UNIVERSITY OF GOTHENBURG

SCHOOL OF BUSINESS, ECONOMICS AND LAW

CAViaR and Cross-sectional quantile regression models to assess risk in S&P500 sectors

Graduate School

School of Business, Economics and Law University of Gothenburg

Master of Science in Finance

Academic year 2022/23

Supervisor:

Adam Farago

Candidate:

Vladyslava Bab'yak

Abstract

The aim of this thesis is to investigate the performance of different models used in risk management to identify and control risks that may negatively impact company operations due to unpredictable events. More specifically, the object of this paper is the discussion of a cross-sectional quantile regression model (CSQR) and the CAViaR model, which is a time series quantile regression model. Additionally, two highly used models were added: the Historical VaR and the Normal VaR. The out-of-sample analysis performed from 2000 to 2021 to the 11 equally sector sectors within the S&P500 index, suggests that the cross-sectional models outperform the time-series models. The outperformance is evident even during periods of stressed market conditions such as the Financial Crisis and the Covid Pandemic. Similar results were found for the value-weighted sectors. The study concludes that the additional incorporation of data from other firms in the same sector allows the risk manager to a more efficient assessment of the market risk and faster adaptation to market shocks.

Keywords: Value-at-Risk, CAViaR, cross-sectional quantile regression, risk

Acknowledgments

Living in Sweden, I discovered a country totally in line with my life principles of stability and integrity. Therefore, I am sincerely grateful for the opportunity to pursue a dual degree at the University of Gothenburg and the University of Rome Tor Vergata. This achievement would not be possible without the support of my lovely parents and my partner, Michele. A huge thank you goes to them for encouraging me and being always by my side. A sincere appreciation to my supervisor Adam Farago, that was always available, patient, and guided me in writing this thesis. Finally, I am immensely grateful to Andrea and Cristiano for offering support and encouragement during the entire master's degree journey.

List of abbreviations

VaR -Value-at-Risk

S&P500 - Standard and Poor 500

CI - Confidence interval

HV - Historical VaR

CAViaR - Conditional autoregressive Value-at-Risk

CSQR - Cross-sectional quantile regression

POF test - Proportion of failures test

TUFF test - Time until first failure test

TBF test - Time between failures test

CCI test - Christofferson test

Table of contents

List of figures

List of Tables

1. Introduction

1.1 Background

One of the most important processes for individuals and organizations, or probably the most important one, is risk management. The aims of risk management are to help companies to identify and to control the risks that may negatively impact their operations. The final goal is to decrease financial losses due to unpredictable events such as financial crises, pandemics, natural disasters, or even scandals. Furthermore, correct management of the risk can help firms to create better strategies and to allocate in a better way their resources. In addition, it can help companies to maintain a competitive position by protecting their assets and reputation. Risk management is important especially for listed companies because they are subject to different capital and regulatory requirements. It is even more important for banks. Indeed, banks must maintain specific minimum levels of capital based on the risk of their assets. Risk management can help them to meet their requirements and to avoid potential penalties and scandals that would damage customers and taxpayers.

In relation to this, in 1974, the central bank governors of all the G10 countries established the Basel Committee on Banking Supervision, or simply Basel. The aim of the Committee is to develop and establish international standards for regulation, supervision, and risk management. These standards consist of three main agreements: Basel I, Basel II, and Basel III. Every agreement is built on the previous one and provides new tools and new requirements. One of the most used measures to assess the market risk introduced by Basel Committee was the Value-at-Risk (Kou et al., 2013). This standardized tool is based on mathematical and statistical models that consider the volatility, the size of the portfolio, and the time horizon. It was developed at the beginning of the 1990s in the financial industry to provide a company's management with a single number that could quickly and easily explain the risk of an asset or portfolio (Holton, 2002). Value-at-risk, or simply VaR, consists in estimating the maximum potential loss of an investment over a specific period and at a specific level of confidence. This method has been criticized and praised. Overall, VaR is a good, standardized approach to assess the risk promoting transparency and consistency across industries. However, it has different limitations. To name one, it does not allow to capture extreme events or unexpected market shocks. This can lead to an underestimation of the risk and the potential losses. However, despite ongoing debates within academic circles, it remains one of the most used tools to capture and mitigate risk. Basel does not provide a specific methodology for the VaR estimation. Instead, it leaves flexibility for each company to adjust its estimation. For this reason, the effectiveness of VaR models in predicting financial risks remains a persistent concern. This topic is particularly noteworthy in this current period where the health of many US banks has raised concerns of worldwide investors and customers. Indeed, on Friday, March 10, 2023, Silicon Valley Bank collapsed, becoming the third-largest bank failure in the United States' history. So now more than ever is important having an efficient and effective risk-management department along with trustworthy methodologies to control the risk.

Therefore, in this thesis, I proceeded by analyzing some of the most common models to assess the market risk and comparing them with a different and innovative methodology. This innovative methodology was proposed in 2022 by Vidal-Llana and Guillén in the paper that I used as a main reference for this analysis (*Cross-sectional quantile regression for estimating conditional VaR of returns during periods of high volatility', 2022).* The innovation of this model is the usage of cross-sectional data in a quantile regression model to provide VaR estimates. A cross-sectional approach indicates a collection of information from multiple firms at a specific point in time. It allows us to analyze different firms at the same time, but it does not account for changes over time.

1.2 Literature

There are different types of Value-at-Risk models commonly used in financial risk management. The simplest method is the Historical VaR which uses historical data to simulate the distribution of potential returns and estimates VaR (Linsmeier and Pearson, 1996). A very common method is the Variance-covariance method which employs statistical techniques to estimate the parameters of a probability distribution. These parameters can include the mean and standard deviation, and the method then computes the VaR based on the estimated distribution. Typically, a normal distribution is assumed for this estimation (Linsmeier and Pearson, 1996). Another popular method is the Monte Carlo method which provides simulations of the returns and generates some possible scenarios to estimate the

VaR based on the distribution of potential returns. Early models that consider autoregression conditional heteroskedasticity (ARCH and GARCH models) were proposed by Engle (1982) and Bollerslev (1986). These models consider volatility as a weighted combination of past stock returns and past variance. Indeed, the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model is a time-series model used to estimate the volatility in stock returns exploiting the concept of volatility clusters. Other authors modified these models by introducing the concept of asymmetry in volatility, or in other words, they introduced the leverage effect. An example of such a model is the GJR-GARCH specification introduced by Glosten (1993). Furthermore, Value-at-Risk models can be additionally modified by adding some assumptions regarding the shape of the conditional distribution, for example, the assumption of a student's t-distribution or normal distribution. However, if the shape of the distribution fluctuates over time, then the forecasts of volatility may result inappropriate. The acknowledgment of this fact induced new studies that led to the development of a few different models that are not based on any assumptions about the underlying conditional distribution. Indeed, in recent times, there have been suggestions to use extreme quantile estimation methods for evaluating VaR (Danielsson and De Vries, 2000). The rationale behind this approach is to leverage insights about the asymptotic behavior of the tail, instead of focusing solely on the entire distribution. However, this method works only with very low probability quantiles and with i.i.d. variables, which is not the case with financial market assets. Then, Manganelli and Engle proposed a Conditional Autoregressive Value-at-Risk (CAViaR) model (2004). The rationale behind this model is that instead of modeling the entire distribution of returns, this model focuses only on the interested quantile. The empirical evidence shows that the stock market tends to create volatility clusters over time, causing an autocorrelated distribution. Then, the intuition of the authors is that also the Value-at-Risk tends to have an autocorrelated behavior. For this reason, the authors propose an autoregressive (AR) specification for the VaR, and the unknown parameters are estimated with Koenker and Bassett quantile regression (1978). A few studies have shown the superior performance of the CAViaR compared to more traditional models. For example, Allen et al. (2012) performed the GARCH(1,1), the RiskMetrics, the t-APARCH(1,1), and the CAViaR model in the Australian market. Their findings show higher forecasting ability of the CAViaR compared to the other models. Similar results were found by Bautista and Mora (2021) in more recent studies. The authors compared CAViaR with GJR-GARCH models under normal, skewed-normal, Student-t, and Student-t normal distribution assumptions. Their result supports the robustness of the CAViaR model in an out-of-sample VaR analysis for emerging market stocks (MILA and ASEAN-5). The success of the CAViaR model over RiskMetrics and Extreme value theory (EVT) models in emerging markets was also confirmed by a study conducted by Bao et al. (2006).

However, the CAViaR model is based on a univariate time-series analysis of a particular company or portfolio, which relies only on past information about that company (or portfolio of companies) and does not consider other companies in the same market. To assess a company's returns based on a particular attribute, it may be necessary to have a cross-sectional sample of firms operating in the same market. To address this issue, the authors Vidal-Llana and Guillén (2022) suggest utilizing a cross-sectional quantile regression model to analyze tail returns for a single company, and then confront this method with the CAViaR model. Both models are based on quantile regression. Their analysis was conducted on 26298 US firms and showed that during high-volatility periods the CSQR model yields enhanced accuracy in predicting low quantiles. There exists a vast multitude of academic studies dedicated to the analysis of cross-sectional asset pricing. There are some papers that provide an overview of this topic. For instance, Nagel (2013) summarized the evidence on cross-sectional return predictability and the failure of standard CAPM models. More recently, Chen and Zimmermann (2021) provided data and code that reproduces cross-sectional stock return predictors. In this thesis, I have developed two models that incorporate some covariates inspired by some different studies on the momentum effect by Jegadeesh and Titman (1993) and Novy-Marx (2012). Other covariates are inspired by the study on the skewness prediction of cross-sectional returns by Amaya et al. (2015).

1.3 Purpose

This thesis seeks to expand upon existing research literature by investigating the viability of quantile regressions for evaluating market risk. The main models analyzed are the CAViaR model, which is based on a time-series quantile regression, and the cross-sectional quantile regression model (CSQR) as proposed by Vidal-Llana and Guillén (2022). By employing quantile regressions, the specific conditional quantile can be directly modeled without any presuppositions about the distribution of the return series. However, a cross-sectional quantile regression incorporates information from other firms, and this may help to obtain a better estimation of the Value-at-Risk, consequently a better assessment of the market risk.

Formally, I aim to answer the following research question:

- Does a cross-sectional quantile regression model exhibit superior market risk prediction capabilities compared to conventional time-series VaR models?

Very few papers investigate the cross-sectional models to estimate VaR, and to the best of my knowledge, is not yet conducted an analysis of cross-sectional quantile regression models on the different S&P500 sectors. As Vival-Llana and Guillén (2022) proposed, I analyze in this paper the CAViaR and a cross-sectional quantile regression model. However, instead of focusing on the entire US market, as the previous authors did, I decided to analyze separately each S&P500 sector. Indeed, I examined the weekly VaR estimates within 11 sector portfolios of the S&P500. This selection was deliberate, as the S&P500 index is built by the largest companies in the US in terms of market capitalization. By focusing on the sector portfolios instead of only the entire S&P500 portfolio, I was able to draw conclusions based on 11 distinct portfolios rather than just one. This approach not only provided a more comprehensive analysis but also underscored the significance of incorporating cross-sectional information from stocks outside the specific sector when predicting VaR. In contrast to Vival-Llana and Guillén's (2022) approach, my study incorporated various confidence interval (CI) levels, namely 95%, 98%, and 99%. Additionally, I introduced a new cross-sectional model (CSQR2) that exhibited superior performance compared to CSQR1 model, which drew inspiration from their work. The effectiveness of various VaR models was assessed using an out-of-sample analysis, which spanned from January 2, 2000, to December 26, 2021. Then, similarly to Vival-Llana and Guillén (2022), two additional out-of-sample analyses were made during stressed market conditions. The first period extended from January 7, 2007, to January 2, 2011, which coincided with the Financial Crisis. The second period extended from June 3, 2018, to December 26, 2021, corresponding to the Covid Pandemic. In contrast to the findings of the authors, who observed improved performance of their cross-sectional model solely during stressed market conditions, my research revealed that both cross-sectional models (CSQR1 and CSQR2) exhibited superior performance across all time periods, including the entire sample and both stressed market conditions. The final aim of my thesis is to show that there are possibilities to implement easy, but better-performing models for managing the market risk. Indeed, the cross-sectional models I presented are relatively easy to implement and understand, and according to the analysis in this paper, they are even performing better than the typical time series models like Historical VaR, Normal VaR, and CAViaR. By doing that, I hope to provide some valuable insights for future analyses, allowing also for the development of even more complex cross-sectional models that the ones I presented in my thesis.

1.4 Structure of the Thesis

The core topic of this thesis is the comparison between the CAViaR and the CSQR models, however, other models, such as the Historical VaR and the Normal VaR, were performed for a more complete analysis. The organization of the thesis is as follows: Section 2 introduces the theoretical background of the different VaR models used in the analysis. This section introduces also the different performance evaluation methods applied to the VaR models. Section 3 involves the data description, the descriptive statistics, and the methodology utilized. Finally, Section 4 includes the empirical results and Section 5 concludes the thesis.

2 Theoretical Background

2.1 Value-at-Risk

Value-at-Risk (VaR) is the most common tool used in risk management to measure an asset or portfolio's potential loss over a specified period. The VaR is a downside risk measure that provides an estimate of the maximum amount of loss that is not expected to be exceeded with a certain level of confidence (CI) and it is typically expressed in a currency or in a percentage of the initial investment. In other words, the Value-at-Risk is defined as the maximal loss that will be exceeded with probability α in the next *l* trading days, i.e.,

$$P(\ R(l) < VaR^{\alpha}\ ) = \alpha \tag{1}$$

- $R(l)$ is the return of the stock or portfolio;
- l is the trading days;
- $1 - \alpha$ is the level of confidence (CI);

To be more precise, the VaR can be expressed in two different ways: in terms of returns, as just discussed, and in terms of losses. Both approaches indicate the same concept but in different ways. In the first case, VaR provides a standardized measure of the potential loss in terms of the return relative to the initial investment made, i.e., the typical VaR values are negative. In the second case, the estimates are made directly on the losses (negative returns), i.e., the resulting VaR values are typically positive. In this thesis, I expressed all the models in terms of the returns. Furthermore, the methods to calculate the VaR are frequently classified into two main categories: parametric and nonparametric methods. Nonparametric models use techniques such as Historical VaR. They are called in this way because it is not required any estimation of parameters. On the other hand, parametric models like the Normal VaR, the CAViaR, and the Quantile Regression approach require parameter estimation.

