Volatility Timing using Machine Learning

An Application to a Signal Based Portfolio

M.Sc. Thesis

Lund School of Economics and Management (LUSEM)
Lund University

Filippa Lövgren

Julian Ulmer

Supervisors:

Hossein Asgharian & Anders Vilhelmsson

24 May 2022

Abstract:

Recent events such as the covid-19 pandemic and the Russian-Ukrainian war have led to a tremendous increase in volatility, making financial markets riskier for investors. To see whether investors can counteract or profit from such risk, we develop a volatility timed trading strategy. To do so, a volatility forecast is used, in our case, the direction of the VIX. We predict the one week ahead movement of the VIX by deploying several machine learning algorithms, resulting in the support vector machine with a linear kernel being the best model. The volatility timed portfolio is based on the forecast acting as a signal, investing in equities if the volatility decreases and in fixed income if it increases. To see whether an investor is better off with that portfolio than a 60% equity and 40% fixed income portfolio, we compare the two using the Sharpe ratio. Furthermore, a split on market regimes is done to see whether there are performance differences between the portfolios for different levels of the VIX. We find that the signal-based portfolio is unable to outperform the 60/40 portfolio over the whole period, while for certain market regimes, it has a substantial advantage.

Keywords: *Machine Learning, Support Vector Machines, VIX, Volatility Timing, Portfolio Construction*

Acknowledgment:

We would like to express our deepest appreciation to our supervisors, Hossein Asgharian and Anders Vilhelmsson. Your knowledge, valuable advice, and comments have been of the utmost importance to this paper. We would also like to extend our sincere thanks to Thorbjörn Wallentin and OQAM Asset Management for your ideas and guidance. Lastly, we want to thank all our friends who supported us in the process of writing this thesis.

Outline

Abbreviations

AR	Autoregressive
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrated Moving Average
CBOE	Chicago Board Options Exchange
CV	Conditional Variance of the Stock Market
ETF	Exchange Traded Fund
FRED	Federal Reserve Economic Data
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GDP	Gross Domestic Product
HAR	Heterogenous Autoregressive
LPM	Linear Probability Model
RBF	Radial Basis Function
ROC	Receiver Operating Characteristic
SABR	Stochastic Alpha Beta Rho
SVM	Support Vector Machines
VAR	Vector Autoregressive
VIX	CBOE Volatility Index
VP	Equity Variance Risk Premium

1. Introduction

In recent years, financial markets have experienced a substantial increase in volatility due to extreme events, such as the covid-19 pandemic and the Russian-Ukrainian war. This higher volatility has increased the investment risk and led to a resumed interest in trading strategies minimizing the potential downside (Taylor, 2022). Most individuals would associate volatility in the financial market with potential losses, while some people could see it as an opportunity. To exploit such an opportunity, the investor would need knowledge about forecasting volatility as well as the necessary skills to form an investment decision out of these forecasts.

Whether this ability yields an investor an advantage over a standard portfolio allocation is the goal to investigate in this paper. To do so, we are applying multiple machine learning algorithms, as well as the probit model, as an econometric benchmark, to predict next week's movement of a volatility proxy, in our case, the CBOE Volatility Index (VIX). Choosing the model with the highest accuracy, we build a signal-based trading strategy that invests in an equity index if our forecast of the VIX decreases and in fixed income securities when our forecast of the VIX increases. The same procedure is done for a benchmark portfolio, using the forecasts of the probit model. Comparing the two portfolios' Sharpe ratios to the one of a standard 60% equity, 40% fixed income (60/40) portfolio (Rekenthaler, 2022) will give us insights into whether the knowledge about forecasting volatility can result in excess returns for an investor.

Trying to forecast volatility is, however, not a new phenomenon; it has long been essential for option pricing (Scott, 1987), portfolio selection (Fleming, Kirby & Ostdiek, 2001; Taylor, 2022), and risk management (Christoffersen & Diebold, 2000; Yang, Chen & Tian, 2015). Even if much research has been done, successfully forecasting volatility is yet a challenge (Wang, 2019). Some of the models mainly used to forecast volatility are the autoregressive conditional heteroscedasticity (ARCH) model (Engle, 1982) and the generalized autoregressive conditional heteroscedasticity (GARCH) model (Bollerslev, 1986). A model-free way of measuring volatility is the VIX, which measures the expected stock market volatility based on the U.S. S&P 500 index options (Gonzalez-Perez, 2015; Duan & Yeh, 2010). There have been many attempts to forecast the VIX (see Aboura, 2003; Ahoniemi, 2007; Fernandes, Medeiros & Scharth, 2007; Konstantinidi, Skiadopoulos & Tzagkaraki, 2008; Qiao, Yang & Li, 2020; Stavros, 2008). As the VIX is highly persistent (Fernandes, Medeiros & Scharth, 2014), most studies forecasting the VIX have used different types of autoregressive (AR) models. When predicting the stock market's direction, a popular machine learning approach has been to use support vector machines (SVM), as the model can solve non-linear classification problems

(Boser, Guyon & Vapnik, 1992). Furthermore, Luong and Dokuchaev (2018) have successfully been forecasting the direction of realized volatility by combining machine learning models and traditional AR models.

Investors want to take advantage of these volatility forecasts by deploying a trading strategy that incorporates these findings and helps them achieve excess returns over a simple market portfolio. Such volatility timed trading strategies depend on the quality of the abovementioned volatility forecasts (Copeland & Copeland, 1999). Resulting dynamic portfolios are usually built by reweighting asset weights depending on the forecasted volatility and thus market riskiness. Assets used vary from only equities, such as small and big market capitalization companies, as well as value and growth companies (Copeland & Copeland, 1999), to bonds and gold (Fleming, Kirby & Ostdiek, 2001). Different studies reach different conclusions about the quality of such volatility timing models (see e.g., Copeland & Copeland, 1999; Fleming, Kirby & Ostdiek, 2001; Fleming, Kirby & Ostdiek, 2003; Taylor, 2022; Liu, Tang, & Zhou, 2019).

This paper can be split into two crucial parts. First, we build a model that predicts the sign of the weekly changes in the market's volatility. We implement a machine learning algorithm to achieve that goal, using the VIX as a proxy for market volatility. We deploy a random forest, multiple support vector machines, as well as the probit model as an econometric benchmark, choosing the machine learning model with the best out-of-sample accuracy as our final setup. The data used consist of U.S. equity indices, commodities, fixed income, macroeconomic variables, and VIX futures. Because of the persistent nature of the VIX, we also choose to include lagged VIX as a variable. Second, we incorporate our forecasted VIX movement, both for our best machine learning model and the probit model, into signal-based trading strategies, with the goal to outperform a conventional 60% equity, 40% fixed income portfolio. To measure whether we outperform such a portfolio, we compare a risk-return measure, namely the Sharpe ratio. To achieve our goal, we define a trading strategy that invests in fixed income if the model predicts an increase in volatility and invests entirely in a stock index if it predicts a decrease in volatility. Neglecting transaction costs, we continuously invest our whole position into one of the two assets, meaning that we will either stay within our position if the forecasted direction of the volatility does not change or sell our whole position and invest into the other option if the forecasted direction of the volatility, and therefore the sign, changes.

We contribute to the literature by implementing machine learning algorithms to a VIX movement prediction, which to our knowledge, has not been done before. We display valuable insights into model choice and variable choice as we investigate which variables significantly impact such a sign prediction. In addition, we add to the research of volatility timing models,

differentiating from the typically reweighted portfolios by investing 100% into either equities or fixed income. Finally, we compare the performance of the portfolios on different levels of the one week lagged VIX (the time the prediction is made). Adding multiple new approaches in one thesis can confuse what the results are related to; therefore, we do all modeling of the movement of the VIX at first and discuss the results on their own. After that, the portfolio construction is evaluated, we also choose to include perfect prediction (the actual observations) besides our existing models. Including perfect prediction allows us to test whether the portfolio construction actually works, independent of our models. With this procedure, we can get valid results for both the sign prediction and the portfolio construction.

Achieving a forecast accuracy of 58.4% using a support vector machine with a linear kernel allows us to slightly outperform our econometric benchmark which has an accuracy of 56.9%. These results lie within an acceptable range compared to other researchers predicting the movement of the VIX. Building a signal-based portfolio incorporating these forecasts, we are unable to beat a standard portfolio of 60% equity and 40% fixed income over the chosen time frame. Splitting the time frame according to market riskiness, the portfolio built on the probit model achieves the highest Sharpe ratio out of all tested portfolios for VIX levels below 20. In contrast, the portfolio based on the support vector machine with a linear kernel achieves the highest Sharpe ratio of all tested portfolios for VIX levels between 20 and 30. Neither signal-based portfolio achieves meaningful results for VIX levels above 30, resulting in the 60/40 portfolio performing best for a high market risk regime. The obtained results with perfect predictions show the lowest Sharpe ratio for VIX levels above 30, compared to the other two market regimes, indicating that our defined trading strategy is worse for times with high market volatility.

The rest of this paper is organized as follows. In section 2, previous literature about volatility modeling, as well as volatility timing, is presented. Section 3 provides the theoretical framework and our model specification for the used machine learning algorithms and the trading strategy. Section 4 displays the used data and the necessary transformations to make modeling possible. In Section 5, the results of this study are presented and discussed. The paper is concluded in section 6, summarizing our findings and giving an outlook on future research.

2. Literature Review

The following section presents the previous literature on the relationship between asset returns and their volatility and how to capture such volatility using parametric models. Afterwards, we introduce the VIX as a market-implied volatility measure, more precisely how it is constructed, forecasted, and broken down into a volatility risk premium. Last, we introduce some volatility timing strategies to improve portfolio selection.

2.1 Volatility Models

Mandelbrot and Hudson (2006) explain that stock market movements do not follow the standard financial theory. Instead, they say that the stock market can move broadly and have much more risk than most people realize. Fama (1970) states that it is impossible to beat the market consistently. He explains that an investor can have big short-term profits; however, the profits will not be much higher than the market average in the long term. High short-term profits can exist because of a stock's volatility, which is the range in which a stock moves around its mean return. Black (1976) introduces the leverage effect, which states that a stock's return is negatively correlated with its volatility. Keeping this in mind, it is no surprise that many researchers are interested in understanding the stock market volatility. Especially so, as volatility plays a considerable role in option pricing (Scott, 1987), investment decisions (Fleming, Kirby & Ostdiek, 2001; Taylor, 2022), and short-horizon risk management (Christoffersen & Diebold, 2000). To understand such volatility, the correlation between financial leverage and volatility in the stock market (Black, 1976; Christie, 1982 & Schwert, 1988) is crucial. Additionally, Schwert (1988) and Christie (1982) find a positive relationship between interest rates and volatility. Furthermore, Schwert (1988) states that the stock market volatility is connected to corporate profitability. To capture such volatility, models like the autoregressive conditional heteroscedasticity (ARCH) model (Engle, 1982), the generalized autoregressive conditional heteroscedasticity (GARCH) model (Bollerslev, 1986) as well as multiple additions to the latter have been deployed successfully. More sophisticated approaches include various stochastic volatility models, including the Heston model (Heston, 1993) and the stochastic alpha beta rho (SABR) model (Hagan, Lesniewski & Woodward, 2015).

