IMPROVING PAIRS TRADING USING NEURAL NETWORK TECHNIQUES AND FUNDAMENTAL RATIOS

Jose Fernando Ospino López

Máster en Banca y Finanzas Cuantitativas

Trabajo de Investigación 019/017

Directores: Dr. Oscar Carchano Alcina

D. Luis Fernández Payeras

Universidad de Valencia

Universidad Complutense de Madrid

Universidad del País Vasco

Universidad de Castilla la Mancha

www.finanzascuantitativas.com

MASTER THESIS

IMPROVING PAIRS TRADING USING NEURAL NETWORK TECHNIQUES

AND FUNDAMENTAL RATIOS

Author: Jose Fernando Ospino López

Supervisor: Oscar Carchano Alcina

Second assessor: Luis Fernández Payeras

Master in Banking and Quantitative Finance

Interuniversity program: UV – UCM – UPV/EHU – UCLM

July 2019

Abstract

Pairs trading is a quantitative trading strategy consisting on identifying two stocks that historically move together and, using the assumption that their prices difference has mean-reverting properties, exploit the deviation from the mean by taking long – short position in the chosen pair to profit. Throughout the years, different approaches have been developed in order to exploit this strategy. However, there is little literature who looks whether the divergences in the prices are generated by poor company results, i.e. whether the deviation from the mean are product of bad (or good) fundamentals and are justified, or if they generate a new equilibrium point for the pair. In addition, since machine learning techniques are becoming more popular in finance, this work aims to analyze the performance of pairs trading strategy using neural network techniques applied to S&P 500 index components, selecting pairs of stocks from same industry and picking up the effects of the fundamental ratios in the pairs before taking a trade decision.

Keywords: Pairs trading, Trading strategy, Cointegration, Mean-reverting process, Neural network, Machine learning, Fundamental ratios.

Resumen

La negociación por pares es una estrategia de negociación cuantitativa que consiste en identificar dos acciones que históricamente se han movido de forma similar y, usando la asunción de que la diferencia de sus precios tiene la propiedad de revertir a una media, explotar las desviaciones de la media tomando posición larga-corta en el par seleccionado para generar ganancias. A lo largo de los años, diferentes metodologías se han desarrollado para explotar esta estrategia. Sin embargo, hay poca literatura que analice si las divergencias en los precios son generadas por resultados deficientes de la compañía, es decir, si las desviaciones de la media son producto de malos (o buenos) fundamentales y están justificados, o si han generado un nuevo punto de equilibrio para el par. Además, desde que las técnicas de aprendizaje automático se han vuelto más populares en finanzas, este trabajo pretende analizar el desempeño de una estrategia de negociación por pares usando técnicas de redes neuronales aplicadas a las acciones que componen el índice S&P 500, seleccionando pares de acciones que pertenezcan a la misma industria y recogiendo los efectos de los fundamentales de las empresas en el par antes de tomar decisiones de negociación.

Palabras claves: negociación por pares, estrategia de negociación, cointegración, procesos con reversión a la media, redes neuronales, aprendizaje automático, ratios fundamentales.

Contents

1 Introduction

Pairs trading is a quantitative trading strategy that exploits inefficient financial markets by selecting a pair of assets that are known to historically move together and have some sort of long-run relationship, in order to take advantage of possible deviations from the mean in the short run to profit. When there is a divergence, we have an overvalued stock and an undervalued stock. Therefore, an investor should take long position in the undervalued asset and short position in the overvalued asset until the equilibrium is reached again.

Even though this may sound as an intuitive approach, throughout the years, several methods have been used to improve the two main aspects of pairs trading: the pairs selection and the trading strategy. Recently, more research has become available on applying machine learning techniques to time series predictions. However, there is little literature about neural network techniques applied to pairs trading. Dunis et al. (2006a, 2006b, 2015)[11][9][10] are the main authors in this area, but their applications are limited to specific spreads, such as the gasoline crack spread or soybean-oil crush spread. Another author in this field is Van Der Have (2017)[35], but his research analyzes the performance of neural network in pairs trading applied to Exchange Traded Funds (ETFs). Therefore, one of our main contributions will be apply neural network techniques to the stock market, using Standard and Poors 500 index components to select pairs and predict each pair spread value.

Moreover, in the standard approaches of pairs trading, even though some literature proposes the use of APT models whose risk factors can take the form of Fama and French three factor model and use the fundamentals to model the pair spread, these works didn't compute any empirical results for test the performance of these methodologies and also use the fundamentals in a complete different way respect to what we want to do with them. Our main objective with the fundamental ratios is make an analysis of the divergences in the pairs spread and if they are justified according to the corporate results of the company, i.e., if a company stock price is going up (down) because the company improved (got worse) his benefits recently, or his debts are decreasing (increasing), etc. which not necessarily implies a long-run mean-reversion in the pair associated with the particular asset and it can be a bad trade. That's why our second main contribution will be build a neural network based on company's fundamental ratios to determine stock trend, allowing us to determine whether the divergences in the pairs spread are justified according to the company's results or not and it can be a trade opportunity to profit. This way will be proposed a pairs trading strategy adding fundamental ratios as trade

criterion and showing empirical results, something that we didn't found in previous literature about pairs trading.

Also, the strategy will be viable by considering trading costs, taking positions with stocks opening prices when a trading signal is detected and allowing the neural networks to predict the pair spread value and stocks trend to anticipate the market dynamic and use this as an advantage to trade.

2 Literature review

In this section, is introduced the main researches over pairs trading. Section 2.1 introduces one of the first pairs trading frameworks in the literature, known as the distance method proposed by Gatev et al. (1999)[18]. Section 2.2 discusses the cointegration method introduced by Vidyamurthy (2004)[36] used as pair selection method in this thesis. In section 2.3 we explain the stochastic spread method whose main author is Elliott et al. (2005)[12] and section 2.4 is about the stochastic residuals spread method proposed by Do et al. (2006)[8] as an extension of the methodology exposed in section 2.3. Section 2.5 introduces the machine learning methods applied to pairs trading, where the main author is Huck who has developed and published on this pairs trading methodology in Huck (2009)[21] and Huck (2010)[22], respectively. Finally, section 2.5.1 will be about neural network approaches applied directly to pairs spread, the main authors in this field are Dunis et al. (2006a, 2006b, 2015)[11][9][10] and Van Der Have (2017)[35].

