PAIRS TRADING: OPTIMAL THERSHOLD STRATEGIES

Alejandro Álvarez Roa

Trabajo de investigación 018/018

Master en Banca y Finanzas Cuantitativas

Director/a: Dr. Federico Platania

Universidad Complutense de Madrid

Universidad del País Vasco

Universidad de Valencia

Universidad de Castilla-La Mancha

www.finanzascuantitativas.com

MASTER THESIS

PAIRS TRADING: OPTIMAL THRESHOLD STRATEGIES

Author: Alejandro Álvarez Roa

Supervisor: Federico Platania

Master in Banking and Quantitative Finance

Interuniversitary program: UPV/EHU - UCM - UV - UCLM

June 2018

Abstract

For already four decades, *pairs trading*has been the subject of study of many theorists. Different approaches have been developed in order to trade with pairs of stocks; i.e. the the *distance approach*, the *cointegration method*, the *stochastic spread method*, or the *stochastic residual spread method* among others. In this scene, the aim of the present work is to propose three different threshold selection strategies to determine the best performing trading trigger, in terms of profitability. We perform an empirical study of cointegrated pairs, using the *supercointegration*[1] concept for the pairs selection, together with the implementing of a robustness test. We thereby compare different strategies, during different time periods, and for different pairs of stocks, so that we determine which strategy fits better.

All those strategies have been selected by studying the spread time series and the previous existing literature. In this sense, we start by using a constant threshold, determined by the standard deviation of the historical spread series. Then, we test how a dynamic threshold works, based on the conditional volatility of the spread series. After previously confirming that the cointegrated series do not have a symmetric behaviour between the positive and the negative part, we then try to solve this problem by fixing different trading triggers; classifying, moreover, per percentile, each part. Furthermore, the study of the stationary spread series suggests that it could have a cyclic component, because of what, we try to model it via spectral analysis, obtaining a smoothed series from which we can compute an optimal trading threshold.

The testing of the above mentioned strategies indicates that the spectral analysis, and the conditional volatility, are indeed significant variables. Therefore, we develop an artificial neural network where these variables are introduced in the model. This machine learning model is able to achieve more profit in the majority of the cases.

Keywords: Trading strategy; Pairs trading; Supercointegration; Threshold selection.

[1] [15] Figuerola-Ferretti et al. (2017)

Acknowledgements

I would first like to thank my thesis advisor Federico Platania, of the Pôle Universitaire Léonard de Vinci (París), for supporting me during my research stay in Paris, for his patience and motivation. He consistently allowed this project to be my own work, but steered me in the right direction whenever I needed it.

I would also like to acknowledge Pedro Serrano, from the university Carlos III of Madrid, who gave us his expert point of view in the Pairs Trading field, and proposed some research paths.

Finally, I would like to thank my family: my parents and my sister for encouraging and supporting me throughout writing this master thesis.

Contents

List of Figures

List of Tables

1. Introduction

Pairs trading was first introduced by a quantitative group, in Wall Street, in the mid '80s. They wanted to find statistical arbitrage opportunities in the equity markets. The aim of the strategy was to trade with a highly-correlated pair of stocks, or some which had similar price co-movements, in the long run, in order to take advantage of the deviations from this equilibrium in the short run. When there is a divergence from the equilibrium, the pair contains an overvalued stock, and an undervalued stock too. Therefore, the strategy will simply consist in going long and short, respectively, on the stocks, until the equilibrium is reached again.

Some authors have developed other methods such as the *cointegration method*. This was introduced by Vidyamurthy (2004) [34]. It contains an important statistical relationship which can capture the stationarity of the spread[2] series. When two stock series are cointegrated it means that, being integrated of order d, they can be linearly combined obtaining a unique time series integrated of order $d - b$, where b>0. Stock prices which are both integrated of order 1, can be combined obtaining an $I(0)$ stationary series. The stationary property allows to trade with mean reverting series, and help decide when to open or close a position.

Another relevant question addresses the trading trigger selection. Once the statistical relationship for the pairs trading is chosen, there are many different ways to decide when to open or close a position. The volatility of the spread series has been determinant for the threshold selection in many important works, such as Gatev et al. (2006) [17]. The latter, established a constant threshold based on the volatility of the spread. Also, Figuerola-Ferretti et al. (2017)[15], introduced the concept of

[2] Relationship between two stocks.

supercointegrated pairs, working with a dynamic threshold based on the conditional volatility.

The main objective of the present pages is, thereby, to propose three different threshold selection methods, performing the analysis under *supercointegrated pairs*, in order to determine if they are more efficient than two of the existing threshold selection strategies.

The work is organized as follows. First, we outline the four main approaches to trade with pairs, passing through the *distance method*, the *cointegration method*, the *stochastic spread method*, and the *residual spread method*. After, adopting the *cointegration method*, we explain in section 3 how he select the pairs. Here, the work of Figuerola - Ferretti et al. (2017) [15] will have especial importance, since we adopt their *pairs selection method*. We perform the *Engle - Granger* and the *Johansen cointegration methods*, under a high confidence level, taking into account the best *p-value* indicators. Next, in section 4, we indicate five different approaches we have studied for the determining of the best performing trading threshold. section 5 explains which of the previous methods has consistently performed better. In section 6, eventually, we outline our main obtained conclusions.

2. Literature Review

In this section we introduce the main existing approaches to implement pairs trading: the distance method proposed by Gatev et al. (1999) [16], the cointegration method introduced by Vidyamurthy (2004) [34], the stochastic spread method outlined in Elliott et al. (2005) [11] and the stochastic residual spread method by Do et al. (2006) [7].

2.1. The Distance Method

Since 1999, some authors have worked under this method. We can see Gatev et al. (1999, 2006) [16], Nath(2003) [28], Engelberg et al. (2008) [13], Papandakis and Wiscocky(2008), Do and Faff (2008) [?], Chen et al. (2012)[5] and Huck (2015) [23] as the most relevant works. The way to trade with this approach is to measure the distance of the co-movement of a pair of stocks, or the sum of squared differences between the two normalized price series. The most referenced paper regarding this method is Gatev et al. (1999, 2006) [16]. They first construct a cumulative total return index (P_{it}) for each stock during a formation period and they set the pairs, to start the trading period as follows:

$$P_{it} = \prod_{s=1}^{t}(1 + r_{is}) \tag{1}$$

where P_{it} is the price of stock i at time t, and r_{is} corresponds to the stock total return, including reinvested dividends.

They the take the normalized prices and compute the squared normalized price difference measure PD, taking the square euclidean distance [3] as the difference.

$$PD_{ji} = \sum_{t=1}^{T}(P_{it} - P_{jt})^2 \tag{2}$$

where i and j represent the stocks in a pair, T is the number of trading days during

[3] Euclidean distance is the ordinary straight-line distance between two points in Euclidean space

the formation period, P is the normalized price for the stock i or j in a day t. Taking N as the total number of stocks, there are $N(N-1)/2$ possible combinations.

They then select the pairs with the lowest PD for the trading period, as they have been more related in the past. They select 20 pairs, for a 6 months trading period, taking two standard deviations of the spread difference as the threshold to open a position, being the standard deviation as follows:

$$\sigma(PD_{ij}) = \frac{1}{T-1} \sum_{t=1}^{T} \left[(P_{ti} - P_{tj})^2 - \overline{(P_{ti} - P_{tj})^2} \right] \tag{3}$$

Opening a position means taking a long position on the undervalue stock of the pair, and taking a short position on the overvalued one. The opened position ends when the prices of both stocks converge.

Gatev et al. (1999, 2006) [16] demonstrated a profitable trading strategy after transaction costs, but the problem with non-cointegrated pairs is that this relation can diverge in the long run. Huck (2015) [23] indicated that cointegrated pairs have a more mean reverting spread than non-cointegrated pairs, which can help us to arrive with more probability to the close position. Therefore, Do and Faff (2009, 2012) [8] showed that the profitability of the strategy was decreasing over time, with a negative return after transaction costs, and Engelberg et al. (2008) [13] also proved an exponentially declining profitability.

The distance trading is a non-parametric method working with a statistical relationship between two stocks, which does not provide a convergence time. There could also be a problem with the assumption of a static distance of two stocks over a long period of time. The selection of the pair regarding the risk-return profiles should be close to identical

2.2. The Stochastic Spread Method

We can find this model in Elliot et al. (2005) [11], where they use the mean reversion of the spread to model the behaviour of the difference between the prices of two stocks, taken as the spread. In this method, the spread is model by a Vasicek process:

$$dx_t = \kappa(\theta - x_t)dt + \sigma dB_t \tag{4}$$

where x is the state variable, κ is the mean reversion speed, θ is the mean of the state variable and dB_t is a standard Brownian motion in some defined probability space. The parameters are estimated applying the Kalman filter. This model has three main advantages:

- It models the mean reversion behaviour of the spread. In this model, x can be either positive or negative, by definition of the spread.

- It can be explicitly obtained the expected time of the spread to revert to its long term mean, as it is a continuous time model. It can be computed the first passage time for the Vasicek process, through the expectation $E[\tau|x_t]$, where τ is the first time the x hits its mean θ, for a time t.

  By the Ornstein-Uhlenbeck process, the first passage time for x_t is:

$$T = inf\left\{t \geq 0, x_t = \theta | x_0 = \theta + \frac{c\sigma}{\sqrt{2\kappa}}\right\} = \hat{t}\kappa \tag{5}$$

  where

$$\hat{t} = 0.5ln\left\{1 + 0.5\left[\sqrt{(c^2-3)^2 + 4C^2} + C^2 - 3\right]\right\}$$

  In Equation 5, T represents the time expended by the spread to hit θ, with $t = 0$, $x_0 = \mu + c\sqrt{\sigma\kappa}$, being c a constant. They settled a strategy based on opening a position when $y_t \geq \theta + c\sigma/\sqrt{2\kappa}$, where $c > 0$ [4] and closing it at time T. Similarly, for the case of the lower boundary, they open a position when $y_t \leq \theta - c\sigma/\sqrt{2\kappa}$.

[4] They don't clearly specify how to choose that value.

- As we explained before, the estimation of the parameters in this model is not a problem, as they estimate them applying the Kalman filter in a state space model. It is an optimal maximum likelihood estimator in terms of MMSE (Minimum Mean Square Error)

Nevertheless, Do et al. (2006) [7] outline that there could be a problem in the long run with this approach. For a chosen pair, both stocks should give us the same return, as any deviation from this return should be corrected in the future. This is a huge limitation for the model as it is not easy to find a pair with same returns.

2.3. The Stochastic Residual Spread Method

We can find the Stochastic Residual Spread approach in Do, Faff and Hamza (2006) [7]. They propose to work with returns instead of using prices during the Pairs Trading. Do et al. (2006) [7] introduce their model as an alternative to the existing purely statistically based approaches. This model assumes an equilibrium between two stocks called spread, and there exists a misprincing when this equilibrium disappears. This mispricing is quantified by a residual spread function $G(R_t^A, R_t^B, U_t)$,where U denotes some exogenous vector potentially present in formulating the equilibrium. That means that the function can identify any deviation form the long term spread. The trade is then opened when there exists such a disequilibrium sufficiently large and the expected correction timing is sufficiently short. Do et al.(2006) [7] base their model in Elliot et al. (2005) [11], explained in the previous section, taking the Vasicek model as the followed process for the state of residual spread:

$$dx_t = \kappa(\theta - x_t)d_t + \sigma dB_t \tag{6}$$

The residual spread equation is:

$$y_t = G_t = x_t + \omega_t \tag{7}$$

Under a measurement of noise, the residual spread function G has to be estimated. They address the APT (Arbitrage Pricing Theory) of Ross(1976), adapted to a pair of stocks. Being the APT for n factor as follows:

$$R^i = R_f + \beta_t^m + \eta^i \tag{8}$$

where R^i is the return for the ith factor, R_f is the risk free return, $\beta = [\beta_1^i \beta_2^i ... \beta_n^i]$ is a vector with the sensitivities of the return for the ith factor to each risk factor and $r^m = [(R^1 - r_f)(R^2 - r_f)...(R^n - r_f)]^T$ and η is the residual with zero mean as APT works on a diversified portfolio.
They can then define the "relative" APT for a pair of stocks:

$$R_t^A = R_t^B + \Gamma r_t^m + e_t \tag{9}$$

Thus, the residual spread equation is:

$$G_t = G(R_t^A, R_t^B, U_t) = e_t = R_t^A - R_t^B - \Gamma r_t^m \tag{10}$$

G_t is observable under a specification of r_t^m, and given the values of Γ. When the values for the vector Γ of exposure differentials are zeros, this model is the same as Elliot et al. (2005) [11].

