon at a given confidence level (Bacon 2008, 100). The time horizon t and the confidence level c are the two fundamental parameters that need to be specified when calculating VaR. The time horizon is defined as the period over which the VaR estimate is calculated, and the confidence level is defined as the level of statistical certainty of the estimate (Illowsky & Dean 2018, 445). Confidence levels can be best described by their relationship to the statistical significance levels. If the significance level α of a statistical estimate implies the probability of an erroneous measure, then the confidence level can be expressed as $c = 1 - \alpha$. In other words, the confidence level describes the probability of the VaR estimate being statistically correct. VaR for a given time period t and a confidence level c is denoted $VaR_{t,c}$. Therefore, a VaR measure with the parameters $t = 1$ and $c = 0.95$, for example, can be written as $VaR_{1,0.95}$.

The actual value of VaR can be expressed both as a monetary value and as a relative percentage of the initial value. If the estimated VaR for a portfolio with an initial value of 100 000 € is 10 000 €, then the value of VaR can be equally expressed as $VaR = 10\ 000$ € or $VaR = 0.1 = 10$ %. Additionally, the role of the parameters t and c can be described by including them in the previous example. If the VaR estimate and the portfolio value are the same as before, the parameters $t = 1$ and $c = 0.95$ can be interpreted as follows: There is a 95 % probability that the maximum loss endured in a single day will not exceed 10 000 € or 10 %. In other words, the loss will be greater than 10 000 € or 10 % in less than five days out of one hundred days.

The mathematical expression of VaR can be defined with the notation provided by Jorion (2011, 108-109). If the initial value of a portfolio is marked as W_0, then the portfolio value at the end of the time horizon t can be written as $W_t = W_0(1 + R_t)$. In this equation, the rate of return R_t is a random variable with an expected value μ and a standard deviation of σ. If the probability density function (pdf) of the future portfolio value W_t is written as $f(w)$, then VaR can be expressed using the following equation:

$$c = \int_{VaR}^{\infty} f(w)\, dw \tag{4}$$

In Equation (4), VaR is the worst possible return that the portfolio will endure at the confidence level c. The pdf $f(w)$ represents the statistical probabilities of the various future outcomes that define the estimated distribution of portfolio values. Therefore, the integral of this pdf can be treated as a cumulative distribution function (cdf) of the future portfolio values. The pdf must be determined in order to calculate the VaR estimate, as the actual value of VaR is denoted in the equation as the lower limit of the integral. In other words, VaR can be only calculated if the distribution of future portfolio values can be statistically estimated or defined by using historical observations. Furthermore, the probabilities of all returns higher than VaR sum up to the confidence level c, which highlights its mathematical role in the notation. Most importantly, this notation provided in Equation (4) emphasizes the connection between VaR estimation and the statistical distributions that are analyzed when calculating the values of VaR. The same equation can also be denoted using the significance level α:

$$\alpha = \int_{-\infty}^{VaR} f(w) dw \tag{5}$$

Based on Equations (4) and (5), the variable VaR can be treated as a quantile of the distribution of future portfolio values (Jorion 2011, 109). It is also important to note that while both equations are used to define the VaR mathematically, they do not offer any insight into the available methods of calculating the probability density function $f(w)$. However, as discussed in the theoretical framework, the underlying statistical distribution is commonly determined by using historical observations. If the distribution of future outcomes is known, the actual calculation of VaR becomes very simple, as it can be determined by looking for the value at the quantile of the distribution that corresponds to the value α. This can be illustrated graphically by utilizing a randomized sample of returns:

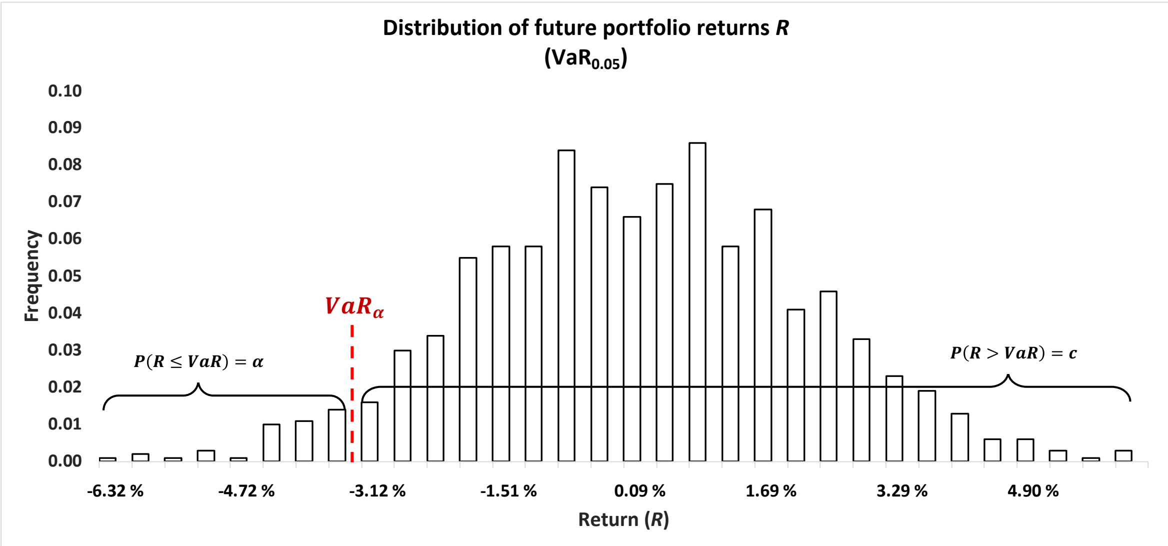

Figure 3: Graphical illustration of VaR using a randomized sample

This visual interpretation of VaR presented in Figure 3 communicates the basic function of the estimates. The x-axis represents the range of future portfolio returns and the y-axis displays the probability of them occurring. Therefore, the histogram presented in Figure 3 can be interpreted as the pdf that is used in the estimation of VaR. In other words, the models aim to determine the cutoff point (VaR_{α}) in the distribution of returns, under which the worst α-quantile of observations lie. In conclusion, VaR models are statistical and analytical tools that utilize statistics and probability theory to calculate the magnitude of potential losses by estimating the distribution of future outcomes. The next chapters will address the actual methods of calculating the underlying distributions and pdfs, which are used to determine the value of VaR.

