Trend-Following, Momentum crashes and high correlations

Maxime Borel and Benjamin Souane
Professor: Fabio Alessandrini
TA : Martina Fraschini

Quantitative Asset Management Project II

HEC Lausanne

UNIL | Université de Lausanne
HEC Lausanne

Abstract

This paper[1] explores trend-following strategies based on the following signals : momentum, moving average crossover, Singular Spectrum analysis, weighted EWMA crossover and support vector machine, and on the following weighting scheme: Equally weighted, Volatility parity and Risk parity. The aim of the research is to find trend-following strategies that are robust to both momentum crashes and high correlation regimes. The authors do find that more advanced signals that capture non-linearity as well as risk based weighting schemes allow to create more robust strategies with significant increase in Sharpe ratio. Results also shows that there is room for improvements since the correlation between signals is not perfect and the signals tends to perform for a few asset class but not all. Therefore, one could diversify a trend-following strategy by combining signals.

[1]You can find the matlab implementation here : Project GitHub

Contents

List of Figures

List of Tables

1 Introduction

Buying rising assets and selling falling assets is the cornerstone of a trend-following strategy. Although easy to understand and used for a long time, even Ricardo in 1838 spoke about it with his famous quote : “Cut short your losses, let your profit run on”, some challenges remind. Indeed, you need to cut your losses but how do you predict the direction of the asset ? And what weight do you put on this asset ?

Trend-Following over performed strongly during the global financial crisis but failed to keep up during it’s aftermath. It is known that trend-following strategies tends to under-perform during high correlation periods between assets which was the case after GFC and suffer during momentum crash. Therefore, one should find a way to predict the direction of the asset and weight it in a way that correct these two downsides.

To answer the first one, we investigated many signals ranging from a simple momentum to a machine learning based signal using support vector machine. Concerning the second problem, we essentially focused ourselves to risk based weighting scheme which have proven efficient in reducing high correlation risk.

We do find that simple signals such as momentum and moving average crossover are heavily exposed both to momentum crash and high correlations regimes. Our more advanced signal such as singular spectrum analysis and Weighted average of EWMA crossover do perform well under high correlation regimes but are still strongly exposed to momentum crash. This brings us to our last signal based on support v,ector machine which improves the performances but do not remedy the problem of correlation regimes, it inverse it.

More formally, we do find that we can increase the Sharpe ratio almost ten times when going from an equally weighted simple momentum signal to a risk parity support vector machine signal. Our intermediate signals are performing three to five times better than the simple momentum signal which is already a strong improvement. Another interesting feature is the relatively low correlation between the strategies and therefore, the signals. We typically found a correlation of around 10% between the singular spectrum analysis and support vector machine signals which means that there is room for improvement either by finding signals incorporating both underlying measure or by combining them[2].

The rest of the paper is divided as follow: section two is the literature review, section three describes the data and section four is about the methodology used for this research. Section five and six are respectively the results and a sensitivity analysis (parameters, factors and correlation). Finally, section seven concludes this paper with the last sections being the tables and figures as well as the appendix.

[2]Indeed, one could think that since the strategies are not entirely correlated, it means that they do not use exactly the same features to generate the signal. Therefore, one could potentially find a signal combining both features.

2 Review of literature

Trend-Following strategies have been widely studied during the last decade but has been used since the beginning of the financial markets [B H]. Our review start with the article from Moskovitz, Ooi and Pedersen [MP12] as well as [Ant] that highlights the significant excess return generated by a global cross-asset time series momemtum portfolio typically showing high Sharpe ratios.

Our study explores several signals, obviously, the most studied ones are simple momentum signals and a moving average cross-over. However, many other more advanced signal have been studied, Bruder, Dao Richard and Roncalli [Bru11] explore dozens of signals which inspired us for the *Singular Spectrum Analysis* and the *Support Vector Machine* based signal. Trend extracting using SSA has been studied by Alexandrov [TAl09]. We did not find evidence on SVM based signal in the literature however, we were inspired by Wang, Smith and Hyndman [XWa05] for the features extraction. An other interesting signal comes from the paper from Baz, Granger, Campbell, Le Roux and Rattray [Jam15] with a CTA-Momentum based on the crossover of exponentially weighted moving average which yield strong results on a cross-asset momentum portfolio. As shown by Levine and Pedersen [Lev16] signals such as momentum or moving average crossovers are capturing the same properties as advanced filters, therefore we will investigate the correlation between our signals.

One of the goal of this paper being to avoid momentum crash, we must cite Daniel and Moskowitz [Dan16] that used a regression model to predict the expected returns, using a bear market indicator and forecast the volatility with a GJR GARCH to avoid such crashes[3]. Xiong and Ibboston [Xio15] also provided interesting comment on the acceleration of the trend which would be an indicator of momentum crash, we will include this intuition into the features for the *Support Vector Machine* based signal.

The second goal of the paper being to develop a strategy that work in all correlation regime can be resolved by risk based weighting schemes as shown by [Bal15] which leads us to the use of risk based weighting schemes. They show that using a risk parity approach compare to a volatility parity or equal weighted approach perform better when the correlation regime is high[4]. Therefore, we expect the same results for our portfolio.

Finally, we also use the intuition provided by Yu and Webb [Yu 16] to see whether our portfolios are exposed to factors. We also have been inspired by many other papers worth noticing such as the paper by F. Nucera on risk-managed momentum [Fed17] as well as the article by N. Baltas and R. Kosowski on momentum strategies [Bal19] which explore continuous signal to lower the turnover of the portfolio.

[3] In undisclosed results, we implemented this concept into our models but it did not yield significant changes out of sample.

[4] Volatility parity is often named inverse volatility or naive parity since it does not include covariances in the scheme. On the other hand, risk parity equalize the total contribution to risk of each asset, taking into account the covariances.

3 Data

The data used in this paper is a sample of 18 futures contract, composed by 5 equity indices, 4 currency pairs, 5 commodities and 4 fixed income instruments as shown in the Table 8.1. All data were retrieved from the provider Thomson Reuters (Datastream).

In terms of geographical allocation, the assets are spread out around the following regions : North America, Europe, United Kingdom and Asia, the goal being to extend the opportunity set of the strategy across the world. However, we have a slight concentration in the euro zone.

Some of our time-series are not in USD, therefore, to get USD prices, we multiply the foreign asset prices by the currency pairs in our dataset. Best practice would be to compute the signals in foreign currencies but assess the performance is USD, however, for simplicity we have measured both the signal and the return on USD basis.

The price series are continuous futures contracts for which the rolling of the contracts are already computed which allows us to use only one time-serie per asset and avoid the task of rolling the contracts. More precisely, the series are constructed as if the investor is always investing into the most liquid contract [5]. One must add that using futures contracts allows the strategy to be almost money neutral since one only needs to pay a small premium to enter into the position. Finally, the time series are not particularly long for all assets, typically the Japanese bond time-series starts very late compared to the others as can be seen from the table 8.1.

We use daily data over a sample period of nearly 22 years from 05 January 1998 to 31 August 2020. Hence, we have 5702 prices and thus 5701 returns. We integrated the recent Coronavirus outbreak period in order to see how our models managed this special period. All assets are not available from the beginning, thus we integrated the missing assets dynamically during the back-test of the strategies.

From this raw price data, we computed the simple returns as in A.1. From that, we also computed monthly prices and returns using 21 data points as month length. One should not that the monthly data are computed before each strategy since the starting date is dependent of the length of the signal and the need to estimate parameters. Moreover, we extracted a vector with the asset class of each asset to compute some statistics by asset class.

Finally, in order to assess the performance in terms of factor analysis, we retrieved the Fama-French 6 Factors dataset which are freely available[6]. As we expect to find almost zero exposition to equity factors since we have a global cross-asset portfolio, we also retrieve data from several indices as shown in Table 8.2. For risk-adjusted performance measures, we use as a risk-free rate of 1% which is conservative because during the period under study the interest rates remains close to zero. Hence, it slightly underestimate our risk-adjusted performance measures.

[5]You can download the Datastream methodology for rolling continuous series here.
[6]You can access the data here.

4 Methodology

Trend-Following strategies use a combination of a signal, a weighting scheme and in general a volatility target as can be seen by equation A.2. In this study, we try to improve the risk-adjusted performance by working both on the return and risk side.

As the name suggests, these strategies are following the trend through a long-short signal that can be binary or continuous. However, there are many ways to compute this trend and we explore some of them[7]. We start with simple signals based on the time-series momentum[8] (eq.B.1) and the moving average crossover (eq.B.2) which gives us an independent signal for each asset. The allocations based on these signals showed that they were not sufficient to either capture the trend or sustain momentum crash. Either the window is too long which allows us to capture the trend but take too much time to return itself, or the window is too short and we do not capture the long-term trend and take a lot of wrong directional bet.

These results yield us to two more advanced signals. The first is a CTA-Momentum based on *Exponentially Weighted Moving average crossover* (subsection A3) and the second one is based on *Singular Spectrum Analysis* (subsection A4). These signals have the great features of taking into account both short term and long-term features of the time series (SSA is based on the entire trajectory matrix). These allocation yielded better results especially in terms of correlation regimes but did not necessarily avoid crashes. This last point made us realize that the relation between the trend and the underlying time-series can be strongly non-linear which yield us to the last signal.

To capture this non-linearity we used a machine learning based signal, more precisely using *Support Vector Machine*(subsection A5) which yielded better results especially when combined with a risk parity weighting scheme and to a trading rule based on the classification score (B.5).

Combined to these signals, we used three weighting schemes, Equal weighted, volatility Parity (subsection A1) and risk parity (subsection A2). A true long-short risk parity is a highly non-linear problem that requires global optimization routine. To avoid such problems, we implemented a "weak form" of the risk parity. Indeed, we compute the weights on a long-only framework starting with the volatility parity weights which should converge to the global optimum since it is a convex optimization problem. As stated previously, we use a target volatility of 10% annualized which is obtained by leveraging our position [9] (eq.A.2).

In order to reduce the fees (by lowering the turnover) and the operational difficulty, we perform monthly rebalancing but with calculations based on daily data. We used fees of 10 basis point which is realistic for a CTA.

Finally, we will perform a sensitivity analysis of our models to see whether our results are robust and are not over-fitted.

[7] For a thorough review of the possible trend extraction methods, see this article from Roncalli & co.

[8] A trend-following momentum is also known as time-series momentum which differs strongly from the momentum based on the cross-section of asset. Indeed, the standard momentum is almost always market neutral whereas the time-series momentum can be all long and all short depending on the signal, taking strong directional bets.

[9] one must note that both the volatility and risk parity will still be optimal since the volatility is homogeneous of degree one (i.e $f(\lambda x) = \lambda f(x)$).

5 Results

This section presents the results of all the models included in this paper. In order to measure accurately the risk-adjusted-performance, we use the Sharpe and the Calmar ratio. Indeed, the Sharpe is widely used and very intuitive while the Calmar gives good insights regarding the crashes by using the Maximum Drawdown. All the performances presented in this report are net of fees.

5.1 Model 1

This section explores standard signals on the time-series momentum with several windows and the moving-average crossovers. The results of these models are shown in Table 8.3.

Firstly, the Sharpe ratio of all strategies goes from -0.1869 to 0.2443. As expected, the risk weighted schemes perform better than the naive in both performance and risk criteria. The Risk parity tend to outperform the volatility parity except for the momentum based on the first 90 days.

Secondly, in Figure (8.2, 8.3, 8.4 and 8.5), we can see that all these strategies suffer from severe crashes, in particular in periods of market rebound[10]. One must note that before fees, the risk parity would have yielded a really strong performance. However, the risk parity benefits come at a cost, the high turnover did eat a huge part of the performance.

Focusing on the Momentum strategy, the 252 days window gives better results in terms of performance but have larger maximum drawdowns. Using a tinier window, 90 days, reduces slightly the returns but reduces as well the drawdown which increases the Calmar ratio. The reduction is even more pronounced among the weighting schemes, the risk parity is able to reduce even more the maximum drawdown. Indeed, there is a trade off between a longer period that captures the trend but take a lot of time to return or a smaller period that is changing faster but tends to make a lot of wrong directional bets. In addition, the momentum 90 days with volatility parity strategy is positively skewed which is a desirable property and it is the strategy that performs the best.

The two other frameworks, the momentum "jump" using the signal based on the return between the 9th to 12th previous to the balancing and the moving average crossovers yielded poor results with close to negative performances. Hence, these standard models were not good overall.

5.2 MBBS Signal

As this model depends on several parameters, namely the forgetting factor, the responseScale and the threshold for the trading rule, we run some empirical backtests on the volatility parity model to find the parameters with relatively good performance. They are set as follows : forgetting factor is 11, the response scale at 0.89 which is standard, price length of 300 days and the trading rule threshold at 70% which is relatively restrictive. It allows to extract very strong trends. The Figure 8.6 shows the all different alternatives of the model performances.

[10] Surprisingly, most crashes in trend-following strategies occurs not necessarily on momentum crashes but more generally on momentum reversals.

The peaks in the cumulative return can be partially explained by an increase of leverage during the period.

