Performance of pairs trading on the S&P 500 index

By: Emiel Verhaert, studentnr: 348122

Supervisor: Dick van Dijk

Abstract

Until now, nearly all papers focused on pairs trading have just implemented the strategy given by Gatev et al. (2006). However, due to many parameters that are a given in that paper, it may be better to perform pairs trading with different parameters. These parameters are:

- **the length of the formation period**
- **the length of the trading period**
- **the standard deviation on which you open a trade**
- **working with a stop loss**

Adjusting parameters in this thesis has showed that opening a pair when it is 1.5 standard deviation from its mean is better than the standard 2 standard deviations which has been used in many papers. The formation and trading period of respectively 250 and 125 days are still good periods to use. Implementing a tight stop-loss however, decreases results. In addition, no pairs trading research has been performed yet on the S&P 500 index, and since there are many combinations possible, this could have produced interesting results. However, when including commissions, pairs trading does not turn out to be a good strategy on the S&P 500 index. In addition, performance of pairs trading has significantly declined in the last 10 years compared to the 10 years before.

Contents

Table of figures

Introduction

Eager to find the holy grail, investors and traders alike have been searching for profitable strategies for investing in the stock market. One of the quantitative strategies being applied is pairs trading. This strategy has almost a 30-year history on Wall Street, used by hedge funds and investment banks alike. The concept is pretty straightforward. The explanation from Gatev et al. (2006) is as follows: find two stocks whose prices have moved together historically. When the spread between them widens, short the winner and buy the loser. If history repeats itself, prices will converge and the trader will profit.

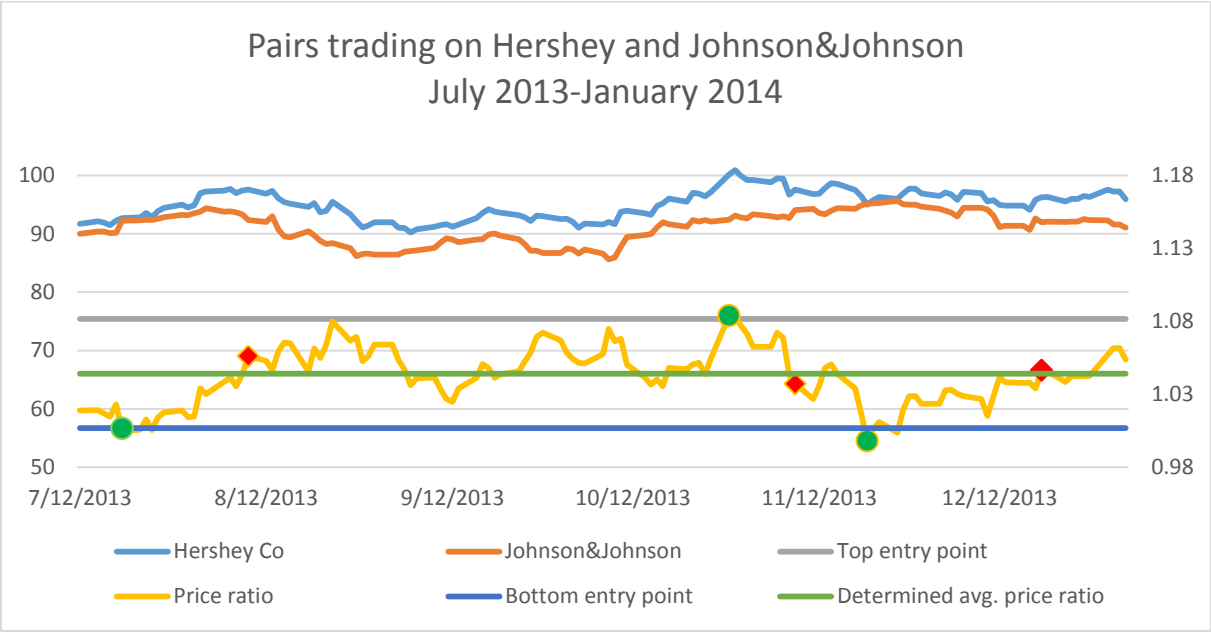

Figure 1: Example of pairs trading strategy on the stock pair Hershey Co. and Johnson&Johnson. Left y-axis: stock price, Right y-axis: Price ratio's

An example of pairs trading can be found in the figure above. The algorithm found that the stock price of Hershey Co and Johnson&Johnson move alike. We then compute a historical mean (in this case the past 250 trading days) of the price ratio (the "Determined average price ratio"). We enter a trade when the current price ratio is either too low (the "Bottom entry point") or too high (the "Top entry point) compared to the historical mean. A dot indicates the opening of trade and a rhombus the closing of a trade. Since the price ratio in this case is the stock price of Hershey divided by the stock price of Johnson&Johnson, we take a long position in Hershey and a short position in Johnson&Johnson if the price ratio is considered to be too low. If the price ratio is considered to be too high we take a short position in Hershey and a long position in Johnson&Johnson.

In this thesis pairs trading strategies will be tested using the S&P 500 stocks, this is chosen because it is the leading market for U.S. equities and it reflects the risk/return characteristics of the large cap universe, which ensures a liquid market and a wide array of companies to choose from. E.g. if the NASDAQ would be selected, the diversity of company sectors would be decreased to only technology companies whereas the S&P 500 lists companies from all kinds of sectors which could create interesting pairs.

