Lund University
Department of Economics
NEKN02 - Master's thesis
Spring semester 2024

The Impact of Implied Volatility on VaR Accuracy

Authors:
Sixten Cederholm
Wilfred Cederholm

Supervisor:
Anders Vilhelmsson

Abstract

The 2007-2008 financial crisis underscored the necessity for enhanced risk measurement tools in financial institutions. Traditional risk modeling approaches, relying on historical data to predict future risks, often fall short during periods of significant market volatility. This thesis explores the potential of incorporating option-implied volatility, specifically the VIX index, to improve the accuracy of Value-at-Risk forecasts. By integrating both non-parametric and parametric models, including Basic Historical Simulation, Volatility Weighted Historical Simulation, and parametric models based on Normal and Student's t-distributions, the study evaluates the effectiveness of these models using data spanning from 2000 to 2024, including periods of extreme volatility like the financial crisis and the COVID-19 pandemic. The results demonstrate that models incorporating VIX can provide as accurate VaR predictions as those of traditional methods using historical data. But the unadjusted VIX index is not an efficient input for parametric models. The thesis concludes with recommendations for further research to refine these models and enhance their practical application in financial risk management.

Key words: Risk management, Value-at Risk, Historical Simulation, EMWA, N-distribution, Student's t-distribution, Implied volatility.

Contents

1 Introduction

The financial crisis of 2007-2008, characterized by the collapse of major financial institutions, the bailout of banks by national governments, and a severe global economic downturn, spurred discussion regarding different approaches to enhance the precision of risk measurement tools. Still, until this very day, a critical aspect remains largely underinvestigated. The prevalent approach in risk modeling, Value at Risk, relies on historical data to forecast future risks and tends to produce backward looking estimates. Based on the assumption that past trends will continue, this method can lead to inaccuracies, especially during significant market volatility when the market's statistical characteristics may change drastically.

Precise Value-at-Risk predictions are not just important, they are the basis of effective bank risk management. They empower banks to foresee potential losses and maintain adequate capital reserves to remain resilient despite volatile markets. The risk measure is also a fundamental part of the Basel framework, enforced by the Basel Committee on Banking Supervision (BCBS). These advancements strive to reduce financial instability by fortifying the banks' ability to withstand economic shocks and maintain global financial stability, thereby enhancing the overall resilience of the financial sector (BCBS, 2009).

Forecasting VaR poses a unique challenge with current methods relying on past return data. When future return distributions deviate significantly from historical ones, the reliance on past data alone can lead to misleading VaR forecasts. This challenge mirrors the difficulties in predicting future stock volatility from historical volatilities. Research in stock volatility forecasting demonstrates that including forward looking information, such as option-implied volatility, can lead to more accurate predictions than methods solely based on historical data (Latane & Rendleman, 1976) and (Beckers, 1981).

Building on this approach, we will explore the potential of using option implied volatility to enhance VaR forecasts. It is important to note that using option implied information comes with limitations, such as the necessity of a liquid option market (Grover & Thomas, 2011) as well as the need for adjustment as there exists a discrepancy between the option implied volatility and the realized volatility (Bekaert & Hoerova, 2014). Christoffersen et al. (2013) conclude from their survey on forecasting with option implied information, that even though a bias exists, it does not prevent the option implied information from being a useful predictor of the future object of interest. This paper examines if VaR prediction accuracy can be improved by leveraging unadjusted future volatility predictions based on the

VIX index, a widely used measure of market volatility, derived from SPX options. The VIX index is often referred to as the 'fear gauge' as it reflects investors' expectations of market volatility over the next 30 days.

Existing research on VaR forecasting is extensive, with many studies using historical data and a variety of different models and complexity. However the existing research using forward looking data in the context of Value at Risk is limited. Nossman and Vilhelmsson (2012) incorporate the VIX index in a volatility weighted historical simulation framework, from here on called HS-VIX, between 1990 and 2010, and find that the performance of their "new" model is at least as good as historical simulation using Generalized Autoregressive Conditional Hetroskedasticity (GARCH). Giot (2003) study the precision of VaR forecast between 1994 and 2003 when the VIX index is utilized as an input to the Skewed Student's t-distribution, and discover that it yields accurate predictions. This paper aspire to expand the research in this area by incorporating the VIX index into the Normal distribution and the Student's t-distribution, whom seems to have been overlooked, perhaps due to their simplistic assumptions. The aim is to determine whether a straightforward approach can produce reliable results. For comparison, we include the HS-VIX model proposed by Nossman and Vilhelmsson (2012), as well as four additional models that rely on historical data. This comprehensive approach will help assess the efficacy of using the VIX index in various VaR forecasting frameworks.

The non-parametric models that will be used are: Basic Historical Simulations (BHS) and Volatility Weighted Historical Simulation (VWHS), with the variance derived from the Exponentially Weighted Moving Average (EWMA). The parametric models that will be used are as mentioned the Normal distribution and the Student's t-distribution. All models will be validated through the Christoffersen (1998) tests, which will assess the accuracy of VaR violations predicted as well as their independence. Additionally, this paper includes data from the financial crisis and the COVID-19 pandemic, highlighting the performance of these models during extreme events. This is crucial, as VaR's primary purpose as a risk metric is to prepare and protect against such economic shocks.

The choice of limiting this thesis by using these relatively accessible models is due to their simplicity in terms of implementation and interpretation of the results. More complex models as well as adjusting for the variance risk premium could potentially yield more accurate results but because of time constraints and the active choice of keeping the thesis uncomplicated we focus on more manageable methods. Despite these limitations we are convinced that our methods will yield insightful results regarding forecasting value at risk.

The thesis is further structured as follows: Section 2 provides an overview of the methods used in this paper. The following section 3 consider the existing research, while section 4 explains the data set used and the applications of the models. In Section 5 the results are outlined straightforwardly, followed by the analysis/discussion Section 6 where different angles of the result is considered. Section 7 summarizes the thesis and provides the main conclusions followed by further research proposals. The thesis ends with the refrences used in Section 8 and an appendix where the majority of graphs and tables used in this paper are presented.

2 Theoretical Framework

2.1 Value-At-Risk

Value at Risk (VaR) is a statistical metric that assesses the financial risk within a firm or investment portfolio over a defined period (often 1, 10, or 22 days). It is referred to as the maximum loss that can be expected with a probability of $1 - \alpha$ over the specified period (Hull, 2023). Because of its simplicity of use and understanding, VaR is a popular risk measure that can be used on any investment portfolio. The risk measure focuses on the main driving risk factors within the portfolio and the interaction between the assets, thus being holistic (Dowd, 2005). By providing the user with a probability of a certain loss and an actual amount of money expected to be lost, the model is very comprehensible, and results are easy to compare. Formally the Value at Risk is defined as:

$$\text{VaR}_\alpha = \min\{\ell : \Pr(L > \ell) \leq 1 - \alpha\} \tag{1}$$

VaR originates from the loss distribution and can be seen as threshold where $1 - \alpha$ of the losses is further in the right tail. While VaR provides an estimate of the maximum potential loss at a given confidence level, it remains silent about the distribution of losses beyond this point, offering no insight into the potential for extreme losses in the right tail (Hull, 2023). For continuous loss distributions, there exists a VaR, so the probability of larger losses than VaR is precisely $1 - \alpha$. Because of this, every loss can be thought of as a VaR for a certain confidence interval. Further, VaR can also be seen as the α-th quantile of the loss distribution.

2.2 Non-Parametric Approaches

Value at Risk can be divided into non-parametric and parametric approaches. The non-parametric approaches utilize solely historical data and its empirical distribution. Therefore, no theoretical assumptions are made regarding the loss distribution, thus avoiding the possibility of distribution restrictions. However, underlying all non-parametric approaches using historical data is the implicit assumption that it can be utilized to forecast future risk. which can be a limitation during periods of changing market conditions or when past volatility significantly differs from future expectations (Dowd, 2005). The non-parametric models that will be used in the research will be Basic Historical Simulation and Volatility Weighted Simulation.

2.2.1 Basic Historical Simulation

Basic Historical Simulation is the most straightforward methods for estimating VaR. It does so by arranging historical losses in ascending order and identifying the threshold loss at a given confidence level. Specifically, given the confidence level α VaR is defined as the $N(1-\alpha)+1$-th largest loss, where N is the sample size (Dowd, 2005). If $N(1-\alpha)+1$ results in a non-integer interpolation can be conducted to estimate VaR and calculate a weighted average between the losses at the two closest integer ranks.

$$\Pr\left(L > \ell^s_{(1-\alpha)N+1}\right) = \frac{(1-\alpha)N}{N} = 1-\alpha \implies \mathrm{VaR}_\alpha = \ell^s_{(1-\alpha)N+1} \tag{2}$$

A rolling window approach is often employed to estimate future VaR by continuously updating the dataset with the most recent losses, replacing the oldest data points. This method assigns equal weights to all losses within the selected window, which, while being straightforward, can introduce inaccuracies. Specifically, when large and potentially outdated losses from past extreme events are given the same weight as more recent observations, potentially skewing the VaR estimates.

The choice of window size is crucial in BHS. A shorter window can lead to more accurate VaR estimates as it focuses on the latest observations and rapidly adapts to new market conditions. Nonetheless, this involves a trade-off, leading to increased volatility in the VaR estimates, as each loss has a higher impact and the dataset changes more rapidly. Conversely, a larger window stabilizes VaR estimates by incorporating a broader dataset, thereby diluting the effects of single-day fluctuations and providing more consistent and predictable risk metrics. However, the downside is that older data, which may be less relevant to current market conditions, can dilute the impact of recent changes, potentially leading to underestimated VaR estimates and reduced sensitivity to market shifts (Stambaugh, 1996).

Balancing the window size is therefore essential, as a shorter window enhances responsiveness to recent market conditions but increases estimate volatility, while a longer window smooths out fluctuations but may be inferior in reflecting current risk dynamics.