The Table 8.4 show all the statistics of the different portfolios. The *MBBS - Risk Parity - Overall* is the best performing strategy. Even though that the strategy has one of the largest volatility (12.6%), the Sharpe ratio is 0.389 exceeding all the others. Hence, the models improve mostly the return side than the risk side compared to the momentum. The performance over time are relatively similar to other models until 2016. Then, models with the trading rule at the portfolio level perform better. The leverage of the models stay relatively acceptable during the whole period. In general, the trading rule reduces the Max drawdown as well as the turnover because it massively smooths the signals. The equivalent model based on the volatility parity shows a slightly better Calmar ratio benefiting from a much lower Maximum drawdown. As expected the turnover of the Risk Parity is higher than the other weighting schemes. However, one can remark that the trading rule is able to reduce it significantly. According to the Figure 8.16, the MBBS Risk Parity with the trading rule at the portfolio level performs better for the equity and commodity asset classes. Compared to the other advanced signal, this model extracts particular nice result for the commodity and can be combined to the others signal for the rest of the asset classes.

In the Tables 8.19, we show a regression of the rolling Sharpe ratio and different types of correlation. In particular, by looking at the columns ALL and InClass, we can see that the Sharpe ratio is positively relationship to the level of correlation. When the correlation among all the assets increase the Sharpe increase. It is also the case when the correlation among assets classes increases, the model perform better. Both parameter and lower than one indicating a low slope. The model seems to perform well also in low correlation regime. It might be a benefit of the risk parity weighting scheme that takes into account not only the volatility but also the correlation structure.

Finally, using a continuous signal with a trading rule and risk parity weighting scheme gives relatively better result compared to a simple momentum strategy. The annualized return almost double for a similar amount of risk. It is also able to reduce the drawdown of the portfolio and perform better in all correlation regimes. However, the improvement seems to be driven by the restrictive trading rule rather than the signal itself as models with different trading rules.

5.3 SSA Signal

A second signal that exhibits interesting properties is the continuous signal based on singular spectral analysis. As previously, we used the trading rule described in the appendix to try and improve the performance.
The best performing model is the *SSA - Risk Parity* model that yields a Sharpe Ratio of 0.68 after fees whereas the two other models yield 0.52 (equally weighted) and 0.40 (volatility parity). The relation is the same for the Calmar ratio, indeed the risk parity allocation also does have the smallest maximum drawdown (24%). However, one must not that the risk parity allocation yields a significantly higher turnover which can be difficult to implement in practice. Finally, looking at the skewness and kurtosis, one can see that all the equally weighted and risk parity weighting scheme are negatively skewed whereas the volatility parity is positively skewed.
One of the prime features of the SSA signal as well as the risk parity weighting scheme is the

strong performance after 2015. Even during the 2020 crisis, the model is quite balanced between the asset classes and therefore, do not suffer from it 8.10. Interestingly, the signal is mostly long and very rarely overall short on an asset class. Indeed, the performance is due not to an amazing prediction but by taking smaller positions in assets where the trend is not strong [11]. The other prime feature of the strategy is the relation to correlation regimes. We can see from the table 8.35 that the return of the strategy is positively correlated to the inter-class correlation which is the most important one but is negatively correlated with the overall pairwise correlation.
Looking at the decomposition of the return by asset class (8.16), we do see that 40% and 50% of the cumulative return are coming from the commodities and equity respectively. However, it does not work at all on the forex with a negative return overall. It stresses the fact that combining signals could really improve the performance of trend-following strategy. Finally, including trading rules on this signal yield negligible changes and should not be included in a portfolio at least with these parameters.

5.4 Support Vector Machine

Our last model improves strongly all the previous results but first we must not that we used around ten years of data to train the algorithm (40%), therefore our back-test starts only mid 2007. Here again, we used the three weighting schemes at our disposal and again, the risk parity is the best performing model with an after-fee Sharpe ratio of 0.88 8.45, however the other weighting schemes also yielded strong performance with Sharpe ratio of 0.70 (8.41) and 0.71 (8.37) respectively. The most spectacular feature of this model is the reduction in maximum drawdown to 22 % for the risk parity weighting scheme which increases the Calmar ratio compared to the other models. The signal yields three allocations that are positively skewed which is an interesting feature. However, looking at the performance and correlation regimes (8.48), we can see that the allocation is positively correlated to every correlation regimes which the opposite of the standard trend-following strategy. However, it is not the exact relationship, we are looking since the best thing would be to be neutral to correlations. We must finally note that the risk parity allocation, here again, yields a massive turnover which can be difficult to implement in practice.
Zooming on a few key periods in our sample starting with the beginning of the GFC[12] which coincides with the beginning of our back-test. We can observe that starting in January, we are short on the equity for almost all 2008 but the algorithm started to capture the crash even in December 2007 being short equity in December. For 2008, we are mainly short the Forex market as well as the equities and long the fixed income and the commodities. The other main market event in our sample is the corona crisis. We can here again observe that our algorithm performs relatively well. Indeed, it is neutral on the equity in March and April 2020 and directly go long on May and June, however it is only slightly long the equities in July and August which is not very a sensitive strategy, we can stipulate that the algorithm did not learn well the *TINA*[13] phenomena and therefore, think that the market will return itself. It is interesting to see that the algorithm acted well on both crises which were quite different in markets reaction. One

[11] Indeed, the SSA is using the entire trajectory matrix therefore the general trend should be positive in general (typically thinking about the CAPM) therefore, the differences are made with the strength of the trend.
[12] Global financial crisis
[13] There Is No Alternative, the fact that with such low yields on bonds, equity is the only place to go for findings profit.

must note from 8.16 that more than of 60% of the return come from the equity which is strongly concentrated. It again stresses the fact that combining signals can yield better performance, we could typically invest on equity with the SVM model and on commodities with SSA model.

5.5 Conclusion

Overall, we can see that the best model is the SVM signal which captures the non-linearity in a better way than the other ones and strongly reduces the drawdowns but as a positive relationship with correlation. An interesting model that could be studied in a following paper would be to combine signals since their correlation are not perfect, typically the SVM and SSA risk parity strategies have a correlation of around 10% and tend not to necessarily work on the same assets.

6 Sensitivity

6.1 MBBS Signal

In this section we will analyze the sensitivity of the MBBS model. We will start to analyze the factor exposure of our models and then varying some parameters. The results are based on the volatility parity weighting scheme with the overall trading rule as it was too much time consuming with the risk parity. While the other parameters such as the length of the EWMA (300 days) or the amount of trading rule are kept constant, we focus the sensitivity analysis on two parameters : the responseScale and the λ forgetting factor. The former is a parameter in the rescaling function $F(X)$ as describe in subsection A3 initially set at 0.89. We test for a range from 0.5 to 1.2. the latter is the forgetting factor use in the EWMA computation yielding x_k in the same section. The default value is $\lambda = 11$.
The factor analysis focuses on Table 8.5, 8.6, 8.14, 8.15, 8.8, 8.9, 8.17, 8.18. In general, the exposition to the 6FF factors as well for the major indices is mostly zero. Nevertheless, the overall trading rule give exposure to MSCI world index which is easily explainable by looking in Figure 8.7. We are highly exposed to stock markets as the signal mainly due to the overall trading rule. Surprisingly, the volatility parity and the risk parity are both exposed to the CMA factor. However, the overall trading rule withdraw the HML exposure of the former and the RMW of the latter. Finally, the α are significant for both strategies.
The Figure 8.8 shows the Sharpe and the Calmar ratios for the above parameters varying. They behave relatively similarly but the Sharpe seems to be more sensitive. There are large peaks on both graphs like at (0.9,11) but also very low points (6, 0.8) yielding a dramatic Sharpe of -0.42. The best Sharpe can be reached at a corner point (10,1.1) reaching a Sharpe of 0.42. High λ indicates higher weight for the recent information, it seems that choice yields better Sharpe ratio in particular for high levels of responseScale. Finally, even though, the performance is better than a simple model, the performance is highly dependent to the choice regarding the parameters. Indeed, by choosing the wrong responseScale and λ the performance can be catastrophic such as -0.42 for the Sharpe ratio. Hence, by choosing the right parameters in-sample, the Sharpe ratio is boosted but there is no strong guarantee that this model will perform in the same way in future market conditions. Therefore, we do not recommend using this model.

6.2 SSA Signal

This model has a lot of parameters to change, indeed, we will vary the latent dimension, length of signal reconstruction, trading rule threshold, shrinkage and rescaling bounds. The most logical parameters to change is the length of the signal which is 90 days on the presented results. As we can see from figure 1 and 4 in 8.11, the length of the signal is quite a robust parameter. Indeed, except for very short signals, the performance is fluctuating around the same value. However, we can see that our initial parameter does not yield the best performance and estimating parameters could be interesting to improve the performance.
Looking at the length of the latent dimension (figure 1 and 3), we can see that the effects are quite undefined. Look at figure 3 on the plot, we can see that our model is not best at all. Here again, estimating the parameters before launching the model could yield a serious increase in the performance.
The two others parameters to change are the trading rule threshold that does not have a clear impact. We do see that with longer signals, the performance is increasing with higher thresholds. On the other hand, the shrinkage and rescaling of the signals have a strong impact. Our based model is with a shrinkage parameter of 2 (we are putting extreme value to 2 and -2) and a rescaling of 1 (our signals are mapped between -1 and 1). We see from 8.11 that the shrinkage is a very important parameter and choosing something else than 1 or 2 would strongly decrease the performance of the strategy. On the other hand, the rescaling does not change a lot since we are leveraging our position afterward to attain a constant volatility.

6.3 Support Vector Machine

Since machine learning models tends to work in a kind of black-box framework, looking at the "factor loading" is relatively interesting. First, one should note that for each weighting scheme the alpha is significant and reach 1% monthly for the risk parity. On the other hand, as expected, all the other factors are not significant since we have a cross-asset portfolio. More interestingly, we are not correlated with major indices of any asset class which means that we do not rely heavily on only one class.
We will now look at the sensitivity of the model parameters. However, there are not a lot of parameters in this model, hence we will only look at the performance while changing the threshold on the rule based on the posterior probability of classification. The previous results are presented with a threshold at 20 %. As we can see with the figure 8.14, the best performance would have been attained with a threshold at 10 %. Taking a smaller threshold yields a significantly smaller performance which means that the trading rule clearly works. Starting at a threshold of 25 % and more, the allocation almost never enters into a position and the Sharpe ratio is converging to 0. Therefore, we do think that the model is relatively robust, there is not a lot of parameters choosing. However, asked on which threshold to choose, we would tell the investment manager to use something smaller than they used one in order to avoid extreme negative performance. The other parameter to choose is the length of the feature and which kernel to use, the model is also robust here with a sweet spot between 45 and 90 days and with both the RBF and Quadratic Kernels.
Finally, changing slightly the trading rule for this strategy by taking the previous signal if we are not confident enough and not 0 dramatically increases the performance, typically yielding a Sharpe Ratio of 1.28 for the risk parity weighting scheme.

7 Conclusion

The aim of the paper was to develop a cross-asset trend-following strategy that would (a.) be resilient to momentum reversals and (b.) work on every correlation regime. We worked on futures contracts continuous prices series for a global investment universe on Equity, Forex, Commodities and Fixed Income. We first investigated the known momentum and moving average crossover that typically yields great performance overall but a really sensitive to both momentum crash and high correlation regimes. Hence, we explore more advanced signals.

The first model is the CTA-momentum based on the EWMA crossover which has interesting properties. However, we do find that it is too much sensible to parameters and may easily be over-fitted. Our implementation improves the performance but still is strongly sensitive to momentum reversal and can yield dramatically low Sharpe ratio with the wrong parameters. Hence, we do not recommend to use this model.

We improve this model with singular spectrum analysis that computes the dimension which explains the biggest variance on the lag of the series. This signals yields very strong performance with a reduction of maximum drawdown to 24 % for the risk parity weighting scheme. However, the relation to correlation regimes is not totally improved which yields us to the last model.

This last model is based on Support Vector Machine, which aims at classifying the features of a time series into a ± 1 signal.This improved strongly our results with a maximum drawdown of 22 % and a positive relation to correlation regimes with is not perfect but is not common. Moreover, having the "Buzzword" machine learning can always be interesting to sell the product.

Our recommendation for our client is to use the SVM model. There are several advantages to do so. It is our best in-class model and it has a marketing advantage by using machine learning which could be a good selling argument. We advice to keep researching and improving the models by using the following arguments.

A version two of the strategies could strongly improve the performance of our model by combining signals together which could yield diversification benefits. Looking more formally to the correlation matrix 8.17, we do see that we have almost zero to negative correlation for both the SVM and SSA compared to other strategies. However, as explained in [Lev16] the momentum, moving average and EWMA are strongly correlated between them. Therefore an potential mix would be to use SVM, SSA and a simple momentum.