Where previous research has focused on restricting pairs within an industry, or the effectiveness of the basic algorithm from Gatev et al. (2006) on other financial markets, no research has yet been done on the many perhaps arbitrarily chosen parameters in the model from Gatev et al. (2006). These parameters are:

- the length of the formation period
- the length of the trading period
- the standard deviation on which you open a trade
- working with a stop loss

It is interesting that nearly all pairs trading research papers assume that these parameters are fine to work with and are integrated in their research without further questioning. While these parameters have not been investigated, they could have a significant influence on pairs trading results. Furthermore, since pairs trading is now widely known in the financial markets, this thesis will investigate whether this has had an influence on the performance of pairs trading. It will compare the period 1994-2004 with 2004-2014 and compare results between the two past decades. Research done by Do and Faff in 2010 reports a continuing deterioration in the performance of pairs trading over time and this thesis will also investigate this. Hence, the question to be answered by this research is: is pairs trading still profitable? And if so, what parameters need to be used in order to gain the best economic performance?

In this thesis, the restriction of pairs within sectors will not be applied. This choice has been made since there could be stocks which are closely related without them being in the same sector. In addition, if stock pairs would match better when being in the same sector, then there is a high chance the same pair will be constructed without this restriction. However, this choice could also decrease returns of the pair trading strategy. This could be investigated in future research.

Pairs will be computed in two different ways. The first will be the same as in Gatev et al. (2006), by computing the sum of squared deviations (SSD) between cumulative return series of two stocks. After creating all pairs, they will be ranked based on the SSD. The pair that has the lowest SSD is the top pair and the pair that has the highest SSD will be the lowest ranked pair.

Also, a different pairs forming method will be used based on a cointegration test. With the second method, pairs will be created if they are considered to be cointegrated. This is computed with the augmented Dickey Fuller test. With this method pairs will be ranked based on the p-value that results from the Dickey Fuller test. The lower the p-value, the more cointegrated a pair is and therefore it will be higher ranked. These methods will be explained further in this thesis.

After forming pairs, different portfolios will be constructed. These will be based on the top 10, top 20 ,top 50 and top 101-120 pairs as a comparing portfolio. With both methods, equal investments will be made in each pair. Additionally, for the portfolios that are constructed with the SSD method, there will also be weighted portfolios based on the SSD between pairs, these will invest more money in pairs that are ranked higher. This weighting scheme is not being applied to the portfolios that are constructed with the cointegration test. This is due to the fact that the p-values of the pairs in the portfolios are so close to each other that the results are almost equal to the portfolios that have an equal investment in each pair. With all these portfolios, the simulation of the strategy of the past 20 years is run for different parameters and then the differences between portfolios as well as general differences between the different parameter settings are evaluated.

Performance of the strategies has been evaluated with a 1-month 95% value at risk measure as well as economically with the monthly return and its standard deviation. To closely resemble returns in the actual market a cost of 0.5% of the trade value is charged when entering as well as exiting a trade.

After evaluation with the measures noted above and comparing with the portfolio that contains pairs 101-120 it turns out that pairs trading does not turn out to be profitable. The best portfolio only has an average monthly return of 0.182% (Note: this is not the excess return, just the return). Even worse, the top-10 and top-20 SSD pairs perform less than pairs 101-120. Only the top-50 pairs portfolios (SSD as well as cointegration test) perform better than the portfolio with pairs 101-120. Another interesting finding is that implementing a stop-loss worsens results significantly. When comparing the periods 1994-2004 with 2004-2014, similar results are found as Do & Faff. The performance of pairs trading continues to deteriorate over time and results in the second decade are significantly less than in the first decade. However, the most useful finding that could be tested in other working pairs trading strategies is that returns are higher when trading takes place on a 1.5 standard deviation from the mean price ratio compared to the 2 deviations that are used in the strategy of Gatev et al. (2006).

Methodology

Forming pairs

In this thesis, pairs will be constructed in two ways. First we use the approach from Gatev et al. to create a cumulative returns index. Our second approach involves performing cointegration tests and choosing pairs that are best cointegrated with the results of the Dickey-Fuller test.

Method 1

The method from Gatev et al. (2006) is as follows: first, construct a cumulative returns index for each stock. Then, we calculate the sum of squared deviations (SSD) between two normalized price series (two stocks x and y) as follows:

$$SSD = \sum_{i=1}^{n} \left(\frac{x_i}{x_1} - \frac{y_i}{y_1}\right)^2 \qquad (1)$$

This value is calculated for all possible pairs, and the pair with the smallest SSD will be selected. Then the same happens with the second stock, et cetera. Finally, the pair that has the smallest SSD will be the first ranked pair. The pair that has the largest SSD will be the last ranked pair. Using these ranks the portfolios will be created. The construction of portfolios will be explained further in this thesis.

Method 2

Secondly, pairs will be selected which are cointegrated. These will be selected with the Engle-Granger two-step method. The price (not returns) of stock X will be regressed on the price of stock Y using ordinary least squares:

$$X_t = c + \beta Y_t + u_t \qquad (2)$$

Then, a stationarity test, the Dickey-Fuller test, will be performed on the residuals u_t of the regression. A significance level of 5% will be applied and pairs that have P-values lower than 5% will be selected for our portfolio. In order to create top-10, top-20 and top-50 portfolios we will rank het pairs based on the p-value that results from the Dickey-Fuller test. The pair that has the lowest p-value will have the highest ranking.