Compared to the abovementioned ways, another more straightforward option to capture volatility is to use a market-implied measure like the VIX, the CBOE volatility index (Gonzalez-Perez, 2015; Duan & Yeh, 2010). The VIX is constructed following the framework proposed by Britten-Jones and Neuberger (2000), backing out information about an underlying price process from option prices. In the case of the VIX, the volatility is computed from the option prices of the S&P 500 index for the next 30 days. The VIX is seen as model-free as it

does not depend on a valuation model for options, like the Black-Scholes model (Gonzalez-Perez, 2015). The use of the VIX is justified as it possesses a significant negative correlation with the S&P 500 returns (Gonzalez-Perez, 2015; Fernandes, Medeiros & Scharth, 2014; Copeland & Copeland, 1999) and therefore is a suitable proxy for volatility.

Fernandes, Medeiros and Scharth (2014) name multiple authors (Konstantinidi, Skiadopoulos & Tzagkaraki, 2008; Carr & Lee, 2009; Clements & Fuller, 2012) who argue that forecasting the VIX is a crucial step when an investor wants to trade in CBOE options or exploit it for hedging purposes. Studies focusing on forecasting the VIX include Nossman and Wilhelmsson (2009), who test how well the VIX can be predicted with VIX future data. They show that when adjusting for a possible risk premium, the VIX future data exceptionally predicts the future VIX; this applies to one-day to one-month forecasts. When not adjusting for the risk premium, the prediction is not as good; however, still above 50% for both one-day and one-month forecasts. Another remarkable finding in predicting the VIX is the industrial production growth (Corradi, Distaso & Mele, 2013), which is related to the gross domestic product (GDP), suggested as a key indicator by Engle and Rangel (2008). To predict VIX, the use of autoregressive models is common; as mentioned in Nossman and Vilhelmsson (2009), several authors, including Aboura (2003); Ahoniemi (2007); Fernandes, Medeiros & Scharth (2007); Konstantinidi, Skiadopoulos & Tzagkaraki (2008) use autoregressive or vector autoregressive (VAR) models. One reason for this is that there is evidence for the VIX to be highly persistent (Fernandes, Medeiros & Scharth, 2014), meaning that the VIX of tomorrow most likely will be similar to the VIX today.

There is a lack of literature on how to predict the sign of the movement of the VIX. However, there exists some literature on how machine learning has been implemented to forecast directions in the stock market. One of the more popular models of predicting movements in the stock market is the support vector machine; the model can solve non-linear classification problems (Boser, Guyon & Vapnik, 1992). Huang, Nakamori and Wang (2005) use support vector machines to predict the weekly movement direction of the NIKKEI 225 index. They outperform linear discriminant analysis, quadratic discriminant analysis, and Elman backpropagation neural networks. Ren, Wu and Liu (2018) use a support vector machine and combine financial market data and sentiment variables such as unstructured news data to predict the direction of the SSE 50 Index with impressive results. Pai and Lin (2005) take advantage of the traditional autoregressive integrated moving average (ARIMA) model and integrate it with a support vector machine to predict stock prices; the results are optimistic. Luong and Dokuchaev (2018) use a similar approach of combining autoregressive models and machine learning algorithms by incorporating the heterogeneous autoregressive (HAR) model and the

random forest algorithm to predict the direction of realized volatility, which shows an improvement from only using the HAR model.

Many researchers suggest that the VIX can be used to compute a volatility risk premium, looking at the difference between realized and expected volatilities (Gonzalez-Perez, 2015; Bekaert & Hoerova, 2014; Duan & Yeh, 2010). Bekaert and Hoerova (2014) do so by breaking down the VIX (squared VIX) into the conditional variance of the stock market (CV), which is the expected realized volatility, and an equity variance risk premium (VP), which is the difference between the VIX and the CV. These two measures are then used to predict stock returns, economic activity, and financial stress. Furthermore, they find that the CV significantly predicts economic activity and has more predictive power than the VP when predicting financial stress, while the VP significantly predicts stock returns. Duan and Yeh (2010) compute a volatility risk premium as well as a jump risk premium, linking the VIX to the latent stochastic volatility under the risk-neutral measure.

2.2 Volatility Timing

When using a market timing strategy based on the VIX like Copeland and Copeland (1999), good forecasting is crucial to ensure a correct market signal. They suggest investing in big-cap companies or value stocks when volatility increases and investing in small-cap companies or growth stocks when volatility decreases. With that approach, they find that following the VIX market signal yields the investor excess returns, making it a suitable method to improve portfolio selection. Fleming, Kirby and Ostdiek (2001) suggest another volatility timing strategy, intending to optimize the mean-variance trade-off in a short time horizon by allowing for daily reweighting of investment positions in a stock index, bonds and gold. They compute the covariance-variance matrix by a nonparametric approach, which nests most autoregressive and stochastic processes and computes the portfolio weights by minimizing the asymptotic mean squared error of the estimator. They find an inverse relationship between asset-specific volatility and the investment size in that asset, which is consistent with the general knowledge. According to Fleming, Kirby and Ostdiek (2001), the average investor would be willing to pay on average 170 basis points to switch towards the dynamic portfolios with a volatility timing strategy from the optimal static portfolios, which underlines the advantage of timing volatility. In 2003, Fleming, Kirby and Ostdiek wrote another paper, comparing their dynamic trading strategy of 2001 with a dynamic trading strategy using realized volatility from intraday returns. They find that this improves the performance of their volatility timing strategy even further. Taylor (2022) investigates which determinants impact the performance of volatility timing strategies. He finds two conditions, the first one being that the investor must be able to predict

future volatility correctly and the second one being the failure of the intertemporal capital asset pricing model. With these two conditions fulfilled, Taylor (2022) states that the investor can achieve an expected benefit of 2-3% per year over the unconditional mean-variance model. Nevertheless, he says that this requires either an extreme risk-return trade-off or the usage of a lot of leverage. Liu, Tang, and Zhou (2019) investigate different volatility timing strategies contradicting the findings of the previous authors. They test four different strategies proposed by different researchers, namely a volatility managed portfolio (Moreira & Muir, 2017), a volatility-trading strategy (Barosso & Santa Clara, 2015), a mean-variance portfolio allocation with estimation risk (Kan & Zhou, 2007), and an unconditional optimal portfolio with conditional information (Ferson & Siegel, 2001). No matter the strategy, Liu, Tang, and Zhou (2019) find that the market possesses a higher Sharpe ratio as well as a smaller maximum drawdown than the portfolios based on volatility timing unless there is a financial crisis and thus extreme levels of volatility.

3. Methodology

Predicting whether the VIX is moving upwards or downwards the following week is a classical binary classification problem. There exist many different models made to solve such a problem; a few generally used are logistic regression, linear discriminant analysis, quadratic discriminant analysis, naive Bayes, K-nearest neighbors, trees, boosting, random forest, and support vector machines (James, Witten, Hastie & Tibshirani, 2013). We choose to focus on the last two, with the main emphasis on different versions of the support vector machine.

In the following section, we will cover how we cross-validate our models, focusing on the time series property of our dataset. Then the different algorithms, as well as evaluation methods, are described. Finally, the relationship of the prediction with the implied market riskiness and thus investment choice is explained using a maximum Sharpe ratio approach.

3.1 Cross-Validation

For building an algorithm, training and test data are needed. One way of doing this is to simply use a certain percentage of the data to train the model (often around 70%) and the other part (often around 30%) to test the model (Liu & Cocea, 2017). A lot of training data is needed to build a good model; therefore, taking away 30% of the data only for testing can have a negative impact on the quality of the model. Cross-validation is a solution to this, using all the data for testing and training. In a simple 10-fold cross-validation, the observations are randomly divided into ten equal partitions. Then the model is trained on nine of these partitions, where the tenth is left out to act as a test set. This process is repeated for all ten groups, meaning each group is left out once for testing (Hastie, Tibshirani & Friedman, 2009).

3.1.1 Time Series Cross-Validation

For time series data, the order of the observations is essential and regular cross-validation does not work. Nevertheless, there exists a solution to this, namely cross-validation with recursive updating, demonstrated in Figure 1. In the first step of cross-validation with recursive updating, the observations of the first time period are used to train the model, and the observations of the time period right after are used for testing the model. In the next step, both the observations of the first and second time periods are used to train the model, and the observations of the third time period are used to test the model (Bergmeir & Benítez, 2012). This process continues until the last time period is tested. In this thesis, we start by using a time period of 50 weeks to predict week 51; after that, we use the first 51 weeks to predict week 52 and so on. This mechanism continues each week until we have used the 913 first weeks to predict week 914 (our last observation).

Cross-Validation with Recursive Updating					
Training	Test				
Training	Training	Test			
Training	Training	Training	Test		
Training	Training	Training	Training	Test	
Training	Training	Training	Training	Training	Test
Week 1-50	Week 51	Week 52	Week 53	Week 54	Week 55

Figure 1: Cross-Validation with Recursive Updating.

3.2 Random Forest

A random forest is an algorithm similar to bagging of trees. Bagging of trees is an algorithm that builds a large set of trees and then takes the average of them. Trees can model complex patterns in data; they have low bias but are often noisy (Hastie, Tibshirani & Friedman, 2009). Therefore, taking the average of all trees reduces the variance while keeping the bias low. When building a tree in bagging, the tree considers all p predictors before choosing the most important one and using that as the first split. In this way, the variables of the first splits will be similar for most trees, making them correlated to each other. Only a random sample m (of fixed size) of the p predictors will be considered when doing the splitting in a random forest. In this way, the predictors, and therefore splits in a tree, will differ from another tree, making them de-correlated. When taking the average of a large set of trees, the variance reduction will be much more significant for de-correlated trees than correlated ones, making the random forest a good variance reduction model (James et al., 2013).

In the random forest algorithm, the following procedure will be repeated for each training set in our cross-validation with recursive updating, thus 864 times. The random forest algorithm repeats the following step for a predetermined number of trees; we use 500 trees, which is also the default setting in R-Studio, meaning that $B = 500$ bootstrap samples will be drawn. First, a bootstrap sample Z_b^* will be drawn from the training data. Second, it grows a random forest tree T_b with the bootstrap data Z_b^*. The tree is grown by repeating three steps until the pre-fixed minimum node size is reached. The first step includes selecting m predictors randomly from the 17 ones in our dataset; m is a hyperparameter that must be tuned in our model. A regular choice of m is the square root of the total predictors, in this case $\sqrt{17}$. In our model, we optimize m, called "mtry" in R, building a range around $\sqrt{17}$, for integer values between two to nine, trying to achieve maximum forecast accuracy. The second step is to pick the best predictor for explaining differences in the predicted variable from the m predictors. The third

9

step is to split the node into two daughter nodes. The algorithm ends when the predetermined number of random forest trees T_b is constructed (Hastie, Tibshirani & Friedman, 2009).