In addition, instead using complex econometric models to solve momentum crashes, one can also combine the trend following strategy to some value factor [CLI13]. In the case of the support vector machine signal, the improvement can take two precise forms. At first, one could dynamically train the algorithm, indeed in our case we are training it once and then back-testing it, however, one could retrain the model every month with the new data that just happened. Secondly, one could improve the feature extractions and even migrate to a labeling strategy which is a current trend in machine learning for finance [BS20]. One would typically create buy/neutral/sell feature instead of the previous period return which would reduce the noise that the return can bring to the algorithm.

8 Figures and tables

Table 8.1: Assets denomination

Equity	Currencies	Commodities	Fixed income
CME-S&P500 INDEX	CME-GBP/USD	ICE-Brent Crude Oil	CBT-10 YRS US TNOTE
OSX-NIKKEI225 INDEX	CME-CHF/USD	CSCE-COCOA	SGXDT-10 YRS JGB
CME-EMINI-NASDAQ100 INDEX	CME-EUR/USD	CBT-CORN	EUREX-EUROBUND
EUREX-EUROSTOXX50 INDEX	CME-MINI-JPY/USD	CMX-GOLD100OZ	LIFFE-LONG GILT
EUREX-SMI INDEX	CSCE-SUGAR11		

Table 8.2: Factors denomination

Equity	Currencies	Commodities	Fixed income
MSCI World	US DOLLAR INDEX Trade Weighted	SP GSCI Commodity	FTSE World Government Bond
MSCI Emerging Market			

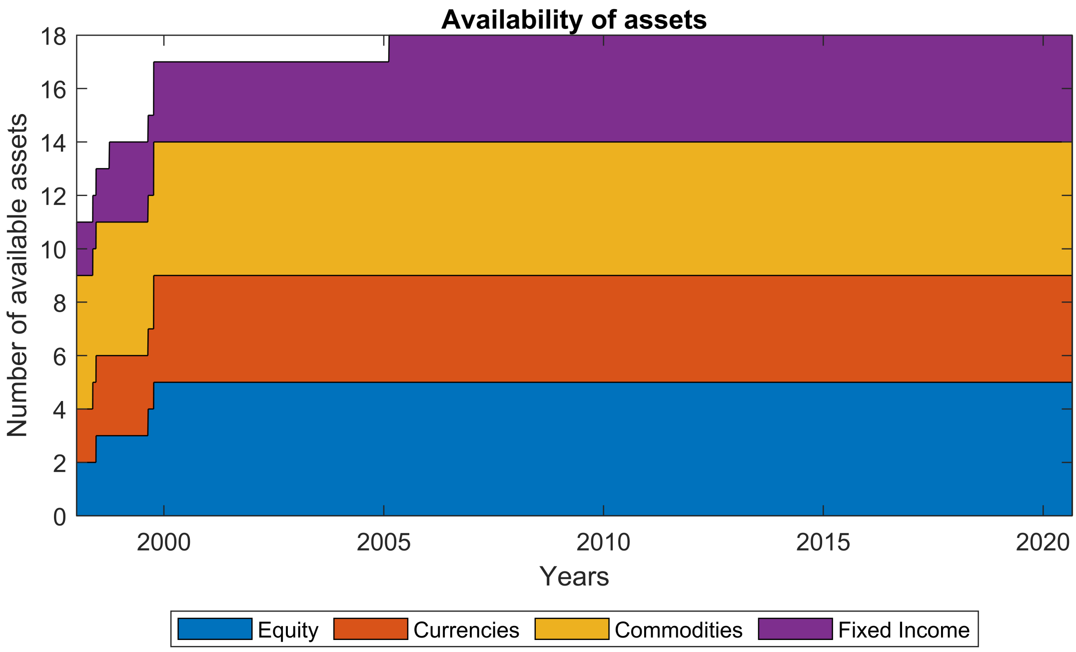


Figure 8.1: Availability of the assets along time

8.1 Basis Model

OriginalVariableNames	AnnualizedMean	AnnualizedVolatility	Kurtosis	Skewness	AverageMonthlyTurnover	SharpeRatio	CalmarRatio	MaximumDrawDown	HH
MOM252VP	0.037962	0.13223	12.8386	-1.3498	0.2863	0.21147	0.09711	0.39092	0.45467
MOM252RP	0.037208	0.12427	14.413	-1.4402	0.6403	0.21895	0.105	0.35436	0.2915
MOM252EW	0.030465	0.1274	11.7614	-1.3243	0.28805	0.16064	0.07682	0.39657	0.14938
MOM90VP	0.038845	0.11805	4.9475	0.17465	0.50116	0.24435	0.11538	0.33666	0.49401
MOM90RP	0.035394	0.1199	5.6399	0.36836	1.2985	0.21179	0.11083	0.31936	0.44493
MOM90EW	0.025926	0.11226	5.0401	0.27166	0.46759	0.14187	0.078043	0.33221	0.1598
MOMJUMPVP	0.01724	0.17837	3.7142	-0.3792	0.48369	0.040588	0.045426	0.37951	1.1204
MOMJUMPRP	0.024128	0.17965	9.8366	-0.95614	1.7484	0.078644	0.046689	0.51678	0.86839
MOMJUMPEW	0.0044955	0.16727	5.1604	-0.72764	0.4604	-0.032908	0.0098275	0.45744	0.42385
MAVP	-0.024697	0.1856	5.8819	-0.52337	0.67294	-0.18695	-0.037271	0.66263	1.1619
MARP	-0.0041053	0.18715	10.6985	-0.86257	2.2355	-0.075371	-0.0074763	0.54911	0.86839
MAEW	-0.018371	0.17188	4.0002	-0.24078	0.6595	-0.16506	-0.035417	0.5187	0.42965

Table 8.3: Descriptive Statistics of Momentum and Moving average models.

8.1.1 252 days Momentum Signal

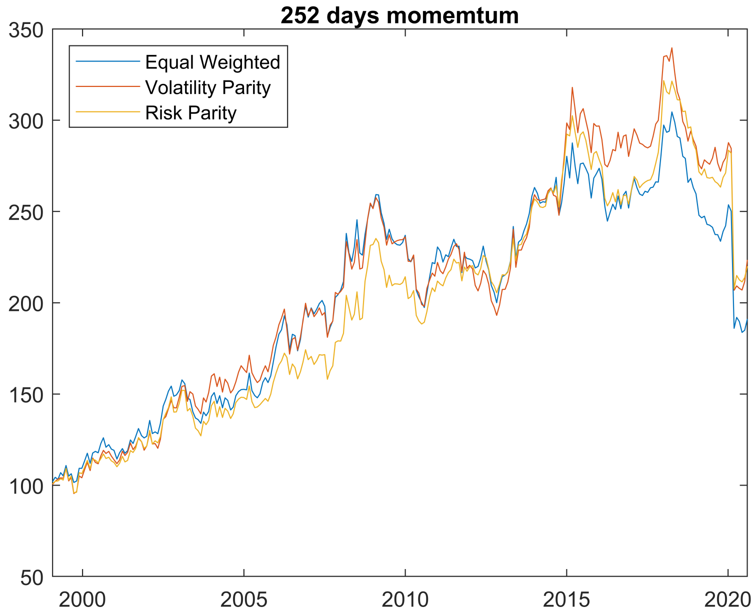


Figure 8.2: Cumulative returns for different allocations using the 252 days momemtum signal.

8.1.2 90 days Momentum Signal

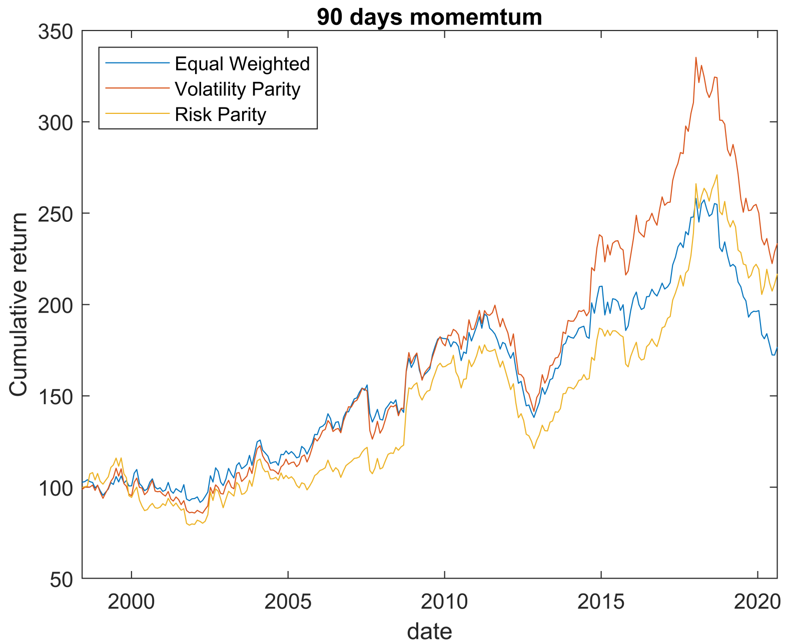


Figure 8.3: Cumulative returns for different allocations using the 90 days momemtum signal.

8.1.3 90 days Momentum Jump Signal

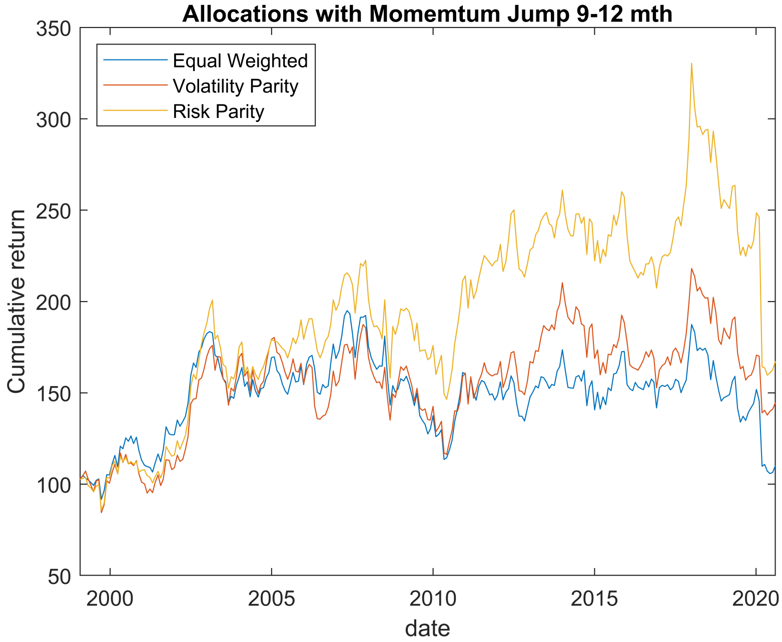


Figure 8.4: Cumulative returns for different allocations using the 90 days momemtum jump signal.

8.1.4 Moving average cross-over signal

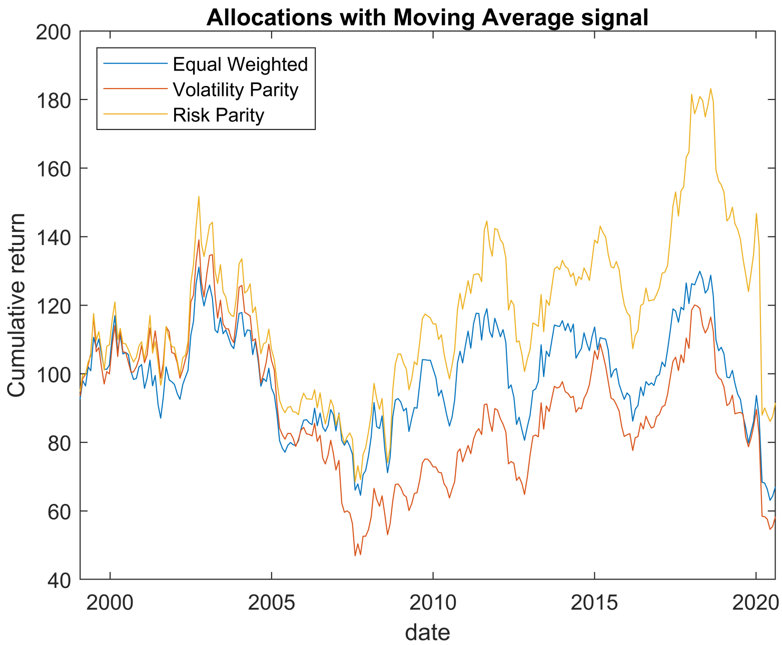


Figure 8.5: Cumulative returns for different allocations using the moving average signal.