Note that the level of confidence (CI) chosen in each model affects the VaR estimates. This level is usually decided by the risk management department based on the risk appetite of the firm. In some cases, it is imposed by the Basel Committee. Indeed, according to Basel II, a

99.9% level of confidence must be used to establish the bank's regulatory market risk capital (Bank for international settlements, 2005). In the literature, it is very common the usage of 95% and 99% as CI. In this paper, I analyzed all the models with 95%, 98%, and 99% CI.

In the following subsections, I am going to introduce the theoretical background behind the widely used models such as Historical VaR and Normal VaR, followed by Manganelli and Engle's model CAViaR (2004), and the cross-sectional quantile regression model by Vidal-Llana and Guillén (2022). At the end of this chapter, I then discuss the theory behind the different evaluation performance techniques applied in this thesis for a comparison of the VaR models.

2.2 Historical VaR

The Historical VaR (HV) is a straightforward and non-theoretical approach that does not involve making assumptions about the statistical distributions of the underlying market factors. In this approach, to estimate the distribution of tomorrow's portfolio returns, the HV method approximates it using the empirical distribution of past returns. The VaR is then the α-quantile of the sequence of observed returns. More precisely, these are the steps to follow for the estimation:

1. Calculate the returns from the stock prices;
2. Choose the desired level of confidence $1 - \alpha$;
3. Determine the α-quantile of the return distribution. The quantile found is the VaR.

Historical VaR is an easy approach to assess market risk and to interpret it. It is frequently used as a benchmark due to its quick calculation. It does not require any estimation of parameters and it does not require to make any assumption on the underlying distribution of the returns. The advantage of HV is that it uses real data. For this reason, it may capture possible unexpected events that cannot be predicted by other more complex theoretical models. However, one of the main disadvantages is that it is usually very slow to react to market shocks and it may undervalue the real risk because it fails to capture the extreme events.

2.3 Normal VaR

The Normal approach to assess Value-at-Risk assumes that the underlying market factors follow a specific distribution, in this case, a normal distribution. This is one of the most used methods to assess market risk in the financial industry.

By making the normality assumption, the method determines the distribution of potential profits and losses of a mark-to-market portfolio. Then, by applying standard mathematical properties of the normal distribution, the VaR can be determined (Linsmeier and Pearson, 1996). The variance and the mean are estimated from the historical data:

$$\hat{\mu}_t = \frac{1}{j}\sum_{j=1}^{J} R_{t-j} \tag{2}$$

$$\hat{\sigma}_t^2 = \frac{1}{J-1}\sum_{j=1}^{J}\left(R_{t-j} - \hat{\mu}_t\right)^2 \tag{3}$$

In the above formulas, R_{t-j} is the return at time *t-j*, $\hat{\mu}_t$ is the sample mean return, $\hat{\sigma}_t^2$ is the sample variance of the return, and *J* is the number of observations in the time window used. Then the Normal Value-at-Risk is estimated as follows:

$$VaR_t^{\alpha} = \hat{\mu}_t + \hat{\sigma}_t * q_z(\alpha) \tag{4}$$

where $q_z(\alpha)$ is the α-quantile of the standard normal quantile at the specified level of confidence, $\hat{\mu}_t$ is the estimated mean and $\hat{\sigma}_t$ is the estimated standard deviation.

This approach is highly used due to its simplicity and intuitive interpretation. It requires a basic statistical computation, and it can be easily integrated with portfolio optimization techniques.

The main disadvantage of this model, however, is that the assumptions of normality in returns and constant moments may not be realistic in real-world scenarios. Indeed, asset returns tend to exhibit high levels of skewness and kurtosis, which all may significantly impact the VaR estimates.

2.4 CAViaR model

The Conditional Autoregressive Value-at-Risk (CAViaR) model was proposed by Manganelli and Engle in 2004. The intuition of the authors was that rather than modeling the entire distribution, the quantile can be modeled directly using a conditional autoregressive quantile specification. The clustering of volatilities of stock market returns over time, which is an empirical observation, implies that the quantiles of the distribution of these returns is autocorrelated. As a result, the VaR, which is closely linked to the standard deviation of the distribution, must show similar autocorrelated behavior. To account for this characteristic, an autoregressive specification is often used, and CAViaR is one such method.

The generic CAViaR specification is a recursive formula for calculating the α-quantile of the portfolio returns at time t, given the past returns. The generic specification of this model can be written as:

$$VaR_t^{\alpha} = \gamma_0 + \sum_{j=1}^{Q} \gamma_j * VaR_{t-j}^{\alpha} + \sum_{j=1}^{H} \gamma_{Q+j} * f(R_{t-j}) \tag{5}$$

where

- VaR_t^{α} denotes the conditional quantile at time *t*;
- VaR_{t-j}^{α} denotes the α-quantile estimate at time *t-j*;
- $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_Q, \dots, \gamma_{Q+H})$ is a vector of parameters to be estimated.

The first term, γ_0, represents the intercept of the model. The second term represents a set of autoregressive components that captures the persistence of the α-quantile estimates over time. The lag structure is specified by the parameter Q, which determines the number of past quantile estimates included in the model.

The third term in equation (5) is a set of lagged past returns that can help improve the accuracy of the quantile estimates. The number of observable variables is denoted by the parameter H, and the coefficients γ_{Q+j} capture the effect of the jth observable variable at time t-j on the current quantile estimate. The term $f(R_{t-j})$ in the CAViaR model plays a crucial role in capturing the effect of observable variables on the conditional quantiles of portfolio returns. This term can be seen as a way to link the quantile estimates to the information set that includes the past values of the observable variables. In other words, $f(R_{t-j})$ allows to incorporate information from the observable returns into the quantile estimation process. This role of $f(R_{t-j})$ is similar to the news impact curve introduced by Engle and Ng (1993) in the context of GARCH models. In the GARCH framework, the news impact curve describes how shocks to the conditional variance of asset returns affect subsequent returns. Similarly, in the CAViaR model, the term $f(R_{t-j})$ captures the impact of observable variables on the conditional quantiles of returns. By incorporating the news impact curve or the observable variables, the model can provide a more accurate description of the dynamics of financial markets and help forecast future returns.

In the stock market, if the past returns of the stock or portfolio were negative, then we would expect the VaR to increase (in absolute value), indicating more risk and higher potential future losses. The explanation is that one bad day may increase the probability of another bad day, leading to a higher loss. On the other hand, very good days may also increase the VaR, as high returns can be followed by high volatility or a significant drop in prices. In other words, the VaR can depend symmetrically on the magnitude of past returns, whether they are positive or negative. This phenomenon of dependence on past returns is very similar to what is observed in volatility models (for example in the GARCH models), where past squared returns are used to predict future volatility. By incorporating the past information into the VaR model, it may be possible to obtain a more accurate estimate of the potential losses of a portfolio and make better investment decisions.

In the CAViaR model, the prespecified α-quantile of the distribution (i.e., the Value-at-Risk) is identified and modeled instead of modeling the entire distribution. This stands in contrast to other approaches such as GARCH models that aim to model the entire distribution (Engle and Manganelli, 2004).

2.4.1 The symmetric absolute value CAViaR approach

In their paper, Engle and Manganelli (2004) discussed several CAViaR processes, including the adaptive process, the symmetric absolute value process, the asymmetric slope, and the Indirect GARCH. This thesis concentrates on the second process (i.e., the symmetric absolute value) due to its straightforwardness and resemblance to the Indirect GARCH model, which models positive and negative returns in a similar way.

The specified process known as the "symmetric absolute value" can be described as follows:

$$VaR_t^{\alpha} = \gamma_0 + \gamma_1 * VaR_{t-j}^{\alpha} + \gamma_3 * |R_{t-1}| \tag{6}$$

It is important to note that CAViaR models can accommodate the non-iid distribution of errors, in contrast to the GARCH models. This indicates that this model is not limited to situations with constant volatilities but can also be applied in scenarios where the error distributions change, or in situations where both the error densities and volatilities are changing.

2.5 Quantile Regression

Koenker and Bassett (1978) introduce a linear estimator that departs from the standard assumption of normality for the error term in linear regressions. The proposed quantile regression model outperforms the least square regression when the error terms are non-normally distributed. In addition, when examining financial time series, there are often abnormal observations during periods of crises and booms. Quantile regression models are recognized for their ability to handle outliers in the sample, making them particularly useful in financial analysis (Uribe and Guillen, 2020). Indeed, quantile regression is a versatile and precise econometric method that can effectively address many of the prevalent challenges faced in modern economics and finance. For instance, it can be used to evaluate the impact of environmental and market factors on the day-to-day decisions made by individuals and firms regarding production and consumption. Additionally, it can be applied to model time series of prices, accounting for market state or seasonality variations.

This method involves the estimation of the coefficients in a linear combination. The linear combination includes some covariates to fit a specific quantile of the response variable's cumulative distribution, at a fixed confidence level. In other words, the method aims to model how changes in the covariates affect a specific α-quantile of the response variable.

Quantile regression aims to achieve different objectives than the classical linear regression. In a classical linear regression, the dependent variable is expressed as a linear combination of some fixed covariates X_i' and a random error term ε_i, and the model aims to fit the mean of the response, assuming the error term has an expected value of zero. That is,

$$Y_i = X_i' \, \beta + \varepsilon_i \tag{7}$$

assuming $E\,[\ \varepsilon_i\ ] = 0$

Then the parameters $\hat{\beta}$ estimated are found with the ordinary least squares (OLS) by minimizing the sum of squared residuals:

$$\hat{\beta} = argmin_{\beta} \sum_{i=1}^{n} [\ Y_i - E(Y_i|X_i)\ ]^{\,2} = \ argmin_{\beta} \sum_{i=1}^{n} [\ Y_i - X_i' \, \beta\ ]^{\,2} \tag{8}$$

where n is the number of observations in the dataset, and the residuals are calculated by subtracting the conditional expected value, $E(Y_i|X_i)$, from the observed response values. The conditional expectation is estimated based on the covariate(s) associated with each observation, represented by X_i.

On the other hand, the quantile regression focuses on studying the quantiles of the response variable, rather than its expected value. The cumulative distribution function forms the basis of quantile regression, making it an excellent tool to investigate the factors that influence the probability of extreme response values. For example, when working with a 50% confidence level, the 50th quantile regression corresponds to a model for the median of the response variable's conditional distribution. In this paper, the level of confidence (CI) used were 95%, 98%, and 99%. For example, if the CI is 95% then the quantile regression represents a model for the 5th quantile, indicating the conditional distribution of the response that should not be surpassed by more than 5% of cases.

The following equation represents the quantile regression:

$$Q_{Y_i}(\alpha) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \cdots + \beta_K X_{Ki} = X_i' \beta \tag{9}$$

where X_i' is the matrix with all the covariates taken into consideration, β is the vector of estimated parameters and Q_{Y_i} is the estimated quantile for a given observation Y_i.

When covariates are taken into account, estimating the α-quantile $Q_{Y_i}(\alpha)$ involves finding a parameter vector that minimizes the expected loss function. This procedure is similar to classical linear regression, where the parameter vector is estimated by minimizing the sum of squared residuals, which corresponds to a quadratic loss function. However, in quantile regression, the objective is to minimize the difference between the predicted values and the actual observations, weighted by a loss function that accounts for the specific α-quantile being estimated. The parameter vector β changes for every confidence level considered and its estimation is done via the following minimization:

$$\hat{\beta} = argmin_\beta \, E \left[\rho_\alpha (Y_i - X_i' \beta) \right] \tag{10}$$

where $\rho_\alpha (u) = (1 - \alpha) \, I(u < 0) \, |u| + \alpha \, I(u > 0) \, |u|$

(11)

When dealing with the median case (α = 0.5), the solution can be obtained by minimizing the sum of the absolute deviations, which is calculated by summing the lengths of the residuals. In other words, the line that best fits the data, in this case, is the one that gives the smallest sum of the absolute differences between the predicted values and the actual observations. The next subsection presents the cross-sectional VaR model, which is based on the quantile regression just discussed.

2.5.1 Cross-sectional quantile regression model

Vidal-Llana and Guillén (2022) propose a cross-sectional approach to model the Value-at-Risk. In this model, the authors fix a point in time *t*, that is between 1 and the number of observations *T*, and they estimate a quantile regression model for the stock returns using specific characteristics of all the firms in the index in that specific time *t.* The result is as many quantile regressions as the number of observations *T*. This methodology, called by the authors the "cross-sectional Quantile Regression model" (CSQR) uses the information of all the firm in the index at a given point in time *t* and it assumes that the quantiles depend on all the companies' characteristics. This is the main difference with the CAViaR model, which assumes that the quantiles of a single company's returns depend only on its own past returns.

Vidal-Llana and Guillen (2022) used a set of seven firm characteristics to compute the cross-sectional quantile regression, such as the company size (MC), the book-to-market ratio (BM), the operating profitability (OP), the growth rate of investment (INV), the 12-month momentum (MOM), the liquidity (LIQ) and the market beta (BETA). The regressors chosen are standard characteristics used to explain the cross-section of expected stock returns. Some of the characteristics, such as MC, BM, OP, and INV are recommended by various previous studies including Fama and French (2015, 2020) and Hou, Xue, and Zhang (2015). Other characteristics, such as BETA, LIQ, and MOM are recommended by Campbell (2017) and Malkiel (2019). In chapter 3 of this thesis, I specify that instead of using seven covariates like Vidal-Llana and Guillen (2022) did, I proposed two models based on only four and five covariates. As specified by the authors, more regressors could be added which would potentially lead to better results. However, my analysis shows that it is sufficient to incorporate just four covariates to obtain better outcomes than other models.

