Homework 2

Filippo Fortugno and Lucas Santos

June 26, 2023

1 Question I

Design ONE trading strategy (i.e., the signaling sequence) different from the MAcross strategy shown in Lecture 6, using any (or all) sentimental indicators $ev \in$ {positiveP, negativeP, BULL, BEAR, BBr, PNlog}, and RVT (volume of news), and test the performance of your strategy for the cases where you either allow long-only or long-short trading, and using the rolling window analysis with window sizes of 254 (1 year) or 127 (6 months). Any other parameters of your strategy must be tuned.

1.1 Initial steps

The first step of the question is decide in which stocks we are going to test our strategy. We decided to restrain our analysis to stocks instead of currencies because there were more data available (602 observations instead of 182), in order to try to create a strategy that work in different scenarios we selected the two stocks that are less correlated. According to the figure 1 the two stocks that are less correlated are Hong Kong and Shanghai Banking Corporation (HSBC) and Microsoft (MSFT) which makes sense since the first one is a financial company, usually with more stability in the results, while the other one is a technology company with high volatility.

Figures 2 and 3 show us that there are no NA's in the *Adj Close* column, just in the indicators column. The missing values of the indicators were filled using the forward filling method that fills the values with the previous data.

1.2 Power of explanatory variables

We opted to perform a test in the bearish and bullish indicators available to see if they increase the forecast power of the models. The logic behind this is, if these variables can help us forecast better they can be used for trading strategies we will expose later on this report.

For test the explanatory power of the bearish and bullish sentiment variables we fit three machine learning models: Support Vector Machine (SVM), Neural Network (NNet), and Gaussian Process (GP). We define two epochs, "epoc1" and "epoc2," which represent different time periods for analysis. The models are trained on the training set, and their predictions are generated for the testing set. The normalized root mean squared error (NRMSE) is calculated to evaluate the performance of each model in each epoch and exposed on tables 1 and 2.

We also included a plot that compares the true values of the price with the predictions made by SVM, NNet, and GP models for the last epoch for the bearish variables (figures 4 and 6) and only the lagged price (figures 5 and 7).

The results shows that the sentiment variables didn't increased the performance of the forecasting task compared to use only the lags of the prices for both stocks, indicating that the lag

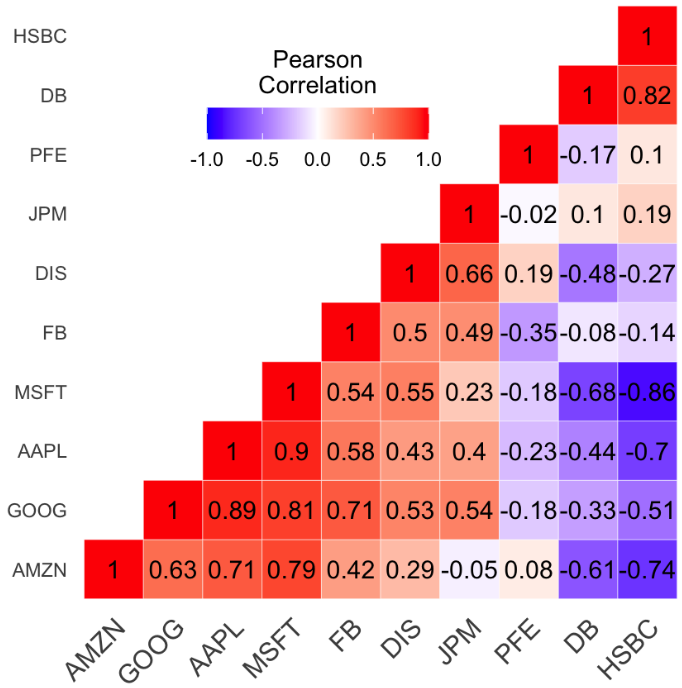

Figure 1: Correlation plot

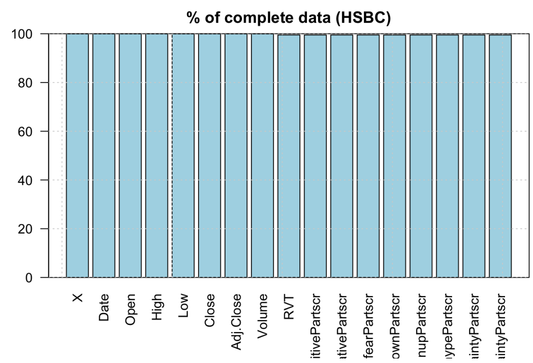

Figure 2: Percentage of complete data of HSBC dataset

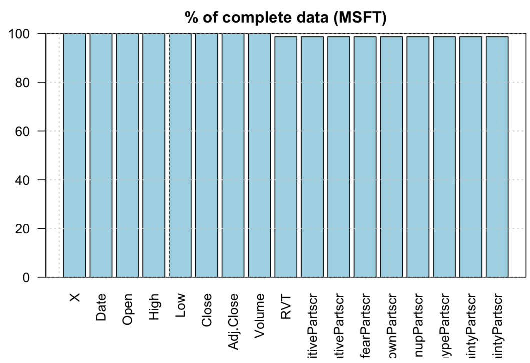

Figure 3: Percentage of complete data of MSFT dataset

	Model	NNET	SVM	GP
	Lags	0.629	0.628	0.700
Epoc 1	Bearish	0.687	0.689	0.683
	Bullish	0.690	0.699	0.700
	Lags	0.217	0.213	0.232
Epoc 2	Bearish	0.235	0.242	0.255
	Bullish	0.247	0.245	0.296

Table 1: NRMSE for HSBC

	Model	NNET	SVM	GP
	Lags	0.312	0.317	0.445
Epoc 1	Bearish	0.327	0.333	0.361
	Bullish	0.350	0.330	0.356
	Lags	0.415	0.412	0.800
Epoc 2	Bearish	0.405	0.401	0.400
	Bullish	0.415	0.414	0.421