8.2 MBBS Model

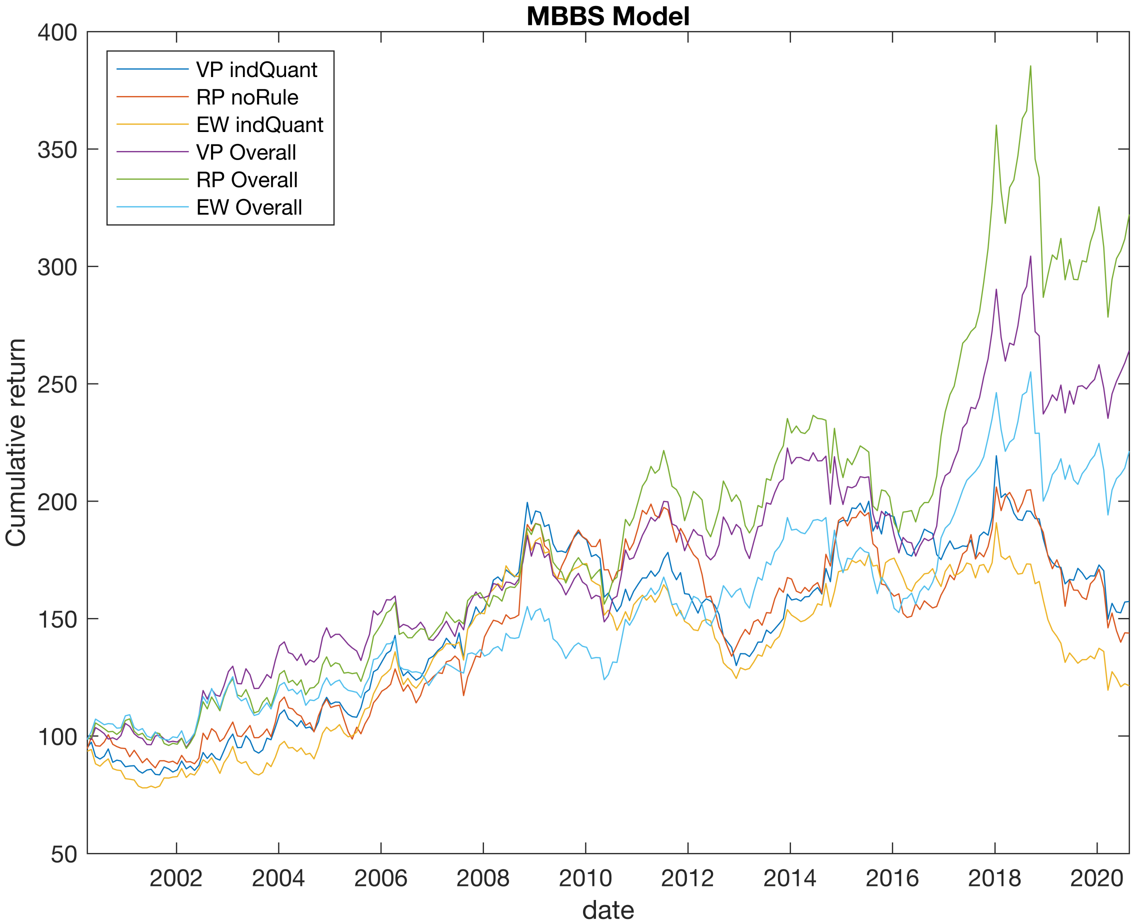


Figure 8.6: Cumulative returns for different allocations using the signal from [Jam15].

OriginalVariableNames	AnnualizedMean	AnnualizedVolatility	Kurtosis	Skewness	AverageMonthlyTurnover	SharpeRatio	CalmarRatio	MaximumDrawDown	HH
Vol.Parity indQuant	0.022445	0.1153	4.6042	0.051006	0.17415	0.10793	0.064439	0.34831	0.49109
R.Parity noRule	0.017969	0.12649	4.9862	0.3537	0.65831	0.063003	0.055113	0.32604	0.52366
EW indQuant	0.0095583	0.10752	3.8996	-0.14476	0.15321	-0.0041086	0.025576	0.37371	0.15529
V.Parity O.quantity	0.048735	0.11531	4.1909	-0.31661	0.092901	0.33591	0.2147	0.22698	0.42671
R.Parity O.Quantity	0.058963	0.12599	4.5551	-0.64774	0.31654	0.38863	0.21237	0.27764	0.44975
EW O.Quantity	0.039683	0.11261	4.2719	-0.61743	0.074025	0.2636	0.16611	0.2389	0.15025

Table 8.4: Descriptive Statistics of the EWMA Crossover signal with a volatility parity weighting scheme.

8.2.1 Rule on individual trend quantity - Volatility Parity

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0022516	0.010366	0	0	0	0	0
pValue CAPM	0.29365	0.82662	0	0	0	0	0
Coefficient FF 3F	0.0024507	0.0022896	0.092758	-0.088128	0	0	0
pValue FF 3F	0.25778	0.96183	0.44398	0.30811	0	0	0
Coefficient FF 6F	0.0021199	0.060463	0.063034	-0.28075	-0.22181	0.44459	0.074784
pValue FF 6F	0.35513	0.3048	0.62252	0.047571	0.17293	0.024223	0.22513

Table 8.5: Factor Analysis of the EWMA Crossover signal with a volatility parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.001183	0.0069027	0.043004	0.0011929	0.036614	0.26567
pValue AF	0.60227	0.85407	0.65425	0.98598	0.84661	0.18374

Table 8.6: Correlation analysis with major indices of the EWMA Crossover signal with a volatility parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass
Coefficient SR	3.2764	42.1729	-11.982	-5.9755	-6.3568	-3.0289	1.2223
pValue SR	3.0428e-09	5.1929e-15	2.1683e-14	1.0788e-08	3.8054e-12	9.4255e-05	0.053669

Table 8.7: Correlation study of the EWMA Crossover signal with a volatility parity weighting scheme.

8.2.2 No Rule on trend quantity - Risk Parity

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0020453	-0.0064169	0	0	0	0	0
pValue CAPM	0.38458	0.90168	0	0	0	0	0
Coefficient FF 3F	0.0016285	-0.0077216	0.19398	0.076531	0	0	0
pValue FF 3F	0.49141	0.8827	0.14397	0.41824	0	0	0
Coefficient FF 6F	0.0013929	0.065345	0.14658	-0.18136	-0.34953	0.60335	0.1028
pValue FF 6F	0.57307	0.30394	0.28907	0.23424	0.046999	0.0047104	0.12255

Table 8.8: Factor Analysis of the EWMA Crossover signal with a risk parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0015618	-0.00036433	0.047953	-0.056823	-0.052708	0.16013
pValue AF	0.53318	0.99298	0.65099	0.44872	0.80084	0.46737

Table 8.9: Correlation analysis with major indices of the EWMA Crossover signal with a risk parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass
Coefficient SR	1.9605	38.8382	-9.7179	-5.9724	-3.6711	-3.5806	1.1318
pValue SR	1.0226e-06	1.7376e-20	2.0085e-16	1.2873e-13	2.3512e-08	1.9364e-09	0.016255

Table 8.10: Correlation study of the EWMA Crossover signal with a risk parity weighting scheme.

8.2.3 Rule on individual trend quantity - Equally Weighted

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0011811	-0.0037082	0	0	0	0	0
pValue CAPM	0.5544	0.93304	0	0	0	0	0
Coefficient FF 3F	0.0013838	-0.013637	0.13083	-0.099824	0	0	0
pValue FF 3F	0.49185	0.75934	0.24622	0.2149	0	0	0
Coefficient FF 6F	0.0011579	0.035559	0.09438	-0.24868	-0.2107	0.36214	0.077937
pValue FF 6F	0.58771	0.51708	0.42929	0.059653	0.16493	0.048645	0.17523

Table 8.11: Factor Analysis of the EWMA Crossover signal with a equally weighted weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.00095885	0.021504	0.024321	0.0006607	-0.10639	0.036246
pValue AF	0.65097	0.53956	0.78615	0.99169	0.54727	0.84571

Table 8.12: Correlation analysis with major indices of the EWMA Crossover signal with a equally weighted weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass
Coefficient SR	3.9144	45.8839	-13.1722	-7.6264	-7.4412	-1.7151	2.1085
pValue SR	5.2913e-09	9.7222e-13	2.1496e-12	2.3716e-09	1.7805e-11	0.063346	0.0065274

Table 8.13: Correlation study of the EWMA Crossover signal with a equally weighted weighting scheme.

8.2.4 Rule on overall trend quantity - Volatility Parity

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0044206	-0.00038004	0	0	0	0	0
pValue CAPM	0.039915	0.99359	0	0	0	0	0
Coefficient FF 3F	0.0042553	-0.0040058	0.14322	0.011949	0	0	0
pValue FF 3F	0.050095	0.93329	0.23775	0.89	0	0	0
Coefficient FF 6F	0.003978	0.03844	0.16418	-0.21761	-0.1386	0.44924	-0.0061621
pValue FF 6F	0.085908	0.51726	0.20407	0.12705	0.39772	0.023904	0.9209

Table 8.14: Factor Analysis of the EWMA Crossover signal with a volatility parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0040306	0.011132	0.20194	-0.070945	-0.28924	0.002942
pValue AF	0.065455	0.75751	0.029351	0.2774	0.1126	0.98775

Table 8.15: Correlation analysis with major indices of the EWMA Crossover signal with a volatility parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R^2
Coefficient SR	0.29453	4.152	-0.95425	-1.5221	1.822	-0.74815	-0.12701	0.42738
pValue SR	0.23467	0.07258	0.15729	0.0014529	1.0664e-05	0.038939	0.67185	0

Table 8.16: Correlation study of the EWMA Crossover signal with a volatility parity weighting scheme.

8.2.5 Rule on overall trend quantity - Risk Parity

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0052961	0.0097748	0	0	0	0	0
pValue CAPM	0.02454	0.85026	0	0	0	0	0
Coefficient FF 3F	0.0052111	0.0046131	0.1439	-0.013408	0	0	0
pValue FF 3F	0.028521	0.92981	0.27829	0.88724	0	0	0
Coefficient FF 6F	0.0051108	0.039401	0.1818	-0.27592	-0.1574	0.49147	-0.039141
pValue FF 6F	0.044051	0.5442	0.19885	0.077375	0.38023	0.023967	0.56458

Table 8.17: Factor Analysis of the EWMA Crossover signal with a risk parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0049503	0.033392	0.2806	-0.091619	-0.26418	-0.039316
pValue AF	0.036424	0.3916	0.0052166	0.19436	0.17945	0.84938

Table 8.18: Correlation analysis with major indices of the EWMA Crossover signal with a risk parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R^2
Coefficient SR	-0.1638	7.0631	-1.2938	-1.8786	2.2299	-2.1591	0.75719	0.52024
pValue SR	0.57442	0.01003	0.10432	0.00087832	5.1501e-06	1.0531e-06	0.033543	0

Table 8.19: Correlation study of the EWMA Crossover signal with a risk parity weighting scheme.

8.2.6 Rule on overall trend quantity - Equally Weighted

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0036464	0.0062991	0	0	0	0	0
pValue CAPM	0.082263	0.89161	0	0	0	0	0
Coefficient FF 3F	0.0036265	0.0028238	0.082013	-0.016572	0	0	0
pValue FF 3F	0.087635	0.95191	0.48919	0.8446	0	0	0
Coefficient FF 6F	0.0036304	0.031274	0.10904	-0.2408	-0.1624	0.42617	-0.028995
pValue FF 6F	0.10911	0.59028	0.3883	0.084696	0.3114	0.02851	0.63314

Table 8.20: Factor Analysis of the EWMA Crossover signal with a equally weighted weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0036711	0.053356	0.20353	-0.065848	-0.25697	-0.11826
pValue AF	0.082588	0.12682	0.023209	0.29708	0.14474	0.52354

Table 8.21: Correlation analysis with major indices of the EWMA Crossover signal with a equally weighted weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R^2
Coefficient SR	0.68429	1.4398	-0.18944	-2.1187	2.1661	-0.035869	-0.17514	0.40945
pValue SR	0.015883	0.58176	0.80417	0.00010989	4.3584e-06	0.93005	0.60719	0

Table 8.22: Correlation study of the EWMA Crossover signal with a equally weighted weighting scheme.