2.1 Distance Method

The distance method, introduced by Gatev et al. (1999)[18] is one of the most popular and still one of the main methods used for pairs trading in practice, for its simplicity and its transparency. As Krauss (2015)[26] said, this method introduced pairs trading to the academic literature and established pairs trading as a true capital market anomaly.

Their strategy consists in compute a cumulative total return index P_{it} for all stocks from the CRSP[1] daily files, removing stocks that have one or more days with no trade to use only the relatively liquid stocks to facilitate the pairs formation. Then, select pairs by minimizing the sum of squared deviations between two cumulative total returns (for now on, lets denote price as the cumulative return index) and trade when the spread between the chosen stocks diverges more than two historical standard deviations from its mean. As Do et al. (2006)[8] point out, the main advantage of the distance method is that it is model-free, which makes the strategy easy to implement. Being model-free also means that the strategy is free from misspecification and mis-estimation. Due to these advantages, the distance method is still most used pairs trading strategy in practice.

---

[1] The CRSP (center for research in security prices) is a provider of historical stock market data designed for research and educational use.

Pair selection: Gatev et al. (1999)[18] found through interviews that this approach approximates the way how traders choose their pairs. First, they split up the sample and define a formation period and a trading period. In the formation period, the pairs are formed, and the historical mean and standard deviation of the spreads are calculated. During the trading period, the trading strategy is executed for the different selected pairs to generate returns that are further analyzed. Gatev et al. (1999)[18] normalize the prices of all the stocks to the first day of the formation period. A matching partner for each stock is found by minimizing the sum of squared distances between the two normalized price series, denoted by P_i and P_j for respectively stock i and j; and use the euclidean squared distance, leading to the average sum of squared distances, $\overline{SSD_{t,i,j}}$, being given as

$$\overline{SSD_{t,i,j}} = \frac{1}{T_f}\sum_{t=1}^{T_f} S_{t,i,j}^2 = V(S_{i,j}) + \left(\frac{1}{T_f}\sum_{t=1}^{T_f} S_{t,i,j}\right)^2 \tag{1}$$

where T_f is defined as the end of the formation period, $S_{t,i,j} = P_{t,i} - P_{t,j}$ is defined as the spread between the two assets at time t, and $V(S_{i,j})$ is the spread variance, that can be written as

$$V(S_{i,j}) = \frac{1}{T_f}\sum_{t=1}^{T_f} S_{t,i,j}^2 - \left(\frac{1}{T_f}\sum_{t=1}^{T_f} S_{t,i,j}\right)^2 \tag{2}$$

This way, the method identifies a matching partner, a certain stock j, for every stock i.

Trading strategy: Following practice, Gatev et al. (1999)[18] based his rules for opening and closing positions based on a standard deviation metric. They open a position in a pair when prices diverge by more than two historical standard deviations, estimated during the pairs formation period. As soon as the spread converges back to its mean, the positions unwind, and profit is made. In case the spread has not converged at the end of the trading period, profits and losses are calculated on the last day of the interval. As a result, the authors get a return of 5.98% over a six-month period with a portfolio of the five best pairs (pairs with the smallest historical distance measure). Including a conservative estimate of transaction costs, this return drops to 3.99%.

However, this method has a few drawbacks, as stated by Krauss (2015)[26]. First Gatev et al. (1999)[18] identify a paired stock for every stock. This implies that could be that one of stocks does not have proper paired stock in the asset universe, but is still paired with one, since the distance method is set up for each stock to have a partner. Therefore, it is likely that using this pair to trade with will result in a loss, because the stock and its partner are not a 'true' pair. Secondly, minimizing (1) results in the variance of spread being minimized, which results in less deviations from the mean and therefore fewer potential profits. Finally, the selection of the pairs

cannot be statistically motivated. It is not tested whether there exists a mean-reverting, long-run relationship between the chosen pairs. This leads to higher divergence risk, which will result in more unprofitable trades with pairs of which the spread does not converge back to the mean. Do and Faff (2010)[7], using an extension of the Gatev et al. (1999)[18] data and the same methodology, confirm that 32 percent of all identified pairs based on the distance metric actually do not converge. As Van Der Have (2017)[35] denotes, these drawbacks indicate that selecting pairs using the distance method is a rather suboptimal selection metric, motivating the choice for pair selection method based on a statistical relationship. Therefore, in this study will be proposed a pair selection method and subsequently more sophisticated trading strategies.

2.2 Cointegration approach

To solve the drawbacks of the distance approach, Vidyamurthy (2004)[36] introduced the cointegration method that it is used for pairs selection in practice and we are going to use it for our pairs trading strategies. This method consists on make the pairs selection based on a statistical relationship between two stocks, when the cointegrated instruments are expected to be mean-reverting and follow the same long-term trend, which is an important statistical characteristic for pairs.

The cointegration method applied by Vidyamurthy (2004)[36] is based on the cointegration relationship outlined by Engle & Granger (1987)[14], which says that two time series are cointegrated of order d whether 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.

As the authors are interested in mean-reverting property, they look for stationary pairs. Let's define p_i to be the natural logarithm of the price series P_i of asset i. Therefore, to test two assets for cointegration, Engle & Granger (1987)[14] provide a procedure consisting of two steps. Given that p_i and p_j are non-stationary, the parameters of the cointegration relation need to be estimated using Ordinary Least Squares (OLS), as follows:

$$p_{t,i} = \alpha + \beta \cdot p_{t,j} + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \qquad (3)$$

where α is the constant (for simplicity, assume $\alpha = 0$), β is the cointegration coefficient and ϵ_t are the residuals. We can rewrite (3) as:

$$\epsilon_t = p_{t,i} - \beta \cdot p_{t,j} \tag{4}$$

Note that ϵ_t is the time series spread $S_{i,j}$ between the two assets.

Thereafter, the computed residuals ϵ_t needs to be tested for stationarity, so they apply the well-known Augmented-Dickey-Fuller (ADF) test published by Dickey and Fuller (1979)[5]. Whether ADF result shows that the residuals of regression (3) are stationary, it can be concluded that there's no unit root, so assets i and j are cointegrated and can therefore be considered as a pair.

