Zurich University
of Applied Sciences

School of
Engineering

IDP Institute of Data Analysis
and Process Design

Project work at Engineering and Management

Allocation tool for asset management from Mobiliar

Author	Pascal Simon Bühler Philipp Rieser
Main Supervisor	Prof. Dr. Marc Wildi
Industrial Partner	Mobiliar
Date	19.12.2020

School of
Engineering

DECLARATION OF ORIGINALITY

Project Work at the School of Engineering

By submitting this project work, the undersigned student confirm that this work is his/her own work and was written without the help of a third party. (Group works: the performance of the other group members are not considered as third party).

The student declares that all sources in the text (including Internet pages) and appendices have been correctly disclosed. This means that there has been no plagiarism, i.e. no sections of the project work have been partially or wholly taken from other texts and represented as the student's own work or included without being correctly referenced.

Any misconduct will be dealt with according to paragraphs 39 and 40 of the General Academic Regulations for Bachelor's and Master's Degree courses at the Zurich University of Applied Sciences (Rahmenprüfungsordnung ZHAW (RPO)) and subject to the provisions for disciplinary action stipulated in the University regulations.

City, Date: ……………………………………………………………

Signature:

……………………………………………………………

……………………………………………………………

……………………………………………………………

The original signed and dated document (no copies) must be included after the title sheet in the ZHAW version of all project works submitted.

Contents

Abstract

1. Introduction

The main purpose of trading is buying and selling stocks, bonds, or other financial instruments with increasing the returns of the investments in mind while maintaining relatively low risk. With the help of a trading strategy, an investor can try to improve his performance. One can simply divide the strategies into passive and active. The praised and well established passive strategy buy-and-hold takes no short price movements into account. Positioning and trading based on these short price movements are considered active trading.

This paper applies time-series analysis to these short price movements to create active trading strategies. The objective of these developed strategies is to outperform the buy-and-hold strategy.

1.1. Data Analysis

The dataset which will be analyzed in this paper contains 4 tradeable indexes, a visualization of the data is shown below in figure 1.

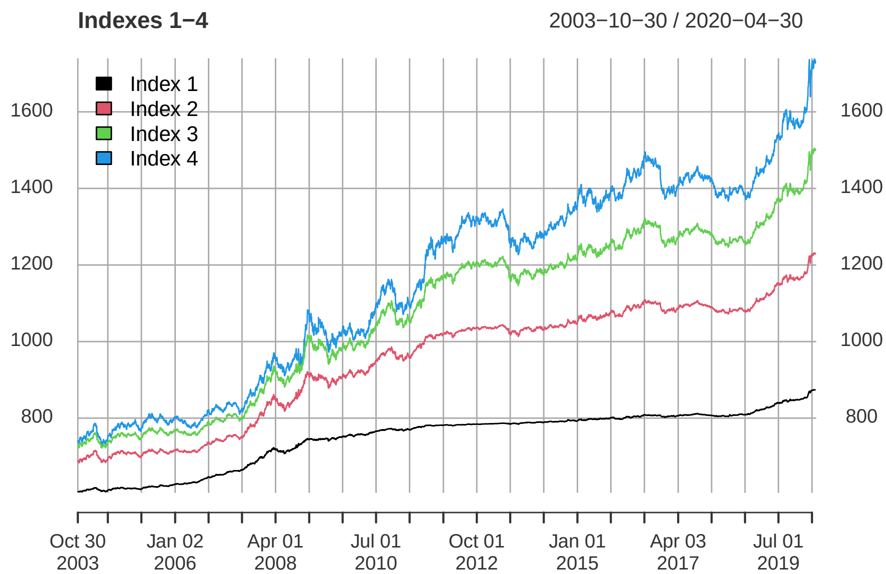

Figure 1: Visualization of the 4 indexes

Each time-series has 4306 observations and starts from October 2003 to April 2020. In all indexes is an upward drift observable, during the time period of the great recession (2008) is a slight bump visible. Also later in 2013 and 2016 are small break-ins evident. More interesting is the up and down behavior at the end of the series during the Covid19 pandemic.

In addition, to the indexes, the dataset contains 8 different interest rates of treasury bonds which will be used for further analysis. A few key-values of the interest rates are shown in the following table 1.

Table 1: Summary of the 8 interest rates.

	Maturity	Mean	Volatility	Min.	Max.
Interest 1	3M	4.09	4.95	-0.28	16.27
Interest 2	6M	4.47	5.05	0.01	16.73
Interest 3	1Y	4.64	4.22	0.20	10.51
Interest 4	2Y	5.42	4.58	0.49	16.58
Interest 5	3Y	6.16	4.33	0.75	16.51
Interest 6	5Y	7.41	3.82	1.06	16.44
Interest 7	7Y	9.50	3.31	1.71	16.63
Interest 8	10Y	11.61	2.93	3.13	17.47

A typical characteristic of interest rates is shown in the given data. A bond with longer maturities is often associated with higher returns compared with those with shorter maturities. An investor which invests in short-term treasury bonds will have his gain earlier but will be confronted with a lower return.

A more in depth analysis of the given dataset will follow in section 3.1.

1.2. Ojective of this paper

The objective of this paper is to trade these 4 indexes with an active trading strategy. The main objective is to outperform the passive buy-and-hold strategy. Methods such as the Moving-Average-Filter or the ARMA-GARCH-Model provide signals for either long or short the position to maximize the return of the investments in these indexes.

The performance of these strategies are build open various different parameters and conditions. The lengths of the filters applied to a Moving-Average may result in different solutions. Models could perform differently for any given length of the in-sample or out-of-sample scope. The necessity of including a historical crisis in the starting-sample can decide if a model performs better or worse than another. The correct validation of model parameters could have a significant impact on the forecasts.

In addition to all criteria and conditions, the strategies can be further adjusted by composing different weighted portfolios. Estimated predicted volatility can be used to modulate the position size to mitigate the risk.

Challenging will be finding the most optimal model in this wide field of conditions and parameters. The buy-and-hold strategy will be used as a benchmark to be compared with the developed active trading strategies. Computing and comparing the Sharpe ratios of each model can serve as an indicator to rely on for better or worse models.

2. Theory

It is assumed that the reader of this paper already has basic knowledge of the mathematical principles of time series analysis. Therefore, this section will only briefly describe the mathematical models and processes.

2.1. Time-Series

Almost anything with a data point to a given timestamp can be named a time-series. The monthly gross domestic product, the weekly US gasoline production, or daily stock price movements. In this paper lies the focus of the analysis of financial time series. Due to trades often only take place during the week, there are gaps in the time series on the weekends, an exception would be the trading of cryptocurrencies like Bitcoin which are also tradeable at the weekends.

A series of data points with more or less equidistant time points t with the sample length of T, is called a time-series $x_t, t = 1, ..., T$ [1]. The analysis of a time-series x_t involves creating a reasonable model that can be utilized to perform forecast predictions.

2.1.1. Stationarity

In order to fit a suitable model with a given time series x_t, the assumptions of stationarity must be met. In this practical application, only the following weak-stationarity properties are required.

$$E[x_t] = \mu \tag{1}$$

$$Var(x_t) = \sigma_x^2 \tag{2}$$

$$Cov(x_t, x_{t-k}) = R(k) \tag{3}$$

Many financial time-series are subject to shift, trends or changing volatility. In figure 2 are the stock prices of Alphabet Inc Class A (Google) visualized. This time-series shows a clear upwards drift and towards the end the volatility increases.

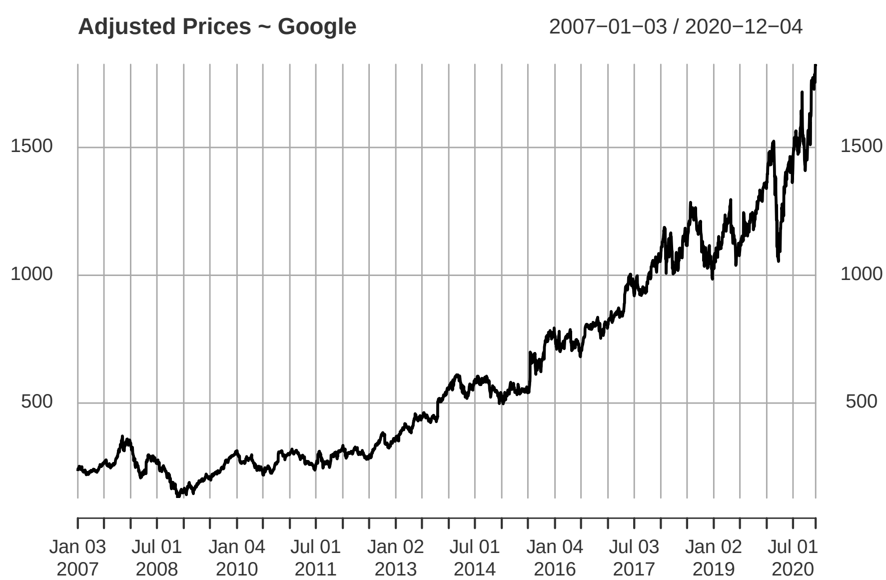

Figure 2: Visualization of the adjusted prices of the Alphabet Inc Class A Stock.

To improve the violated properties the first difference can be applied and additionally a logarithmic transformation can be performed [2]. The log-returns transformation can only performed to strict positive data.

$$\text{LogReturn} = \log(x_t) - \log(x_{t-1})$$

The result is the so-called log-returns.

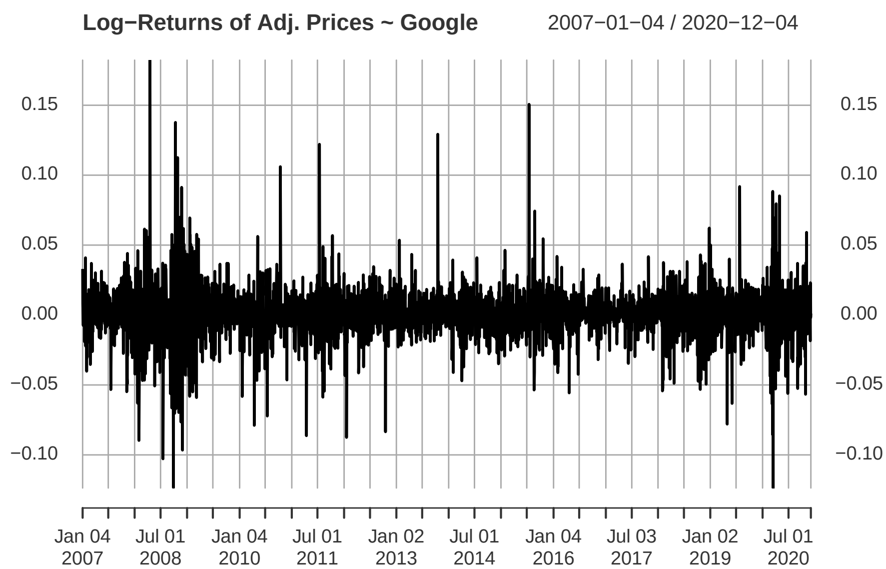

Figure 3: Visualization of the Log-Returns

Applying the transformation to the data causes the drift to disappear, but the series still contains stronger and weaker volatile phases. This effect often occurs in non-stationary financial data and is called volatility cluster. This special property is used for the modeling of forecast models, which will be discussed in chapter 2.2.