Do et al. (2006) [7] have formulated a mean reverting model in continuous time, through the relative pricing model from APT. Nevertheless, they stress that this approach fails to make any assumption on the validity of the APT model. It is not clear if the APT factor model is rigorously applicable in this price scheme.

2.4. The Cointegration Method

This is the method we are going to use for our pairs trading strategies. We can find cointegration approach in Vidyamurthy (2004) [34]. He proposes a non-empirical pairs trading method based on the cointegration relationship outlined by Engle and Granger (1987) [12]. Cointegration occurs when two time series that are integrated

of order d can be linearly combined obtaining a unique time series integrated of order $d-b$, where $b>0$. A variable x_t is integrated of order d ($x_t \sim I(d)$) if it becomes stationary after differentiating it d times and it has a number of unit roots equal to d. Thus, a variable x_t is stationary when it is integrated of order 0 ($x_t \sim I(0)$) so that it does not have a unit root.
Thanks to the stationary property, there will be a mean reverting cointegration relationship, which is going to be decisive for the strategy within the cointegration approach.

Vidyamurthy (2004) [34] uses the Engle-Granger test, following an ordinary least squares regression, to identify a cointegration relationship between two stocks price time series, as follows:

$$log(P_t^A) = \mu + \beta log(P_t^B) + \epsilon_t \qquad \epsilon \sim N(0, \sigma_\epsilon^2) \tag{11}$$

where P_t^A and P_t^B are the prices for the stocks A and B respectively, μ is a constant, β is the cointegration coefficient and ϵ are the residuals.

Therefore, we can rewrite the cointegration relationship, or spread, as follows:

$$log(P_t^A) - \hat{\beta} log(P_t^B) = \hat{\mu} + \hat{\epsilon}_t \tag{12}$$

Within the Engle-Granger approach, the residuals (ϵ) of the cointegration relationship have to be $I(0)$, so they apply the Augmented Dickey Fuller Test (ADF) in order to verify the stationarity. If the test assert that that the relationship is stationary, it means that we have an $I(0)$ linear combination of two $I(1)$ series. The selection of the dependant and the independent variable will change the residual series (i.e. if $log(P_t^B)$ is regressed against $log(P_t^A)$). The t-statistics in Engle and Yoo (1987) [12] can solve this problem.

Vidyamurthy's work studies the mean reversion property of the spread. The spread in a cointegration relationship must have a long run mean of μ, where the mean of the residuals ϵ is equal to zero:

$$E(log(P_t^A) - \hat{\beta} log(P_t^B)) = E(\hat{\mu} + \hat{\epsilon_t}) \tag{13}$$

$$E[log(P_t^A) - \hat{\beta} log(P_t^B)] = \hat{\mu} \tag{14}$$

Thus, when the mean is not μ there is no equilibrium and there is a trade opportunity, consisting of long 1 unit of A and short $\hat{\beta}$ units of B. For the mean reversion analysis of the spread, this author proposes two different methodologies. An ARMA process is applied, as a parametric approach, in order to capture the mean reversion property of the residuals ($\hat{\epsilon}$), and Rice's (1945) formula is used to capture the optimal trading trigger (Δ) that maximizes the profit. Regarding this approach, the first passage time expression, explained in section 2.3 would be more appropriate as Rice's method does not take into account the restriction when the trade is opened.

The second method is a non-parametric empirical approach, consisting of a distribution of zero crossings which can capture the mean reversion in order to set an optimal trading trigger (Δ).

There are some other works where the authors perform the trading strategy based on the cointegration methodology, like Lin et al. (2006) (applying the Cointegration Coefficients Weighted in the cointegration approach), Chen et al. (2012) or Puspaningrum (2012). Do et al. (2006) [7] find a problem related with the APT of Ross (1976), regarding the economic theory and not only the pure statistical results. Bierens and Martins (2009) propose a time varying cointegration vector error correction model in which the cointegrating relationship varies smoothly over time.

Regarding the four main methods we have outlined, we are going to base our study in the cointegration approach. Therefore, the proposed strategies are going to be developed considering the cointegration property. This allows us perform an investigation based on some assumptions we can deduce from the cointegration property.

3. Pairs Trading: Supercointegration method for pairs selection

In this section we follow the work of Figuerola-Ferretti et al. (2017) [15] to the pairs selection. The authors of this study tested that the level of cointegration had a positive impact on the pairs trading profitability, thus, we search *supercointegrated pairs* in the SP100 index. To do so, we perform Johansen and Engle - Granger tests, restricting the results to a high confidence level.

3.1. Cointegration

As we advanced in the previous section, we are going to base the pairs trading strategy in the cointegration method. It provides the characteristics we are going to focus on during the establishment of the proposed threshold strategies.

Within the cointegration method, we can find pair of stocks satisfying a relationship in the long run, that can be identified following different methods, such as the Johansen Test or the Engle - Granger Test.

In Figure 1 we can see an example of two cointegrated stocks:

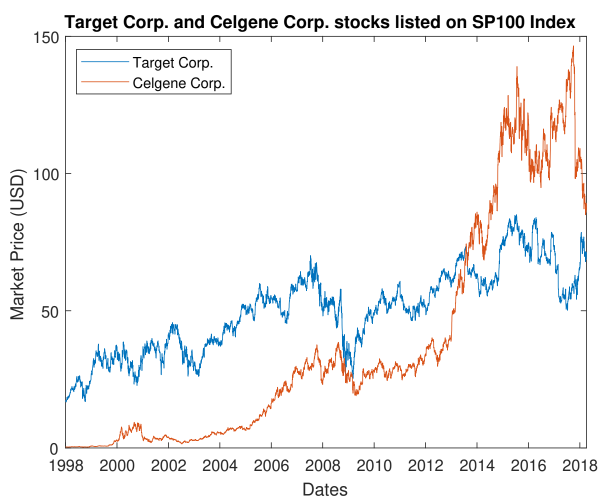
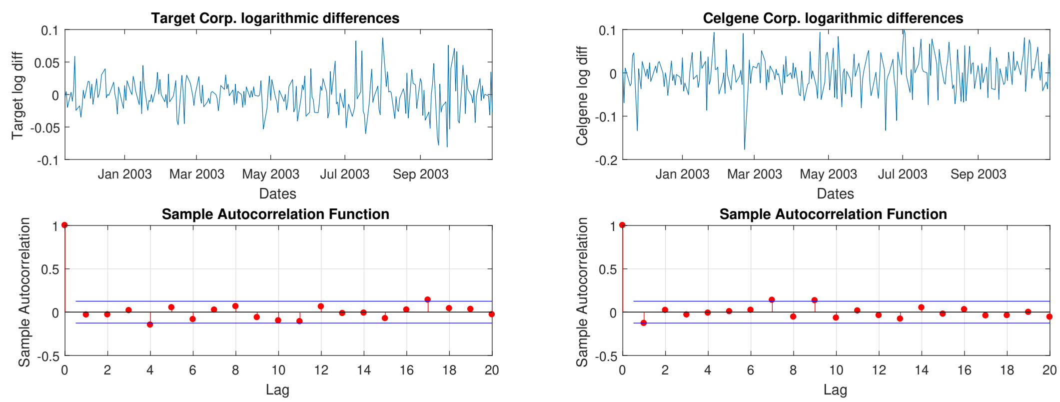


Figure 1: Cointegrated stocks: Target Corp. and Celgene Corp. stocks prices.

When two series of stocks are cointegrated, it means that, being separately integrated of order d, they can be linearly combined obtaining a unique time series integrated of order $d-b$, where $b>0$. As we outlined in the subsection 2.4, a variable x_t is integrated of order d ($x_t \sim I(d)$) if it becomes stationary [5] after differentiating it d times and it has a number of unit roots equal to d. Stock logarithmic prices are commonly integrated of order 1, therefore, differentiating them one time, we can obtain an $I(0)$ series.

The Figure 2 below shows, for a selected pair of stocks, a graphical analysis of the stationary property of the series. After differentiating the logarithmic prices one time, there is no evidence of the existence of a unit root. Nonetheless, we can run the Augmented Dickey Fuller Test (ADF) to verify the stationarity of the series. We will apply ADF test later on during the analysis of the cointegrated pairs.



(a) Target Corp. (b) Celgene Corp.

Figure 2: Stationarity analysis of the stocks

3.2. Data selection

We have daily data of stock assets included in the S&P 100 from 1st January 1998 to 4th April 2018. There are 5285 daily close observations for 76 assets, after remov-

[5] Stationary series are integrated of order 0.

ing missing value stocks. [6]

3.3. Johansen Method

The first step in the selection of *supercointegrated pairs* is to apply the Johansen Test to identify which pairs are cointegrated in the SP100 Index. The level of cointegration will be restricted to a confidence level of 99% and an accurate *p-value*.

Johansen test, based on the vector autoregressive (VAR), analyse all the existing relationships between the series of stocks, finding a maximum number of $n-1$, where n is the number of variables in the sample.

Considering an autorregressive vector $VAR_k(\rho)$

$$y_t = \mu_t + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_\rho y_{t-\rho} + Bx_t + \epsilon_t \tag{15}$$

where y_t a vector containing k non-stationary variables $I(1)$, x_t is a vector of deterministic variables and ϵ is the vector of innovations.

We can rewrite the VAR model as the Vector Error Correction Model (VECM):

$$\Delta y_t = \mu_t + \Pi y_{t-1} + \sum_{i=1}^{\rho-1} \Gamma_i \Delta y_{t-i} + Bx_t + \epsilon_t \tag{16}$$

where

$$\Pi = \sum_{i=1}^{\rho} A_i - I_k, \quad \Gamma_i = -\sum_{j=i+1}^{\rho} A_j, \quad i = 1, 2, ..., \rho - 1 \tag{17}$$

Hence, Johansen establishes that one option to test for the existence of cointegration is to analyse the rank of the matrix Π of long-term multiplier. The VAR model is estimated by maximum likelihood and then the estimation rank of the matrix Π is

[6] AbbVie, Biogen Indec, Bank of New York Mellon, Caterpillar, Exelon, Facebook, Fedex, General Electric, Gilead Sciences, General Motors, Google, Johnson & Johnson, JP Morgan Chase, Lockheed Martin, Mastercard, Mondelez International, 3M, Nexterea Energy, Oracle, Pfizer, Procter & Gamble, United Health Group, US Bancorp Visa.

analysed. The are two contrasts: maximum eigenvalue test and trace test, that are executed sequentially.

Within the trace test, the null hypothesis is:

$$H_0 : Rank(\Pi) \leq m$$

$$H_1 : Rank(\Pi) > m$$

H_0 indicates that the variables belonging the vector y_t have a maximum number of m cointegration relationships, while H_1 express that the variables belonging the vector y_t have more than m number of cointegration relationships.

Where the statistic is defined as:

$$LK_{tr}(m) = -(T-\rho)\sum_{i=m+1}^{k} ln(1-\hat{\lambda}_i)$$

where λ_i are the generalized eigenvalues estimated for a given matrix arising in the estimation process by Maximum Likelihood

Within the maximum eigenvalue test, the null hypothesis is:

$$H_0 : Rank(\Pi) = m$$

$$H_1 : Rank(\Pi) = m+1$$

Where the statistic is defined as:

$$LK_{max}(m) = -(T-\rho) ln(1-\hat{\lambda}_i)$$

After analysing our data, we compute the model considering intercepts in the cointegrating relations and trends in the data. The AIC and BIC criterion suggest to use 2 lags in the VAR order.

Running the Johansen test, with $\alpha = 0.01\%$ and selecting the pairs with the best *p-value* indicators, we find 31 cointegrated pairs, which are displayed in Table 5 in the annex.

3.4. Engle - Granger Method

The next step in the selection of *supercointegrated pairs* is to apply the Engle - Granger methodology to identify which stocks from the precedent list are cointegrated. We will perform this test under a confidence level of 99%. As we remark in the literature review section, this method was first applied by Vidyamurthy (2004) [34] in the pairs trading scenario following the cointegration approach.

Engle - Granger method search for cointegrated pars running two steps. Firstly, it has to find a relationship by performing an OLS regression. The residuals of this regression are then tested for the stationarity property. If both conditions are fulfilled, the pair is identified.

Equation 18 shows the OLS regression, performed in the first step of the Engle - Granger method, to find the relationship between two stocks.

$$log(P_t^A) = \mu + \beta log(P_t^B) + \epsilon_t \qquad \epsilon \sim N(0, \sigma_\epsilon^2) \tag{18}$$

where P_t^A and P_t^B are the prices for the stocks A and B respectively, μ is a constant, β is the cointegration coefficient and ϵ are the residuals.