2.2.2 Volatility Weighted Historical Simulation

In volatile market conditions clustering behaviour is not uncommon. In these conditions when the returns are not independently and indentically distributed (IID) scaling the estimation of the risk with more weight of more recent observations can be more efficient. In order to accommodate changing market conditions, where clustering behaviour is common and returns are not IID the extension of volatility weighting can be applied to the historical simulation procedure (Pritsker, 2006). In essence the approach adjusts historical losses to reflect changes in future volatility relative to its past values. Specifically, the Volatility Weighted Historical Simulation model compares the most recent volatility forecast to historical volatility forecasts and scales the historical losses accordingly. This adjustment provides a more accurate VaR estimate that is aligned with the most recent market conditions. In the model, the ℓ_t is the historical loss of the portfolio at time t. σ_t is the forecasted volatility of the loss at time t which is estimated at the end of day $t-1$. σ_{T+1} is the most recent forecast of the volatility of the portfolio usually estimated either with a GARCH-model or EWMA-model (Hull & White, 1998a) As can be seen in the below the adjusted return will increase or decrease in relation to the actual return. This depends on whether the volatility forecast for $T+1$ is higher or lower than the historical volatility forecasts for day t. In other words, if the future volatility is forecasted to be higher than what has been historically forecasted this will be accounted for leading to higher losses and a higher estimated VaR. The formula for the adjusted loss ℓ_t^* is as follows:

$$\ell_t^* = \left(\frac{\sigma_{T+1}}{\sigma_t}\right)\ell_t$$

After the losses have been adjusted the VaR is estimated according to the BHS procedure as outlined in section 2.2.1. The VWHS-model has shown to be superior to the BHS-model providing more accurate VaR estimates (Hull & White, 1998a)

2.2.3 Exponentially Weighted Moving Average

The Exponentially Weighted Moving Average model prioritizes more recent data points, thereby diminishing the impact of outdated information and older extreme events when forecasting volatility. The model utilizes σ_{t-1}^2 the volatility estimate from the previous day, and x_{t-1}^2 which represents the return from the previous day. The decay factor, λ dictates the extent to which past values influence current volatility estimates. A high λ causes the weights to decline more slowly, hence past values exert a stronger influence compared to a lower λ (Dowd, 2005). RiskMetrics, which originated the EWMA model, recommends an industry standard decay factor of 0.94 as established by J.P. Morgan in 1996, and this standard will be utilized in this study. The EWMA model has two main advantages. Firstly, it assigns greater weight to the most recent observations, allowing the model to react more quickly to recent market shocks compared to an equally weighted model. Secondly, by gradually reducing the weight of older data through its decay factor, the EWMA model minimizes the impact of past shocks that may no longer be relevant. This approach prevents abrupt changes in the estimated VaR when past extreme events drop out of the sample as these events already has been "scaled down" weight wise, ensuring a smoother and more accurate risk assessment. (J.P. Morgan, 1996).

$$\sigma_t^2 = (1-\lambda)x_{t-1}^2 + \lambda\sigma_{t-1}^2 \tag{3}$$

2.3 Parametric Approaches

Unlike non-parametric approaches, parametric models require assumptions about the probability distribution and potential losses within that distribution. This makes them potentially more powerful than non-parametric models because they incorporate more detailed information. Parametric methods are often favored for their convenience, as they provide straightforward value at risk formulas. However, the effectiveness of a parametric model heavily depends on how accurately the probability density function is estimated in relation to the data. To obtain reliable parameter estimates, sufficient data must be utilized and accurately modeled to reflect the true distribution of the data. The parameters are estimated with Maximum likelihood estimation (MLE). In essence the MLE method estimates the values of the parameters by maxing the likelihood of the observations occurring given the certain statistical model (Hull, 2023). The parametric approaches considered in this research paper are the Normal distribution and Student t-distribution.

2.3.1 Normal Distribution

When assuming that the losses are independently normally distributed, the Value at Risk calculation can be straightforwardly implemented using the basic statistical properties of the normal distribution. The VaR is computed based on the mean $\mu_{P/L}$ and the standard deviation of the loss distribution $\sigma_{P/L}$ along with a quantile of the normal distribution z_α, which depends on the chosen confidence level (Dowd 2005).

$$\mathrm{VaR}_\alpha = -\mu_{P/L} + \sigma_{P/L} z_\alpha \tag{4}$$

However, its reliability hinges on the assumption of normality in the distribution of returns. According to Hull and White (1998b) the normality assumption may not hold in practical scenarios, especially for financial returns which often display excess kurtosis (fat tails), indicating that extreme events occur more frequently than predicted by a normal distribution. This misalignment with empirical reality can lead to significant underestimation of VaR, potentially resulting in inadequate risk management strategies. Therefore, while the normal distribution model's simplicity makes it attractive for calculating VaR, it is crucial for risk management to consider more robust models or to employ additional risk management tools that can better account for the fat-tailed nature of financial data enhancing VaR models to include non-normal distributions or using complementary risk assessment methods.

2.3.2 Student's t-Distribution

In cases of extreme events, applying a t-distribution can effectively address excess kurtosis, which is common in financial return data. The t-distribution is similar to the Normal distribution, in terms of using the mean and standard deviation as input variables. However, the t-distribution differs significantly due to its fatter tails, which result in higher quantile values and, consequently, a higher VaR. This adjustment is reasonable as it accurately reflects the increased risk of extreme losses. The degrees of freedom, denoted as ν, in the t-distribution, should be chosen to reflect the level of kurtosis of the data: a lower number of degrees of freedom is used for higher kurtosis to capture the heavier tails, and vice versa. As ν increases, the t-distribution converges towards the normal distribution, illustrating its flexibility in modeling different levels of tail risk (Dowd, 2005). T-distribution can be seen as a generalization of normal distribution. This makes the t-distribution a versatile tool in risk management, particularly valuable when dealing with datasets prone to outliers or heavy tails.

$$\mathrm{VaR}_\alpha = -\mu_{P/L} + \sqrt{\frac{\nu - 2}{\nu}}\sigma_{P/L} t_{\alpha,\nu} \tag{5}$$

2.4 Backtesting

The effectiveness of a VaR model rely of its capacity to provide reasonably accurate estimates. Backtesting is one way of validating the adequacy of the VaR predictions by examining whether the number of losses exceeding the predicted VaR level, ie., violations, are in line with what is expected (Novales & Garcia Jorcano, 2019). Given a sample of n observed losses and a confidence level of α percent an accurate model is expected to have $n * (1 - \alpha)$ VaR violations. As banks partially determine their capital requirements based on VaR predictions the accuracy of the forecasts is vital. A backtesting result yielding significantly fewer or more than expected results in a rejected VaR model. If the number of observed exceptions exceeds the predicted amount, the model underestimates risk, leading to insufficient capital and reserve allocations. Conversely, if there are fewer exceptions than predicted, the model overestimates risk, thus inefficiently allocating excessive funds to reserves (Jorion, 2007). Backtesting is not only a method for validation but additionally provides feedback for model refinement. Continuous monitoring and backtesting allow financial institutions to adjust their models and improve their accuracy.

2.4.1 Christoffersen test

The Christoffersen backtesting framework (1998) consist of two parts. A test for unconditional coverage examining if the model yields the correct number of VaR violations. And a test for independence of violations, determining if the violations are clustered and thus if the model can account for changing market conditions. The tests can be conducted separately or combined in a joint test, allowing to test for coverage and independence simultaneously.

The Christoffersen test of unconditional coverage, investigate if the failure rate of the data is consistent with the predicted probability of exceedances $p = (1 - \alpha)$ of the VaR model. Where the failure rate is x/n given a sample size of n and x number of days where VaR is violated. In the form of a likelihood ratio, the Christoffersen test formally compare if the number of empirical exceptions is approximately p. The statistic's values are elevated and the null hypothesis rejected when the number of violations is either significantly low or extremely high. The likelihood test follows a chi-square distribution with one degree of freedom, with critical values of 3,43 (0,05) and 6,64 (0,01) is defined as follows: (Jorion, 2007):

$$LR_{uc} = -2\ln\left[(1-p)^{n-x}p^{x}\right] + 2\ln\left[\left(1-\frac{x}{n}\right)^{n-x}\left(\frac{x}{n}\right)^{x}\right] \sim \chi^2(1) \tag{6}$$

The subsequent Christoffersen test of independence, examines whether the violations occur independently. In the test n_{ij} represents the number of days that a violation, j, occurred following a non-violation i on the previous day. Accordingly, π_{ij} is the likelihood of a violation j occurring given that a non-violation i was observed the preceding day. The independence test displayed below also follows an chi-square distribution with one degree of freedom, with critical values of 3,43 (0,05) and 6,64 (0,01) (Jorion, 2007):

$$LR_{ind} = -2\ln\left[(1-\pi_2)^{n_{00}+n_{11}}\pi_2^{n_{01}+n_{10}}\right] + 2\ln\left[(1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}}\right] \sim \chi^2(1) \tag{7}$$

$$\text{Given by: } \pi_{01} = \frac{n_{01}}{n_{00}+n_{01}}, \quad \pi_{11} = \frac{n_{11}}{n_{10}+n_{11}}, \quad \pi_2 = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}} \tag{8}$$

As earlier mentioned combining test (1) and (2) creates a joint test. Consequently, this test follows a chi-square distribution with two degrees of freedom, with critical values of 5,99 (0,05) and 9,21 (0,01).

$$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2) \tag{9}$$

2.5 Implied Volatility

Within an option pricing model like the Black-Scholes model from 1973, the volatility anticipated over the option's duration is reflected in the option's price. Given the same formula, the volatility is the only parameter that cannot be observed. When the prices for a call or put option are accessible, it's possible to back out the volatility variable in the Black-Scholes pricing equation. Allowing us to derive the expected volatility over the option's term from the market prices observed for a call or put option, ie., the implied volatility (Hull, 2012)(Christoffersen et al. 2013).

Implied volatility, derived from option prices observed in the market, is considered as an appropriate estimate of the market's notion of the underlying assets future volatility, conditional on the remaining life span of the option contract. Additionally, because option traders incorporate both past and future information when pricing options, the option market encompasses a wider set of data in relation to the underlying asset's market. Assuming the option market operates efficiently, the implied volatility should serve as an unbiased and applicable forecast of the future realized volatility of the underlying asset. The implied volatility measure should encompass all available information, (Le, 2020) and several studies have shown that it outperforms historical volatility as a predictor of future volatility (e.g., (Li & Yang, 2009; (Christensen & Prabhala, 1998), (Wang & Wang, 2016)).

More specifically, Latane and Rendleman (1976) and Beckers (1981) conclude that future volatility forecasts based on implied volatility are superior to those based on historical volatility for individual stocks. Mayhew and Stivers (2003) and Brous et al. (2010) further find this dominance prominent for larger firms and for firms with higher option trading volume. Suggesting that a higher option volume incurs more efficient option pricing and, thus, more accurate implied volatility predictions. The estimation of the implied volatility can be noisy if the liquidity is low in the option market, the liquidity is best measured as the difference in bid-ask spread rather than the trade volume, thus, the higher the liquidity, the more information the implied volatility contains (Grover & Thomas, 2012).