In the following examples, we will only work with a section of the time series, as it often makes no sense to look long into the past. The further a value lies in the past, the smaller its influence on a future value will be.

2.1.2. Autocorrelation

The autocorrelation function (ACF) reveals how the correlation between any two data points of the time series changes as their separation changes [3]. More precisely, acf measures the dependence between x_t and $x_{t\pm k}$ at lag k. The partial autocorrelation (PACF) measures the dependency between x_t and x_{t-k} at lag k [1]. For stationary time series, ACF can be used to identify the model order of a MA-process, PACF for AR-processes.

In the following figure 4 are ACF and PACF of the non-stationary adjusted Google stock visualized. Both graphics show the typical pattern of a non-stationary time series. The plot above shows the dependence structure of the time series. This means that it takes a long time until the series changes. Often a large value is followed by another large value, which indicates a strong trend. This property of the series can be seen in figure 2 as the long upward drift. The plot below indicates a significant partial autocorrelation at lag $k = 1$.

In the following section 2.2. the characteristics of the autocorrelation function can be used for the verification of ARIMA and ARCH-processes.

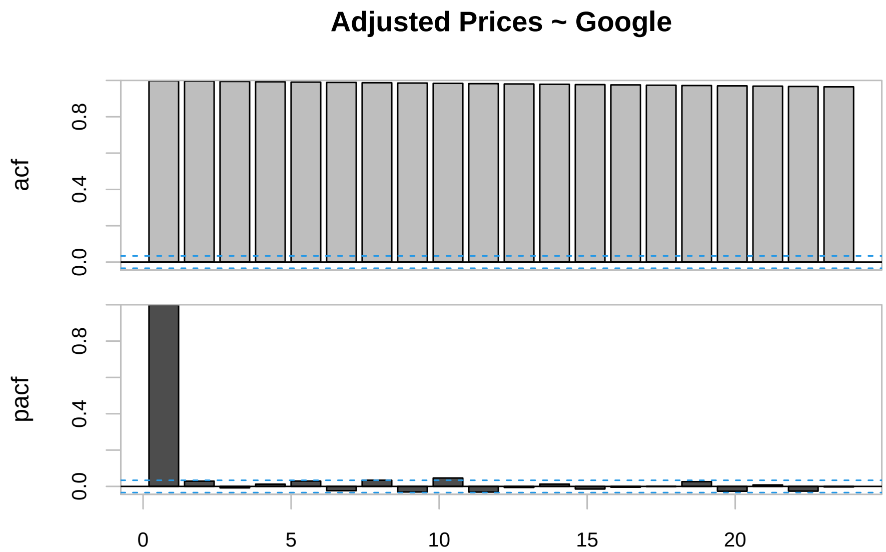

Figure 4: Acf and Pacf of the Adjusted Prizes of Google.

2.2. Models

The following processes are used to determine certain properties and characteristics of a time series so that they are transformed into a model. The goal is to fit the time series as well as possible in order to create reliable forecasts.

2.2.1. ARIMA

An ARIMA(p,d,q) process is defined as follows.

$$x_t = c + a_1 x_{t-1} + ... + a_p x_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + ... + b_q \epsilon_{t-q} \tag{4}$$

- p and q are the AR- and MA-model orders
- a and b are the AR- and MA-model parameters
- d is the differential parameter
- ϵ_t is a white noise sequence
- x_t is the given data $x_1, ..., x_T$

The mean of an ARIMA-process can be computed as:

$$\mu = \frac{c}{1 - a_1 - ... - a_p}$$

ARIMA processes can be divided into 4 different models. Choosing a model that best represents the time series is a difficult task. The goal is to find the best possible model with as few parameters as possible.

The previously introduced ACF and PACF can help to determine the orders of simple models. Provided that the time series is stationary, the model orders can be determined directly. For an AR(p)-process (ARIMA(p,0,0)), the ACF plot will gradually decrease and simultaneously the PACF should have a sharp drop after p significant lags. For an MA(q)-process (ARIMA(0,$0.q$)) the opposite is true, the ACF should show a sharp drop after a certain q number of lags while PACF should show a gradual decreasing trend. If both ACF and PACF show a gradual decreasing pattern, then the ARIMA(p,0,q)-process should be considered for modeling [4]. If the time series is not stationary, differentiation can be considered (ARIMA(p,d,q)).

The application of an analysis method to Google prices finds an ARIMA($1,1,0$) as the optimal model. This makes sense if you look back at figure 4. Long dependency structures in the ACF plot indicating an AR(p) process and at the same time after lag=1=p the PACF has a strong drop. The differential operator d=1 transforms the non-stationary series into a stationary one.

To convince yourself of the quality of the model, you can use the Ljung-Box statistics shown in figure 5. For the lags where the p-values are above the 5% line, the forecasts are reliable. Here the values from the 6 lag are below the significant line, but since one only wants to make short-term forecasts (for example 1 day into the future), the lag is sufficient up to 5. If you want to forecast further than 5 days into the future, you might have to adjust the ARIMA model.

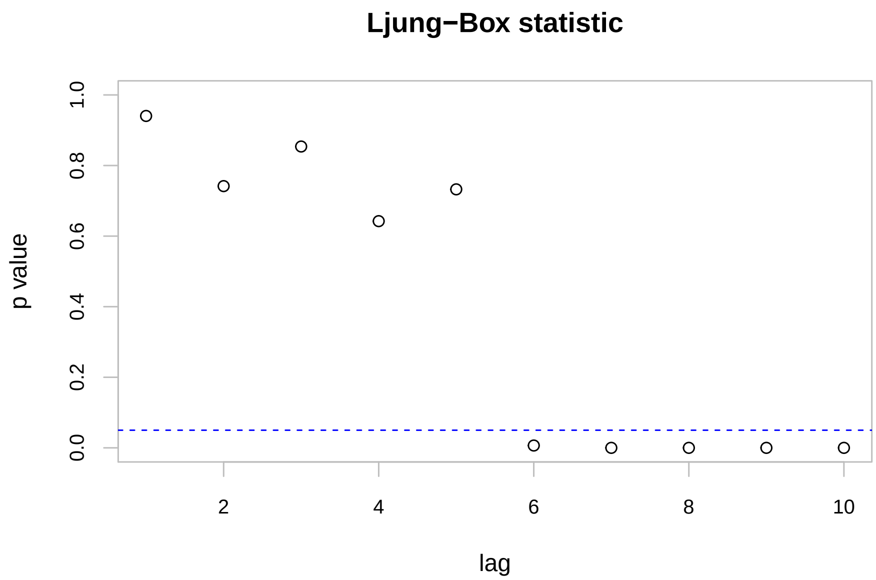

Figure 5: Ljung-Box statistic of Google log returns.

In figure 6 you can see the prediction of the model. The whole representation is shifted by one day so that one can compare the model with a true value. The green dot is the actual value of the time series. The red dots indicate the upper and lower 95% interval limits respectively. These indicate that a future value will be within this band. The blue dot is the point forecast predicted by the model.

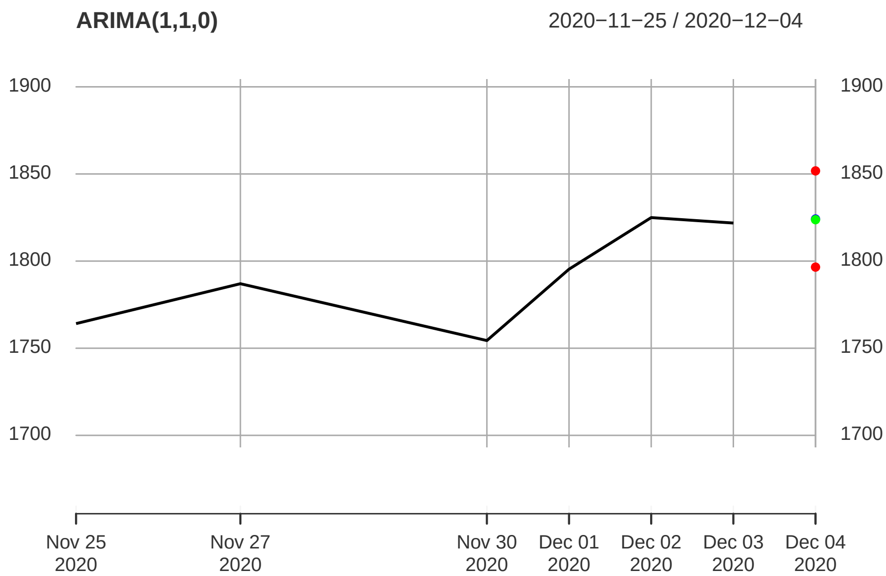

Figure 6: ARIMA-Forecast.

2.2.2. GARCH

The volatility clustering mentioned in section 2.1.1 can be handled with an generalized auto-regressive conditional heteroscedastic process.

$$\begin{aligned} \epsilon_t &= \log(x_t) - \log(x_{t-1}) \\ \epsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= c\sigma^2 + \sum_{j=1}^{n} \alpha_j \sigma_{t-j}^2 + \sum_{k=1}^{m} \beta_k \epsilon_{t-k}^2 \end{aligned} \tag{5}$$

with

- x_t is the original data (often non-stationary)
- ϵ_t is the stationary log-return
- u_t is independent and identically distributed (iid) and a standardized random variable
- σ^2 is the unconditional variance of the process ϵ_t.
- σ_t^2 is the conditional variance of the process ϵ_t.

With a GARCH(n,m)-process it is possible to model the volaclusters of a time series. The GARCH(1,1) model has become widely used in financial time series modeling and is implemented in most statistics and econometric software packages. Those models are favored over other stochastic volatility models by many economists due to their relatively simple implementation [5].

Table 2: Coefficients GARCH(1,1).