Trading pairs

Pairs trading involves a formation period of pairs, and a trading period of pairs. The formation period of pairs is the period over which the SSD or p-values are computed and the pairs are formed. The trading period immediately follows the formation period and is the period in which the chosen pairs from the algorithms are traded. In the paper from Gatev et al. (2006), these periods are chosen arbitrarily. They work with a 12-month formation period and a 6-month trading period. In this thesis it will be investigated whether results differ when we adjust the formation as well as the trading period.

We open a position in a pair when prices diverge by more than 2 historical standard deviations from the price ratio as estimated during the pairs formation period. The price ratio is calculated as follows:

$$PR_t = \frac{Price_X}{Price_Y} \qquad (3)$$

Where PR_t is the pair ratio and $Price_X$ and $Price_Y$ are respectively the normal price (again, not returns) of stock X and stock Y at time t. When this ratio is 2 standard deviations above or below the mean, we will enter a trade. When the current price ratio is 2 standard deviations above the historical mean, the price ratio is considered to be too high. This either means that the price of stock X is too high, or that the price of stock Y is too low, or both. Therefore in this case, a short position is opened in stock X and a long position is opened in stock Y.
When the current price ratio is 2 standard deviations below the historical mean, the price ratio is considered to be too low. This either means that the price of stock X is too low, or that the price of stock Y is too high, or both. Therefore in this case, a long position is opened in stock X and a short position is opened in stock Y. An example of this can be found in the introduction.

Perhaps when trading with a 2 standard deviation difference we will trade too often for little profits, and therefore it will also be examinined to open trades with a 2.5 or 3 standard deviation difference from the mean. A smaller standard deviation difference will also be considered. Since 2 standard deviations is considered to be an industry standard, it might turn out to be profitable to enter positions before a difference of 2 standard deviations. However, that will also naturally result in performing more trades which will raise the transaction costs.

Entering and exiting trades

We open a position in a pair when the price ratio diverges by more than two historical standard deviations from the historical mean price ratio as estimated during the pairs formation period.

We unwind the position when the price ratio crosses the historical mean again. This is because this historical mean is the point at which you expect the price ratio would be. If the price ratio does not return to the historical mean before the end of the trading interval, gains or losses are calculated at the end of the last trading day of the trading interval. The position will also be closed when one of the stocks in the pair is delisted from the S&P 500 index.

In addition, compared to Gatev et al., this research will forbid opening a position within one trading week from closing of the trading window, this will prevent additional commissions since a pair is unlikely to close again if it is opened late in the trading period. In the last five days of the trading period, only the closing of positions will be allowed. The real effectiveness of this restriction will not be measured as it is not the main goal of this research, however it is a sensible restriction to put in place. This is how Gatev et al. (2006) mentioned this issue: “There is a reason why our trading strategies require ‘‘too much’’ trading. We open pairs at any point during the trading period when the normalized prices diverge by two standard deviations. This is not a sensible rule toward the end of a trading interval. For example, suppose that a divergence occurs at the next to last day of the trading interval. The convergence has to be substantial to overcome the transaction cost that will be incurred when we close out the position on the next day (the last day of the trading interval). Unreported results suggest that this is also an important source of excess trading.”

Also, the possibility of a stop-loss will be investigated. Perhaps it is better to close a position after having an x% loss since the pairs drift away from each other.

Pair portfolios

In this thesis there are nine portfolios which will be examined. 6 portfolios will have an equal investment in each pair. The other three portfolios will have a weighted investment in each pair.

These weights are calculated as follows for each pair:

$$b_i = \frac{1}{SSD_i} \qquad (4)$$

Then,

$$w_i = \frac{b_i}{\sum b_i} \qquad (5)$$

Where SSD_i is the sum of squared deviations for pair i, and w_i turns out to be our determined weight for pair i.

These are the following:

- Portfolio 1: The top-10 pair portfolio as determined by the SSD-method, with the same investment in each pair
- Portfolio 2: The top-20 pair portfolio as determined by the SSD-method, with the same investment in each pair
- Portfolio 3: The top-50 pair portfolio as determined by the SSD-method, with the same investment in each pair
- Portfolio 4: The top-10 pair portfolio as determined by the SSD-method, with a weighted investment in different pairs
- Portfolio 5: The top-20 pair portfolio as determined by the SSD-method, with a weighted investment in different pairs
- Portfolio 6: The top-50 pair portfolio as determined by the SSD-method, with a weighted investment in different pairs
- Portfolio 7: The top-10 pair portfolio as determined by the Engle-Granger method, with the same investment in each pair
- Portfolio 8: The top-20 pair portfolio as determined by the Engle-Granger method, with the same investment in each pair
- Portfolio 9: The top-50 pair portfolio as determined by the Engle-Granger method, with the same investment in each pair
- Portfolio 10: The pairs 101-120 as determined by the SSD-method, with the same investment in each pair. With this portfolio it can be examined whether the ranking of pairs works.

Evaluating portfolios

Portfolios will be evaluated on three criteria. These are the following:

1. Monthly returns
2. Monthly standard devation
3. 1-month 95% VaR of $1 million

Within each portfolio, these statistics will be calculated for the entire period, the first 10 years and the second 10 years. Then we can examine whether the performance of pairs trading has declined in the last decade.