3.1 Historical Value-at-Risk

Historical Value-at-Risk (HVaR) is the most widely used form of VaR modeling (Christoffersen 2012, 21). It is relatively simple to use and interpret, as it does not require heavy computing or complex stochastic calculations. It estimates the maximum probable loss for a given asset or a portfolio of assets by examining the distribution of their historical returns (Jorion 2011, 262-263; Alexander 2008, 42-43). This means that the estimate is based completely on historical data and observations. Consequently, it relies strongly on the assumption that history will repeat itself, and that past performance is linked to future outcomes (Cheung & Powell 2012a).

The calculation of HVaR is done by taking the range of historical returns of a given time series and sorting them in ascending order. This creates the historical distribution that the estimate will be determined from. After sorting the returns based on their magnitude, the HVaR estimate can be calculated by finding out the specific quantile value that corresponds to the value of α. This process of calculating HVAR can be illustrated with an example that incorporates the same randomized returns that were used in the graphical illustration of VaR in Figure 3. The first step of calculating the estimate is to collect the historical observations into a sorted table. This sample of historical returns presented in Table 1 consists of 1000 observations that range from -6.32 % to 6.10 %.

Number of sample	*Sample return*
1	-6.32 %
2	-6.14 %
3	-5.99 %
⋮	⋮
999	6.01 %
1000	6.10 %
Total samples (N)	1000

Table 1: Sorted sample returns

After sorting the data, the calculation of the HVaR estimate can be completed by searching for the sample with a quantile that corresponds to the prespecified value of α. In this example $\alpha = 0.05$ and $N = 1000$. Therefore, the sample number and the index of the sample that matches the α-quantile is $\alpha * N = 50$. After determining the corresponding index, the VaR estimate can be determined by looking for a return with this index. In this example, the lowest α-quantile of returns can be found below the 50th sample. The actual value of VaR is therefore the return associated with the 50th lowest sample, which is -3.31 %. However, in the common case where $\alpha * \mathrm{N}$ is not an integer that can be matched to a specific index, the α-quantile is calculated using interpolation. This study also utilizes a method of calculation that excludes the quantiles from 0 to $\frac{1}{N+1}$ and $\frac{N}{N+1}$ to 1 to better incorporate the traditional definition of quantiles. The process of finding the value of VaR also becomes easy to interpret, as it can be phrased into a practical question: "Under what value, lie $\alpha * 100$ percent of the returns?" Finally, the relationship between the example HVaR estimate and α is illustrated in Equation (6) using the general mathematical notation of VaR given in Equation (5):

$$0.05 = \int_{-\infty}^{-3.31\%} f(w)dw \tag{6}$$

3.2 Monte Carlo Value-at-Risk

Monte Carlo Value-at-Risk (MCVaR) modeling is based on Monte Carlo simulation. The core idea of Monte Carlo simulation is to solve mathematical and statistical problems by modeling a random process repeatedly in order to approximate the solution (Dimov 2008, 1). MCVaR applies this method of simulation to financial variables and their risk factors to produce risk estimations (Jorion 2011, 307). Therefore, the basic idea of MCVaR is to design a random stochastic process that represents the risk factors of an asset or a portfolio. The random process then draws its values from a prespecified probability distribution, which is sampled a large number of times. In the context of VaR estimation, Monte Carlo simulations are used to construct a set of artificial future outcomes that represent the distribution of the portfolio value W_t. By this definition, MCVaR approximates the probability density function of potential profits and losses through a very large number of iterations. The number of repeated iterations and simulations has a crucial role in reducing errors and increasing the accuracy of the models (Jorion 2011, 316). In most cases, a high number of simulations such as 10 000 (or higher depending on the variables and confidence level) will reduce the role of inaccuracies sufficiently (Jorion 2011, 316-319). After calculating the simulations, the actual VaR estimate can be discovered by examining the simulated distribution and finding out the value that corresponds to the quantile specified by α. The last step is, therefore, similar to the HVaR estimation, as the only difference is that the source of the distribution of future portfolio values is based on simulations instead of historical observations.

MCVaR has recently become more popular, as the computation of complex processes has become more available due to technological advancements. The main benefit of Monte Carlo simulation is the possibility to design a random process that can represent almost any type of risk factor or a combination of risk factors. However, from this possibility, emerges the greatest challenge of using MCVaR. The actual construction of the various stochastic processes can be challenging when complex risk factors are present. Common variations of MCVaR utilize the concept of geometric Brownian motion (GBM) to estimate the future values and price paths of financial assets and portfolios (Cheung and Powell 2012b). The mathematical notation of GBM is presented in the following equation:

$$S_{\mathrm{t}+\Delta\mathrm{t}} = S_{\mathrm{t}}\, e^{\mathrm{k}\Delta t+\sigma\varepsilon_t\sqrt{\Delta t}} \tag{7}$$

In Equation (7), S_t is the price of the asset at time t, Δt is the added time increment, e is the natural logarithm, k is the expected return of the asset, and ε_{t} is the added randomness at time t that affects the assets market price. Most commonly the value of ε_{t} is drawn from the standard normal distribution with parameters μ = 0, σ = 1. However, as discussed in Chapter 2, the asset returns rarely follow a normal distribution. Therefore, the MCVaR models can be modified to incorporate various statistical distributions from which the variable ε_t is drawn. This flexibility highlights the main advantage of MCVaR; it is customizable to account for the specific risk of a given scenario. To conclude, this process of calculating the changes in the assets market price with GBM can be used to simulate the probabilities of future returns. This study will utilize a simple form of GBM and MCVaR simulation to approximate the returns of the indices in a single variable estimation, which will be described in the next chapter.