Once identified the residuals in the OLS regression, the second step for the test of a unit root is conducted, applying the Augmented Dickey Fuller Test:

$$\Delta\epsilon_t = \alpha + \beta t + \gamma\epsilon_{t-1} + \delta_1 \Delta\epsilon t - 1 + ... + \delta_{\rho-1}\Delta\epsilon_{t-\rho+1} + u_t \tag{19}$$

where α is a constant, β is the trend coefficient, ρ is the number of lagged difference terms and u is a mean zero innovation process.

With the model above, we can make the testing for the null hypothesis of a unit root:

$$H_0 : \hat{\gamma} = 1$$

$$H_1 : \hat{\gamma} < 1$$

where the statistic is:

$$DF_\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

If the test fails to reject the null hypothesis, it fails to reject the possibility of a unit root.

Therefore, we apply the explained methodology to the previous selected pairs with Johansen test. As we did in the subsection 3.3,and following the work of Figuerola-Ferretti et al. (2017) [15], the Engle-Granger method is restricted to a confidence level of 99%, with $\alpha = 0.01$. We find 30 cointegrated pairs, which are displayed in Table 6 in the annex.

4. Optimal Threshold Selection

In this section, we show five different strategies we have performed, in order to determine the best operating threshold for the pairs trading. There will be an in sample period of 3 years in which we will show our strategy motivation, and 1 year out of sample to test the performance of each strategy. All the strategies are illustrated for a particular pair of stocks and period of time, in order to see how the different trading triggers can adapt the spread series, maintaining the same sample data. The selected pair of stocks are *Target Corporation* and *Celgene Corporation*, during the period 1999 to 2003 [7].

The importance of the threshold selection falls on the profit we can obtain during each trade. We must identify the highest point the spread series will reach before returning to the mean, in order to get the maximum profit per operation. We will not focus our work on the closing threshold selection, as we stablish it at the mean line of the spread due to the mean reverting property. We explain two traditional approaches based on the unconditional and conditional volatility, and we then try to propose three different alternative methods to select the optimal threshold.

The first approach relies on Gatev et al. (2006) [17], who stablish a constant threshold based on the unconditional standard deviation of the spread series.

The second method shows a dynamic trading trigger as Figuerola-Ferretti et al. (2017) [15] do, determining the threshold as the conditional standard deviation of the spread.

The third strategy try to decompose the spread series into the positive and the negative part, in order to demonstrate the different distributions they follow. The

[7] Note that the selected example is showing a high concentration of data in the positive part of the spread series. This will allow us be more clear in the explanation of the bounds, in prejudice of showing how the thresholds adapt the negative part. In section 5 we will implement a robustness test in order to determine which strategy is consistently performing better.

constant trading trigger will be determined regarding some different levels of percentile for each part.

Within the fourth method we try to forecast the spread series for a period of time thanks to a spectrum analysis. The mean reverting process followed by the spread suggests a cyclic behaviour. We propose a dynamic trading threshold.

The fifth approach makes use of an artificial neural network model to predict a dynamic trigger. We want to combine two of the previous techniques within a non-parametric approach. The chosen techniques have been the conditional variance and the spectral analysis, since they can provide important information to the model.

4.1. Fixed Threshold

In this section we perform a a constant trading trigger strategy. Gatev et al.(2006) [17] and some others like Do and Faff (2009) [8] work with that kind of strategy. They set a constant threshold defined as two unconditional standard deviations from the in sample spread. In this work, we notice that the historical spread series is not able to reach that level of two standard deviations in the most of the cases. Therefore, after making an analysis of the available data, we set this constant threshold as one standard deviation. In the Figure 3 we can see an example of the *Target Corp.* and *Celgene Corp.* spread series, showing the two different standard deviation levels.

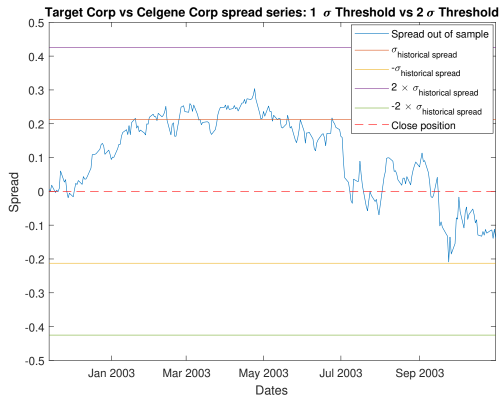


Figure 3: $1 \times \sigma_{hist.spread}$ vs $2 \times \sigma_{hist.spread}$

In the upper bound, the two standard deviations level is too high for the sample, while one standard deviation level is closer to this spread series. In the case of the lower part, the cointegrated series do not reach the minus one standard deviation level. This is an optimal example to introduce the subsection 4.3, where we try to avoid the assumption that both positive and negative parts have the same distribution. Furthermore, we show that taking the lower trade trigger as two standard deviations would not be appropriate, since we could not open a position the vast majority of the cases.

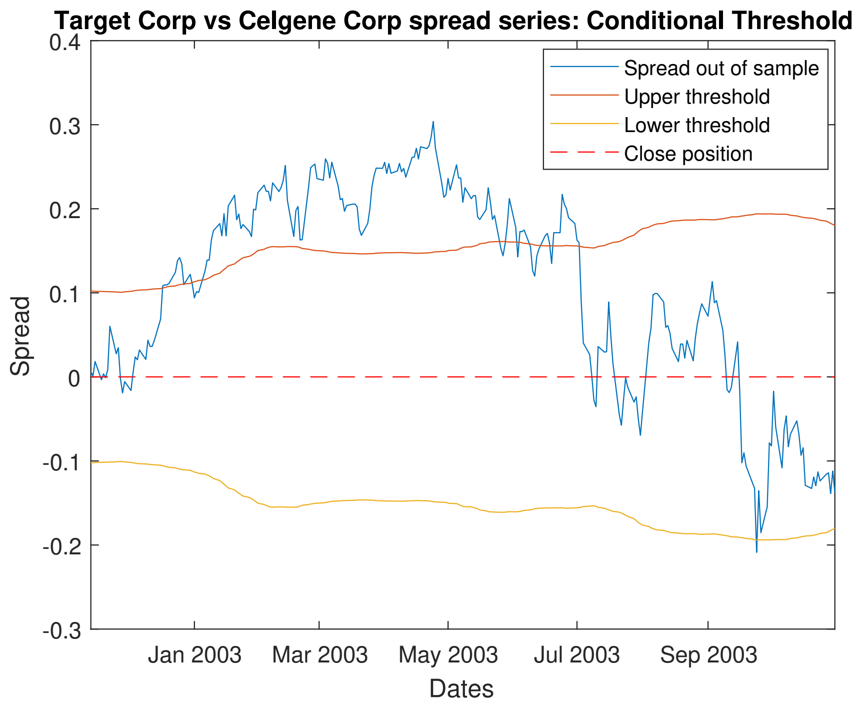
4.2. Conditional Volatility Threshold

The previous approach shows a invariant trading trigger, based on the historical spread series and remaining constant during the whole testing period. In this section, we perform a dynamic threshold taking advantage of the most recent information (Figuerola-Ferretti et al. (2017) [15]).

As the referenced paper does, we set the threshold regarding the conditional volatility of the spread time series. The chosen criteria is to define this dynamic threshold taking a 1 year rolling window to capture the time-varying nature of price spreads.

In the Figure 4 we can see an example of this conditional volatility threshold, for the same pair as before[8].

Figure 4: Conditional Threshold



We can appreciate how, for this particular case, this moving threshold can adapts the spread series better than the previous constant trading trigger in the lower part of the sample.

Nevertheless, we continue observing a non-symmetric performance of the spread series. For this reason, in the next section we try to develop a strategy taking into account this issue.

[8]Spread series containing the pair of stocks of *Target Corp.* and *Celgene Corp.*

4.3. Percentile Based Threshold

Within this approach, we try to adapt the strategy to the the asymmetric distribution that spread series present in the short run.

As we mentioned in the precedent sections, a symmetric trading trigger could not be efficient. There are some times that, while one (upper or lower) part of the spread performs good when one standard deviation of the hole sample is set as the threshold, the opposite part could not reach such a "high" level.

In Figure 5, we can see how the pair *Target Corp.* and *Celgene Corp.* is distributed. In this figure, we compare the sample distribution with a standard normal. We can observe two different distributions corresponding to this spread series.

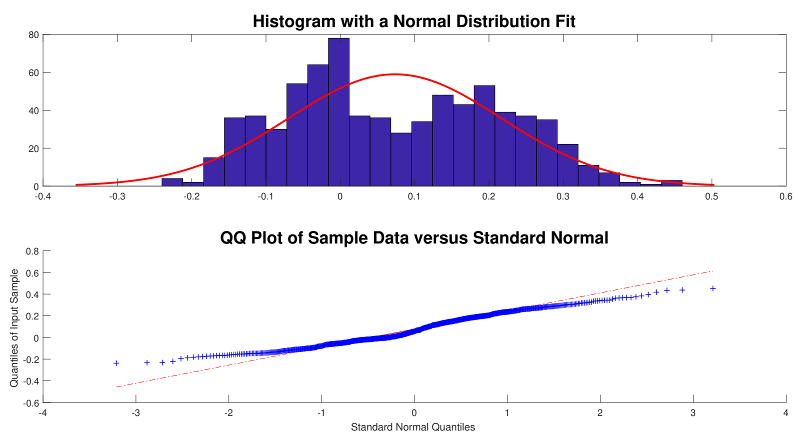


Figure 5: Sample vs Standard Normal

Analysing the whole sample, we test that this characteristic is repeated over the spread series.

As an example, in this particular case, the mean and standard deviation of the positive values of the spread time series is $\mu_{spread\ positive} = 0.1661$, the correspondent mean and standard deviation for the negative part is $\sigma_{spread\ positive} = 0.0986$ and $\mu_{spread\ negative} = 0.0736$ and $\sigma_{spread\ negative} = 0.1432$

Therefore, in this section we separate each positive and negative part of the spread, regarding different percentile levels.

Assuming a normal distribution, when we set the threshold as one standard deviation, we leave the 34.1% of the sample out of the threshold, which corresponds to a percentile 65.9. We are going to examine different percentile levels, separating the positive and negative part, in order to determine which one can achieve better results. These results appear in section 5.

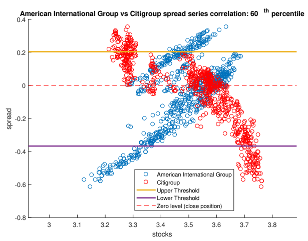
In the figures below, we can observe, for a particular example, how the stock prices are correlated with it corresponding spread. Once again, we can notice that the majority of the sample data is accumulated in one of the two regions.

The displayed example shows three different levels of constant threshold. For the figure a) in the left, we have chosen the percentile 60, leaving the 40% of the whole sample out of the threshold [9]. The figure b) in the right displays a classification for a 70^{th} percentile, and the figure c) is configured with respect to the 80^{th} percentile.

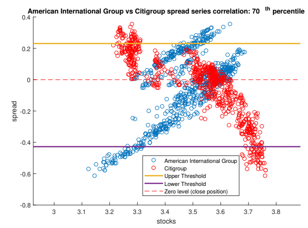
[9] The positive threshold is set on the 80^{th} percentile and the negative threshold is set on the 20^{th} percentile.

Figure 6: Scatter stocks

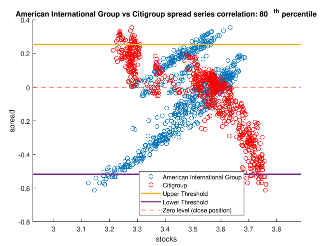
(a) 60^{th} Percentile



(b) 70^{th} Percentile



(c) 80^{th} Percentile



Once we have identified the trading triggers for the different percentile classifications, we can apply it to the spread series we are trading with. In the figure below, we have illustrated how these different thresholds have performed. In this particular case, in the upper bound, the three trading triggers are hit by the spread series before turning back to the closing point.