When calculating the monthly returns, we assume that one month has 22 trading days.

For Value at Risk the returns first all get standardized to monthly returns, then the mean and standard deviation of the montly returns are used to calculate the 1-mont 95% VaR of $1 million. The Value at Risk measure is useful because it shows a loss that can happen in the worst 5% of cases, and it can provide an estimate of the potential leverage that could be applied to the strategies, if they were working. Even though history has taught us that Value-at-Risk measures should not be too heavily relied upon, the pairs trading strategy seems to be exposed to relatively little risk since a position is long as well as short in the market, or in other words, delta-neutral.

The Sharpe ratio has also been considered as an evaluation measure, however, due to the negative nature of our results these ratios have been left out of this research since differences between negative Sharpe ratios do not make sense. For example, if you have two strategies that have the same negative excess return, but two different standard devations, the strategy that has a higher standard deviation will have a higher Sharpe ratio.

Data

In this thesis, a 20-year historical end-of-day data-set of the S&P 500 will be used. This set includes 787 stocks that are or have been present on the S&P 500 index. Due to the restriction of the stock actual being on the S&P 500 on a current moment, data has been cleared of the stocks before they list on the index as well as after they leave the index. This means that an open position must be closed when a stocks gets delisted. Stocks get delisted due to mergers or bankruptcies and including all past stocks on the index prevents a survivorship bias in this thesis.

The cumulative returns for determining the SSD between pairs have been constructed by dividing the current price by the price on the start of the formation period.

Results

First a simulation with the parameters from Gatev et al. (2006) is being performed. This leads us to the following results for our 10 portfolios portfolios (the highlighted columns are the most profitable options in each table):

Portfolio properties	Portfolio 1	2	3	4	5	6
Formation period	250	250	250	250	250	250
Trading period	125	125	125	125	125	125
Trading St.dev.	2.0	2.0	2.0	2.0	2.0	2.0
Avg. montly return	-0.139%	-0.183%	0.098%	-0.024%	-0.034%	0.078%
St.dev of return	0.993%	0.811%	0.550%	0.815%	0.609%	0.579%
# trades performed	536	1110	2958	536	1110	2958
Value at risk	97769	83871	46452	76623	58790	50737
Results 1994-2004						
Avg. monthly return	-0.07%	-0.28%	0.17%	0.20%	0.08%	0.17%
St.dev of return	1.09%	0.93%	0.55%	0.65%	0.53%	0.57%
Value at Risk	101737	99025	41740	51147	47049	45279
Results 2004-2014						
Avg. monthly return	-0.21%	-0.10%	0.03%	-0.24%	-0.14%	-0.01%
St.dev of return	0.92%	0.70%	0.56%	0.91%	0.67%	0.58%
Value at Risk	95295	69266	51452	96005	69246	55982

Portfolio properties	Portfolio 7	8	9	10
Formation period	250	250	250	250
Trading period	125	125	125	125
Trading St.dev.	2	2	2	2
Avg. montly return	0.132%	0.074%	0.007%	-0.010%
St.dev of return	1.18%	0.88%	0.84%	0.85%
# trades performed	348	689	1732	1183
Value at risk	100520	77213	77126	79010
Results 1994-2004				
Avg. monthly return	0.19%	0.19%	0.10%	0.01%
St.dev of return	1.34%	0.97%	0.78%	0.96%
Value at Risk	100520	77213	77126	79010
Results 2004-2014				
Avg. monthly return	0.08%	-0.03%	-0.08%	-0.03%
St.dev of return	1.04%	0.79%	0.91%	0.74%
Value at Risk	90472	75244	87233	72369

Figure 2: simulation of pairs trading over 20 years with the parameters given by Gatev et al. (2006)

The results found performing pairs trading on the S&P 500 index are rather disappointing, out of the basic strategy there are only five portfolios (the highlighted ones) that manage to show positive average monthly returns, yet even these portfolios have such a small average monthly return, that simply allocating your capital in t-bills is a much better option from a risk/reward perspective.

Another observation is that when we compare the returns from 1994-2004 with the returns from 2004-2014, all monthly returns are decreased, except for portfolio 2. This means that even if one would like to apply this strategy, the returns are declining. In terms of volatility, the results suggest that they are pretty safe to work with and only have a standard deviation on the return of around one percent. The 95% 1 million USD Value-at-Risk measure is also around or below 100.000 USD in each of the portfolios and if the worst case is a 10% loss, it seems to be a relatively safe strategy.

In addition, it appears that of all the SSD-portfolios, the one that has the top-50 pairs seems to perform the best and since it includes the top-10 and top-20 portfolio, pairs 21-50 make more profitable trades. Therefore portfolio 3 also performs better than portfolio 6 since portfolio 6 invests more money in higher ranked pairs. Pairs ranked 101-120 in the SSD-method also appear to perform better than the top-10 and top-20 pair portfolio.
When comparing the portfolios with pairs created by the SSD method compared to the cointegration test method, the cointergrated pairs have a larger standard deviation. This is especially true when the best SSD portfolio (the top-50 SSD pairs) is compared with the best cointegration test portfolio (the top-10 cointegrated pairs). For a not-so significant improvement in monthly returns (around 0.04%) one has to deal with twice as much volatility. In terms of risk/reward therefore portfolio 3 is the winner. The top-50 SSD pairs portfolio (portfolio 3) performs best on the S&P 500.

