Statistical Arbitrage - Backtest a generalized pairs-trading strategy on the stocks of the S&P500 index.

DEQUIDT Simon

May 2024

Abstract

This report presents a backtesting analysis of a generalized pairs trading strategy applied to stocks in the S&P 500 index. The methodology involves data preparation, hierarchical clustering, spread estimation, determination of trading signals, and performance evaluation. Different strategies are tested, and their performance metrics are compared to identify the most effective approach. The results provide insights into the profitability and risk associated with pairs trading strategies in different market conditions.

Contents

1 Introduction **3**

2 Methodology **4**

2.1 Data preparation 4

2.2 Hierarchical clustering 4

2.3 Spread estimation 5

2.4 Determination of trading signals 6

2.5 Performances of the strategies 7

3 Results & Analysis **8**

3.1 Basic strategies - review of the performances 8

3.1.1 Strategy 1 8

3.1.2 Strategy 2 11

3.2 More advanced strategies - improving backtest performances 15

3.2.1 Calibration thresholds 15

3.2.2 Revisited clustering approach 16

3.2.3 Revisited OLS estimation window for spread forecast 18

4 Conclusion **22**

5 Annex **24**

1 Introduction

Pairs trading is a statistical arbitrage strategy based on the notion of two stock prices "co-moving" with each other. When these prices diverge from their long-term equilibrium, the strategy involves shorting the overvalued stock and longing the undervalued one, thus profiting from the reconvergence of the prices. This relies on the principle of mean-reversion, which states that the spread (or price difference) between the pair will eventually revert to its historical mean. In a highly volatile market, pairs trading offers a compelling opportunity to exploit asset correlations without taking a directional market stance, thereby mitigating systemic risk.

Hedge funds have employed pairs trading for years, continuously reinventing their strategies to stand out and generate profit. With advancements in stock prediction methods, pairs trading has evolved, and many are now incorporating machine learning to enhance the base strategy and add value.

This report aims to evaluate the performance of a generalized pairs trading strategy on the stocks that compose the S&P 500 index. Unlike classic pairs trading, this approach involves trading one stock against a benchmark portfolio of stocks rather than another single stock. The methodology is straightforward: various backtest configurations will be implemented to identify an outstanding strategy capable of delivering excess returns.

We begin with data preparation, considering all S&P 500 stocks and assuming consistent data despite potential anomalies (e.g., due to index rebalancing). Next, we proceed with the clustering of the stocks in our universe to identify group of stocks that are affected similarly by some risk factors. Following this, we build spreads and optimize the model to forecast them. Once completed, we establish trading rules with diverse entry, exit, and stop-loss thresholds to compute and compare performances.

To assess returns effectively, we use performance metrics such as annualized returns and volatility, Sharpe ratio, maximum drawdown, and expected shortfall. Ultimately, this report aims to communicate the best backtesting configuration for a generalized pairs trading strategy.

2 Methodology

2.1 Data preparation

The investment universe for the implementation of the generalized pairs trading strategy is the S&P500 which is composed of the 500 stock with the highest market capitalization in the United States of America. Data spans from 01/01/2015 to 04/30/2024, and represents the daily returns across the global sectors of the S&P500. The data will be split into two subsets, one used for the calibration of the model (2015 – 2019) and another used for the out-of-sample analysis (2020-2024). Our Universe is composed of 428 stocks from the following sectors of the S&P500:

- Communication services
- Consumer Discretionary
- Consumer Staples
- Energy
- Financials
- Healthcare
- Industrials
- Information Technology
- Materials
- Utilities

2.2 Hierarchical clustering

As the starting point of the strategy, various methodologies have been implemented in the literature: distance approach, cointegration, stochastic modelling, copula methods, or some new machine learning techniques like clustering. In this work, we decide to operate hierarchical clustering based on historical returns of the stocks that form the universe. The clusters will be estimated over the learning period and is assumed to remain unchanged out-of-sample. To do so, we first normalized stock exposures to the first five principal components derived from the covariance matrix of the historical returns. Using Euclidean distance and Ward's linkage method, we cluster the stocks to minimize within-cluster variance. The resulting dendrogram helps identify closely related stocks that perform similarly, ideal for generalized pair-trading strategies. We then used the Gap statistic in order to select the right number of clusters. This method allowed us to select 7 cluster of stocks (Figure 1). Eventually, for each cluster of stocks, we extract the most important risk factor which is more precisely the first eigen-portfolio from PCA. The latter will be used to compute spreads.