Table 2: NRMSE for MSFT

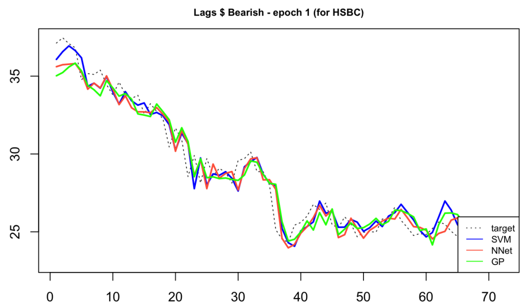

Figure 4: Forecasting using the lags of bearish variables for HSBC

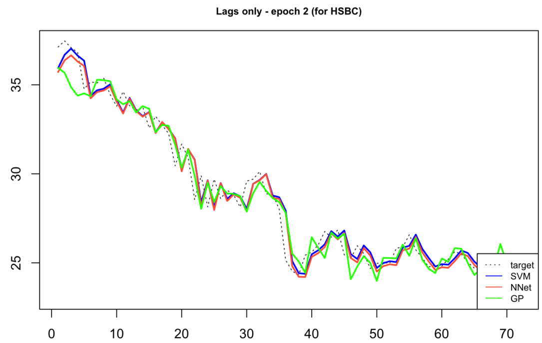

Figure 5: Forecasting using the lags of price for HSBC

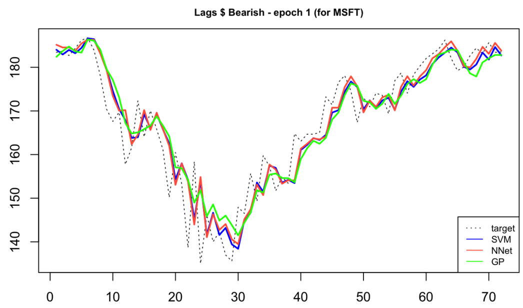

Figure 6: Forecasting using the lags of bearish variables for MSFT

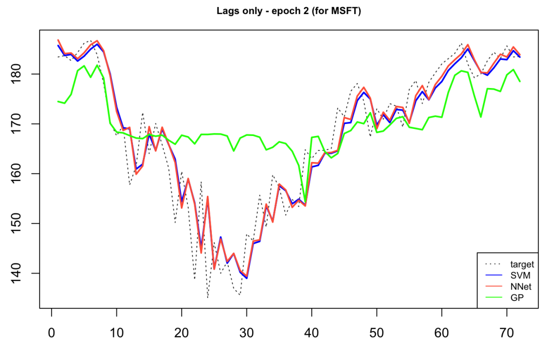

Figure 7: Forecasting using the lags of price for MSFT

of the prices is the one that carry the most relevant information for predict the stock prices in comparison to the sentiment indexes. In the next session we will compare the performance of strategies build with only the price and the sentiment indexes and observe if the pattern persists.

1.3 Strategies

As proposed in the challenge, we backtested strategies existent in literature on the stocks we selected. Below is the intuition behind it and how we constructed.

1.3.1 Moving Average Convergence Divergence - MACD

The MACD (Moving Average Convergence Divergence) strategy is a popular technical analysis indicator used in trading. The strategy aims to identify potential buy and sell signals based on the convergence and divergence of two moving averages of a stock's price.

In the testMACDStrategy function specified below, the MACD indicator is calculated using the MACD function from an underlying package. The MACD consists of three components: the MACD line, the signal line, and the MACD histogram. The MACD line is calculated as the difference between two exponential moving averages (EMAs) of different periods, typically the short period and the long period. The signal line is an EMA of the MACD line with a specific period.

To generate trading signals, the function compares the MACD line and the signal line. If the MACD line is greater than the signal line, it suggests a bullish signal, and a value of 1 is assigned to the macd_signal variable. Otherwise, if the MACD line is less than or equal to the signal line, it suggests a bearish signal, and a value of longshort (defaulted to 0) is assigned to the macd_signal. The Lag function is then used to shift the signals by one day to avoid lookahead bias.

The resulting macd_signal is a series of buy (1) and sell (0 or longshort) signals based on the MACD strategy. The function calculates the performance of the MACD strategy by multiplying the macd_signal with the daily returns of the stock. The Performance function calculates the performance metrics based on the trading signals, considering transaction costs (tcost) if provided.

The testMACDStrategy function merges the MACD returns with the benchmark (BH) returns and returns a performance summary that includes the MACD strategy's performance metrics (performance_macd) and the benchmark's performance metrics. The performance metrics can include measures such as cumulative returns, annualized returns, risk-adjusted returns, and drawdown analysis. We let the function available below to those interested in apply on their trading strategies.

```
  testMACDStrategy <- function(myStock, ts = myStock, short_period = 12,
                                     long_period = 26, signal_period =
                                     9, longshort = 0, tcost = 0) {

 # Calculate MACD
 macd <- MACD(ts, nFast = short_period, nSlow = long_period, nSig =
                                    signal_period)
 macd_line <- macd$macd
 signal_line <- macd$signal
```

```r
  # Create signal: MACD strategy
  macd_signal <- Lag(ifelse(macd_line > signal_line, 1, longshort), 1)
  macd_runs <- rle(as.vector(na.omit(macd_signal)))
  macd_lruns <- length(macd_runs$lengths)

  # Calculate returns
  bmkReturns <- dailyReturn(myStock, type = "arithmetic")
  macd_returns <- bmkReturns * macd_signal

  # Set column names
  names(bmkReturns) <- 'BH'
  names(macd_returns) <- 'MACD'

  # Merge returns and calculate performance
  tt <- na.omit(merge(bmkReturns, macd_returns))
  performance_macd <- Performance(tt$MACD, macd_lruns, tcost)

  # Return performance results
  result <- cbind(performance_macd, BH = Performance(tt$BH, 2, tcost))
  return(result)
}
```

The intuition behind **the MACD strategy is that it seeks to capture trends and potential reversals in a stock's price by identifying when the shorter-term moving average crosses above or below the longer-term moving average**. The crossover of the MACD line and the signal line is often considered a signal of a potential change in the stock's direction. Traders may interpret a bullish signal as an opportunity to buy or hold the stock, while a bearish signal may indicate a potential sell or short-selling opportunity.