	Estimate	Std. Error	p-Value
ω	1.094153e-05	1.850391e-06	3.357512e-09
α_1	7.379778e-02	1.081436e-02	8.850698e-12
β_1	8.947058e-01	1.382004e-02	0.000000e+00

For an optimal model some conditions must be fulfilled. Suppose you want to model the Google time series with a GARCH(*1*,*1*).

In table 2 the estimated coefficients of the process can be seen. The p-values are all lower than 0.05 and thus indicate that they are essential for the model. (Note: $\omega = c\sigma^2$)

The following parameter restrictions are also examined:

$$c + \sum_{j=1}^{n} \alpha_j + \sum_{k=1}^{m} \beta_k = 1 \tag{6}$$

with

$$c > 0, \alpha_k \geq 0, j = 1, ..., n, \beta_k \geq 0, 1, ..., m$$

To satisfy formula 6, c needs to be determined from ω. First calculate the unconditional variance.

$$\sigma^2 = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

Calculate c with:

$$c = \frac{\omega}{\sigma^2}$$

and then check for the restriction in 6.

For the coefficients of GARCH(*1*,*1*) the restrictions are fulfilled. You can see that $c = 1 - \alpha_1 - \beta_1$. So this restriction can be determined easily with:

$$\sum_{j=1}^{n} \alpha_j + \sum_{k=1}^{m} \beta_k < 1 \tag{7}$$

If the parameter restrictions are not fulfilled, complications may arise, the forecast of the conditional variance $\hat{\sigma}_t^2$ may diverge to the unconditional variance σ^2 of the process.

Furthermore, the Ljung-Box statistics are important for the standardized residuals. Looking back at formula 5 standardized residuals u_t are proportional to the conditional volatilities σ_t, which should lead to the log returns ϵ_t. The conditional volatilities map the volacluster in the time series. To achieve the best possible model, one does not want to find these volacluster effects in the standardized residuals, but only in the conditional volatilities. The Ljung-Box statistics check this property. In figure 7 Ljung-Box statistics of the $\hat{u}_t$ and the $\hat{u}_t^2$ are shown.

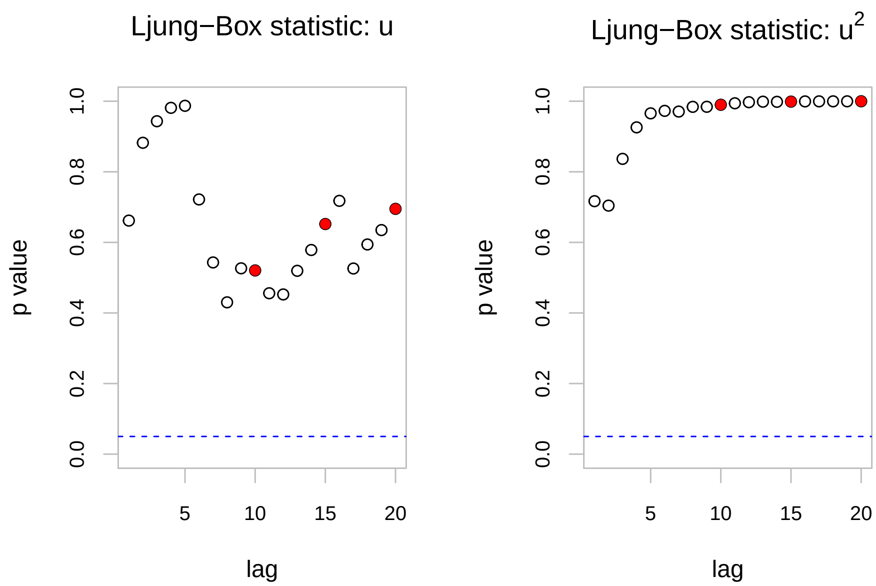

Figure 7: Ljung-Box statistic of the standardized residuals.

The plot shows reliable statistics for forecasts from the Google-GARCH(*1*,*1*) model. For all lags up to 20, the p-values are above the 5% line and thus hypothesis tests are discarded. However, if for a given lag=k the p-values would fall below the 5% line, then forecasts would only be reliable up to a forecast horizon k.

If one wants to improve the the standardized residuals $\hat{u}_t$ (if autocorrelations exists in the $\hat{u}_t$), an ARMA part would have to be added to the existing model (see 2.2.3). This can again be optimized with different model orders. If you want to improve the squared standardized residuals $\hat{u}_t^2$ (if volaclustering exists within the $\hat{u}_t^2$), then you should modify the GARCH model order.

Now an optimal model has been found and a forecast can be made. Since a GARCH(n,m) process is white noise sequence the expected value $E[\epsilon_{T+h}|\epsilon_T, ..., \epsilon_1] = \mu = 0$ can be assumed (if the mean value in the fit object was also estimated and is significant, then the expected value is the estimated μ).

Calculating the forecast variance is a recursive process. With increasing model order the calculation becomes more and more difficult. For this work the rather simple calculation for a GARCH(*1*,*1*) model is sufficient. One receives:

$$\hat{\sigma}_{T+h}^2 = \omega + (\alpha_1 + \beta_1)^2 \hat{\sigma}_{T+h-1}^2$$

If the parameter restriction from formula 7 is true, the forecast variance converges with the increasing forecast horizon to the unconditional variance of the process.

$$\hat{\sigma}_{T+h}^2 = \frac{\omega}{1 - \alpha_1 - \beta_1} = \sigma^2$$

The 95% forecast interval is calculated as follows:

$$E[\epsilon_{T+h}|\epsilon_T,...,\epsilon_1] \pm 1.96\sqrt{\hat{\sigma}^2_{T+h}}$$

Figure 8 shows the GARCH(*1,1*) forecast for 20 days. The blue line represents the expected value of the time series. Reds are the two 95%-interval limits and green shows the actual values of the series.

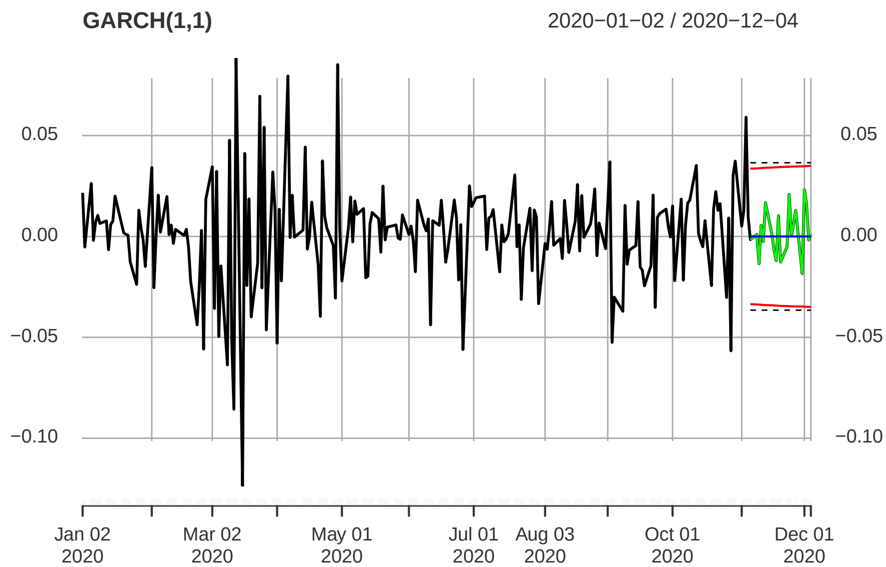

Figure 8: GARCH-Forecast.

You can see how the red lines are approaching the black dotted lines ($0 \pm 1.96 * \sqrt{\sigma^2}$) as the forecast horizon increases. This means that the forecast variance converges to the unconditional variance and thus there is another proof for the satisfaction of the parameter restrictions.

2.2.3. ARMA-GARCH

If you want to adopt a GARCH model and you discover non-vanishing autocorrelations in the standardized residuals $\hat{u}_t$, the model can be extended with an ARMA part to counteract the autocorrelations.

$$\begin{aligned} y_t &= \mu + a_1 y_{t-1} + ... + a_p y_{t-p} + \epsilon_t + b_1\epsilon_{t-1} + ... + b_q\epsilon_{t-q} \quad (8) \\ \epsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= c\sigma^2 + \sum_{j=1}^{n}\alpha_j\sigma^2_{t-j} + \sum_{k=1}^{m}\beta_k\epsilon^2_{t-k} \quad (9) \end{aligned}$$

This model contains now 4 different model orders: p,q,n and m. Experience shows that for financial time series, a model order of n=m=1 and $p,q\leq 1$ is often sufficient to fit the data to an ARMA-GARCH model.

The equation 8 is called the *mean-equation* and describes the conditional mean of the process, which is used to obtain optimal point forecast. The equation 9 defines the variance of the process and is called *variance-equation*, which is used for the forecast error variance [6].

2.2.4. M-GARCH

The last model that is considered as a forecast model in this thesis is the so-called MGARCH model. Highly volatile phases indicate downturns. Market participants tend to oversell during these downturns. Overselling leads to inflated volumes and this in turn leads to inflated volatilities. This model can be helpful since it emphasizes recession and crisis dynamics.

$$\begin{aligned} r_t &= \log(x_t) - \log(x_{t-1}) \\ r_t &= \mu + e\sigma_t + \epsilon_t \\ \epsilon_t &= \sigma_t u_t \end{aligned} \tag{10}$$

where

- x_t is the original data (typically non-stationary)
- r_t are the log-returns (stationary)
- μ is the long-term drift
- ϵ_t is a volacluster process (GARCH)
- e is a constant (a parameter to be estimated), $e > 0$ implies a larger expected return. $e < 0$ would imply a smaller expected return. If $e = 0$ then the MGARCH-effect vanishes [7].

For this model, the in-sample conditional standard deviations (volatilities) from any GARCH process are determined and the out-of-sample conditional standard deviation for obtaining forecasts of the future returns is then calculated by regression.

$$\hat{r}_{t+1} = \hat{\mu} + \hat{e}\hat{\sigma}_{t+1}$$

2.3. Moving Average Filters

Moving average filters are basically used to identify trends and smooth out price fluctuations. As a commonly used tool, moving average filters are very simple in their usage, historical data from a timeframe L gets summarized and divided by the length of the filter (L). Depending on the length of the filter, the application to the timeseries the MA gets shifted, longer filters have a higher shift then shorter ones we visualize this behavior in 9. Many different indicators are built upon the Moving Average principle, mostly they are used in combinations of different lengths to create signals. In the following section, we introduce some of the most popular indicators based on the MA principle.

The actual challenge in using Moving average filters, is to figure out which length of the filter brings the most useful information.