The impact of commissions is also pretty large on these portfolios. In table 2 below you can clearly see the impact of commissions on the monthly returns. If there were no commission, all our returns have been positive. In the table below one can find the differences in monthly returns due to commission in the first six porfolios.

	Portfolio 1	2	3	4	5	6
Incl commissions	-0,14%	-0,18%	0,10%	-0,02%	-0,03%	0,08%
Excl commissions	0,11%	0,08%	0,38%	0,20%	0,19%	0,31%

Figure 3: performance difference due to commissions

In the chart below one can find the returns per period of portfolio 7 from 1994-2014.

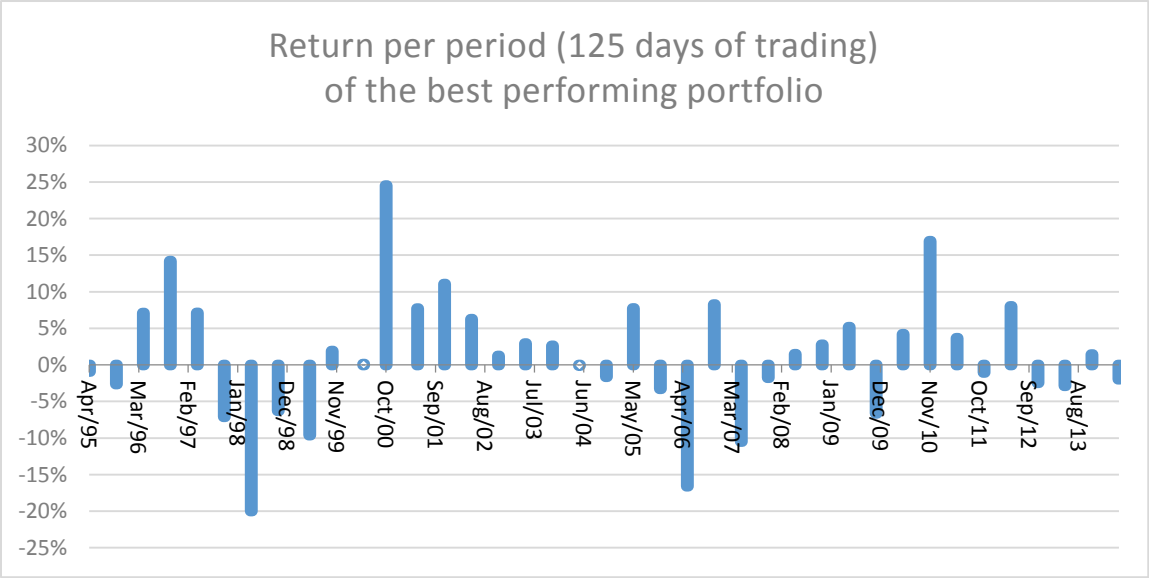

Figure 4: Performance of portfolio 7 with the basic parameters given by Gatev et al.

It appears that the economic crisis in 2008 barely had an impact on the strategy and that the profits as well as the losses tend to get smaller over time. The winning and losing months also seem to change arbitrarily and that there is not really a pattern of winning/losing months.

Comparing standard deviations

Now we have established that portfolio 3 is our best option, the influence of changing the trading standard deviation will now be looked at. We will use only portfolio 3 and will look at the influence of adjusting the trading standard deviation:

Portfolio properties	Portfolio 3	3	3	3	3
Formation period	250	250	250	250	250
Trading period	125	125	125	125	125
Trading St.dev.	1.5	1.7	2.0	2.5	3.0
Avg. montly return	0.135%	0.132%	0.098%	0.067%	0%
St.dev of return	0.569%	0.570%	0.550%	0.514%	0.422%
# trades performed	3840	3450	2958	2380	1991
Value at risk	46227	46558	46452	44715	36296
Results 1994-2004					
Avg. monthly return	0.22%	0.21%	0.17%	0.15%	0.15%
St.dev of return	0.59%	0.58%	0.55%	0.52%	0.45%
Value at Risk	43030	43006	41740	40680	33720
Results 2004-2014					
Avg. monthly return	0.05%	0.06%	0.03%	-0.01%	-0.03%
St.dev of return	0.55%	0.56%	0.56%	0.51%	0.39%
Value at Risk	49426	50417	51452	48762	37851

Figure 5: Comparison of performance when changing the trading standard deviation

After further looking into portfolio 3 and simulating the 20-year period again with different standard deviations, it turns out that there are better settings for the standard deviation than the 2 standard deviations used by Gatev et al. and in many other papers. This could be due to the fact that since many arbitrageurs already act at the 2 standard deviations point or before, which decreases the profit opportunities. At a standard deviation of 1.5, a position will be already opened before other traders act which could cause the extra profitability. Even though more trades are performed, the profits from opening a position earlier cover the transaction costs. Yet, still all the portfolios are coping with lower returns than a risk-free interest rate which implies that allocating capital can be done in much more profitable strategies or other exchanges than the S&P 500. The results of the other portfolios are similar except for the performance of the top 101-120 pairs. In this portfolio the returns also increase when choosing higher standard deviations. This result could be due to the 'random' nature of the pairs as it is our only portfolio where this happens. The results of these pairs, as well as the top-10 cointegration pairs portfolio can be found in the appendix.