The MACD strategy and its variations have been widely studied and discussed in technical analysis literature, some of the seminal references about MACD are (APPEL; DOBSON, 2007) and (MURPHY JOHN, 1999).

1.3.2 TMA

We also make available the testTMAStrategy function, which implements the Triple Moving Average (TMA) strategy to generate trading signals. The TMA strategy uses three moving averages of different periods to determine buy and sell signals. Here's how it works:

1. Calculation of TMA:
   - Three moving averages are calculated using the Simple Moving Average (SMA) function: tma_short, tma_medium, and tma_long.
   - tma_short represents the moving average with a shorter period.
   - tma_medium represents the moving average with a medium period.
   - tma_long represents the moving average with a longer period.
2. Generation of signals:
   - The TMA strategy compares the values of the three moving averages.
   - If tma_short is greater than tma_medium, and tma_medium is greater than tma_long, it suggests a bullish signal.

- In this case, a value of 1 is assigned to the tma_signal variable, indicating a buy signal.
- Otherwise, if the condition is not met, a value of longshort (defaulted to 0) is assigned to the tma_signal, indicating a sell signal or no position.

3. Performance calculation:
   - The function multiplies the tma_signal series with the daily returns of the stock to calculate the returns of the TMA strategy.
   - Transaction costs (tcost) can be considered in the performance calculation if provided.
   - The Performance function calculates the performance metrics based on the trading signals, taking into account the length of consecutive signals (tma_runs) and transaction costs.

```
    testTMAStrategy <- function(myStock, ts = myStock, tma_short_period = 5,
                                       tma_medium_period = 20,
                                      tma_long_period = 50, longshort =
                                      0, tcost = 0) {

  # Calculate TMA
  tma_short <- SMA(ts, n = tma_short_period)
  tma_medium <- SMA(ts, n = tma_medium_period)
  tma_long <- SMA(ts, n = tma_long_period)

  # Create signal: TMA strategy
  tma_signal <- Lag(ifelse(tma_short > tma_medium & tma_medium > tma_long, 1
                                      , longshort), 1)
  tma_runs <- rle(as.vector(na.omit(tma_signal)))
  tma_lruns <- length(tma_runs$lengths)

  # Calculate returns
  bmkReturns <- dailyReturn(myStock, type = "arithmetic")
  tma_returns <- bmkReturns * tma_signal

  # Set column names
  names(bmkReturns) <- 'BH'
  names(tma_returns) <- 'TMA'

  # Merge returns and calculate performance
  tt <- na.omit(merge(bmkReturns, tma_returns))
  performance_tma <- Performance(tt$TMA, tma_lruns, tcost)

  # Return performance results
  result <- cbind(performance_tma, BH = Performance(tt$BH, 2, tcost))
  return(result)
}
```

The resulting performance metrics for the TMA strategy are calculated and compared to the benchmark returns (BH) as we did in the MACD strategy function.

Regarding academic references for the TMA strategy, while specific academic works were not mentioned in the code, the concept of using multiple moving averages to generate trading signals is a common approach in technical analysis. You can refer to textbooks and academic literature on technical analysis, such as "Technical Analysis of the Financial Markets" by John J. Murphy, "Evidence-Based Technical Analysis" by David Aronson, or "Encyclopedia of Chart Patterns" by Thomas N. Bulkowski, for more in-depth information and studies related to moving average-based trading strategies.

A non-exhaustive list of references that talk about the TMA strategy and moving average-based trading strategies in general are (ARONSON, 2011), (MURPHY JOHN, 1999) and (BULKOWSKI, 2021).

1.4 Results

In the following we just display the strategies and the parameters that obtained the best results. We simulated with different parameters and indicators both strategies, MACD and TMA, for both stocks, HSBC and MSFT.

We tested the following indicators:

- **BULL** - Bullish sentiment indicator
- **BEAR** - Bearish sentiment indicator
- **Bull Bear Ratio (BBr)** - Sentiment indicator that captures factors of bullish and bearish sentiments

1.4.1 HSBC

For the MACD all the strategies that had a trading cost on it performed worse than the benchmark. It's possible to see from table 5 that this strategy is characterized by a big amount of trades, so any increase in the trading cost has a significant impact on the return which makes the strategy not suitable for real world applications. All the strategies that allowed long and short performed better for the MACD in the HSBC.

The TMA results displayed in table 4 were better in comparison to the ones obtained with the MACD strategy. Even using the same parameters (indicator BBr, short allowed and trading cost = 0) the cumulative return obtained jumped from 10% to 0.94%.

1.4.2 MSFT

The best MACD strategy in terms of performance for the MSFT stock was the one that use the BEAR indicator, not allowed short selling and without trading costs. However, even with a cumulative return of 70% the strategy didn't outperformed the buy-and-hold strategy which returned 104% in the period. Following the pattern observed with HSBC the strategies with trading cost, even though minimal ($tcost = 0.005$) performed really badly compared to those without trading costs.

For TMA the results were even worse than the MACD for the MSFT stock, losing for the buy-and-hold strategy by even more. Both strategies selected didn't contain trading costs so the bad performance of the strategy it's attributed to the poor explanatory power of the indicators on this strategy.

	performance_macd	BH
Cumulative Return	0.02	-0.49
Annual Return	0.01	-0.26
Annualized Sharpe Ratio	0.04	-1.12
Win %	0.51	0.49
Annualized Volatility	0.23	0.23
Maximum Drawdown	-0.21	-0.49
Max Length Drawdown	353.00	566.00
n.trades	131.00	2.00

	performance_macd	BH
Cumulative Return	0.10	-0.49
Annual Return	0.04	-0.26
Annualized Sharpe Ratio	0.18	-1.12
Win %	0.52	0.49
Annualized Volatility	0.23	0.23
Maximum Drawdown	-0.27	-0.49
Max Length Drawdown	169.00	566.00
n.trades	144.00	2.00