Unfortunately, as stated by Van Der Have (2016)[35], Vidyamurthy (2004)[36] does not provide any empirical results of the cointegration method. However, subsequent research does, using the trading rule proposed by Gatev et al. (1999)[18]. For example, Dunis et al. (2010)[31] apply the cointegration approach to high frequency data. They find that the cointegration method can give a good indication of the profitability of the pair in a high frequency setting.
Caldeira & Moura (2013)[3] use cointegration to select pairs on a Brazilian stock index, using as trading rule a two standard deviation threshold from the z-score defined as:

$$z_t = \frac{e_t - \mu_e}{\sigma_e}$$

where μ_e is the long-term spread mean, and σ_e is the long-term spread standard deviation. The trading signals used are:

$$Buy\ to\ open\ if\ z_t < -2.00$$

$$Sell\ to\ open\ if\ z_t > 2.00$$

The position is closed the z-score approaches zero again. They find excess returns of more than 16% per year for the identified pairs. This return includes transfer costs equal to 0.1% of brokerage fees for each share, plus 0.05% from slippage for each stock and 0.2% of rental cost for short positions. This shows that the cointegration approach can result in large profits. Furthermore, Huck & Afawubo (2015)[23] show that the cointegration approach can generate high, stable and robust returns of up to 5% per month, while the distance method is unable to provide significant returns over the same time period. This indicates that the cointegration method is a better pair selection method than the distance method. Additionally, this shows that employing the cointegration method for pairs selection can serve as a foundation in order to optimize the trading strategy.

2.3 The Stochastic Spread Method

The main author in this approach is Elliot et al. (2005)[12], modeling the pairs spread in state space at the price level. They describe the spread with a mean-reverting Gaussian Markov Chain, observed in Gaussian noise. They assume that the spread is driven by a latent state variable x_k follows a mean-reverting process, defined by:

$$x_{k+1} - x_k = (a - b \cdot x_k)\tau + \sigma \cdot \sqrt{\tau} \cdot \epsilon_{k+1} \tag{5}$$

where $\sigma \geq 0$, $b > 0$, $a \in \mathcal{R}^+$ and ϵ_k is iid Gaussian $N(0,1)$. This process reverts to $\mu = \frac{a}{b}$ with mean-reversion strength b. We can also write (5) as

$$x_{k+1} = A + B \cdot x_k + C \cdot \epsilon_{k+1} \tag{6}$$

with $A = a\tau \geq 0$, $0 < B = 1 - b\tau < 1$ and $C = \sigma\sqrt{\tau}$. In continuous time, it is possible the state process follows the well-known Ornstein-Uhlenbeck process:

$$dx_t = \rho(\mu - x_t)dt + \sigma \cdot dW_t \tag{7}$$

where dW_t is a standard Brownian motion defined on some probability space, and $\rho = b$ describes the speed of mean-reversion. The parameters are estimated applying the Kalman filter, and the observation process, i.e., the observed spread is defined by:

$$y_k = x_k + D \cdot \omega_k, \tag{8}$$

where ω_k are iid gaussian $N(0,1)$ and independent of the ϵ_k in (5) and $D > 0$. According to this model, a trade position is taken as follows:

$$Short\ position\ when\ y_k \geq \mu + c\left(\frac{\sigma}{\sqrt{2 \cdot \rho}}\right)$$

$$Long\ position\ when\ y_k \leq \mu - c\left(\frac{\sigma}{\sqrt{2 \cdot \rho}}\right)$$

where $\sigma = C/\sqrt{\tau}$, $\rho = b = (1 - B)/\tau$ and $c > 0$. The position is reversed at time T, which denotes the first passage time result for the Ornstein-Uhlenbeck process, and is given by:

$$T = \inf\left\{t \geq 0, x_t = \mu \middle| x_0 = \mu + \frac{c \cdot \sigma}{\sqrt{2k}}\right\} = \hat{t}k \tag{9}$$

where:

$$\hat{t} = \frac{1}{2}\ln\left[1 + \frac{1}{2}\left(\sqrt{(c^2 - 3)^2 + 4 \cdot c^2} + c^2 - 3\right)\right] \tag{10}$$

Following Do et al. (2006)[8], this approach has three key advantages: First, the model is fully tractable, meaning that its parameters can be estimated applying the Kalman Filter in a state space model. It is an optimal maximum likelihood estimator in terms of MMSE (Minimum Mean Square Error). Second, the continuous time model can be exploited for forecasting purposes. Critical questions about pairs trading such as the expected holding times and the expected returns can be explicitly answered, provided the fact that the spread really follows this rigid model. Third, the approach is fundamentally based on mean-reversion, which is key to pairs trading.

However, Do et al. (2006)[8] also criticize the model of Elliot et al. (2005)[12] and propose an extension of this, consisting on working with returns (difference of log prices) instead of the prices, that's because only then, the mean of the spread remains the same when the two stocks produce exactly the same return (except when they trade at similar price points).

2.4 The Stochastic Residual Spread Method

As we say before, Do et al. (2006)[8] criticize the methodology proposed by Elliot et al. (2005)[12]. Do et al. (2006)[8] state that the model restricts the long-run relationship, since the two assets chosen must provide the same return such that any departure from it will be expected to be corrected in the future. To overcome this issue, Do et al. (2006)[8] suggest the so-called stochastic residual spread method, which uses the mispricing on return level instead of price level, which means that the spread will be defined as the difference in returns instead of difference in prices. Also, this gives the possibility to adjust the spread for differences in exposure to risk factors.

They established x as the state of mispricing or residual spread, with respect to a given equilibrium relationship, whose dynamic is governed by a Vasicek process:

$$dx_t = \rho(\mu - x_t)dt + \sigma \cdot dB_t \tag{11}$$

The observed mispricing is:

$$y_t = G_t = x_t + \omega_t \tag{12}$$

Where $G(R_t^A, R_t^B, U_t)$ is the residual spread function, and U_t denotes some exogenous vector potentially present in formulating the equilibrium. These two equations constitute a state space model of relative mispricing, defined with respect to some equilibrium relationship between two assets. As set by Do et al. (2006)[8], the state of mispricing is not fully observed, rather it is observed up to some measurement noise. Under the measurement of noise, the residual spread function G must be estimated. They address the APT (Arbitrage Pricing Theory) of Ross(1976)[30], adapted to a pair of stocks to create a completely tractable model of mean-reverting relative pricing for two asset based on their returns. Being the APT for a pair of stocks A and B at time t written as:

$$R_t^A = R_t^B + \Gamma \cdot r_t^m + e_t \tag{13}$$

where r_t^m is the raw return on the ith factor. Thus, the residual spread equation is:

$$G_t = G(R_t^A, R_t^B, U_t) = e_t = R_t^A - R_t^B - \Gamma \cdot r_t^m \tag{14}$$

where the return R_t^i for asset i at time t is defined as:

$$R_t^i = \log\left(\frac{P_{t,i}}{P_{t-1,i}}\right),$$

and Γ is a vector of exposure differentials to risk factors r_t^m. This way, the mean-reverting process is adjusted for the differences in exposure to risk factors.