2.3.1. Equally-weighted Moving Average or SMA

SMA stands for Simple Moving Average, depending on the length of the filter(L), L observations since the last noted observation will be considered. The observations are getting summarized and divided by the filterlength L. As stated in the name, all past observations are weighted equally. For every timestep, a new observation is considered and the last one eliminated [8]. SMA's are very easily customized by changeing the length of the filter.

EqMA

$$y_t = \frac{1}{L} \sum_{k=0}^{L-1} x_{t-k} \tag{11}$$

- L = filterlength
- x = original series price e.g.

2.3.1.1 Momentum Momentum is an indicator wether a market is bullish or bearish, it measures the "speed" of the trend direction in the market. for a timespan k the last price p_k, k timesteps ago is subtracted from the last price. This is equivalent to applying and EqMA on the returns of a series.

Momentum

$$y_t = p_t - p_{t-K} \tag{12}$$

- p_t = prices of the series
- K = Lag

2.3.2. Exponentially-weighted Moving Average

Since not all observations are having the same influence on the future value, we can apply a weight to past observations. One method will be exponentially weighted Moving average. So we chose an optimal parameter to give past observations weights decreasing by α^k. In comparison to the SMA 2.3.1. [9] ,a EMA from the same length L, reacts faster than to price changes.

A skillful trader chooses an optimal α to increase the performance of the measurement. Weights could also be given individually by adding a weight vector to the filter.

EMA

$$y_t = \frac{1}{\sum_{k=0}^{m} \alpha^k} \sum_{k=0}^{m} \alpha^k x_{t-k} \tag{13}$$

- m = filterlength
- α = Parameter to weigh the observations

2.3.3. Moving Average Crossings

Moving average crossings are basically just different MA's with different lengths applied to a time-series. The points the filters then cross, will be used as a trading signal to go long, short or hold. The "death" and "golden" cross are very popular trading patterns [9]. If a shorter MA crosses the longer MA from above, its called a "golden cross" it is an indicator that the price will rise in the future and can be used to created the buy signal. In contrast stands the "death cross" vise versa, a shorter MA (popular L = 50) crosses a longer MA(popular L = 200) from below, signalizing that further losses are in store.

Moving Average Crossing

$$y_t = \frac{1}{L_1} \sum_{k=0}^{L_1 - 1} x_{t-k} - \frac{1}{L_2} \sum_{k=0}^{L_2 - 1} x_{t-k} \tag{14}$$

- L_1 = filterlength 1
- L_2 = filterlength 2
- $0 < \alpha < 1$ = Parameter to weigh the observations

An example of MA average crossings with 2 SMAs of different length is visualized in figure 9. The prices are in green while the blue line represents an 50 day SMA an the red line a 250 (1 year) SMA. The crossing points of those two SMAs could now be used as trading signals. In this example these crossings wouldn't perform very well, therefore as mentioned earlier, finding the right length of the filter depends on each timeseries, their behavior and the preferences of the trader.

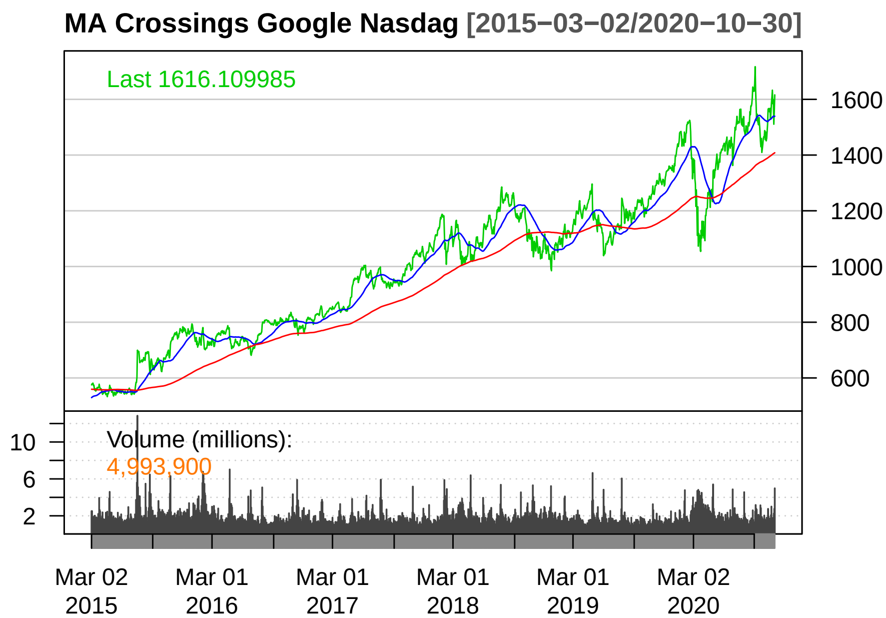

Figure 9: Moving Average Crossing

2.4. Relative Strength Index

The Relative strength index is a tool to measure momentum, the value indicates if positions are overvalued or undervalued. the scale goes from 0 to 100. [10].

$$U_t = \begin{cases} 1, \text{if } x_t \geq x_{t-1} \\ 0, \text{otherwise} \end{cases} \quad D_t = \begin{cases} 1, \text{if } x_t < x_{t-1} \\ 0, \text{otherwise} \end{cases} \tag{15}$$

- X_t =Original timeseries

We then apply SMA or EMA of length N to U_t and U_t which converts them in $up_t(N)$ and $down_t(N)$ The RSI is now computet by:

$$RSI_t(N) = 100 \frac{up_t(N)}{up_t(N) + down_t(N)} \tag{16}$$

The original developer J. Welles Wilder Jr. proposed a length of N =14 in its work from 1978 [11]. Traditional tradings signals based on the RSI are the upper 70 or lower 30 limit to buy or sell. Usually when the RSI is going over 70 it suppose that the asset is overvalued and in contrast when its under 30 then its undervalued.

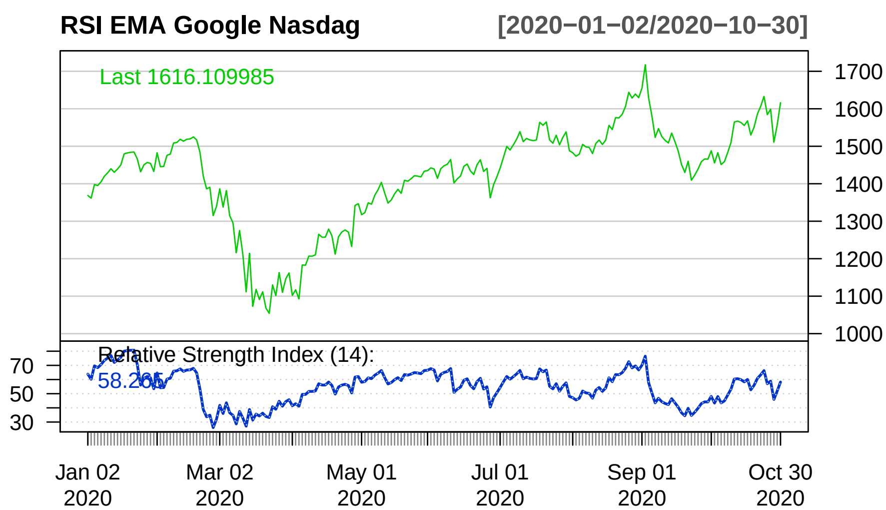

Figure 10: RSI

2.5. Moving Average Convergence Divergence

The MACD is also a commonly used filter. The basic principle is to Subtract a longer EMA with length L as in section 2.3.3 from a shorter EMA from length S then smooth the result with another EMA with length R. As a result with can use the crossing of the 2 generated curves for trading. As an alternative we could use SMA's instead of EMA's

$$macd_t = \frac{1}{\sum_{k=0}^{t-1} \alpha^k} \sum_{k=0}^{t-1} \alpha^k x - t - k - \frac{1}{\sum_{k=0}^{t-1} \beta^k} \sum_{k=0}^{t-1} \beta^k x - t - k \tag{17}$$

$$MACDsignal_t = \frac{1}{\sum_{k=0}^{t-1} \gamma^k} \sum_{k=0}^{t-1} \gamma^k x - t - k \tag{18}$$

- x_t = prices or log prices
- S = length of the $short_1$ EMA usually 12
- L = length of the $long_1$ EMA usually 26
- R = length of the "double smoothing ema" usually 9
- $\alpha = 1 - \frac{1}{S}$, $\beta = 1 - \frac{1}{L}$, $\gamma = 1 - \frac{1}{R}$

$_1$ short and long in the meaning s<l, not buy sell

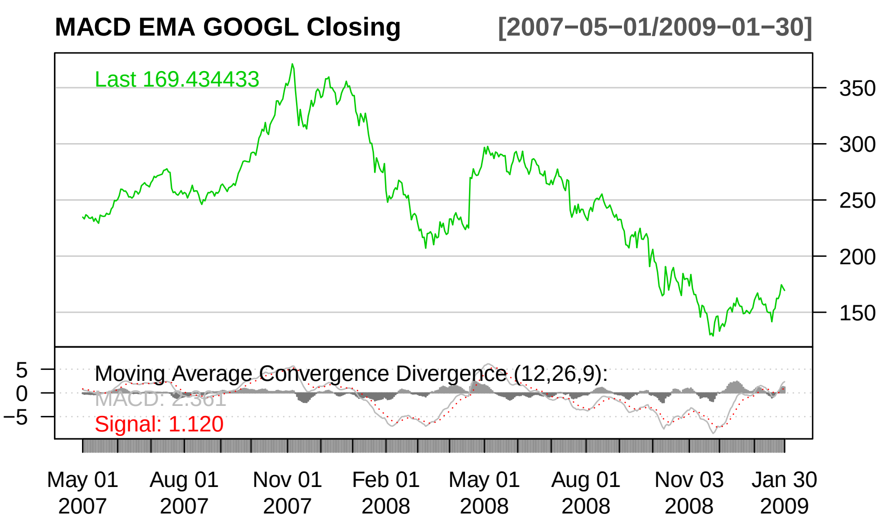

Figure 11: MACD

2.6. Bollinger bands

Bollinger bands are a analysis tool founded by John Bollinger. It contains a moving average and an upper and lower band . The bands are defined by adding a constant K times a standard deviation σ_t to the *Moving Average* for the upper , and subtracting it for the lower band.

$$\mathrm{U_t} = MA_t + K\sigma, L_t = MA_t - K\sigma \tag{19}$$

the variance from bollingers theory is calculated by:

$$\sigma_{\mathrm{t}}^2 = \frac{1}{N}\sum_{k=0}^{N-1}(x_{t-k} - MA_t)^2 \tag{20}$$

The calculated σ_t could be problematic because its derived from the original series and increases with the level, its non stationary. Therefore an other method to calculate the standard deviation could be used. As done in section 2.2.2. σ_t could be provided by a GARCH, which would handle the increasing volatility.