The class prediction for a single tree b, inside the random forest is defined in equation (1):

$$\hat{C}_b(x) \tag{1}$$

As we work with a classification problem, the result of the random forest for each training set of the 864 is a majority vote of all B trees in the forest. The final prediction of the model is then given by equation (2):

$$\hat{C}_{rf}^B(x) = Majority\ vote \sum_{b=1}^{B} \hat{C}_b(x) \tag{2}$$

3.3 Support Vector Machines[1]

A support vector machine is a machine learning classification method motivated geometrically by fitting a separating hyperplane between the observations of (two) different classes within a dataset (Deisenroth, Faisal & Ong, 2020). In a two-dimensional space, such a hyperplane is displayed as a straight line separating the two classes based on two input features, as shown in Figure 2. In our case, these two classes are whether the following week's VIX moves upwards or downwards compared to this week. Mathematically a hyperplane is defined in equation (3):

$$\{x : f(x) = x^T\beta + \beta_0 = 0\} \tag{3}$$

where β is the direction vector of the hyperplane and β_0 is the position of the hyperplane. The optimal hyperplane, and thus support vector machine, is found by maximizing the margin M, the distance between the closest observations to the hyperplane and the hyperplane, with respect to the direction vector and position parameter (Hastie, Tibshirani & Friedman, 2009). The observations closest to the hyperplane are called support vectors and are the only observations that influence the position and direction of the hyperplane. With that in mind, the goal of the support vector machine is to train the hyperplane on our training dataset to maximize forecast accuracy on the classification in the test set. As previously explained, this is done multiple times as we use time series cross-validation with recursive updating. Using the resulting classification rule (equation 4), which depends on the margin maximizing parameter estimates $\hat{\beta}$ and $\widehat{\beta_0}$ as well as the feature data, we can obtain predictions of the direction of the VIX and compute the accuracy.

---

[1] For support vector machines we follow the notation of the book "Elements of Statistical Learning" (Hastie, Tibshirani & Friedman, 2009).

$$G(x) = sign\left(x^T \hat{\beta} + \widehat{\beta_0}\right) \qquad (4)$$

With that classification rule at hand, it is straightforward to decide whether a point is predicted to belong to class 1 or class -1. For a perfectly separable dataset, this can be seen in Figure 2, where the hyperplane is displayed as the thick black line, and the support vectors are the points on the dotted black line, which have the distance M to the hyperplane.
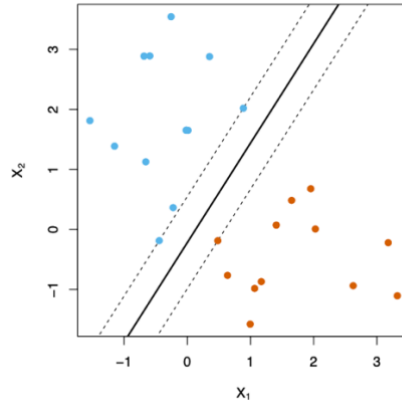

Figure 2: Visualization of an optimal separating hyperplane in a perfectly separable dataset (Efron & Hastie, 2021).

If a dataset is not separable, as in our case, it is impossible to fit a perfectly separating hyperplane between the observations of the two classes. To adjust for that problem, a budget is introduced, which will allow for a certain amount of margin violations (Efron & Hastie, 2021). A margin violation occurs whenever a support vector is inside the margin or on the opposite side of the hyperplane. Therefore, support vectors are now all observations that are either the margin M or less away from the hyperplane. The higher the budget, our new hyperparameter, the more margin violations are allowed. For the sake of computation, the budget is not used directly but rather a hyperparameter inverse to the budget. This so-called constant governing aversion C, also known as cost, is chosen via cross-validation.

Graphically this extension to non-separable data is visible in Figure 3, where in both plots, the same non-separable dataset is displayed. The margin violations are the support vectors that are connected to the black dotted lines with a light green line. In the left plot, there are fewer margin violations as the budget is lower (the constant governing aversion is higher). Therefore also, the margin distance is smaller compared to the right plot.
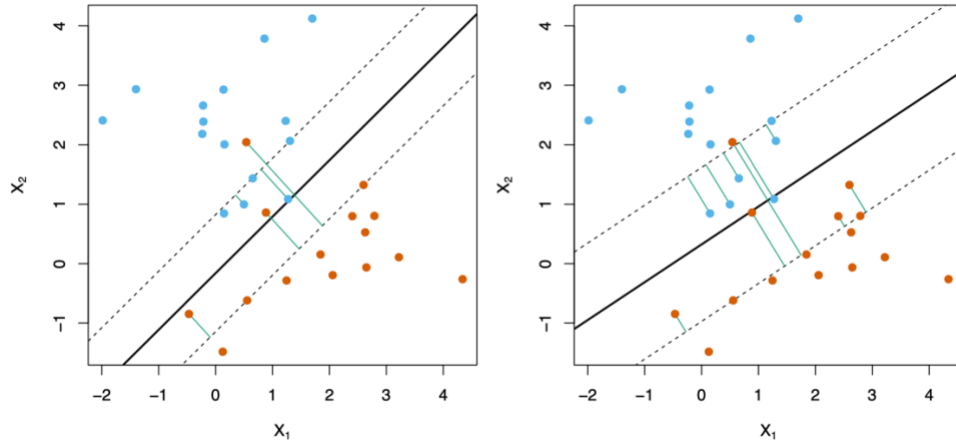
Figure 3: Hyperplanes for a non-separable dataset with a low budget (left) and a high budget (right) (Efron & Hastie, 2021).

It is evident that the hyperplane becomes more robust with a larger budget, which means it will be less affected by changes in the dataset (Efron & Hastie, 2021). Still, a budget too big will result in many misclassifications, so a trade-off between robustness and overfitting to the training set is essential. This is achieved by choosing the optimal constant governing aversion parameter C (also known as cost) via cross-validation to achieve the abovementioned goal of maximum forecast accuracy.

3.3.1 Kernel functions

As with other linear methods, it is possible to introduce nonlinearities into support vector machines. Instead of entering basis expansions, like polynomials or non-linear basis functions, support vector machines use kernel transformations to introduce nonlinearity (Deisenroth, Faisal & Ong, 2020). The abovementioned optimization problem can be rewritten so that predictors enter as an inner product. With that in mind, such regressors, entering as an inner product, are replaced by kernel functions, which measure the similarity between the two observations that enter the kernel function as an inner product (Deisenroth, Faisal & Ong, 2020). The kernel functions alter the feature space into a higher dimension, in which the margin maximizing hyperplane is then to be found. This optimal hyperplane through the kernel function altered feature space will still be a straight line, but if transformed back into the original feature space will look like a curve, thus accounting for nonlinearities. We will deploy three different kernel functions, each as its own model, to see which one performs best for our task at hand.

Linear Kernel:

A linear kernel is a kernel function that does not introduce nonlinearities as the inner product that enters the kernel is unchanged. Therefore, running a support vector machine with a linear

kernel will not change the result from running a support vector machine without a kernel. The linear kernel is defined in equation (5):

$$K(x, x') = x^T x' \tag{5}$$

Polynomial Kernel:

A (dth-degree) polynomial kernel is a kernel function that works similar to introducing polynomials as a basis function into the features, making the feature space non-linear. The dth-degree polynomial kernel is defined in equation (6):

$$K(x, x') = (1 + s\langle x, x'\rangle)^d \tag{6}$$

The hyperparameter d in equation (6) is the optimal degree of the polynomial. The scale hyperparameter s, is used to normalize patterns in the dataset to ensure data is on the same scale. The optimal value for both hyperparameters is found by cross-validation, maximizing a chosen performance measure, in our case, accuracy.

Gaussian Radial Basis:

The current standard for introducing nonlinearities into support vector machines is the radial kernel (RBF) (Efron & Hastie, 2021). The radial kernel is defined in equation (7).

$$K(x, x') = exp\left(-\gamma ||x - x'||^2\right) \tag{7}$$

The term $||x - x'||^2$ is known as the squared Euclidean distance. The hyperparameter γ, which is the kernel scale parameter, is again chosen via cross-validation (Hastie, Tibshirani & Friedman, 2009). This scale parameter can be further parameterized by defining it in equation (8).

$$\gamma = \frac{1}{2\sigma^2} \tag{8}$$

With that, the hyperparameter of the RBF changes to σ, the sigma parameter.

3.3.2 Penalization

Penalization is a method to avoid overfitting by adding a penalty term to a regression, which will pull the coefficients towards zero. The two most common penalization methods for a regression problem are ridge, where a quadratic penalty term is added, and lasso, where an absolute penalty term is added. In ridge regression, the parameters will never be zero, whereas lasso will produce a sparse solution, which can act as a variable selection method. The higher the hyperparameter λ of such a penalty term, the higher the effect of the penalty term and, therefore, the bigger the shrinkage of the coefficients towards zero.

A similar approach can be implemented into support vector machines, where again, a penalty term is included in the optimization problem, shrinking the direction vector of the hyperplane towards zero (Hastie, Tibshirani & Friedman, 2009). This penalty term is included in the minimization of loss, and the penalty term takes the form of the L2 penalty term and thus is quadratic in nature. The main difference between support vector machines with penalization and the earlier mentioned ridge and lasso regularization is the type of loss function. The support vector machine uses the hinge loss function, which estimates the classifier $G(x)$ directly, whereas ridge and lasso use squared loss functions for linear regression problems or the binomial log-likelihood for logistic regression problems (Hastie, Tibshirani & Friedman, 2009). The abovementioned minimization of loss for the support vector machine is displayed in equation (9).

$$\min_{\beta_0, \beta} \sum_{i=1}^{N} [1 - y_i(x^T\beta + \beta_0)]_+ + \lambda \|\beta\|_2^2 \qquad (9)$$

The optimal value of the hyperparameter λ (called "tau" in R) in equation (9) is chosen via cross-validation to maximize accuracy. This hyperparameter replaces the constant governing aversion parameter C, in the models where a penalty term is added. With the addition of this penalty term, we will test three more models, one for each different kernel type, leaving us with six different support vector machines.

In our framework, the six models and their according hyperparameters are optimized to maximize each model's forecast accuracy. For the SVM with a linear kernel, the cost C is optimized between the values 0.1 and 2, with a step size of 0.1. For the SVM with a linear kernel and penalization, tau is optimized between the values 0.1 and 2, with a step size of 0.1. In the SVM with a polynomial kernel, the cost is optimized between 0 and 2, with a step size of 0.2; the scale between 0.1 and 1, with a step size of 0.1 and the degree between 2 and 4, with a step size of 1. For the SVM with a polynomial kernel with penalization, tau is optimized for values 0.1, 0.4 and 0.7; scale from 0.1 to 0.5 with a step size of 0.1 and degree between 2 and 4 with step size 1. For the SVM with a radial basis function, C is optimized from 0.1 to 1.7 with a step size of 0.1 and sigma from 0.1 to 0.5 with a step size of 0.1. Lastly, for the SVM with a radial basis function and penalization, tau is optimized for the values 0, 0.3 and 0.7, and sigma is optimized for the values 0.1, 0.4 and 0.7.

3.4 Model Evaluation

To help us evaluate the model performance, we include a model built on econometric theory rather than machine learning, namely the probit model. The reason for this is to evaluate how

much we can increase the accuracy by using machine learning compared to a benchmark, in our case, the more traditional probit model.

The probit model predicts a binary outcome by deploying a dummy variable as a dependent variable. In the simplest binary choice model, the linear probability model (LPM), it is assumed that there is a linear relationship between explanatory variables and the estimated probability of event 0 or 1 happening. Because of that linear relationship, the LPM does not restrict the estimated probability to lie between zero and one, which can result in severe limitations (Brooks, 2008). The probit model solves these limitations by transforming the regression problem with the cumulative normal distribution to ensure that the estimated probabilities lie between 0 and 1 (Brooks, 2008). To estimate this non-linear regression problem, maximum likelihood is used.