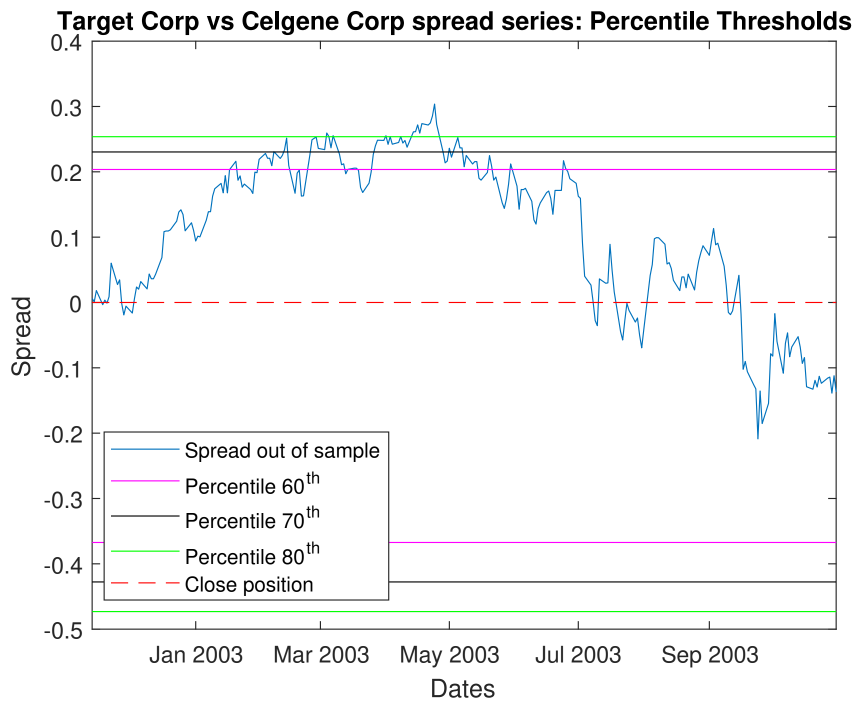


Figure 7: Comparison three thresholds by percentile

Hence, the percentile 80^{th} will obtain the maximum profit from the trading pair. Nevertheless, we will see in the section 5 that for different pairs and periods of time, this percentile configuration is not always the best performing.

Note that in this particular case, the established triggers are not able to capture the lower part of the spread series. This section of the spread, in the displayed figure, has not a high concentration of data in comparison to the historical series. In consequence, the lower trading triggers are not capturing the performance of this part of the figure. With the robustness test, in section 5, we will check the strategy for different pairs and for different periods of time, in order to eliminate the dependence of a particular period or spread time series.

4.4. Spectral Analysis Threshold

In this section, we propose an approach based on the cyclical component of the spread series. Assuming that the spread has embedded a periodic behaviour, we can easily identify the main cycles via spectral analysis. Filtering the relevant frequencies allow us to smooth the spread time series and extrapolate the cyclical behaviour. Identifying the peaks and troughs we can optimize the threshold providing a more reliable and accurate trading signal. As stated above, the spread follows a stationary process; moreover, choosing "supercointegrated" pairs reinforces this condition providing a perfect environment to perform a frequency domain analysis.

First, applying a fast Fourier transform allow us to work on the frequency domain, then via spectral analysis we can identify the relevant frequencies for each pair. Considering the pair composed by *Target Corp.* and *Celgene Corp.*, Figure 8 presents an example where the upper graph presents the time domain signal, that is the spread evolution, and the lower graph presents the spectra:

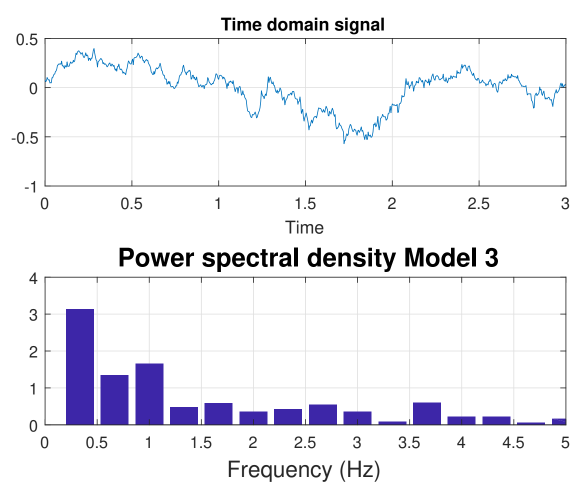


Figure 8: Spectral Analysis

Once we have successfully identified the relevant set of frequencies, we compute an OLS regression to obtain estimates of the amplitude parameter as follows:

$$spread_t = \mu + \sum_{i=1}^{N} \{\alpha_i sin(2\pi\omega_i t/252) + \delta_i cos(2\pi\omega_i t/252)\} + \epsilon_t \tag{20}$$

where μ is a constant, ϵ corresponds to the residuals and α and δ are the amplitudes for the frequency $i-th$.

In the Figure 9 we can observe the in-sample fitting of the spread series.

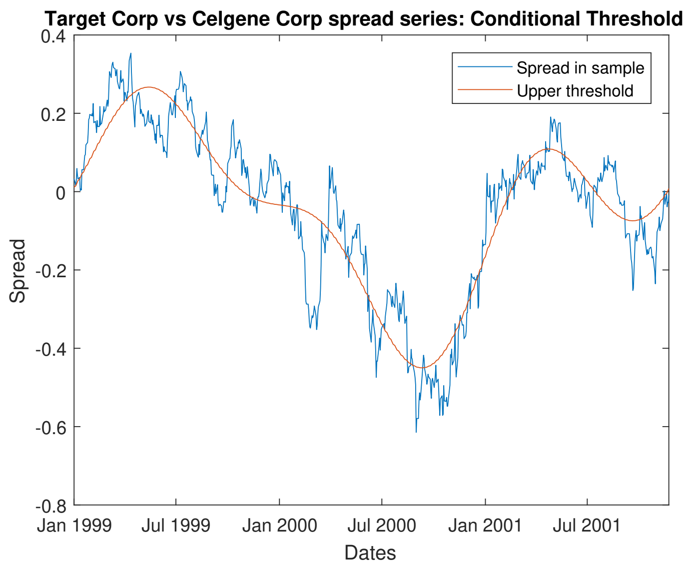


Figure 9: Spread in sample Spectral Fitting

Using the frequencies and the estimated amplitudes, α_i and δ_i , we can extrapolate the series and considering the peaks and troughs we can predict the dynamic threshold. We can easily compute an algorithm to obtain the optimal thresholds from this smoothed series. For a time t < h, where t is the current time and h is the next maximum, we have taken the next maximum (for the upper threshold) or minimum (for the lower threshold) point h of the series, and we have set the threshold at this point h, until the point h is reached. When the series arrive to this point, we must take the series between h and 0 as the threshold until the close position is hit (the zero point in our case). Figure 10 provides an example of the extrapolated smooth series and the forecasted threshold.

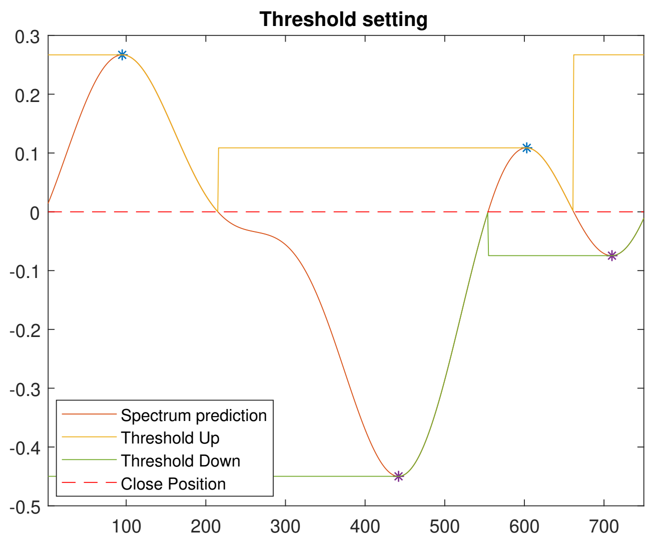


Figure 10: Threshold Setting

Once the threshold for a particular spread series is established, we test the strategy out of sample. The following figure displays an example of how the called "spectral threshold" is capturing the performance of a series in the positive region, for a testing period of one year.

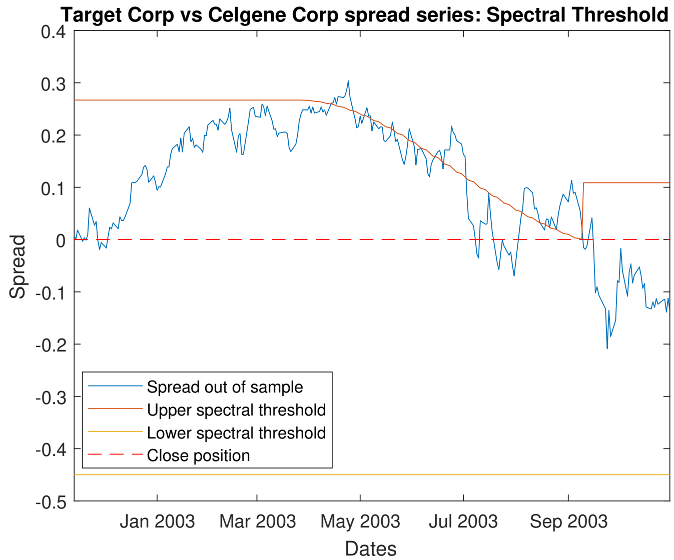


Figure 11: Spectral Threshold

In this example, the spectrum analysis has been able to adapt the trading trigger in the upper bound to the performance of the spread series. The strategy consists in opening a position when the spread hits the upper dynamic threshold. The important fact here is to identify the highest point the cointegrated series can reach, in order to gain leverage with the strategy. As in the previous mentioned strategies, the position is closed when the zero line is crossed. Here, the lower bound is just an extension of what the spectral analysis is predicting as the lowest point the spread can reach during the followings days.

4.5. Neural Network Threshold

Within this approach, we are trying to improve the results of our strategy with respect to the previous methods. We want to combine two of the previous techniques that have performed in quite an acceptable way.

One of this techniques included in the model is going to be the spectral analysis. We have noticed that the calculated threshold has reported good results. In the

Spectral Analysis explained above, we have adjusted the series with a parametric regression, but this time it is going to be estimated using an Artificial Neural Network, a non-parametric method.

In addition, we are going to include in this model the adjustment error of the spectral analysis, in order to provide important information to the network.

Furthermore, regarding the results, we have considered that the conditional volatility of the spread has relevant importance in the threshold selection. Hence, it is going to be another input in the neural network.

Given these three input variables, the predicted output is going to be focused in the optimal threshold selection. This selection is biased by the assumptions we want to take. In this case, we have chosen the higher (lower) value of the spread, during 30 days, actualising it each week (5 days), in order to capture the maximum (minimum) values reached by the spread, which can maximise the profits.

We are going develop two neural networks for each pair of stocks; one for the upper bound and one for the lower bound.

Summarising, the model is going to be composed by three variables as the input, and one variable as the output. It should be pointed out that we want to predict the threshold for day $t+1$ given the information of the day t. Therefore, the matrix x of inputs is going to be composed by the variables $spectral\_analisis_t$, $spectral\_error_t$, $conditional\_volatility_t$. The vector y of the output variable is going to be determined by the variable $optimal\_threshold_{t+1}$.

4.5.1. Backpropagation Network

Rumelhart, Hinton and Williams (1986) [32] developed a neural network method which can learn the existing association between the inputs and the corresponding classes. This network works as follows: First, an input pattern is applied as a stimulus for the first layer of the neurons, and it is then propagated through all the upper layers until an output is generated. It compares the result of the output neurons with the output that is desired to obtain, and it calculates an error value for each output neuron. Then, these errors are transmitted backwards, starting from the output layer to all the neurons in the intermediate layer that contribute directly to the output. This process is repeated, layer by layer, until all the neurons in the network have received an error value which can describe their relative contribution to the total error. Regarding the error values, it can adjust the connexion weights between each neuron, so that it can better approximate the obtained output to the real desired output.

Backpropagation network can auto-adapt the weights of the intermediate layers to learn the existing relationship between a group of inputs and outputs. It must have the generalization capacity, in order to be able to make an optimal output for an input that it has never seen before during the training.

The structure of this network is composed by an input layer with n neurons and an output layer with m neurons, and at least one hidden of inner neurons. Each neuron of the layer receives an input from all the neurons of the previous layer (except from the first layer), and it sends its outputs to the next layer (except for the last layer).

We can divide the training process into two sections. First, the algorithm has a forward advance, and then it is tested in a backward way. During the forward phase, the inputs are propagated through the layers until it arrives to the last layer. Once the values of the network are obtained, it starts the backward testing, where the values are compared with the target values in order to calculate the error and adjust the weights, until it arrives to the first layer.

The steps to apply the Backpropagation method are the followings:

1. Generate small random weights (w).
2. Set the input pattern and the target outputs.
3. Calculate the output of the network:

The inputs are calculated for the hidden neurons from the input neurons. For a hidden neuron j:

$$net_{pj}^{h} = \sum_{i=1}^{N} w_{ij}^{h} x_{pi} + \theta_j^h$$

where h is the magnitude of the hidden layer, p is the $p-th$ training vector, j is the $j-th$ hidden neuron, with an optional θ which can be interpreted as another hidden layer.