- N = usually the filterlength and the length considered for σ are the same
- K = Constant usually equals 2
- σ_p = standard deviation of the series
- U_t = upper band
- L_t = lower band

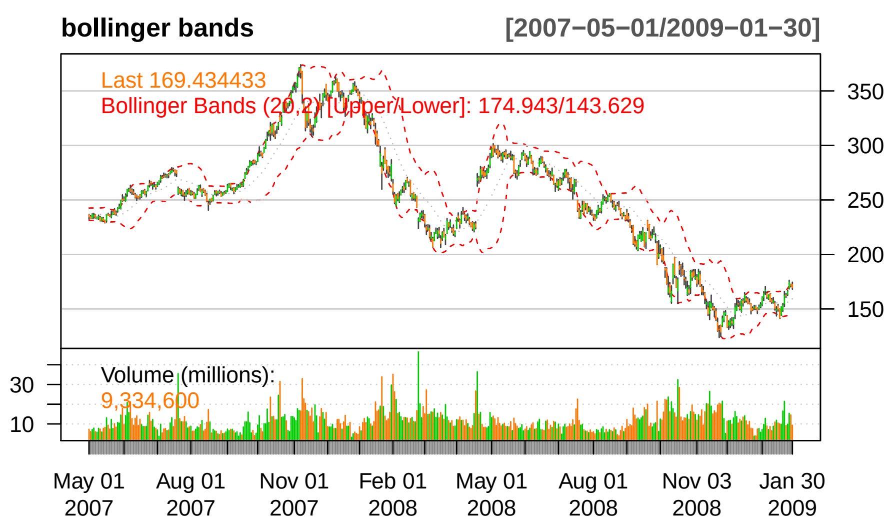

Figure 12: Bollinger Bands

2.7. Performance Indicators

2.7.1. Sharpe Ratio

Sharpe ratio is a very powerful and widely used ratio to measure performance. It describes return per risk.

$$\text{SharpeRatio} = \frac{R_p - R_f}{\sigma} \tag{21}$$

- R_p = Return of Portfolio
- R_p = Risk free Rate, mostly treasury bonds
- σ_p = standard deviation of portfolios excess return (risk)

2.7.2. Drawdown

2.7.3. MSE

2.8. Carry

Carry trades are trading strategies were usually money is borrowed at a lower interest rate, than the investment is giving in return. the risk of this strategy is based in the currency risk.

3. Methodology

In this section, different models are created and compared with the buy and hold strategy. You start with pure data analysis and then work your way from simple models to more and more complex ones.

3.1. Data Analysis

As mentioned in section 1.1 we are now going to analyze the data further to gain as much information as possible just by using some simple tools and comparisons.

3.1.1. Correlation

One could nearly tell just by looking at the indexes how strong they're correlated. The correlation matrix in table 3 confirms the assumption, the correlation is nearly 1 for every index to each other.

Table 3: Correlations oft the four indexes

	Index 1	Index 2	Index 3	Index 4
Index 1	1.0000000	0.9899111	0.9788826	0.9672956
Index 2	0.9899111	1.0000000	0.9975499	0.9921171
Index 3	0.9788826	0.9975499	1.0000000	0.9983460
Index 4	0.9672956	0.9921171	0.9983460	1.0000000

3.1.2. transformation, volatility and clusters

Applying the natural logarithm to the series is an approach to cancel out increasing volatility 13. The strong upward drift is still visible, the original series has more than doubled to its original price over the whole timespan (index 4) 1.

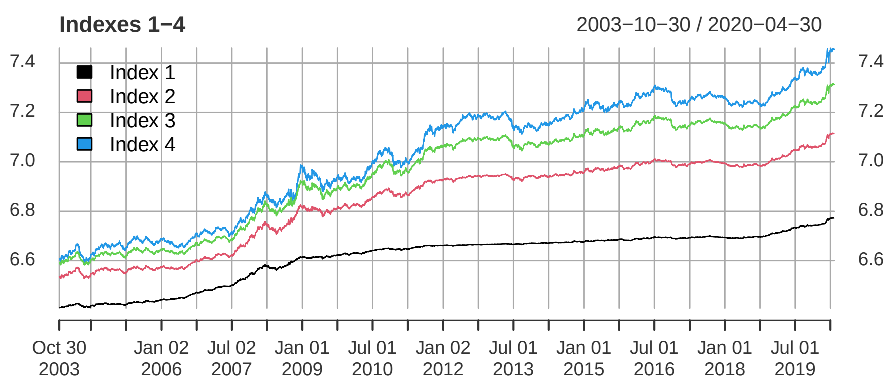

Figure 13: Visualization log indizes

By taking the returns of the transformed series we can visualize volatility clusters as seen in figure 14. The first value of the series is eliminated because of the differences, Clearly visible are the high spikes in the times of the financial crisis 2007-2009. Also at the end of the series, the impact of COVID19 in march 2020 is remarkable.

Table 4: Sharpe Ratio and Volatility of the 4 indexes (log-returns).

	Index 1	Index 2	Index 3	Index 4
Volatility	0.0008	0.0020	0.0029	0.0039
Sharpe-Ratio	1.6710	1.0601	0.9143	0.7922

The unconditional volatility of the indexes are 0.64e-06, 4..1e-06, 8.5e-06, 15.6e-06. Which means the first index is 24 times more volatile than the fourth index.

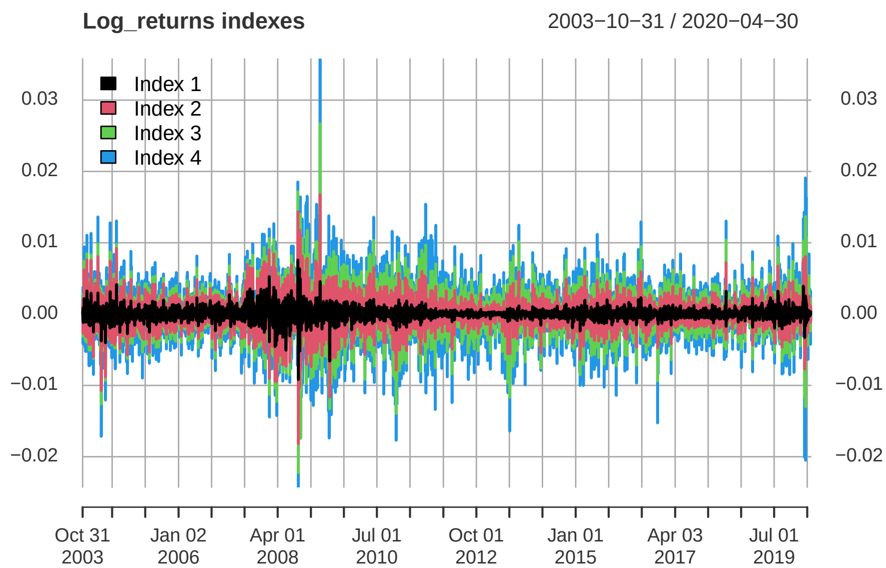

Figure 14: Log returns

3.1.2.1. Autocorrelation of log returns

By computing the ACF of the squared log-returns we see that the volatility cluster has very long dependency structures.

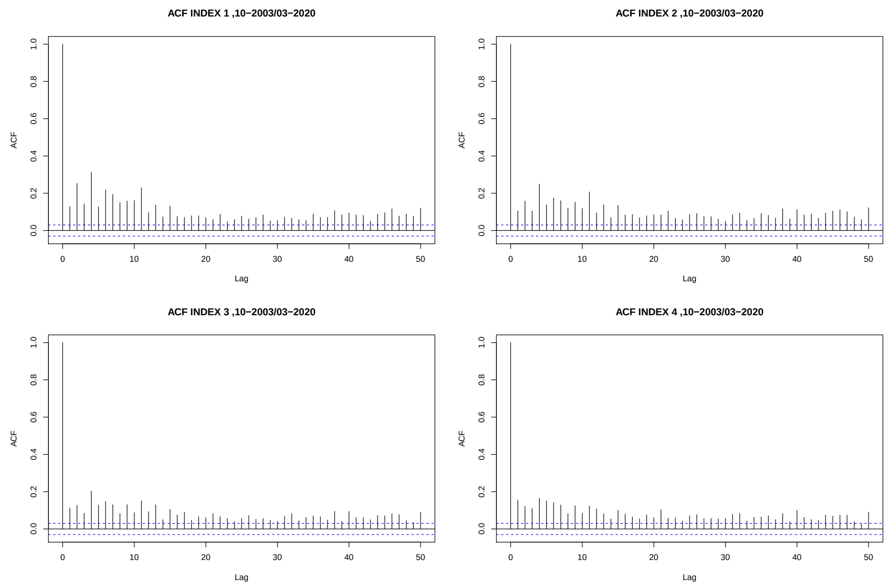

Figure 15: ACF log returns

For further analysis, we concentrate on different timespans, because in most models it makes little sense to consider the whole time-series from 2003 till today.

3.2. Trading

In section 2 we have learned different indicators and models for time-series-analysis. Some of these models and indicators are now used to trade the indexes we've introduced in the previous section.

To do so we need to create trading signals based on the models and indicators. For example, we are using the MA Crossings, as mentioned 2.3.3. the points where the two MAs cross, are now used to create a trading signal. when the longer MA comes from below to the crossing we are going long the asset and if it approaches the point from above we're shorting the position. Technically we apply a 1 to a vector at each crossing, where we intend to buy and apply a -1 at the points we want to sell.

3.2.1. Buy and Hold Performance

As mentioned earlier the goal of this work is trying to outperform the buy and hold strategy. Because the series all have a strong upward trend this task is very tricky. Buy and hold has very low trading costs because the underlying is just bought once. According to swissquote [12] costs for asset trades over 50k, are 190 USD per trade $_1$, so these costs should also be taken into consideration for choosing the strategy. In this work, we've excluded these costs and only concentrate on the trades themselves.

$_1$ notice: This fee is only for private investors, conditions may differ for institutions.

For the following models a good comparability should be achieved. To ensure this, all models are created with the same out-of-sample range, with the start date 2019-01-01.