Adjusting the trading period and formation period

After the interesting findings with tweaking the stop loss, there still are other parameters which we can tweak. Our next step is adjusting the formation and trading period. Again, research will continue on the portfolio that displays the best performance which is the top-50 SSD pair portfolio that trades on a change of 1.5 standard deviations. Because a smaller formation period could also result in less accuracy on a long-term trading period, the formation period and the trading period will be scaled down by the same factor. Because working with a trading period of six months is considered as long term by many investors, and since pairs trading also happens with shorter terms, this research will only look at shorter trading periods.

Portfolio properties	3	3	3	7	7	7
Formation period	120	200	250	120	200	250
Trading period	60	100	125	60	100	125
Trading St.dev.	1.5	1.5	1.5	1.5	1.5	1.5
Avg. montly return	-0.352%	-0.055%	0.135%	0.001%	-0.122%	0.182%
St.dev of return	0.861%	0.612%	0.569%	2.17%	1.62%	1.47%
# trades performed	7409	4656	3840	905	561	468
Value at risk	47879	48216	46227	93994	121352	124844
Results 1994-2004						
Avg. monthly return	-0.39%	0.03%	0.22%	0.32%	0.27%	0.31%
St.dev of return	0.92%	0.60%	0.59%	2.65%	1.78%	1.67%
Value at Risk	51052	43391	43030	93994	121352	124844

Results 2004-2014						
Avg. monthly return	-0.31%	-0.14%	0.05%	-0.31%	-0.52%	0.06%
St.dev of return	0.81%	0.62%	0.55%	1.53%	1.36%	1.29%
Value at Risk	45065	53046	49426	75083	116825	114052

Figure 6: Performance difference of the top-50 SSD pairs portfolio (3) and the top-10 cointergration test portfolio(7)

In the table above, we see that decreasing the formation and trading period also decreases our already small returns. Therefore, the periods suggested by Gatev et al. (2006) remain the best option if one would like to do pairs trading on the S&P 500. Results of the other portfolios are similar and can be found in the appendix.

Implementation of a stop-loss

Until now, we have tried to tweak the standard deviation, the formation period and the trading period. One more option worth evaluating is the implementation of a stop-loss. Again, with portfolio 3 as our basis, the results in table 6 are obtained.

These results are very interesting. The tighter we set the stop-loss, the more badly the pairs trading strategy performs. However, this is similar to the results of Ruiter (2011) and show that it is not beneficial to set a tight stop-loss on a pairs trading strategy. Given these results, especially the result when applying the 20% stop-loss, these results imply that probably before a pair is closed, sometimes that position reaches a loss of more than 20%, and when it does, traders need to hold tight because closing the position at that moment will give a lower results in the end. The results of our best cointegration pairs portfolio, and our 'random pairs' portfolio (with pairs 101-120) are shown in another table (figure 8 and 9) on the next page. With those pairs similar results are generated.

Portfolio properties	3	3	3	3	3
Formation period	250	250	250	250	250
Trading period	125	125	125	125	125
Trading St.dev.	1.5	1.5	1.5	1.5	1.5
Stoploss	None	2%	5%	10%	20%
Avg. montly return	0.135%	-1.214%	-0.462%	-0.140%	0.062%
St.dev of return	0.569%	0.788%	0.675%	0.606%	0.580%
# trades performed	3840	17795	9896	6385	4518
Value at risk	46227	134759	86727	63869	50981

Results 1994-2004					
Avg. monthly return	0.22%	-1.28%	-0.46%	-0.07%	0.17%
St.dev of return	0.59%	0.66%	0.70%	0.61%	0.59%
Value at Risk	43030	128037	89525	60157	46174

Results 2004-2014					
Avg. monthly return	0.05%	-1.15%	-0.46%	-0.21%	-0.04%
St.dev of return	0.55%	0.91%	0.67%	0.61%	0.56%
Value at Risk	49426	140944	85618	68039	55220

Figure 7: Results of the top-50 SSD-pairs portfolio when applying a stop-loss

Portfolio properties	7	7	7	7	7
Formation period	250	250	250	250	250
Trading period	125	125	125	125	125
Trading St.dev.	1.5	1.5	1.5	1.5	1.5
Stoploss	None	2%	5%	10%	20%
Avg. montly return	0.182%	-0.995%	-0.436%	-0.139%	0.060%
St.dev of return	1.47%	1.76%	1.61%	1.59%	1.57%
# trades performed	468	2821	1622	994	632
Value at risk	124844	207105	165580	151135	138694

Results 1994-2004					
Avg. monthly return	0.31%	-1.02%	-0.37%	-0.06%	0.17%
St.dev of return	1.67%	2.14%	1.89%	1.93%	1.90%
Value at Risk	124844	207105	165580	151135	138694

Results 2004-2014					
Avg. monthly return	0.06%	-0.97%	-0.50%	-0.22%	-0.04%
St.dev of return	1.29%	1.37%	1.34%	1.24%	1.22%
Value at Risk	114052	174079	147988	126066	115132

Figure 8: Results of the top-10 cointegration pair portfolio when applying a stop-loss.