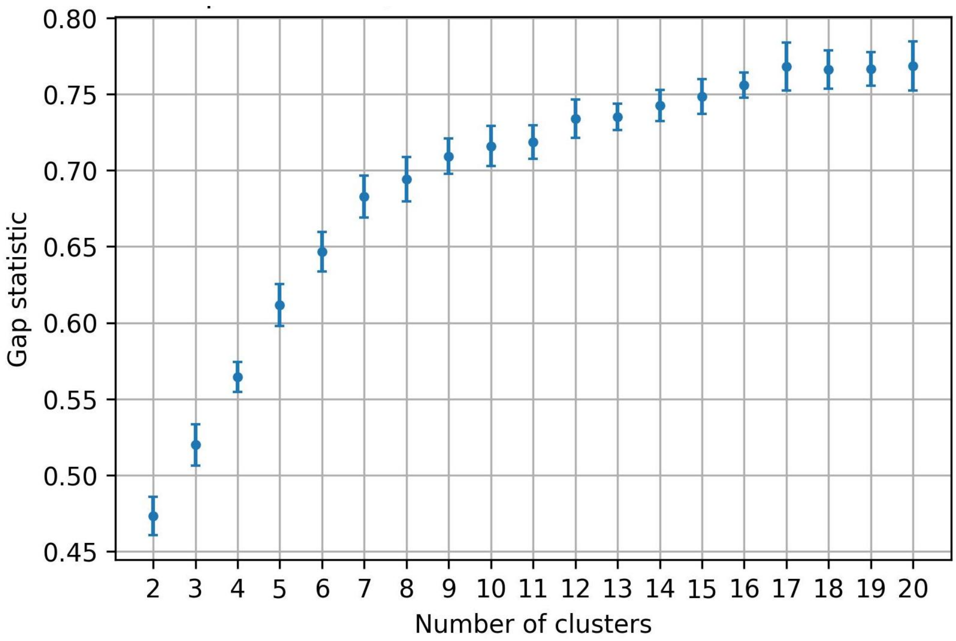

Figure 1: Gap Statistic - Euclidean distance

2.3 Spread estimation

Spreads estimation and forecasting are crucial to the strategy because it is the root of the signal generation. In our work, spreads will be computed as follows:

$$S_t = \sum_{t=t^*_{m-1}-T+1}^{t^*_{m-1}} \hat{\epsilon}_{i,t}$$

With

$$r_{i,t} = \hat{a} + \hat{b}F_{j,t} + \hat{\epsilon}_{i,t}$$

being the return of stock $stock_{i,t}$ at time t and $F_{j,t}$ the return of the main eigen-portfolio of cluster j at time t, over a given month, using an estimation window of T trading days.

Then, during month m, the forecasted spread, denoted as :

$$\tilde{S}_t = \sum_{t=t^*_{m-1}+1}^{t^*_m - 1} \tilde{\epsilon}_{i,t}$$

is computed using the parameters $\hat{a}$ and $\hat{b}$ estimated in $t = t^*_{m-1}$. The forecasted residuals are defined as:

$$\tilde{\epsilon}_{i,t} = r_{i,t} - \hat{a} - \hat{b}\tilde{F}_{j,t}$$

with $\tilde{F}_{j,t}$ the forecasted returns of the main eigen-portfolio of cluster j, assuming that the weights of the stocks in the eigen-portfolio are fixed over the forecast period, given by the eigenvector.

The computation of the parameters for the model is run through an OLS estimation which seems to be the most appropriate one. The coefficients $\hat{a}$ and $\hat{b}$ are estimated using $\beta = (X^T X)^{-1} X^T Y$, with Y the return of a given stock and X the forecasted returns of the main eigen-portfolio.

2.4 Determination of trading signals

After we have built forecasted spreads over the test period, it is essential to normalize them before using them through the determination of the signals. Normalization can be run using various methods, but we choose to use the most common one which subtracting each forecasted spread value with the fixed mean of the foreacted spread estimated over the estimation window and then dividing by its respective standard deviation.

Setting up trading rules is fundamental to improving the performance of a strategy. High turnover in trades leads to high transaction costs, so finding the right trade-off to maximize profit is essential. The best way to identify the most appropriate trading rules is by testing different thresholds. We experimented with two setups (see Table 1), varying the estimation window to see its impact on results. The periods used (in trading days) are: 252, 504, and 1260, resulting in six backtests.

After we defined the thresholds, below are the four trading signals that will allow us to run the strategy:

Table 1: Trading Signal Setups

	Short Stop	Short Open	Short Close	Long Close	Long Open	Long Stop
Setup 1	2.5	1.5	0.5	-0.5	-1.5	-2.5
Setup 2	-2.5	-1.5	-0.5	0.5	1.5	2.5

1. A short-spread position is opened if:
   (a) $ShortOpen < S_t^* < ShortStop$
2. A short-spread position is closed if:
   (a) $ShortStop < S_t^*$
   (b) $ShortClose > S_t^*$
3. A long-spread position is opened if:
   (a) $LongOpen < S_t^* < LongStop$
4. A long-spread position is closed if:
   (a) $LongClose < S_t^*$
   (b) $LongStop > S_t^*$

2.5 Performances of the strategies

Performance evaluation is the last step of the report. To make things easier, we assume a simplified allocation framework. At each date, the return of each cluster benchmark $r_{j,t}$ is computed as the mean return of the stocks that compose the cluster. The daily return of each traded spread position is then given by $r_{i,t} - r_{j,t}$ with $stock_i$ in $cluster_j$. The daily return of the global portfolio is finally determined by averaging the returns of all traded spread positions. This approach lets us evaluate the strategy's performance by assessing the collective behavior of the spread positions within the portfolio.

To evaluate the effectiveness of the strategies, we employ several key performance metrics. The annualized returns represent the annualized portfolio's compounded returns over the backtest period, providing a measure of overall performance. Annualized volatility calculates the annualized standard deviation of the returns, indicating the level of risk associated with the portfolio. The Sharpe Ratio, defined as the ratio of excess return (risk-free rate assumed to be 0, but can be discussed) to volatility, measures the risk-adjusted performance of the portfolio. Maximum drawdown highlights the largest peak-to-trough decline during the portfolio's history, illustrating a worst-case loss scenario. Lastly, Expected Shortfall (95%) evaluates the average loss in the worst 5% of scenarios, offering insight into the portfolio's tail risk. These metrics collectively provide a comprehensive assessment of the strategy's performance, balancing returns with associated risks.

3 Results & Analysis

In this part, we present the principal results of the backtest of our generalized pairs trading strategy. In the Basic strategies section, we will present the results for the 2 setups presented above using 3 different estimation window, 252 days, 504 days and 800 days.