One of the main advantages of the CSQR method to assess the VaR is the simplicity of computational requirements, which is easier than fitting a CAViaR model for example. It uses covariates, that allow to include exogenous characteristics to predict the return quantiles for external firms that were not initially in the dataset. It is robust to outliers since it is based on a quantile regression instead of a mean-based regression. However, the main disadvantages are the requirement of large data sets of different covariates and the limitation of the cross-sectional analysis that does not account for time series dynamics.

2.6 Performance evaluation measures

It is of vital importance to evaluate the performance of the VaR estimates. In order to assess the efficacy of various models, it is possible to use different evaluation measures. In this thesis, the following widely used techniques are performed: the percentage of VaR violations, the score function, the Kupiec test, and the time between failures likelihood ratio test. An additional test is performed to verify the independence of the VaR estimate series, the so-called conditional coverage test of Christoffersen.

The percentage of VaR violations represents the proportion of times that the actual losses exceeded the estimated Value-at-Risk threshold. For example, if the estimated Value-at-Risk corresponds to the 95% level of confidence, then a higher percentage of violations than 5% indicates that the VaR estimate is too low, meaning the method used underestimates the market risk of the portfolio. On the other hand, a lower percentage of violations than 5% suggests that the VaR estimate is too conservative, meaning that the method used overestimates the risk. For this reason, ideally, the percentage of violations should be close to the chosen significance level.

A common way to evaluate the VaR models is through the so-called "Hit function". The Hit function $I_t(\alpha)$ is an indicator function indicating the ex-post observations R_t of VaR violations. That is,

$$I_t\,(\alpha) = \begin{cases} 1 & if\ R_t < -VaR_{t|t-1}(\alpha) \\ 0 & otherwise \end{cases} \tag{12}$$

Christoffersen (1998) stressed that Value-at-Risk forecasts are valid if and only if the violation process $I_t(\alpha)$ satisfies two hypotheses: the unconditional coverage and the independence hypothesis.

- The *unconditional coverage* (UC) hypothesis indicates that the probability of the return exceeding the VaR forecast must be equal to the coverage rate (level of confidence stated before). That is,

$$\text{Prob}\,[I_t(\alpha) = 1] \;=\; E[I_t(\alpha)] \;=\; \alpha \tag{13}$$

Consequently, if the frequency of observed violations is significantly lower than the coverage rate, then the used model to estimate the VaR is overestimating the true level of risk. On the other hand, if the observed violations are significantly higher than the coverage rate, then the model is not efficiently assessing the real risk. However, this hypothesis does not tell anything regarding the level of dependence on the VaR violations.

- The *independence hypothesis* indicates that the VaR violations, with the same coverage rate observed on two different dates, must be independently distributed. Therefore, past VaR violations must not have any additional information about the present and/or future VaR violations.

$$I_t(\alpha) \perp\!\!\!\perp I_{t-k}(\alpha)$$

2.6.2 Kupiec POF test

To test the first hypothesis (unconditional coverage), Kupiec (1995) proposed the Proportion of failure test (POF). This test examines whether the empirical coverage matches the theoretical coverage. For instance, if the Value-at-Risk has an α-value of 5%, this test assesses whether the coverage obtained when applying the VaR to the data is truly 5%. The Kupiec test statistic is:

$$LR_{POF} = 2\ln\left(\left(\frac{1-\hat{\alpha}}{1-\alpha}\right)^{n-I(\alpha)}\left(\frac{\hat{\alpha}}{\alpha}\right)^{I(\alpha)}\right) \tag{14}$$

where

$$\hat{\alpha} = \frac{1}{n}\,I(\alpha) \tag{15}$$

and

$$I(\alpha) = \sum_{t=1}^{n} I_t(\alpha)$$

(16)

where *n* is the number of observations. The test statistic approximately follows a Chi-squared distribution with one degree of freedom. The null hypothesis assumes that the excess rate obtained by the model is identical to the assumed excess rate. Consequently, the test aims to determine whether there exists a significant difference between the observed and assumed excess rates.

When testing the 95% VaR, if the proportion of violations is exactly equal to 5% then the Kupiec test takes the value zero. This indicates that there is no evidence of inadequacy in the model used to assess the Value-at-Risk, implying that the model was able to accurately capture the risks associated with the investment. Therefore, the VaR estimate may be considered reliable. When the proportion of violations is different from 5%, then the test statistics grows and indicates that there is evidence that the model either systematically understates or overstates the risk of the stock or portfolio (Zhang and Nadarajah, 2016).

2.6.3 Christofferson test

Christoffersen (1998) developed a likelihood test to assess the frequency of consecutive exceedances. I will refer to this as the CCI test in the thesis. This test verifies the independence hypothesis, discussed earlier in this section. The test statistics follows a Chi-squared distribution, and it is calculated as follows:

$$LR_{CCI} = -2\log\left(\frac{L(\alpha; I_1, I_2, \dots, I_T)}{L(\hat{\Pi}_1; I_1, I_2, \dots, I_T)}\right)$$

(17)

where the likelihood under the null hypothesis is the following:

$$L(\alpha; I_1, \dots, I_t) = (1-\alpha)^{n_0}\,\alpha^{n_1}$$

(18)

And the approximate likelihood function for this process is:

$$L(\hat{\Pi}_1; I_1, \dots, I_t) = (1-\hat{\pi}_{01})^{n_{00}}\,\hat{\pi}_{01}^{n_{01}}\,(1-\hat{\pi}_{11})^{n_{10}}\,\hat{\pi}_{11}^{n_{11}}$$

$$\hat{\pi}_{01} = \frac{n_{01}}{n_0} \tag{19}$$

$$\hat{\pi}_{11} = \frac{n_{11}}{n_1} \tag{20}$$

(21)

In the following equations, α indicates the probability of exceedance for a significance level of 5%, n_1 represents the number of exceedances present in the time series and n_0 represents the number of predictions that were not exceeded. The number of misses that are followed by another miss is indicated with n_{00} and the number of hits followed by another hit is indicated with n_{11}. Finally, the number of hits followed by a miss is indicated with n_{10}.
The null hypothesis of the Christofferson test states that the exceedances are independent of each other, while the alternative states the non-independence. If the observed test statistics surpasses the critical value, then the null hypothesis is rejected, i.e., the exceedances are not independent or/and the exceedance rate is not correct.

2.6.4 Kupiec TUFF test

Kupiec (1995) introduced another test based on the same assumptions as the POF test, the Time until first failure test (TUFF). This test measures the time until the first violation, and it is mainly used as a preliminary to the POF test. The null hypothesis of the test is the following:

$$H_0 : p = \frac{1}{\upsilon} \tag{22}$$

where υ represents the time until the first violation.
From Zhang and Nadarajah (2016), the test statistic is the following:

$$LR_{TUFF} = 2 \ln \left(\frac{p(1-\alpha)^{\upsilon \ \ -1}}{\hat{p}(1-\hat{\alpha})^{\upsilon \ \ -1}} \right) \tag{23}$$

where $\hat{\alpha}$ is given by equation (15).

As in the POF test, the test statistic approximately follows a Chi-squared distribution with one degree of freedom. Note that this test relies only on the number of violations, but does not consider the time dynamics of the violations.

2.6.5 TBF test

Haas (2001) presented another test to verify the hypothesis of independence, the Time between failures test (TBF), which is an extension of the Kupiec TUFF test. This statistical test assess that the violations are not correlated, and this is following likelihood ratio statistic:

$$LR_{TBF} = \sum_{i=2}^{m} \left(-2\, ln \left(\frac{\alpha(1-\alpha)^{v_{i-1}}}{\hat{\alpha}(1-\hat{\alpha})^{v_{i-1}}} \right) \right) - 2 \ln \left(\frac{\alpha(1-\alpha)^{v_{i-1}}}{\hat{\alpha}(1-\hat{\alpha})^{v_{i-1}}} \right) \tag{24}$$

The likelihood ratio statistic proposed evaluates the null hypothesis that violations are independent of each other. It employs the time gaps (v_i) between successive violations, and the number of violations (m), with the resulting statistic having a Chi-squared distribution with m degrees of freedom. Rejecting the null hypothesis requires exceeding the critical value for the chi-squared distribution.

An advantage of this test is its robustness since it can detect issues with violation dependencies and their count.

2.6.6 Score function

The scoring function is used in decision theory to assess the accuracy of a forecast with a probability distribution, such as a quantile. The score function is a strictly consistent function (Fissler and Ziegel, 2016), meaning that it can be used as the loss function in a forecast estimation (Gneiting and Raftery, 2007). The following type of scoring function is a consistent form for the VaR estimation:

$$S(Q_t, R_t) = \left(\alpha - I(R_t \leq Q_t)\right) * \left(G(R_t) - G(Q_t)\right) \tag{25}$$

In this equation, Q_t is the quantile with probability level α, I is an indicator function that assumes value 1 when the return R_t is lower than the quantile Q_t and value 0 vice versa, R_t are the index returns and G is a weakly increasing function. If the function G is strictly increasing, then it can be proven that the scoring function is strictly consistent (Gneiting, 2011). Then:

$$S(Q_t, R_t) = \left(\alpha - I(R_t \leq Q_t)\right) * (R_t - Q_t) \tag{26}$$

Due to its simplicity and familiarity as the quantile regression loss function, the score function is extensively used in the academic literature in the evaluation of Value-at-Risk models. By taking the average of the score, it is possible to derive a measure to evaluate α-quantile forecasts.

Vidal-Llana and Guillen (2022) introduce the Acerbi and Szekely (2014) scoring function to evaluate the performance of the CAViaR and CSQR models for Value-at-Risk:

$$Q_\alpha^0 = \frac{1}{N}\sum_{i=1}^{N}\left[\alpha - I\left(R_i \leq \tilde{R}_i(\alpha)\right)\right] * (R_i - \widetilde{R_i}\,(\alpha)) \tag{27}$$

The following function measures the discrepancy between the observed value y_i and the predicted quantile value $\tilde{R}_i(\alpha)$, that it is found as $\tilde{R}_i(\alpha) = X_i' * \widehat{\beta_\alpha}$. The score function Q_α^0 is simply a weighted average of the absolute distance between the observed value and the fitted α-quantile. The lower the score, the better the approximation of the quantile.

3 Data

3.1 Data description

The objects of the study are the different sector portfolios in the S&P500 index, which contains the biggest 500 US companies by market capitalization. The sectors are defined using the Global Industry Classification Standard (GICS) taxonomy, which was developed in 1999 by MSCI and Standard & Poor's. The GICS categorizes all companies into 11 sectors: Information Technology, Health Care, Consumer Discretionary, Communication Services, Financials, Industrials, Consumer Staples, Energy, Utilities, Real Estate, and Materials. The first two digits a firm's GICS code indicates which sector the company belongs to. These two-digit sector codes are reported in the appendix, table 18.

The sector that had the smallest weight in the S&P500 index on 05/05/2023 was Real Estate, consisting only of 22 companies and having a market capitalization of around $1.38T. The largest sector in the index was Information technology, composed of 66 companies and a market capitalization of around $12.24T. From 2000 to 2021, this sector is the one that gave the highest weekly returns on average (*0.32*%), and it is composed of firms in Software & Services, Technology Hardware & Equipment, Semiconductors & Semiconductor Equipment. The main companies representing this sector are Apple Inc., Microsoft Corp, Nvidia Corp, Broadcom Corp, Salesforce Inc., Cisco Systems Inc., Accenture plc A, Adobe Inc., Texas Instruments Inc., and Oracle Corp. The second largest sector in the S&P500 index was Financials, with $8.69T of market cap. It is composed of banks, financial services firms, and insurance firms. The main companies in the sector are Berkshire Hathaway B, JP Morgan Chase & Co, Visa Inc A, Mastercard Inc A, Bank of America Corp, Wells Fargo & Co, S&P Global Inc, Morgan Stanley, Goldman Sachs Group Inc, and BlackRock Inc. The market capitalization data for the sectors of the S&P500 Index reported above were computed as of the 5th of May 2023, as reported by Fidelity.com.

In this thesis, all the Value-at-Risk models were performed separately for each S&P500 sector portfolio. The overall sample period covered in the analysis is from 1 January 1995 to 26 December 2021. The list of historical S&P500 index constituents throughout this period was obtained from Wharton Research Data Services (WRDS). During this period, a total of 1210 companies were included in the S&P 500 index at some point. The individual weekly return series for each of these companies was downloaded from the Center for Research in Security Prices (CRSP) database through WRDS. The GICS codes of the companies (to be able to form the sector portfolio) were obtained from the *Compustat* database through WRDS. According to the sector classification provided by GICS, I filtered and collected the stocks belonging to each sector. Finally, I created the weekly return series of a specific sector portfolio from the individual returns on the stocks that are included in the S&P500 index during that week and belong to the given sector.

The out-of-sample evaluation performance was computed from 2 January 2000 to 26 December 2021. While the out-of-sample evaluation performance during stressed market conditions was computed first from 7 January 2007 to 2 January 2011 (Financial Crisis period) and secondly from 3 June 2018 to 26 December 2021 (Covid Pandemic period).

The results of the different Value-at-Risk methodologies applied to the 11 sectors in the S&P500 index under consideration were largely consistent, except for the Information Technology sector. The latter showed a slightly different outcome due to the Dot-com bubble in the late 1990s. For this reason, the graphical illustration of the results in this thesis featured the Financials sector, which was the second-biggest sector in the S&P500 index and served as a representative example.

The source of data to collect some of the covariates for the cross-sectional analysis was obtained from the replication dataset connected to Chen and Zimmermann (2021)[1]. These involve the following firm characteristics: market capitalization (*SIZE*), the 1-year beta (*BETA*), the profitability factor (*ROE*), the asset growth factor (*GROWTH*), and the book-to-market (*BM*). All the variables are expressed in US dollar currency (*USD*).

For the main part of the analysis, the sector returns were calculated as the equally-weighted average of the returns of all companies in each sector. This decision is based on the significant variation of the market capitalization between all the companies listed in the

[1] The data are available from https://www.openassetpricing.com/.

index, i.e., many companies have very significant market capitalization compared to other companies in the same index. This variation may lead to an overemphasis on companies with bigger market capitalization and can potentially distort the analysis. However, also note that the main results are replicated using value-weighted sector portfolios as well. The equal- and value-weighted approaches lead to similar conclusions.