Then the outputs for the hidden layers are calculated:

$$y_{pj} = f_j^h(net_{pj}^h)$$

where f is the activation function. In this method it is used a sigmoid function, which can be differentiable:

$$y = \frac{1}{1+e^{-z}}$$
$$\frac{dy}{dz} = y(1-y)$$

Each neuron $i-th$ is characterised by a numeric value called activation state $a_i(t)$. The activation function can transform the actual state $a_i(t)$ into a new activation state $a_t(t+1)$

For the output neurons, the calculus are similar:

$$net_{pk}^{o} = \sum_{j=1}^{L} w_{kj}^{o} y_{pj} + \theta_k^o$$

$$y_{pk} = f_k^o(net_{pk}^o)$$

4. Calculate the error terms (δ).

   Given that the error function is:

$$E_p = \frac{1}{2} \sum_{k=1}^{M} \delta_{pk}^2$$

   in (δ), the output (y) of each neuron is compared with the target output (d).

$$\delta_{pk} = \frac{\partial E}{\partial y_{pk}} \frac{\partial y_{pk}}{\partial net_{pk}}$$

If j is an output neuron:

$$\delta^o_{pk} = (d_{pk} - y_{pk}) y_{pk} (1 - y_{pk})$$

If j is an inner neuron:

$$\delta^h_{pj} = y_{pj}(1 - y_{pj}) \sum_{k=1}^{M} \delta^o_{pk} w^o_{kj}$$

The partial derivative of the error is:

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial net} \frac{\partial net}{\partial w} = \delta y$$

that we will consider later.

5. Weights actualising: it uses a backward algorithm, going from the output neurons to the input neurons, adjusting the weights as follows: For the output layers:

$$w^o_{kj}(t+1) = w^o_{kj}(t) + \Delta w^o_{kj}(t+1)$$

For the hidden layers:

$$w^h_{ji}(t+1) = w^h_{ij}(t) + \Delta w^h_{ji}(t+1)$$

where

$$\Delta = -\alpha \frac{\partial E}{\partial w} = -\alpha \delta y$$

with $\alpha > 0$ the chosen learning rate. In our case is set 0.01 as the default value for a sigmoid function.

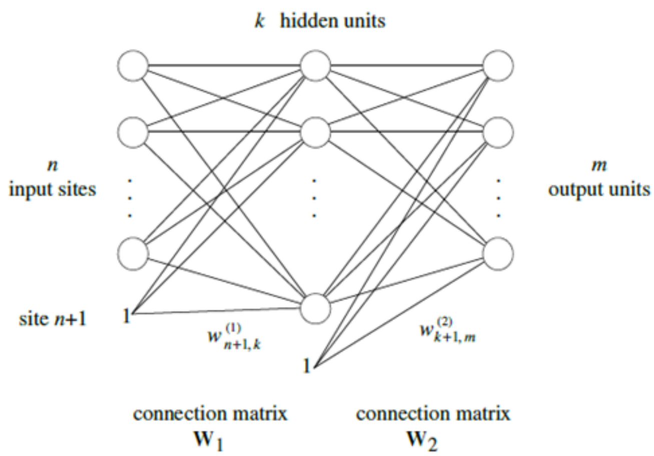
6. The process is repeated until the error is minimised.

4.5.2. Network Performance

In this subsection, we try to illustrate how the neural network works for the same example we have been studying in the previous sections.

In Figure 12 we can see, within an graphical explanation, how a general neural network works and connects its layers. The complexity of the neural network changes when more hidden layers are added.

Figure 12: Example of Neural Network Connections



Source: Rojas (1996) [31]

As we explained in the subsubsection 4.5.1, the applied methodology try to identify the point where the mean squared error is minimised, in order to assign the proper weights to the intermediate layers, to learn the existing relationship between inputs and outputs.

We can divide the process into three different steps: training, validation and testing period. The training period uses the 70 percent of the input data, to adjust the parameters. Then, in the validation process, it takes the 15 percent of the input data to to estimate how well the model has been trained. Finally, it takes out of the sample the other 15 percent of the input data, to test how the model is performing.

The figure below shows the best validation performance for the pair of stocks we are studying.

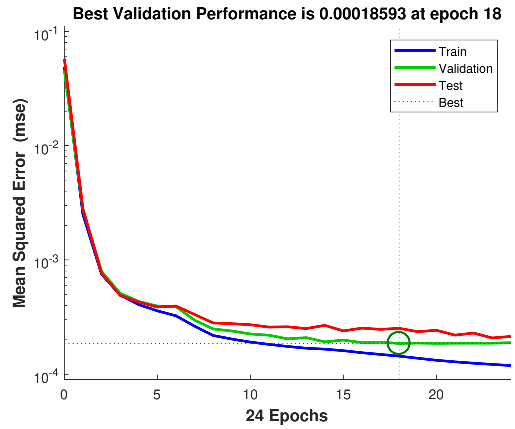


Figure 13: Validation performance

In this particular example, the model has identified the optimal validation performance at epoch 18, with a MSE of 1.8593e-04. [10]

Moreover, the coefficient of determination of the whole process (including train, validation and test) has been 0.96633. It means that the input variables have been able to predict the 96.633% of the target variable. Figure 14 displayed below, shows the regression of the model, including the whole process (train, validation and test).

[10]The error histogram is displayed in the Annex section (Figure 21).

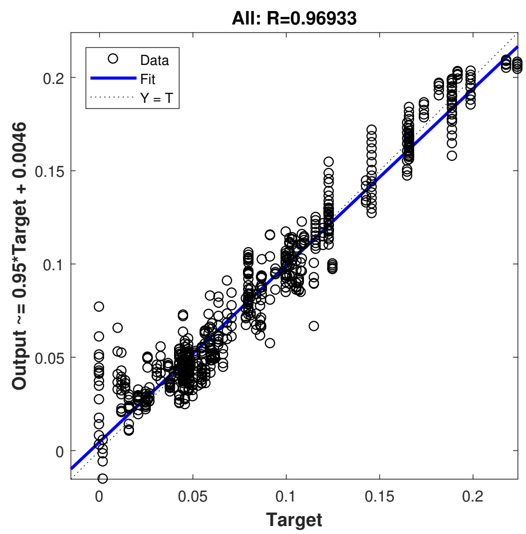


Figure 14: Coefficient of determination

During the training process, the neural network learn how to fit the input data. Hence, as we have introduced a previously computed optimal threshold as the output variable, the model try to fit the series from the input variables. The figure below shows how the optimal threshold has been adapted by an smoothed threshold series. The precision level of the fitting will depend on how accurate we want to be. Selecting more hidden layers could help us adjust a complex model, but here appears the overfitting problem. [11] The number of hidden layers has to be selected regarding the complexity of the model, in order avoid more computation time that necessary and the mentioned overfitting problem (Maciel and Ballini (2008)[27]).

The fitted series are displayed in the figure below, also showing the optimal threshold we have computed for the in sample data.

[11] After testing the results for different number of hidden layers, we have selected 5 hidden layers for every training set. The error histogram, displayed in Figure 21 in the annex, can also help us decide how accurate is the model.

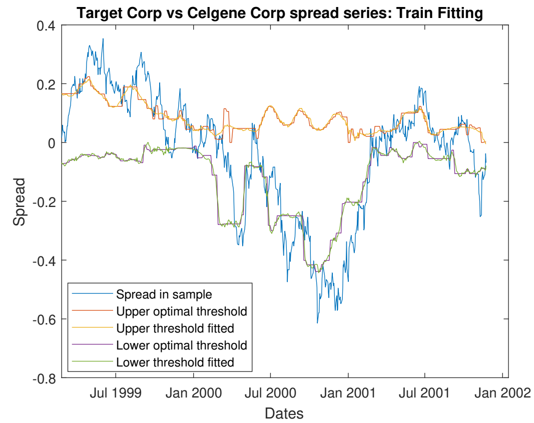


Figure 15: Training Fitting

Finally, the predicted thresholds are displayed in the Figure 15 for the spread series out of sample.

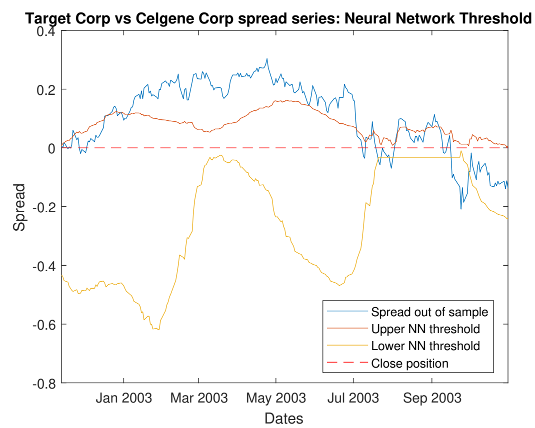


Figure 16: Threshold Prediction

In this example, as we did in the precedent sections, we should focus on the upper section of the sample. The predicted trading trigger has been able to identify the trend of the spread series, but not always reaching the highest points.

In the following section, we will test this strategy for other samples, so that we can determine if it is consistently adapting the trading threshold to the spread series.

5. Results and discussion

In this section we want to perform a robustness test for each strategy, in order to determine which of them has been consistently obtaining better results during different periods and for different pair of stocks. The main assumptions of this studies are that there are no transaction costs for any of the strategies, and the assets are perfectly divisible. [12]

We must be consistent with the formation period of 3 years we have taken for the pairs selection in the previous sections. Therefore, for each approach, there will be an in sample period of 3 years, and a testing period of 1 year out of sample. We have selected 4 different relevant periods, in order to test the robustness of the strategies during a non plane economic situation:

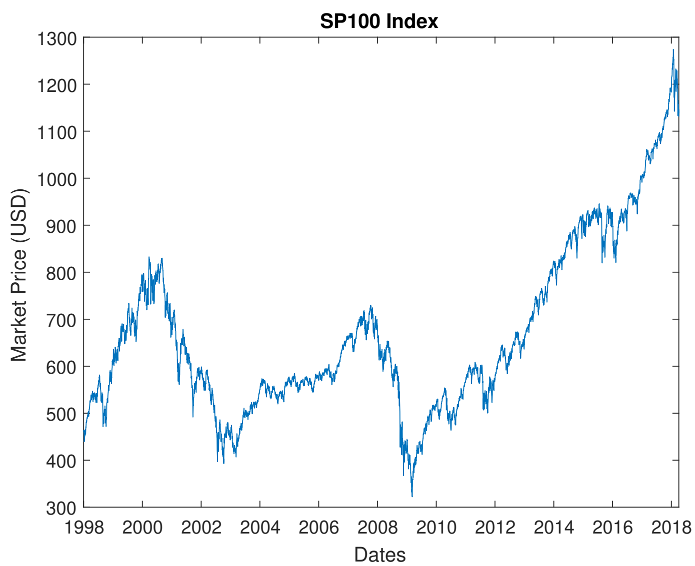


Figure 17: SP100 Index

Regarding the Figure 17 above, we have focused our analysis during the following

[12] To avoid this problem, we could adapt the number of shares to buy or sell to approximately replicate the established spread relationship of the stocks.

periods:

- 2000 Dot-com bubble.
- 2008 Leheman Brothers crisis.
- 2012 Growing period.
- 2009 Growing period.

For this robustness test, we have selected 5 different pair of stocks from the previously identified cointegrated pairs in section 3 [13]. Thus, the data for each strategy in the following tables correspond to the trade of 5 pair of stocks, during a testing period of 1 year [14]. In the annex we can see the results for each pair and period, as well as the detail of the trade positions for the pair containing TGT Corp. and CELG Corp. (Table 12).

The first column of the tables has the information corresponding to the log-return sum. We consider the profit sum as the most determinant result to decide which strategy works better. We also display the mean per operation, the median and the number of operations, to observe which strategy is more active. We can also examine the existing relationship between the total profit and the number of operations.

In the annex we can see an extended version of these tables, showing the principal statistics of the operations.

Table 1 displays the results for the period 2002 - 2003. The best performing strategy has been the Neural Network approach. The spectralis in the second position, registering the also the second highest number of operations. The selected thresh-

[13] Some of the strategies require an individual analysis for each pair of stocks, therefore, we use 5 of the 30 cointegrated pairs.

[14] We define as the Fixed Threshold the one calculated via unconditional volatility.

old via percentile have acceptable results, noting that with the same number of operations, the percentile 70^{th} can achieve higher profits. Conditional volatility is not getting good results, compared to the other strategies.