Apart for the requirement of sufficient liquidity, there is also an discrepancy between the implied volatility, as it is obtained under the risk-neutral measure (Q) and the forecasted realized volatility observed under the physical measure (P), because the risk-neutral measure also includes the market price of risk, i.e., the volatility premium. Conduction direct comparisons between these two measures implies a assumption of a volatility risk premium equating to zero. Empirically, this assumption does not seem to hold, as Prokopczuk and Wese Simen (2014) finds that adjusting the implied volatility for the volatility risk premium significantly outperform its unadjusted counterpart. It is widely known that the volatility risk premium is a function of the level of the implied volatility Bams et al. (2017) argues that a bias between the two measures is anticipated, but it may be minimal. Additionally, the adjustment intended to correct this bias might introduce more problems than it resolves. This bias is a decreasing function of the time horizon in consideration, as the ratio between the risk-neutral measure and physical measure converge to 1 as the maturity of the option decreases (Molino & Sala, 2020)

The empirical research regarding implied volatility is quite extensive, overall, the studies indicate that predicting future volatility based on implied volatility is superior to historical data (Granger & Poon, 2005), whether the implied volatility has been adjusted or not. In adherence to this, there is also a significant number of papers use index option implied volatility, such as the VIX index to accurately predict market (stock index) volatility.

2.6 VIX Index

In 1993, the Chicago Board Options Exchange introduced the first version of the VIX index, known as the fear gauge, which measures the market´s expectation of 30-day volatility. Initially based on Black-Scholes implied volatility of at-the-money options on the S&P 100 index, the VIX was revised in 2003 and 2014 to be model-free and no longer depend on an option pricing framework. Instead, it is based on a combination of price quotations of S&P 500 index out-of-the-money puts and calls centered around an at-the-money strike price.(Cboe, 2022). Also SPX options with weekly expiration is hereby included, allowing the VIX index to be calculated from the option series to more precisely match the 30-day target time frame for expected volatility that the VIX is intended to represent. (Moran and Liu, 2020)

The VIX index inherently has limitations and is known to exhibit a volatility premium bias, thus systematically overestimating the future realized volatility. The VIX is seen as the risk-neutral expected stock market variance for the S&P500 contract and as earlier mentioned derived from option prices. The index contains both the stock market uncertainty (which is the realized volatility) as well as a variance risk premium (Bekaert & Hoerova, 2014). The premium is likely to arise because market participants generally exhibit risk aversion, particularly under conditions of uncertainty and economic instability. As a result they are willing to pay more for options as a protective measure against anticipated swings in the market, pushing the implied volatility indicated by the VIX above the future realized volatility. Furthermore, the VIX is commonly understood to have an asymmetrical relationship with current market volatility. Typically, it experiences a more pronounced increase in declining markets than its relatively milder decrease in rising markets. This asymmetry is likely due to the increased hedging activities during market downturns, and thus the VIX can be seen as an indicator reflecting the price of portfolio insurance (Bekaert et el., 2020)(Bekeart & Hoerova, 2014) Despite this, the VIX index is widely used as a benchmark for expected stock market volatility (Sharma et al., 2019) and is considered being the market's best estimation of future volatility. (Nossman & Wilhelmsson, 2012)

Most studies utilizing index option implied volatility still though disclose its superior forecasting ability compared to forecasts based on historical data. (Le, 2020). Christensen and Hansen (2002) find that implied volatility subsumes vital information about future volatility movements and that the VIX index is an efficient forecast of index return volatility. Banerjee et al. (2007) and Gospodinov et al. (2006) find similar results, underpinning its adequate predictive ability.

3 Existing litterature

A few studies have documented the efficiency of different volatility forecasts within a Value at risk context. Rather than directly comparing these forecasts to observed volatility, they are employed to estimate potential maximum losses on assets or portfolios given a certain probability. Importantly, the primary concern in these empirical investigations is not the precise prediction of volatility levels but rather the accuracy in capturing the distribution's tails, representing extreme value occurrences.

Giot (2003) analyzes that information content of VIX and VXN implied volatility indexes in a VaR framework. They divide their sample ranging from 1994 to 2003 into three sub-periods that differs in respect to volatility characteristics (bull/bear markets and high/low volatility) thus, allowing assessment of their models performance in different market conditions. Using a parametric approach, they start with the skewed Student's t-distribution incorporating three different volatility methods; VIX, EWMA and GARCH. To evaluate the VaR estimations they employ the Christoffersen (1998) tests, including testing for correct coverage, independence and conditional coverage. The test result shows that the VaR models generate the correct number of violations and the independence and joint test are generally not rejected. Furthermore, they find that the VaR models performs equally well during all sub-periods. Thus, they conclude that volatility forecasts based on VIX is an effective input for value at risk measurement as it provides accurate forecasts even under difficult market conditions marked characterized by declining prices and increased volatility.

Chong (2004) compares daily exchange rate VaR forecasts between 1993 and 1999 derived from GARCH-models, simple historical methods and exponentially weighted moving average methods with a model using implied volatility derived from OTC-options. Documenting implied volatility, derived from currency options, as an ineffective input for estimating VaR, as it tends to overestimate risk during stable market conditions and conversely underestimate risk in more volatile periods.

Nossman and Vilhelmsson (2012) look at the period between 1990 and 2010 for the S&P 500 and compare traditional VaR-forecast methods such as standard historical simulation and GARCH historical simulation against an option implied volatility using the VIX-index(historical simulation VIX-model). They expand their study further by applying several forecast horizons (1, 10 and 22 days). Around the date of bankruptcy of the Lehman Brothers the paper shows that the VIX-model adjust its forecasts more efficiently than the standard HS-model.

If a bank would have used the VIX-model to determine their market induced capital requirements they would have been shielded by 36.7% more capital compared to the HS-GARCH model and 40.4% compared to the HS-model. When backtesting the models with Christoffersen joint tests the VIX-model is rejected less often compared to the others potentially making it superior. As a result, they conclude further that if financial institutions had based their capital requirements on the estimates of the VIX model they would generally have lower capital requirements but an approximately 40 procent higher capital requirement just before the financial crisis.

Schindelhauer and Zhou (2018) utilize the maximum entropy method to estimate the risk-neutral probability distribution of future returns given the market prices of options for the S&P 500 during the period 1996 to 2016. They incorporate implied moments (mean, variance, skew, and kurtosis) and quantile into well-established VaR prediction models like covariate and GARCH and find that including only implied volatility significantly improves the performance of the models compared to only using historical returns as input. Furthermore, the results are compared with and concluded equally as accurate as those of implied volatility measures like VIX and Black-Scholes.

Molino and Sala (2020) compare the performance of VaR forecasts using four econometric approaches based on stock returns and four using option market data. The data used in the research is based on the European options on the S&P 500 from 2000 to 2019. They conclude that, independently of the method used, inferring option implied volatility is superior or at least as good as those based solely on stock returns. The results demonstrate robustness and intriguingly suggest that the primary factor influencing the outcomes is the incorporation of option market data into the estimation rather than the econometric method employed.

In summary, several studies where the incorporation of implied volatility has been conducted. Most use the VIX index and the S&P 500, they are usually tested against a Christoffersen (1998) framework. The sample periods spans from 5 to 20 years, and commonly include periods of high volatility. No matter how the VIX index is incorporated in the study, whether it is in parametric or historical simulation and if it is adjusted or not, it seems to be producing accurate better or at least as good VaR predictions as traditional models. Both in terms of correct coverage or independence of violations.

4 Data and Methodology

This paper utilizes a comprehensive data set comprising daily observations of the VIX and S&P 500 indices collected from Bloomberg between 2000 and 2024, resulting in a total of 6219 daily observations. The extensive nature of the data set enhances the statistical robustness of our analysis, minimizing the influence of anomalies and outliers and ensuring enough data points for reliable and consistent estimates. The selected sample period covers both stable and turbulent market conditions, making it ideal for evaluating the effectiveness of various Value-at-Risk methodologies across different market dynamics. This is particularly pertinent to our study's aim, as forecasting VaR in tranquil conditions is relatively straightforward. At the same time, what really puts the model to the test is an extreme change in volatility. Including implied volatility in our analysis challenges the models to accurately predict VaR in the face of significant market volatility shifts.

The loss distribution is simply the inverse distribution regarding the gains, where the largest losses can be found in the right tail (Dowd, 2005). In this paper the loss is defined as the negative geometric return of the underlying asset or investment, as the equation below displays it equals the logarithm of the ratio of asset's value at the end of time t and the value at the end of time t-1.

$$\ell_t = -R_t = -\ln\left(\frac{P_t}{P_{t-1}}\right) \tag{10}$$

In accordance with the methodology of (Nossman & Wilhelmsson, 2012) we will analyze three forecasting horizons, namely 1 day, 10 days, and 22 days. To obtain out-of-sample results, the dataset is divided into an estimation sample and a forecast sample. The size of the estimation sample varies with the length of the forecast horizon, as they argue that longer horizons require more observations to construct the empirical return distribution accurately. Specifically, the estimation sample includes the first 500 daily observations for the 1-day forecast, 1,000 daily observations for the 10-day forecast, and 2,200 daily observations for the 22-day forecast. For instance, a 22-day VaR forecast uses a single-step-ahead forecast from the empirical distribution of 22-day returns. All forecasts employ a rolling window, allowing the models to adapt to evolving market conditions and ensure accurate parametric estimation. The rolling window maintains a constant size for the estimation sample while advancing the start and end points by an increment equal to the forecast horizon after each

prediction. This strategy is vital for capturing the nuanced behavior of financial markets and refining the forecast accuracy of our VaR models.
As mentioned earlier, the VIX index is, by definition, a forecast of the volatility in the next 30 days expressed annually. A typical application for time scaling of volatilities is the square root of time rule. To accommodate this the square-root-of-time rule, which is commonly used in financial risk aggregation, where risk estimates are scaled to a lower (higher) frequency T by multiplying (dividing) by the square root of T. The models utilize the square root of the time rule to transform the VIX index into being expressed in equal terms/days as the VaR forecasting horizon. Thus, for the one day VaR forecast the VIX employed represent the markets estimation of the volatility for the next 30 days expressed in daily terms. Notably this is an approximation since the rule requires returns to be independent and identically distributed.

$$\mathrm{VIX}_{t,\delta} = \frac{\mathrm{VIX}_t}{\sqrt{252}} \cdot \sqrt{\delta} \tag{11}$$