The software used to conduct the analysis was Matlab R2022a. The code for the CAViaR model (2004) was taken from Manganelli's website and then adapted to this paper's specific analysis.

3.2 Descriptive statistics

The following table displays the descriptive statistics of the equal-weighted weekly index returns of the 11 S&P500 sectors from 1995 to 2021.

	Mean %	*Median %*	*Std %*	*Skewness*	*Kurtosis*	*Jarque-Bera*	*Max %*	*Min %*	*Obs.*
Energy	0.21	0.40	4.21	-0.54	9.04	2211.4	21.40	-30.82	1409
Materials	0.21	0.30	3.14	-0.07	8.13	1545.2	22.39	-14.82	1409
Industrials	0.23	0.34	2.88	-0.35	8.89	2068.1	15.35	-19.45	1409
Consumer Discretionary	0.22	0.27	3.35	0.43	16.66	1098.9	0.30	-24.12	1409
Consumer Staples	0.21	0.24	1.96	-0.64	9.37	2472.9	11.83	-15.85	1409
Health Care	0.30	0.40	2.44	-0.66	8.26	1727.3	12.36	-18.63	1409
Financials	0.25	0.33	3.72	-0.58	15.78	9661.3	31.05	-25.04	1409
Information Technology	0.32	0.44	3.88	-0.12	5.76	449.9	18.99	-21.24	1409
Communication Services	0.20	0.29	2.97	-0.25	9.92	2820.9	23.99	-21.44	1409
Utilities	0.15	0.21	2.61	-0.32	12.51	5332.7	19.0	-19.45	1409
Real estate	0.16	0.27	3.79	0.08	15.61	9342.1	29.27	-25.40	1409

Table 1: Descriptive statistics S&P500 sectors

According to table 1, the sector that experienced the biggest drop was Energy with a weekly -*30.82%*. On the other hand, the sector that had the highest weekly return was the Financials with *31.05*%. An important consideration is that all the sectors in the S&P500 index experienced a downturn during the Financial Crisis in 2008 and during the Covid Pandemic in 2020. The expectation was to find high estimated values (in absolute value) of the Value-at-Risk during stressed conditions.

A way to verify if the returns followed a normal distribution is by employing a Jarque-Bera test, as suggested by Cryer and Chan (2009). The following test is based on the fact that returns with a normal distribution have zero kurtosis and zero skewness. Once having assumed that the data are independent and identically distributed, test statistics is defined as follows:

$$JB = \frac{nS^2}{6} + \frac{n(K-3)^2}{24} \tag{28}$$

Under the null hypothesis, JB follows approximately a Chi-squared distribution with two degrees of freedom. S is sample skewness and K is the sample kurtosis of the returns considered. Table 1 indicates the Jarque-Bera statistics for the different S&P500 sectors. For each sector this value exceeded the critical value of a Chi-Squared distribution with a 5% confidence level $\chi^2_{2,0.05} = 5.991$, indicating that the null hypothesis was rejected and that the time series did not show evidence to follow a normal distribution.

3.3 Methodology

This paragraph explains the empirical methodology used for each model in the out-of-sample analysis. The choice of the window length for the rolling or expanding windows were made with the aim of achieving more accurate and reliable results from the model.

3.3.1 Historical VaR

The Historical VaR is a non-parametric approach to estimate Value-at-Risk. This implies that it does not assume any specific distribution of the asset returns. Rather than making assumptions, HV estimates VaR by finding the α-quantile of the previous N returns. The estimation covered a period from 2 January 2000 to 26 December 2021. More precisely, the VaR estimate for week t, i.e., VaR_t^α was found in the following way:

1. Calculate the α-quantile of the weekly returns from the previous 104 weeks (2 years), i.e., using the returns from week *t-104* to *t-1*.

3.3.2 Normal VaR

The idea behind the Normal VaR is that if the normality assumption about the return distribution is acceptable, then the model gives an accurate estimation of the Value-at-Risk. Starting from 2 January 2000 to 26 December 2021, these were the following steps to find the estimates:

1. Estimate the mean (μ_t) and variance (σ_t^2) of the weekly returns using the previous 52 weeks (1 year), i.e., using the returns from week *t-52* to *t-1* in the formulas in the equations (2) and (3).
2. Plug the estimated mean and variance from the previous step into formula (4) to arrive at the predicted Value-at-Risk for week *t*, VaR_t^α.

3.3.3 CAViaR model

The CAViaR model was estimated using the original code from Simone Manganelli's website[2]. As before, the estimation of parameters γ started from 2 January 2000 to 26 December 2021 using an expanding. Better results justify the choice of using an expanding window instead of a rolling window (the results from the rolling window estimation are omitted from the thesis). More precisely, these were the steps followed:

1. Estimate the model in equation (6) using weekly returns from January 1995 (*t=1*) to week *t-1* to obtain $\hat{\gamma}_{0,t-1}$, $\hat{\gamma}_{1,t-1}$, $\hat{\gamma}_{2,t-1}$, and $\widehat{VaR}^{\alpha}_{t-1}$ (expanding window). Note that the *t-1* subscript indicates that the estimation uses data up until week *t-1*.
2. The predicted Value-at-Risk for week *t* is:

$$VaR_t^{\alpha} = \hat{\gamma}_{0,t-1} + \hat{\gamma}_{1,t-1} * \widehat{VaR}^{\alpha}_{t-1} + \hat{\gamma}_{2,t-1} * |R_{t-1}| \tag{29}$$

3.3.4 Cross-sectional quantile regression model

The CSQR model was estimated weekly, using every firm in the S&P500 index for that week, and then the predicted stock-specific quantiles were averaged for each sector separately (using only the stocks in that sector) to obtain a VaR prediction for the sector portfolios. This method involved a cross-sectional regression using the characteristics of all S&P500 firms starting from 2 January 2000 to 26 December 2021 using a 3 weeks as a rolling window. To be more precise, these are the steps followed:

1. Estimate the cross-sectional quantile regression in each week throughout the sample using *all the stocks* that are in the S&P500 index during that week:

$$Q_{R_{it}}(\alpha) = \beta_{0t} + \beta_{1t}X_{1it-1} + \cdots + \beta_{Kt}X_{Kit-1} \tag{30}$$

It should be emphasized that the X values are taken from week *t-1*.

[2] Simone Manganelli's website: http://www.simonemanganelli.org/Simone/Research.html

2. To predict VaR_t^α of a sector portfolio, the average of the β estimates from the *previous three weeks* are used. In particular, let

$$\bar{\beta}_{k,t-3:t-1} = \frac{1}{3}\sum_{j=1}^{3}\hat{\beta}_{kt-j} \tag{31}$$

   where the $\hat{\beta}$ values are the estimates from the quantile regressions in step 1. It is important to emphasize that information from week *t-4* to *t-1* is used to calculate $\bar{\beta}_{k,t-3:t-1}$.

3. Let S_{st} denote the set of stocks that belong to sector *s* during week *t*, and n_{st} the number of such stocks. Then,

$$VaR_t^\alpha = \sum_{i \in S_{st}} w_{it} * \left(\bar{\beta}_{0,t-3:t-1} + \bar{\beta}_{1,t-3:t-1} * X_{1it-1} + \cdots + \bar{\beta}_{K,t-3:t-1} * X_{Kit-1}\right) \tag{32}$$

   where w_{it} is the weight of stock *i* in the portfolio. For the equal-weighted case $w_{it} = \frac{1}{n_{st}}$, and for the value-weighted case w_{it} capitalization share of stock *i* in the sector portfolio. Note how only information up to week *t-1* is used to calculate VaR_t^α.

In this thesis, two cross-sectional models were proposed. The first model, referred to as CSQR1, was computed with the following covariates: company size (*SIZE*), one-year beta (*BETA*), Book-to-market ratio (*BM*), asset growth or investment (*GROWTH*), and profitability (*ROE*). The following equation represents the first cross-sectional quantile regression model:

$$\begin{aligned} R_{it}^{CSQR1}(\alpha) = \beta_{0t} + \beta_{SIZEt}SIZE_{it-1} + \beta_{BMt}BM_{it-1}\beta_{GROWTHt}GROWTH_{it-1} \\ + \beta_{ROEt}ROE_{it-1} + \beta_{BETAt}BETA_{it-1} \end{aligned} \tag{33}$$

The covariate *SIZE* represents the market capitalization. The regressor *BM* is the annual book value of equity divided by the market value of equity. Finally, the covariate *BETA* is simply the market beta calculated over a window of 52 weeks.

I also proposed a second CSQR model, which relied on a different set of cross-sectional variables.. This new model, called CSQR2, involved the following covariates:

- $MEAN1_{it-1}$ is the mean of the weekly returns of stock *i* calculated over weeks t-26 to t-1;
- $MEAN2_{it-1}$ is the mean of the weekly returns of stock *i* calculated over weeks t-52 to t-27;
- VOL_{it-1} is the volatility of the weekly returns of stock *i* calculated over weeks t-52 to t-1;
- $SKEW_{it-1}$ is the skewness of the weekly returns of stock *i* calculated over weeks t-52 to t-1.

The values of these covariates were calculated by me using the weekly return series of the individual stocks. The same methodology was used to perform the out-of-sample analysis, and the following equation represents the new model:

$$R_{it}^{CSQR2}(\alpha) = \beta_{0t} + \beta_{MEAN1_t} MEAN1_{it-1} + \beta_{MEAN2_t} MEAN2_{it-1} + \beta_{VOL_t} VOL_{it-1} + \beta_{SKEW_t} SKEW_{it-1} \tag{34}$$

These covariates were inspired by the cross-sectional asset pricing literature. *MEAN1* and *MEAN2* were inspired by momentum effect, introduced by Jegadeesh and Titman (1993), who show that stocks' past return positively predicts their future return (winners continue to win and losers continue to lose). Separating past year's returns into two half-year periods is inspired by Novy-Marx (2012), who showed that momentum is primarily driven by firms' performance 12 to seven months prior to portfolio formation. *VOL* is inspired by Ang et al. (2006), who showed that the cross-section of past volatility negatively predicts expected returns. Finally, *SKEW* is inspired by Amaya et al. (2015), who showed that the skewness of stocks' past returns negatively predicts future returns in the cross-section.

Note that the above cited studies establish these covariates as good predictors of future expected returns. The question that I study in this thesis is whether these characteristics are also helpful for predicting left-tail quantiles (i.e., the Value-at-Risk) of the return distribution.

Note that all the covariates in both the CSQR1 and CSQR2 models were normalized cross-sectionally each week, meaning that each variable was subtracted from its mean and then divided by the standard deviation. This procedure can help to reduce the influence of outliers and can help to interpret the coefficients. The formula for normalizing a variable is:

$$z = \frac{x - \bar{x}}{std(x)} \tag{35}$$

After the normalization, each variable has a mean of 0 and a standard deviation of 1.

3.3.6 Limitations

The results obtained from the application of the models presented in this thesis are strictly related to the length of the rolling or expanding window used. Setting a different window length, may led to slightly different results. However, the purpose of the thesis is to show how straightforward cross-sectional models can leverage information from other firms in the market and become accurate predictors of risk compared to more complex models. Indeed, cross-sectional models are easy, fast to implement, very intuitive to understand, and they led to better results than the more complex model CAViaR.

4 Results

4.1 Forecast

The following chapter shows a graphical illustration of the out-of-sample application of the different Value-at-Risk methodologies object of this thesis: Historical VaR, Normal VaR, CAViaR, CSQR1, and CSQR2.

As previously stated, the graphical illustration of the results features the Financials sector, however, the results are very similar for all sectors with the exception of the Technology sector. The appendix reports the graphical results of all the sectors, including this latter, to illustrate the significant Value-at-Risk forecast at the beginning of the 2000s due to the Dot-com bubble.

The next figures display the weekly returns of the Financials sector and the Value-at-Risk at a 95% level of confidence. However, it will be discussed later that the results are similar also for 98%, and 99% CI. For an intuitive representation, all the figures the returns are represented by a black line and the VaR by a red line.

4.1.1 Historical VaR

The Historical VaR (HV) method is a non-parametric approach to estimate Value-at-Risk. As discussed in the theoretical background chapter, this method does not presuppose any specific distribution of the asset returns. Instead, HV estimates VaR by fitting the α-quantile of the previous N returns. For this reason, the expectation was not to find highly accurate results, but to find very approximative estimations of the market risk.

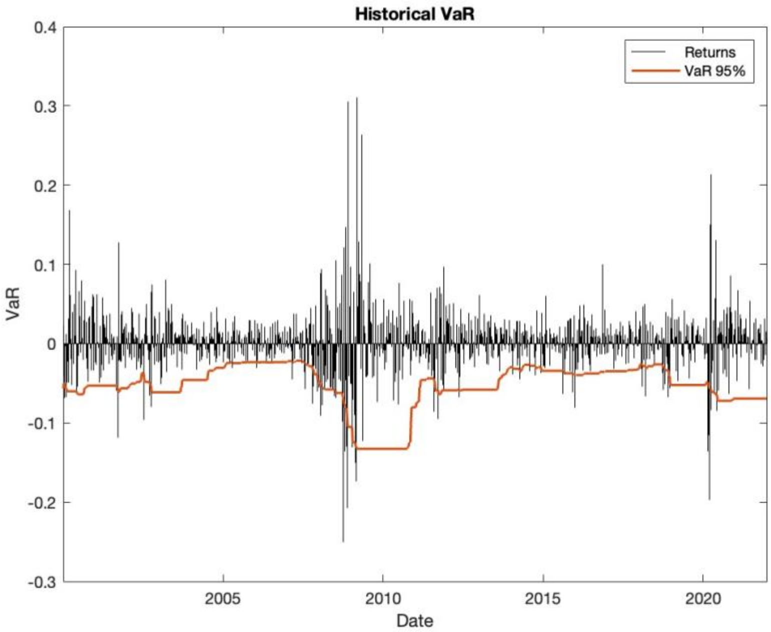

Figure 1: Plot of Historical VaR (Financials)

From figure 1, one may notice that the accuracy of the Historical method is uncertain. This approach provides a curve that tends to behave as a piecewise constant curve. This is due to the fact that quantiles tend to remain fixed for several weeks until extreme events occur. Consequently, the HV method is slow in responding to changes in volatility. This phenomenon can be easily seen during the Financial Crisis, where the returns in the Financials sector showed, for example, very high volatility starting from September 2008 and a weekly return of *-0.25*% in October 2008, but the historical VaR adjusted this information only after a few months. The same performance is seen in 2020, when the Financials sector experienced a significant drop of *-13.6*% in February 2020 due to the Covid Pandemic. The Historical VaR, however, did not incorporate this information immediately. Indeed, a few months of delay are always present after a significant drop in returns.