Similar to the stochastic spread method, the approach of Do et al. (2006)[8] has not been used on empirical data in the academic literature as well. Krauss (2015)[26] describes the return restriction Do et al. (2006)[8] propose as difficult, since it is rather odd to find two assets with identical returns. However, the idea is to have a spread that is mean-reverting, which is the case for two returns series and therefore, this method can be valuable for pairs trading. As described by Cummins & Bucca (2012)[4], one of the major limitations of the stochastic residual spread method is the fact that the Gaussian properties of the Ornstein-Uhlenbeck process is not in line with the stylized facts of financial data, such as the assumption of linearity or constant volatility the methods use. However, it can be argued that the analytic simplicity of the Ornstein-Uhlenbeck process offsets this limitation.

2.5 Machine learning approach

Huck (2009)[21] and Huck (2010)[22] based his framework on three steps: Forecasting, outranking and trading. In the forecasting step, he considered n potential stocks, so $\frac{n(n-1)}{2}$ combinations of pairs can be constructed.

Forecasting: Huck (2009)[21] propose the use of a decision matrix, created using multiple spreads and forecasts of each series defined conditionally to all possible bivariate information sets, to deal with this large volume of information, where there is no restrictions about the forecasting method: in the decision matrix, he just assume the use of a quantitative forecasting model, which is based on Elman neural networks proposed in Elman (1990)[13] for simplicity and efficiency reasons, and generates one week ahead return forecasts for each security i conditional to the past returns of securities i and j, with $i, j \in [1, \dots, n]$.

As stated by Huck (2009)[21], "*Elman neural network allows the neurons to depend not only on the input variables but also on their own lagged values. Elman networks, compared to feedforward networks, are specific to data that have a time dimension. It is a way of capturing memory in financial markets*". Therefore, "*a recurrent Elman type network, it is characterized by a dynamic structure where the hidden layer output feeds back into the hidden layer with a time delay*".

Outranking: Huck (2009)[21] uses a Multi-Criteria Decision Method (MCDM) called ELECTRE III. This method ranks a set of alternatives according to a set of criteria. In this particular case, the n stocks represent the alternatives as well as the criteria. The performance x_{ij} of each stock i relative to criterion j can be calculated as follows:

$$x_{ij} = \hat{x}_{|X_{i,t}, X_{j,t}}^{i,t+1} - \hat{x}_{|X_{i,t}, X_{j,t}}^{j,t+1}, \tag{15}$$

being $X_{i,t}$ the past returns of asset X_i until date t and $\hat{x}_{|X_{i,t}, X_{j,t}}^{i,t+1}$ the one period ahead return forecast of X_i conditionally to $X_{i,t}$ and $X_{j,t}$.

These performance values are collected in an antisymmetric $n \, x \, n$ matrix, where the rows correspond to the n alternatives and the columns to the n criteria. In each cell, we find the anticipated spread of stock i versus criterion j. Next, the difference in performance between two alternatives is then compared with three thresholds associated to each criterion j, the indifference threshold (q_j), the preference threshold (p_j) and the veto threshold (v_j) with $q_j \leq p_j \leq v_j$, and that allows an outranking of the pairs which helps to detect potentially under (top of the ranking) and overvalued (bottom of the ranking) stocks.

Trading: After the phases of forecasting and ranking, the top m stocks are bought ("undervalued stocks") whereas the bottom m stocks are sold short ("overvalued stocks") and assume equally-weighted portfolios, i.e., all positions/pairs have the same weight and are closed out after one week, a new ranking is created and new positions are then initiated, and so on. An empirical application on the S&P 100 constituents from 1992 to 2006, using 8 years forecasting period (416 weeks), and selecting 5 pairs (top 5 and bottom 5 stocks of the ranking) has a rate of success of 54% and more than 0.8 percent weekly excess returns.

However, it is important to notice that Huck (2009)[21] results did not include roundtrip costs, so these findings should be handled with care. Moreover, trading actions are simply triggered based on the position in the final ranking shown by ELECTRE III outranking method, and the Elman neural network is used for forecasting purpose only, using historical stocks returns as network inputs.

In our analysis it will be considered roundtrip costs in the trading strategy to show transparency in the results, the pairs will be selected using supercointegration method introduced by Figuerola-Ferretti et al. (2017)[17], neural network predictions will be used as trading actions trigger and stocks trend will be predicted by a neural network using fundamental ratios inputs.

As Krauss (2015)[26] states, the following are other relevant machine learning approaches, most of them in an experimental setup and with limited applications to only a few selected securities: Thomaidis et al. (2006)[34] propose an experimental statistical arbitrage system based on neural network GARCH models. Lin and Cao (2008) and Huang et al. (2015)[20] use genetic algorithms for pairs mining. Finally, Montana and Parrella (2009)[27] use an ensemble of support vector regressions to develop a pairs trading strategy for the iShares S&P 500 ETF.

2.5.1 Neural network approach

In the academic literature, the most relevant authors are Dunis et al. (2006a, 2006b, 2015)[11][9][10] and Van Der Have (2017)[35] applying neural networks directly to pairs trading. The first authors develop a trading framework applied to specific spreads, such as a gasoline crack and corn/ethanol crush spread. They apply to gasoline crack spread three neural network types to predict the spread, a multilayer perceptron (MLP), recurrent neural network (RNN) and high order neural network (HONN). Using threshold, asymmetric threshold and correlation trading filters, and evaluating all models by Calmar ratio, defined as follows:

$$Calmar\ Ratio = \frac{Return}{|MaxDD|}$$

where $return$ is the annualized return of trading model and $MaxDD$ is the maximum drawdown of the trading model, they conclude than the best model in terms of out-of-sample results is the HONN with threshold filter. However, a drawback of this approach is the high number of trades, leading to an extremely large effect of transaction costs.

On the other hand, they apply to corn/ethanol crush spread once again MLP and HONN, and these methods are gauged against a genetic programming algorithm (GPA). They find better results with the GPA strategy (33.92% annualized returns including transaction and leverage costs) using a leverage structure to trade with more leverage during low volatility periods and avoid trades during times of high volatility. In second place, the HONN model was also profitable, with 33.01% annualized returns including costs.