3.2.1.1. Portfolios

By focussing on trading we build a portfolio for the 4 indexes. One approach would be the equally weighted portfolio with weights for every index of $\frac{1}{4}$. As we've seen in figure 14 The volatility of the indexes strongly differs, meaning that index 4 would have the most impact, nearly 50 %, on the portfolio by weighing it equally. To cancel this effect, it may be useful to size the position with the inverse volatility, so the indexes have the weights seen in the table on the right side.

$$procentualshare\sigma_k = \frac{\sigma_k}{\sum_{k=1}^{4} \sigma_k} \tag{22}$$

Table 5: shares

	percentage of vola when equally weighted	weights when weigthed with inverse vola
Index 1	3.70980	70.246503
Index 2	16.15568	16.130577
Index 3	31.55567	8.258436
Index 4	48.57885	5.364485

3.2.2. AR trading

As with the theory data set, these 4 indexes are non-stationary series. For the AR models, the stationary log-returns of the time series are used. First we look at index 1, the series with the lowest volatility but with the highest Sharpe ratio (highest return with lowest risk). To find the best possible AR model, different AR(p) models are fitted with different model orders p for different in-sample ranges. For each model the out-of-sample sharpe is calculated. The execution of this calculation leads to the solution shown in table 6.

Table 6: Solution of the AR-Model calculation.

StartDate	AR-Sharpe	p	|	StartDate	AR-Sharpe	p
2003-01-01	3.296	1	|	2011-01-01	2.004	3
2004-01-01	3.037	1	|	2012-01-01	1.852	3
2005-01-01	3.146	1	|	2013-01-01	2.033	4
2006-01-01	2.957	1	|	2014-01-01	1.954	3
2007-01-01	2.722	1	|	2015-01-01	1.932	5
2008-01-01	1.867	1	|	2016-01-01	3.449	2
2009-01-01	2.473	1	|	2017-01-01	2.163	4
2010-01-01	2.651	1	|	2018-01-01	3.765	2

With these calculations, the optimal model is an AR(2) with the start date 2018-01-01. Note that the Index 1 series has a very flat rising trend and a rising upward trend towards the end of 2018, so only a short in-sample is needed for this optimal model. It is interesting that the long in-sample ranges (2003 to 2007) also lead to relatively good performances. This probably has something to do with the fact that trend at the beginning of the series has a greater influence on the model than the middle part.

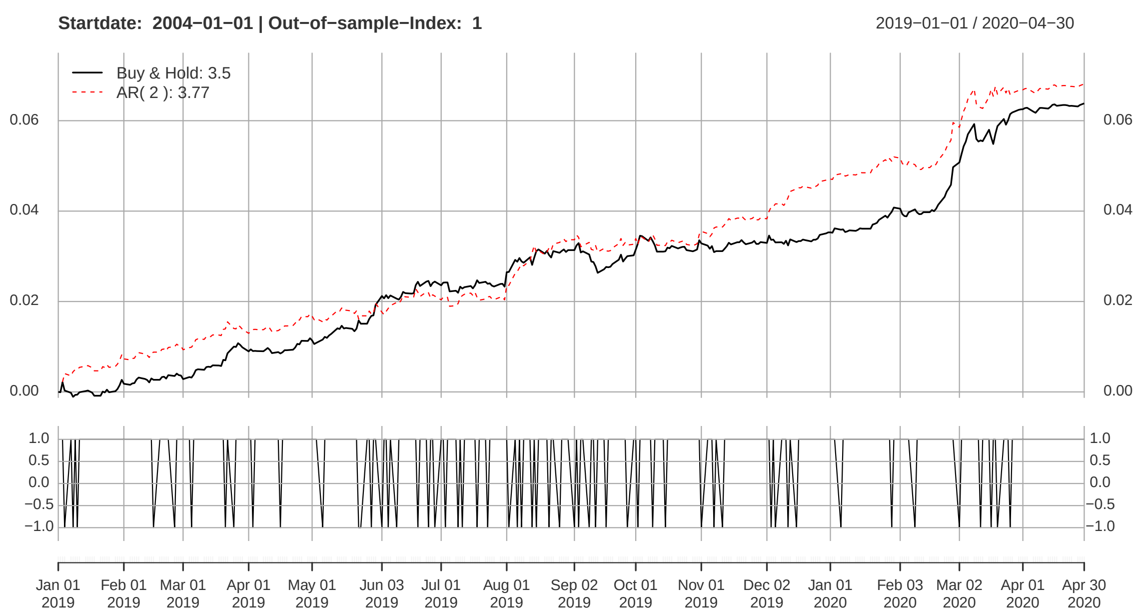

Figure 16: Visualization of the optimal AR(2)-Model.

The upper part of figure 16 shows the optimal model compared to buy and hold. The AR(2) has a higher sharpe than buy and hold. Regarding performance, it should be mentioned that the sharpe optimization of the algorithm is based on daily trades. The signals for a trade are newly generated with each daily forecast and could lead to high trading costs (positive signals = long, negative signals = short).

The lower part of figure 16 shows the trading signals visually. If one were to trade exactly according to the AR(2) forecasts, one would make a trade at each vertical line, a total of 111 trades would have to be made during out-of-sample periods. If you want to include trading costs in the performance, an AR model with daily trades would have to be much better than this one to outperform buy-and-hold.

But what influence do the trading costs really have? If you assume a fictitious investment amount of 1 million USD and you have to pay 190 USD for each trade, as mentioned in section 3.2.1. Invest this amount at the beginning of the out-of-sample period. With the buy-and-hold strategy, a return of USD 63832 is achieved by the end of the out-of-sample range. Without trading costs the AR(2) would yield USD 68130, with trading costs only USD 47040. To get at least the same return on the investment, the trading costs should not exceed USD 39. Note that in addition to the trading costs, the initial investment amount also matters.

Thus it can be said that for index 1, the buy-and-hold strategy cannot be outperformed by an AR(2) model. Nevertheless, one would like to apply this procedure also to the other indexes, possibly the generated models perform better for the other time-series.

Table 7: Optimal AR-Models.

	StartDate	AR-Sharpe	p	B&H-Sharpe	Trades
Index 1	2018-01-01	3.765	2	3.505	111
Index 2	2004-01-01	3.100	1	2.476	19
Index 3	2005-01-01	2.528	1	2.155	21
Index 4	2003-01-01	2.218	1	1.945	9

Table 7 shows the solutions of all 4 indexes. The algorithm finds an optimal AR model with model order 1 for all 3 remaining time-series. In contrast to index 1, it is noticeable that a much longer in-sample range is required for the optimal model. This also makes sense because the indexes show an increasing trend during the complete in-sample range. Index 1 is less volatile and has a less strong trend compared to the others, so a strong increase in the trend (spring 2018) has a greater effect on model performance. The other 3 series are more volatile, have larger peaks and therefore the model performance is less influenced by single events (like the increasing trend in spring 2018). The number of trades is much smaller for all models than for the first one. The out-of-sample performances are better than the corresponding buy-and-hold performances for all AR models.

Because of the trading costs, it has been seen with index 1 that the AR model, despite better performance, generates a lower return than buy-and-hold. But what about the other indexes. You choose the same investment amount and trading costs as before and you get the following table 8.

Table 8: Out-of-Sample Returns.

	Buy & Hold	AR without tradingcost	AR with tradingcost
Index 1	63832	68130	47040
Index 2	106055	131723	128113
Index 3	143291	167284	163294
Index 4	184657	209858	208148

Despite the trading costs, daily trading with the indexes 2-4 generates a higher return than the passive buy and hold strategy.

Figure 17 shows the out-of-sample performance of the second index. Right at the beginning of the series, the AR model signals a trade (one goes short and immediately long again). This trade immediately leads to an increase in performance which already outperforms the passive strategy. Further signaled trades will only occur again later. During the Corona crisis (spring 2020), one can see several close consecutive trades. The volatility increases sharply during this period, the AR model catches this event well and remains positive.

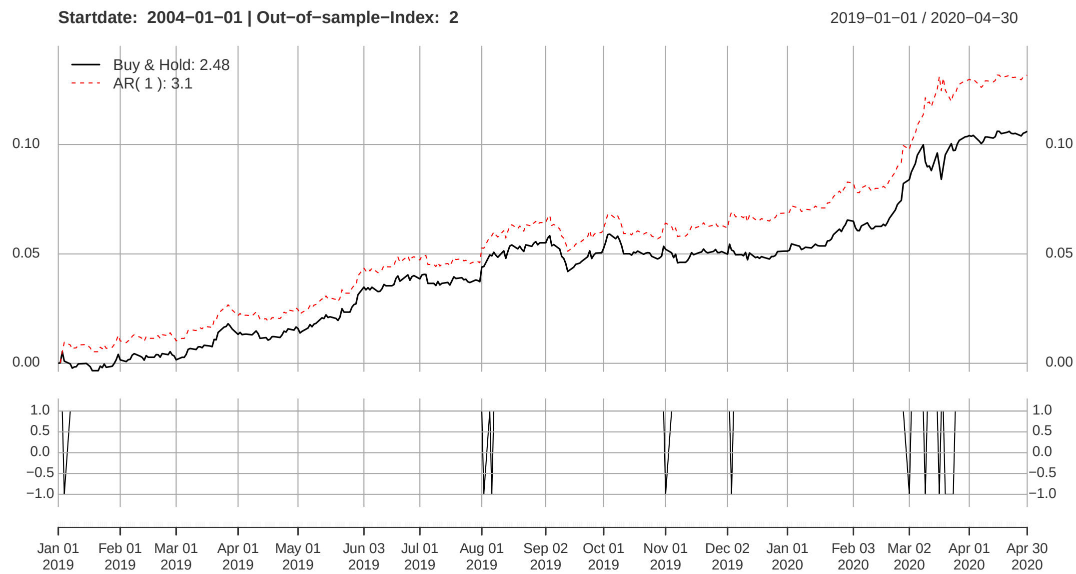

Figure 17: Visualization of the optimal Index 2 AR(1)-Model.

The optimal AR model for the third index shows almost the same behavior as index 2 (as seen in figure 18), the only difference being an additional trade in spring 2019.