Where δ is the length of forecast horizon

This thesis will use two non-parametric and two parametric econometric methodologies to estimate value at risk: Basic Historical Simulation and Volatility Weighted Historical Simulation incorporating EMWA, respectively normal distributed and Student's t-distributed VaR relying exclusively on historical stock return data. The BHS method is widely used and considered an industry standard due to its simplistic implementation further the VWHS is slightly more complex model frequently adjust for volatility changes but still being completely non-parametric. As the non-parametric does not make any assumptions about the loss distribution making them flexible and adaptable to changing market conditions. As for the parametric models the t-distribution is employed to due its ability to account for fat tail, which is common for financial data. Further the N-distribution is a comparable baseline model. The aim is to determine whether incorporating forward-looking information from option data; specifically, option implied volatility, enhances VaR estimates and reduces the incidence of VaR violations when backtesting. Additionally it will be interesting to see which models, non-parametric or parametric, that produce the most precise VaR estimates during a period of several extreme events. We will compare the performances of these historically conducted methods with three VaR methods where the VIX index is incorporated. The first method, as proposed by Nossman and Wilhelmsson (2012), is the historical simulation VIX model:

$$r^*_{t,t-\delta} = r_{t,t-\delta}\frac{\text{VIX}_T}{\text{VIX}_t} \tag{12}$$

Where the nominator VIX_T is the closing value of the VIX index which is the last observation in every rolling sample and the denominator VIX_t is the closing value of the index of day t. This model functions similarly to the Volatility Weighted Historical Simulation described in section 2.2.2, with a key modification: the scaler in equation (12) employs the VIX index instead of estimates of conditional volatility. This adjustment allows the scaler to modulate historical losses based on the market's anticipation of future volatility. Consequently, if the market expects increased volatility in the near future, the VaR adjusts accordingly, preemptively accounting for potential market fluctuations. Nossman & Wilhelmsson (2012) further argue that as long as the variance premium is proportional to the VIX level and thus can be expressed as VRP * VIX the model is unbiased toward it as the premium will be canceled out. If the same assumption apply for the time horizon of the VIX index, the model is also unbiased in the time aspect. For the parametric models; the N-distributed and the t-distributed, in this paper called N-dist(VIX) and t-dist(VIX), we similarly to the regular parametric models estimate the parameters mentioned in section 2.3 with the maximum likelihood method for every rolling sample but we replace the estimated standard deviation with the VIX_t, yielding the following methods:

$$\text{VaR}^n_\alpha = -\mu_{P/L} + \text{VIX}_T z_\alpha \tag{13}$$

$$\text{VaR}^t_\alpha = -\mu_{P/L} + \sqrt{\frac{\nu-2}{\nu}}\text{VIX}_T t_{\alpha,\nu} \tag{14}$$

Backtesting will be conducted using the Christoffersen likelihood ratios to evaluate these methods' performance, encompassing the unconditional, conditional, and joint test variants. Summing up the different VaR horizons and confidence intervals, 15 joint tests and 30 separate tests will be employed for each method, allowing for analysis of the correct number of violations and independence of violations collectively and separately. Additionally, we will take a closer look at the different models' performance in two periods of extreme market volatility: the financial crisis and the Covid-19 crisis, allowing for further nuancing.

5 Result

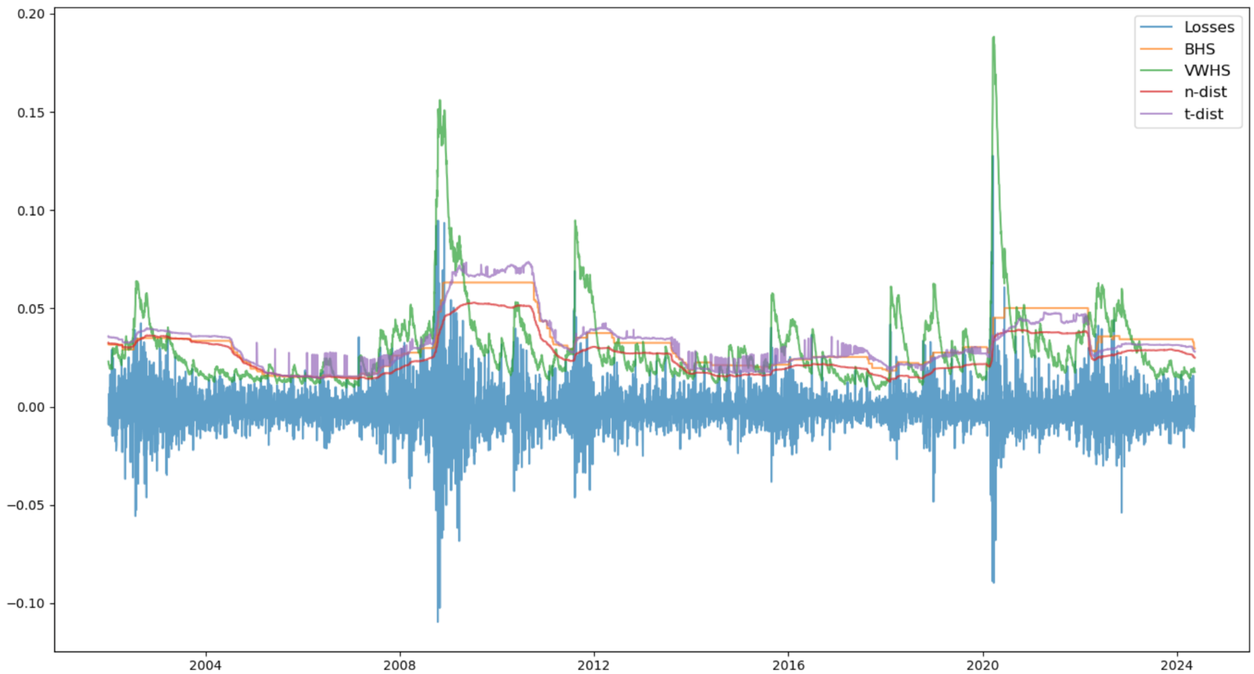

Figure 1: Comparison of Losses (1day99), VaR - BHS, VaR - ewma, VaR - n, VaR - t

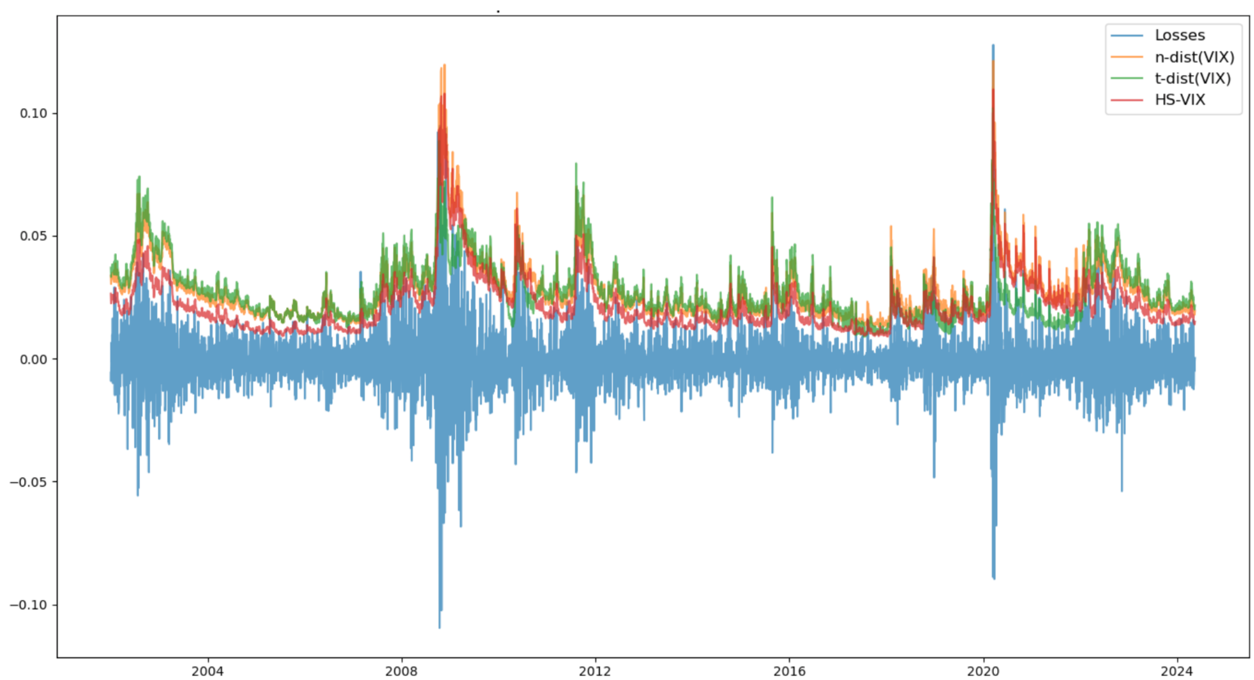

Figure 2: Comparison of Losses(1day99), VaR - tVIX, VaR - nVIX, VaR - VIX,

Observing Figure 1 and figure 2, we can visually conclude that the models incorporating historical data, with the exemption of the VWHS model, are generally slow to respond to the changing dynamics of the loss distribution, and tend to have many VaR exceedances at the start of periods of high market volatility. Contrarily, the VWHS-model and in particular the models incorporating the VIX index are more volatile and seem to be faster and more accurate in capturing changing dynamics of the loss distribution. Begging the question if incorporating some smoothing mechanism or constructing a moving average of the VIX index would produce more stabile VaR forecasts.

In general, the N-dist and t-dist models produce more VaR violations compared to other models. However, the N-dist and t-dist models using the VIX index produce relatively fewer VaR exceedances than those using historical data. When testing for the correct number of violation predictions, both the historical and VIX simulations of the N-dist and t-dist models show lower p-values (< 0.01 and 0.05) for most cases, indicating over- or underestimation of VaR forecasts. This is confirmed by the higher-than-expected VaR violations for historical N-dist and t-dist models and lower-than-expected violations for the VIX-based models.

The VWHS (using the EWMA approach) and HS-VIX models show the most accurate VaR predictions, with p-values above 0.01 and 0.05 in all cases, indicating that the null hypothesis of correct VaR predictions cannot be rejected. This accuracy is reflected in their VaR violations closely matching the respective VaR levels. Regarding the independence of violations, both the N-dist (VIX) and t-dist (VIX) models perform better, showing high p-values for most cases. Conversely, the historical N-dist and t-dist models often show lower p-values (< 0.01), resulting in rejection of the null hypothesis of independent VaR violations. The VWHS and HS-VIX models again show high p-values across different VaR levels and forecast horizons, supporting the null hypothesis of independent violations.