3.1 Basic strategies - review of the performances

3.1.1 Strategy 1

Figure 2: Global Portfolio Cumulative Returns - Strategy 1 - 252D

Metric	Value
Annualized Return	-0.0118
Annualized Volatility	0.0772
Sharpe Ratio	-0.1525
Maximum Drawdown	0.1871
Expected Shortfall (95%)	-0.0448

Table 2: Performance Metrics - Strategy 1 - 252D

Estimation window: 252 days

The first strategy uses the first trading setup and a 252 days estimation windown. This strategy displays strong negative results with a high maximum drawdowns and very few upward moments.

Figure 3: Global Portfolio Cumulative Returns - Strategy 1 - 504D

Metric	Value
Annualized Return	-0.0085
Annualized Volatility	0.0776
Sharpe Ratio	-0.1092
Maximum Drawdown	0.1140
Expected Shortfall (95%)	-0.0494

Table 3: Performance Metrics - Strategy 1 - 252D

Estimation window : 504 days

The second strategy uses the first trading setup and a 504 days estimation windown. Again, this strategy displays significantly negative results, with a slight improvement however on the maximum drawdowns.

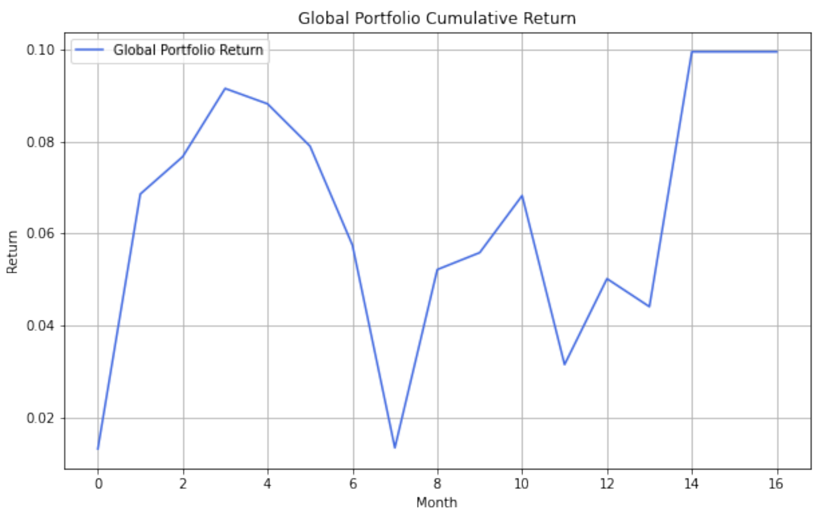

Figure 4: Global Portfolio Cumulative Returns - Strategy 1 - 800D

Metric	Value
Annualized Return	0.0725
Annualized Volatility	0.0915
Sharpe Ratio	0.7925
Maximum Drawdown	0.0833
Expected Shortfall (95%)	-0.0439

Table 4: Performance Metrics - Strategy 1 - 800D

Estimation window : 800 days

The third strategy uses the first trading setup and a 800 days estimation window. This strategy shows the highest results above all the basic strategies that are mentioned here. In fact with a 7.2% annualized return and a sharpe ratio of approximately 0.8, this strategy confirms that a longer window can lead to a better normalisation of the forecasted spreads.

3.1.2 Strategy 2

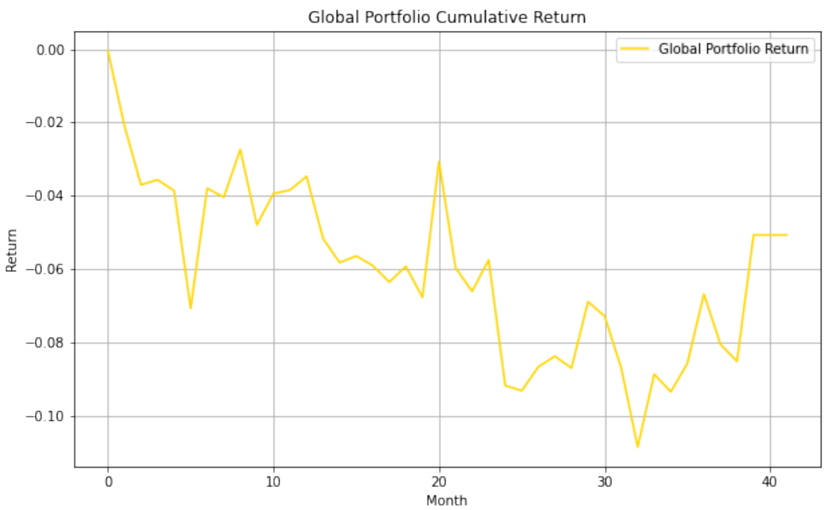

Figure 5: Global Portfolio Cumulative Returns - Strategy 2 - 252D

Metric	Value
Annualized Return	-0.0144
Annualized Volatility	0.0553
Sharpe Ratio	-0.2602
Maximum Drawdown	0.1064
Expected Shortfall (95%)	-0.0318

Table 5: Performance Metrics - Strategy 2 - 252D

Estimation window : 252 days

The fourth strategy uses the second trading setup and a 252 days estimation window. This strategy shows very similar results to the first strategy. However the results of this strategy are even worse, which introduce the fact that a tighter trading setup might not be a good idea.