4. Data and the research method

This chapter describes the data and the research method used in the study. It will also provide additional discussion about the statistical properties of the data. After introducing the data, there will be a theoretical overview of the backtesting method. The chapter will end with a summary of the calculation processes and a brief commentary on the construction of the VaR models.

4.1 Market data

The data used in the study consists of daily market prices of three popular stock indices ranging from 31.12.2009 to 31.12.2020. These prices are used to calculate the returns that are used in the construction of the models and backtesting. These returns range from 1.1.2010 to 31.12.2020, as the calculation of a return requires the price of that day, as well as the price of the previous day. The indices chosen for this study are S&P 500, STOXX Europe 600, and MSCI Emerging Markets Index. These indices are chosen based on their location and market capitalization. Each of the three indices tracks the performance of a geographically distinct market, which helps with reducing the potential problems that emerge from the geographical properties of the markets. The markets included in the data are the U.S. (S&P 500), Europe (STOXX 600), and emerging markets (MSCI EM). This data selection offers a comprehensive picture of the overall market reaction to covid-19, as it includes the unique characteristics of the individual markets. Also, the chosen indices together cover a high percentage of the total global market capitalization, which reduces the randomness caused by various anomalies that may affect a single market sector.

The S&P 500 consists of the 500 largest companies in the U.S. exchanges. The stocks in the S&P 500 index are weighted by their market capitalizations and they collectively cover approximately 80 % (13.50 trillion USD) of the U.S. stock market (S&P Global 2021). The STOXX 600 tracks the performance of a wide range of companies listed in European exchanges. It includes a fixed number of small-, mid-, and large-cap companies that collectively cover approximately 90 % (12.96 trillion USD) of the market capitalization of the European stock exchanges (Qontigo 2021). The MSCI EM index tracks the performance of various companies listed in emerging markets' exchanges. The index covers approximately 85 % (7.94 trillion USD) of the market capitalization of the various emerging markets countries. The countries that have the largest weights in the emerging markets index are China, Taiwan, South Korea, India, and Russia (MSCI 2021). The training data will consist of the daily geometric returns of the three indices ranging from 1.1.2010 to 31.12.2019. This data will be used to model VaR estimates during the testing period ranging from 1.1.2020 to 31.12.2020. Descriptive statistics for the testing data consisting of the daily returns of the three indices are presented in Table 2:

Returns (1.1.2010 - 31.12.2019)			
Index	*S&P 500*	*STOXX 600*	*MSCI EM*
Mean	0.042 %	0.019 %	0.005 %
Median	0.060 %	0.053 %	0.049 %
Minimum	-6.896 %	-7.293 %	-6.524 %
Maximum	4.840 %	6.907 %	4.810 %
Standard Deviation	0.932 %	0.998 %	0.955 %
Kurtosis	4.600	4.308	2.733
Skewness	-0.497	-0.343	-0.394
Count	2516	2566	2607

Table 2: Descriptive statistics of the testing data

These descriptive statistics provide interesting information about the indices' historical returns and their statistical distributions. Each of the three indices has a different number of observations, as the different markets are open and closed on different days. This difference is small, and it will not negatively affect the quality of data. All three indices have slightly positive mean and median values, which implies that the returns have on average been positive during the period of observation. The S&P 500 and MSCI EM have similar minimum and maximum values, while the STOXX 600 index has produced both the smallest minimum return and the largest maximum return. Another important observation can be made from the kurtoses of the indices. Both S&P 500 and STOXX 600 have a very high kurtosis, which implies that they carry significant tail risk. For example, the normal distribution has a kurtosis value of three, which indicates that both S&P 500 and STOXX 600 have a higher probability of returns with extreme values (Bacon 2008, 85). This is not the case with MSCI EM, as it has a smaller kurtosis than the standard normal distribution. The role of kurtosis is also important in the context of VaR analysis, as the high chance for extreme events can produce inaccuracies in the models. Finally, each of the three indices has negative skewness. This observation along with the kurtoses highlights that the models are dissimilar to the normal distribution, which cannot be relied on when estimating the risk measures.

By examining the descriptive statistics of the indices presented in Table 2, it can be concluded that the historical returns and statistical distributions of the three indices exhibit similar characteristics. This similarity is also illustrated graphically in Figure 4:

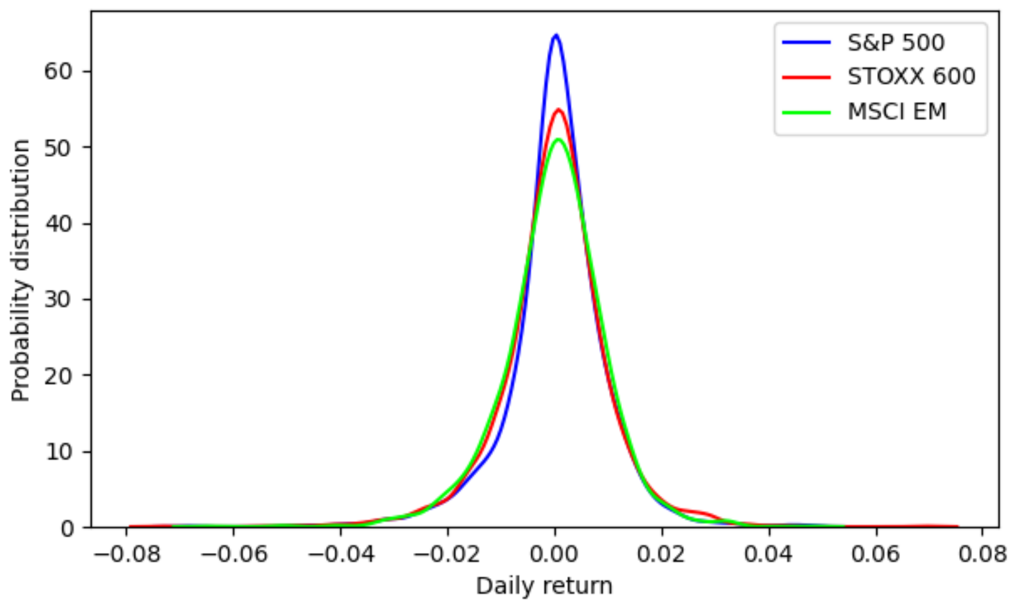

Figure 4: Probability distributions of S&P 500, STOXX 600, and MSCI EM

The three different graphs represent the statistical distributions of the three indices during the period of 1.1.2010 – 31.12.2020. These distributions are created by estimating the probability density functions of each index using kernel density estimation. In conclusion, the three distributions are visually very similar.