It's important to mention that each network prediction used by Dunis et al. (2006a, 2006b, 2015)[11][9][10] are computed as the average estimation of ten networks, because the starting point for each network is a set of random weights. As stated by the authors, this helps to overcome the problem of local minima[2] affecting the training procedure. Therefore, the average result of all ten neural network models was used as the forecast of the change in the spread for each network type.

Moreover, Van Der Have (2017)[35] develop a framework applied to ETFs listed on the NYSE Arca. He applies feedforward neural network (FNN) and one recurrent neural network type called LSTM (Long short-term memory). The motivation to use LSTM instead of standard RNN is capture the long-term behavior, since this is important for pairs trading by the mean-reverting property, and the EBP algorithm (Error Backpropagation) for RNN results in the vanishing or exploding gradient problem. To solve this issue, Hochreiter & Schmidhuber (1997)[19] propose a special kind of RNN, named LSTM network whose dynamic in the hidden layer of the LSTM change in order to overcome vanishing gradient problem. The RNN and LSTM will be explained in depth later in section 4.2.1 and 4.2.2 respectively, and the LSTM network will be used to predict the pairs spread in this work.

In his work, Van Der Have (2017)[35] select pairs per sector based on their cointegration relation, i.e., he apply the cointegration approach explained in section 2.2. With cointegrated pairs, the

---

[2] Local minima are suboptimal equilibrium points and it's not the global optimum cost function value. Local minima points can be reached according to initial cost function value given by network training starting weights.

performance of the distance method, stochastic spread method, stochastic residual spread method, feedforward neural network and recurrent neural network are evaluated in terms of returns and economic measures. The author concludes that RNN outperform the traditional methods and the FNN in terms of return, sharpe ratio and sortino ratio, and can consistently generate a two years total return of around 11%.

However, this author doesn´t make assumptions on the amount of trading costs per trade, instead, the author calculates a break-even giving an indication for investors whether it would be profitable after transaction costs to use the strategy, according to their personal trading costs. Even though the RNN has less trades than the other strategies (except distance method), are 52 trades and the inclusion of real transfer costs may cause an important drop in the return shown by the author.

3. Data

The chosen data used in this study are all S&P500 index components listed until May 2019 (505 stocks in total) and get 21 years historical daily opening and closing prices from Datastream, since January 1998 to May 2019. Stocks from companies with no complete historical data[3] will be dropped for the following reasons: reduce the sample, robustness in the econometric tests results and to have the same data amount for the neural network training, validation and test. As a result, 324 stocks remain in the sample.

The observed data period spanning from 1 January 1998 to 13 May 2019 has been segregated into in-sample and out-of-sample data as used during the modelling process (Table 1).

Table 1 Prices data segregation for the full sample period.

Period	Trading days	Beginning	End
Total dataset	5370	02 January 1998	13 May 2019
Training dataset (in-sample)	3267	02 January 1998	31 December 2010
Validation dataset (in-sample)	502	01 January 2011	31 December 2012
Test set (out-of-sample)	1601	01 January 2013	13 May 2019

In the functionality of neural networks, the above in-sample period was segregated once more into two sub-periods to compensate for the training and validation of each network to avoid "over-fitting" with the dataset. Also, we let the training dataset include the subprime crisis period to considered in the cointegration tests and to enable the network to learn in stress scenarios. This way, it is pretended that the neural network might be able to produce accurate forecasts during periods with high volatility.

Moreover, we also get from Datastream the S&P 500 index components historical market capitalization, price to book value, enterprise value to EBITDA and net debt to enterprise value since January 2001 to May 2019. Start date is January 2001 because is the oldest date with historical fundamental data available in Datastream. These fundamental ratios are used because they have a daily frequency[4], so the sample data is greater and allow us to use more data to train the network than the fundamental ratios with lower frequency.

---

[3] A list of the dropped stocks can be found in appendix 1.
[4] Notice that daily frequency is given by one element in fundamental ratios formula, the stock price. Ratios formulas can be found in table 5.

This data is also segregated like the historical open and close prices to be used as neural network training, validation and test datasets. Table 2 shows the data segregation for fundamental ratios historical data.

Table 2 Fundamental ratios data segregation for the full sample period.

Period	Trading days	Beginning	End
Total dataset	4614	02 January 2001	13 May 2019
Training dataset (in-sample)	2511	02 January 2001	31 December 2010
Validation dataset (in-sample)	502	01 January 2011	31 December 2012
Test set (out-of-sample)	1601	01 January 2013	13 May 2019

4. Methodology

In this section, it will be described the pairs selection methodology, the LSTM network used for pairs spread forecast and the MLP network used for stocks signal (or trend) predictions. For both networks is explained the network configuration, like number of layers, number of neurons per layer, activation functions, cost function used to compute and minimize the errors and optimizers used for the network training. Finally, the proposed trading strategy used in this work is described, from trade criteria to open a position until criteria to close positions and the trading costs.

4.1 Pairs selection: Supercointegration method

Using the training dataset, it is calculated the supercointegrated pairs based on Figuerola-Ferretti et al. (2017)[17] to the pairs selection and consisting in a two-step procedure. This two-step selection procedure combines Johansen (1991)[24] method and Engle and Granger (1987)[14] method and its applied only to same industry stocks, i.e., it is only looked cointegration relationships among same industry stocks, because financially can be more reliable a long-run relationship between companies with similar productive activities than the ones with different activities.

Johansen test: First, it is tested the existence of cointegration among possible pairs by industry by conducting the Johansen (1991)[24] cointegration test at the 1% significance level. This methodology in based on the vector autoregressive (VAR) of order ρ, represented as follows:

$$y_t = \mu_t + A_1 \cdot y_{t-1} + A_2 \cdot y_{t-2} + \cdots + A_\rho \cdot y_{t-\rho} + B \cdot x_t + \epsilon_t \tag{16}$$

Where y_t is a vector containing k non-stationary variables $I(1)$, x_t is a vector of deterministic variables and ϵ_t is a vector of innovations. Its representation as Vector Error Correction Model (VECM) is:

$$\Delta y_t = \mu_t + \prod \cdot y_{t-1} + \sum_{i=1}^{\rho-1} \Gamma_i \cdot \Delta y_{t-i} + B \cdot x_t + \epsilon_t \tag{17}$$

Where

$$\prod = \sum_{i=1}^{\rho} A_i - I_k, \qquad \Gamma_i = -\sum_{j=i+1}^{\rho} A_j, \qquad i = 1, 2, \dots, \rho - 1 \tag{18}$$

VAR model is estimated by maximum likelihood and then there are two contrast to test the existence of cointegration: trace test and maximum eigenvalue test.