3.5 Variable Importance

Variable importance is used to understand which variables have the most considerable effect on a model. Understanding which variables are the most important for the prediction can give valuable insights into which variables to choose when training the model. With a random forest, variable importance is constructed by first taking the test accuracy from the bth tree, then randomly permuting the values of the relevant variable in the test data and comparing this accuracy with the previous one. The change in accuracy is calculated for all trees, and the average is the variable importance. Unfortunately, this is not as simple for a support vector machine. Instead, we can take advantage of the receiver operating characteristic (ROC) curve. For a binary classification like ours, a sequence of cutoffs is used on the data for prediction. For each cutoff, the sensitivity and specificity are calculated, and the area under the ROC curve is obtained. The area under the ROC curve is used as a measure of variable importance (Kuhn, 2008).

3.6 Portfolio Construction

To further implement our prediction of the weekly changes in the VIX, we build four different portfolios and compare their results. Since our last prediction predicts the VIX's movement for a date not included in our timeframe, we drop the last prediction when constructing the portfolios. Therefore, all portfolios are built on 863 instead of 864 observations.

The first portfolio, a standard 60% equity and 40% fixed income portfolio (Rekenthaler, 2022), is one benchmark we aim to beat. The second portfolio we build is a signal-based portfolio that incorporates our VIX forecast from the best-performing ML algorithm. In this portfolio, we invest in our equity index if the forecasted VIX decreases and into fixed income if the forecasted

VIX increases. The third portfolio is equivalent to the second portfolio, with the difference that the forecast of the VIX stems from the probit model. We include this third portfolio to see whether the machine learning approach has an advantage over the traditional econometric theory approach regarding portfolio performance. To confirm that the signal-based approach can work, we also test it with the observed directions of the movement of the VIX, as then we should see a very high Sharpe ratio.

To further investigate the difference between the portfolios, we decided to split up each portfolio further, based on the level of the one week lagged VIX (the time the prediction is made). We hope to gain further insight into whether a signal-based trading strategy works better for specific market regimes than others. We split the VIX level into three categories when the VIX is below 20, the VIX is between 20 and 30, and the VIX is above 30.

A risk-return measure is used to compare all our different portfolios, namely the Sharpe ratio (Sharpe, 1966). The one-period Sharpe ratio, defined in equation (10), measures the excess return per unit of risk for a given portfolio i.

$$SR_i = \frac{E[R_i] - R_f}{\sigma_i} \tag{10}$$

where $E[R_i]$ is the expected return of portfolio i, R_f is the risk-free rate, in our case the one-month Treasury bill, transformed to a weekly rate and σ_i the standard deviation of the returns of portfolio i.

As it is common practice to compare Sharpe ratios on a yearly basis, we multiply the Sharpe ratio obtained using weekly data by the square root of 52 to obtain a yearly Sharpe ratio.

4. Data

The following section covers the selection of relevant data, giving insights into the data properties and data sources for our prediction. Furthermore, we present the motivation behind the variable selection to allow the reader to follow our line of thought. Then, the necessary data transformations to fit the data to the model and solve problems like outliers and non-stationarity are displayed. Lastly, we cover the data used for the portfolio construction and computation of Sharpe ratios.

4.1 Prediction Data

The data used for the prediction consist of different equity indices, commodities, and macroeconomic variables. The equities and commodities are all constructed using future data, which is automatically back adjusted since a moving average over future contracts is used by Bloomberg (Roll Method BBG Bloomberg default) and FactSet. For our macroeconomic variables, we argue on behalf of market expectation. We assume that the market has an expectation of the value; therefore, using revised macroeconomic data should work as reliable predictors. We make sure not to have any missing values within our dataset, as that can have a negative effect on our prediction. Therefore, we decide to put our timeframe to 02.04.2004 – 01.10.2021, as within that time frame we get a complete dataset for all our predictors and the dependent variable.

An overview of the different variables used, in which frequency they are obtained and where the data is collected from is displayed in APPENDIX A.

In the following section, we cover the different predictors in detail. Starting with the different equity indices, we mainly focus on U.S. indices since the VIX measures the volatility of the U.S. market. Therefore, we include the S&P 500 Index and the NASDAQ 100, two of the most prominent U.S. based indices. We furthermore include the lagged weekly change of the VIX and a VIX future contract, as we suspect both might be good indicators of the VIX movement in the following week. As the VIX is a measure of the volatility in the U.S. market, we include the U.S. Dollar Index to capture the general value of the U.S. Dollar compared to other general currencies.

Secondly, we test different commodities since they often have unique characteristics. Oil, wheat, and corn are strategic resources that can capture specific geopolitical movements as they are not found/produced everywhere in the world. With the Russian-Ukrainian war, we currently see how much of an influence the supply of wheat and oil have on the economy and how dependent other countries are on imports of those resources. Therefore, we believe it is crucial

to include the impact of such commodities on the VIX movement. Other commodities we include are gold and copper, as both are base metals and can act as a hedge against inflation. Additionally, we include the U.S. Treasury Bond and a 10-year U.S. Treasury Note, both fixed income securities issued by the U.S. government. Therefore, we believe it is interesting to see how these might influence the volatility, and thus the VIX, in the U.S. market.

Lastly, we choose to add multiple macroeconomic variables into our prediction, as they can capture the current state of the economy. We hereby rely on widespread variables like Inflation, Federal Funds Rate (Interest Rate), GDP growth, and the Industrial Production Index. To directly account for the unemployment rate, one of the essential variables to capture current economic movements, we include the U.S. Insurance Claims. Having a direct measure of changes in the U.S. labor market gives us further information about the U.S. economic growth, which we believe to have a significant impact on the volatility of the U.S. stock market and thus our proxy, the VIX.

Since we are using almost 17 years of data, we find it reasonable to obtain all the data of indices and commodities weekly to avoid high computational costs that increase with frequency. The VIX itself is obtained daily and then transformed into weekly data by taking the average over the trading days from Monday to Friday. By obtaining daily data, we ensure outlier resistance, as taking the average is a lot more robust than looking at weekly data, which is simply Friday's level of the asset in question. Macroeconomic variables, like the 10-Year Breakeven Inflation Rate and the Federal Funds Effective Rate, are also obtained daily, the latter even on weekends. As in the transformation of the VIX, an average is taken; either over all trading days for the 10-Year Breakeven Inflation Rate or over all weekdays for the Federal Funds Effective Rate. One of our macroeconomic variables, namely the Initial Claims, is obtained weekly, making it easy to implement. The Industrial Production Index is obtained monthly; therefore, we calculate the percentage change (further explained in section 4.1.1 Data Transformation) between every two months and break that down to a weekly change by dividing by 4.33. With that change, the monthly data is transformed into weekly data, with weeks within the same month having the same percentage change. The annual rate of change in the GDP, obtained quarterly, must also be transformed into a weekly rate. To do so, we divide the annual rate by 52 to approximate the weekly rate. Therefore, each week within the same quarter will have the same weekly rate.

4.1.1 Data Transformation

Two things need to be discussed in order to decide which data transformations to use. First, how we can transform the variables to fit the data to the model. Second, which transformations are necessary to solve problems such as extreme values and non-stationary series.

This thesis aims to predict the movement of the VIX in the following week. We use 17 variables that we believe have a valuable effect on how the VIX moves. All variables except the GDP growth and inflation are given in real numbers or indices, resulting in a level. The level of these variables would be interesting if our goal were to predict the level of VIX, but as it is not, we will use transformations to make the variables more relevant for the sign prediction. As the sign of the VIX is calculated through the first difference, it makes sense to transform the other variables similarly since we use the first difference of the VIX to predict the direction of its movements.

As we work with time series data, we also have to check for non-stationarity. Using non-stationary variables can lead to spurious regressions, meaning that two variables seem to be dependent on each other while in reality, it is a third unnoticed aspect, such as time, making them move in the same direction (Brooks, 2008). To check for non-stationarity, the augmented Dickey-Fuller test is used on the non-transformed variables. The null hypothesis states that the data includes a unit root and therefore is non-stationary (Maechler, Rousseeuw, Struyf, Hubert & Hornik, 2013). The test is done for a significance level of 0.05, resulting in the following variables showing non-stationarity: CL1, C 1, W 1, US1, TY1, GC1, HG1, ES1, NQ1, and Industrial Production. To handle these variables correctly, we need to evaluate which type of non-stationarity they are characterized by; we do this by plotting the data (APPENDIX B). The plots show that all non-stationary variables include a random walk, and some variables also include a drift term. Different types of non-stationarity require different solutions; using the wrong transformation will lead to additional errors in the model. For a random walk, both with and without a drift term, the first difference is used (Wooldridge, 2013).

The transformation for fitting the variables to the model and to avoid the problems of non-stationarity is unanimous; therefore, we transform all variables except the GDP growth and inflation into the first difference. To ensure the effectiveness of our transformations, we do the augmented Dickey-Fuller test on the transformed variables, which all are stationary.

As the variables used for our prediction are on different scales, it is necessary to transform the abovementioned variables into percentage changes, giving them the same scale. This is done using the following equation (11), where the percentage change is defined as the first difference divided by the previous week's value.

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} \tag{11}$$

where r_t is the percentage change from period $t-1$ to period t, p_t is the value or price in period t and p_{t-1} is the value or price in period $t-1$.

Without this transformation, the machine learning algorithms would assign a higher impact weight to the variables whose first differences are of larger values; the models would falsely believe some variables have a high impact based on high first difference values, resulting in training the model into false patterns. As GDP growth and inflation are already defined as percentage differences, we do not need to transform these further. Note that for Industrial Production, the transformation into percentage change is done before making the division into weeks, earlier described in the data chapter.

4.2 Portfolio Data

As the VIX is a volatility measure for the U.S., we construct all portfolios using assets from the U.S. As an equity index we use the ES1 contract, an S&P 500 future. We do this as we cannot buy the S&P 500 directly, meaning we have to invest in an ETF of the S&P 500 instead. We use the TY1 for the fixed income, a future contract of the 10-year U.S. Treasury note that is again easily tradable. To compute the Sharpe ratio of each portfolio, we need to use the risk-free rate. Since we are using weekly data, we also need to obtain a weekly risk-free rate, which can easily be collected from the Kenneth R. French – Data Library. This weekly risk-free rate stems from the one-month Treasury bill, which acts as a proxy for the monthly risk-free rate, transformed into a weekly rate.

5 Empirical Results and Discussion

In the following section, we present the results of our sign prediction as well as the portfolio construction. The result of the sign prediction is divided into model accuracy, confusion matrices, and variable importance; each part is further analyzed and discussed. We display the Sharpe ratios of the different portfolios for three market regimes, discussing the performance differences of each of the portfolios. Additionally, limitations regarding data, the optimization of parameters, the prediction, and the portfolio construction are mentioned. We conclude this section by giving an outlook on possible future research.

5.1 Sign Prediction Results

Eight different models are used to predict the sign of the VIX in the following week: Probit, Random Forest, Linear SVM without penalization, Linear SVM with penalization, SVM with the polynomial kernel, SVM with polynomial kernel and penalization, SVM with radial basis function, and SVM with radial basis function and penalization. The optimal parameters and the maximized accuracy are shown in Table 1.

Table 1: Models with optimized parameters and accuracy.