4.1.2 Normal VaR

The idea behind the Normal model is that if the assumption of normality on the distribution is acceptable, then the model gives accurate estimation values of the Value-at-Risk.

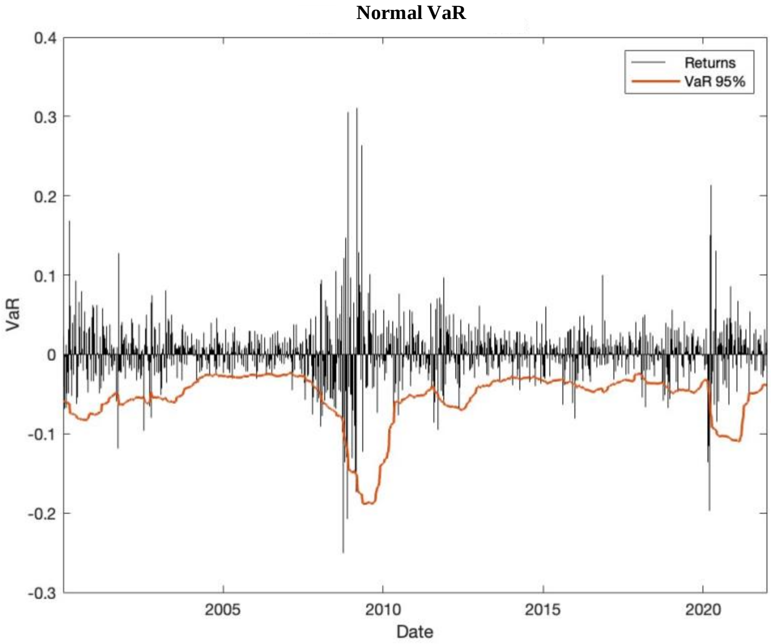

Figure 2: Plot of Normal VaR (Financials)

From figure 2, one may say that the Normal VaR method is still not providing accurate estimates, but it must be noticed that it reacted quicker to market shocks compared to the previous method. Indeed, both during the Financial Crisis in 2009 and during the Covid Pandemic crisis, the VaR reacted quicker to the negative drops. However, it still provided slow and smooth estimates, and during periods of not stressed markets, like for example in 2013 or in 2017 the Normal VaR overestimated the market risk. Even in this case, it is evident from the figure that this method adjusted the new information with a significant delay.

4.1.3 CAViaR method

Next figure displays the results from CAViaR model.

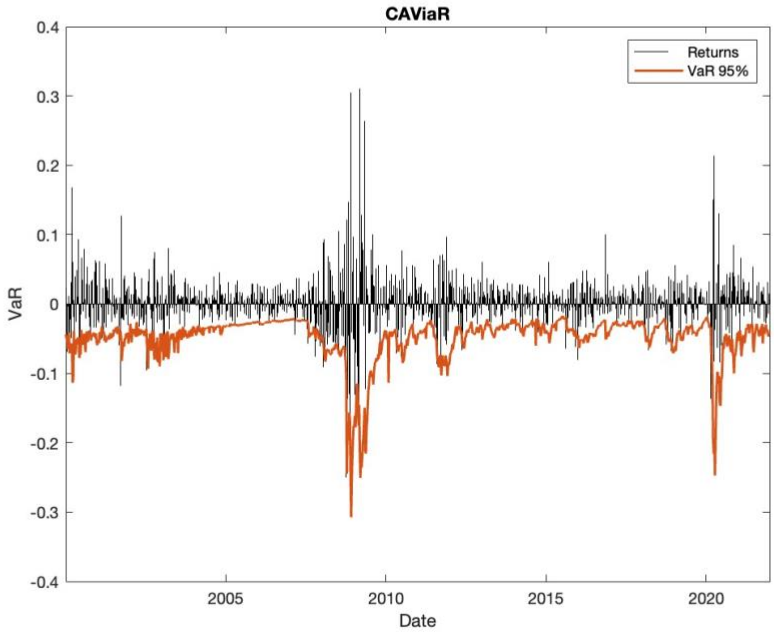

Figure 3: Plot of CAViaR (Financials)

From figure 3, the CAViaR method provided better VaR estimates. It can be noticed that this model quickly reacted to market shocks both during the Financial Crisis in 2009 and during the Covid Pandemic crisis. In addition, during periods of not stressed markets, the CAViaR model does not overestimated the market risk as much as the Normal VaR model. For example, during the entire period from 2004 to 2007, the VaR values were extremely low. In other words, in normal times, the CAViaR model did not provide high VaR estimates (in absolute value), but when a shock occurs, the model adjusted quite quickly to that shock and provided higher estimates of the risk.

4.1.4 CSQR models

As previously stated, the CSQR1 model was computed with different covariates: company market capitalization (*SIZE*), one-year beta (*BETA*), Book-to-market ratio (*BM*), asset growth (*GROWTH*), and profitability (*ROE*). Instead, the CSQR2 model was computed with the following covariates: *MEAN1*, *MEAN2*, *VOL*, and *SKEW*. The next figures display the results obtained from both cross-sectional methods.

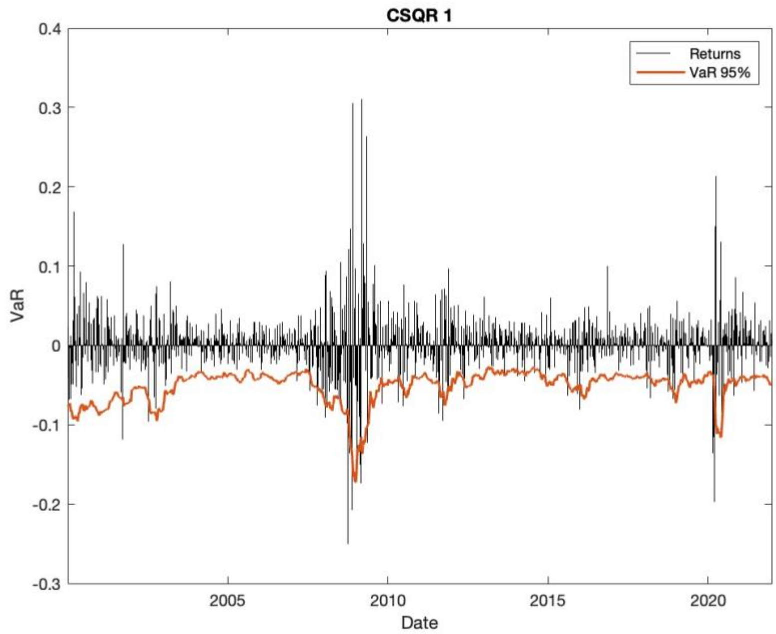

Figure 4: Plot of CSQR1 (Financials)

Figure 4 shows the VaR values predicted by the CSQR1 method. It is not intuitive to see that it provides even more accurate VaR estimates than the CAViaR. However, this is the result that will be found in the next section. Indeed, by incorporating the data of other companies in the cross-sectional regression, this model rapidly adapted its VaR estimates to market shocks. Also, during the Financial Crisis and the Covid Pandemic, this model provided quickly adjusted estimates.

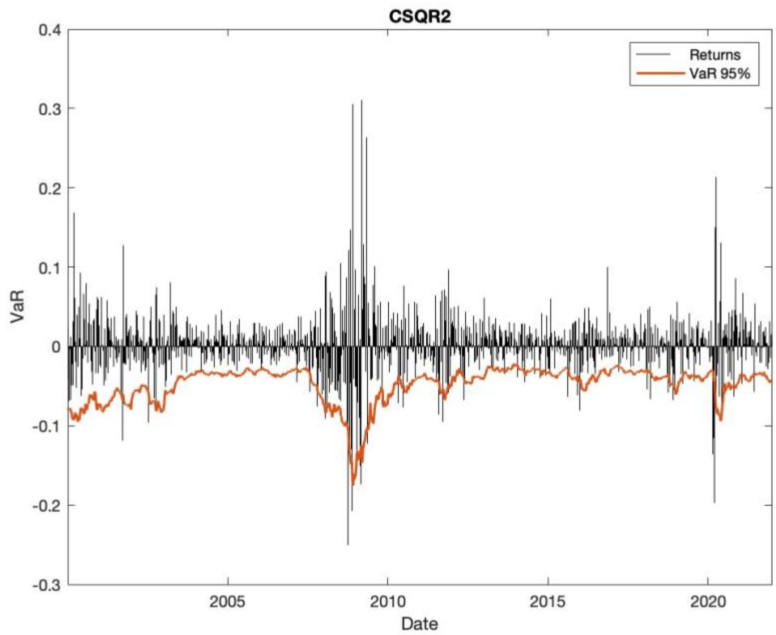

Figure 5: Plot of CSQR2 (Financials)

From figure 5, the CSQR2 method provided very similar estimation results to CSQR1. In the evaluation performance paragraph will be discussed which of the two models provided the best estimates.

4.1.5 Comparison of CAViaR and CSQR methods:

The following paragraph compares the plots of the CAViaR model and the CSQR models in the estimation of Value-at-Risk. As before, the Financials sector was taken as an illustrative example. The dataset includes both the Financial Crisis (2007-2011) and the Covid Pandemic (2020-2021). As shown in the figure below, an important consideration is that during stressed market conditions, the CAViaR model presented higher forecasts (in absolute value) of the fitted quantile VaR compared to the CSQR models. More precisely, during the Financial Crisis (when the sector reached -*25*%), the CAViaR reached forecasts of -*30.7*%, while the CSQR models reached only -*17.5*%. The same happened during the Covid period (when the sector reached -*19.7*%), the first model reached -*24.7*% while the cross-sectional ones reached -*11.7*%.

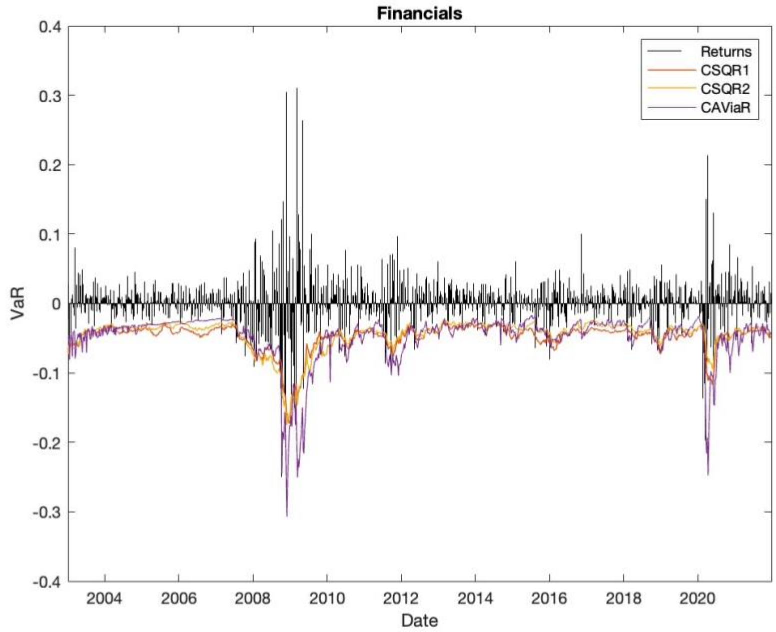

Figure 6: CAViaR, CSQR1, CSQR2 comparison (Financials)

Note that this phenomenon happened in the Financials sector, but it was not the case for all the sectors (see subsection 6.1 appendix). In fact, for example, the Consumer Staples and Health Care sectors had the opposite results during the Financial crisis, i.e., the CSQR models gave higher forecasts (in absolute value) than other models (figures 11 and 12, 6.1 appendix). Indeed, as it is described in the next paragraph (Evaluation performance subsection), according to the different evaluation tests, the model that on average performed better during volatile periods, was not the CAViaR, but CSQR1 and CSQR2 instead.

The final consideration is that the CAViaR model provides higher estimates in absolute value during stressed market conditions, however, the CSQR models did not overestimate the market risk as much as other models when the market was not stressed, and they were quicker to react and adjust new information from the market.