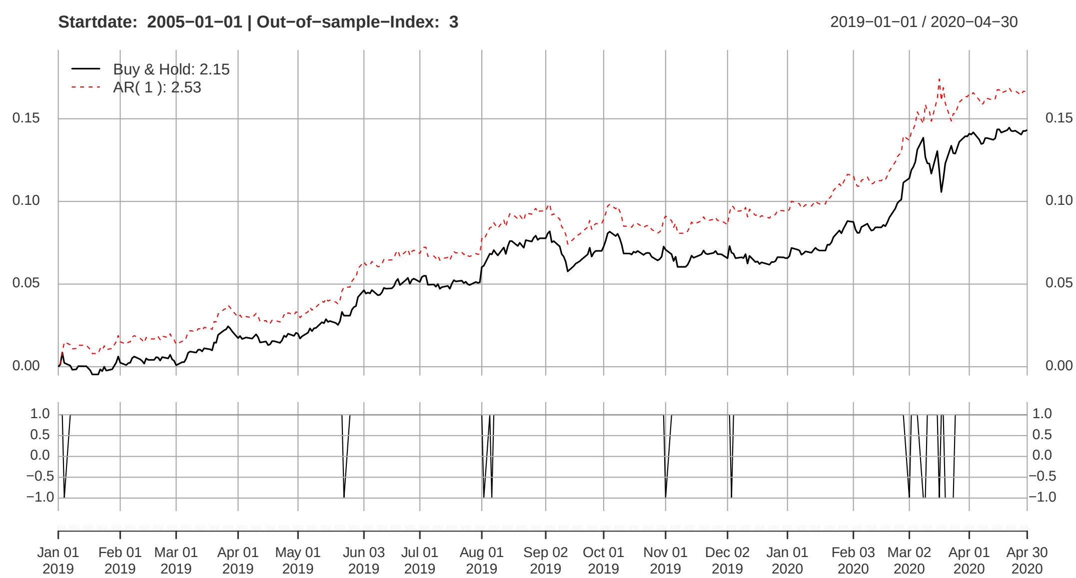

Figure 18: Visualization of the optimal Index 3 AR(1)-Model..

The fourth index, which is shown in the lower figure 18, is interesting again. The optimal AR model predicts almost the same behavior as buy-and-hold for almost the entire duration. Only during the highly volatile phase of the Corona crisis are the same trades listed as for index 2 and 3.

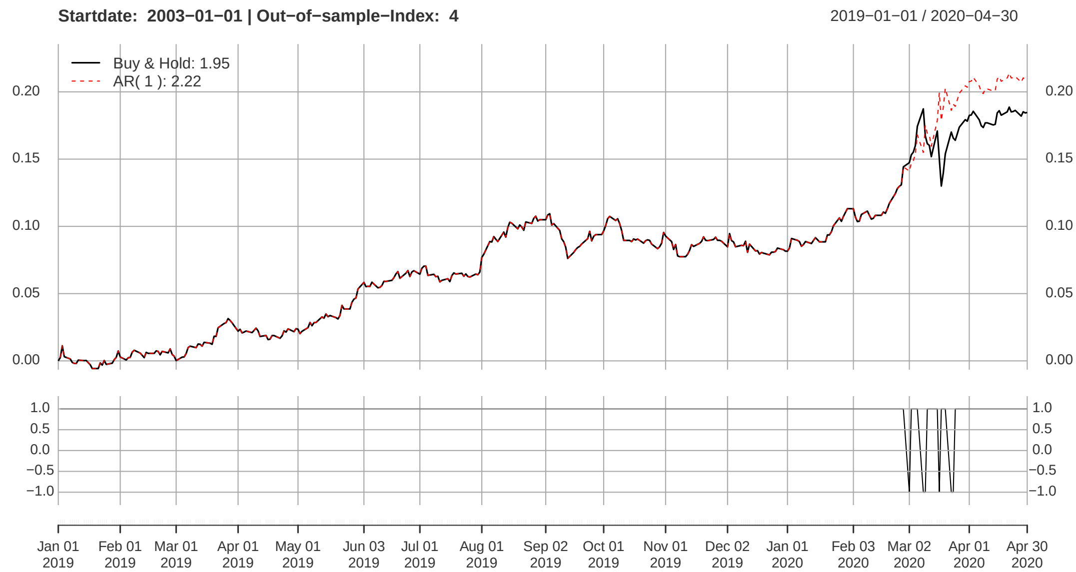

Figure 19: Visualization of the optimal Index 4 AR(1)-Model.

So it can be said that with very simple autoregressive models, a tool has been found with which it is possible to outperform the popular buy-and-hold strategy. Consider that this is only the case in this example with defined conditions. Depending on how high the trading costs and the amount of investment are, the returns can vary greatly. In addition to the buy-and-hold strategy, where you only buy and hold, you have to trade effectively with an active trading strategy.