Portfolio properties	10	10	10	10	10
Formation period	250	250	250	250	250
Trading period	125	125	125	125	125
Trading St.dev.	1.5	1.5	1.5	1.5	1.5
Stoploss	None	2%	5%	10%	20%
Avg. montly return	-0.016%	-1.566%	-0.745%	-0.299%	-0.134%
St.dev of return	0.85%	0.73%	0.73%	0.77%	0.68%
# trades performed	1479	8656	4922	3021	2003
Value at risk	80395	147853	105775	87428	70344

Results 1994-2004					
Avg. monthly return	0.08%	-1.71%	-0.73%	-0.28%	-0.03%
St.dev of return	0.80%	0.78%	0.84%	0.83%	0.78%
Value at Risk	80395	147853	105775	87428	70344

Results 2004-2014					
Avg. monthly return	-0.10%	-1.43%	-0.76%	-0.31%	-0.23%
St.dev of return	0.90%	0.67%	0.62%	0.73%	0.57%
Value at Risk	91615	136452	98727	85914	65756

Figure 9: Results of the portfolio with SSD pairs that are ranked 101-120 when applying a stop-loss

Comparing 1994-2004 with 2004-2014

As can be seen in all of the tables above, as well as in the appendix, in nearly all cases, the monthly return on pairs trading is significanty lower in the second decade. This is in accordance with research done by Do & Faff (2010). However, they found that this decline is not due to the increased competition. Rather, the decline is due to worsening arbitrage risks. These are the possibilities of unexpected disruptions between paired securities, noise-trading risk (when irrational trading causes further divergence) and synchronization risk (the uncertainty about when other arbitrageurs will exploit a common mispricing). In addition, these lower results also come with the same amount of volatility which means the risk/return characteristics of pairs trading also decline in the second decade. This perhaps earlier winning strategy is now losing profitability due to its success and follows the same path many other trading strategies have followed.

Conclusions

Unfortunately, a winning strategy has not been found on the S&P 500. The algorithm by Gatev et al. (2006) does not generate monthly positive returns. In addition, profitability has declined in the last decade which means nowadays it is even harder to create a profitable pairs trading portfolio on the S&P 500. An interesting feature of this study however, is that adjusting the standard deviation of the price ratio can significantly increase results. A 1.5 standard deviation difference from the mean is the best option for trading in the past 20 years and could be applied in the future.

Another interesting point is that the saying "Stop your losses and let your profits run" does not turn out to be true when performing pairs trading. The tighter a stop-loss, the worse are the results. Even when faced with an open position that has a more than 20% loss it is better to hold tight until the position is closed at the end of the trading interval, or by crossing the historical price ratio.

When looking at the riskiness of this strategy, this strategy turns out to be pretty good. The one-month 1 million USD value-at-risk is about 10% of the capital which isn't an immediate disaster if things could turn to the wrong side. The volatility is also around 1% a month which is not a lot for a trading strategy.

In conclusion, the biggest winner in this strategy in the way it has been performed in this research is the broker. I strongly advise that if you are planning to invest on the S&P 500 index, you should not implement this strategy as there are plenty of other strategies which can be much more profitable. Simply investing your capital in the index itself or in a t-bill provides better returns over 20 years.

References

- **Gatev et al. (2006). Pairs Trading: Performance of a Relative-Value Arbitrage Rule.** ***The review of financial studies, v 19, n 3, 2006, p. 797 – 827***
- **Ruiter, Hugo Jan (2011). The Performance of a Pairs Trading Strategy in Asian Markets for 2002 to 2009**
- **Do & Faff (2010). Does simple pairs trading still work?** ***The Financial Analysts Journal, Volume 66, number 4***

Appendix

Portfolio properties	7	7	7	7	7
Formation period	250	250	250	250	250
Trading period	125	125	125	125	125
Trading St.dev.	1.5	1.7	2	2.5	3
Avg. montly return	0.182%	0.149%	0.132%	0.115%	0.082%
St.dev of return	1.47%	1.42%	1.18%	0.93%	0.82%
# trades performed	468	408	348	262	191
Value at risk	124844	122744	100520	78429	71416
Results 1994-2004					
Avg. monthly return	0.31%	0.25%	0.19%	0.14%	0.09%
St.dev of return	1.67%	1.65%	1.34%	1.13%	1.04%
Value at Risk	124844	122744	100520	78429	71416
Results 2004-2014					
Avg. monthly return	0.06%	0.05%	0.08%	0.09%	0.07%
St.dev of return	1.29%	1.19%	1.04%	0.72%	0.57%
Value at Risk	114052	106195	90472	60643	48227

Figure 10: Returns on the top-10 cointegration pairs portfolio with different standard deviations

Portfolio properties	10	10	10	10	10
Formation period	250	250	250	250	250
Trading period	125	125	125	125	125
Trading St.dev.	1.5	1.7	2	2.5	3
Avg. montly return	-0.016%	-0.019%	-0.010%	0.038%	0.134%
St.dev of return	0.85%	0.88%	0.85%	0.88%	0.88%
# trades performed	1479	1343	1183	972	815
Value at risk	80395	82952	79010	79433	74890
Results 1994-2004					
Avg. monthly return	0.08%	0.08%	0.01%	0.02%	0.12%
St.dev of return	0.80%	0.93%	0.96%	1.01%	0.87%
Value at Risk	80395	82952	79010	79433	74890
Results 2004-2014					
Avg. monthly return	-0.10%	-0.11%	-0.03%	0.06%	0.15%
St.dev of return	0.90%	0.84%	0.74%	0.76%	0.91%
Value at Risk	91615	86309	72369	68355	78333

Figure 11: Returns of the portfolio with pairs 101-120 as determined by the SSD portfolio with different standard deviations.