Table 3: MACD results for HSBC - BULL indicator (left) and HSBC - BBr indicator (right), short allowed, and trading cost = 0

	performance_tma	BH
Cumulative Return	0.57	-0.47
Annual Return	0.22	-0.25
Annualized Sharpe Ratio	0.95	-1.06
Win %	0.50	0.49
Annualized Volatility	0.23	0.23
Maximum Drawdown	-0.19	-0.49
Max Length Drawdown	140.00	505.00
n.trades	68.00	2.00

	performance_tma	BH
Cumulative Return	0.94	-0.51
Annual Return	0.34	-0.27
Annualized Sharpe Ratio	1.48	-1.18
Win %	0.52	0.49
Annualized Volatility	0.23	0.23
Maximum Drawdown	-0.25	-0.51
Max Length Drawdown	160.00	567.00
n.trades	77.00	2.00

Table 4: TMA results for HSBC - BEAR indicator (left) and HSBC - BBr indicator (right), short allowed, and trading cost = 0

	performance_macd	BH
Cumulative Return	0.50	1.04
Annual Return	0.20	0.37
Annualized Sharpe Ratio	0.84	1.13
Win %	0.55	0.57
Annualized Volatility	0.23	0.33
Maximum Drawdown	-0.18	-0.28
Max Length Drawdown	138.00	113.00
n.trades	114.00	2.00

	performance_macd	BH
Cumulative Return	0.70	1.04
Annual Return	0.27	0.37
Annualized Sharpe Ratio	1.19	1.13
Win %	0.59	0.57
Annualized Volatility	0.22	0.33
Maximum Drawdown	-0.16	-0.28
Max Length Drawdown	70.00	113.00
n.trades	115.00	2.00

Table 5: MACD results for MSFT - BULL indicator (left) and MSFT - BEAR indicator (right), short not allowed, and trading cost = 0

	performance_tma	BH
Cumulative Return	0.32	1.16
Annual Return	0.13	0.40
Annualized Sharpe Ratio	0.76	1.22
Win %	0.56	0.57
Annualized Volatility	0.17	0.33
Maximum Drawdown	-0.17	-0.28
Max Length Drawdown	171.00	113.00
n.trades	63.00	2.00

	performance_tma	BH
Cumulative Return	0.31	1.03
Annual Return	0.13	0.37
Annualized Sharpe Ratio	1.27	1.11
Win %	0.61	0.57
Annualized Volatility	0.10	0.33
Maximum Drawdown	-0.10	-0.28
Max Length Drawdown	164.00	113.00
n.trades	60.00	2.00

Table 6: TMA results for MSFT - BEAR indicator (left) and MSFT - BBr indicator (right), short not allowed, and trading cost = 0

1.4.3 Rolling window

We selected the best parameters for each strategy and tested it with a rolling window of 252 days. The results for HSBC are displayed on tables 7 and 8, while the results for the MSFT stock are displayed on tables 9 and 10. The table shows the average of the results in different rolling windows.

For the HSBC stock the mean performance was positive for both strategies, showing that these strategies are effective to obtain better results than buy-and-hold. However, it is important to emphasize that both strategies are without trading costs, which has a significantly impact specially on the MACD strategy.

For the MSFT stock both strategies performed poorly than the simple buy-and-hold strategy as we observed in the non-rolling window simulations. However, the gap between the strategies is smaller in the rolling-window average than in the simple strategies

1.4.4 Conclusion

General conclusions:

- The parameters and variables that perform one in one strategy also performs well on the other strategy;
- Both strategies are really sensitive to trading costs, especially the MACD because it's a strategy with high number of trades;
- The strategies tended to performed better than the buy-and-hold strategy for HSBC and worse for MSFT. One of the possible explanations is the poorly performance of HSBC Buy-and-Hold and the good performance of this strategy on MSFT;
- The rolling window simulations doesn't change a lot the general trend we observed in the plain simulations. Strategies that outperform the buy-and-hold continue to outperform (HSBC case) while strategies that don't outperform (MSFT case) continue to don't outperform buy-and-hold, this shows how robustness in the strategies;
- While TMA strategy performed better for HSBC stock, the MACD was more profitable for the MSFT compared to the TMA. The fact of selecting stocks with low degree of correlation could explain the difference of the results.

	Cumulative Return	Annual Return	Annualized Sharpe Ratio	Win %	Annualized Volatility	Maximum Drawdown	Max Length Drawdown	n.trades
BH	-0.12	-0.12	-0.57	0.50	0.18	-0.23	200.39	1.00
MACD	0.18	0.18	1.06	0.54	0.18	-0.13	90.90	1.00

Table 7: MACD Rolling window for HSBC - BBr indicator, short allowed (316 windows)

	Cumulative Return	Annual Return	Annualized Sharpe Ratio	Win %	Annualized Volatility	Maximum Drawdown	Max Length Drawdown	n.trades
BH	-0.12	-0.12	-0.57	0.50	0.18	-0.23	200.26	1.00
TMA	0.08	0.08	0.35	0.51	0.18	-0.16	136.61	1.00

Table 8: TMA Rolling window for HSBC - BBr indicator, short allowed (314 windows)

	Cumulative Return	Annual Return	Annualized Sharpe Ratio	Win %	Annualized Volatility	Maximum Drawdown	Max Length Drawdown	n.trades
BH	0.38	0.38	1.53	0.57	0.27	-0.18	79.68	1.00
MACD	0.29	0.29	1.53	0.60	0.20	-0.09	58.51	1.00

Table 9: MACD Rolling window for MSFT - BEAR indicator, short not allowed (316 windows)

	Cumulative Return	Annual Return	Annualized Sharpe Ratio	Win %	Annualized Volatility	Maximum Drawdown	Max Length Drawdown	n.trades
BH	0.38	0.38	1.54	0.57	0.27	-0.18	79.88	1.00
TMA	0.15	0.15	1.06	0.56	0.15	-0.11	115.13	1.00

Table 10: TMA Rolling window for MSFT - BEAR indicator, short not allowed (320 windows)

2 Question II

Choose 9 stocks from the dataset dataset.rds for Lec. 8 on Factor Models and Sentiment, and compute for these the Robust (ellipsoid) Global Maximum Return Portfolio using as perturbation matrix Sigma (S) corresponding to the factor models:1) the Fama- French 3-factors returns 3FF; and 2) the Sentiment indicator PNlog factor model. Use a two consecutive years of data for experiments selected from 2015-01 to 2020-06; register this period of data in column Prob. 2 of sheet HW2 of the spreadsheet ML4Finance23 HWteams in the drive. Try to select different 2-years period from those selected by others (not mandatory). Try different kappas (at least 3 values in (0,1)) and multiple robust noisy solutions to check for sensitivity. Comment on the differences/similarities of results for both cases of Sigma.