3.2.3. Single Moving average trading

SMA undso

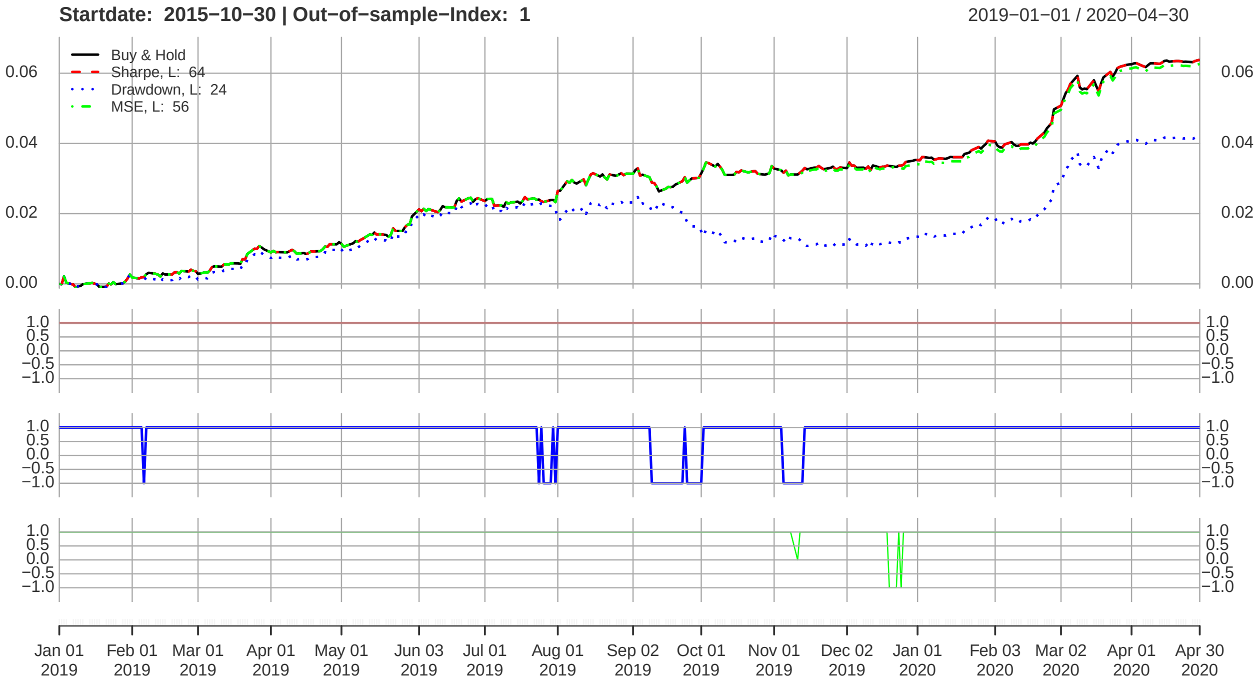

Figure 20: Placeholder

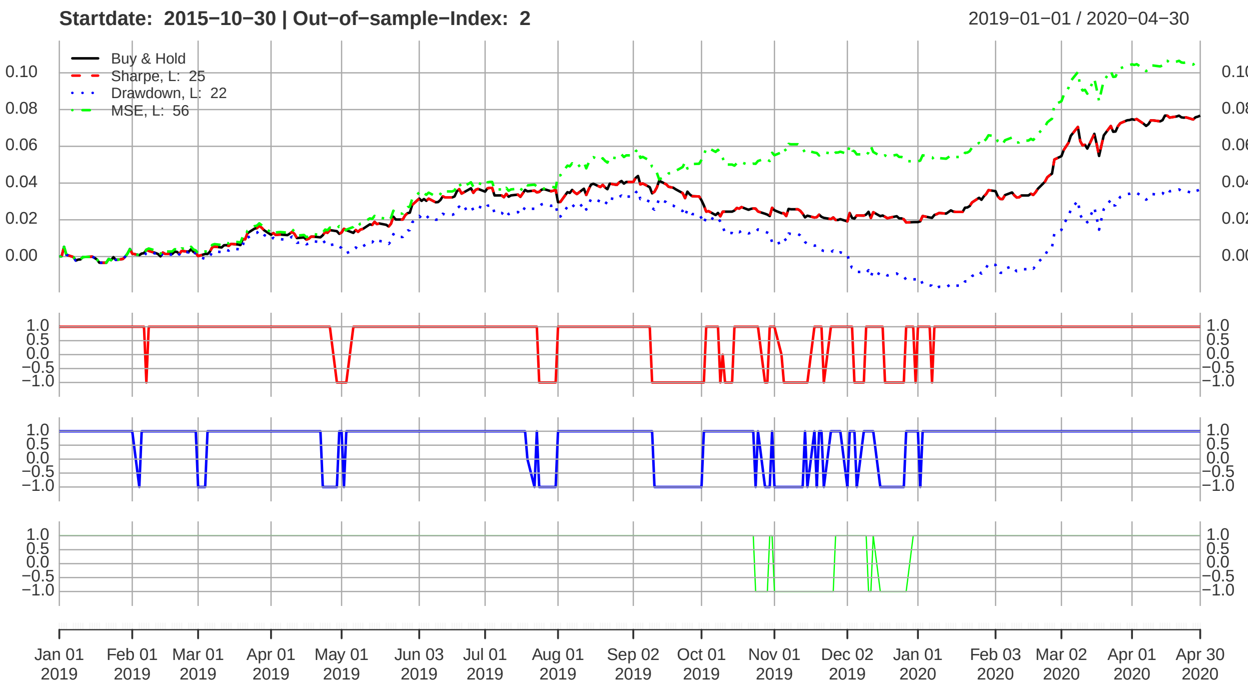

Figure 21: Placeholder

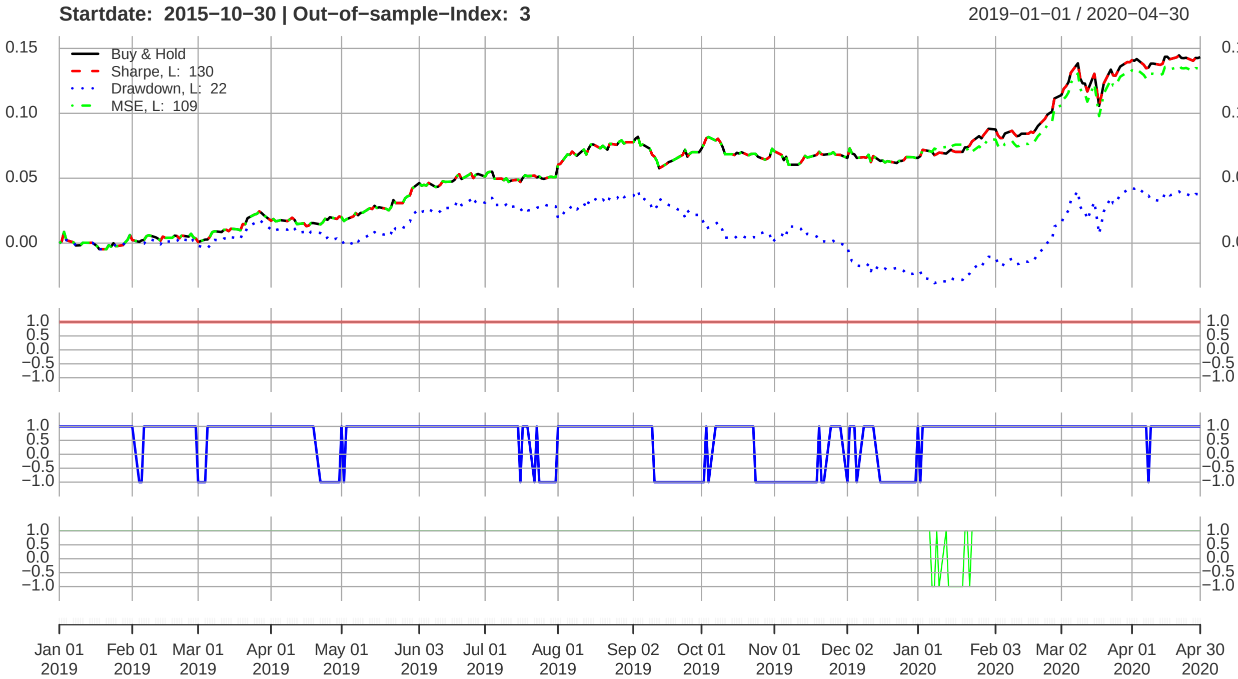

Figure 22: Placeholder

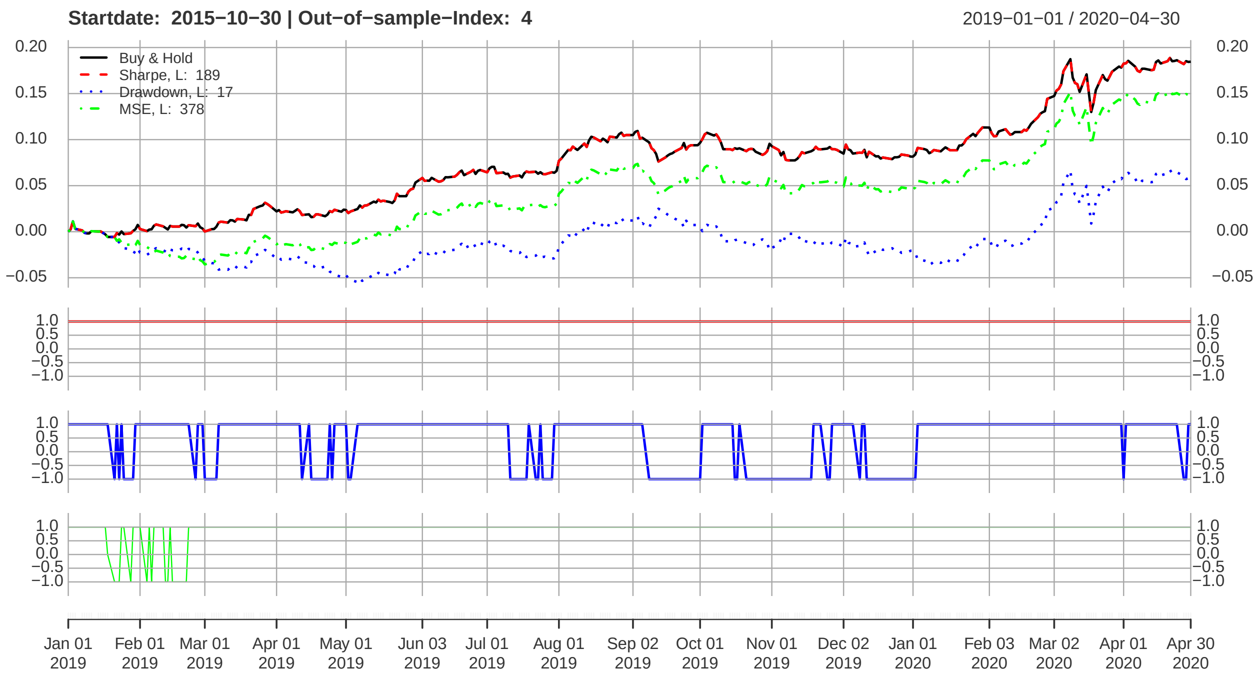

Figure 23: Placeholder

3.2.4. Moving average crossings trading

As a second approach we apply 2 SMAs for every time series, trade them with the common Ma crossings rules and optimize them with sharpe and max drawdown For the optimization we consider the out of sample timespan from 2019-01-01 to march 2020. The initialization of a filter has the same timespan as the length of filter itself, therefore we add an in-sample timespan to iniciate the filters. The insample

The intuition says that this strategy may not be so bad , why leaving a trend market instead of buy and hold.

3.2.2. Moving average crossings to trade

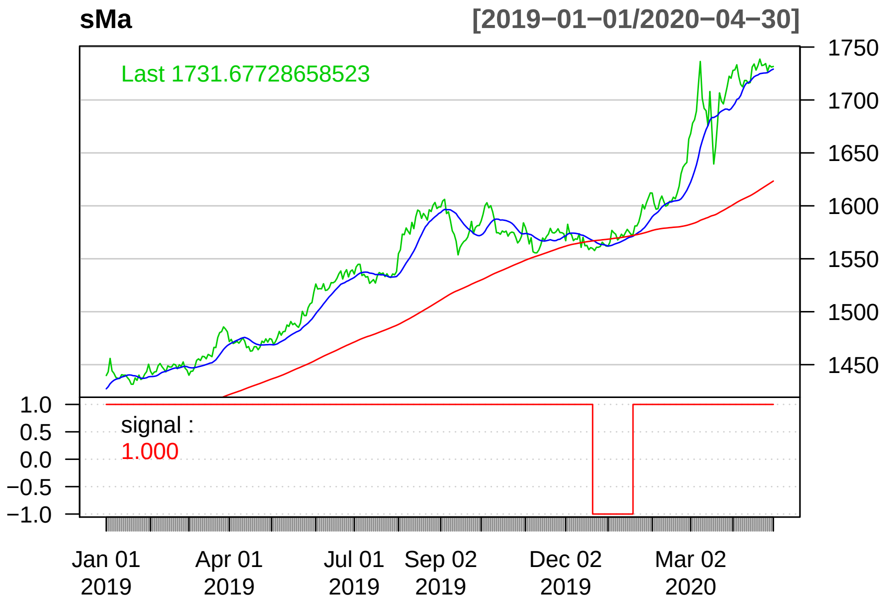

Figure 24: conversion data

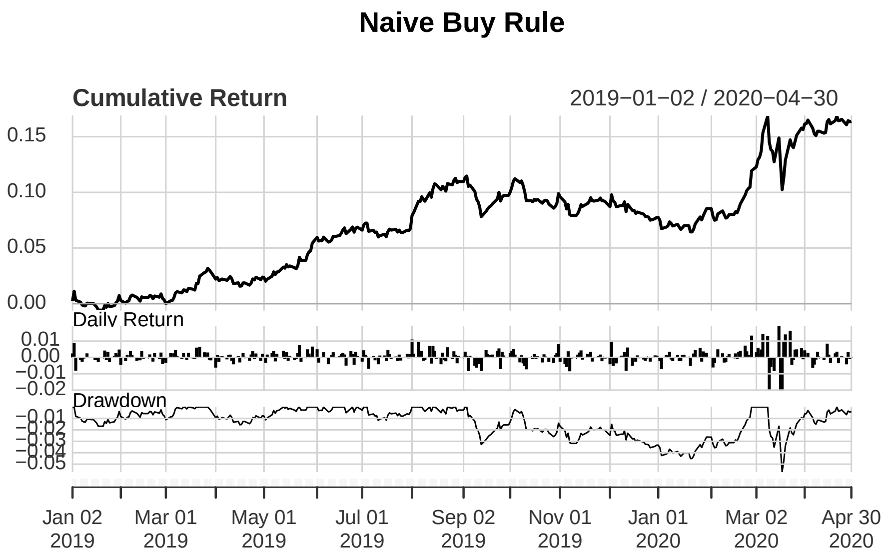

Figure 25: conversion data

4. Results

4.1

4.1 Key Take Away

Here's a short list of key takeaways from this project work

5. References

[1] M. Wildi, *Econometrics 1: Time series analysis*. Winterthur: ZHAW, 2017, p. 221.

[2] M. Wildi, *Econometrics 3: Conditional heteroscedasticity models, part 1*. ZHAW, 2020, p. 56.

[3] A. H. A. L. Kholkin D. A. Kiselev, *Encyclopedia of materials: Science and technology (second edition)*. Elsevier, 2001, p. 10388.

[4] M. Masum, "Time series analysis: Identifying ar and ma using acf and pacf plots." https://towardsdatascience.com/identifying-ar-and-ma-terms-using-acf-and-pacf-plots-in-time-series-forecasting-ccb9fd073db8 (accessed Nov. 16, 2020).

[5] B. Williams, *GARCH(1,1) models*. Ruprecht-Karls-Universität Heidelberg, 2011, p. 42.

[6] M. Wildi, *An introduction to conditional volatility models*. Winterthur: ZHAW, 2020, p. 32.

[7] M. Wildi, *Econometrics 3: Conditional heteroscedasticity models, part 3*. ZHAW, 2020, p. 53.

[8] A. Hayes, "Simple moving average." https://www.investopedia.com/terms/s/sma.asp (accessed Sep. 22, 2020).

[9] A. Hayes, "Exponential moving average." https://www.investopedia.com/terms/e/ema.asp (accessed Sep. 10, 2020).

[10] M. Wildi, *Econometrics 2: Trading indicators*. ZHAW, 2020, p. 78.

[11] J. Fernando, "Relative strength index (rsi)." https://www.investopedia.com/terms/r/rsi.asp (accessed Nov. 30, 2020).

[12] swissquote, "Trading fees." https://library.swissquote.com/shared-images/brochure-trading-pricing-bank-de?_ga=2.183413157.222455671.1606673517-1327743316.1606673517 (accessed Nov. 29, 2020).

Attachment