A brief correlation analysis was also conducted to further explore the role of similarity between the S&P 500, STOXX 600, and MSCI EM. The correlation coefficients of the indices' daily returns are presented in Table 3:

Correlation of returns (1.1.2010 - 31.12.2019)			
Index	*S&P 500*	*STOXX 600*	*MSCI EM*
S&P 500	1	0.6293	0.4696
STOXX 600	0.6293	1	0.6297
MSCI EM	0.4696	0.6297	1

Table 3: Correlation coefficients of the testing data

The correlation coefficients of the indices fall between the range of 0.40-0.69. This means that the correlations between the indices can be generally interpreted as moderate (Schober, Boer & Schwarte 2018). The role of coefficients is significant when analyzing the backtesting results, as they can account for the similarities between the performances of the VaR models.

4.2 Backtesting methodology

This study uses the Proportion of Failures (PoF) backtesting method first introduced by Kupiec (1995) to analyze the performance of the VaR models. The basic idea of the PoF method is to compare observed estimation failures x and accumulated sample size n against the theoretical distribution of failures. Following the notation provided by Kupiec (1995), the hypotheses for PoF backtesting can be written as:

$$H_0: p = p^*$$

$$H_1: p \neq p^*$$

where p^* is the probability of a failure under the null hypothesis and p is the ratio of observed estimation failures and total samples x/n. In the context of VaR analysis, the variable p^* is specified by the significance level α, and it follows the χ^2 distribution with 1 degree of freedom. Therefore, the acceptance of the null hypothesis H_0 assumes that the theoretical failure ratio p^* and the observed failure ratio p are statistically equal. Conversely, if H_0 is rejected, the alternative hypothesis H_1 will be accepted and the variables are concluded as statistically unequal. It is important to note that the statistical test used in the backtesting is two-sided. Therefore, the null hypothesis is rejected if the observed variable p is either too large or too small. In practice, this means that the accurate VaR models cannot underestimate or overestimate the market risk, which creates additional requirements for the models' performance. Kupiec (1995) concludes in the same study that the most powerful test statistic for testing the null hypothesis H_0 that consists of two ratios is the likelihood ratio (LR) test, which is given in Equation (8):

$$LR = -2\text{Log}[(1 - p^*)^{n-x}(p^*)^x] + 2\text{Log}[(1 - [x/n])^{n-x}(x/n)^x] \tag{8}$$

As the test statistic of PoF backtesting is based on the χ^2 distribution, the acceptance and rejection of the null hypothesis can be determined by comparing the LR test statistic to the corresponding critical value calculated from the χ^2 distribution. In other words, the null hypothesis is rejected if:

$$LR \geq F^{-1}(1 - \alpha) \tag{9}$$

and accepted if:

$$LR < F^{-1}(1 - \alpha) \tag{10}$$

where the F^{-1} is the inverse cumulative distribution function of χ^2 with one degree of freedom.

In the context of the backtesting conducted in this study, the process of calculating the test statistics and determining the acceptance or rejection of the null hypothesis is as follows. First, the VaR model will be constructed and used to estimate risk over a given period of days. After the estimated risk measures are collected, they are compared to the actual returns observed during the same period. These steps determine the number of failures x, which can be divided by the total observations n in order to calculate the variable p. After calculating the variables, they are then used along with $p^* = \alpha$ to calculate the test statistic LR for the same period. This test statistic is then compared to the theoretical χ^2 statistic to determine whether or not the VaR model was able to measure and forecast accurately and reliably the daily market risk during the period of backtesting.

4.3 Construction of the VaR models

The VaR models are constructed using a time horizon $t = 1$ and a confidence level $c = 0.95$. The significance level α and the probability of failure p^* used in the calculation of the backtesting statistics are both 0.05. There are three indices, which will be tested using both HVaR and MCVaR during the year 2020.

The VaR measures used in the backtesting are obtained by calculating an individual VaR estimate for each testing day using the previous $N = 1000$ days as historical samples. This sample size of 1000 observations is chosen based on the recommendation made in Basel II, which states that at least three years of historical data should be used to cover a sufficient range of economic conditions (BCBS 2005). Furthermore, the accuracy of the VaR measures can be also increased by using a high number of samples, as the calculation of quantiles becomes more precise when the tails of the distribution include more data points (Alexander 2008, 145). Additionally, to explore the effect of sample size, the models will be backtested using smaller sample sizes ranging from six months to two years of historical observations. The smaller sample sizes used in the alternative models are $N = 125, 250 \,\&\, 500$. This allows for testing the performance of the models more comprehensively while assessing how the reduction of available data affects the models' accuracy and ability to react to changes in market conditions.

The practical implementation of the sampling method is as follows. If the first testing day is t_0 and the historical samples used in the model range from t_{-N} to t_{-1}, then the next testing day t_1 will use data ranging from t_{-N+1} to t_0. The previous testing day will, therefore, always become the most recent historical observation included in the calculation of the next VaR estimate. This sampling process forms a rolling sampling window with N observations that moves along with the testing day. This method of modeling by using rolling windows represents the real-world use of VaR, as the calculation of the measures is done daily by using the latest data available. The historical observations used in the estimation of each VaR measure will be different. This process of sampling provides flexibility for the models, as they can react to new information provided by the market.