Firstly, we select 9 different stocks from the given dataset. The selection is based on the rate of missing values in the sentiment scores data (PNlog). We aim to have our data as complete as possible to minimize bias in the results. Accordingly, we chose the following 9 stocks: "AAPL", "AMZN", "DB", "DIS", "FB", "GOOG", "PFE", "JPM", "MSFT". With the same criterion, we select the range from 2015-01-02 to 2016-12-31.

The log returns of the selected stocks are modelled using both the Fama-French 3-factor model and a 1-factor model based on the Sentiment Index. For each model, we estimate the parameters (alphas and betas), calculate residuals, fitted values, and estimate the covariance of the residuals and the expected returns. After that, we calculate the perturbation matrix Sigma using the residuals obtained from the previous step. We report the results for the perturbation matrix Sigma for the Fama-French 3-factor model and a 1-factor model based on the Sentiment Index.

	AAPL	AMZN	DB	DIS	FB	GOOG	PFE	JPM	MSFT
AAPL	2.51129e-04	1.13606e-04	1.32521e-04	8.00053e-05	1.08780e-04	1.01524e-04	7.86677e-05	1.03730e-04	1.16648e-04
AMZN	1.13606e-04	3.89568e-04	1.24837e-04	8.35905e-05	1.20694e-04	1.12797e-04	8.44553e-05	1.00153e-04	1.26445e-04
DB	1.32521e-04	1.24837e-04	7.49738e-04	1.23987e-04	1.22111e-04	1.15936e-04	1.05116e-04	2.18376e-04	1.55045e-04
DIS	8.00053e-05	8.35905e-05	1.23987e-04	1.61594e-04	8.04354e-05	7.55252e-05	6.11577e-05	9.53016e-05	9.07705e-05
FB	1.08780e-04	1.20694e-04	1.22111e-04	8.04354e-05	2.85565e-04	1.07210e-04	8.14336e-05	9.65080e-05	1.20254e-04
GOOG	1.01524e-04	1.12797e-04	1.15936e-04	7.55252e-05	1.07210e-04	2.44999e-04	7.54353e-05	9.31371e-05	1.13720e-04
PFE	7.86677e-05	8.44553e-05	1.05116e-04	6.11577e-05	8.14336e-05	7.54353e-05	1.48230e-04	8.07152e-05	8.72657e-05
JPM	1.03730e-04	1.00153e-04	2.18376e-04	9.53016e-05	9.65080e-05	9.31371e-05	8.07152e-05	2.18660e-04	1.23113e-04
MSFT	1.16648e-04	1.26445e-04	1.55045e-04	9.07705e-05	1.20254e-04	1.13720e-04	8.72657e-05	1.23113e-04	2.58962e-04

Table 11: Perturbation Matrix Sigma for Fama French 3 factors model

	AAPL	AMZN	DB	DIS	FB	GOOG	PFE	JPM	MSFT
AAPL	2.50752e-04	-4.52381e-08	-5.45397e-07	8.85638e-09	3.96154e-08	1.19751e-07	4.00349e-07	-1.13631e-07	2.07374e-07
AMZN	-4.52381e-08	3.88772e-04	1.79321e-07	-2.91189e-09	-1.30251e-08	-3.93730e-08	-1.31631e-07	3.73607e-08	-6.81826e-08
DB	-5.45397e-07	1.79321e-07	7.48492e-04	-3.51062e-08	-1.57033e-07	-4.74687e-07	-1.58696e-06	4.50427e-07	-8.22020e-07
DIS	8.85638e-09	-2.91189e-09	-3.51062e-08	1.61340e-04	2.54997e-09	7.70816e-09	2.57697e-08	-7.31421e-09	1.33483e-08
FB	3.96154e-08	-1.30251e-08	-1.57033e-07	2.54997e-09	2.85116e-04	3.44793e-08	1.15270e-07	-3.27171e-08	5.97082e-08
GOOG	1.19751e-07	-3.93730e-08	-4.74687e-07	7.70816e-09	3.44793e-08	2.44624e-04	3.48445e-07	-9.88989e-08	1.80489e-07
PFE	4.00349e-07	-1.31631e-07	-1.58696e-06	2.57697e-08	1.15270e-07	3.48445e-07	1.47995e-04	-3.30636e-07	6.03404e-07
JPM	-1.13631e-07	3.73607e-08	4.50427e-07	-7.31421e-09	-3.27171e-08	-9.88989e-08	-3.30636e-07	2.18783e-04	-1.71264e-07
MSFT	2.07374e-07	-6.81826e-08	-8.22020e-07	1.33483e-08	5.97082e-08	1.80489e-07	6.03404e-07	-1.71264e-07	2.58722e-04

Table 12: Perturbation matrix Sigma for Sentiment indicator PNlog factor model

From the perturbation matrices generated for both Fama-French 3-factor model and Sentiment Indicator (PNlog) factor model, we can see that they exhibit different characteristics. As regard to the first one (Perturbation Matrix Sigma for Fama-French 3-factor model), the matrix values are all non-zero which indicate that there is a covariance among all stocks. This is intuitive as these stocks are expected to react to the market factors captured in the Fama-French model. The diagonal elements represent the variance of each individual stock's returns. For instance, Apple (AAPL) has a variance of 2.511293e-04, indicating a lower risk in terms of volatility, as compared to Deutsche Bank (DB) that has a variance of 0.000749738, indicating a higher risk. The off-diagonal elements represent the covariance between each pair of stocks, providing information on how these stocks move together. This off-diagonal elements seem larger than the one of the Sentiment indicator model, suggesting that Fama French 3 factors model sees a higher degree of interdependence between different stocks, which could be due to the model considering multiple factors (market risk premium, SMB, HML) that commonly affect all stocks.