It is evident that the volatility-weighted historical simulation method (VWHS) stands out as the superior approach. The VWHS method cannot be rejected a single time by the joint Christoffersen test, making it statistically robust. This can be attributed to its simplistic yet effective approach. When looking at the results for the models incorporating the VIX index, the HS-VIX model is particularly noteworthy. It is only rejected in 1 out of 15 joint tests at a significance level of 1 percent and cannot be rejected for any confidence level or forecast horizon at a significance level of 5 percent, implying that this model accurately predict the risk associated with the S&P 500.

Conversely, the N-dist model emerges as the poorest performer, failing all 15 joint tests. This under performance is due to the model's systematic overestimation of loss magnitudes, leading to excessively high VaR estimates. While this conservative approach minimizes the likelihood of breaches, ensuring sufficient reserves to cover losses, it results in inefficient capital allocation. The bank holds more reserves than necessary, which could otherwise be deployed more productively.

The t-dist model, which integrates the VIX index, initially seemed promising due to its ability to account for fat tails, a common characteristic in empirical return distributions as well as our particular sample. However, the joint tests still reject the t-dist model 9 out of 15 times, highlighting its practical shortcomings.

The percentage of VaR violations in Table 1 further supports these findings. The parametric VaR models incorporating VIX, such as the N-dist (VIX) and t-dist (VIX), produce VaR forecasts that are generally too high thus, yielding too few violations. The N-dist (VIX) consistently falls short of the significance level, and the t-dist (VIX) does so for low significance levels and across all significance levels for longer forecast horizons.

One probable reason for the poor performance of parametric models incorporating the VIX index is the inherent variance premium within the VIX. This has been demonstrated in studies by Bekaert and Hoerova (2014) and Zhou (2018), showing that the variance premium inflates the VaR predictions. This theory is supported by the strong performance of the HS-VIX model, which uses a scalar of the VIX index rather than its absolute value. Given that the variance premium is proportional to the VIX level, it is effectively canceled out in the HS-VIX model, not affecting its performance. Future research should investigate an adequate way of scaling the VIX or integrating it into a GARCH framework with the t-distribution.

Table 1: VaR Forecast performance

Model	Forecast horizon 1 day					Forecast horizon 10 days					Forecast horizon 22 days				
	VaR Level														
	1%	2%	3%	4%	5%	1%	2%	3%	4%	5%	1%	2%	3%	4%	5%
	VaR Violations														
HS	1.57%	2.80%	3.59%	4.35%	5.24%	1.49%	3.11%	3.61%	4.35%	5.22%	3.01%	3.34%	4.68%	5.35%	5.69%
VWHS	1.19%	2.27%	3.12%	4.03%	5.15%	1.74%	2.36%	3.48%	4.35%	5.22%	2.34%	3.29%	5.02%	5.69%	6.69%
N-dist	2.55%	3.49%	4.10%	4.77%	5.40%	3.23%	4.23%	4.85%	5.47%	5.72%	3.68%	4.01%	4.35%	5.35%	6.02%
t-dist	1.68%	2.93%	4.23%	5.40%	6.27%	2.49%	3.73%	5.10%	5.97%	6.59%	3.01%	4.01%	5.35%	5.35%	7.69%
HS-VIX	1.35%	2.37%	3.20%	3.99%	5.09%	1.74%	2.24%	3.61%	4.60%	5.85%	1.67%	3.34%	4.01%	5.35%	6.35%
N-dist (VIX)	0.16%	0.49%	0.79%	1.32%	1.88%	0.00%	0.12%	0.25%	0.50%	0.62%	0.00%	0.00%	0.33%	0.33%	0.33%
t-dist (VIX)	0.54%	2.19%	3.58%	5.13%	6.68%	0.00%	0.12%	0.75%	2.11%	2.74%	0.00%	0.33%	1.34%	2.01%	3.68%
	LR UNC														
HS	15.55	11.58	6.32	1.74	0.65	1.71	4.33	0.96	0.25	0.08	7.94	2.30	2.49	1.29	0.28
	< 0.01	< 0.01	0.01	0.19	0.42	0.19	0.04	0.33	0.61	0.77	< 0.01	0.13	0.11	0.26	0.59
VWHS	1.94	0.77	0.29	0.01	0.25	3.65	0.51	0.61	0.25	0.08	3.93	2.59	3.49	1.96	1.63
	0.16	0.38	0.59	0.92	0.62	0.47	0.47	0.43	0.61	0.77	0.05	0.11	0.06	0.16	0.20
N-dist	93.68	51.16	20.70	8.03	1.82	25.52	15.49	8.01	4.09	0.84	12.86	4.80	1.64	1.29	0.62
	< 0.01	< 0.01	< 0.01	0.00	0.18	< 0.01	< 0.01	< 0.01	0.04	0.36	< 0.01	0.03	0.20	0.26	0.43
T-dist	21.45	21.22	25.45	25.56	17.38	13.38	9.82	10.11	7.10	3.92	7.94	4.80	4.63	4.63	3.95
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.05	< 0.01	0.03	0.03	0.03	0.05
HS-VIX	6.33	3.57	0.72	0.00	0.10	3.65	0.23	0.96	0.72	1.15	1.14	2.30	0.96	1.29	1.07
	0.01	0.06	0.40	0.97	0.75	0.47	0.64	0.33	0.40	0.28	0.29	1.13	0.33	0.26	0.30
N-dist (VIX)	60.43	92.55	130.07	139.03	147.67	16.16	25.27	35.46	41.38	52.06	7.42	12.24	11.77	17.37	23.17
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
t-dist (VIX)	14.09	0.94	5.95	16.93	29.98	16.16	24.89	19.96	8.94	10.31	6.01	6.47	3.56	3.77	1.20
	< 0.01	0.33	0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.01	0.01	0.06	0.05	0.27
	LR IND														
HS	16.10	19.71	33.39	38.10	28.17	1.94	4.14	5.55	3.24	5.33	1.34	1.02	0.27	0.13	0.12
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.16	0.04	0.02	0.07	0.02	0.25	0.31	0.61	0.71	0.73
VWHS	3.78	7.73	2.21	1.09	0.14	0.53	0.51	0.97	3.24	3.14	0.38	0.78	0.18	1.09	1.96
	0.05	< 0.01	0.14	0.30	0.71	0.47	0.47	0.32	0.07	0.08	0.54	0.38	0.67	0.30	0.16
N-dist	28.84	21.62	33.07	33.53	33.71	11.05	12.17	16.92	16.42	14.79	0.76	0.55	0.39	0.13	0.13
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.38	0.46	0.53	0.71	0.72
t-dist	17.93	23.91	27.69	36.04	34.16	6.25	8.38	15.22	13.29	12.82	1.34	0.55	0.13	0.13	0.62
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.25	0.46	0.71	0.71	0.43
HS-VIX	2.66	3.89	1.87	2.92	1.06	0.53	0.87	0.82	0.15	0.15	0.20	0.76	1.09	1.93	2.72
	0.10	0.05	0.17	0.09	0.30	0.47	0.35	0.36	0.70	0.70	0.65	0.38	0.30	0.17	0.10
N-dist (VIX)	0.03	0.27	0.72	2.89	1.76	0.00	1.00	0.01	0.05	0.07	0.00	0.00	0.01	0.01	0.01
	0.86	0.60	0.40	0.09	0.18	1.00	0.94	0.90	0.82	0.79	1.00	1.00	0.91	0.91	0.91
t-dist (VIX)	0.34	1.80	11.08	14.43	23.93	0.00	0.00	0.11	3.96	2.25	0.00	0.01	0.14	0.29	0.76
	0.56	0.18	< 0.01	< 0.01	< 0.01	1.00	0.94	0.75	0.05	0.13	1.00	0.91	0.71	0.59	0.38
	LR CC														
HS	31.64	31.29	39.71	39.85	28.82	3.65	8.47	6.50	3.49	5.41	9.28	3.32	2.76	1.42	0.40
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.16	0.01	0.04	0.17	0.07	< 0.01	0.19	0.25	0.50	0.82
VWHS	5.72	8.50	2.50	1.10	0.39	4.19	1.09	1.58	3.49	3.22	4.33	3.37	3.68	3.06	3.59
	0.06	0.01	0.29	0.58	0.82	0.12	0.58	0.45	0.17	0.20	0.12	0.19	0.16	0.22	0.17
N-dist	122.51	72.77	53.76	41.57	35.53	36.56	28.65	24.93	20.51	15.64	13.62	5.35	2.03	1.42	0.75
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07	0.36	0.50	0.69
T-dist	39.38	45.13	53.14	61.61	51.54	20.13	18.21	25.33	20.38	16.73	9.28	5.35	4.77	4.77	4.57
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07	0.09	0.09	0.10
HS-VIX	8.99	7.47	2.59	2.92	1.16	4.19	1.10	1.78	0.87	1.30	1.34	3.06	2.05	3.22	3.79
	0.10	0.02	0.27	0.23	0.56	0.12	0.58	0.41	0.65	0.52	0.51	0.22	0.36	0.20	0.15
N-dist (VIX)	60.46	92.83	130.79	141.92	149.43	16.16	25.28	35.48	41.43	52.13	7.42	12.24	11.78	17.38	23.18
	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.02	< 0.01	0.00	< 0.01	< 0.01
t-dist (VIX)	14.42	2.74	17.03	31.36	53.91	16.16	24.89	20.07	12.90	12.55	6.01	6.48	3.70	4.05	1.96
	< 0.01	0.25	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.05	0.04	0.16	0.13	0.37

Table 2: The table explains how the different models used in this thesis perform when forecasting VaR for different confidence levels (α). The models are forecasted over 1, 10, and 22-day horizons. The percentage of violations for each model is displayed, as well as test results to determine if the violations are consistent with the expected proportion (LR UNC). Additionally, the independence of each model's VaR violations is tested (LR IND). Most importantly, a joint test combining the predictions of the correct number of exceptions and the independence of these exceptions is shown (LR CC). The first number for each model is the test statistic value while the number below is the p-value.

6 Discussion

When analysing the distribution of the S&P 500 in our sample period, the kurtosis amounts to 10, indicating a leptokurtic distribution with higher probability mass in the tails, i.e., higher likelihood of extreme losses. This implies that a VaR model assuming normal distribution should according to theory be more likely to underestimate the VaR level and thus produce more violations than a VaR model assuming a t-distribution that accounts for the fat tails. This is not strictly in line with what we observe in the empirical test conducted in this study, as the N-dist only yields a higher percentage of violations than the t-dist for alpha values of 1% & 2% across all forecast horizons. Statistically, the N-dist model and t-dist model are rejected the same amount of times in all the test and exhibit too many violation in general over all confidence levels and time horizons, indicating that given our sample period, the property of accounting for fat tails is not enough to produce accurate VaR estimations. Furthermore, as the distribution of S&P 500 in our sample period also exhibits a negative skew, affecting the performance of the parametric models employed in this study negatively as it is not accounted for. Bams et al. (2014) argues that to obtain reliable VaR estimates, the necessity to include some non-normality and negative skewness in the innovation distribution.