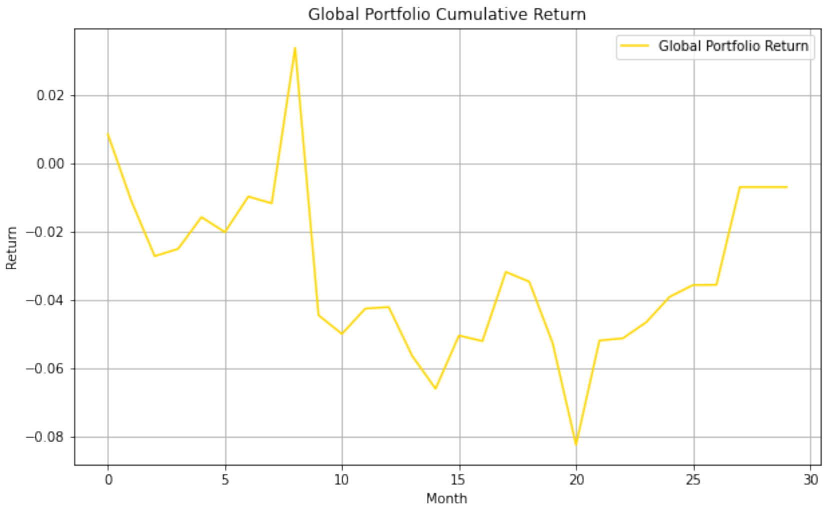

Figure 6: Global Portfolio Cumulative Returns - Strategy 2 - 504D

Metric	Value
Annualized Return	-0.0028
Annualized Volatility	0.0728
Sharpe Ratio	-0.0388
Maximum Drawdown	0.1174
Expected Shortfall (95%)	-0.0541

Table 6: Performance Metrics - Strategy 2 - 252D

Estimation window : 504 days

The fifth strategy uses the second trading setup and a 504 days estimation window. This strategy shows very similar results to the second strategy. However, this time the second setup performed better than the first one in terms of sharpe ratio, but has a worse maximum drawdown.

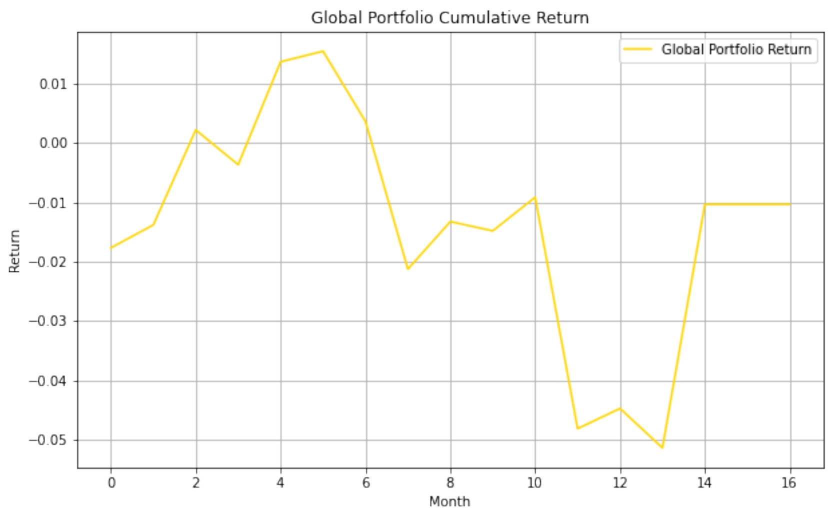

Figure 7: Global Portfolio Cumulative Returns - Strategy 2 - 800D

Metric	Value
Annualized Return	-0.0073
Annualized Volatility	0.0596
Sharpe Ratio	-0.1217
Maximum Drawdown	0.0668
Expected Shortfall (95%)	-0.0390

Table 7: Performance Metrics - Strategy 2 - 800D

Estimation window : 800 days

The last strategy uses the second trading setup and a 800 days estimation window. The results for this strategy are definitely worse than the results of the third strategy. This strategy confirms that using a tighter thresholds for trading signal is not a good idea. Our hypothesis for this is that the distribution of spread returns might not be normal and could display significant tails and skew.

To conclude about those basic strategies, most of them show significantly negative results. The only strategy that showed positive returns with a 0.79 sharpe ratio is the strategy using the 1st setup and a 800 days rolling window for the normalisation of the spreads forecasts. It seams that larger thresholds for the trading signal are helping the performance of the strategy. Therefore we will try to enlarge the thresholds furthermore in the section 3.2 More Advanced Strategies. Overall, the maximum drawdowns and expected shortfalls of all the strategies are not attractive.

For these reasons, we decided to enlarge our backtest to test different parameters. In the next section, we will study the impact of enlarging and creating an asymmetry in the trading signal. We will then assess if the clustering method is robust and if choosing other parameters could lead to a different clustering and therefore different results. Finally, we will assess if the estimation window for the spread estimation has an impact on the strategy's returns. In fact, until now, we estimated the OLS regression model for spread forecasting using the previous month data, however we found this choice arbitrary, this is why we will test if a longer estimation period could lead to a more robust model.

3.2 More advanced strategies - improving backtest performances

3.2.1 Calibration thresholds

Asymetrical and widen thresolds

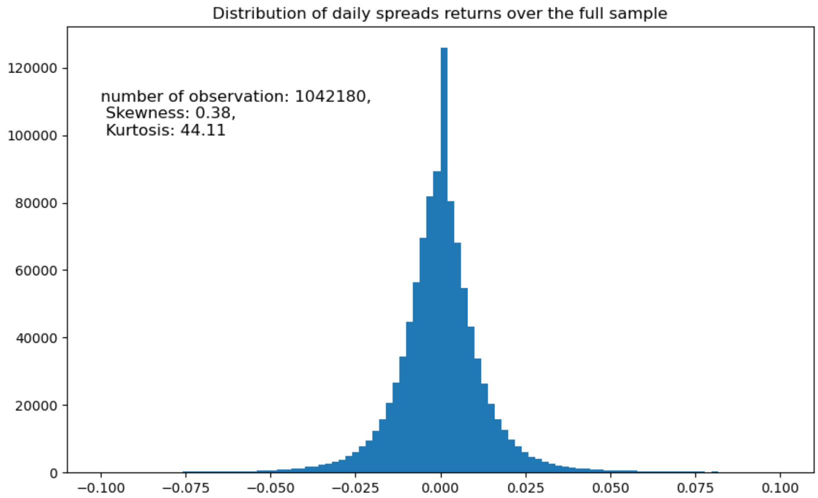

Figure 8: Distribution of the daily spread returns over the full sample

The figure above displays the daily spread returns for the whole period over the 428 spreads that were computed. As we can see, the distribution is positively skewed and presents heavy tails, this is why we decided to implement a slight skew and to enlarge the thresholds in the trading setup that you will find below, in order to take advantage of this distribution non-normality.