In the Sentiment Indicator model, these covariances are generally smaller in magnitude. Some are even close to zero. This could indicate that the Sentiment model perceives that the stock returns are less dependent on each other compared to the Fama-French model. It might also be due to the fact that the Sentiment model considers only a single factor which might not be influencing all the stocks in the same way.

Next, we compute the Robust (ellipsoid) Global Maximum Return Portfolio using the estimated perturbation matrices. We do this for different values of 'kappa' and multiple robust noisy solutions to check for sensitivity. In particular, 'kappa' is the risk aversion parameter in the portfolio optimization model. It balances the trade-off between the portfolio's expected return (which we aim to maximize) and the portfolio's risk, as measured by the ellipsoidal uncertainty set. By selecting different values of kappa, it allows us to examine how this trade-off plays out for different levels of risk tolerance: a low kappa value reflects an investor who is less risk averse; this scenario may lead to a portfolio that is more weighted in higher risk and return assets. On the contrary, a high value of kappa means an investor who is more risk-averse. This will likely result in a portfolio weighted more with lower risk assets.

The Robust (ellipsoid) Global Maximum Return Portfolio is computed for both the Fama-French 3-factors model and Sentiment Indicator PNlog Factor Model. In particular, we select the following values of kappa: kappa = 0.1, kappa = 0.33, kappa = 0.66 and kappa = 0.9.

Regarding the Fama-French 3-factors model, we present the outcomes in the subsequent figures for each selected kappa value.