Surprisingly, when incorporating the VIX index the N-distribution produce significantly fewer violations than what is expected both in absolute terms and relative to the t-distribution. While the N-dist (VIX) and t-dist (VIX) for most confidence intervals and horizons are rejected by the joint Christoffersen test, the LR IND test statistic in table 1 showcase that the models perform very well in terms of testing for independent violations. This would theoretically imply that the model successfully accounts for time-varying volatility, but unfortunately a more probable explanation for this result is that the violations are quite few and scattered. Worth noticing is that the t-dist (VIX) cannot by rejected by any test for the 22-day horizon. Since the VIX is an expectation of 22-day volatility, it should be most accurate when used to forecast volatility over a similar time horizon. Consequently, the VaR forecast for the 22-day horizon should be more accurate than those of shorter horizons as the wedge between implied volatility and realized volatility is smaller, ceteris paribus. Without intention to diminishing this observation, it is also important to point out that seemingly all models with exception of the N-dist (VIX) perform well the 22-day forecasting horizon, even the N-distribution.

The parametric VaR models employed in this thesis, simply and perhaps naively, incorporate the VIX index without adjustments. Prokopczuk and Wese Simen (2014) analyses the role of volatility risk premium for volatility forecasting accuracy, i.e., the discrepancies between realized volatility and the implied volatility. They utilize a model-free adjustment to account for the volatility risk premium and conclude that it result in significantly improved volatility forecasts. Bams et al. (2014) use a similar adjustment methodology but apply it in the context of VaR estimations. In their study they compare the forecasts accuracy of non-parametric VaR models based on implied volatility and conditional volatility. They find that when accounting for the volatility risk premium the VaR forecast immediately improve, but not enough to be superior to the model using historical data in a GARCH framework. The parametric models employed in this study incorporate the unadjusted VIX, and thus, the findings of Bams et al. (2014) i.e, that the VIX includes a volatility risk premium, being a plausible explanation to why the VaR forecasts are systematically overestimated and often rejected when backtested. Additionally, all models incorporating the VIX-index tend to exhibit high volatility, posing a challenge for banks seeking stable forecasts to avoid frequent adjustments to their reserves.

Given the results obtained from this study, we can conclude that the HS-VIX model is performing well even when tested on more recent data. This is in line with the findings of Nossman & Wilhelmsson (2012). Regarding our parametric models incorporating options implied data, we find that they cannot produce accurate VaR predictions, with the exemption for the 22-day horizon. This is in contrast to the study conducted by Giot (2003), using the skewed student's t-distribution and definitely in divergence to the findings of Molino and Sala (2020). In common with all the other studies is that we use same stock and volatility indeces.

Given the finding and the continuous learning process in this study, suggestions for future research could be to investigate the VaR performance of a parametric model, accounting for negative skewness and non-normality, for example the skewed student t-distribution and the incorporation of adjusted VIX index. As those have been studied separately (Giot, 2003) and (Bams et al. 2014) but not in conjunction and as earlier mentioned there is also a shortage of research regarding parametric VaR models and VIX. Futhermore, methods to improve and appropriately incorporate adjustments to the VIX-index to better align with realized volatility and thus mitigate the overestimation of VaR predictions and enhance their accuracy could also be explored. For instance, integrating the VIX index within the GARCH framework can more effectively capture volatility dynamics, potentially leading to more accurate and stable VaR estimates. Additionally, using a t-distribution within the GARCH framework can address the presence of fat tails in the distribution of returns, thereby improving the robustness of VaR estimates by better capturing extreme market movements. To further stabilize VaR forecasts, smoothing mechanisms or moving average techniques can be developed and implemented to the VIX index. These approaches help reduce the volatility of VaR estimates, providing banks with more consistent and reliable forecasts. Conducting a detailed analysis of the variance premium and its impact on VaR forecasts is also crucial. Understanding the relationship between the variance premium and realized volatility can lead to better adjustments in parametric models, enhancing the accuracy and reliability of VaR estimates.

7 Conclusion

This thesis is based on the premise that the VIX index is the market's best estimation of volatility for the next 22 trading days and that it can easily be incorporated into the VaR framework, yielding potentially accurate forecasts or at least being quicker in reacting to changes in the market conditions and therefore be able to rise before losses occur.

The answer to whether incorporating implied volatility can improve forecasting VaR is not completely black on white. Since the time frames include two extreme events, the financial crisis and the Covid-19 pandemic making VaR forecasts especially challenging. As have been seen in this study the parametric models that incorporate the VIX overestimate the VaR forecasts. This could be a result of the variance risk premium which makes the implied volatility higher than the realized volatility thus resulting in too high forecasts. This leads to too few violations and rejected Christoffersen tests regarding correct proportions of violations and independence.

The VWHS and the HS-VIX are arguably most successful models when it comes to performance of joint tests. The HS-VIX could arguably be the most successful of the two as it follows the loss curves more closely than the VWHS thus would result in a more effective capital requirement allocation. That the HS-VIX model is the most dominant model of the volatility implied models is in line with (Nossman and Villhelmsson, 2012) and further shows that incorporating implied volatility could be advantageous regarding forecasting VaR but further implementations might be needed to get even more efficient models.

8 References

Bams, D., Blanchard, G., & Lehnert, T. (2017). Volatility measures and Value-at-Risk. *International Journal of Forecasting*, vol 33, no. 3, pp. 848-863, https://www.sciencedirect.com/science/article/pii/S0169207017300511?ref=pdf_download&fr=RR-2&rr=887dde68a852abcf

Banerjee, P. S., Doran, J. S., & Peterson, D. R. (2007). Implied volatility and future portfolio returns. *Journal of Banking and Finance*, vol. 31, no. 10, pp. 3183-3199, https://doi.org/10.1016/j.jbankfin.2006.12.007.

BCBS. (2009). Revisions to the Basel II market risk framework [PDF], https://www.bis.org/publ/bcbs158.pdf.

Beckers, S. (1981). Standard deviation implied in option prices as predictors of future stock price variability. *J Bank Finance*, vol. 5, no. 3, pp. 363-381.

Bekaert, G., Engstrom, E., & Ermolov, A., (2020). The Variance Risk Premium in Equilibrium Models [pdf], https://business.columbia.edu/sites/default/files-efs/pubfiles/26214/Variance_risk.pdf

Bekaert, G., & Hoerova, M. (2014). The VIX, the variance premium and stock market volatility. *Journal of Econometrics*, vol. 183 no. 2, pp. 181-192. https://www.sciencedirect.com/science/article/pii/S0304407614001110

Cboe. (2022). Volatility Index Methodology: Cboe Volatility Index [pdf], https://cdn.cboe.com/api/global/us_indices/governance/VIX_Methodology.pdf

Christensen, B. J., & Hansen, C. S. (2002). New evidence on the implied-realized volatility relation. *The European Journal of Finance*, vol. 8, no. 2, pp. 187-205. https://doi.org/10.1080/13518470110071209.

Christensen, B., & Prabhala, N. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, vol. 50, no. 2, pp. 125–150, https://doi.org/10.1016/S0304-405X(98)00034-8.

Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review*, vol. 39, no. 4, pp. 841-862, https://doi.org/10.2307/2527341

Christoffersen, P., Hahn, J., & Inoue, A. (2001). Testing and comparing value-at-risk measures. *Journal of Empirical Finance*, vol. 8, no. 3, pp. 325-342, https://www.sciencedirect.com/science/article/pii/S0927539801000251

Christoffersen, P., Jacobs, K., & Chang, B. (2013). Forecasting whith Option-Implied Information. *Handbook of Economic Forecasting*, vol. 2, pp. 581-656, https://www.sciencedirect.com/science/article/pii/B9780444536839000104

Chong, J. (2004). Value at risk from econometric models and implied from currency options. *Journal of Forecasting*, vol. 23, no. 8, pp. 603-620, https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.934

Dowd, K. (2005). *Measuring Market Risk* (2nd ed.). Chichester: John Wiley Sons.

Giot, P. (2003). Implied volatility indexes and daily value-at-risk models. *Journal of Financial Research*, vol. 26, no. 3, pp. 353-376, https://web.archive.org/web/20050519031631id_/http://www.core.ucl.ac.be:80/econometrics/Giot/Papers/IMPLIED3_g.pdf

Gospodinov, N., Gavala, A., & Jiang, D. (2006). Forecasting volatility. *Journal of Forecasting*, vol. 25, no. 6, pp. 381-40, https://doi.org/10.1002/for.993.

Granger, C., & Poon, S-H. (2003). Forecasting Volatility in Financial Markets: A review, *Journal of economic Literature*, vol. 41, no. 2, pp. 478-539, https://www.jstor.org/stable/3216966?seq=7

Grover, R., & Thomas, T. (2011). Liquidity considerations in estimating implied volatility, 24th Australasian Finance and Banking Conference 2011 Paper, https://ssrn.com/abstract=1832236 [Accessed 28 April 2024]

Hull, J. C. (2012). *Options, Futures, and Other Derivatives: Global Edition* (8th ed.), Harlow: Pearson Education Limited.

Hull, J. C. (2023). *Risk Management and Financial Institutions* (6th ed.), NJ: John WIley & Sons, Inc., Hoboken.

Hull, J. C., & White, A. (1998a). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, vol. 1, no. 1, pp. 5-19, https://www.researchgate.net/profile/John-Hull-6/publication/2645882_Incorporating_volatility_updating_into_the_historical_simulation_method_for_VaR/links/00b7d5335d8e2394d0000000/Incorporating-volatility-updating-into-the-historical-simulation-method-for-VaR.pdf.

Hull, J. C., & White, A. (1998b). Value at risk when daily changes in market variables are not normally distributed. *Journal of Derivatives*, vol. 5, no. 3, pp. 9-19, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=6b1c6e7cee3f756be964f3c9cee9848e2fcf7646.

Jorion, P. (2007) *Value at Risk: The New Benchmark for Managing Financial Risk*. 3rd edn. New York: McGraw-Hill Professional

J.P. Morgan & Reuters. (1996). RiskMetrics™- Technical Document, https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a [Accessed 16 April 2024].