Table 8: Advanced Trading Signal Setup

	Short Stop	Short Open	Short Close	Long Close	Long Open	Long Stop
Strategy 1	3.2	1.8	0.8	-0.6	-1.6	-2.8

This advance Trading signal setup will help us in testing 2 significant takeaways from our previous backtest, the first one being that larger thresholds tend to outperform, and the second one being that as the distribution of spread returns is positively skewed, an asymmetric signal could also bring more positive returns

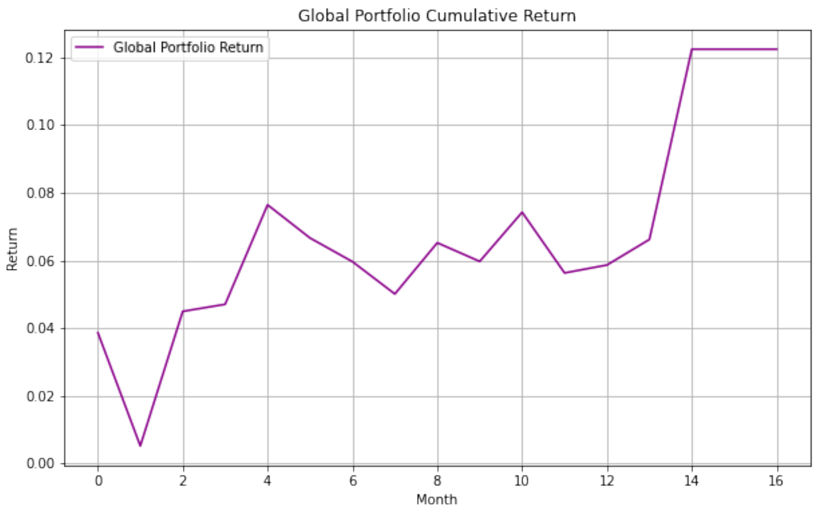

Figure 9: Global Portfolio Cumulative Returns - Advanced Threshold Strategy - 800D

Metric	Value
Annualized Return	0.0899
Annualized Volatility	0.0771
Sharpe Ratio	1.1658
Maximum Drawdown	0.0347
Expected Shortfall (95%)	-0.0334

Table 9: Performance Metrics - Advanced Thresholds Strategy- 800D

We can conclude that enlarging the thresholds and tilting them towards the left brought a significant sharpe ratio increase to the backtest of the strategy.

3.2.2 Revisited clustering approach

In this section we have modified the clustering approach for the index stocks. We have chosen to test the robustness of hierarchical clustering to the chosen distance.

Robustness of hierarchical clustering to the distance 5 distances were tested, resulting in the formation of clusters differing in size and composition. These distances are:

- Correlation

- Cityblock
- Chebyshev
- Minkowski
- Cosinus

The "correlation" distance results in the formation of 3 clusters, Cityblock and Chebyshev 6 clusters, and Minkowski and Cosinus 8 clusters. For all distances other than Chebyshev, the classic performance evaluation metrics are presented in the appendices (Table 15, Table 16, Table 17). For Chebyshev, which shows the best performance in the light of these metrics, you will find the results below (Table 10)

Metric	Strategy 1	Strategy 2	Adv Strat
Ann Return	0.0819	-0,0454	0.2050
Ann Volatility	0.0685	0.0591	0.1147
Sharpe Ratio	1.1947	-0.7683	1.7865
Maximum Drawdown	0.0435	0.0968	0.0609
Expected Shortfall (95%)	-0.018	-0.0416	-0.0180

Table 10: Performance Metrics - Advanced Clustering Approach - number of clusters : 6 - distance : Chebyshev - 800D

Firstly, it is interesting to note that compared with the Euclidean distance, performance remains consistent, with Sharpe Ratio of the same sign. It would seem that changing the Euclidean distance at Chebyshev amplifies the performance of the strategies. The Sharpe Ratio increases from 0.7925 to 1.1947 , -0.1217 to -0.7683, and 1.1658 to 1.7865. In terms of risk, it would appear that moving to Chebyshev's distance systematically increases the risk of loss (MD, ES) for all strategies.

Another interesting fact that emerges from the results obtained and presented (Table 10 and, annex Table 15, Table 16, Table 17) on clusters is that increasing the number of clusters detected from a certain distance is not systematically synonymous with better performance of the strategy. This leads us to the conclusion that the composition of the clusters is probably the most important factor beyond the number of clusters itself.

3.2.3 Revisited OLS estimation window for spread forecast

In this section, we try to evaluate the impact of the size of the estimation window of the OLS model for spread forecasting impact on the strategies results. To do this we decided to try 2 different estimation windows, 2 month and 6 month.