The data consist of 914 different weekly observations for all our variables. The first 50 observations from the 02.04.04 - 11.03.05 are purely used for training, and predictions are obtained for the 864 last observations. With these predictions, the forecast accuracy for each model can be computed and is displayed under the column "accuracy". The different hyperparameters, which are optimized to obtain the maximum forecast accuracy, are given in the first six columns. Their optimal values are displayed for each model, or if not used in the model, a hyphen is shown.

Model	tau	C	mtry	sigma	scale	degree	accuracy
Probit	-	-	-	-	-	-	56.9%
Random Forest	-	-	3	-	-	-	57.1%
SVM - linear kernel	-	0.3	-	-	-	-	58.4%
SVM – linear kernel with penalization	0.9	-	-	-	-	-	54.7%
SVM – polynomial kernel	-	0.8	-	-	0.1	3	57.0%
SVM – polynomial kernel with penalization	0.1	-	-	-	0.3	3	55.4%

SVM – RBF kernel	-	1.5	-	0.1	-	-	56.4%
SVM – RBF kernel with penalization	0.3	-	-	0.4	-	-	54.1%

After a look at the obtained results, it becomes clear that the best performing model is the SVM with a linear kernel, as it achieves 58.4% accuracy. The random forest and SVM with a polynomial kernel are close in performance with 57.1% respective 57.0% accuracy. Worst off are the linear SVM with penalization and the SVM with radial basis function and penalization, resulting in 54.7% respective 54.1% accuracy. The probit model results in 56.9% accuracy, 1.5 percentage units lower than our best model. Having a traditional econometrics model performing this close to our best machine learning approach questions the relevance of machine learning within this field.

For further evaluation, we compare our results with previous studies; Nossman and Wilhelmsson (2009) predict the sign of the VIX 5 days ahead with 63% accuracy. Ahoniemi (2007) predicts the sign of the VIX 1 day ahead with 57.2% accuracy. These studies indicate that our results are within reasonable limits and that it is rather difficult to predict the sign of the VIX.

To continue evaluating each model's performance, we use confusion matrices displayed in Figures 4 to 11. The prediction shows on the left side, and the actual value on the bottom. The value 1 indicates a negative sign of VIX movement, meaning that the index will decrease in the following week. In contrast, the value 2 indicates that the VIX will increase in the following week. The color scale goes from pink for low values to blue for high values.
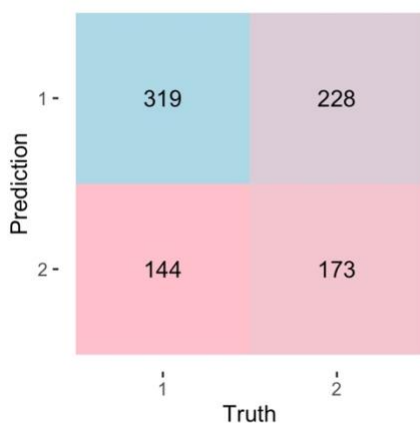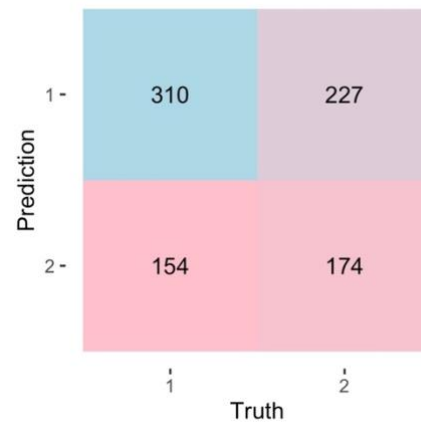

Figure 4: Confusion matrix probit.

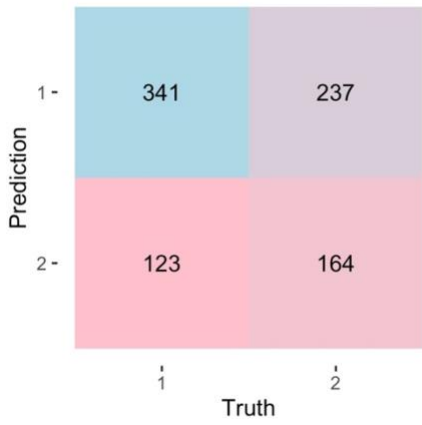Figure 5: Confusion matrix random forest.

Figure 6: Confusion matrix SVM with a linear kernel.

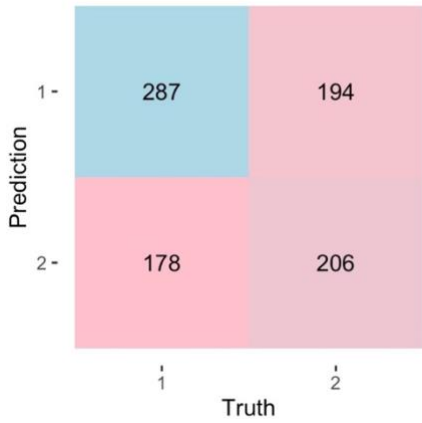Figure 7: Confusion matrix SVM with a linear kernel and penalization.

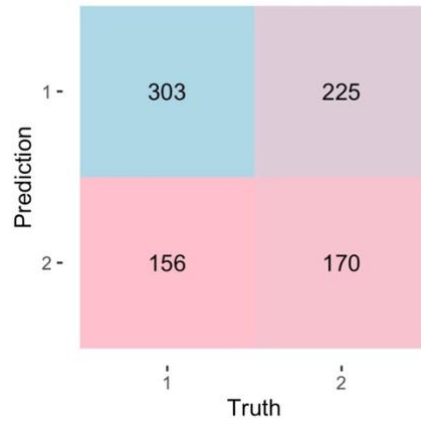Figure 8: Confusion matrix SVM with a polynomial kernel.

Figure 9: Confusion matrix SVM with a polynomial kernel and penalization.
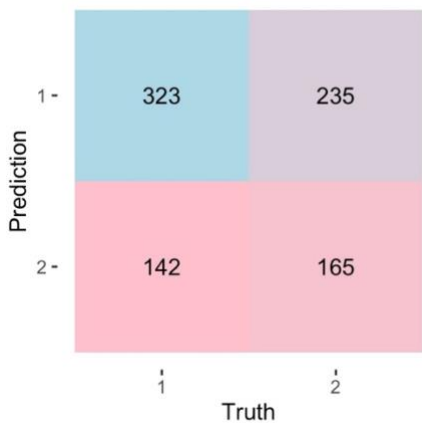

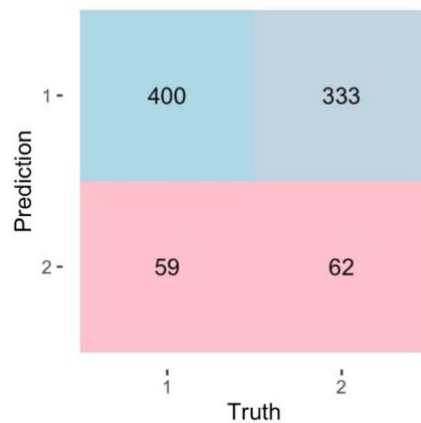Figure 10: Confusion matrix SVM with a radial basis function.

Figure 11: Confusion matrix SVM with a radial basis function and penalization.

We can see that all models show similarities in the pattern of the confusion matrix. Mainly, all of them predict the VIX to go down in the following week (value 1) more often than the

opposite; thus, the likelihood of a type I error is relatively high. For the predictions of our best performing model, the SVM with a linear kernel (Figure 6), 66.8% percent of the test data is predicted to be in class 1. The SVM with RBF and Penalization (Figure 11) is even worse, with 84.7% of the data being predicted to be in class 1. Finding reasoning as to why this is the case will be done in the following paragraph.

A possible explanation could be a class imbalance problem in our data set. However, our data set is relatively balanced, with roughly 53.6% of the observations classified as class 1 and 46.4% of the test data set classified as class 2. Therefore, the possibility of a class imbalance problem is negligible. Another option would be that we are overfitting, but since we use time series cross-validation, which typically is a technique for preventing overfitting, this should not be a problem. We choose to leave the investigation explaining this problem for future research, while coming back to the influence of this problem when we discuss the limitations of our portfolio construction.

Revisiting the best performing model and assuming a normal distribution of the movement in the VIX, the model performs with 8.4 percentage units higher accuracy than random guessing. The result is an interesting finding as it indicates that the relationship between our two classes, the VIX moving up or down the following week and our predictors, is best modeled as linear. This could again be explained by the leverage effect, where we see an inverse relationship between asset returns, which are part of our predictors, and asset volatility, which is captured in our dependent variable. Since not all variables are assets/stocks, we will investigate how each variable affects the quality of the predictions. To do so, we examine the variable importance of the best model, SVM with a linear kernel (Figure 12).
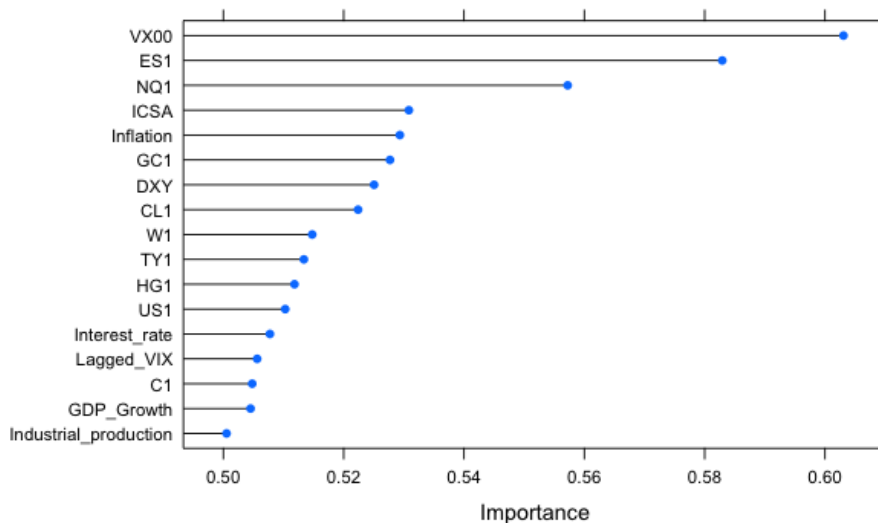

Figure 12: Variable importance plot of SVM with a linear kernel.