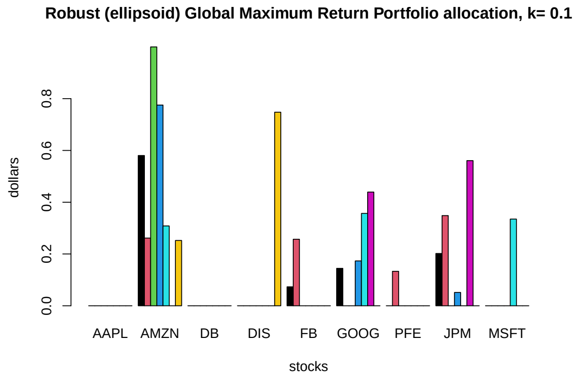

Figure 8: FF3 Bar plot representation for k=0.1

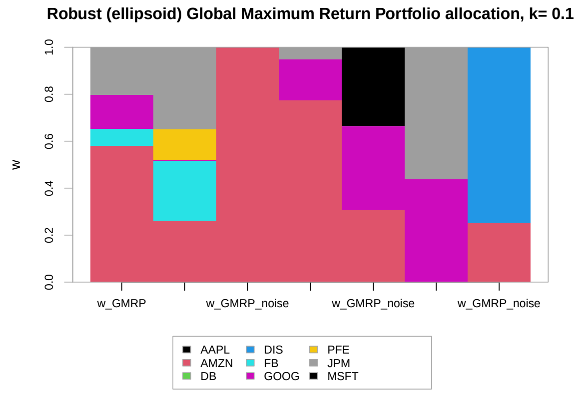

Figure 9: FF3 Stacked bar representation for k=0.1

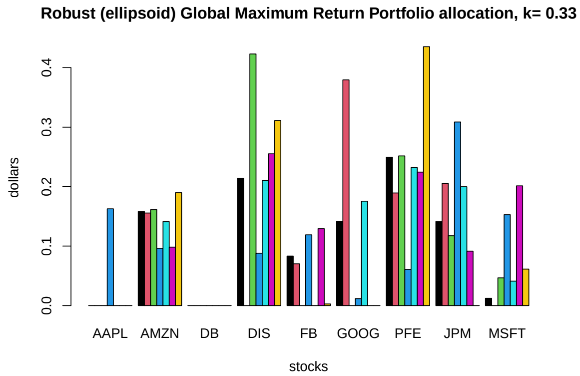

Figure 10: FF3 Bar plot representation for k=0.33

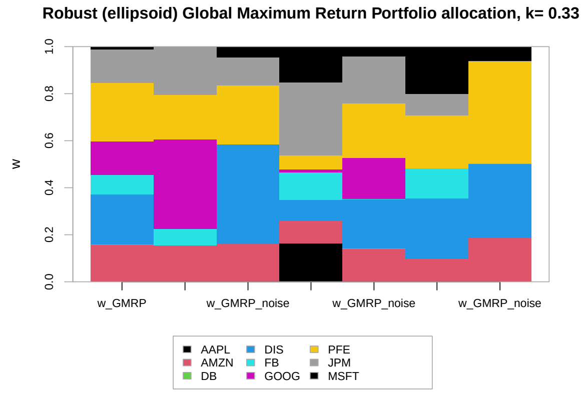

Figure 11: FF3 Stacked bar representation for k=0.33

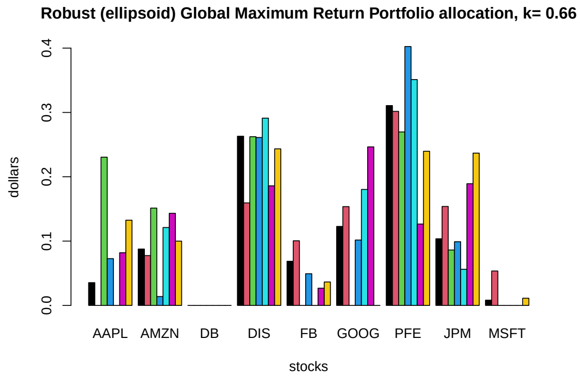

Figure 12: FF3 Bar plot representation for k=0.66

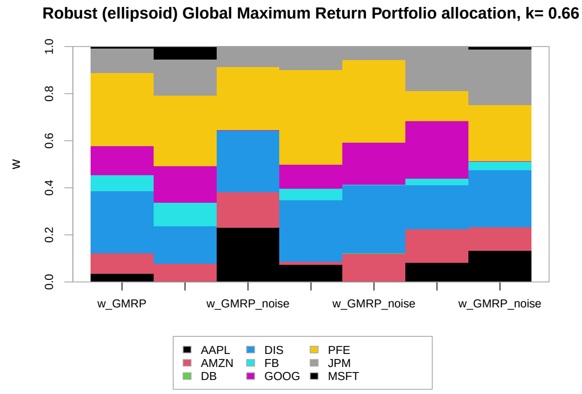

Figure 13: FF3 Stacked bar representation for k=0.66

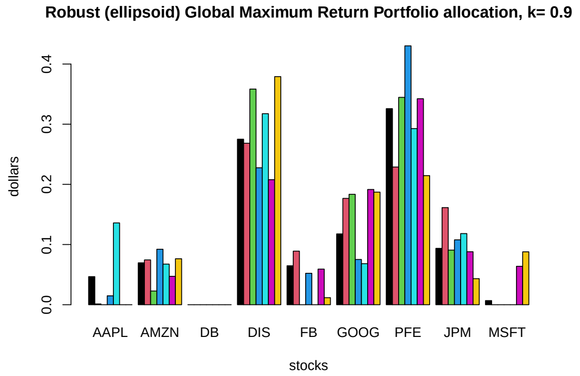

Figure 14: FF3 Bar plot representation for k=0.9

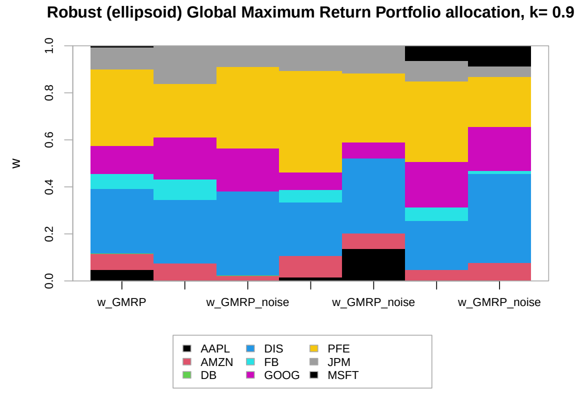

Figure 15: FF3 Stacked bar representation for k=0.9

The analysis of the graphs reveals a clear relationship between the value of kappa and the preference for lower variance in portfolio allocations. As kappa increases, the optimization problem tends to select portfolios with reduced volatility, resulting in greater stability. This observation is consistent across all figures, where the results for different values of kappa are depicted.

Now, we present the outcomes for the Sentiment Indicator PNlog Factor Model for each selected kappa value.