Latane, H. A., & Rendleman, R. J. (1976). Standard deviations of stock price ratios implied in option prices. *The Journal of Finance*, vol. 31, no. 2, pp. 369–381, https://doi.org/10.2307/2326608.

Le, T. (2020). *Analyzing Intraday Implied Volatility For Pricing Currency Options*, Cham: Springer

Li, S., & Yang, Q. (2009). The relationship between implied and realized volatility: evidence from the Australian stock index option market. *Review of Quantitative Finance and Accounting*, vol. 32, pp. 405-419, https://www.researchgate.net/publication/46547743_The_relationship_between_implied_and_realized_volatility_Evidence_from_the_Australian_stock_index_option_market

Molino, A., & Sala, C. (2021). Forecasting value at risk and conditional value at risk using option market data. *Journal of Forecasting*, vol. 40, no. 7, pp. 1190-1213, https://onlinelibrary.wiley.com/doi/full/10.1002/for.2756

Moran, M. T., & Liu, B. (2020). The VIX index and volatility-based global indexes and trading instruments. *CFA Institute Research Foundation*. https://rpc.cfainstitute.org/-/media/documents/article/rf-brief/rfbr-moran-vix-volatility.pdf

Novales, A. & Garcia-Jorcano, L. (2019). Backtesting extreme value theory models of expected shortfall, *Quantitative Finance*, vol. 19, no. 5, pp. 799–825, https://www.tandfonline.com/doi/full/10.1080/14697688.2018.1535182

Nossman, M., & Vilhelmsson, A. (2012). Non-parametric forward looking value-at-risk [PDF]. https://www.kyos.jp/wp-content/uploads/2016/06/Forward_looking_JoR.pdf.

Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking & Finance*, vol. 30, no. 2, pp. 561-582, https://doi.org/10.1016/j.jbankfin.2005.04.013

Prokopczuk, M., & Wese Simen, C. (2014). The importance of the volatility risk premium for volatility forecasting. *Journal of Banking & Finance*, vol. 40, pp. 303–320, https://www.sciencedirect.com/science/article/pii/S0378426613004640?ref=pdf_download&fr=RR-2&rr=887dd78bc9db92c1

Schindelhauer, K., & Zhou, C. (2018) Value-at-Risk prediction using option-implied risk measures. [PDF], https://www.dnb.nl/media/hordszwd/working-paper-no-613_tcm47-379849.pdf

Sharma, G., Kayal, P., Pandey, P. (2019). Information Linkages Among BRICS Countries: Empirical Evidence from Implied Volatility Indices. *Journal of Emerging Market Finance*, vol. 18, no. 3, pp. 263-289, https://doi.org/10.1177/0972652719846315

Stambaugh, F. (1996). Risk and Value at Risk. *European Management Journal*, 14(6), 612-621. https://www.sciencedirect.com/science/article/pii/S0263237396000576.

Wang, Y. H., & Wang, Y. Y. (2016). The information content of intraday implied volatility for volatility forecasting. *Journal of Forecasting*, vol. 35, no. 2, pp. 167-178,

Whaley, R. E. (2009). Undertstanding the VIX. *Journal of Portfolio Management*. Spring 35 (3), pp. 98-105, https://www.researchgate.net/publication/277429711_Understanding_the_VIX

Zhou, H. (2018) Variance Risk Premia, Asset Predictability Puzzles, and Macroeconomic Uncertainty. *Annual Review of Financial Economics*, Vol. 10, pp. 481-497, https://www.jstor.org/stable/26774130?seq=1

9 Appendix

Figure 3: Comparison of Losses (1day95), VaR - BHS, VaR - ewma, VaR - n, VaR - t

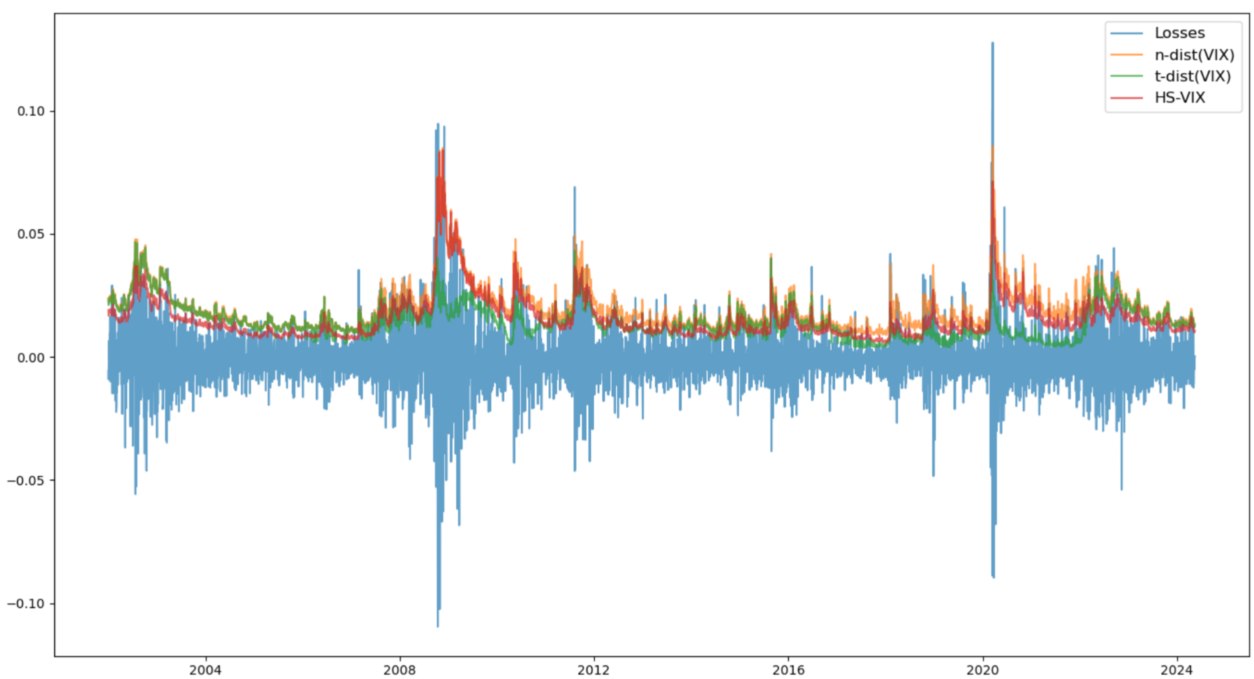

Figure 4: Comparison of Losses(1day95), VaR - tVIX, VaR - nVIX, VaR - VIX,

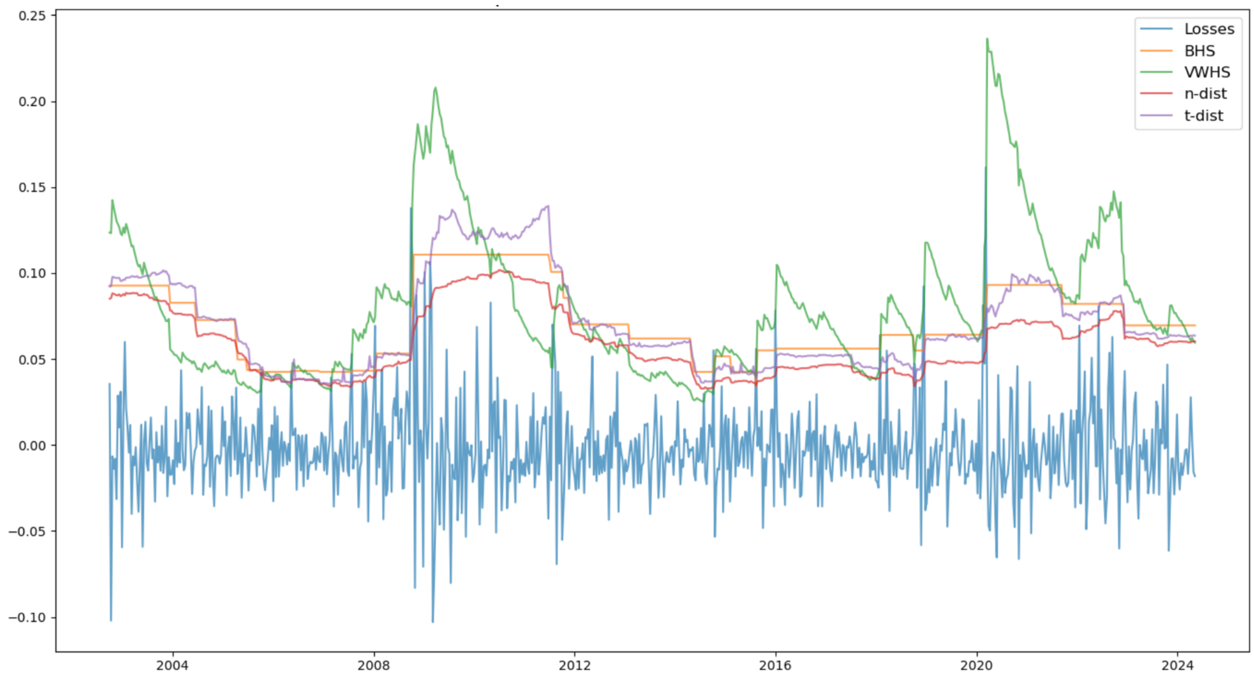

Figure 5: Comparison of Losses (10day99), VaR - BHS, VaR - ewma, VaR - n, VaR - t

Figure 6: Comparison of Losses(10day99), VaR - tVIX, VaR - nVIX, VaR - VIX,

Figure 7: Comparison of Losses (10day95), VaR - BHS, VaR - ewma, VaR - n, VaR - t