24

As earlier described, the variable importance for the SVM is obtained by the area under the ROC curve. The variables with the most significant effect on the prediction are the percentage difference of VX00, ES1, and NQ1. It is not surprising that VX00, which is the future contract of the VIX, has a significant impact on the prediction, as this implies that the level of VIX will increase or decrease in the future. This indicates that the VIX futures market, to some degree, is able to predict the VIX. However, knowing that the VIX will increase in the future will not automatically give good predictions of the sign of movement in the next week, as the VIX fluctuates a lot and therefore can go both down and up before the contract ends. As the ES1 is obtained by E-mini S&P futures and the VIX by options on the S&P 500 Index, we expect the variables to correlate. Furthermore, the NQ1 is constructed by the NASDAQ 100 E-mini futures, which partly overlap with the S&P 500; with this in mind, we believe both the ES1 and the NQ1 to have a significant impact on our prediction. Additionally, the weekly claims – ICSA – has a relatively high impact on the model, which could be explained by the fact that job losses directly impact the U.S. economy and, therefore, can affect the stock market's volatility. With this result, it would make sense for Industrial Production and GDP growth, which are direct measures of the U.S. economy, to have a similar impact; however, this is not true. One reason behind this is the frequency of observations on different variables; as previously explained in Data Sources, Industrial Production and GDP growth are given monthly respectively quarterly, and as we assume equal change for each week within that time, the variables have the same value for 4.33 respective 12 weeks. When having the same value for multiple weeks in a row, it is hard for the model to find a pattern, and the variables are given a lower impact. Other variables with low impact on prediction are the Interest Rate, the lagged VIX, and C 1. The Interest Rate has a negligible variation over time. Therefore, even when taking the percentage difference, it has minimal values, which could lead to the variable not having a significant impact on the prediction. The level of lagged VIX has in previous research shown a good impact on predicting the level of the VIX, which makes sense as the VIX this week most likely is similar to the VIX next week. However, in this thesis, we use the sign of the lagged VIX movements to predict the sign of the future VIX movements. As the VIX fluctuates a lot, the sign of the lagged VIX movements does not tell much of the sign of next week's movements. The VIX will often keep a stable level, but the slight movements up and down are hard to predict even with last week's information. We included the C 1, which is the corn future, as we thought it would capture movements in the economy. However, both C 1 and W 1, which is the wheat future, perform poorly. The crude oil future, CL1, has a more considerable impact, but still not a specifically good one. Therefore, for our chosen model, the specific food and oil commodities do not significantly impact the predictions. Gold is the commodity that impacts

the prediction the most; this could be because gold is seen as a safe asset that is often bought when the economy is unstable and therefore has a more direct impact on the economy and the movement of the VIX.

With the results of variable importance, a simpler model could be suggested. For example, only using the ten most important variables and deleting the other seven ones, resulting in a simpler model but with 54.2% accuracy, 4.2 percentage units lower than when including all variables. The decrease in accuracy has to be weighed against computational costs gained and time and money spent collecting the last eight variables. As we aim to forecast with the highest accuracy possible, simplifying the model is not beneficial in our case.

5.2 Portfolio Results

The results from the portfolio building can be seen in Table 2. We start by analyzing the three first models: 60% equity and 40% fixed income, as well as the signal-based portfolios based on the predictions of the SVM with a linear kernel and the probit model. Later we discuss the effect of using a signal-based strategy with observed values instead of the predictions.

Table 2: Sharpe-Ratios for the different portfolios at different levels of the VIX.

The table displays Sharpe ratios for different portfolios and levels of the VIX. The sample consists of weekly trading during the period 18.03.05 - 01.10.21. The Sharpe ratio is constructed from a signal-based strategy that invests in the equity index ES1 if our forecast of the VIX decreases and in the fixed income security TY1 when our forecast of the VIX increases. The risk-free rate stems from the one-month Treasure bill transformed into a weekly rate. The level of the VIX is the one-week lagged VIX (available at the time the prediction is made). The "60/40" is a standard 60% equity and 40% fixed income portfolio. "SVM – linear" is a support vector machine with a linear kernel. "Probit" is an econometric probit model. "Observed values" is a model with perfect prediction.

Portfolio	All levels of VIX	VIX < 20	$20 \leq$ VIX < 30	VIX $\geq$ 30
60/40	0.479	0.416	1.236	0.037
SVM – linear	0.425	0.406	1.312	-0.827
Probit	0.360	0.522	0.821	-0.729
Observed values	3.04	3.485	3.743	2.536

Looking at all levels of the VIX, the 60% equity and 40% fixed income portfolio gives a Sharpe ratio of 0.470, outperforming both the portfolio built on the linear SVM and the portfolio built on the probit model. Given that our goal is to outperform such a 60/40 portfolio, we can see that either the signal-based portfolio does not work and the sign of the VIX is not a good indicator for investment decisions or that our predictions are not accurate enough. To test for that, we rerun our signal-based portfolio with the observed values, which equals predicting with

an accuracy of 100% and obtain a Sharpe ratio of 3.04. This result indicates that the signal-based portfolio is a valid approach to achieve excess returns as using the observed values, we outperform the 60/40 portfolio regarding the Sharpe ratio. Therefore, if our models were more accurate, the sign-based portfolio-building strategy could generate Sharpe ratios high enough to beat the traditional 60/40 portfolio.

Nevertheless, we investigate whether there is a scenario in which our models can beat the standard 60/40 portfolio even without a better accuracy. To do so, the portfolios are compared under different market regimes and thus in between different VIX levels. For a VIX level below 20, the portfolio built on the probit model gives the highest Sharpe ratio with a value of 0.522, above the Sharpe ratios of the 60/40 portfolio and the portfolio built on the SVM with a linear kernel with Sharpe ratios of 0.479 and 0.425, respectively. Only looking at the level of the VIX between 20 and 30, the portfolio based on the linear SVM has the highest Sharpe ratio with a value of 1.312. The 60/40 portfolio yields a Sharpe ratio of 1.236 for that range, with the portfolio based on the probit model far behind with a Sharpe ratio of 0.821. For a VIX level above 30, the 60/40 portfolio generates the highest Sharpe ratio with a value of 0.037, whereas both signal-based portfolios have negative Sharpe ratios. This suggests that either our predictions for the high market risk regime are worse compared to the other regimes or that our trading strategy performs worse in such a high volatility environment. For a detailed discussion, the sign predictions needed to be split up into the same market regimes to account for possible accuracy, and thus performance, differences in the prediction. Due to limited time, we could not perform this task but instead, gain more insights about the performance from a portfolio built on perfect predictions later on. Generally, we can see that the standard 60/40 portfolio will be the best when it is impossible to divide the VIX into different levels. However, investors who have the possibility to form different portfolios based on the level of the VIX can increase their profitability using this method. 587 of the 863 predictions used have a level of VIX beneath 20, meaning that in a method where the levels of the VIX are divided, the signal-based portfolio based on the probit model should be used most often. For the 197 predictions having a level of the VIX between 20 and 30, the linear SVM should be used. For the 79 remaining predictions where the VIX is above 30 and the market thus classified as risky, the 60/40 portfolio should be used. This result contradicts the results from not dividing the VIX into levels, resulting in only using the 60/40 portfolio.

To get a general idea of how the sign-based portfolio selection works when splitting the data into different market regimes, we also split up the signal-based portfolio using observed values. Using observed values is the same as perfect prediction, implying that we can test the trading

strategy independent of different models' accuracy within the different regimes. We get high Sharpe ratio values in all three market regimes, with the one where the VIX is above 30 being the least successful with a Sharpe ratio of 2.536, while in the low and medium risk regime, Sharpe ratios of 3.485 respectively, 3.743 are reached. These results confirm our earlier findings that the signal-based strategy performs worst in high-risk scenarios and, to some extent, gives insight into the cause of that gap in performance between different regimes.

5.3 Limitations and Future Research

Whenever one uses a machine learning algorithm, the quality and quantity of data plays a crucial role in the success of the model deployed. In our case, we have a trade-off between the length of the data series and the number of predictors, as many predictors we thought about using are not available for a long enough timeframe. To overcome this, we decide to stick with the predictors that appeared to be of the utmost importance to us while still maintaining a relatively high number of observations. With more time, this trade-off could be looked at in more detail, and a possibly better choice of variables could have been made. Adding to that, we also acknowledge the possibility of us missing important predictors, which might have further improved our obtained accuracy.

Achieving high accuracy in the forecast of the movement of the VIX is one of the two goals of this thesis. Model-specific hyperparameters greatly influence the obtained accuracy, as they tweak the model to fit the data better. We optimize these hyperparameters to achieve maximum forecast accuracy when training our seven machine learning models in question. The optimization is done using a grid for each hyperparameter, meaning the more refined the grid and the more hyperparameters there are, the larger the number of possible combinations of hyperparameters. Since we have no access to cloud computing services, we cannot run very fine grids of hyperparameters as even a few combinations can already take multiple hours to run for just a single model. Therefore, we acknowledge that our optimal parameter choice might not optimize the models to the best degree possible, meaning a higher accuracy could have been possible with more computational power at hand.

As previously mentioned, the imbalance in the predictions is something we do not find an explanation for with limited time at hand. It still is important to keep this problem in mind as it might severely influence the success of the signal-based portfolio strategy. This is due to the framework in which we create our trading strategy, namely, with the predicted VIX decreasing, we invest in equities, and with the predicted VIX increasing, we invest in fixed income. Overpredicting one class over the other might significantly alter the Sharpe ratios found and raise questions about whether overall accuracy is the best metric to optimize or whether

sensitivity or specificity should be considered. The abovementioned fact might be the cause of another limitation that we acknowledge, namely the relatively poor performance of the signal-based portfolio on a general level using our machine learning forecasts.

Despite these shortcomings, we still believe that the insights are valuable and can only be improved by future research. It would be interesting to see how well the chosen models can perform simply by finding better optimized hyperparameters. Furthermore, the question arises whether more sophisticated models like a neural net or a language-based model could improve forecasting accuracy, replacing the model suggestion of us. In our opinion, one of the most relevant points for future research is to discover the breakeven point between the standard 60/40 portfolio and our signal-based portfolio, displaying the needed level of accuracy for the signal-based portfolio to be at least as good as the 60/40 portfolio. That would allow investors to take advantage of previous research and models as it directly shows whether these insights could be exploited to ensure excess returns. Other possibilities for future research are different trading strategies for the signal-based portfolios, whether it is to test the hypothesis with different equity indices and fixed incomes, allowing for different strategies, including short selling, or applying the framework to a different country.

6 Conclusion

Out of our eight models, a support vector machine with a linear kernel gives us the best results in terms of forecast accuracy of the movement of the VIX. Nevertheless, the difference between the accuracy of our chosen model (58.4%) and a simple probit model (56.9%) is not as big as we would have suspected it to be. Classifying the direction of the VIX correctly in more than 50% of the cases means we are better than under a random walk hypothesis, therefore justifying the use of our machine learning model. Looking at the confusion matrices of our eight models, we see a clear pattern that all of them predict the following week's VIX to decrease much more often than the opposite. This might have severe implications for our signal-based portfolio, as with a decreasing VIX forecast, the portfolio will invest in equities. Given the same accuracy and most predictions to predict the following week's VIX to increase, we would have a significantly different portfolio outcome. At the current stage, we have no clear explanation as to why all our models predict in such a way, but we firmly believe that there are important insights to be found if this problem were to be further investigated. The variable importance plot of the SVM with a linear kernel gives insights into the effect single predictors have on the performance of our prediction model. VIX futures and U.S. equity indices have the most significant impact on the VIX, which seems logical as the VIX is constructed from options of the S&P 500, one of the most extensive U.S. equity indices. Macroeconomic variables, which are not available on a frequent basis, force us to transform monthly or quarterly rates into weekly rates, which to no surprise, do not perform well, as many of them are constant over multiple weeks. Thus, such macroeconomic variables have no significant impact on the prediction accuracy. To conclude the prediction of next week's VIX movement, we believe that our model offers a good trade-off between complexity and applicability, given the time constraints. Comparing these results to other researchers, we believe that we are at a reasonable level of one-week ahead forecast accuracy of the VIX.