Table 1: Results of the different strategies for the period 2002 - 2003

	Profit Sum	**Mean**	**Median**	**Num operations**
NN	1,69	0,08	0,08	21
Spectrum	1,41	0,12	0,10	12
Percentile(70%)	1,20	0,15	0,16	8
Percentile(60%)	1,14	0,14	0,16	8
Fixed	1,13	0,13	0,13	9
Conditional	1,09	0,10	0,12	11
Percentile(80%)	0,72	0,18	0,20	4

Table 2 indicates that, during the out of sample period between 2009 and 2010, the spectral analysis haven been able to predict better than the other strategies. Conditional volatility threshold is ranked in the second position, improving the results from the previous period. Neural Networks is not obtaining the best profit results in this period.

Table 2: Results of the different strategies for the period 2009 - 2010

	Profit Sum	Mean	Median	Num operations
Spectrum	3,541	0,272	0,164	13
Conditional	3,361	0,134	0,090	25
NN	3,124	0,130	0,078	24
Fixed	2,827	0,283	0,154	10
Percentile(60%)	2,749	0,275	0,249	10
Percentile(70%)	2,460	0,273	0,174	9
Percentile(80%)	2,342	0,293	0,264	8

Table 3 shows that, once again, the trading trigger calculated using the machine learning technique, has obtained higher profits. The conditional volatility threshold and the spectral analysis are also obtaining good results. Regarding the inputs of the neural network, conditional volatility and spectral analysis could introduce important information to the model.

Table 3: Results of the different strategies for the period 2011 - 2012

	Profit Sum	Mean	Median	Num operations
NN	0,468	0,033	0,018	14
Conditional	0,447	0,041	0,050	11
Spectrum	0,290	0,041	0,019	7
Fixed	0,250	0,083	0,093	3
Percentile(60%)	0,195	0,065	0,072	3
Percentile(80%)	0,146	0,073	0,073	2
Percentile(70%)	0,113	0,056	0,056	2

Table 4 illustrates, as the previous tests, how the neural network threshold has achieved better results. The spectral analysis is losing importance in this period, unlike previous periods. Thresholds calculated via unconditional volatility has achieved good results, but the conditional volatility has not been able to make an optimal trade.

Table 4: Results of the different strategies for the period 2014 - 2015

	Profit Sum	Mean	Median	Num operations
NN	0,358	0,018	0,018	20
Fixed	0,346	0,058	0,083	6
Spectrum	0,278	0,031	0,025	9
Percentile(70%)	0,238	0,048	0,035	5
Percentile(80%)	0,231	0,058	0,043	4
Conditional	0,227	0,057	0,055	4
Percentile(60%)	0,219	0,044	0,034	5

Summarising, we can notice that the trading trigger calculated via Artificial Neural Networks, has been consistently performing better. Compared to the other applied strategies, this one has a higher number of operations, with a lower mean and median. Nevertheless, the sum of the obtained profits is significantly better. As we are no considering transaction costs, the number of operations just help the model achieve better results. Nonetheless, the model could be adapted to the transaction costs.

Besides, spectral analysis could be a good option for the threshold selection. It has shown acceptable results, compared to the other strategies.

Trading triggers calculated via conditional volatility has different results depending on the examined period. Hence, we can not conclude that it is consistently better. Nevertheless, in two of the analysed periods, the profit sum obtained via conditional volatility has been in the second place regarding the ranking tables.

Assuming the unconditional volatility as the simplest strategy, it has been ranked in the middle of the table during the different periods of time. Moreover, in Table 4, this fixed strategy is showing important results.

In the Figure 18, we can appreciate, for the particular pair we have been analysing in the previous sections, what is the differences between the unconditional and conditional volatility. [15]

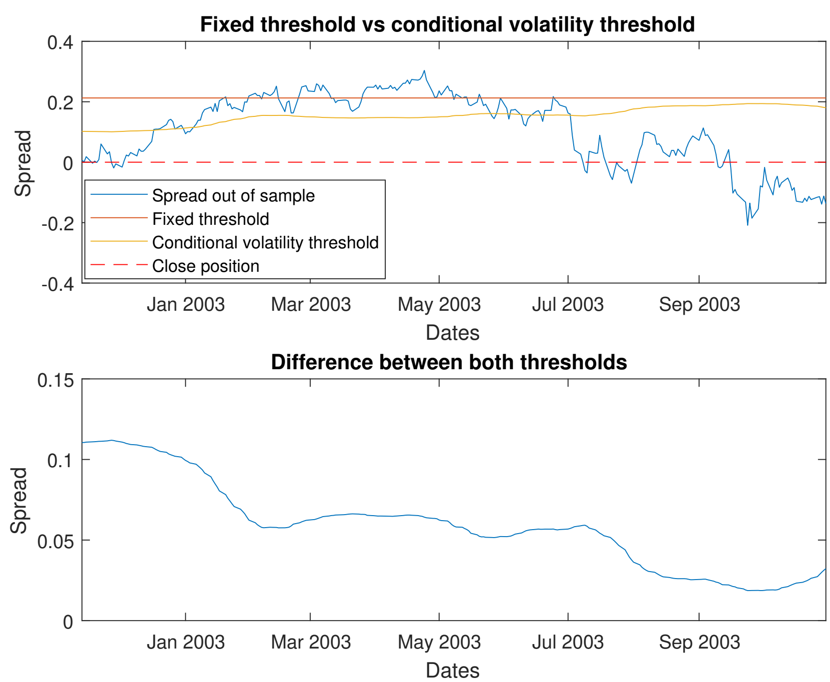


Figure 18: Comparison between unconditional and conditional volatility

We can see that both trading triggers differs at the beginning of the testing period. As expected, both threshold should not have a very high difference, since they

[15] Figures below are only displaying the upper part of the trading, in order to focus the graphical analysis on this part.

are calculated in a similar way.

Strategies calculated via percentile selection, have not achieve the expected results. Been too much accurate, separating positive and negative parts, could create a dependence from the historical data.

The Figure 19 below, exhibit, for the same pair of stocks as before, the differences with the other fixed threshold strategy: the unconditional volatility threshold.

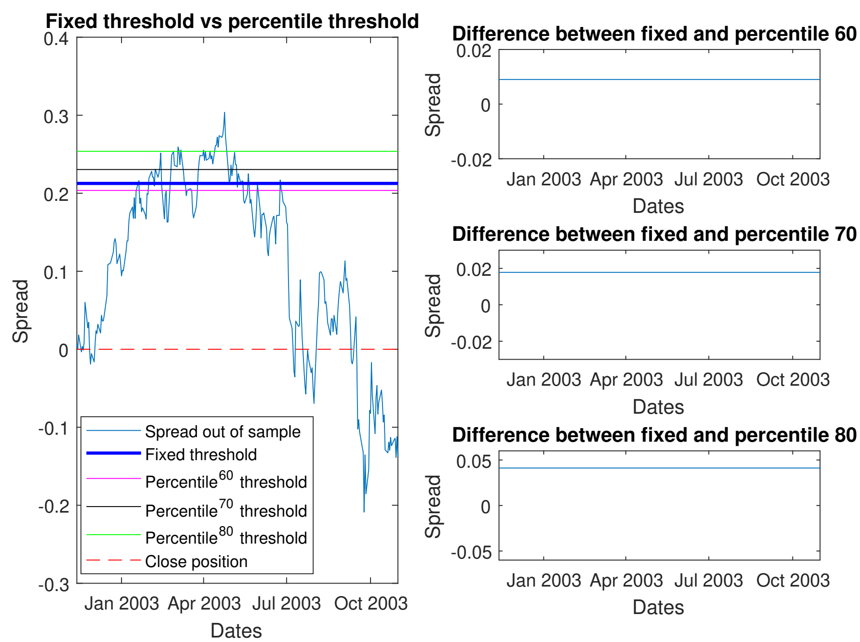


Figure 19: Comparison between fixed threshold and threshold by percentile

We can see that, as the spread series is reaching a high point, the trigger calculated from the percentile 80^{th} is obtaining better results. Notwithstanding, when the tested series are not able to reach such a high point observed during the historical data, the strategy fails.

The following figure presents the three dynamic strategies we have performed.

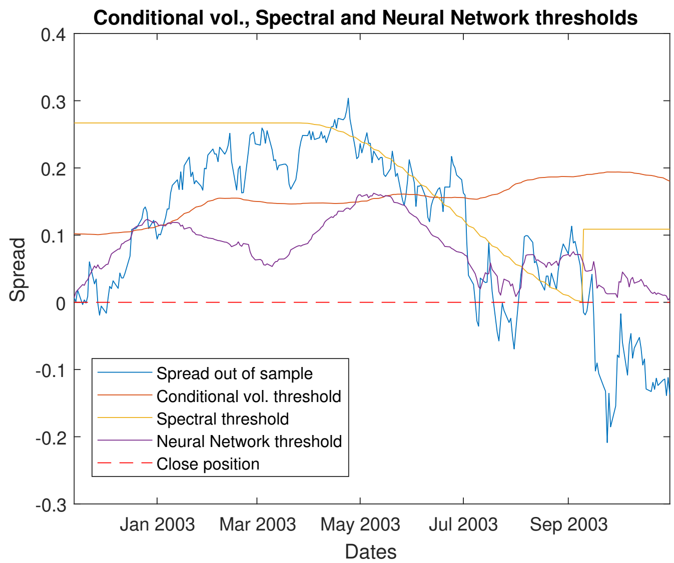


Figure 20: Threshold comparison between conditional volatility, spectral analysis and Neural Network

We can note that, whereas the spectral analysis is only focusing on the highest point a series could reach, the neural network try to identity the optimal points for every moment.

6. Conclusions

The aim of this work was to determine if our thresholds proposals are more efficient than two of the existing threshold selection strategies. We use highly cointegrated pairs, testing the proposed strategies during several periods and for different pair of stocks.

The first proposed strategy has been a non-dynamic trading trigger, whose selection was based on the asymmetric property of the spread time series. Classifying it by percentile levels, we determine three different threshold configurations. The result has been worse than expected, achieving just one time better results than the traditional strategies.
The second proposed method has been a threshold based on the spectral analysis of the spread series. The results have been positive, as the profit sum of the operations under this strategy has beaten the traditional methods in four of the five presented scenarios.
The third strategy, containing a Neural Network prediction model, has outperform the other strategies (traditional and proposed) during three of the four analysed periods.

In consequence, regarding the results, we could state that the Neural Network could help us develop an optimal trading threshold. Taking into account that we have not include operating costs in the analysis, we could consider training the model to operate less times, reaching higher entry points. The percentile based threshold strategy has not been able to obtain better results than the traditional strategies. Hence, we think incorporating the last observations into the model, could help obtain better results, changing from a fixed threshold to a dynamic one.

7. Bibliography

[1] Alexander, C. O., *Market models: a guide to financial data analysis.* John Wiley & Sons, Chichester. 2001

[2] Alexander, C.O. and Dimitriu, A., *The Cointegration alpha: Enhanced Index Tracking and long-short equity market neutral strategies.* Discussion papers in Finance. ISMA Centre, University of Reading. UK. 2002.

[3] Alexander, C.O. and Dimitriu, A., *Indexing and statistical Arbitrage: Tracking error or cointegration?*. Journal of Portfolio Management, 32, pp. 50-63. 2005.

[4] Bertoneche, M. L., *Spectral Analysis of Stock Market Prices*, Journal of Banking and Finance. 1979.

[5] Chen, H., Chen, S.J., and Li, F. *Empirical investigation of an equity pairs trading strategy.* SSRN Electronic Journal. 2012.

[6] Desai, J., Trivedi, A., and Joshi, N., *The Case of Gold and Silver: A New Algorithm for Pairs Trading*, Shri Chimanbhai Patel Institutes, Ahmadabad. 2012.

[7] Do, B., Faff, R. and Hamza, K., *A New Approach to Modelling and Estimation for Pairs Trading.* Monash University. 2006.

[8] Do, B., and Faff, R., *Does naïve pairs trading still work.* Working Paper, Monash University. 2009.

[9] Do, B., and Faff, R., *Are Pairs Trading Profits Robust to Trading Costs?*, Journal of Financial Research, 35(2), pp. 261-287. 2012.

[10] Dwyer, G. P.,*The Johansen Tests for Cointegration*, White Paper. 2015.

[11] Elliott, R. J., Van Der Hoek, J. and Malcolm, W. P. *Pairs Trading*, Quantitative Finance 5(3), 271–276. 2005.

[12] Engle, R. F. and Granger, C. W., *Cointegration and Error Correction: Representation, Estimation, and Testing*, Econometrica: journal of the Econometric Society pp. 251–276. 1987.