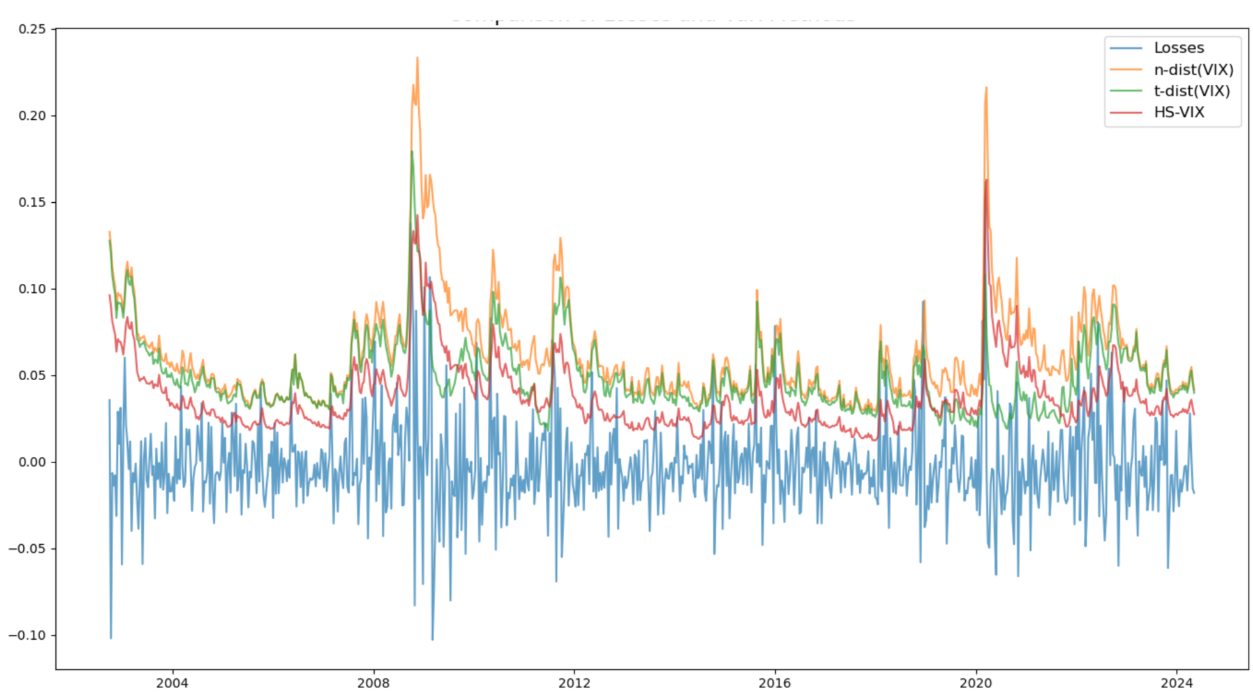

Figure 8: Comparison of Losses(10day95), VaR - tVIX, VaR - nVIX, VaR - VIX,

Figure 9: Comparison of Losses (22day99), VaR - BHS, VaR - ewma, VaR - n, VaR - t

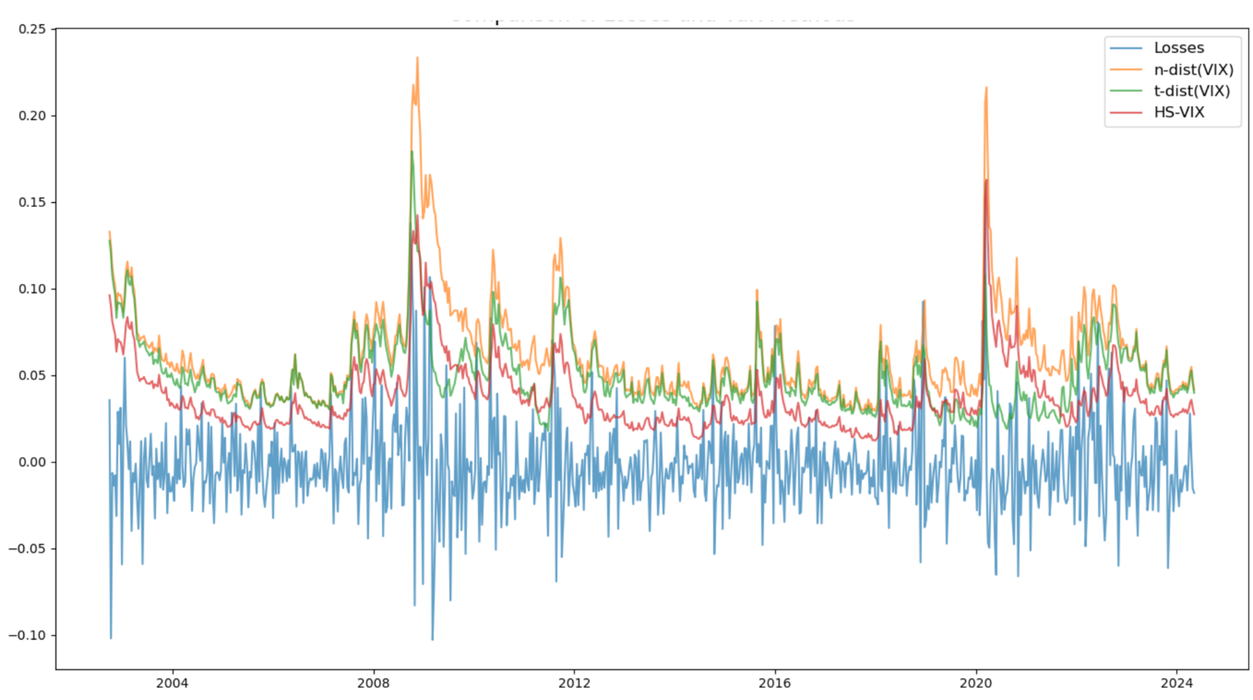

Figure 10: Comparison of Losses(22day99), VaR - tVIX, VaR - nVIX, VaR - VIX,

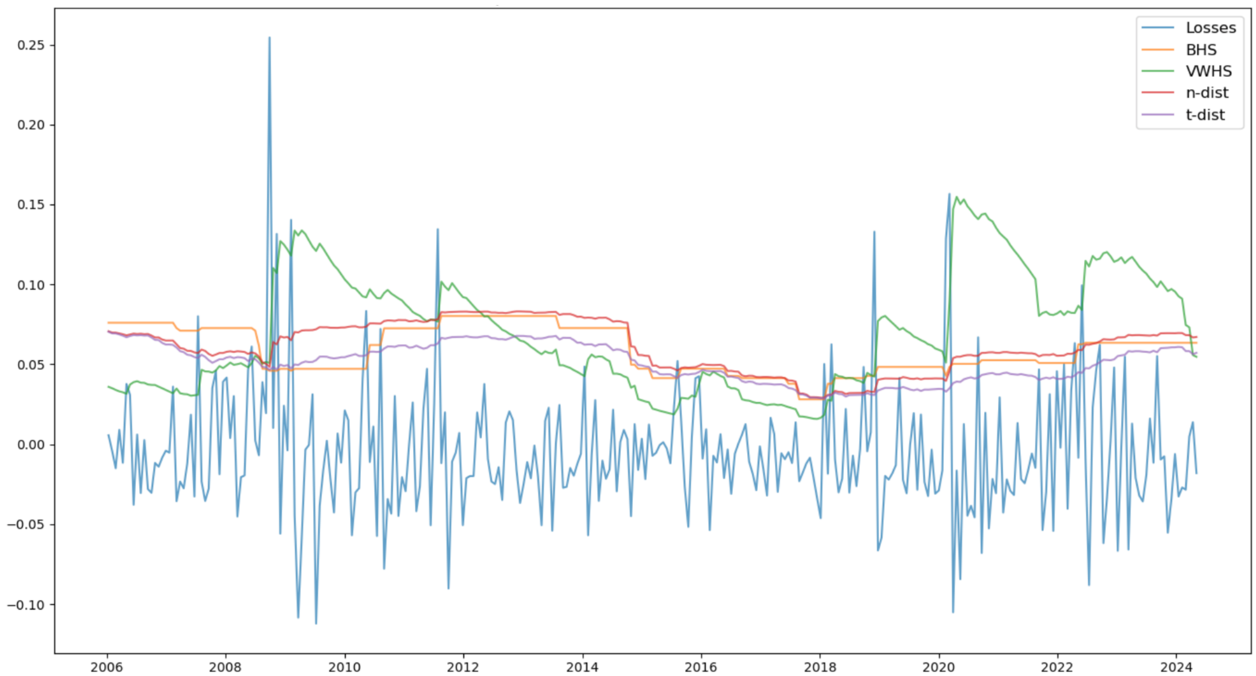

Figure 11: Comparison of Losses (22day95), VaR - BHS, VaR - ewma, VaR - n, VaR - t

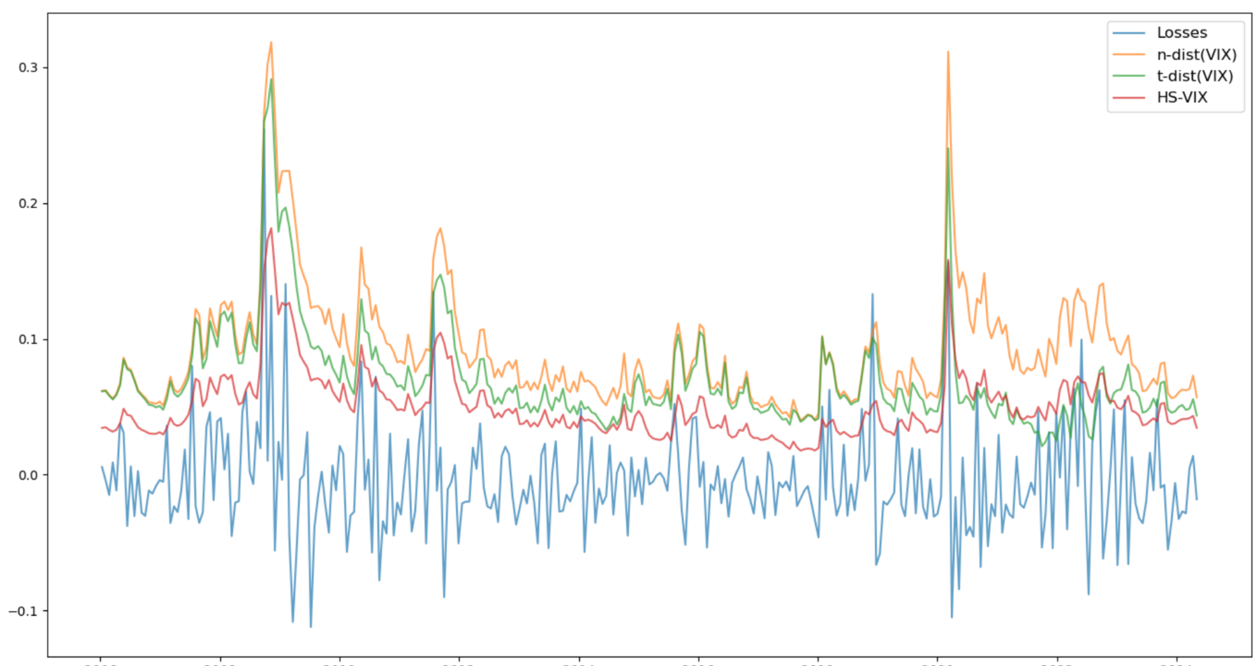

Figure 12: Comparison of Losses(22day95), VaR - tVIX, VaR - nVIX, VaR - VIX,