Incorporating the forecast of the movement of the VIX into our portfolio building yields some interesting insights. For one, we confirm the inverse relationship between the VIX movement and returns; as with observed values, our signal-based portfolio yields a very high Sharpe ratio. Knowing this relationship holds lets us believe that we should be able to beat the standard 60/40 portfolio with our predictions. Unfortunately, we could not do so, which shows that our prediction accuracy is not high enough. Nevertheless, when further splitting the portfolio into three subparts differing in lagged levels of the VIX and thus in market riskiness, we can outperform the 60/40 portfolio under certain conditions. The signal-based trading strategy using the probit model outperforms the 60/40 portfolio when the level of VIX is below 20, while the

one using the SVM with a linear kernel outperforms the 60/40 portfolio for a VIX level between 20 and 30. For a VIX level higher than 30, the 60/40 portfolio prevails, being the only one to display a positive Sharpe ratio. Combined with the results of the portfolio built on perfect predictions, we find that our trading strategy performs worse during high-risk regimes.

With these insights, we believe that an investor can take advantage under the right circumstances, allowing him for a substantial higher Sharpe ratio in specific market regimes than with a 60/40 portfolio. Despite that, we note that while we still neglect all transaction costs, the effort put in might not be worth pursuing for monetary gain.

References

Aboura, S. (2003). International transmission of volatility: A study on the volatility indexes VX1, VDAX and VIX, working paper, ESSEC Business School. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=514282 [Accessed 30 March 2022]

Anhoniemi, K. (2007). Modeling and Forecasting the VIX Index, working paper, no.129, Helsinki School of Economics; Helsinki Center of Economic Research (HECER). Available online: https://helda.helsinki.fi/bitstream/handle/10138/16698/modeling.pdf?sequence=1 [Accessed 30 March 2022]

Baltagi, H. B. (2011). *Econometrics*, vol. 5, p.383, Springer Heidelberg Dordrecht London New York.

Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments, *Journal of Financial Economics*, vol. 116, no. 1, pp.111-120. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Bekaert, G., & Hoerova, M. (2014). The VIX, the variance premium and stock market volatility, *Journal of econometrics*, vol. 183, no.2, pp.181-192. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation, *Information Sciences*, vol. 191, pp.192-213. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 28 April 2022]

Black, F. (1976). Studies of stock market volatility changes. 1976 Proceedings of the American statistical association business and economic statistics section, *American Statistical Association*, pp.177-181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of econometrics*, vol. 31, no. 3, pp.307-327. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers, *In Proceedings of the fifth annual workshop on Computational learning*

theory, pp.144-152. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Britten-Jones, M., & Neuberger, A. (2000). Option prices, implied price processes, and stochastic volatility, *The Journal of Finance*, vol. 55, no. 2, pp.839-866. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 7 April 2022]

Brooks, C. (2008). *Introductory Econometrics for Finance,* vol. 2, pp.319-322, 512-518, Cambridge University Press.

Carr, P., & Lee, R. (2009). Volatility derivatives, *Annual Review of Financial Economics*, vol. 1, pp.319-339. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects, *Journal of financial Economics*, vol. 10, no. 4, pp.407-432. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Christoffersen, P.F., & Diebold, F.X. (2000). How relevant is volatility forecasting for financial risk management?, *Review of Economics and Statistics*, vol. 82, no. 1, pp.12-22. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Clements, A.C., & Fuller, J. (2012). Forecasting increases in the VIX: A time-varying long volatility hedge for equities, NCER Working Paper Series, no. 88. Available online: http://www.ncer.edu.au/papers/documents/WP88.pdf [Accessed 16 May 2022]

Copeland, M.M., & Copeland, T.E. (1999). Market timing: Style and size rotation using the VIX, *Financial Analysts Journal*, vol. 55, no. 2, pp.73-81. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Corradi, V., Distaso, W., & Mele, A. (2013). Macroeconomic determinants of stock volatility and volatility premiums, *Journal of Monetary Economics*, vol. 60, no. 2, pp.203-220. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Degiannakis, S. A. (2008). Forecasting Vix, *Journal of Money, Investment and Banking*, no. 4, pp.5-18. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1806044. [Accessed 30 March 2022]

Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). *Mathematics for machine learning*. pp.371, 388, Cambridge University Press.

Duan, J.-C., & Yeh, C.-Y. (2010). Jump and volatility risk premiums implied by VIX, *Journal of Economic Dynamics and Control*, vol. 34, no. 11, pp.2232–2244. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Efron, B. and Hastie, T. (2021). *Computer Age Statistical Inference, Student Edition: Algorithms, Evidence, and Data Science*, vol. 6, pp.376-382, Cambridge University Press.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica: Journal of the econometric society*, vol. 50, no. 4, pp.987-1007. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Engle, R. F., & Rangel, J. G. (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes, *The review of financial studies*, vol. 21, no. 3, pp.1187-1222. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work, *The Journal of Finance*, vol. 25, no. 2, pp.383-417. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 7 April 2022]

Fernandes, M., Medeiros, M. C., & Scharth, M. (2007). Modeling and predicting the CBOE market volatility index, working paper, no.548, Queen Mary University. Available online: https://www.researchgate.net/publication/24124312_Modeling_and_predicting_the_CBOE_market_volatility_index [Accessed 16 May 2022]

Fernandes, M., Medeiros, M. C., & Scharth, M. (2014). Modeling and predicting the CBOE market volatility index, *Journal of Banking & Finance*, vol. 40, pp.1-10. Available

through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Ferson, W.E., & Siegel, A.F. (2001). The efficient use of conditioning information in portfolios, *The Journal of Finance*, vol. 56, no. 3, pp.967-982. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Fleming, J., Kirby, C., & Ostdiek, B. (2001). The Economic Value of Volatility Timing, *The Journal of Finance*, vol. 56, no. 1, pp.329–352. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using "realized" volatility, *Journal of Financial Economics*, vol. 67, no. 3, pp.473-509. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Gonzalez-Perez, M.T. (2015). Model-free volatility indexes in the financial literature: A review, *International Review of Economics & Finance*, vol. 40, pp.141-159. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Hagan, P., Lesniewski, A., & Woodward, D. (2015). Probability distribution in the SABR model of stochastic volatility, *In Large deviations and asymptotic methods in finance*, pp.1-35, Springer, Cham. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Hastie, T., Tibshirani, R., Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, vol. 2, pp.174, 241-245, 418, 427, 587-588, New York: springer.

Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies*, vol. 6, no. 2, pp.327-343. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine, *Computers & operations research*, vol. 32, no. 10,

pp.2513-2522. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013) *An introduction to statistical learning: with applications in R,* vol. 2, pp.129, 343-345, New York: springer.

Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis*, vol. 42, no. 3, pp.621-656. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Konstantinidi, E., Skiadopoulos, G., & Tzagkaraki, E. (2008). Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices, *Journal of Banking & Finance*, vol. 32, no. 11, pp.2401-2411. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package, *Journal of Statistical Software*, vol. 28, no. 5, pp.1-26. Available online: https://www.jstatsoft.org/article/view/v028i05 [Accessed 2 May 2022]

Liu, F., Tang, X., & Zhou, G. (2019). Volatility-managed portfolio: Does it really work?, *The Journal of Portfolio Management*, vol. 46, no. 1, pp.38-51. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Liu, H., & Cocea, M. (2017). Semi-random partitioning of data into training and test sets in granular computing context. *Granular Computing*, vol. 2, no. 4, pp.357-386. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 25 April 2022]

Luong, C., & Dokuchaev, N. (2018). Forecasting of realized volatility with the random forests algorithm, *Journal of Risk and Financial Management*, vol. 11, no. 4, pp.61-76. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. (2013). *aTSA: adf.test: Augmented Dickey-Fuller Test*. R package version 3.1.2.

Mandelbrot, B., & Hudson, R. L. (2006). The Misbehavior of Markets: A fractal view of financial turbulence, Basic books.

Moreira, A., & Muir, T. (2017). Volatility-managed portfolios, *The Journal of Finance*, vol.72, no. 4, pp.1611-1644. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Nossman, M., & Wilhelmsson, A. (2009). Is the VIX futures market able to predict the VIX index? A test of the expectation hypothesis, *The Journal of Alternative Investments*, vol. 12, no. 2, pp.54-67.

Qiao, G., Yang, J., & Li, W. (2020). VIX forecasting based on GARCH-type model with observable dynamic jumps: A new perspective, *The North American Journal of Economics and Finance*, vol. 53, p.101186. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Pai, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting, *Omega*, vol. 33, no. 6, pp.497-505. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Rekenthaler, J. (2022). Why the 60/40 Portfolio Continues to Outlast Its Critics, Available online:https://www.morningstar.com/articles/1077576/why-the-6040-portfolio-continues-to-outlast-its-critics [Accessed 10 May 2022]

Ren, R., Wu, D. D., & Liu, T. (2018). Forecasting stock market movement direction using sentiment analysis and support vector machine, *IEEE Systems Journal*, vol. 13, no. 1, pp.760-770. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 30 March 2022]

Scott, L.O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application, *Journal of Financial and Quantitative analysis*, vol. 22, no. 4, pp.419-438. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Schwert, G. W. (1989). Why does stock market volatility change over time?, *The Journal of Finance*, vol. 44, no. 5, pp.1115-1153. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

Sharpe, W. F. (1966). Mutual fund performance, *The Journal of Business*, vol. 39, no. 1, pp.119-138. Available online: https://www.jstor.org/stable/2351741?seq=6 [Accessed 12 May 2022]

Taylor, N. (2022). The Determinants of Volatility Timing Performance, *Journal of Financial Econometrics*, pp.1-30. Available online: https://academic.oup.com/jfec/advance-article/doi/10.1093/jjfinec/nbac002/6522743?login=true [Accessed 30 March 2022]

Wang, H. (2019). VIX and volatility forecasting: A new insight, *Physica A: Statistical Mechanics and its Applications*, vol. 533, p.121951. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 31 March 2022]

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach*, vol. 5, p.396, South Wester Cengage learning.