Within the trace test, the null hypothesis is:

$$H_0: Rank(\textstyle\prod) \leq m$$

$$H_1: Rank(\textstyle\prod) > m$$

That means: H_0: variables in vector y_t has a maximum of m cointegration relationship, against the alternative hypothesis: H_1: the variables in vector y_t has more than m cointegration relationships.

Within the maximum eigenvalue test, the null hypothesis is:

$$H_0: Rank(\textstyle\prod) = m$$

$$H_1: Rank(\textstyle\prod) = m + 1$$

After analyzing the data, it is calculated the model considering intercept in the cointegration relations and trends in the data, because we are working with financial series and they usually have trend. Johansen methodology is applied and are selected the pairs that at least approve one of the two contrast tests explained above. If the pair shows cointegration with both tests, the pair is "strongly cointegrated", while if the pair only pass one of the tests, the pair is "weakly cointegrated". The tests are restricted to a 99% confidence level. It was found 145 cointegrated pairs by Johansen approach.

Engle-Granger test: In a second step, it is refined our selection of pairs by running the OLS regression in equation [(3)](#) and once identified the residuals in the OLS regression, the test of a unit root is conducted, applying the Augmented Dickey Fuller Test:

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

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:

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

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

If the test rejects the null hypothesis, there is no unit root and the stocks are cointegrated, which implies the existence of a long-run stochastic common trend between the variables and his first differences reverts throughout the time.

Therefore, we apply the Engle-Granger methodology to the previous selected pairs with Johansen test restricted to a confidence level of 99%, with $\alpha = 0.01$. We find a total of 21

supercointegrated pairs (3 weakly cointegrated and 18 strongly cointegrated), which are displayed in appendix 2.

It is important to mention that supercointegration based on the two-step analysis explained before did not implies strongest cointegration between the pair or additional statistic properties respect to pairs which pass one cointegration test, but Figuerola-Ferretti et al. (2017)[17] shows that pairs trading profitability is monotonically associated with the significance level of cointegration and it called supercointegrated pairs the ones which do not reject standard cointegration tests at the 1% significance level.

4.2 Pairs spread prediction: LSTM network

Once the supercointegrated pairs are selected, will be applied neural network techniques to predict the pair spread for each selected pair and their results will be used to trade following the trading criteria explained later in this work. The network type used for these predictions is the LSTM method, that is a type of recurrent neural network enable to capture long-run dependencies in the data, something useful for pairs trading.

The recurrent network is able to remember important things about the inputs they received, which enables them to produce precise predictive results in sequential data and that's why this type of network is widely applied to information with sequential structure, like time series and financial data, as can be observed with the RNN applications made by Dunis et al. (2006a, 2006b, 2015)[11][9][10] and Van Der Have (2017)[35] to pairs trading. However, it will be explained the reason to use LSTM instead of standard RNN below, but first will be explained how it works each network.

4.2.1 Recurrent Neural Network (RNN)

Recurrent neural network is a type of neural network where hidden layer outputs from previous period are feed as input to the current period. This characteristic allows the information to persist over time and it's the main difference with feedforward neural networks (FNN). The RNN save $t-1$ hidden layer(s) outputs in a state unit(s) and it is used as new input for the hidden layer(s) in order to train itself at time t. An example of RNN with single hidden layer can be found in figure 1.

Where a FNN only uses the information at the current time to predict the value at the next time, a RNN has the advantage that its recurrent properties allow the network to also look at and remember the information from the past (Sundermeyer et al., 2014[33]). This works, as it was explained before, by taking hidden layer output from previous period and use it as input data to the current period.

Given that we are going to predict time series with mean-reverting properties, being able to use long-term dependencies should be useful. This is the main reason to use recurrent networks instead of feedforward neural networks in this work.
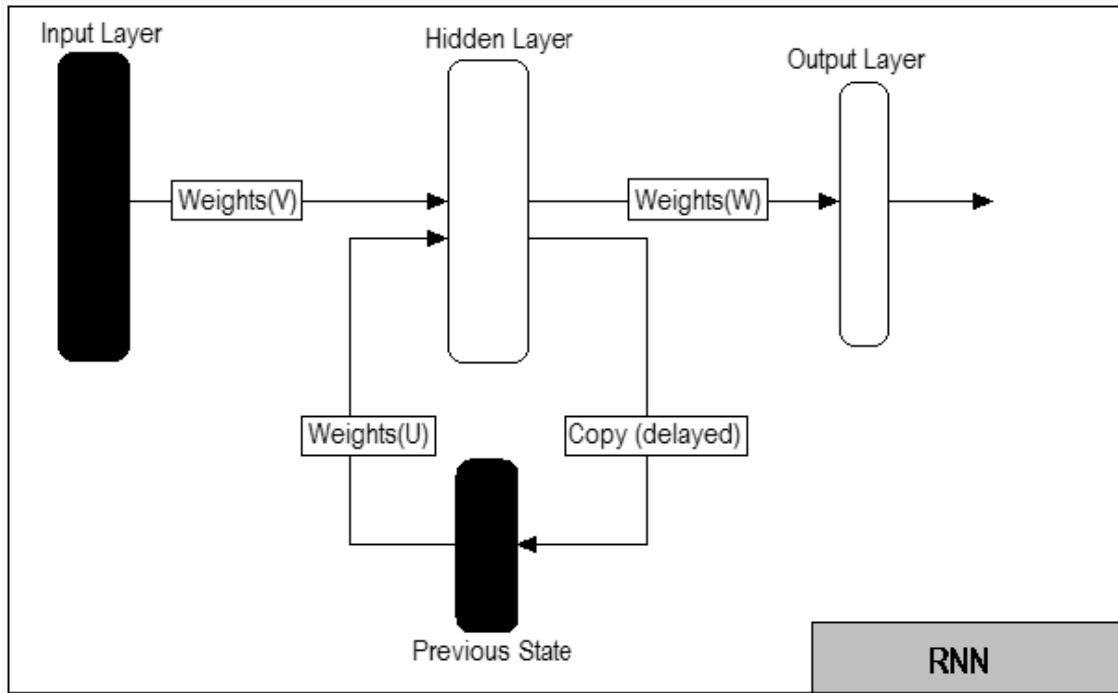
Figure 1 **Graphical illustration of a recurrent neural network.** RNN with one hidden layer, whose inputs are the data from input layer and the previous hidden layer data saved in previous state.
Source: Dunis et al. (2006)[9].