[13] Engelberg, J., Gao, P. and Jagannathan, R., *An anatomy of Pairs Trading: The Role of Idiosyncratic News, Common Information and Liquidity*. University Of Notre Dame. 2008.

[14] Faber, M.T., *A Quantitative Approach to Tactical Asset Allocation*. The Journal of Wealth Management. 2007 (2013, update).

[15] Figuerola-Ferretti, I., Serrano, P., Tang, T., and Vaello, A., *Supercointegrated*, Working paper. 2017.

[16] Gatev, E., Goetzmann, W. and Rouwenhorst, K. G., *Pairs Trading: Performance of a Relative Value Arbitrage Rule*. Working papers, Yale School of Management. 1999.

[17] Gatev, E., Goetzmann,W. and Rouwenhorst, K. *Pairs Trading: Performance of a Relative-Value Arbitrage Rule*, Review of Financial Studies 19(3), 797–827. 2006.

[18] Granger, C. W. J., and Morgenstern, O.*Spectral Analysis of New Your Stock Market Prices*, Kyklos. 1963.

[19] Hamilton, J. D., *Time Series Analysis* Princeton University Press. 1994.

[20] Herlemont, D., *Pairs Trading, Convergence Trading, Cointegration*, YATS Finances and Technology, vol. 33, p. 1-31. 2003.

[21] Hecht-Nielsen, R., *Theory of the Backpropagation Neural Network* Neural Networks for Perception Computation, Learning, and Architectures pp. 65–93. 1992.

[22] Huafeng, C., Shaojun, C. and Feng, L., *Empirical Investigation of an Equity Pairs Trading Strategy*. Working paper. 2012.

[23] Huck, N. *Pairs trading: does volatility timing matter?*. Applied Economics, pages 1–18. (2015).

[24] Iacobucci, A., *Spectral Analysis for Economic Time Series*, Observatoire Francais des Conjonctures Economiques (OFCE). 2003.

[25] Johansen, S., *Statistical Analysis of Cointegration Vectors*,Journal of Economic Dynamics and Control 12, pp. 231-254. North-Holland. 1988

[26] Johansen, S., *Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models*, Econometrica: Journal of the Econometric Society. pp. 1551–1580. 1991.

[27] Maciel, L. S., Ballini, R.,*Design a neural network for time series financial forecasting: accuracy and robustness analysis.* Anales do 9º Encontro Brasileiro de Finanças, Sao Pablo, Brazil. 2008.

[28] Nath, P., *High Frequency Pairs Trading with U.S Treasury Securities: Risks and Rewards for Hedge Funds.* Working paper, London Business School. 2003.

[29] Pizzutilo, F., A Note on the Effectiveness of Pairs Trading For Individual Investors. International Journal of Economics and Financial Issues, Vol. 3, No. 3, pp.763-771. Bari, Italy. 2013.

[30] Rampertshammer, S., *An Ornstein-Uhlenbeck Framework for Pairs Trading*,Preprint. Available at http://www. ms. unimelb. edu. au/publications/RampertshammerStefan. pdf 2007

[31] Rojas R., *The Backpropagation Algorithm*, Neural Networks. Springer, Berlin, Heidelberg. 1996.

[32] Rumelhart D. E., Hinton G. E., Williams R. J., *Learning representations by back-propagating errors*, Nature Publishing Group. 1986.

[33] Specht, D. F., *A General Regression Neural Network*, IEEE Transactions on Neural Networks, Vol. 2. 1991.

[34] Vidyamurthy, G., *Pairs Trading: Quantitative Methods and Analysis*, Vol. 217, John Wiley & Sons. 2004.

8. Annex

Table 5: Cointegrated pairs by Johansen Test

Stock 1	Stock 2
C UN Equity	AIG UN Equity
QCOM UW Equity	AMZN UW Equity
WMT UN Equity	BA UN Equity
WMT UN Equity	BIIB UW Equity
TWX UN Equity	BMY UN Equity
VZ UN Equity	BMY UN Equity
LMT UN Equity	BRK/B UN Equity
MO UN Equity	BRK/B UN Equity
NEE UN Equity	BRK/B UN Equity
TGT UN Equity	CELG UW Equity
WMT UN Equity	CELG UW Equity
COST UW Equity	CL UN Equity
UNP UN Equity	CL UN Equity
WMT UN Equity	CL UN Equity
DWDP UN Equity	COF UN Equity
UTX UN Equity	DHR UN Equity
WFC UN Equity	FDX UN Equity
WFC UN Equity	GD UN Equity
TGT UN Equity	GILD UW Equity
WFC UN Equity	JNJ UN Equity
PFE UN Equity	LLY UN Equity
MMM UN Equity	LOW UN Equity
WFC UN Equity	LOW UN Equity
WFC UN Equity	MMM UN Equity
TWX UN Equity	QCOM UW Equity
WMT UN Equity	QCOM UW Equity
TGT UN Equity	SBUX UW Equity
USB UN Equity	SBUX UW Equity
WFC UN Equity	SBUX UW Equity
TGT UN Equity	SPG UN Equity
VZ UN Equity	TWX UN Equity

Table 6: Cointegrated pairs by Engle-Granger Test

Stock 1	Stock 2
C UN Equity	AIG UN Equity
WMT UN Equity	BA UN Equity
WMT UN Equity	BIIB UW Equity
TWX UN Equity	BMY UN Equity
VZ UN Equity	BMY UN Equity
LMT UN Equity	BRK/B UN Equity
MO UN Equity	BRK/B UN Equity
NEE UN Equity	BRK/B UN Equity
TGT UN Equity	CELG UW Equity
WMT UN Equity	CELG UW Equity
COST UW Equity	CL UN Equity
UNP UN Equity	CL UN Equity
WMT UN Equity	CL UN Equity
DWDP UN Equity	COF UN Equity
UTX UN Equity	DHR UN Equity
WFC UN Equity	FDX UN Equity
WFC UN Equity	GD UN Equity
TGT UN Equity	GILD UW Equity
WFC UN Equity	JNJ UN Equity
PFE UN Equity	LLY UN Equity
MMM UN Equity	LOW UN Equity
WFC UN Equity	LOW UN Equity
WFC UN Equity	MMM UN Equity
TGT UN Equity	SBUX UW Equity
USB UN Equity	SBUX UW Equity
WFC UN Equity	SBUX UW Equity
TGT UN Equity	SPG UN Equity
VZ UN Equity	TWX UN Equity
WFC UN Equity	UNH UN Equity
WMT UN Equity	UNP UN Equity

Table 7: Selected pairs

Stock 1	Stock 2
C UN Equity	AIG UN Equity
WMT UN Equity	BA UN Equity
WMT UN Equity	BIIB UW Equity
TGT UN Equity	CELG UW Equity
VZ UN Equity	BMY UN Equity

Error histogram in the Neural Network model, for the upper bound of the pair *Target Corporation* and *Celgene Corporation*, during the period 2002 - 2003:

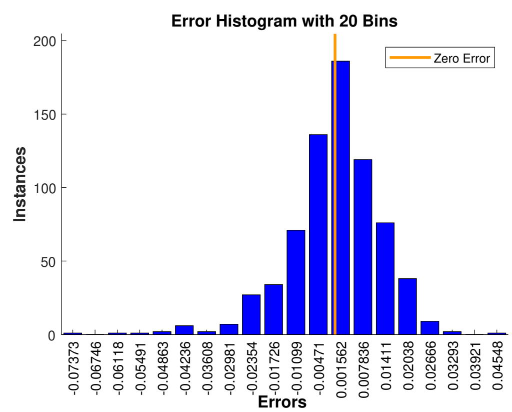


Figure 21: Error histogram in the Neural Network model

Table 8: Period 2002 - 2003

	Profit Sum	**Mean**	**Median**	**Standard Deviation**	**Skewness**	**Kurtosis**	**Max**	**Min**	**Num operations**
NN	1,69	0,08	0,08	0,07	-0,35	2,85	0,21	-0,09	21
Spectrum	1,41	0,12	0,10	0,07	0,99	3,31	0,29	0,03	12
Percentile(70%)	1,20	0,15	0,16	0,09	-0,22	1,63	0,26	0,02	8
Percentile(60%)	1,14	0,14	0,16	0,08	-0,23	1,55	0,24	0,02	8
Fixed	1,13	0,13	0,13	0,11	-0,86	2,57	0,24	-0,09	9
Conditional	1,09	0,10	0,12	0,09	-0,88	2,94	0,21	-0,09	11
Percentile(80%)	0,72	0,18	0,20	0,10	-0,54	2,04	0,29	0,04	4

Table 9: Period 2009 - 2010

	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	**Num operations**
Spectrum	3,541	0,272	0,164	0,260	0,841	2,600	0,832	-0,026	13
Conditional	3,361	0,134	0,090	0,145	1,739	5,863	0,606	-0,050	25
NN	3,124	0,130	0,078	0,157	1,420	4,761	0,606	-0,051	24
Fixed	2,827	0,283	0,154	0,237	0,379	1,503	0,606	-0,022	10
Percentile(60%)	2,749	0,275	0,249	0,182	0,364	2,167	0,606	0,011	10
Percentile(70%)	2,460	0,273	0,174	0,188	0,495	2,041	0,606	0,030	9
Percentile(80%)	2,342	0,293	0,264	0,192	0,271	1,945	0,606	0,030	8

Table 10: Period 2011 - 2012

	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
NN	0,468	0,033	0,018	0,059	0,628	3,625	0,162	-0,078	14
Conditional	0,447	0,041	0,050	0,039	-1,148	3,248	0,083	-0,041	11
Spectrum	0,290	0,041	0,019	0,054	1,309	3,417	0,151	-0,008	7
Fixed	0,250	0,083	0,093	0,057	-0,300	1,500	0,135	0,022	3
Percentile(60%)	0,195	0,065	0,072	0,032	-0,383	1,500	0,093	0,031	3
Percentile(80%)	0,146	0,073	0,073	0,001	0,000	1,000	0,074	0,072	2
Percentile(70%)	0,113	0,056	0,056	0,022	0,000	1,000	0,072	0,040	2

Table 11: Period 2014 - 2015

	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
NN	0,358	0,018	0,018	0,041	0,790	5,359	0,136	-0,066	20
Fixed	0,346	0,058	0,083	0,052	-1,322	3,173	0,092	-0,041	6
Spectrum	0,278	0,031	0,025	0,039	-0,049	2,875	0,092	-0,041	9
Percentile(70%)	0,238	0,048	0,035	0,031	1,452	3,185	0,103	0,030	5
Percentile(80%)	0,231	0,058	0,043	0,039	0,958	2,152	0,115	0,030	4
Conditional	0,227	0,057	0,055	0,049	0,065	1,442	0,113	0,003	4
Percentile(60%)	0,219	0,044	0,034	0,024	1,424	3,143	0,086	0,030	5

Table 12: Trading Positions for TGT and CELG pair - Period 1

Strategy	*Open Position*	*Close Position*	*Hour*	*Opening TGT*	*Opening CELG*	*Closing TGT*	*Closing CELG*	*Stock to buy when opening*
Fixed	20-Jan-2003	09-Jul-2003	Closing price	3,79 €	1,23 €	3,44 €	0,61 €	TGT
Conditional	17-Dec-2002	09-Jul-2003	Closing price	3,72 €	1,42 €	3,44 €	0,61 €	TGT
	24-Sep-2003	31-Oct-2003	Closing price	3,26 €	0,66 €	3,40 €	1,03 €	CELG
Fourier	14-Apr-2003	09-Jul-2003	Closing price	3,78 €	0,91 €	3,44 €	0,61 €	TGT
	05-Aug-2003	10-Sep-2003	Closing price	3,57 €	0,88 €	3,46 €	0,70 €	TGT
NN	15-Nov-2002	18-Nov-2002	Closing price	3,64 €	1,48 €	3,62 €	1,49 €	TGT
	22-Nov-2002	27-Nov-2002	Closing price	3,68 €	1,46 €	3,60 €	1,38 €	TGT
	17-Dec-2002	09-Jul-2003	Closing price	3,72 €	1,42 €	3,44 €	0,61 €	TGT
	11-Jul-2003	21-Jul-2003	Closing price	3,50 €	0,61 €	3,45 €	0,77 €	TGT
	22-Jul-2003	04-Aug-2003	Closing price	3,43 €	0,77 €	3,56 €	0,90 €	CELG
	05-Aug-2003	10-Sep-2003	Closing price	3,57 €	0,88 €	3,46 €	0,70 €	TGT
	17-Sep-2003	31-Oct-2003	Closing price	3,39 €	0,74 €	3,40 €	1,03 €	CELG
Percentile60	20-Jan-2003	09-Jul-2003	Closing price	3,79 €	1,23 €	3,44 €	0,61 €	TGT
Percentile70	07-Feb-2003	09-Jul-2003	Closing price	3,81 €	1,22 €	3,44 €	0,61 €	TGT
Percentile80	04-Mar-2003	09-Jul-2003	Closing price	3,81 €	1,07 €	3,44 €	0,61 €	TGT