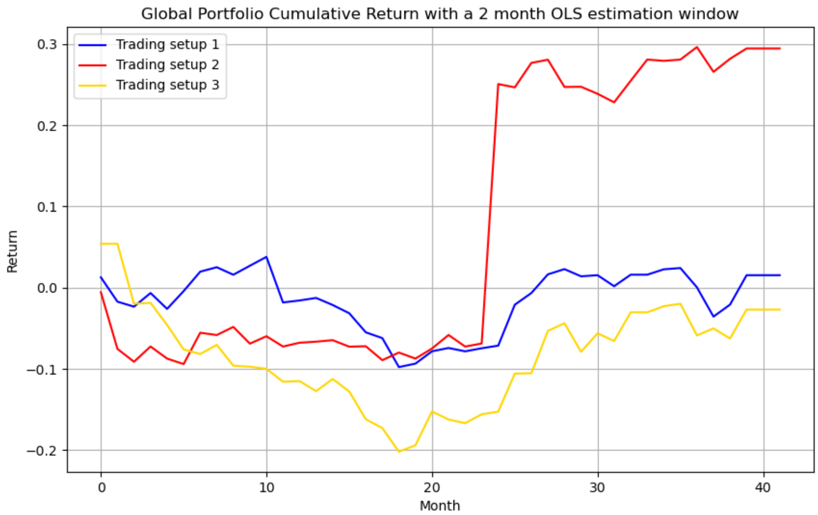

Figure 10: Global Portfolio Cumulative Returns with 2 month OLS estimation window - All setups - 252D

Metric	Strategy 1	Strategy 2	Advanced Threshold Strategy
Annualized Return	0.0044	0.0874	-0.0077
Annualized Volatility	0.0674	0.1808	0.0879
Sharpe Ratio	0.0650	0.4836	-0.0874
Maximum Drawdown	0.1339	0.0868	0.2425
Expected Shortfall (95%)	-0.0426	-0.0446	-0.0493

Table 11: Performance Metrics - Enlarged OLS estimation window - 2 month - 252D

Above we can see the results for all the different trading setups using a 252 days estimation window for spread normalisation and a 2 month estimation window for the estimation of the OLS model for spread forecasting. As we can see the results of these strategies are a lot better than the ones using a 1 month estimation window. We can therefore make the hypothesis that using a longer estimation window can lead to an improvement of returns and sharpe ratios of the strategy. The only surprising result here is for the second trading setup which clearly outperforms the other. We will now test using the 800D estimation window to see if we can confirm this hypothesis.

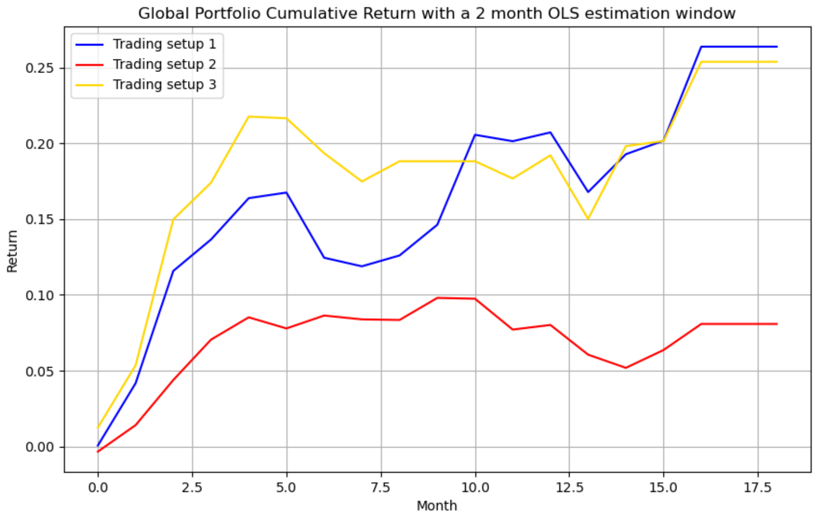

Figure 11: Global Portfolio Cumulative Returns with 2 month OLS estimation window - All setups - 800D

Metric	Strategy 1	Strategy 2	Advanced Threshold Strategy
Annualized Return	0.1800	0.0523	0.1726
Annualized Volatility	0.1028	0.0467	0.1080
Sharpe Ratio	1.7507	1.1196	1.5992
Maximum Drawdown	0.0570	0.0502	0.0825
Expected Shortfall (95%)	-0.0430	-0.0204	-0.0420

Table 12: Performance Metrics - Enlarged OLS estimation window - 2 month - 800D

Above are the results for the 3 trading setups using a 800D estimation window for spread normalisation and a 2 month window for OLS estimation. Those results confirm that using a larger estimation window to build the OLS model for spread forecasting is directly link to higher returns and sharpe ratios. Here we even achieve to build 3 strategies with more than 1 of sharpe ratio which is significantly better than what we found before. We will now conduct the same analysis using a 6 month estimation window for the OLS model.

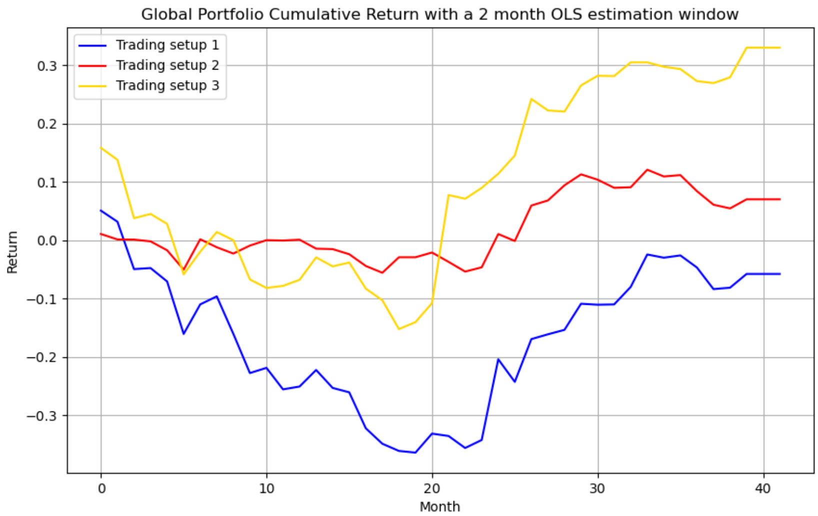

Figure 12: Global Portfolio Cumulative Returns with 6 month OLS estimation window - All setups - 252D