The input used in the input layer are lags of the time series of the pair spread, lags from each stock returns, one year mean and standard deviation. The output produced is the forecast of the spread at time t + 1. The RNN output can be written as follows:

$$\hat{y}_t = G_1(W_{ya}a_t + b_y) \tag{20}$$

Where G_1 is the output layer activation function, W_{ya} is the weights vector between the hidden and output layer, b_y is the bias vector and a_t is the hidden layer output defined as

$$a_t = G_2(W_{aa}a_{t-1} + W_{ax}x_t + b_a) \tag{21}$$

Where G_2 is the hidden layer activation function, W_{aa} is the weights vector applied to the previous hidden layer values in state neuron a_{t-1}, W_{ax} is the weights vector between input layer and hidden layer, x_t is the input layer values and b_a is the bias vector applied in the hidden layer. Note that the activation functions in the hidden and output layer can be different and depends of the problem which we are working.

However, the information persistence only allows short-term memory and it is difficult learn long-term dependencies in the standard RNN approach as stated by Bengio et al. (1994)[2]. Also, as mentioned in section 2.5.1, this method has as disadvantage the vanishing (1) or exploding (2) gradient problem when the EBP (error backpropagation) algorithm is executed. Hochreiter & Schmidhuber (1997)[19] explained that in case 1, learning to bridge long time lags takes a prohibitive amount of time or does not work at all; case 2 may lead to oscillating weights. Both cases makes difficult the neural network training.

4.2.2 Long Short-Term Memory (LSTM)

To solve gradient problems from standard RNN, Hochreiter & Schmidhuber (1997)[19] proposed the LSTM methodology, which main difference respect to standard RNN is the incorporation of an additional unit called cell state used to store and process the information from all previous periods, instead of just the previous hidden layer data. This way, the LSTM network is designed to avoid the long-term dependency problem.

The cell state units have linear interactions with the inputs and previous hidden layers transformations through so-called gates. The gates can be seen as the memory of the network and are used to optimally let information through, i.e., allowing the network to decide whether or not to store or delete information based on the importance it assigns to the information. Figure 2a illustrates a hidden layer in a regular RNN, whereas figure 2b illustrates a hidden layer in LSTM network.

As can be observed in Figure 2b, the LSTM network consists of three gates, the forget gate, input gate and output gate. The forget gate, illustrated in figure 3, determines the information to be removed from previous cell state C_{t-1}, applying the sigmoid function to new input data x_t and previous hidden layer h_{t-1}. The sigmoid function is given by:

$$f_t = \sigma\left(W_{fh} \cdot h_{t-1} + W_{fx} \cdot x_t + b_f\right)$$

$$f_t = \frac{1}{1+e^{-(W_{fh} \cdot h_{t-1}+W_{fx} \cdot x_t+b_f)}} \tag{22}$$

Where W_{fh} and W_{fx} represents the weights vector in forget gate of previous hidden layer h_{t-1} and the input x_t respectively, and b_f is the bias vector. f_t output is between 0 and 1 for each number in the cell state C_{t-1}, where a value close to 0 represents a specific element to be removed because it isn't important, while a number close to 1 represents an important specific element to be kept in the next cell state.

(a) Hidden layer of a regular RNN   (b) Hidden layer of a LSTM
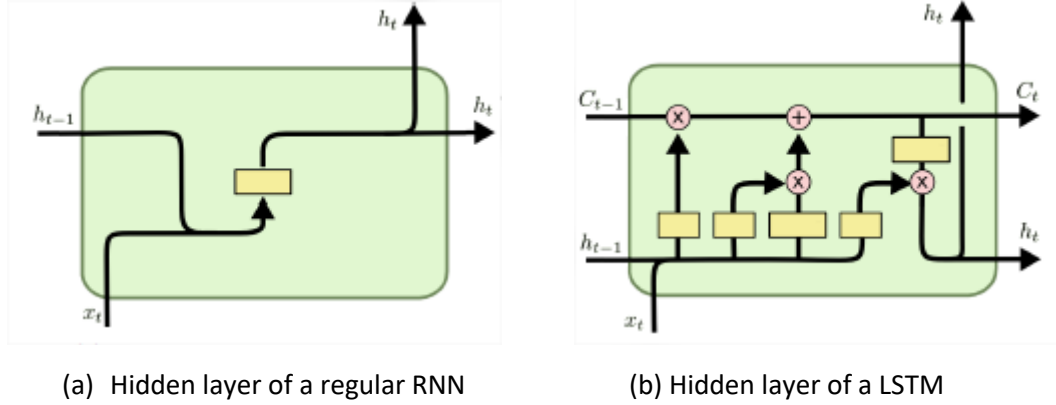
Figure 2 **Hidden layer comparison between a regular RNN (a) and LSTM (b).** The yellow rectangle represents nonlinear functions where the new input x_t and output from the previous hidden layer h_{t-1} are processed through. The red circles containing either a multiplication or addition sign represent the corresponding linear operation. Furthermore, the variable C_t denotes the cell state at time t.
Source: Van Der Have (2017)[35]
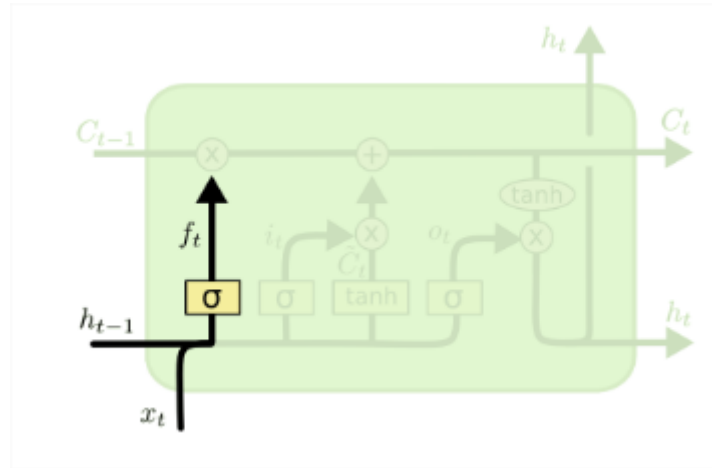
Figure 3 **LSTM forget gate.** The yellow rectangle containing σ represents the sigmoid function written in equation (22) applied to input data x_t and previous hidden layer h_{t-1}, returning as output f_t.
Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