The HVaR model utilizes the rolling window of samples as historical returns that are used in the calculation of the sorted α-quantile. In the case of MCVaR, the same window of returns is used in performing the simulations. The actual method chosen for simulating the distributions of future returns with MCVaR utilizes kernel density estimation (KDE). The concept of KDE is based on estimating the pdf of a variable in various univariate and bivariate applications (Weglarczyk 2018). In the context of VaR modeling, KDE can be utilized in estimating the pdf of future portfolio returns. In practice, this means that the rolling sample window will be used to calculate an empirical pdf, from which the future returns are sampled. The number of simulations used in the sampling of each MCVaR estimate is 100 000. This is done to ensure a smooth distribution of estimated returns and to reduce possible model inaccuracies.

5. Results

This chapter presents the results of the study. It also provides commentary on the most significant insights gained from these results. The limitations of the study are also presented at the end of this chapter. The process of calculating the VaR estimates and test statistics along with the collection of data is described in Chapter 4.

The HVaR and MCVaR models were backtested using the index returns of S&P 500, STOXX 600, and MSCI EM. These index returns were observed during a period ranging from 1.1.2020 to 31.12.2020. The significance level used throughout the testing was $\alpha = 0.05$. Both models were tested individually with each of the three indices, using the same sampling technique. This resulted in six statistical tests, which consisted of PoF test statistics that were compared to the corresponding χ^2 values to either reject or accept the null hypothesis. These tests aimed to examine the accuracy and the performance of the VaR models in order to determine whether they were usable and reliable risk measures during a period of stressed market conditions. The primary backtesting results are presented in Tables 4 and 5. The tables contain the backtesting results that were obtained during the year 2020 using a sampling window of $N = 1000$ for both the HVaR and MCVaR.

Sampling window N = 1000, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	32	0.126	22.298	3.841	Rejected
STOXX 600	258	31	0.120	19.530	3.841	Rejected
MSCI EM	262	27	0.103	12.045	3.841	Rejected

Table 4: HVaR backtesting results 1.1.2020 – 31.12.2020

Sampling window N = 1000, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	30	0.119	18.396	3.841	Rejected
STOXX 600	258	30	0.116	17.660	3.841	Rejected
MSCI EM	262	26	0.099	10.526	3.841	Rejected

Table 5: MCVaR backtesting results 1.1.2020 – 31.12.2020

Table 4 provides the backtesting results that were obtained using HVaR with the parameters $t = 1$ and $\alpha = 0.05$. Table 5 presents the testing results calculated using MCVaR with the same parameters. None of the six statistical tests was accepted, which indicates that both models failed to model the market risk during the period of backtesting. The PoF statistics calculated for each of the tests were significantly larger than the corresponding χ^2 values, which were determined using the inverse χ^2-function with one degree of freedom. The column x/n in both Tables 4 and 5 represents the ratio between the VaR model failures and the number of observations. By examining the values in this column, it could be concluded that the failure rates of the VaR measures exceed the prespecified value of $p^* = 0.05$. This observation implies that both HVaR and MCVaR underestimated market risk during the testing period of 1.1.2020 – 31.12.2020.

The performance of the models was overall very similar. However, during the primary backtesting using the sampling window $N = 1000$, the risk estimates provided by MCVaR had slightly better accuracy. This was determined by looking at the number of failures presented in column x in Tables 4 and 5. In each of the primary backtests, MCVaR had one or two fewer failures compared to HVaR. This observation was not particularly significant, as the improvements obtained using the MCVaR model were relatively small compared to the total number of failures. Furthermore, the reductions in the model's failures did not lead to the acceptance of any of the null hypotheses. In conclusion, both models exhibited poor performance and did not appear to be usable or reliable in the stressed market conditions observed in 2020.

5.1 Alternative sampling windows and testing periods

Both HVaR and MCVaR were also backtested using smaller sampling windows. The alternative windows used in the secondary backtests were $N = 125, 250$ & 500. The smaller window sizes were chosen to represent the varying availability of historical data used in market risk estimation. The sampling windows approximated trading periods of six months, one year, and two years. The secondary backtesting was conducted similar to the first set of tests by comparing the estimation failures of the VaR models to the theoretical distribution of expected failures. The secondary VaR estimates and test statistics were also obtained using the same calculation methodology that was described in Chapter 4. The only changing parameter between the secondary and the primary tests was the size of the sampling window N. The full results of the 18 statistical tests obtained using the alternative sampling windows are presented in more detail in Appendices A and B. They include the number of failures and the values of the PoF test statistics in the same format as Tables 4 and 5. Table 6 provides an overview of these secondary backtesting results using the modified sampling windows. It also illustrates whether or not the null hypothesis of a given test was accepted or rejected.

Sampling window N = {125, 250, 500}, significance level α = 0.05						
Method	HVaR			MCVaR		
N	125	250	500	125	250	500
S&P 500	Accepted	Rejected	Rejected	Accepted	Rejected	Rejected
STOXX 600	Accepted	Accepted	Rejected	Accepted	Accepted	Rejected
MSCI EM	Accepted	Accepted	Rejected	Accepted	Accepted	Rejected
Total failures	177			170		
Sum of PoF	35.90			29.31		

Table 6: HVaR & MCVaR backtesting results 1.1.2020 – 31.12.2020

The results presented in Table 6 showed an inverse relationship between the size of the sampling window and the accuracy of the model. These results were surprising, as they indicated that limiting the amount of market data seemed to improve the performance of the models' significantly. By only examining whether a given backtesting hypothesis was accepted or rejected, another interesting phenomenon could be observed from the results;

the size of the sampling window seemed to have a greater impact on the performance of the risk models compared to the type of model that was used. This observation was in line with the case study provided by Alexander (2008), where he showed that the chosen sample size affected the estimated VaR measures more than the decision between using a historical or a parametric model. While the results obtained from this study did not provide very specific insight into the actual function of the sampling window, the overall effect of the window size was noteworthy. The models were also compared by examining the total failures and the sum of the PoF statistics during the backtesting period. Again, the results indicated that MCVaR seemed to produce slightly more accurate risk estimates, as both the number of failures and the sum of the PoF statistics were smaller. To conclude, the differences highlighted between the models were relatively small and this type of comparison of the test statistic did not have any statistical significance outside of speculation.