Yang, K., Chen, L., & Tian, F. (2015). Realized volatility forecast of stock index under structural breaks, *Journal of Forecasting*, vol. 34, no. 1, pp.57-82. Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 March 2022]

APPENDIX A

| Variables:	Description:	Source:
Dependent Variable: VIX Index	The VIX Index is a volatility estimate of the S&P 500 Index, obtained by using put and call options on the S&P 500 Index. The VIX Index is obtained daily.	YF
Predictors: <u>*Indices:*</u>		
ES1 Index	E- Mini S&P 500 Futures is an index that at the end of the contract line has the S&P 500 Index, which covers roughly 80% of the US market cap, as an underlying. The ES1 Index is obtained weekly.	Bloomberg
NQ1 Index	The NASDAQ 100E-Mini Futures is an index that, at the end of the contract line, has the NASDAQ 100 Stock Index as an underlying. The NASDAQ 100 consists of the 100 largest non-financial companies on the NASDAQ. The NQ1 Index is obtained weekly.	Bloomberg
VX00 Index	The CBOE Volatility Index Continuous Futures Contract is a future contract for the VIX. The VX00 is obtained weekly.	FactSet
DXY Curncy	The US Dollar Index measures the global value of the US Dollar, as it is computed by taking the average over the exchange rates of US Dollar against other major currencies. The DXY Curncy is obtained weekly.	Bloomberg
VIX[t-1] Index	The same index as the VIX but lagged by one period.	YF
<u>*Commodities:*</u>		
W 1 Comdty	The Generic 1st Wheat Future is obtained weekly.	Bloomberg
C 1 Comdty	The Generic 1st Corn Future is obtained weekly.	Bloomberg

|		
CL1 Comdty	The Generic 1st Crude Oil Future is obtained weekly.	Bloomberg
GC1 Comdty	The Generic 1st Gold Future is obtained weekly.	Bloomberg
HG1 Comdty	The Generic 1st Copper Future is obtained weekly.	Bloomberg
US1 Comdty	The U.S. Treasury Long Bond Future is obtained weekly.	Bloomberg
TY1 Comdty	The 10-Year U.S. Treasury Note Future is obtained weekly.	Bloomberg
Macroeconomic Variables:		
ICSA	The Initial Claims are filled out if an employee loses his job and applies for the Unemployment Insurance program in the US. The ICSA is obtained weekly.	FRED
A191RP	The Gross Domestic Product measures the market value of goods and services produced in the US. The A191RP measures the annual rate of the change in GDP at a quarterly frequency.	FRED
T10YIE	The 10-Year Breakeven Inflation Rate measures the average expected inflation over the next ten years. Inflation is defined as the percentage change in the consumer price index. The T10YIE is obtained daily.	FRED
DFF	The Federal Funds Effective Rate is the average interest rate at which banks borrow money overnight from another bank. The DFF is obtained daily for all seven weekdays.	FRED
INDPRO	The Industrial Production Index measures the real output for plants located in the US by industry groups (mining, manufacturing, utilities), market groups or aggregated. The INDPRO is obtained with a monthly frequency.	FRED

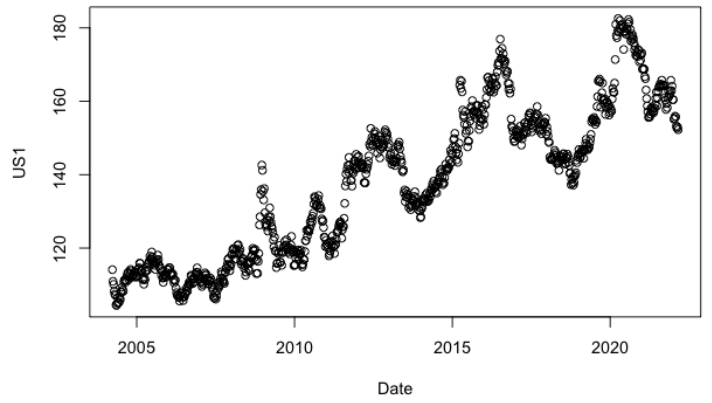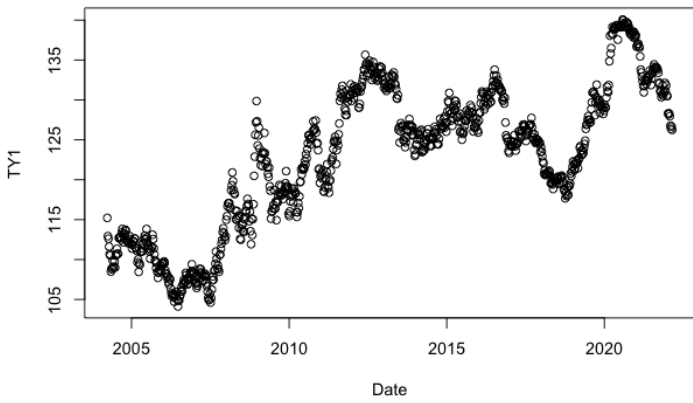## APPENDIX B

Plots of non-stationary variables:

CL1

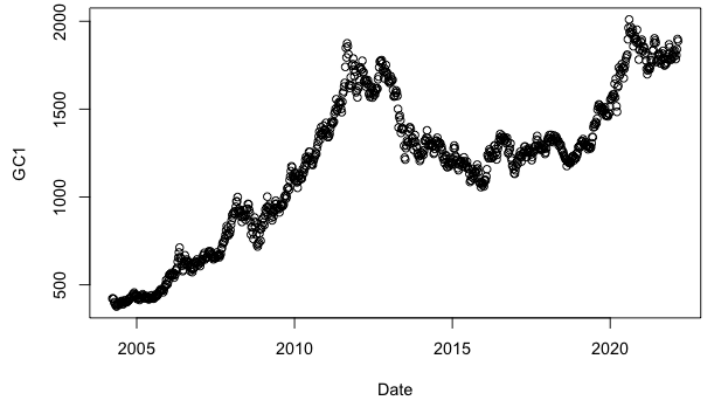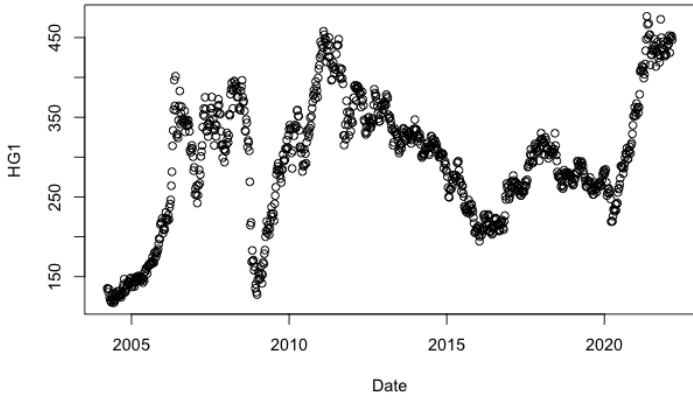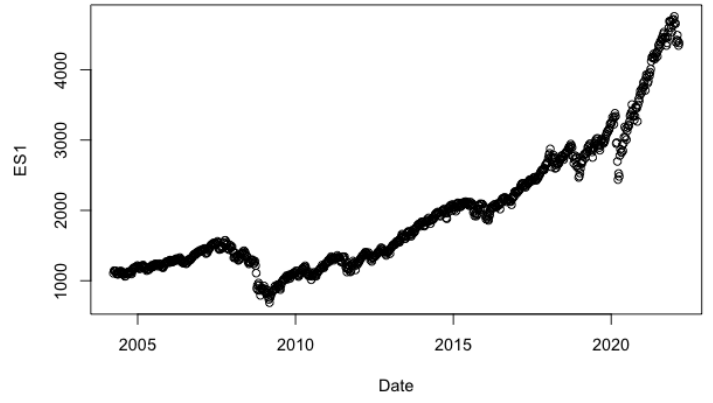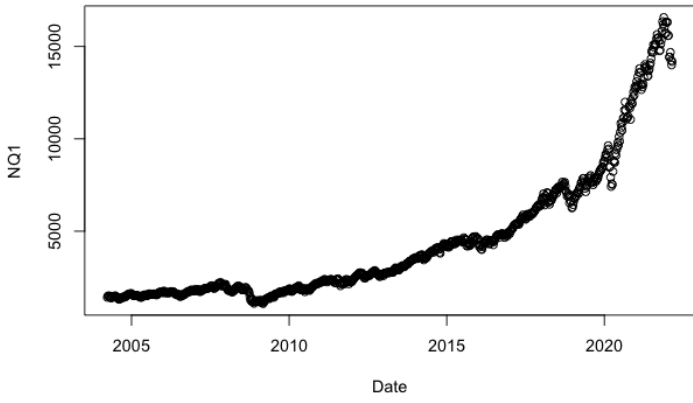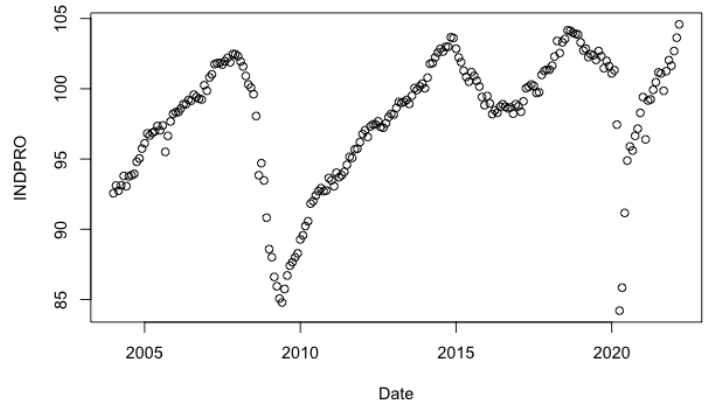

C 1

W1

US1
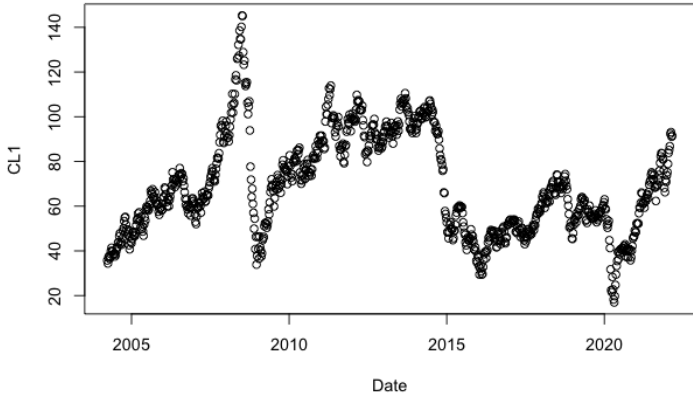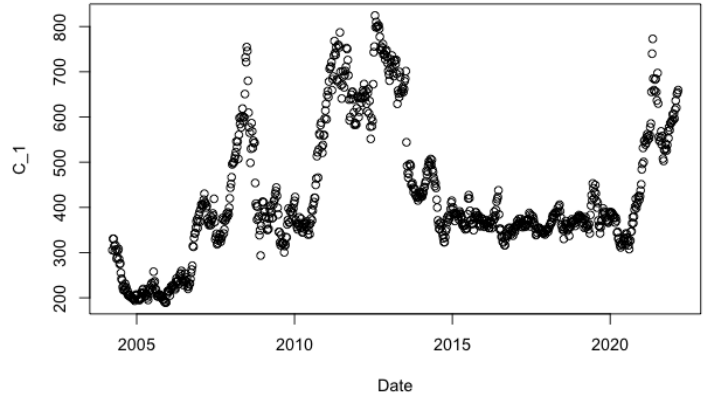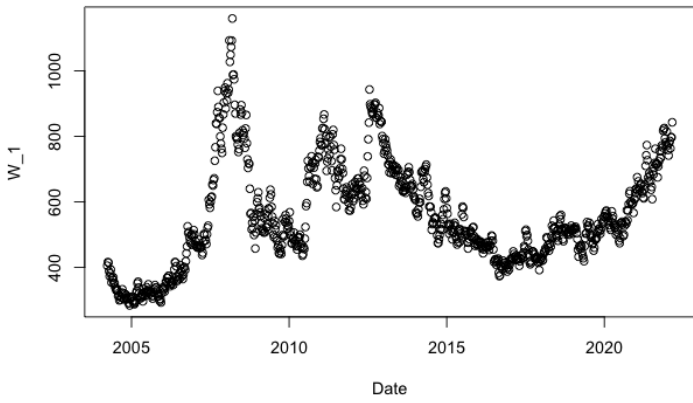
TY1

GC1

HG1

ES1

NQ1

Industrial Production

42