Table 13: Results per strategy - Period 1

FIXED	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	-0,009	-0,009	-0,009	0,000	-	-	-0,009	-0,009	1
Pair 2	0,110	0,055	0,055	0,200	0,000	1,000	0,197	-0,086	2
Pair 3	0,210	0,210	0,210	0,000	-	-	0,210	0,210	1
Pair 4	0,244	0,244	0,244	0,000	-	-	0,244	0,244	1
Pair 5	0,570	0,143	0,123	0,045	1,109	2,296	0,209	0,115	4

CONDITIONAL	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,083	0,041	0,041	0,101	0,000	1,000	0,113	-0,030	2
Pair 2	0,110	0,055	0,055	0,200	0,000	1,000	0,197	-0,086	2
Pair 3	0,160	0,160	0,160	0,000	-	-	0,160	0,160	1
Pair 4	0,208	0,104	0,104	0,046	0,000	1,000	0,136	0,072	2
Pair 5	0,533	0,133	0,123	0,055	0,614	2,071	0,209	0,078	4

SPECTRUM	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,333	0,111	0,076	0,066	0,701	1,500	0,187	0,070	3
Pair 2	0,238	0,119	0,119	0,110	0,000	1,000	0,197	0,041	2
Pair 3	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 4	0,362	0,181	0,181	0,153	0,000	1,000	0,289	0,073	2
Pair 5	0,478	0,096	0,115	0,041	-1,160	2,736	0,127	0,026	5

NN	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,093	0,031	0,052	0,053	-0,620	1,500	0,070	-0,030	3
Pair 2	0,231	0,058	0,060	0,117	-0,072	1,930	0,197	-0,086	4
Pair 3	0,348	0,116	0,129	0,065	-0,350	1,500	0,173	0,046	3
Pair 4	0,438	0,063	0,073	0,055	-0,573	2,581	0,136	-0,035	7
Pair 5	0,581	0,145	0,127	0,043	1,125	2,313	0,209	0,119	4

Table 14: Results per strategy - Period 1

PERCENTILE 60%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,325	0,108	0,070	0,068	0,706	1,500	0,187	0,068	3
Pair 2	0,235	0,117	0,117	0,142	0,000	1,000	0,218	0,017	2
Pair 3	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 4	0,244	0,244	0,244	0,000	-	-	0,244	0,244	1
Pair 5	0,335	0,168	0,168	0,059	0,000	1,000	0,209	0,126	2

PERCENTILE 70%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,333	0,111	0,076	0,066	0,701	1,500	0,187	0,070	3
Pair 2	0,258	0,129	0,129	0,159	0,000	1,000	0,241	0,017	2
Pair 3	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 4	0,259	0,259	0,259	0,000	-	-	0,259	0,259	1
Pair 5	0,350	0,175	0,175	0,048	0,000	1,000	0,209	0,141	2

PERCENTILE 80%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,187	0,187	0,187	0,000	-	-	0,187	0,187	1
Pair 2	0,041	0,041	0,041	0,000	-	-	0,041	0,041	1
Pair 3	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 4	0,287	0,287	0,287	0,000	-	-	0,287	0,287	1
Pair 5	0,209	0,209	0,209	0,000	-	-	0,209	0,209	1

Table 15: Results per strategy - Period 2

FIXED	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	1,766	0,589	0,583	0,015	0,615	1,500	0,606	0,577	3
Pair 2	0,307	0,154	0,154	0,014	0,000	1,000	0,164	0,143	2
Pair 3	0,226	0,113	0,113	0,017	0,000	1,000	0,125	0,101	2
Pair 4	0,414	0,414	0,414	0,000	-	-	0,414	0,414	1
Pair 5	0,114	0,057	0,057	0,112	0,000	1,000	0,136	-0,022	2

CONDITIONAL	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	2,221	0,171	0,099	0,166	1,442	4,589	0,606	-0,019	13
Pair 2	0,443	0,074	0,056	0,045	1,563	3,792	0,164	0,040	6
Pair 3	0,216	0,072	0,073	0,024	-0,092	1,500	0,095	0,048	3
Pair 4	0,414	0,414	0,414	0,000	-	-	0,414	0,414	1
Pair 5	0,067	0,033	0,033	0,117	0,000	1,000	0,117	-0,050	2

SPECTRUM	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	2,582	0,516	0,577	0,228	0,215	1,774	0,832	0,265	5
Pair 2	0,238	0,119	0,119	0,064	0,000	1,000	0,164	0,074	2
Pair 3	0,095	0,095	0,095	0,000	-	-	0,095	0,095	1
Pair 4	0,414	0,414	0,414	0,000	-	-	0,414	0,414	1
Pair 5	0,212	0,053	0,052	0,067	0,080	1,935	0,136	-0,026	4

NN	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	2,180	0,218	0,189	0,169	1,174	3,783	0,606	0,039	10
Pair 2	0,218	0,036	0,022	0,069	1,026	3,178	0,164	-0,042	6
Pair 3	0,195	0,049	0,045	0,061	0,201	1,718	0,125	-0,019	4
Pair 4	0,414	0,414	0,414	0,000	-	-	0,414	0,414	1
Pair 5	0,117	0,039	0,051	0,084	-0,254	1,500	0,117	-0,051	3

Table 16: Results per strategy - Period 2

PERCENTILE 60%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	1,718	0,430	0,394	0,132	0,572	1,725	0,606	0,324	4
Pair 2	0,307	0,154	0,154	0,014	0,000	1,000	0,164	0,143	2
Pair 3	0,125	0,125	0,125	0,000	-	-	0,125	0,125	1
Pair 4	0,414	0,414	0,414	0,000	-	-	0,414	0,414	1
Pair 5	0,185	0,092	0,092	0,115	0,000	1,000	0,174	0,011	2

PERCENTILE 70%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	1,394	0,465	0,456	0,137	0,119	1,500	0,606	0,332	3
Pair 2	0,307	0,154	0,154	0,014	0,000	1,000	0,164	0,143	2
Pair 3	0,141	0,141	0,141	0,000	-	-	0,141	0,141	1
Pair 4	0,414	0,414	0,414	0,000	-	-	0,414	0,414	1
Pair 5	0,204	0,102	0,102	0,101	0,000	1,000	0,174	0,030	2

PERCENTILE 80%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	1,394	0,465	0,456	0,137	0,119	1,500	0,606	0,332	3
Pair 2	0,307	0,154	0,154	0,014	0,000	1,000	0,164	0,143	2
Pair 3	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 4	0,414	0,414	0,414	0,000	-	-	0,414	0,414	1
Pair 5	0,227	0,114	0,114	0,118	0,000	1,000	0,197	0,030	2

Table 17: Results per strategy - Period 3

FIXED	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 2	0,093	0,093	0,093	0,000	-	-	0,093	0,093	1
Pair 3	0,022	0,022	0,022	0,000	-	-	0,022	0,022	1
Pair 4	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 5	0,135	0,135	0,135	0,000	-	-	0,135	0,135	1

CONDITIONAL	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,065	0,065	0,065	0,000	-	-	0,065	0,065	1
Pair 2	-0,041	-0,041	-0,041	0,000	-	-	-0,041	-0,041	1
Pair 3	0,178	0,059	0,050	0,021	0,655	1,500	0,083	0,045	3
Pair 4	0,042	0,042	0,042	0,000	-	-	0,042	0,042	1
Pair 5	0,204	0,041	0,052	0,038	-0,889	2,622	0,080	-0,021	5

SPECTRUM	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,0121	0,0121	0,0121	0,0000	-	-	0,0121	0,0121	1
Pair 2	0,0000	0,0000	0,0000	0,0000	-	-	0,0000	0,0000	0
Pair 3	0,0227	0,0113	0,0113	0,0273	1,0000	0,0000	0,0306	-0,0079	2
Pair 4	0,0910	0,0455	0,0455	0,0376	1,0000	0,0000	0,0721	0,0189	2
Pair 5	0,1639	0,0820	0,0820	0,0971	1,0000	0,0000	0,1507	0,0133	2

NN	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,199	0,050	0,057	0,098	-0,270	2,016	0,162	-0,078	4
Pair 2	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 3	0,022	0,022	0,022	0,000	-	-	0,022	0,022	1
Pair 4	0,067	0,022	0,009	0,024	0,707	1,500	0,050	0,009	3
Pair 5	0,180	0,030	0,015	0,053	1,567	3,872	0,135	-0,014	6

Table 18: Results per strategy - Period 3

PERCENTILE 60%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 2	0,093	0,093	0,093	0,000	-	-	0,093	0,093	1
Pair 3	0,031	0,031	0,031	0,000	-	-	0,031	0,031	1
Pair 4	0,072	0,072	0,072	0,000	-	-	0,072	0,072	1
Pair 5	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0

PERCENTILE 70%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 2	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 3	0,040	0,040	0,040	0,000	-	-	0,040	0,040	1
Pair 4	0,072	0,072	0,072	0,000	-	-	0,072	0,072	1
Pair 5	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0

PERCENTILE 80%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 2	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 3	0,074	0,074	0,074	0,000	-	-	0,074	0,074	1
Pair 4	0,072	0,072	0,072	0,000	-	-	0,072	0,072	1
Pair 5	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0

Table 19: Results per strategy - Period 4

FIXED	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,092	0,092	0,092	0,000	-	-	0,092	0,092	1
Pair 2	-0,041	-0,041	-0,041	0,000	-	-	-0,041	-0,041	1
Pair 3	0,090	0,090	0,090	0,000	-	-	0,090	0,090	1
Pair 4	0,086	0,086	0,086	0,000	-	-	0,086	0,086	1
Pair 5	0,119	0,060	0,060	0,030	0,000	1,000	0,080	0,039	2

CONDITIONAL	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	-	-	0,000	0,000	0,000	0
Pair 2	0,083	0,042	0,042	0,000	1,000	0,054	0,080	0,003	2
Pair 3	0,031	0,031	0,031	-	-	0,000	0,031	0,031	1
Pair 4	0,000	0,000	0,000	-	-	0,000	0,000	0,000	0
Pair 5	0,113	0,113	0,113	-	-	0,000	0,113	0,113	1

SPECTRUM	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 2	-0,041	-0,041	-0,041	0,000	-	-	-0,041	-0,041	1
Pair 3	0,083	0,083	0,083	0,000	-	-	0,083	0,083	1
Pair 4	0,092	0,092	0,092	0,000	-	-	0,092	0,092	1
Pair 5	0,145	0,024	0,022	0,009	0,538	1,972	0,039	0,014	6

NN	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	-0,001	0,000	0,000	0,032	0,000	1,000	0,023	-0,023	2
Pair 2	-0,041	-0,041	-0,041	0,000	-	-	-0,041	-0,041	1
Pair 3	0,054	0,006	0,014	0,029	-1,892	5,644	0,034	-0,066	9
Pair 4	0,136	0,136	0,136	0,000	-	-	0,136	0,136	1
Pair 5	0,210	0,030	0,025	0,025	1,050	3,458	0,080	-0,001	7

Table 20: Results per strategy - Period 4

PERCENTILE 60%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 2	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 3	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 4	0,086	0,086	0,086	0,000	-	-	0,086	0,086	1
Pair 5	0,133	0,033	0,032	0,004	0,580	1,733	0,039	0,030	4

PERCENTILE 70%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 2	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 3	0,000	0,000	0,000	0,000	-	-	0,000	0,000	0
Pair 4	0,103	0,103	0,103	0,000	-	-	0,103	0,103	1
Pair 5	0,134	0,034	0,033	0,004	0,322	1,420	0,039	0,030	4

PERCENTILE 80%	Profit Sum	Mean	Median	Standard Deviation	Skewness	Kurtosis	Max	Min	Num operations
Pair 1	0,0000	0,0000	0,0000	0,0000	-	-	0,0000	0,0000	0
Pair 2	0,0000	0,0000	0,0000	0,0000	-	-	0,0000	0,0000	0
Pair 3	0,0000	0,0000	0,0000	0,0000	-	-	0,0000	0,0000	0
Pair 4	0,1148	0,1148	0,1148	0,0000	-	-	0,1148	0,1148	1
Pair 5	0,1162	0,0387	0,0335	0,0119	0,6485	1,5000	0,0524	0,0303	3