The results provided in Tables 4, 5, and 6 suggested that the reliability and predictive ability of the models depended heavily on the size of the sampling window. When a sample size of four years ($N = 1000$) was used, the models underestimated market risk heavily. This could be explained in part by the reduced flexibility and the decreased ability to react to new information caused by the relatively large sample size in both HVaR and MCVaR. In practice, this inflexibility could lead to inaccuracies and create problems for various parties that utilize VaR in their risk estimation. The models' reduced ability to react to changing market conditions with sufficient speeds would lower their reliability and usability in day-to-day market risk estimation. Also, another problem caused by the same mechanism of large sampling windows is that the VaR measures could be elevated for a long time after the actual condition that caused the increase in market risk would be gone. This would lead to overly conservative VaR measures, which is also an undesirable property of a risk measurement tool when managing market risk. The concept of overestimation was also emphasized by Berkowitz and O’Brien (2002) in their study of the six large U.S. banks, in which the VaR models did not fail due to underestimation, but overestimation of market risk during a long period of observation. On the contrary, the inclusion of a large sample has also its positives. It may reduce the negative effects of short-term anomalies and stabilize the risk estimates calculated over a long horizon. Large sampling windows can also capture a wider range of economic conditions that may better represent the riskiness of financial markets.

Finally, to finish the overview of the backtesting results and to provide a more comprehensive understanding and a context for the earlier backtesting results, the study will provide a third set of results showcasing the performance of HVaR and MCVaR outside of the year 2020. This set of results provides an overview of models' performance during the years 2015 – 2019. The sampling window $N = 1000$ and significance level $\alpha = 0.05$ were used in the backtesting. The full results of the 30 statistical tests are again provided in more detail in Appendices C and D, while Table 7 presents an overview of the results.

Sampling window N = 1000, significance level α = 0.05						
Method	HVaR			MCVaR		
Year	*S&P 500*	*STOXX 600*	*MSCI EM*	*S&P 500*	*STOXX 600*	*MSCI EM*
2019	Accepted	Accepted	Rejected	Accepted	Rejected	Rejected
2018	Rejected	Accepted	Rejected	Rejected	Accepted	Accepted
2017	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected
2016	Accepted	Accepted	Accepted	Accepted	Accepted	Accepted
2015	Accepted	Rejected	Accepted	Accepted	Rejected	Accepted
Total failures	192			179		
Sum of PoF	96.78			91.15		

Table 7: HVaR & MCVaR backtesting results 2015 – 2019

The results presented in Table 7 illustrate the performance of the models during the years 2015 – 2019. The table contains information on whether the null hypothesis of a given statistical test was accepted or rejected. Additionally, the rejections have been divided into two categories using red and blue colors. If the null hypothesis was rejected due to exceeding the allowed number of failures during the testing period, the result was highlighted red. However, if the individual test was rejected due to the number of failures being too small, the result was highlighted in blue. This emphasized a critical mechanism of the VaR models over a longer time period; the models tended to more commonly overestimate than underestimate the market risk. The backtesting indicated that the HVaR model had eight accepted and seven rejected null hypotheses. Of these seven rejections, four were due to an overestimation, and three were due to underestimation of the risk. In the case of MCVaR, the results were once again marginally better. It had the same number of accepted hypotheses, however, it had one less year of exceeding the prespecified limit of failures. On the other hand, it had one more year compared to HVaR where the model overestimated market risk and produced too pessimistic risk estimations. Furthermore, the rejections that were caused by breaching the maximum limit of model failures could generally be perceived as more negative than the rejections that are caused by very small numbers of failures. Despite this, it is still important to treat the overestimation of market risk as a negative feature of the VaR models. In conclusion, the models exhibited similar performance during the years 2015 – 2019. The results were also in line with the two earlier sets of results, where the MCVaR model showcased a slightly better ability to estimate the market risk compared to HVaR.

5.2 Limitations of the study

This section provides a brief discussion about the limitations of the study. It includes a short commentary about the possible restrictions that were created by the data and the methods used in the study. It also touches on the reliability of the testing results.

The study was conducted using the index values of three popular stock indices during a specified time period. Even though the indices collectively covered a large proportion of the global financial markets, this selection of data still created the most significant limitation of the study. The three indices used in the study did not provide enough evidence to form conclusive statements about the results. Also, as highlighted in Chapter 4, the indices had moderate levels of correlation. Therefore, similarity in the backtesting results was expected. The negative impact of the correlations could have been minimized by including a larger sample of testing data. For example, the performance of the VaR models might have varied if they would have been used to estimate the risk of a portfolio with a different selection of assets. Furthermore, it would have been interesting to compare portfolios with significantly different levels of volatility.

The method of utilizing the different sampling windows also created some additional limitations in the results. VaR models at their core, capture the volatility and the distribution of returns preceding each day of estimation. Therefore, the market conditions leading to the estimation period have a significant effect on the value of the measure. This inherent property of VaR is also tied to the size of the sampling window. For example, volatile market periods that contained unusually large negative returns would affect the subsequent VaR estimates by lowering them. This effect was more pronounced with models that used a smaller sampling window, as any short period of increased volatility could cover large part of the data used in the estimation. Therefore, analyzing the market conditions in order to optimize the method of sampling could have provided more accurate estimates. However, due to the scope of this study, the alternative sampling windows were only provided as an additional point of view.

Another limitation of the study emerged from the fact that the performance of the models was only tested during a single stressed market period. Therefore, it could not be concluded if the VaR models would always underestimate market risk when stressed market conditions are present, or if the results were only applicable to the year 2020. The final point regarding the reliability of the results had to do with the backtesting methodology. This study utilized only a single backtesting methodology that examined the ratio of model failures to the total observations. More comprehensive and reliable results could have been obtained by using multiple backtesting methods. For example, a duration-based approach suggested by Christoffersen and Pelletier (2003), could have given a different point of view into the accuracy of the models. This testing methodology could have been used to assess the performance of the models by calculating the duration between the failures and testing it against the hypothesis of expected durations.