Metric	Strategy 1	Strategy 2	Advanced Threshold Strategy
Annualized Return	-0.0163	0.0203	0.0985
Annualized Volatility	0.1450	0.0717	0.1759
Sharpe Ratio	-0.1125	0.2827	0.56
Maximum Drawdown	0.3692	0.0726	0.3236
Expected Shortfall (95%)	-0.0794	-0.0280	-0.0846

Table 13: Performance Metrics - Enlarged OLS estimation window - 6 month - 252D

The figures above show the result for the 3 different trading setup using a 252 days estimation window for the spread normalisation and a 6 month estimation window for the OLS model. Those results are mixed. In fact the drawdowns and overall volatility of the strategies are higher. However we see that the setup 3 for example performs a lot better than the other setups in this configuration. We will now try to use a 800 days window for spread normalisation to conclude about the use of longer estimation windows for OLS estimation.

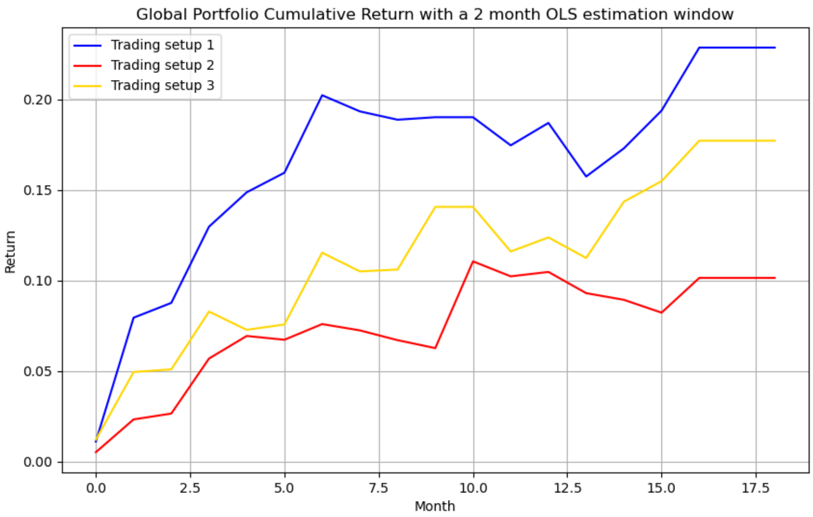

Figure 13: Global Portfolio Cumulative Returns with 6 month OLS estimation window - All setups - 800D

Metric	Strategy 1	Strategy 2	Advanced Threshold Strategy
Annualized Return	0.1544	0.0660	0.1179
Annualized Volatility	0.0774	0.0502	0.0631
Sharpe Ratio	1.9966	1.3145	1.8695
Maximum Drawdown	0.0544	0.0313	0.0324
Expected Shortfall (95%)	-0.0296	-0.0117	-0.0246

Table 14: Performance Metrics - Enlarged OLS estimation window - 6 month - 800D

Above are the results for the 3 trading setups using a 800 days estimation window for the spread normalisation and a 6 month estimation window for the OLS model. Those results confirm that using a larger estimation window to construct the OLS model benefits to this generalized pairs trading strategy as all trading setups show a positive relation between the size of the estimation window and sharpe ratios. This is by far the best setup we have found for the strategy in terms of sharpe ratio, drawdowns and expected shortfall.

4 Conclusion

In this paper, we discussed the performance of a quantitative based generalized pairs trading strategy on the stocks of 10 sectors of the S&P 500 from 2020 to 2024.
Firstly our clustering analysis showed the existence of 7 different clusters of stocks. Using those 7 clusters, we constructed a spread model using linear regression to forecast monthly spread between single stocks and their respective cluster. Using those forecasted spread, we built monthly re-balanced portfolios composed of long-short positions of single stocks versus their corresponding clusters.

The first results we found using basic parameters for the trading setup, hierarchical clustering method and OLS estimation window were deceiving. In fact most strategies were not able to deliver a positive performance and the only performing strategy displayed large drawdowns. This is why we decided to change the parametrization of the strategy to see if we could improve those results. We first decided to look at the spread returns distribution tails, looking for non-normality signs. We found a positive skewness and a very high kurtosis. This is why we decided to test a different trading setup with larger thresholds in absolute values but also to skew the signals towards more positive values. The results we got confirmed that using larger and skewed thresholds could lead to an improvement of results but not in all case, however for the best strategy, the sharpe ratio was above 1. Then we decided to try a different distance metric for the hierarchical clustering of stocks. We used the Chebyshev distance approach which lead to the discovery of an six cluster. Using this method lead to significantly improve the sharpe ratio for 2 strategies out of 3, but increase the tail risks. Finally, we tried 2 new estimation windows for the estimation of the OLS model for spread forecasting, the 2 previous month and the 6 previous months. This change of parameter showed some significant improvement on the strategies' results. In fact, it seems that the sharpe ratios of the strategies are positively correlated with the size of the estimation window for the OLS model. With this parameter change, we were able to find sveral strategies with a sharpe higher than 1. To conclude on the advanced method, the best strategy we found is using an 800 days window for spread normalisation, the first trading setup and a 6 month window for the OLS estimation.

Our final word on the strategy is that the choice of a great parameter set is necessary for the generalized pairs trading strategy to be attractive in terms of risk/reward profile. This paper showed that some parametrization techniques can be used to significantly improve backtest results and could lead to significantly high sharpe ratios .