8.2.7 Sensitivity Analysis

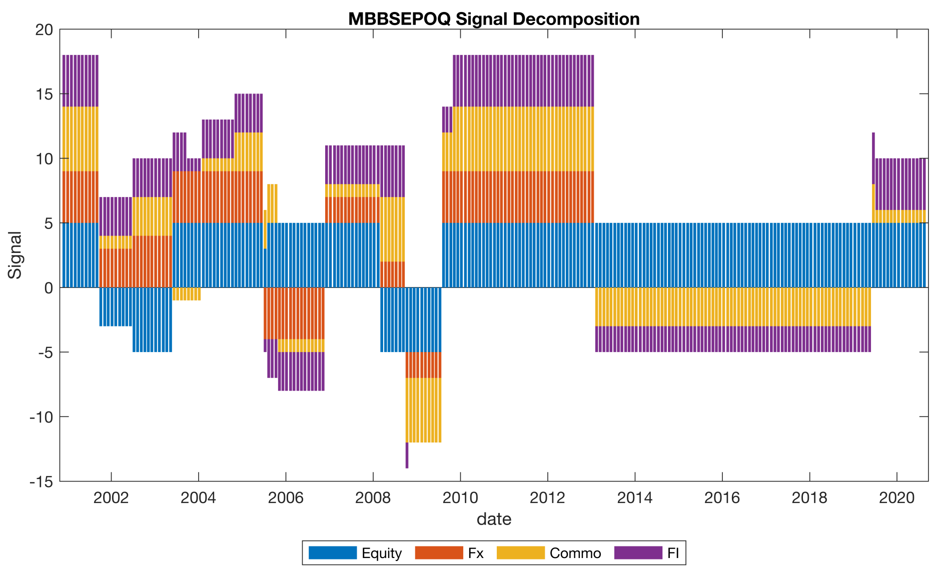
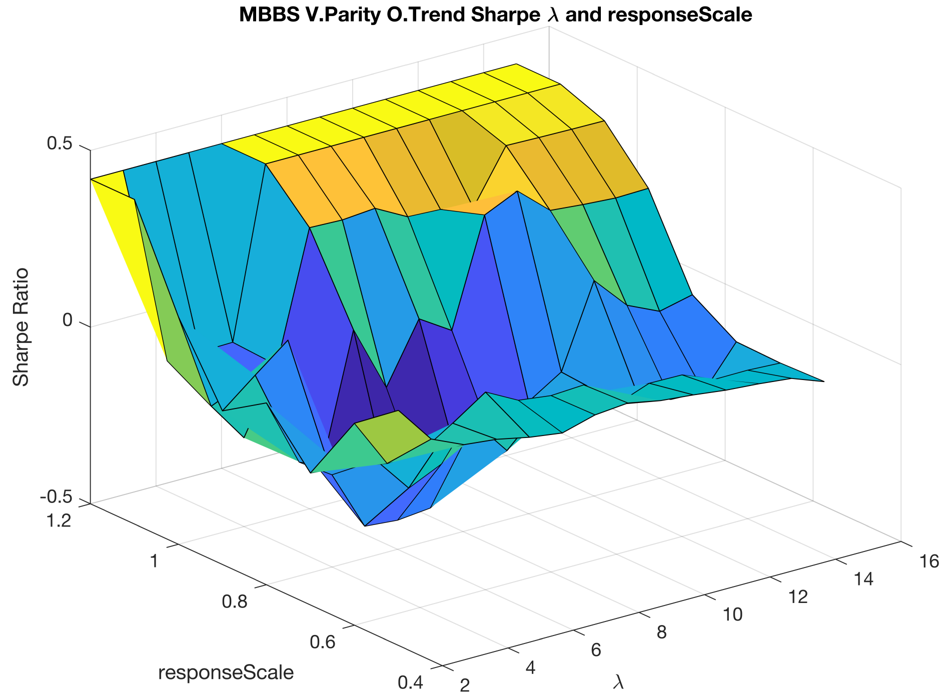


Figure 8.7: Decomposition of MBBS Signal by asset class, it is the sum of : $S_i = 1|S_i > 0$ and -1 otherwise. Therefore, it is not signal weighted.



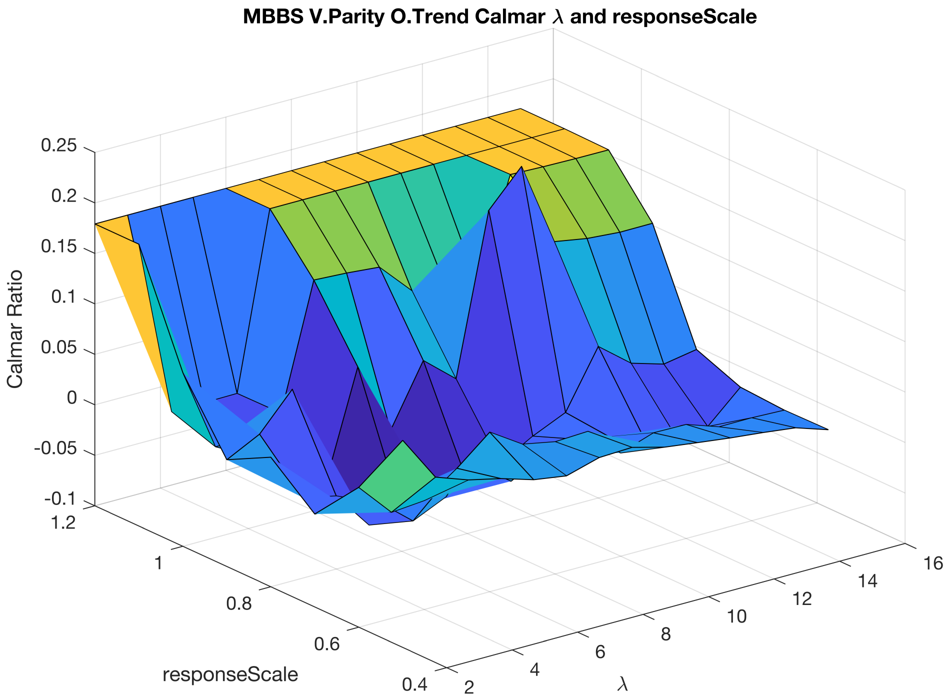


Figure 8.8: MBBS Volatility Parity strategy Sharpe and Calmar with λ and responseScale varying

8.3 SSA Model

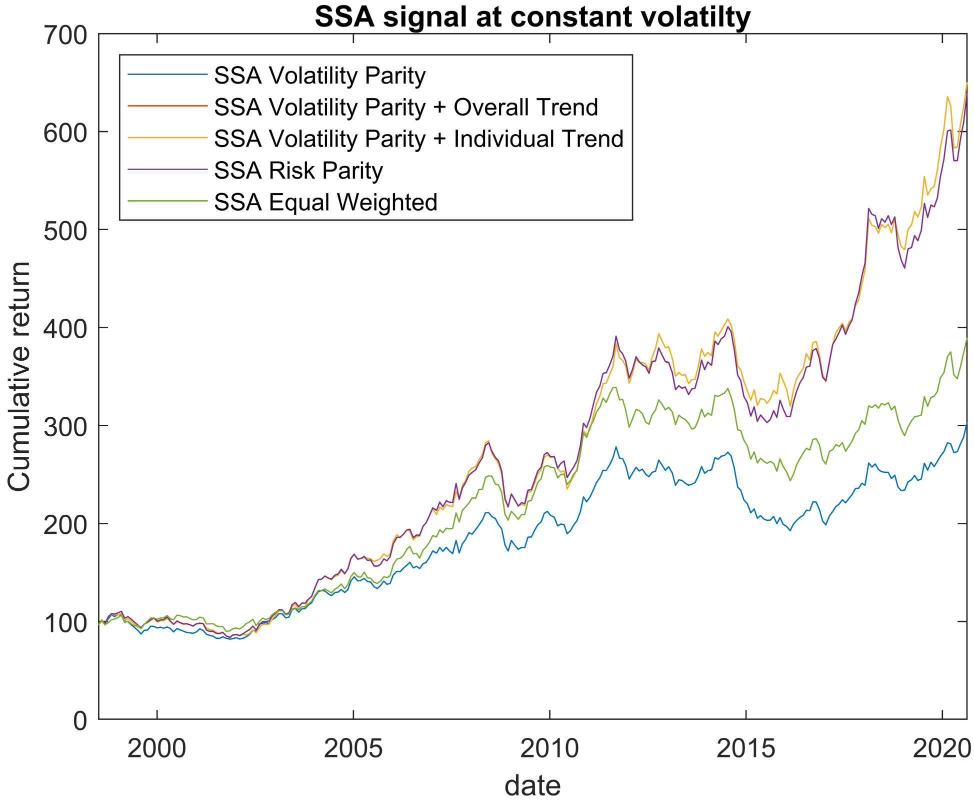


Figure 8.9: Cumulative returns for different allocations using the Singular Spectrum Analysis signal.

OriginalVariableNames	AnnualizedMean	AnnualizedVolatility	Kurtosis	Skewness	AverageMonthlyTurnover	SharpeRatio	CalmarRatio	MaximumDrawDown	HH
Vol.Parity	0.051354	0.10197	3.128	0.045493	0.45056	0.40556	0.16701	0.30748	0.34222
R.Parity	0.087406	0.11302	3.3763	-0.099217	0.97909	0.68492	0.35781	0.24428	0.50366
EW	0.063257	0.1021	3.1236	-0.22371	0.44673	0.52164	0.22536	0.2807	0.11349
R.Parity O.Quantity	0.087406	0.11302	3.3763	-0.099217	0.97909	0.68492	0.35781	0.24428	0.50366
R.Parity I.Quantity	0.088263	0.11327	3.2167	-0.024299	0.78733	0.69094	0.37268	0.23683	0.47696

Table 8.23: Descriptive Statistics of the Singular Sprectal Analysis signal with a volatility parity weighting scheme.

8.3.1 Volatility Parity

	Var1
Annualized Mean	0.051354
Annualized Volatility	0.10197
Kurtosis	3.128
Skewness	0.045493
Average Monthly Turnover	0.45056
Sharpe Ratio	0.40556
Calmar Ratio	0.16701
Maximum DrawDown	0.30748
HH*	0.34222

Table 8.24: Descriptive Statistics of the Singular Sprectal Analysis signal with a volatility parity weighting scheme.

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0032838	0.25827	0	0	0	0	0
pValue CAPM	0.049955	1.4758e-11	0	0	0	0	0
Coefficient FF 3F	0.0032346	0.257	0.055594	-0.0042547	0	0	0
pValue FF 3F	0.054778	3.1044e-11	0.52269	0.9464	0	0	0
Coefficient FF 6F	0.0020843	0.30324	0.083754	-0.060182	0.18279	0.10342	0.02131
pValue FF 6F	0.24785	6.1631e-10	0.37195	0.58046	0.15291	0.48804	0.64181

Table 8.25: Factor Analysis of the Singular Sprectal Analysis signal with a volatility parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0046734	0.015443	-0.053539	0.0048319	-0.062849	-0.05995
pValue AF	0.016822	0.62331	0.49384	0.9278	0.69159	0.721

Table 8.26: Correlation analysis with major indices of the Singular Sprectal Analysis signal with a volatility parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R^2
Coefficient SR	-3.585	-2.555	0.17087	2.8483	3.551	-0.46535	2.1382	0.7508
pValue SR	1.4391e-23	0.22063	0.78826	8.5147e-07	2.1716e-12	0.3176	6.8906e-07	0

Table 8.27: Correlation study of the Singular Sprectal Analysis signal with a volatility parity weighting scheme.

8.3.2 Equally Weighted

	Var1
Annualized Mean	0.063257
Annualized Volatility	0.1021
Kurtosis	3.1236
Skewness	-0.22371
Average Monthly Turnover	0.44673
Sharpe Ratio	0.52164
Calmar Ratio	0.22536
Maximum DrawDown	0.2807
HH*	0.11349

Table 8.28: Descriptive Statistics of the Singular Sprectal Analysis signal with a equally weighted weighting scheme.

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0038838	0.33247	0	0	0	0	0
pValue CAPM	0.013575	3.2805e-19	0	0	0	0	0
Coefficient FF 3F	0.0038175	0.32959	0.091298	-0.017572	0	0	0
pValue FF 3F	0.015629	1.3016e-18	0.26232	0.76668	0	0	0
Coefficient FF 6F	0.0032229	0.34012	0.095782	0.046083	0.13899	-0.10799	0.029257
pValue FF 6F	0.056932	2.8947e-13	0.27577	0.65129	0.24551	0.43949	0.49536

Table 8.29: Factor Analysis of the Singular Sprectal Analysis signal with a equally weighted weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0063623	0.013305	-0.050503	-0.017683	-0.2301	-0.21677
pValue AF	0.0011836	0.67183	0.51819	0.73992	0.14673	0.19678

Table 8.30: Correlation analysis with major indices of the Singular Sprectal Analysis signal with a equally weighted weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R
Coefficient SR	-3.8482	-5.9402	1.1037	3.0426	4.5571	-3.0286	3.6526	0.
pValue SR	4.1023e-24	0.0075389	0.10174	6.3936e-07	1.6178e-16	6.1359e-09	3.7527e-14	0

Table 8.31: Correlation study of the Singular Sprectal Analysis signal with a equally weighted weighting scheme.

8.3.3 Risk Parity

	Var1
Annualized Mean	0.087406
Annualized Volatility	0.11302
Kurtosis	3.3763
Skewness	-0.099217
Average Monthly Turnover	0.97909
Sharpe Ratio	0.68492
Calmar Ratio	0.35781
Maximum DrawDown	0.24428
HH*	0.50366

Table 8.32: Descriptive Statistics of the Singular Sprectal Analysis signal with a risk parity weighting scheme.

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0058901	0.32647	0	0	0	0	0
pValue CAPM	0.001172	5.8233e-15	0	0	0	0	0
Coefficient FF 3F	0.0058116	0.32414	0.093046	-0.0099118	0	0	0
pValue FF 3F	0.0014266	1.6304e-14	0.32006	0.8842	0	0	0
Coefficient FF 6F	0.0044745	0.37278	0.12085	-0.038017	0.22742	0.057473	0.033216
pValue FF 6F	0.021356	2.4354e-12	0.23051	0.745	0.097889	0.71959	0.4996

Table 8.33: Factor Analysis of the Singular Sprectal Analysis signal with a risk parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0081732	0.015208	-0.0539	0.00044193	-0.19207	-0.20496
pValue AF	0.00018332	0.66261	0.53434	0.99404	0.27465	0.27126

Table 8.34: Correlation analysis with major indices of the Singular Sprectal Analysis signal with a risk parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	*R*
Coefficient SR	-3.8482	-5.9402	1.1037	3.0426	4.5571	-3.0286	3.6526	0.
pValue SR	4.1023e-24	0.0075389	0.10174	6.3936e-07	1.6178e-16	6.1359e-09	3.7527e-14	0

Table 8.35: Correlation study of the Singular Sprectal Analysis signal with a risk parity weighting scheme.