6. Conclusions

The purpose of this study was to examine the usability and reliability of Value-at-Risk (VaR) modeling in stressed market conditions. The study utilized Historical VaR (HVaR) and Monte Carlo VaR (MCVaR) calculation methods to estimate the market risk of three popular stock indices during the year 2020. The performance of these models was analyzed using a statistical backtesting method, which compared the ratio of observed model failures to a prespecified theoretical distribution.

The results obtained from the backtesting indicated that the VaR models were not able to accurately estimate the increased risk and volatility present in stressed market conditions during the year 2020. The results were unambiguous, as both models surpassed every prespecified statistical failure limit in the first set of backtesting results. These results provided an answer to the first research question: "*Is VaR a reliable risk measure when stressed market conditions are present?*" In 2020, using the selected methods of calculation and data, VaR proved not to be a reliable measure of risk. Noteworthy, these conclusions were not necessarily generalizable to all potential stressed market conditions and methods of VaR modeling.

The performance of both HVaR and MCVaR was overall very similar. However, MCVaR performed, in some cases, slightly better, as it produced a smaller number of total failures. Additionally, the aggregated Proportion of Failures (PoF) test statistics of MCVaR were also smaller compared to HVaR, which indicated that the model produced more accurate estimates during the backtesting. These results answered the second research question: "*Do the Historical VaR and Monte Carlo VaR models perform equally well during the testing?*" However, the differences between the models were, in most cases, very marginal and statistically insignificant. The results also highlighted an interesting relationship between the size of the sampling window and the performance of the models; reducing the availability of data and using smaller sampling windows resulted in more accurate estimates of market risk. It indicated that the smaller sample sizes increased the models' ability to adapt to changing market conditions. The increased flexibility and reactiveness of the models had a significant impact on their performance during the onset of the global pandemic in 2020. This was due to rapidly changing market conditions that required fast adaptation of the VaR models.

In conclusion, the results of this study provided an interesting starting point for analyzing the performance of risk estimation tools in the context of market risk caused by covid-19. Most importantly, it could be concluded, that market risk estimation became increasingly challenging during the increased volatility and unpredictability observed in 2020. In the following studies, it would be insightful to explore the usability of other methods of calculating VaR. The role of time-varying volatility, exponential weighted moving averages, and ARCH/GARCH -models would provide interesting avenues for researching the subject. These concepts could provide the needed accuracy and reliability for the VaR models and make them sufficiently reliable and usable in stressed market conditions.

References

Adamko, P., Spuchľakova, E. & Valášková, K. (2015) The History and Ideas Behind VaR. *Procedia Economics and Finance*. 24. 18-24.

Alexander, C. (2008) Market Risk Analysis. Volume IV, Value-at-Risk-Models. Chichester, England: Wiley.

Artzner, P., Delbaen, F., Eber, J. & Heath, D. (1999) Coherent Measures of Risk. *Mathematical Finance*. 9(3). 203-228.

Bacon, C. (2008) Practical Portfolio Performance: Measurement and Attribution. Second Edition. Chichester, England: Wiley.

Basuony, M.A.K., Bouaddi, M., Ali, H. & EmadEldeen, R. (2021) The effect of COVID-19 pandemic on global stock markets: Return, volatility, and bad state probability dynamics. *Journal of Public Affairs*. 10.1002. 2761.

BCBS. (2005) Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Basel, Switzerland: BIS.

Berkowitz, J. & O'Brien, J. (2002) How Accurate Are Value-at-Risk Models at Commercial Banks? *The Journal of Finance (New York)*. 57(3). 1093-1111.

Cheung, Y.H. & Powell, R.J. (2012a) Anybody can do Value at Risk: A Nonparametric Teaching Study. *Australasian Accounting, Business & Finance Journal*. 6(1). 111-123.

Cheung, Y.H. & Powell, R.J. (2012b) Anybody can do Value at Risk: A Teaching Study using Parametric Computation and Monte Carlo Simulation. *Australasian Accounting, Business & Finance Journal*. 6(5). 101-118.

Christoffersen, P. (2012) Elements of Financial Risk Management. Amsterdam: Elsevier, Academic Press.

Christoffersen, P. & Pelletier, D. (2003) Backtesting Value-at-Risk: A Duration-Based Approach. Rochester, New York.

Claessens, S. (2014) Financial Crises: Causes, Consequences, and Policy Responses. Washington, District of Columbia: International Monetary Fund.

Consoli, S., Recupero, D.R. & Saisana, M. (2021) Data Science for Economics and Finance Methodologies and Applications. Cham: Springer International Publishing.

Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*. 1(2). 223-236.

Cvjetković, M., Stepanov, S. & Radivojević, N. (2016) The new hybrid value at risk approach based on the extreme value theory. *Estudios De Economia*. 43(1). 29-52.

Dimov, I.T. (2008) Monte Carlo Methods for Applied Scientists. Hackensack, New Jersey: World Scientific.

Engle, R. (2001) GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *The Journal of Economic Perspectives*. 15(4). 157-168.

ESMA. (2020) Guidelines on Liquidity Stress Testing in UCITS and AIFs. ESMA34-39-897. European Securities and Market Authority.

FCIC. (2011) The Financial Crisis Inquiry Report: Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States. New York, NY: Public Affairs.

Gabrielsen, A., Kirchner, A., Liu, Z. & Zagaglia, P. (2015) Forecasting Value-at-Risk with Time-Varying Variance, Skewness and Kurtosis in an Exponential Weighted Moving Average Framework. *Annals of Financial Economics*. 10(01). 1550005.

Horcher, K.A. (2005) Essentials of Financial Risk Management. Hoboken, New Jersey: Wiley.

Illowsky, B. & Dean, S.L. (2018) Introductory Statistics. Houston, Texas: OpenStax, Rice University.

ILO. (2010) Global Employment Trends: January 2010. Geneva: International Labour Office.

Jain, E. (2014) Impact of Law of Demand & Supply on Stock Market: A Study of Most Active BSE Indices with the Help of RSI. *IOSR Journal of Economics and Finance*. 5. 6.