4.2 Beta estimates

This subsection briefly presents the average β estimates and the corresponding t-statistics found in the two cross-sectional models. As clarified before, these models involved cross-sectional quantile regressions where the parameters β were estimated weekly. It may be useful, as other authors did in the past, to examine these estimates, their signs, and the respective t-statistics. Table 2 reports the average estimates of the parameters β given by model CSQR1, and table 3 those by the model CSQR2. These results were obtained considering a 95% level of confidence, however similar results were found with 98% and 99% CI.

	constant	*SIZE*	*BM*	*GROWTH*	*ROE*	*BETA*
average β	-0.0544***	0.0022***	-0.0020***	-0.0025***	-0.0003	-0.0095***
t-stat	-25.1996	8.5049	-3.9910	-6.6139	-0.9435	-12.1791

Table 2: Summary statistics model CSQR1

From this table, one may notice that all the covariates in model CSQR1 are significantly different from zero at a 99% level of confidence, except for the covariate *ROE*. For the correct interpretation, recall that the Value-at-Risk in this thesis is expressed in terms of returns (i.e., negative values of VaR). The average parameter sign of the factor *SIZE* is positive (*0.0022*), indicating that bigger companies by market capitalization, tend to have higher Value-at-Risk threshold (i.e., decrease the VaR in magnitude). On the other hand, the smaller companies have lower Value-at-Risk, and this is due to the fact the smaller companies are usually riskier than the bigger ones. The average parameter of factor *BM* showed a negative sign (-*0.0020*), indicating that value companies (those with high book-to-market ratios) tend to have lower VaR, or in other words, tend to have higher risk. This means that higher values of book-to-market ratio are perceived as riskier. From the same table, Value-at-Risk tend to decrease with *GROWTH*, since the average parameter sign is negative (-*0.0025*). This intuitively indicates that the firms that direct their revenues towards high-growth projects are more likely to experience monetary losses, or in other words, they are riskier. Finally, the average VaR decreases with *BETA* (-*0.0095*). A higher *BETA* indicate that the firm's returns are more sensitive to market movements, and for this reason, it is intuitive to expect a negative sign for this parameter. Another consideration is regarding the size of the parameters. Indeed, since all the variables were normalized, one may compare their size. *BETA* has the highest coefficient, and *SIZE*, *BM*, and *GROWTH* have similarly sized coefficients.

	constant	*MEAN1*	*MEAN2*	*VOL*	*SKEW*
average β	-0.0539***	0.0026***	0.0004	-0.0179***	0.0001
t-stat	-26.8517	5.5991	1.3982	-21.2443	0.2826

Table 3: Summary statistics model CSQR2

Table 3 reports the average of the parameters estimated in model CSQR2 and the respective t-statistics. The covariates *MEAN1* and *VOL* resulted to be significantly different from zero at 99% level of confidence. The variable *VOL* is the one with the biggest highest size, while *MEAN2* and *SKEW* have similar size coefficients. The sign of the average parameters of *MEAN1* is positive (*0.0026*), meaning that companies with higher mean (calculated over weeks t-26 to t-1) tend to have higher VaR (i.e., smaller VaR in magnitude). On the other hand, companies with lower mean tend to have lower VaR, meaning they are perceived as riskier. Finally, as expected, the average parameter sign of *VOL* is negative (*-0.0179*), meaning that firms with higher standard deviation (over weeks t-52 to t-1) tend to decrease the Value-at-Risk, or in other words, they are perceived as riskier.

4.3 Evaluation performance

The following section presents the results obtained after having performed the most common Value-at-Risk evaluation techniques: percentage of violations, Score value, POF test, TUFF test, TBF test, and CCI test. The out-of-sample analysis covered the period from 2 January 2000 until 26 December 2021. Weekly VaR estimates were considered.

4.3.1 Performance summary

All the tests were performed for each model and for each individual sector. All detailed results can be found in table 13 of the appendix. Instead, table 4 shows a summary of the results obtained. Table 4 displays the results with an analysis of the VaR models with a 95% of confidence level.

2000-2021	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	5.5908	0.3955	0	**11**	9	0	2
NORMAL	5.0919	0.38378	0	**11**	9	0	3
CAVIAR	5.3532	0.3644	4	10	9	0	10
CSQR1	4.4663	0.3566	2	7	8	1	10
CSQR2	**4.9572**	**0.3530**	**5**	7	**10**	**3**	**11**

Table 4: Summary out-of-sample S&P500 sectors VaR evaluation (95% CI)

The first column called *Viol.%* represents the mean of the proportion of times that the actual losses of the 11 S&P500 sectors exceeded the estimated Value-at-Risk threshold, or in other words, it represents the percentage of violations given on average by each method across the 11 S&P500 sectors. Ideally, the percentage of violations should be close to the chosen confidence level. In this case, it should be close to 5%. It can be observed that the model that adapts best to the data on average is the CSQR2, with a percentage of violations closest to 5%. The worse method is the Historical one, with *5.59*% of violations, meaning that it underestimates the market risk.

The second column, called *Average score*, represents the average of the Score value across the 11 sectors. According to the Score function, the best estimate model is still the CSQR2 (with a *0.353* average score) followed by the CSQR1 (with a *0.356* average score), meaning that it provides estimates with the lowest distance between the fitted quantile and the actual data. A lower score indicates a better performance of the VaR model in estimating the desired α-quantile. Therefore, when the score function is at its lowest for a specific VaR model, it means that the model has achieved the best possible fit to the data, and it is more accurate in forecasting the quantile at the given probability level. In other words, a lower score implies that the predicted quantiles are closer to the observed values, indicating better forecasting accuracy of the VaR model.

The third column, named *Score winner,* represents the number of times that that model was the best-performing one in a specific sector. For example, the Historical VaR was never the best model, therefore the number of times it resulted to be the best-performing model is zero. As shown in the table, the CSQR2 was the best model in 5 sectors out of 11, while the CAViaR was the best model in 4 sectors.

Columns 4 to 8 present the results obtained from each test. To be more precise, they report the number of sectors where that specific model was accepted. For example, table 4 reports a number 11 in the CCI test for the CSQR2 model. This means that in 11 out of 11 models that specific model was accepted. Therefore, it suggests that the model was good.
According to the TUFF, TBF, and the CCI test, the model that performed the best most of the time was the CSQR2. This model was accepted 10 out of 11 times in the TUFF test, 3 out of 11 times in the TBF test and 11 out of 11 times in the CCI test. When the POF test is accepted, it means that the test statistic is less than the critical value and the VaR model accurately estimates the VaR. Accepting the CCI test means failing to reject the null hypothesis, which states that the exceedances are independent of each other. This implies that there is no evidence to suggest that the observed exceedances deviate significantly from what is expected under the assumption of independence. In other words, the test indicates that the exceedance rate and the patterns of exceedances in the data are consistent with what would be expected under the null hypothesis.

A final consideration is that almost no model passed the TBF test. Indeed, the cross-sectional models were the only ones to be accepted by the TBF test. CSQR2 is still the model that was accepted the most among all the models. When a model fails to reject the TBF test, it means that the violations do not seem to be correlated. This means that the cross-sectional models are the only models that showed uncorrelated results, and the CSQR2 is the model that showed more times uncorrelated results with respect to other models.

According to this analysis, the CSQR2 is the most accurate model to predict market risk. This can be attributed to the additional information provided by the cross-sectional regression which is then used in the forecast of the Value-at-Risk.

Tables 5 and 6 below performed all the tests for the 98% and 99% VaR-s.

2000-2021	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	2.9142	0.2203	0	5	8	0	0
NORMAL	**3.1517**	0.2110	0	1	10	0	0
CAVIAR	2.6449	0.1948	1	**7**	7	0	1
CSQR1	**1.6155**	0.1831	3	5	10	4	**6**
CSQR2	3.1517	**0.1806**	**7**	6	**11**	**5**	**6**

Table 5: Summary out-of-sample S&P500 sectors VaR evaluation 2000-2021 (98% CI)

2000-2021	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	1.7342	0.1358	0	3	9	0	5
NORMAL	2.3282	0.1366	0	0	8	0	1
CAVIAR	1.5996	0.1202	1	5	5	1	**7**
CSQR1	0.7602	0.1120	3	**7**	**9**	**6**	5
CSQR2	**0.8394**	**0.1096**	**7**	**7**	**9**	5	6

Table 6: Summary out-of-sample S&P500 sectors VaR evaluation 2000-2021 (99% CI)

Similar results are found with 98% and 99% level of confidence. Indeed, according to both tables, the model with the lowest average score was CSQR2 followed by CSQR1. As before, according to TUFF, TBF, and CCI tests, the CSQR2 model is the one that was accepted most of the time across all the S&P500 sectors. Even according to the percentage of violations, the cross-sectional models are the ones that gave closer results to the specific significance level.

4.3.2 Performance summary during Financial Crisis

Table 7 summarizes the evaluation performance of the models during the financial crisis. The out-of-sample analysis covers the period from 7 January 2007 to 2 January 2011.

2007-2011	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	8.0035	0.6020	0	6	11	0	7
NORMAL	6.9595	0.5543	0	**10**	11	1	7
CAVIAR	7.7860	0.5107	0	6	11	3	**10**
CSQR1	**6.6551**	0.4519	1	7	11	6	9
CSQR2	7.0030	**0.4315**	**10**	8	11	**8**	**10**

Table 7: Summary out-of-sample S&P500 sectors VaR evaluation 2007-2011 (95% CI)

Based on these findings, even in times of high volatility periods like during the Financial crisis period, the cross-sectional methods show better results. Indeed, according to the *Average score* and the *Score winner*, for all the 11 sectors, the cross-sectional methods were the best to predict the risk. More precisely, in 10 out of 11 sectors, the best model resulted to be the CSQR2, and the CSQR1 was the best one for the remaining sector. Furthermore, the CSQR2 was the best model even according to the TBF, and CCI tests. Indeed, the TBF test was accepted in 8 out of 11 sectors, and the CCI test in 10 out of 11 tests. According to the POF test, the CSQR2 showed good results, with 8 out of 11 acceptances. Also, an important consideration from *Viol.%* is that all the models showed a higher percentage of violations than 5%, meaning that all the models underestimated the risk. The additional findings presented in the appendix, corresponding to confidence levels of 98% and 99%, align closely with those reported above for the 95% confidence level (subsection 6.3 appendix).

4.3.3 Performance summary during Covid Pandemic

The next table summarizes the evaluation performance of the models during the Covid period. The out-of-sample analysis is performed on the weekly returns of each S&P500 sector, and it covers the period from 3 June 2018 to 26 December 2021.

2018-2021	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	5.5907	0.5030	0	**11**	11	3	9
NORMAL	**4.6667**	0.5048	0	**11**	11	1	9
CAVIAR	6.1254	0.4405	2	10	11	9	**11**
CSQR1	5.9796	**0.4160**	**8**	10	11	**10**	**11**
CSQR2	6.3685	0.4304	1	9	11	**10**	**11**

Table 8: Summary out-of-sample S&P500 sectors VaR evaluation 2018-2021 (95% CI)

According to this analysis, the best-performing models during the Covid period were still the cross-sectional methods. Similar results were found in Table 4. Indeed, the CSQR1 presents on average the lowest Score values, and it is the method that is accepted the most across the sectors in TBF and CCI tests. To be more precise, CSQR1 and CSQR2 were the best-performers in 9 out of 11 sectors. Subsection 6.3 of the appendix shows the results during the Covid Pandemic for 98 and 99% CI.

To conclude, based on the out-of-sample analysis performed from 2000 to 2021, the cross-sectional methods were the models that performed better. Furthermore, during highly stressed market conditions like the Financial Crisis and the Covid Pandemic, the cross-sectional methods were even more reliable.

4.3.4 CAViaR, CSQR1, and CSQR2 for each sector

This paragraph covers the evaluation performance of CAViaR and CSQR models for each equally weighted sector in the S&P500 index. The aim is to investigate the sectors that showed the best results, those that showed the worst results, and see if there is consistency among the different models.

2000-2021	*Viol. %*	*Score value*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
Energy	5.2265	0.4954	accept	accept	reject	accept
Materials	4.8781	0.3745	accept	accept	reject	accept
Industrials	6.3589	0.3572	reject	accept	reject	accept
Consumer Discretionary	6.1847	0.3788	accept	accept	reject	accept
Consumer Staples	**4.9652**	**0.2359**	accept	accept	reject	reject
Health Care	5.5949	0.2940	accept	accept	reject	accept
Financials	5.8362	0.3923	accept	accept	reject	accept
Information Technology	4.6167	0.4233	accept	reject	reject	accept
Communication Services	5.1394	0.3602	accept	reject	reject	accept
Utilities	5.5749	0.3080	accept	accept	reject	accept
Real estate	4.5296	0.3889	accept	accept	reject	accept

Table 9: CAViaR method for each S&P500 sector 2000-2021 (95% CI)

According to these findings, the CAViaR model performed the best in the Consumer Staples sector, with a Score value of *0.2359* and with a percentage of violations very close to the desired 5% (*4.97*%). The worst performance was for the Energy sector, with a Score value of *0.4954*.

2000-2021	*Viol. %*	*Score value*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
Energy	9.5819	0.5136	reject	accept	reject	accept
Materials	4.4425	0.3631	accept	accept	reject	accept
Industrials	3.8328	0.3322	accept	accept	accept	accept
Consumer Discretionary	4.0070	0.3536	accept	reject	reject	accept

Consumer Staples	1.8293	**0.2596**	reject	accept	reject	accept
Health Care	2.6132	0.3016	reject	accept	reject	reject
Financials	**4.7038**	0.3719	accept	reject	reject	accept
Information Technology	4.3554	0.4081	accept	reject	reject	accept
Communication Services	3.5714	0.3277	reject	accept	reject	accept
Utilities	3.9199	0.3018	accept	accept	reject	accept
Real estate	6.2718	0.3900	accept	accept	reject	accept

Table 10: CSQR1 method for each S&P500 sector 2000-2021 (95% CI)

Even according to these results, the sector where the CSQR1 model performed better in the Consumer Staples sector, with a Score value of *0.2596*. However, the percentage of violations is *1.83*% which is very far from the desired 5%. Even the worst performance was again for the Energy sector, with a Score value of *0.5136*.

2000-2021	*Viol. %*	*Score value*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
Energy	8.6237	0.5055	reject	accept	reject	accept
Materials	**5.0523**	0.3532	accept	accept	accept	accept
Industrials	5.1394	0.3293	accept	accept	reject	accept
Consumer Discretionary	4.1812	0.3534	accept	accept	accept	accept
Consumer Staples	2.2648	**0.2507**	reject	accept	reject	accept
Health Care	2.7875	0.3060	reject	accept	reject	accept
Financials	5.8362	0.3591	accept	accept	reject	accept
Information Technology	4.2683	0.4102	accept	accept	reject	accept
Communication Services	3.8328	0.3307	accept	reject	reject	accept
Utilities	4.6167	0.2940	accept	accept	accept	accept
Real estate	7.9268	0.3912	reject	accept	reject	accept

Table 11: CSQR2 method for each S&P500 sector 2000-2021 (95% CI)

Finally, even according to the last table, the CSQR2 model performed better in the Consumer Staples sector, with a Score value of *0.2507*. However, even in this case, the percentage of violations is far from the desired 5% (*2.26*%). As before, the worst performance was for the Energy sector, with a Score value of *0.5055*.

Summarizing the last three tables, the final consideration is that all three models (the CAViaR, the CSQR1, and the CSQR2) performed better in the Consumer Staples sector and performed the worst in the Energy sector.