8.3.4 Sensitivity Analysis

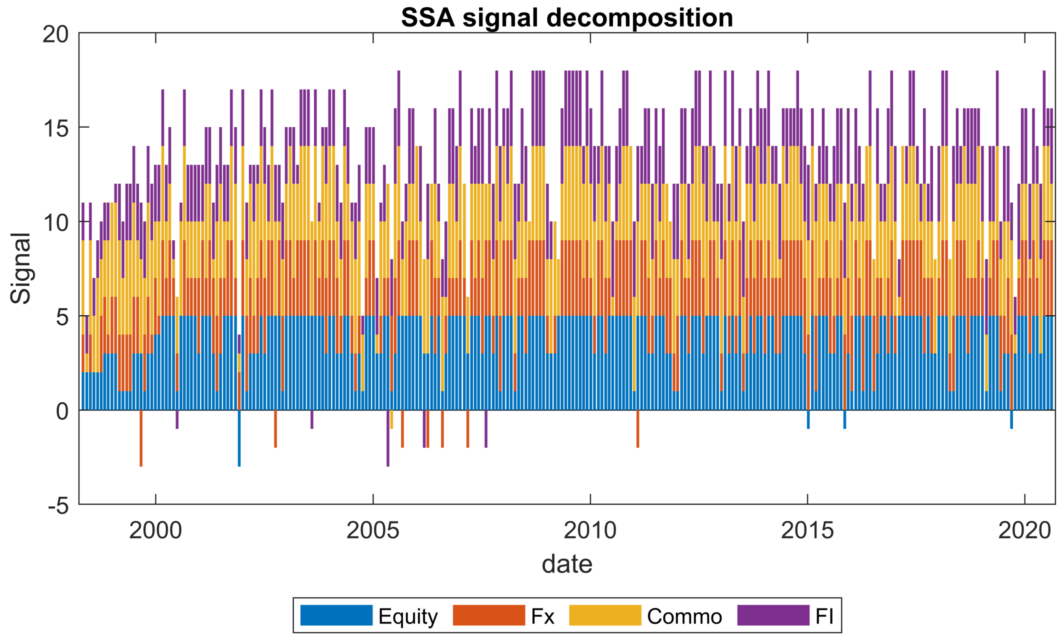


Figure 8.10: Decomposition of SSA Signal by asset class, it is the sum of : $S_i = 1|S_i > 0$ and -1 otherwise. Therefore, it is not signal weighted.

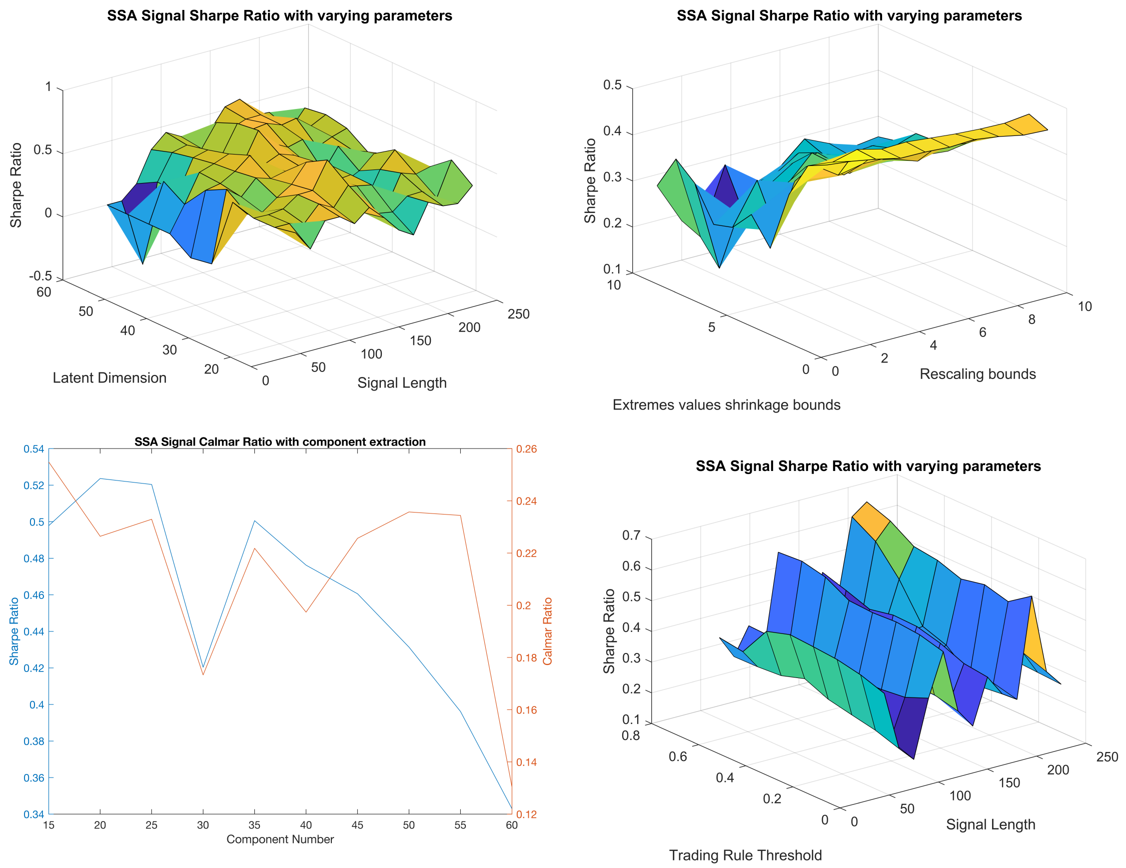


Figure 8.11: SSA Volatility Parity strategy Sharpe Ratio with varying parameters

8.4 SVM Model

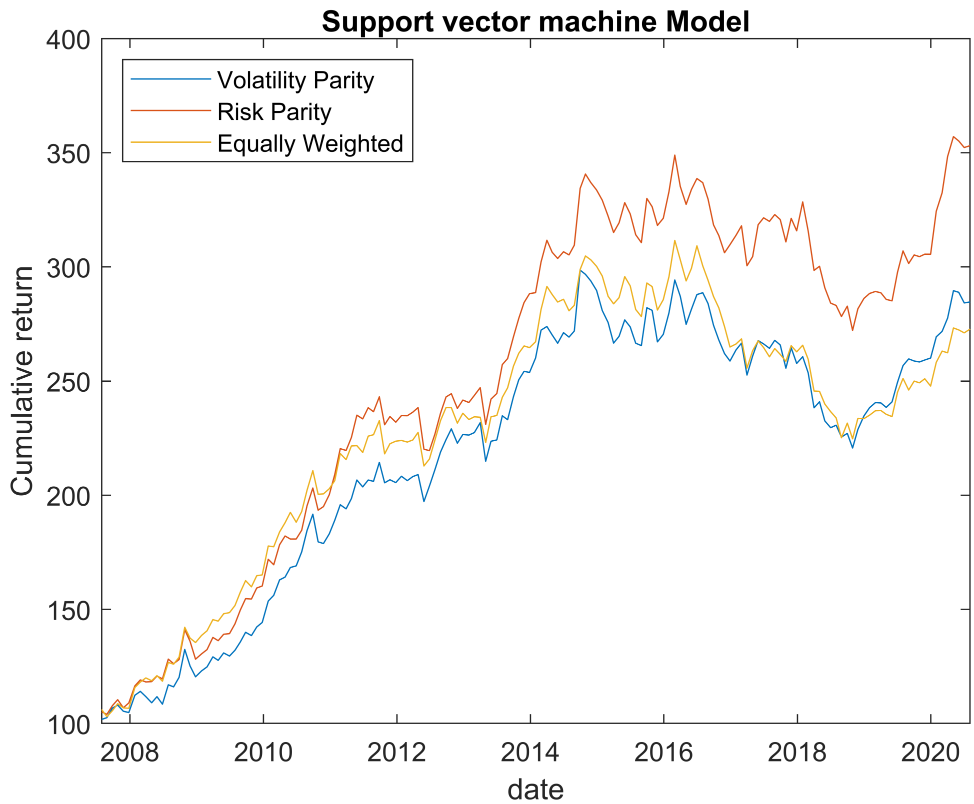


Figure 8.12: Cumulative returns for different allocations using the Support Vector machine based signal.

OriginalVariableNames	AnnualizedMean	AnnualizedVolatility	Kurtosis	Skewness	AverageMonthlyTurnover	SharpeRatio	CalmarRatio	MaximumDrawDown	HH
Vol.Parity	0.083236	0.10291	3.6089	0.084397	0.59256	0.71167	0.31908	0.26086	1.0339
R.Parity	0.10103	0.10337	3.3495	0.036027	1.3115	0.88069	0.45962	0.21982	0.86007
EW	0.079714	0.098564	3.4089	0.26096	0.57877	0.70729	0.28588	0.27884	0.3016
Vol.Parity NR	0.067012	0.10087	3.1783	0.21191	0.64434	0.56522	0.30981	0.2163	0.93022
R.Parity NR	0.07242	0.10061	3.0118	0.24974	1.5004	0.62042	0.32402	0.2235	0.77236
EW NR	0.037308	0.10534	3.8701	-0.061114	0.63081	0.25922	0.13229	0.28202	0.23652

Table 8.36: Descriptive Statistics of the Support Vector Machine signal with a volatility parity weighting scheme.

8.4.1 Volatility Parity

	Var1
Annualized Mean	0.083236
Annualized Volatility	0.10291
Kurtosis	3.6089
Skewness	0.084397
Average Monthly Turnover	0.59256
Sharpe Ratio	0.71167
Calmar Ratio	0.31908
Maximum DrawDown	0.26086
HH*	1.0339

Table 8.37: Descriptive Statistics of the Support Vector Machine signal with a volatility parity weighting scheme.

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0072949	-0.046707	0	0	0	0	0
pValue CAPM	0.0026043	0.34721	0	0	0	0	0
Coefficient FF 3F	0.0074721	-0.044742	-0.17744	0.094853	0	0	0
pValue FF 3F	0.0025698	0.38046	0.28703	0.42826	0	0	0
Coefficient FF 6F	0.0077723	-0.035282	-0.17868	-0.069945	-0.19595	0.18574	-0.042865
pValue FF 6F	0.0032635	0.57372	0.29923	0.72354	0.46654	0.46121	0.61899

Table 8.38: Factor Analysis of the Support Vector Machine signal with a volatility parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0048126	0.037599	-0.042541	0.069216	0.46117	0.64347
pValue AF	0.057302	0.42234	0.68365	0.38978	0.023072	0.0023984

Table 8.39: Correlation analysis with major indices of the Support Vector Machine signal with a volatility parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R^2
Coefficient SR	-4.9222	8.0561	2.1611	2.4828	0.98672	1.9226	2.1271	0.869
pValue SR	4.5018e-14	0.0080458	0.0096389	5.5141e-08	0.25562	0.0031228	0.010322	0

Table 8.40: Correlation study of the Support Vector Machine signal with a volatility parity weighting scheme.

8.4.2 Equally Weighted

	Var1
Annualized Mean	0.079714
Annualized Volatility	0.098564
Kurtosis	3.4089
Skewness	0.26096
Average Monthly Turnover	0.57877
Sharpe Ratio	0.70729
Calmar Ratio	0.28588
Maximum DrawDown	0.27884
HH*	0.3016

Table 8.41: Descriptive Statistics of the Support Vector Machine signal with a equally weighted weighting scheme.

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0069833	-0.046451	0	0	0	0	0
pValue CAPM	0.0026112	0.32897	0	0	0	0	0
Coefficient FF 3F	0.0073976	-0.054188	0.010136	0.10072	0	0	0
pValue FF 3F	0.0018962	0.26894	0.9494	0.38109	0	0	0
Coefficient FF 6F	0.006886	-0.026309	0.040671	-0.012101	0.065663	0.23455	-0.038696
pValue FF 6F	0.0065648	0.66235	0.80544	0.94921	0.7994	0.33319	0.64035

Table 8.42: Factor Analysis of the Support Vector Machine signal with a equally weighted weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0051006	0.026242	-0.13763	0.10521	0.36126	0.54437
pValue AF	0.03539	0.55796	0.16929	0.17247	0.061977	0.0070505

Table 8.43: Correlation analysis with major indices of the Support Vector Machine signal with a equally weighted weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R^2
Coefficient SR	-4.5644	10.8347	1.6739	2.1351	1.5084	2.5722	0.40884	0.8571
pValue SR	4.7974e-10	0.0024842	0.084726	4.1149e-05	0.13911	0.00079883	0.66987	0

Table 8.44: Correlation study of the Support Vector Machine signal with a equally weighted weighting scheme.