Portfolio properties	1	1	1	2	2	2
Formation period	120	200	250	120	200	250
Trading period	60	100	125	60	100	125
Trading St.dev.	1.5	1.5	1.5	1.5	1.5	1.5
Stoploss						
Avg. montly return	-0.268%	-0.271%	-0.125%	-0.437%	-0.189%	-0.169%
St.dev of return	1.417%	1.187%	1.085%	1.013%	0.872%	0.920%
# trades performed	1475	928	724	2971	1828	1461
Value at risk	70580	96299	106932	56708	72440	93575
Results 1994-2004						
Avg. monthly return	-0.27%	-0.38%	-0.03%	-0.47%	-0.23%	-0.26%
St.dev of return	1.42%	1.27%	1.19%	1.09%	0.91%	1.06%
Value at Risk	70219	107053	112045	61216	76615	111420
Results 2004-2014						
Avg. monthly return	-0.26%	-0.16%	-0.21%	-0.40%	-0.15%	-0.08%
St.dev of return	1.43%	1.11%	0.99%	0.94%	0.85%	0.78%
Value at Risk	71679	86163	103054	52537	69410	76279

Figure 12: Results of the top-10(1) and top-20 SSD portfolios (2) when trading different periods.

Portfolio properties	4	4	4	5	5	5
Formation period	120	200	250	120	200	250
Trading period	60	100	125	60	100	125
Trading St.dev.	1.5	1.5	2.0	1.5	1.5	2.0
Stoploss						
Avg. montly return	-0.381%	-0.289%	-0.024%	-0.434%	-0.230%	-0.034%
St.dev of return	1.594%	1.239%	0.815%	1.244%	0.912%	0.609%
# trades performed	1475	928	536	2971	1828	1110
Value at risk	81574	99699	76623	66914	76855	58790
Results 1994-2004						
Avg. monthly return	-0.55%	-0.41%	0.20%	-0.51%	-0.30%	0.08%
St.dev of return	1.71%	1.38%	0.65%	1.44%	1.00%	0.53%
Value at Risk	89971	114557	51147	77357	86018	47049
Results 2004-2014						
Avg. monthly return	-0.21%	-0.17%	-0.24%	-0.36%	-0.16%	-0.14%
St.dev of return	1.47%	1.09%	0.91%	1.03%	0.83%	0.67%
Value at Risk	73259	84726	96005	55811	68184	69246

Figure 13: Results of the weighted top-10 (4) and top-20 SSD portfolios (5) on different trading periods

Portfolio properties	6	6	6	8	8	8
Formation period	120	200	250	120	200	250
Trading period	60	100	125	60	100	125
Trading St.dev.	1.5	1.5	2.0	1.5	1.5	1.5
Stoploss						
Avg. montly return	-0.373%	-0.162%	0.078%	-0.034%	-0.081%	0.127%
St.dev of return	1.098%	0.705%	0.579%	1.52%	1.20%	1.10%
# trades performed	7409	4656	2958	1817	1115	938
Value at risk	58954	59767	50737	67890	91994	94962
Results 1994-2004						
Avg. monthly return	-0.45%	-0.17%	0.17%	0.20%	0.13%	0.30%
St.dev of return	1.31%	0.79%	0.57%	1.80%	1.40%	1.21%
Value at Risk	70072	66198	45279	67890	91994	94962
Results 2004-2014						
Avg. monthly return	-0.30%	-0.16%	-0.01%	-0.26%	-0.29%	-0.04%
St.dev of return	0.85%	0.62%	0.58%	1.15%	0.94%	0.98%
Value at Risk	46497	53741	55982	57955	81231	92458

Figure 14: Results of the weighted top-50 SSD-portfolio (6) and the top-20 cointegration portfolio (8) on different trading periods

Portfolio properties	9	9	9	10	10	10
Formation period	120	200	250	120	200	250
Trading period	60	100	125	60	100	125
Trading St.dev.	1.5	1.5	1.5	1.5	1.5	1.5
Stoploss						
Avg. montly return	-0.085%	-0.199%	0.015%	-0.132%	0.016%	-0.016%
St.dev of return	1.26%	0.97%	0.96%	1.16%	1.21%	0.85%
# trades performed	4502	2794	2317	2969	1870	1479
Value at risk	58215	79442	87481	55128	88911	80395
Results 1994-2004						
Avg. monthly return	0.06%	-0.07%	0.07%	-0.10%	0.13%	0.08%
St.dev of return	1.31%	0.95%	0.92%	1.14%	1.38%	0.80%
Value at Risk	58215	79442	87481	55128	88911	80395
Results 2004-2014						
Avg. monthly return	-0.22%	-0.33%	-0.03%	-0.16%	-0.10%	-0.10%
St.dev of return	1.20%	0.99%	1.02%	1.18%	1.03%	0.90%
Value at Risk	59479	86552	94491	56951	80484	91615

Figure 15: Results of the top-50 cointegration portfolio (9) and the 101-120 ranked SSD pairs portfolio (10) on different trading periods