4.3.5 Summary results value-weighted

This subsection reports the summary results across the 11 value-weighted S&P500 sectors, meaning that each firm in the sector was weighted by its market capitalization. The aim is to shows that the previous results can be extended also in the value-weighted case.

2000-2021	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	5.5908	0.3605	0	**11**	9	0	3
NORMAL	**5.0998**	0.3524	0	**11**	**10**	0	2
CAVIAR	5.1156	0.3406	3	**11**	9	1	8
CSQR1	4.6246	0.3348	1	6	9	2	**10**
CSQR2	5.1156	**0.3289**	**7**	4	**10**	**4**	**10**

Table 12: Summary S&P500 sectors VaR evaluation (95% CI) value-weighted

According to the average score, the score winner, the TUFF, TBF, and CCI test, the most accurate model to predict the risk from 2000 to 2021 was the CSQR2 model. Indeed, this model was the one presenting the lowest score on average across all the sectors (*0.3289*, second column), and it was the model with the lowest score function in 7 out of 11 sectors (see third column '*Score winner*'). According to the TUFF and CCI tests (fifth and seventh column respectively), CSQR2 was accepted in 10 out of 11 sectors. Finally, according to the TBF test, it was accepted the most among all the other models.

5 Conclusions

The aims of risk management are to help companies to identify and to control the risks that may negatively impact their operations. The final goal is to decrease monetary losses due to unpredictable events. The most popular and used method to evaluate market risk is the Value-at-Risk. This standardized tool is based on a mathematical and statistical model developed at the beginning of the 1990s in the financial industry to provide a company's management with a single number that could quickly and easily explain the risk of an asset or portfolio. From those years, many different models were proposed to find the Value-at-Risk. This thesis analyzed the performance of five different methods: Historical VaR, Normal VaR, CAViaR model, and two cross-sectional quantile regression models. The first two models are the most common and simple models. CAViaR is a model introduced by Simone Manganelli (2004), where rather than modeling the entire distribution, the quantile is modeled directly using a conditional autoregressive quantile specification. The last two models are simple models based on quantile regression with a cross-sectional approach. In other words, to find the VaR of a specific firm were used data of other firms in the same sector. This additional information from other companies in the same sector may help the risk manager to assess the market risk in a more efficient way.

An out-of-sample analysis was performed from 2 January 2000 until 26 December 2021, with 95%, 98%, and 99% levels of confidence. The objects of the analysis were the 11 S&P500 equally weighted sectors. The models that gave better results for all the levels of confidence were the cross-sectional ones. More importantly, the same analysis was conducted individually during two subperiods, during the Financial Crisis and the Covid Pandemic, finding the same outperforming results by the cross-sectional models. Indeed, in both periods, the cross-sectional methods showed a highly superior forecasting performance with respect to the other models. Due to the incorporation of the data from other firms in the cross-sectional regression, this model rapidly adapted its VaR estimates to market shocks, leading to a more accurate prediction of market risk. Finally, the same results were consistent with the value-weighted S&P500 sectors.

6 Appendix

6.1 CAViaR, CSQR1, CSQR2 plots S&P500 sectors

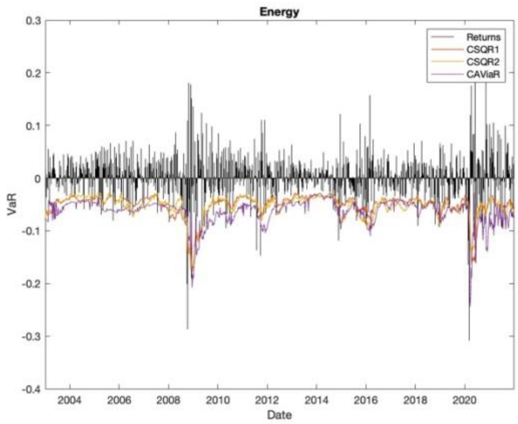

Figure 7: Energy

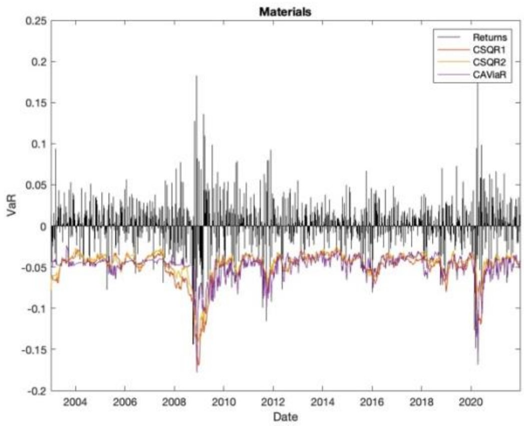

Figure 8: Materials

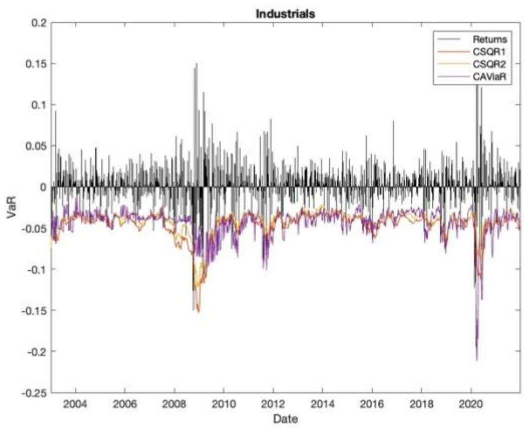

Figure 9: Industrials

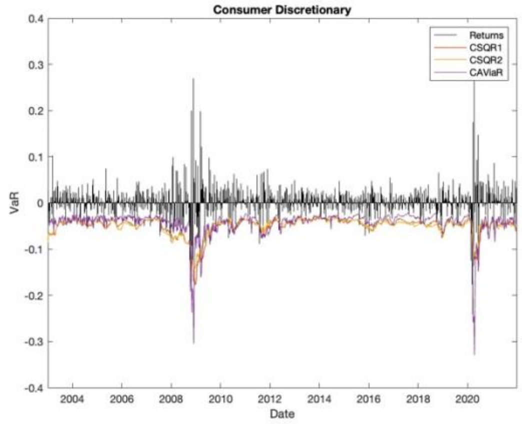

Figure 10: Consumer Discretionary

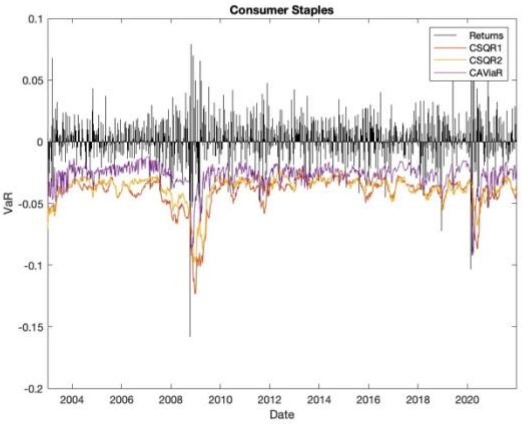

Figure 11: Consumer Staples

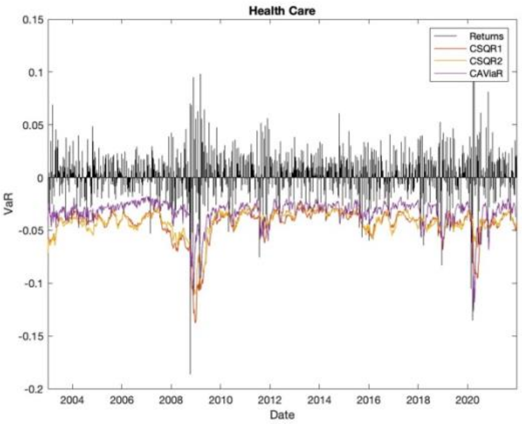

Figure 12: Health Care

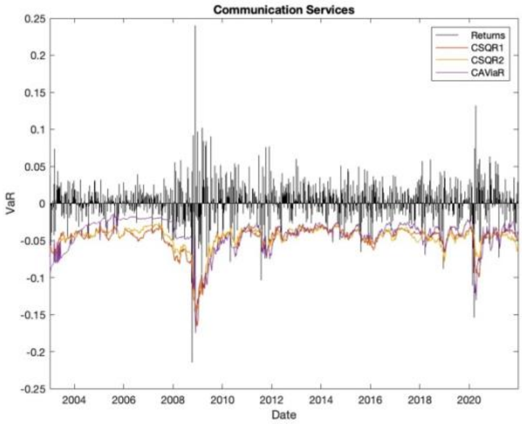

Figure 13: Communication Services

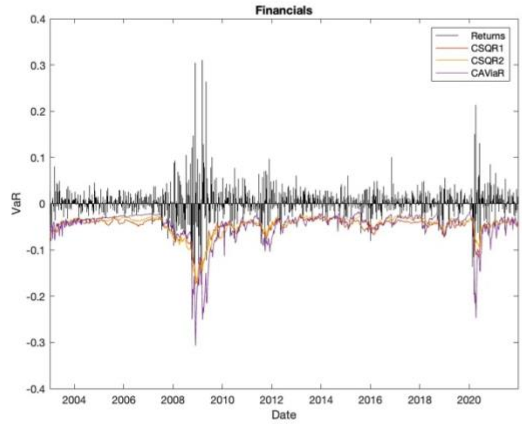

Figure 14: Financials

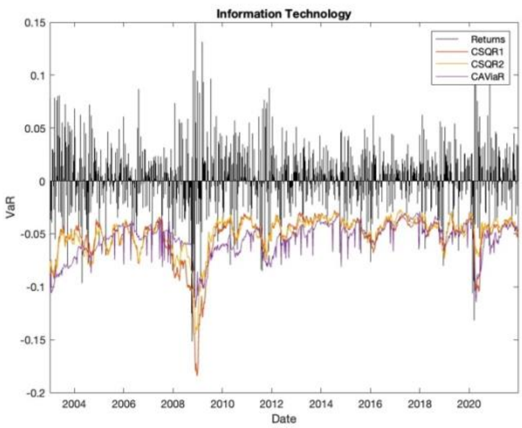

Figure 15: Information Technology

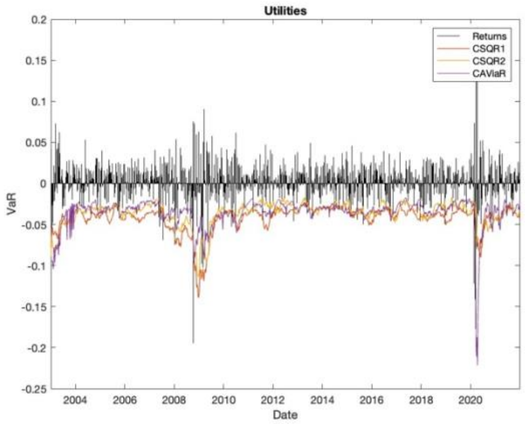

Figure 16: Utilities

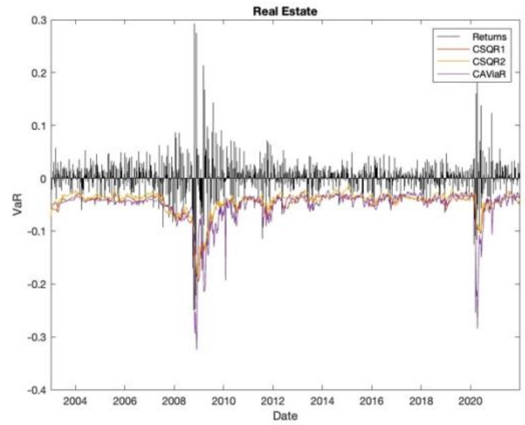

Figure 17: Real Estate

6.2 Detailed results

	2000-2021	*Viol. %*	*Score value*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
Energy	HS	6.0105	0.5468	accept	accept	reject	reject
	NORMAL	5.0523	0.5219	accept	accept	reject	reject
	CAVIAR	5.2265	0.4954	accept	accept	reject	accept
	CSQR1	9.5819	0.5136	accept	accept	reject	accept
	CSQR2	8.6237	0.5055	reject	accept	reject	accept
Materials	HS	5.7491	0.3974	reject	accept	reject	accept
	NORMAL	5.0523	0.3909	accept	accept	reject	reject
	CAVIAR	4.8780	0.3745	accept	accept	reject	accept
	CSQR1	4.4425	0.3631	accept	accept	reject	reject
	CSQR2	5.0523	0.3532	accept	accept	reject	accept
Industrials	HS	5.4878	0.3756	accept	accept	reject	accept
	NORMAL	4.9652	0.3676	accept	accept	accept	accept
	CAVIAR	6.3589	0.3572	reject	accept	reject	accept
	CSQR1	3.8328	0.3322	accept	accept	reject	reject
	CSQR2	5.1394	0.3292	accept	accept	reject	accept
Consumer Discretionary	HS	5.1394	0.4157	reject	accept	reject	accept
	NORMAL	4.9652	0.4068	accept	accept	accept	accept
	CAVIAR	6.1847	0.3788	accept	accept	reject	accept
	CSQR1	4.0070	0.3536	accept	accept	reject	accept
	CSQR2	4.1812	0.3534	accept	accept	reject	reject
Consumer Staples	HS	5.7491	0.2551	accept	accept	reject	reject
	NORMAL	4.7038	0.2480	accept	accept	reject	accept
	CAVIAR	4.9652	0.2359	accept	accept	reject	reject
	CSQR1	1.8293	0.2596	accept	reject	reject	accept
	CSQR2	2.2648	0.2507	accept	accept	accept	accept
Health Care	HS	5.2265	0.3162	accept	accept	reject	reject
	NORMAL	4.6167	0.3067	accept	accept	reject	reject
	CAVIAR	5.5749	0.2940	accept	accept	reject	accept
	CSQR1	2.6132	0.3016	accept	accept	reject	reject
	CSQR2	2.7875	0.3060	reject	accept	reject	accept
Financials	HS	5.6620	0.4419	reject	accept	reject	accept
	NORMAL	5.2265	0.4206	accept	accept	reject	reject
	CAVIAR	5.8362	0.3923	accept	accept	reject	accept
	CSQR1	4.7038	0.3719	accept	accept	reject	reject
	CSQR2	5.8362	0.3591	accept	accept	reject	accept