8.4.3 Risk Parity

	Var1
Annualized Mean	0.10103
Annualized Volatility	0.10337
Kurtosis	3.3495
Skewness	0.036027
Average Monthly Turnover	1.3115
Sharpe Ratio	0.88069
Calmar Ratio	0.45962
Maximum DrawDown	0.21982
HH*	0.86007

Table 8.45: Descriptive Statistics of the Support Vector Machine signal with a risk parity weighting scheme.

	intercept	MktRF	SMB	HML	RMW	CMA	Mom
Coefficient CAPM	0.0086302	-0.039085	0	0	0	0	0
pValue CAPM	0.00042528	0.43374	0	0	0	0	0
Coefficient FF 3F	0.0087319	-0.036143	-0.16442	0.072272	0	0	0
pValue FF 3F	0.00049653	0.48134	0.32691	0.54858	0	0	0
Coefficient FF 6F	0.0090321	-0.033694	-0.17497	0.037879	-0.14858	0.024343	0.022654
pValue FF 6F	0.00074788	0.59451	0.31383	0.84945	0.58424	0.92371	0.79452

Table 8.46: Factor Analysis of the Support Vector Machine signal with a risk parity weighting scheme.

	intercept	GSCICommo	MSCIWORLD	MSCIEM	USDindex	GlobalBonds
Coefficient AF	0.0064246	-0.0022975	-0.022312	0.04134	0.27223	0.55652
pValue AF	0.012271	0.9612	0.8322	0.61038	0.1812	0.008974

Table 8.47: Correlation analysis with major indices of the Support Vector Machine signal with a risk parity weighting scheme.

	intercept	All	Equity	Fx	Commo	FI	InClass	R^2
Coefficient SR	-4.5307	10.4817	1.441	2.3095	0.7381	1.5181	1.8217	0.84675
pValue SR	2.5558e-11	0.0014043	0.10429	1.7018e-06	0.427	0.028126	0.039458	0

Table 8.48: Correlation study of the Support Vector Machine signal with a risk parity weighting scheme.

8.4.4 Sensitivity analysis

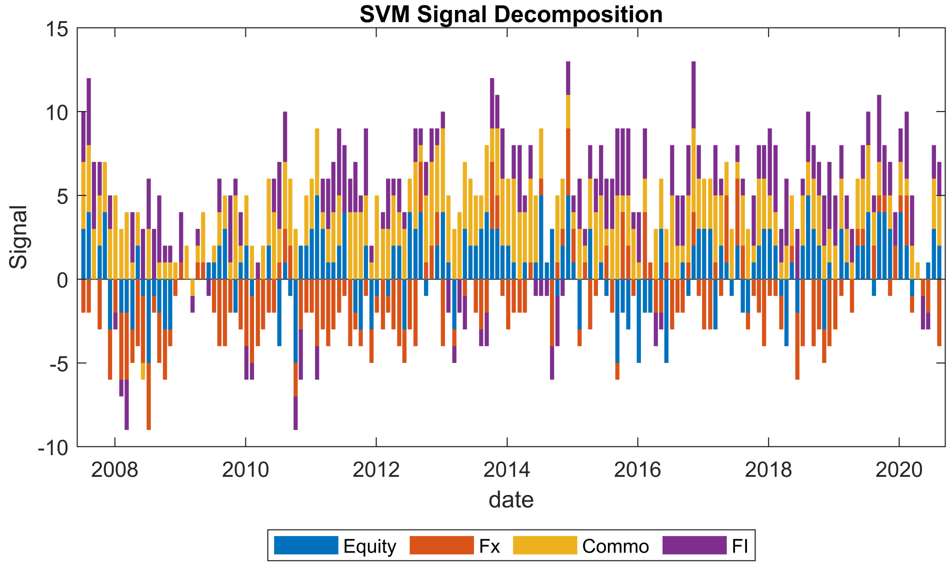


Figure 8.13: Support Vector Machine signal decomposition, it is the sum of : $S_i = 1|S_i > 0$ and -1 otherwise. Therefore, it is not signal weighted.

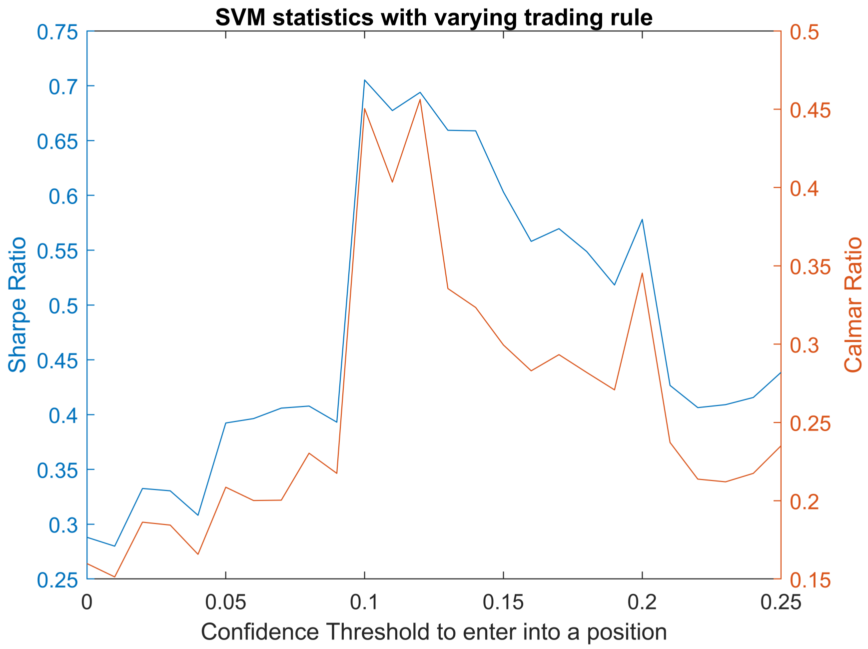


Figure 8.14: Sensitivity of the SVM strategy to the trading rule threshold

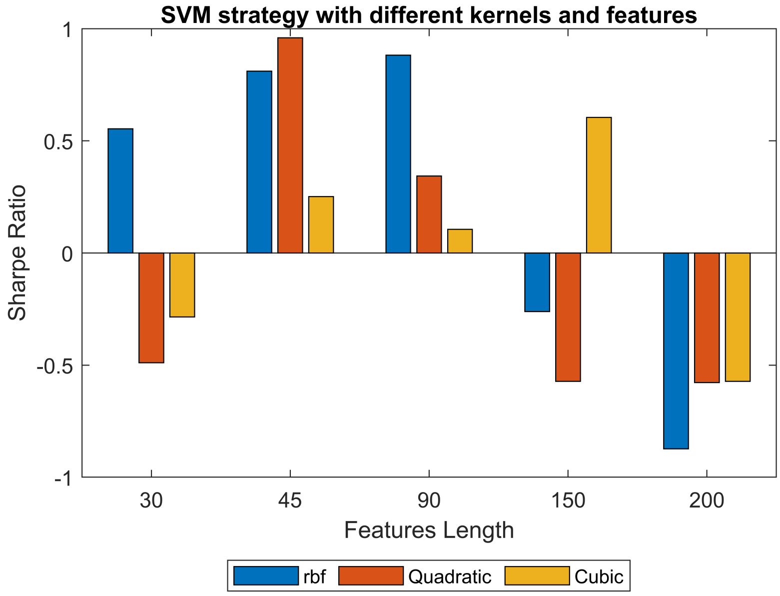


Figure 8.15: Sensitivity of the SVM strategy to the features and Kernels

8.5 Return Decomposition

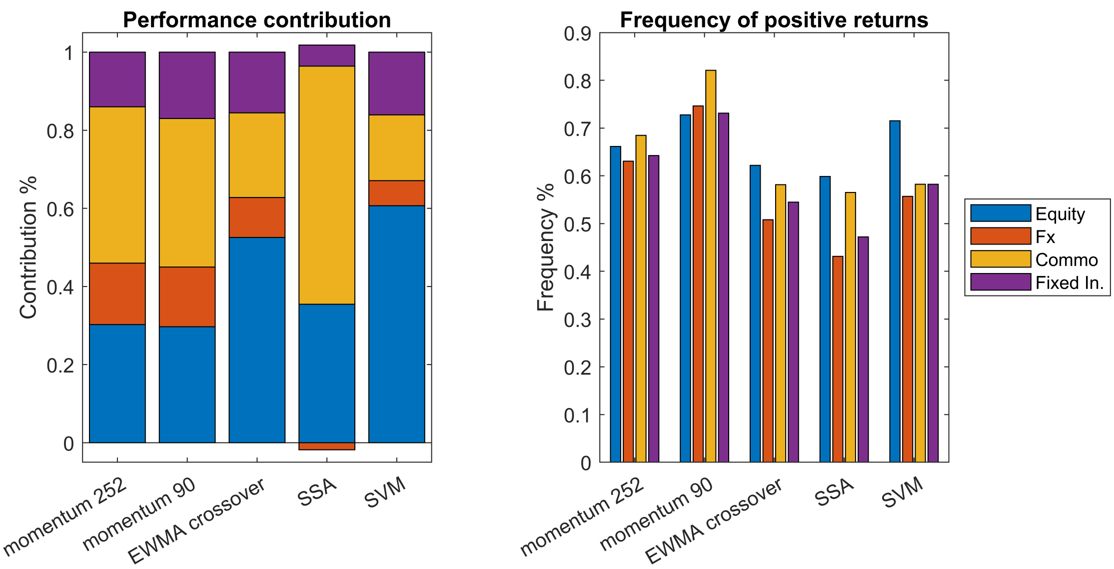


Figure 8.16: Decomposition of the total return of the strategies by asset class as well "precision" of the model, i.e the percentage of periods for which the strategy yielded positive return.

8.6 Correlation Matrix

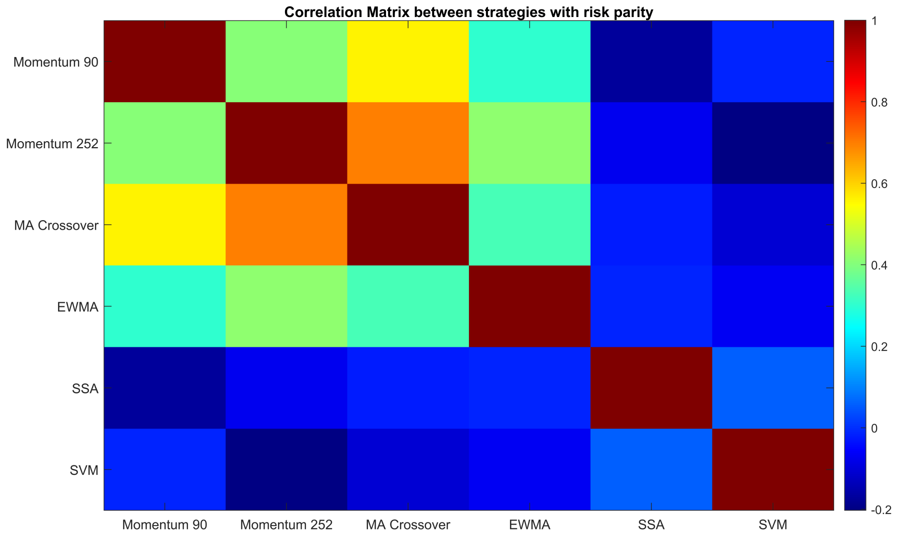


Figure 8.17: Correlation matrix between the strategies

References

[XWa05] X.Wang, A.Smith, R.Hyndman. "Dimension Reduction for Clustering Time Series Using Globl Characteristics". In: *Springer* (2005).

[TAl09] T.Alexandrov. "A Method of trend extraction using singular spectrum analysis". In: *Statistical Journal* 7.1 (2009), pp. 1–22.

[Bru11] Bruder, Dao, Richard and Roncalli. "Trend Filtering Methods for Momentum Strategies: Lyxor Asset Management". In: (2011).

[MP12] Ooi Moskowitz and Pedersen. "Time series momentum". In: *Journal of financial Economics* 104 (2012), pp. 228–250. DOI: 10.1016/j.jfineco.2011.11.003. URL: doi:10.1016/j.jfineco.2011.11.003.

[Bal15] Baltas. "Trend-Following, Risk-Parity and the Influence of Correlations". In: *Risk Based and Factor Investing* (2015).

[Jam15] Jamil Baz, Nick Granger, Campbell R. Harvey, Nicolas Le Roux and Sandy Rattray. "Dissecting Investment Strategies in the Cross Section and Time Series." In: (2015). DOI: 10.2139/ssrn.2695101. (Visited on).

[Xio15] Xiong and Ibbotson. "Momemtum Acceleration and Reversal". In: *Journal of Investment Management* 13.1 (2015), pp. 84–95.

[Dan16] Daniel and Moskowitz. "Momentum Crahses". In: *Journal of financial Economics* 122 (2016), pp. 221–247. DOI: 10.1016/j.jfineco.2015.12.002.