5 Annex

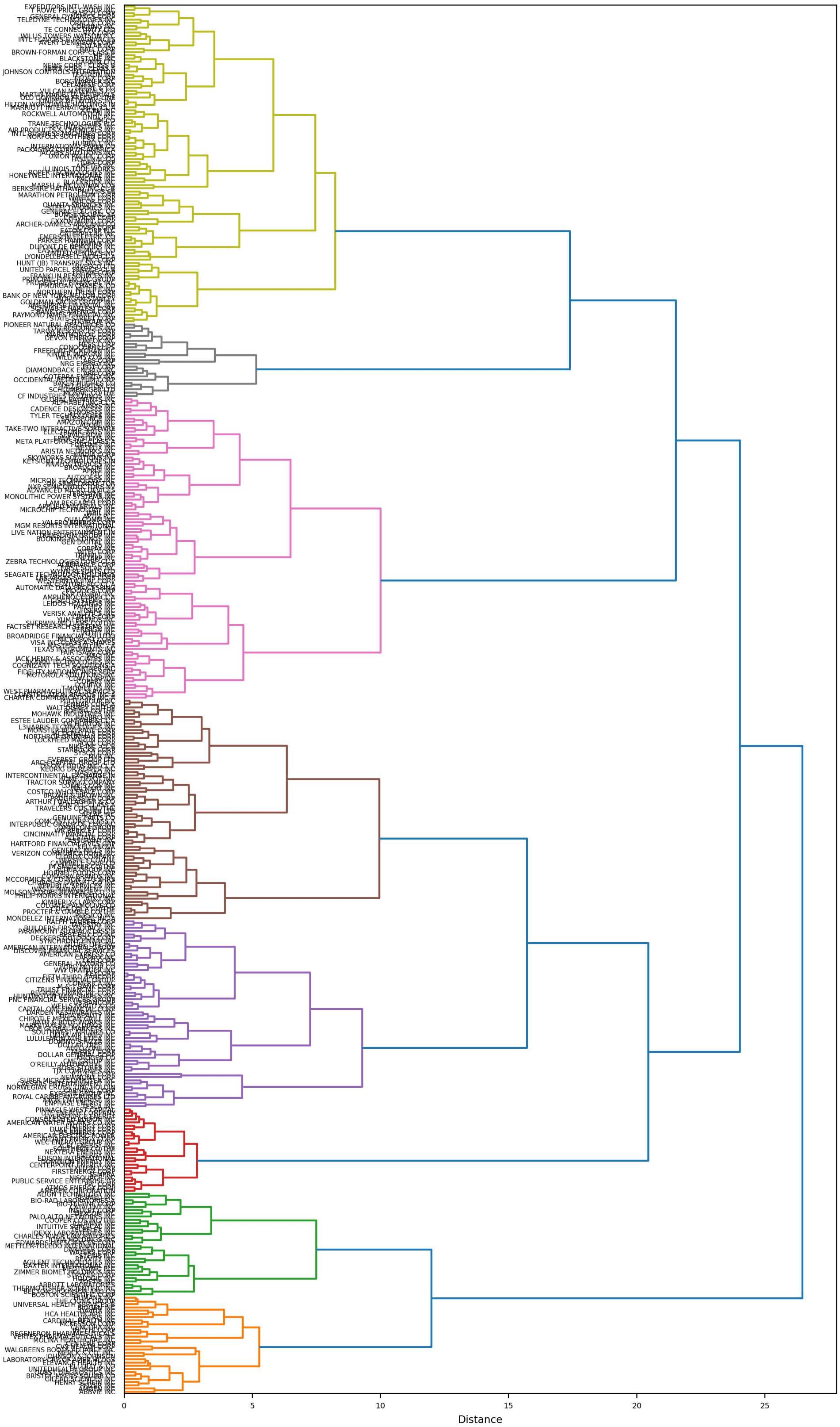

Figure 14: Dendrogram of the clusters of stocks

Metric	Strategy 1	Strategy 2	Adv Strat
Annualized Return	0.0308	0.0187	-0.0100
Annualized Volatility	0.1049	0.0506	0.1016
Sharpe Ratio	0.2939	0.3688	-0.0980
Maximum Drawdown	0.1290	0.0477	0.1427
Expected Shortfall (95%)	-0.0564	-0.0306	-0.0614

Table 15: Performance Metrics - Advanced Clustering Approach - number of clusters : 3 - distance : Correlation - 800D

Metric	Strategy 1	Strategy 2	Adv Strat
Ann Return	0.0572	0.0043	0.0967
Ann Volatility	0.0980	0.0555	0.0810
Sharpe Ratio	0.5836	0.0767	1.1947
Maximum Drawdown	0.0632	0.0573	0.0786
Expected Shortfall (95%)	-0.0359	-0.026	-0.0238

Table 16: Performance Metrics - Advanced Clustering Approach - number of clusters : 6 - distance : Cityblock - 800D

Metric	Strategy 1		Strategy 2		Adv Strat	
	Minkowski	Cosinus	Minkowski	Cosinus	Minkowski	Cosinus
Ann Return	0.0269	0.0463	-0.0027	0.0125	0.0554	0.0469
Ann Volatility	0.0892	0.0764	0.0626	0.0397	0.0709	0.0962
SR	0.3010	0.6056	-0.0429	0.3157	0.7822	0.4872
MD	0.1162	0.0692	0.0820	0.0302	0.0695	0.1001
ES (95%)	-0.0536	-0.0267	-0.0337	-0.0155	-0.0481	-0.0351

Table 17: Performance Metrics - Advanced Clustering Approach - number of clusters: 8 - 800D