Information Technology	HS	5.2265	0.4384	reject	accept	reject	reject
	NORMAL	4.7909	0.4330	reject	accept	reject	accept
	CAVIAR	4.6167	0.4233	accept	reject	reject	accept
	CSQR1	4.3554	0.4081	accept	accept	reject	reject
	CSQR2	4.2683	0.4102	accept	accept	reject	reject
Communication Services	HS	5.7491	0.3822	accept	accept	reject	accept
	NORMAL	5.5749	0.3650	accept	reject	reject	accept
	CAVIAR	5.1394	0.3602	accept	reject	reject	accept
	CSQR1	3.5714	0.3277	accept	accept	reject	accept
	CSQR2	3.8328	0.3307	accept	reject	reject	accept
Utilities	HS	6.1847	0.3284	accept	reject	reject	accept
	NORMAL	5.6620	0.3259	accept	reject	reject	accept
	CAVIAR	5.5749	0.3080	accept	accept	reject	accept
	CSQR1	3.9199	0.3018	accept	reject	reject	accept
	CSQR2	4.6167	0.3307	accept	accept	reject	accept
Real estate	HS	5.3136	0.4527	accept	reject	reject	accept
	NORMAL	5.4007	0.4351	accept	reject	reject	accept
	CAVIAR	4.5296	0.3887	accept	accept	reject	accept
	CSQR1	6.2718	0.3900	accept	reject	reject	accept
	CSQR2	7.9268	0.3912	reject	accept	reject	accept

Table 13: HV, NORMAL, CSQR1, CSQR2 for all S&P500 sectors 2000-2021 (95% CI)

6.3 Summary results 98% and 99% CI during stressed market conditions

2007-2011	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	4.3497	0.3401	0	5	11	0	8
NORMAL	4.6977	0.3115	0	4	11	1	9
CAVIAR	4.3062	0.2817	0	5	11	0	8
CSQR1	**2.1749**	0.2121	1	**10**	11	**11**	**11**
CSQR2	2.5228	**0.1990**	**10**	8	11	10	**11**

Table 14: Summary out-of-sample S&P500 sectors VaR evaluation 2007-2011 (98% CI)

2007-2011	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	2.5228	0.2116	0	8	11	0	9
NORMAL	3.8278	0.2031	0	0	11	0	9
CAVIAR	2.3923	0.1707	0	8	11	3	9
CSQR1	0.8699	0.1188	2	6	11	**10**	**11**
CSQR2	**1.0439**	**0.1111**	**9**	**9**	11	**10**	**11**

Table 15: Summary out-of-sample S&P500 sectors VaR evaluation 2007-2011 (99% CI)

2018-2021	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	3.0141	0.3170	0	**11**	11	0	10
NORMAL	3.0627	0.3122	0	**11**	11	0	10
CAVIAR	3.1113	0.2363	3	**11**	11	4	10
CSQR1	**2.8683**	**0.2202**	**8**	10	11	8	**11**
CSQR2	2.9655	0.2298	0	10	11	**10**	**11**

Table 16: Summary out-of-sample S&P500 sectors VaR evaluation 2018-2021 (98% CI)

2018-2021	*Viol. %*	*Average score*	*Score winner*	*POF test*	*TUFF test*	*TBF test*	*CCI test*
HS	2.0418	0.2017	0	**11**	11	3	11
NORMAL	2.5766	0.2266	0	9	11	0	10
CAVIAR	1.9932	0.1445	**5**	10	11	**9**	10
CSQR1	**1.5071**	**0.1358**	4	10	11	**9**	10
CSQR2	1.6529	0.1404	2	10	11	7	**11**

Table 17: Summary out-of-sample S&P500 sectors VaR evaluation 2018-2021 (99% CI)

6.4 GICS codes

Energy	*Materials*	*Industrials*	*Consumer Discretionary*	*Consumer Staples*	*Health Care*	*Financials*	*Information Technology*	*Communication Services*	*Utilities*	*Real Estate*
10	15	20	25	30	35	40	45	50	55	60

Table 18: GICS codes for S&P500 sectors

7 Bibliography

Acerbi, C., and Szekely, B. (2014), ‘*Backtesting Expected Shortfall’*. Msci.
Available at:
https://www.msci.com/documents/10199/22aa9922-f874-4060-b77a-0f0e267a489b

Allena, D.E., and Singh, A.K. (2012), ‘*A gourmet’s delight: caviar and the australian stock market’*. Applied Economics Letters Vol. 19: 1493 – 1498.
Available at:
https://www.tandfonline.com/doi/full/10.1080/13504851.2011.636017?scroll=top&needAccess=true&role=tab&aria-labelledby=full-article

Amaya, D., Christoffersen, P., Jacobs, K. and Vasquez, A. (2015), ‘*Does realized skewness predict the cross-section of equity returns?*’. Journal of Forecasting Vol. 25: 101 – 128.
Avaiable at: https://www.sciencedirect.com/science/article/pii/S0304405X15001257

Ang, A., Hodrick, R.J., Xing, Y., and Zhang, X. (2006), ‘*The Cross-Section of Volatility and Expected Returns’.* The Journal of Finance Vol. 61: 259 – 299.
Available at: https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2006.00836.x

Bao, Y., Lee, T., and Saltoglu, B. (2006), ‘*Evaluating predictive performance of value-at-risk models in emerging markets: a reality check’*. Journal of forecasting.
Available at: https://onlinelibrary.wiley.com/doi/10.1002/for.977

Basel Committee on Banking Supervision (2005), ‘*An Explanatory Note on the Basel II IRB Risk Weight Functions* ‘. Bank for international settlements.
Available at: https://www.bis.org/bcbs/irbriskweight.pdf

Bautista, R.S., and Mora J.A.N. (2021), *‘Value-at-risk predictive performance: a comparison between the CaViaR and GARCH models for the MILA and ASEAN-5 stock markets’*. Journal of Economics, Finance and Administrative Science Vol. 26: 197 – 221.

Available at:
https://www.econstor.eu/bitstream/10419/253819/1/1789176263.pdf

Bollerslev, T. (1986), '*Generalized Autoregressive Conditional Heteroskedasticity'*. Journal of Econometrics Vol. 31: 307 – 327.
Available at: http://public.econ.duke.edu/~boller/Published_Papers/joe_86.pdf

Campbell, J.Y. (2017), '*Financial decisions and markets: a course in asset pricing'*. Princeton University Press.
Available at: https://www.perlego.com/book/740026/financial-decisions-and-markets-a-course-in-asset-pricing-pdf

Campbell, R., Yan, L., and Heqing, Z. (2016), '*...and the Cross-Section of Expected Returns'*. The Review of Financial Studies Vol. 29: 5 – 68.
Available at: https://www.jstor.org/stable/43866011

Chen, A.Y., and Zimmermann, T. (2021) '*Open Source Cross-Sectional Asset Pricing'*. Critical Finance Review. Vol. 11: 207 – 264.
Available at: https://ssrn.com/abstract=3604626 or http://dx.doi.org/10.2139/ssrn.3604626

Christoffersen, P.F. (1998), '*Evaluating Interval Forecasts'*. International Economic Review. Vol. 39: 841 – 862.
Available at: https://www.jstor.org/stable/2527341

Cryer, J.D., and Chan, K. (2008) '*Time Series Analysis With Applications in R'*. Springer Texts in Statistics. Pag. 284.
Available at: https://mybiostats.files.wordpress.com/2015/03/time-series-analysis-with-applications-in-r-cryer-and-chan.pdf

Danielsson, J., and De Vries, C.G. (2000), *'Value-at-Risk and Extreme Returns'.* Annales d'Économie et de Statistique N°60: 239 – 70.
Available at: https://www.jstor.org/stable/20076262

Engle, R.F. (1982), '*Autoregressive Conditional Heteroskedasticity with Estimates of the Variance f United Kingdom Inflation*'. Econometrica Vol. 50: 987 – 1008.
Available at: http://www.econ.uiuc.edu/~econ536/Papers/engle82.pdf

Fama, E. F., and French K. R.. (2015), '*A five-factor asset pricing model*.' Journal of Financial Economics Vol. 116: 1–22.
Available at: https://www.sciencedirect.com/science/article/pii/S0304405X14002323

Fama, E.F., and French, K.R. (2020), '*Comparing Cross-Section and Time-Series Factor Models'*. The Review of Financial Studies. Vol. 33: 1891–1926,
Available at: https://academic.oup.com/rfs/article/33/5/1891/5555879

Fissler, T., and F. Ziegel, J.F. (2015), *'Higher order elicitability and Osband's principle'.*
The Annals of Statistics Vol. 44: 1680 - 1707.
Available at: https://arxiv.org/abs/1503.08123

Glosten, L.R., Jagannathan, R., Runkle, D.E. (1993), '*On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks'*. The Journal of Finance Vol. 48: 1779-1801.
Available at: https://faculty.washington.edu/ezivot/econ589/GJRJOF1993.pdf

Gneiting, T. (2011), *'Making and Evaluating Point Forecasts'.*
Journal of the American Statistical Association Vol. 106: 746 – 762.
Available at: https://www.tandfonline.com/doi/abs/10.1198/jasa.2011.r10138

Gneiting, T., and Raftery, A.E. (2007), *'Strictly Proper Scoring Rules, Prediction, and Estimation'.* Journal of American Statistical Association Vol. 102: *359 – 378.*
Available at: https://sites.stat.washington.edu/raftery/Research/PDF/Gneiting2007jasa.pdf

Haas, M. (2001), *'New methods in backtesting'*. Communications in Statistics Vol. 47: 3616 – 3639.

Available at: https://www.ime.usp.br/~rvicente/risco/haas.pdf

Holton, A. (2002), '*History of Value-at-Risk:1922-1998*'. Contingency analysis.
Available at:
http://stat.wharton.upenn.edu/~steele/Courses/434/434Context/RiskManagement/VaRHistlory.pdf

Jegadeesh, N., and Titman, S. (1993), '*Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency'*. The Journal of Finance Vol.48: 65 – 91.
Available at: https://www.jstor.org/stable/2328882

Hou, K., Xue, C., and Zhang, L. (2015), Digesting Anomalies: An Investment Approach, The Review of Financial Studies Vol. 28: 650 – 705.
Available at: https://academic.oup.com/rfs/article/28/3/650/1574802

Koenker, R., and Bassett, G. (1978), '*Regression quantiles'*. Econometrica
Vol. 46: 33 – 50.
Available at: https://www.jstor.org/stable/1913643

Koenker R., and Hallock, K.F. (2001), '*Quantile regression'*. Journal of Economic Perspectives Vol.15: 143 – 156.
Available at: http://www.econ.uiuc.edu/~roger/research/rq/QRJEP.pdf

Kou, S., Peng, X., and Heyde, C.C. (2013), '*External Risk Measures and Basel Accords*'. Mathematics of Operations Research Vol. 38: 393 – 417.
Available at: https://www.jstor.org/stable/24540860

Kupiec, P. (1995), *'Techniques for Verifying the Accuracy of Risk Measurement Models'*. The Journal of Derivatives Vol. 3: 73 – 84.
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6697

Linsmeier, T.J., and Pearson, N.D. (1996), '*Risk Measurement: An Introduction to Value at Risk'*.
Available at: https://exinfm.com/training/pdfiles/valueatrisk.pdf

Malkiel, B.G. (2019), '*A random walk down Wall Street: the time-tested strategy for successful investing'*. WW Norton & Company. Princeton University.
Available at: https://wwnorton.com/books/9780393358384

Manganelli, S., and Engle, R.F. (2004), '*CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles'*. Journal of Business & Economic Statistics. Journal of Business & Economic Statistics Vol. 22: 367 – 381.
Available at:http://www.simonemanganelli.org/Simone/Research_files/caviarPublished.pdf

Nagel, S. (2013), '*Empirical Cross-Sectional Asset Pricing'*. Annual Review of Financial Economics Vol. 5: 167 – 199.
Available at:
https://www.annualreviews.org/doi/abs/10.1146/annurev-financial-110112-121009

Novy-Marx, R. (2009), '*Is momentum really momentum?*'. Journal of Financial Economics Vol. 103: 429 – 453.
Available at: https://www.sciencedirect.com/science/article/pii/S0304405X11001152

Taylor, J.W. (2020), '*Forecast Combinations for Value at Risk and Expected Shortfall'*. International Journal of forecasting Vol.36: 428 – 441.
Available at: https://users.ox.ac.uk/~mast0315/CombiningVaRES_IJF.pdf

Uribe, J.M., and Guillén, M. (2020), '*Quantile Regression for Cross-Sectional and Time Series Data*'. SpringerLink.
Available at: https://link.springer.com/book/10.1007/978-3-030-44504-1

Vidal-Llana, X., and Guillén, M. (2022), '*Cross-sectional quantile regression for estimating conditional VaR of returns during periods of high volatility'*. The North American Journal of Economics and Finance Vol. 63: 101835.
Available at: https://www.sciencedirect.com/science/article/pii/S106294082200170X

Zhang, Y., and Nadarajah, S. (2016), '*A review of backtesting for value at risk'*. Communications in Statistics Vol. 47: 3616 – 3639.
Available at: https://pure.manchester.ac.uk/ws/portalfiles/portal/60673220/back4.pdf

8 Website references

CRSP database: https://www.crsp.org/

Global Industry Classification Standard (GICS):
https://en.wikipedia.org/wiki/Global_Industry_Classification_Standard

Investopedia, '*Book-to-market ratio*':
https://www.investopedia.com/terms/b/booktomarketratio.asp#:~:text=The%20book%2Dto%2Dmarket%20ratio%20compares%20a%20company's%20book%20value,the%20number%20of%20shares%20outstanding

Market capitalization S&P500 sectors:
https://eresearch.fidelity.com/eresearch/markets_sectors/sectors/sectors_in_market.jhtml

Mathworks, '*Risk Management Toolbox*':
https://www.mathworks.com/help/risk/var-backtest.html?s_tid=CRUX_lftnav

Simone Manganelli, Matlab codes CAViaR model:
http://www.simonemanganelli.org/Simone/Research.html

S&P500 sectors data: https://www.openassetpricing.com/