Jorion, P. (2011) Value at Risk: The New Benchmark for Managing Financial Risk. New York: McGraw Hill Professional.

Koulafetis, P. (2017) Modern Credit Risk Management Theory and Practice. London: Palgrave Macmillan UK.

Kupiec, P. (1995) Techniques for Verifying the Accuracy of Risk Measurement Models. Rochester, NY.

Lemons, D.S. (2002) An Introduction to Stochastic Processes in Physics. Baltimore, Maryland: Johns Hopkins University Press.

Longerstaey, J. & Spencer, M. (1996) RiskMetrics™ – Technical Document. New York: J.P. Morgan & Reuters.

Miskolczi, P. (2017) Note on simple and logarithmic return. *Applied Studies in Agribusiness and Commerce*. 11(1-2). 127-136.

MSCI. (2021) MSCI Emerging Markets Index (USD).

Qontigo. (2021) STOXX® EUROPE 600 INDEX. Broad Indices. STOXX Ltd.

S&P Global. (2021) S&P 500®. S&P Dow Jones Indices.

Schmidt, U. & Zank, H. (2005) What is Loss Aversion? *Journal of Risk and Uncertainty*. 30(2). 157-167.

Schober, P., Boer, C. & Schwarte, L. (2018) Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*. 126(5).

Schwert, G.W. (2011) Stock Volatility during the Recent Financial Crisis. *European Financial Management*. 17(5). 789-805.

Taleb, N. (1997) The Jorion-Taleb Debate. *Derivatives Strategy*.

Voit, J., Thirring, W., Beiglböck, W., Grosse, H. & Balian, R. (2005) The Statistical Mechanics of Financial Markets. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Weglarczyk, S. (2018) Kernel density estimation and its application. *ITM Web of Conferences*. 23. 00037.

Appendix A. HVaR backtesting results using different sampling windows

Sampling window N = 125, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	16	0.063	0.865	3.841	Accepted
STOXX 600	258	16	0.062	0.731	3.841	Accepted
MSCI EM	262	19	0.073	2.470	3.841	Accepted

A.1: HVaR backtesting results 1.1.2020-31.12.2020

Sampling window N = 250, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	21	0.083	4.882	3.841	Rejected
STOXX 600	258	16	0.062	0.731	3.841	Accepted
MSCI EM	262	17	0.065	1.122	3.841	Accepted

A.2: HVaR backtesting results 1.1.2020-31.12.2020

Sampling window N = 500, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	23	0.091	7.253	3.841	Rejected
STOXX 600	258	27	0.105	12.512	3.841	Rejected
MSCI EM	262	22	0.084	5.333	3.841	Rejected

A.3: HVaR backtesting results 1.1.2020-31.12.2020

Appendix B. MCVaR backtesting results using different sampling windows

Sampling window N = 125, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	16	0.063	0.865	3.841	Accepted
STOXX 600	258	15	0.058	0.343	3.841	Accepted
MSCI EM	262	16	0.061	0.633	3.841	Accepted

B.1: MCVaR backtesting results 1.1.2020-31.12.2020

Sampling window N = 250, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	20	0.079	3.850	3.841	Rejected
STOXX 600	258	17	0.066	1.252	3.841	Accepted
MSCI EM	262	17	0.065	1.122	3.841	Accepted

B.2: MCVaR backtesting results 1.1.2020-31.12.2020

Sampling window N = 500, significance level α = 0.05						
Index	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	253	22	0.087	6.017	3.841	Rejected
STOXX 600	258	26	0.101	10.958	3.841	Rejected
MSCI EM	262	21	0.080	4.274	3.841	Rejected

B.3: MCVaR backtesting results 1.1.2020-31.12.2020

Appendix C. HVaR backtesting results during the years 2015 – 2019

Sampling window N = 1000, significance level α = 0.05							
Index	Year	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	2019	252	9	0.036	1.197	3.841	Accepted
	2018	251	26	0.104	11.749	3.841	Rejected
	2017	251	3	0.012	10.891	3.841	Rejected
	2016	252	11	0.044	0.223	3.841	Accepted
	2015	252	16	0.063	0.893	3.841	Accepted
STOXX 600	2019	256	7	0.027	3.288	3.841	Accepted
	2018	256	9	0.035	1.319	3.841	Accepted
	2017	256	0	0.000	26.262	3.841	Rejected
	2016	257	17	0.066	1.287	3.841	Accepted
	2015	257	26	0.101	11.068	3.841	Rejected
MSCI EM	2019	261	5	0.019	6.765	3.841	Rejected
	2018	261	21	0.080	4.339	3.841	Rejected
	2017	260	3	0.012	11.601	3.841	Rejected
	2016	261	20	0.077	3.374	3.841	Accepted
	2015	261	19	0.073	2.519	3.841	Accepted

Appendix D. MCVaR backtesting results during the years 2015 – 2019

Sampling window N = 1000, significance level α = 0.05							
Index	Year	n	x	x/n	PoF statistic	$F^{-1}(1-\alpha)$	$H_0: p = p^*$
S&P 500	2019	252	9	0.036	1.197	3.841	Accepted
	2018	251	25	0.100	10.219	3.841	Rejected
	2017	251	3	0.012	10.891	3.841	Rejected
	2016	252	11	0.044	0.223	3.841	Accepted
	2015	252	16	0.063	0.893	3.841	Accepted
STOXX 600	2019	256	6	0.023	4.696	3.841	Rejected
	2018	256	7	0.027	3.288	3.841	Accepted
	2017	256	0	0.000	26.262	3.841	Rejected
	2016	257	15	0.058	0.360	3.841	Accepted
	2015	257	24	0.093	8.203	3.841	Rejected
MSCI EM	2019	261	5	0.019	6.765	3.841	Rejected
	2018	261	19	0.073	2.519	3.841	Accepted
	2017	260	3	0.012	11.601	3.841	Rejected
	2016	261	20	0.077	3.374	3.841	Accepted
	2015	261	16	0.061	0.657	3.841	Accepted