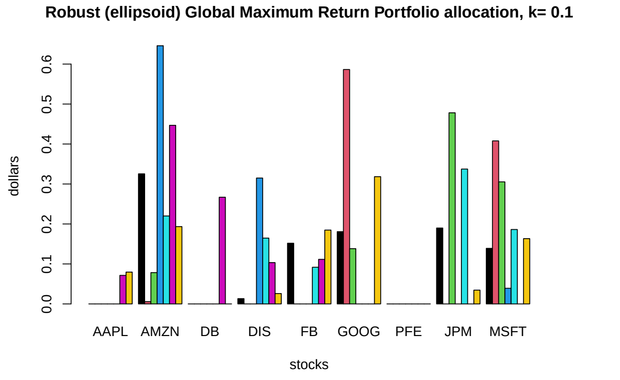

Figure 16: SentIndx Bar plot representation for k=0.1

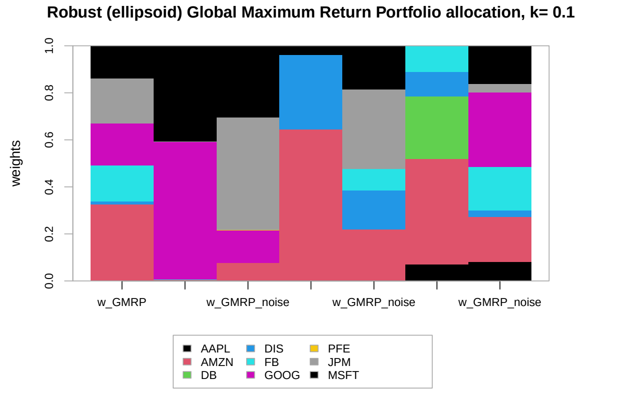

Figure 17: SentIndx Stacked bar representation for k=0.1

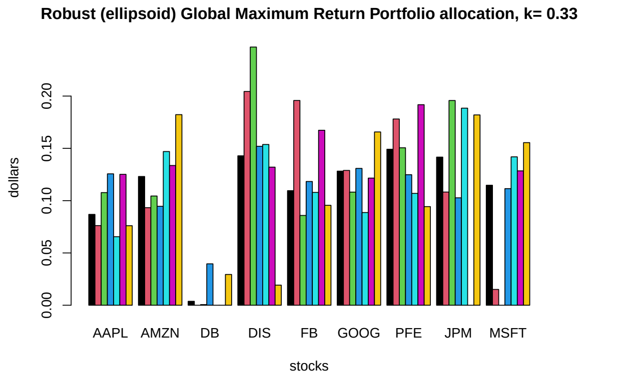

Figure 18: SentIndx Bar plot representation for k=0.33

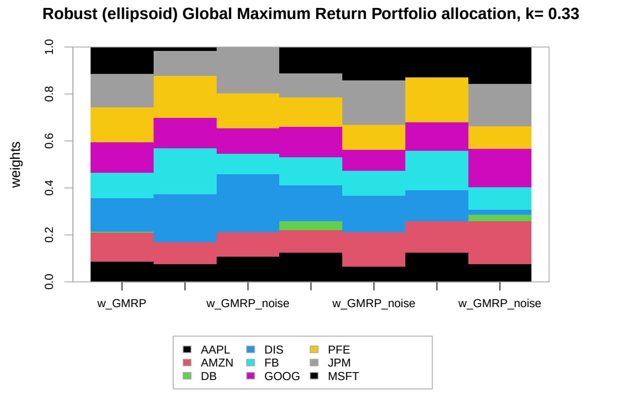

Figure 19: SentIndx Stacked bar representation for k=0.33

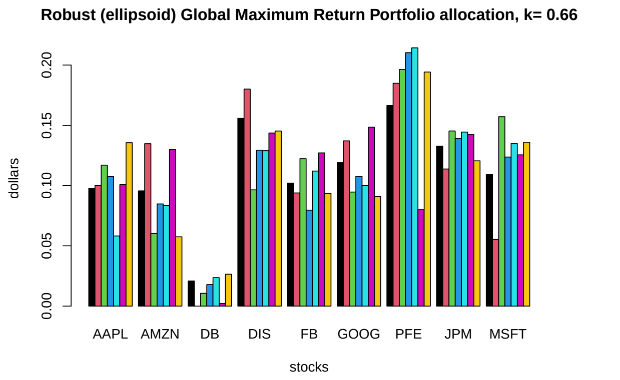

Figure 20: SentIndx Bar plot representation for k=0.66

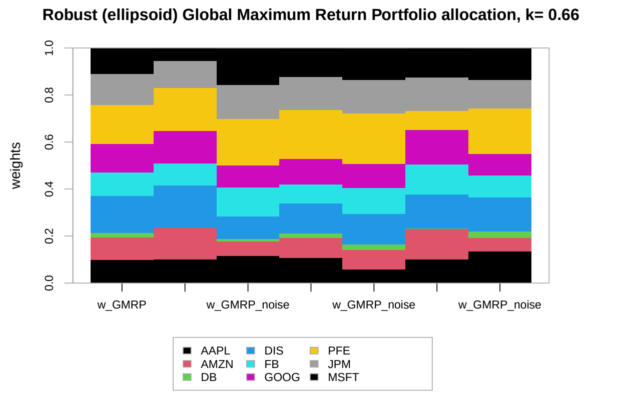

Figure 21: SentIndx Stacked bar representation for k=0.66

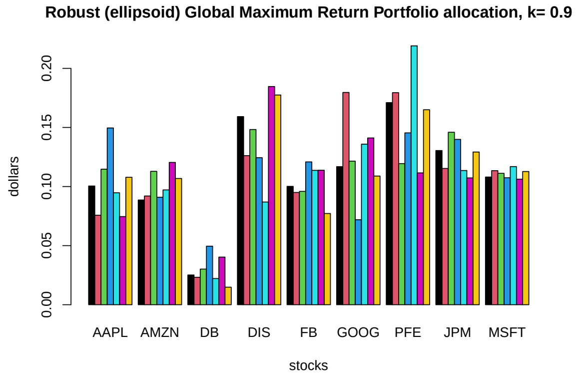

Figure 22: SentIndx Bar plot representation for k=0.9

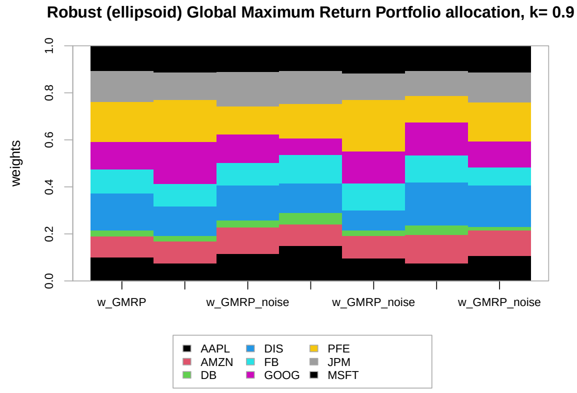

Figure 23: SentIndx Stacked bar representation for k=0.9

When comparing the Sentiment Indicator Pnlog Factor Model with the Fama-French 3-factors model, we observe a distinct characteristic. Despite considering a single factor, the Sentiment Indicator Model consistently exhibits more stable portfolio asset allocations, even at the same value of kappa. This finding underscores the importance of factor selection in portfolio optimization. While the Fama-French model captures multiple dimensions of risk, the Sentiment Indicator Model's focus on a specific factor enables it to generate more stable and consistent portfolio allocations. These insights emphasize the need for careful consideration of factors when designing and implementing portfolio optimization strategies.

3 Acknowledgments and Disclaimer

3.1 Disclaimer

The research conducted and any information provided herein are solely for academic purposes and educational discussions. The content should not be interpreted as financial advice or recommendations for trading strategies in real-world scenarios. The discussions, analyses, and findings presented are purely hypothetical and may not reflect the actual market conditions or trading outcomes.

Trading and investing in financial markets involve significant risks, and it is essential to carefully consider your own financial situation, risk tolerance, and objectives before making any investment decisions. It is strongly recommended to consult with a qualified financial advisor or professional who can assess your specific circumstances and provide personalized advice.

The information and analyses provided in this context are based on historical data, assumptions, and various methodologies that may have limitations and uncertainties. Past performance is not indicative of future results, and there is no guarantee that any trading strategy or approach will be successful or profitable.

By accessing or using this information, you acknowledge and agree that you will not rely on it as a basis for making financial decisions. You assume all risks associated with your use of the information and release the provider from any liability for any direct, indirect, or consequential damages that may arise from your reliance on the information.

Please note that financial markets are dynamic and subject to various factors, including economic, political, and market conditions, which may significantly impact the performance of trading strategies. It is crucial to conduct thorough research, seek professional advice, and stay informed about current market developments before engaging in any trading activities.

3.2 Acknowledgments

Both authors contributed equally to this work, we also used as reference the chapters available of the (ARRATIA, 2014) book and the material available (codes and lectures) in the Machine Learning for Finance course.

The codes are also available following this link on GitHub.

References

APPEL, Gerald; DOBSON, Edward. **Understanding MACD**. [S.l.]: Traders Press, 2007. v. 34.

ARONSON, David. **Evidence-based technical analysis: applying the scientific method and statistical inference to trading signals**. [S.l.]: John Wiley & Sons, 2011.

ARRATIA, Argimiro. Computational finance. **An Introductory Course with R, Atlantis Studies in Computational Finance and Financial Engineering**, Springer, v. 1, 2014.

BULKOWSKI, Thomas N. **Encyclopedia of chart patterns**. [S.l.]: John Wiley & Sons, 2021.

MURPHY JOHN, J. Technical analysis of the financial markets. **New York Institute of Finance**, 1999.