[Lev16] Levine and Pedersen. "Which Trend Is Your Friend?" In: *Financial Analyst Journal* 72.3 (2016), pp. 51–66.

[Yu 16] Yu and Webb. "CAN FUNDAMENTAL FACTORS ENHANCE THE PERFORMANCE OF TRADITIONAL MOMENTUM STRATEGIES?" In: *Journal of Investment Management* 14.4 (2016), pp. 28–43.

[Fed17] Federico Nucera. "Risk-Managed Momentum: The effect of leverage constraints". In: (2017). URL: https://ssrn.com/abstract=3008530.

[Bal19] Baltas and Kosowski. "Demystifying Time-series Momentum Strategies: Volatility Estimators, Trading Rules and Pairwise Correlations". In: *SSRN Electronic Journal* (2019). DOI: 10.2139/ssrn.2140091. URL: https://ssrn.com/abstract=2140091.

[BS20] Michal Balcerak and Thomas Schmelzer. "Constructing trading strategy ensembles by classifying market states". In: (2020). arXiv: 2012.03078 [q-fin.TR].

[Ant] Antonacci. "Optimal Momentum: A Global Cross Asset Approach". In: (). URL: https://ssrn.com/abstract=1833722.

[B H] B. Hurst, Y.Ooi, L.Pedersen. "A Century of Evidence on Trend-Following Inesting". In: (). URL: https://ssrn.com/abstract=2993026.

[CLI13] CLIFFORD S. ASNESS TOBIAS J. MOSKOWITZ LASSE HEJE PEDERSEN. "Value and Momentum Everywhere". In: *The Journal of Finance* (30 January 2013). URL: https://doi.org/10.1111/jofi.12021.

Appendix

A General formula

This formula allows us to compute the return of the strategy at teach time using :

- $$R_{i,t} = \frac{F_{i,t} - F_{i,t-1}}{F_{i,t-1}} \quad \text{the simple return} \tag{A.1}$$
- $\frac{\sigma_{Target}}{\sigma_{t,t+1}^{Portfolio}}$ the constant volatility ratio.
- $signal(r_{t-j,t}^{i})$ the signal generated by the chosen formula.
- $GrossWeight_{i,t}^{WeightingScheme}$ the weights generated by the weighting scheme.

$$r_{t,t+1}^{WeightingScheme} = \frac{\sigma_{Target}}{\sigma_{t,t+1}^{Portfolio}} * \sum_{i=1}^{N_t} signal(r_{t-j,t}^{i}) * GrossWeight_{i,t}^{WeightingScheme} * r_{t,t+1}^{i} \tag{A.2}$$

B Signals

A1 Momemtum

We are computing the cumulative return for the M previous days. The signal is simply $+1$ for the positive returns and -1 for the negatives ones.

$$\begin{aligned} & signal_{t,i}^{M} = F(\prod_{j=t-M}^{t} (1 + r_j) - 1) \\ & where : \\ & F(X) = [-1 : X \leq 0, 1 : X > 0] \end{aligned} \tag{B.1}$$

A2 Moving Average Crossover

We use a simple moving avergae crossover with the following rule :

$$\begin{aligned} & \Delta = MA_{ShortTerm} - MA_{LongTerm} \\ & \text{if} \Delta <= 0, Signal = -1 \\ & \text{else} Signal = 1. \end{aligned} \tag{B.2}$$

A3 Exponentially Weighted Moving average crossover

We use a signal based on [Jam15] which is a CTA-momentum signal based on the crossover of Exponentially Weighted Moving average:

- We need three short / long moving average time scale. In our case, we use $S_k = [8, 16, 32]$ and $L_K = [24, 48, 96]$.
- For each $K \in 1, 2, 3$, we compute :
  $x_k = EWMA[P|S_k] - EWMA[P|L_K]$
- We then perform a two step normalization:
  $y_k = \frac{x_k}{\sigma_{P-63:P}}$
  $z_k = \frac{y_k}{\sigma_{y_{k,1:252}}}$
- The next step is to compute an intermediate signal u_k via an activation function:

$$\begin{cases} F(X) = \frac{x * exp(\frac{-x^2}{4})}{0.89} \\ u_k = F(z_k) \end{cases}$$

- Finally, we compute the average of this average signal (we could put different weights for each length). $S_{CTA} = \sum_{k=1}^{3} w_k * u_k$

The strong advantage of this signal is that it combines short term and long term moving average. Indeed, we have three dynamics of short/long term pair of values which are aggregated in the end.This signal is continuous between $\in [-1, 1]$ which can express strongly trending markets with close to ± 1 or non-trending market with a signal close to 0. This allowed us to implement the trading rule as presented in section A6 where the concept is that if the market is not trendy enough, we do not change the signal in order to avoid useless fees.

A4 Singular Spectrum Analysis

SSA decomposes components of a time series into trend, oscillation and noise components using a Principal Component Analysis (PCA) on the trajectory matrix of the underlying time-series. Starting with a time series $X_N = [X_1, X_2, ..., X_N]$, we create the trajectory matrix over M days that will have a size of $M * K$ with $K = N - M + 1$:

$$A_{m,n} = \begin{pmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_M & x_{M+1} & \cdots & x_N \end{pmatrix}$$

We then compute $C_{M,N} = COV(A_{M,N})$, the covariance matrix of the trajectory matrix of size MXM. We can then extract the eigenvalues (λ) and eigenvectors (ρ) of this matrix. Finally, we can reconstruct the components:

$$PC = \rho \cdot Y \tag{B.3}$$

This allows us to extract the signal which is the first principal components. Indeed, this first principal components will contain the bigger trend in the data. In the following example, the first PC is the orange curve which follows the general trend of the data[14]:

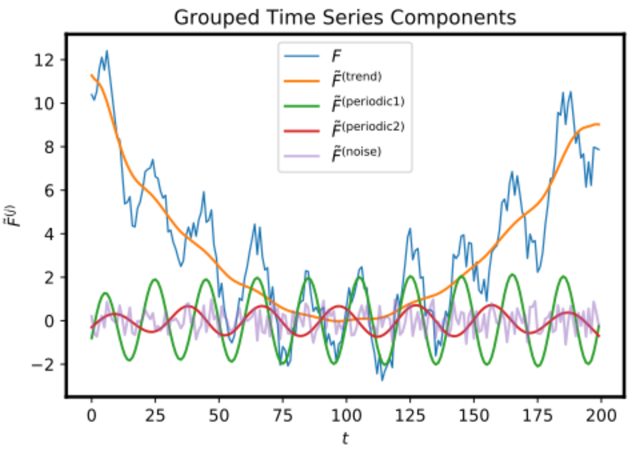


Figure A4.1: Principal Components of the SSA algorithm.

The last parameters to figure out is the length of the component we are extracting. Indeed, we need to compute the trend of this component. For our base model, we choose 90 days however, we will change it in the sensibility analysis. In order to have a finite interval of value, we rescale the signal so that $S_i \in (-1, +1)$.

A5 Support Vector Machine

We use support vector machine (SVM) to classify the time series into a binary signal $(+1/-1)$. The SVM algorithm allows for both classification and regression, for the purpose of this paper, we will use the classification framework[15]. More formally, the SVM algorithm is finding the hyperplane (a N-1 dimensional plane for N dimensional features) that is dividing the sample into to categories, with gap between the categories being as wide as possible. The support vectors are then the closest points to this boundary, one can see the following two dimensional example with kernels:

[14] Obviously, one could use either another component or a combination of components but taking the first component is the logical choice since it should express the general trend in the data, using another component could quickly become overfitting.

[15] One should note that a SVM regression based signal could as well work and would yield a continuous signal. However, it has been extensively shown that predicting the level of return of the following days is a hard task, hence, it seems more stable to use it in a classification framework.

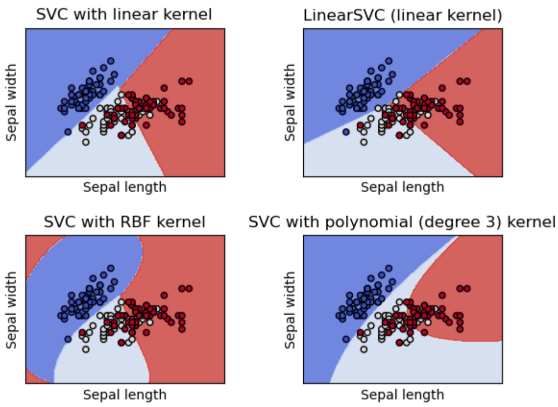


Figure A5.1: Two dimensional example of the SVM classification

However, in most real life application, the data are not linearly separable, nor separable on higher dimensional space, therefore one must use a Kernel function to add non-linearity into the model. More formally, the optimization for SVM classification for non separable data is the following[16]:

Starting with a linear model, we can define the margin as:
$y(x) = w^t * \phi(x) + b$
Which can be transformed assuming the data are separable:
$\arg\max_{w,b} \left(\frac{1}{w} min_n [t_n (w^T \phi(x_n) + b)] \right)$
Where the min values are the support vectors which allows us to transform the problem as follow:
$\arg\min_{w,b} \frac{1}{2} w^2$
$s.t. \forall n, t_n (w^T \phi(x_n) + b) \geq 1$
Which can be solved by Lagrangian method using Kuhn-Tucker.

(B.4)

In our case, we use the RBF[17] and we extract from the 90 days return series some properties that we then feed to the SVM algorithm with the corresponding signal which is the momentum signal of the next month. These properties are:

- Mean
- Standard deviation
- Kurtosis
- Skewness
- Mean[0, 1/3]

[16]Obviously the purpose of this paper is not propose a formal demonstration of the optimization problem. For the interested reader in a thorough and complete demonstration see.

[17]Radial Basis Function which one of the most used kernel for support vector machines.

- Mean[1/3, 2/3]
- Mean[2/3, 1][18]
- One hot encoded indicator of asset class belonging.

One of the prime feature of the SVM algorithm is that we can extract the posterior probability that an observation is in a given class (given the features).

$$P\left(S_j\right) = \frac{1}{1 + exp(A_{S_j} + B)} \tag{B.5}$$

This allows us to derive a simple trading rule such as "if the probability is smaller than 60%, don't enter into a position".For ease of computations, we directly use the classification score in the trading rule. A fine approximation of the score transform is $P(S_j) = \frac{1+S_j}{2}$. Therefore, a probability of 60% is a score of 20%. We choose the 20% heuristically, however we are testing this parameter in the sensitivity analysis.

A6 Trading rules on the quantity of trend

It is well know that when market exhibit no trend, trend following strategies are not working well. Indeed, to work these strategies need to have large market moves and therefore, tend to trade for nothing during these periods. Therefore, for some strategies, we have add a rule on total quantity of trend:

Using continuous signals that are mapped between -1 and + 1, we can define the total quantity of absolute trend as :

$$Q_{Trend} = \sum_{i=1}^{n} |S_i| \tag{B.6}$$

Where as the maximal quantity of signal is $n * 1 = n$. Therefore using a threshold $T \in [0, 1]$, we can express the following rule :

$$\begin{aligned} &\text{if} Q_T rend > T * n \\ &\text{Use the normal strategy} \\ &\text{else} \\ &\text{Go only long with the weighting scheme} \end{aligned} \tag{B.7}$$

Going only long allows for more stable weights and therefore, no useless fees and should not under perform since there is not a lot of trend. Moreover, by the CAPM, we can think that being long is the right choice since there should be a general return. We also added a rule on the individual quantity of trend which the simple following one : if $S_i < T$ then $W_i = 0$, we simply go out of this market when it does not have enough trend.

[18] The three means are proxies for the acceleration of the trend.

C Weighting Schemes

A1 Volatility Parity

Often called Naïve Risk-Parity because it ignores the non-diagonal value of the covariance matrix. The weighting scheme is the following :

$$W_t^i = \frac{(\sigma_t^i)^{-1}}{\sum_{i=1}^{N_t} (\sigma_t^i)^{-1}} \tag{C.1}$$

A2 Risk Parity Portfolio

Risk parity aims to distribute the weights in such a way that each assets contribute equally to the risk:

$$w_t^{RP,i} MRC_t^i(P) = constant, \forall i \tag{C.2}$$

This can be solve for any or any differentiable risk measure ρ and portfolio P using the following constrained optimization problem [19]:

$$L(w_t) = \sum_{i=1}^{N_t} log(w_t^i) - \lambda(\rho_{w_t} - \rho_{Target}) \tag{C.3}$$

Taking all the partial derivatives :

$$\frac{\partial L(w_t)}{\partial w_t^i} = \frac{1}{w_t^i} - \lambda \frac{\partial \rho_{w_t}}{\partial w_t^i}, \forall i, t \tag{C.4}$$

And setting these partial derivatives to zero :

$$w_t^{RP,i} MRC_t^i(P) = \frac{1}{\lambda} = constant, \forall i \tag{C.5}$$

And finally, since, in our case we are using the volatility that which homogeneous of degree one, we can rescale the weights to target any level of volatility.

[19] This approach is inspired from Baltas, 2015 which introduce this procedure for a long-short risk parity portfolio