The second gate is the input gate, which makes two operations with previous hidden layer h_{t-1} and input data x_t. First, a sigmoid function is applied to decide the values to be updated, resulting in output i_t, given by:

$$i_t = \sigma(W_{ih} \cdot h_{t-1} + W_{ix} \cdot x_t + b_i)$$

$$i_t = \frac{1}{1+e^{-(W_{ih} \cdot h_{t-1}+W_{ix} \cdot x_t+b_i)}} \tag{23}$$

Where W_{ih} and W_{ix} represents the weights vector in input gate of previous hidden layer h_{t-1} and the input x_t respectively, and b_i is the bias vector. i_t output is between 0 and 1 for each number in the cell state C_{t-1}, where a value close to 0 represents a specific element to not be updated, while a number close to 1 represents a specific element to be updated in the cell state.

At same time, a hyperbolic tangent function is used to create a vector of possible new values $\widetilde{C}_t$, and is given by:

$$\widetilde{C}_t = \tanh(W_{Ch} \cdot h_{t-1} + W_{Cx} \cdot x_t + b_C)$$

$$\widetilde{C}_t = \frac{e^{2(W_{Ch} \cdot h_{t-1}+W_{Cx} \cdot x_t+b_C)}-1}{e^{2(W_{Ch} \cdot h_{t-1}+W_{Cx} \cdot x_t+b_C)}+1} \tag{24}$$

Where W_{Ch} and W_{Cx} represents the weights vector of previous hidden layer h_{t-1} and the input x_t respectively, and b_C is the bias vector. $\widetilde{C}_t$ output is between -1 and 1, these values are multiplied by i_t and gives a set of candidate values scaled by how much we decided to update each state value. Next, this result is added to the forget gate value to obtain the new cell state value:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \tag{25}$$

To summarize, Figure 4a shows the nonlinear operations applied to x_t and h_{t-1}, while figure 4b shows the linear operation between the nonlinear output and the state cell unit C_{t-1} to finally get the new cell state value given by (25) in input gate.
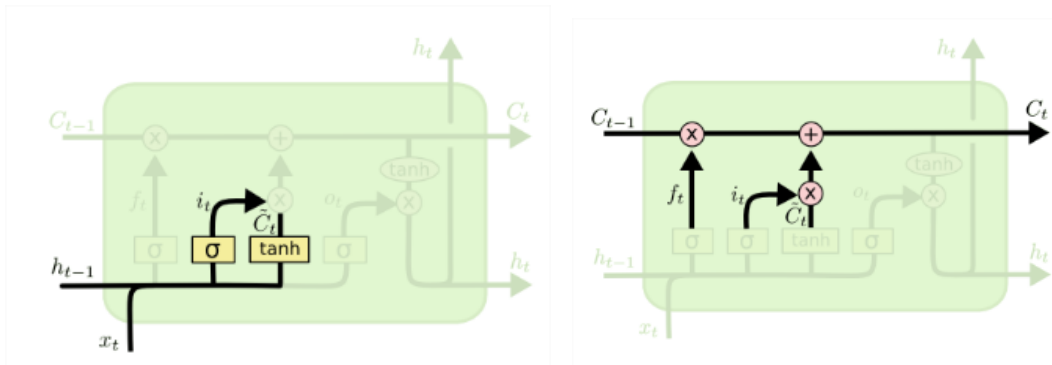
(a) Non-linear functions   (b) Linear operations

Figure 4 **LSTM input gate**. Figure (a) indicates the non-linear functions applied to input data x_t and previous hidden layer h_{t-1}. The yellow rectangle containing σ represents the sigmoid function as in equation (23) and generates as output i_t. The yellow rectangle containing $tanh$ represents the hyperbolic tangent function written in equation (24), resulting in output $\widetilde{C}_t$. Figure (b) shows the linear operations between the forget and input gates, resulting in new cell state C_t given by equation (25).
Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Finally, the output gate is computed based on a sigmoid layer which decides the part of the cell state we are going to output, given by

$$o_t = \sigma(W_{oh} \cdot h_{t-1} + W_{ox} \cdot x_t + b_o)$$

$$o_t = \frac{1}{1 + e^{-(W_{oh} \cdot h_{t-1} + W_{ox} \cdot x_t + b_o)}} \tag{26}$$

Where W_{oh} and W_{ox} represent the weights vector of previous hidden layer h_{t-1} and the input x_t respectively, and b_O is the bias vector in the output gate. Next, this result is multiplied by the cell state value given by equation (25) passing through a hyperbolic tangent function. The output gate is illustrated in figure 5 and the final output of the hidden layer h_t is:

$$h_t = o_t \cdot \tanh(C_t)$$

$$h_t = o_t \cdot \frac{e^{2(C_t)} - 1}{e^{2(C_t)} + 1} \tag{27}$$

Given the fact that o_t is between 0 and 1, and $\tanh(C_t)$ is between -1 and 1, the hidden layer output range is $[-1,1]$. For this work, this is useful given that our LSTM output will be a pair spread return prediction.
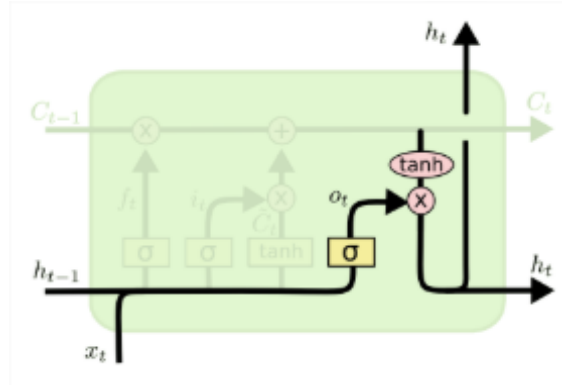
Figure 5 **LSTM Output gate:** The yellow rectangle containing σ represents the sigmoid function applied to previous hidden layer h_{t-1} and input data x_t as in equation (26). The red circle containing *tanh* represents the hyperbolic tangent function applied to cell state value C_t computed in equation (25), and the other red circle containing multiplication signal indicates the linear operation expressed in equation (27).
Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

The general form for the spread model predicted by the LSTM neural network with M hidden layers is the following:

$$S_{t,i,j} = \phi' \cdot \tilde{x}_t + \sum_{m=1}^{M} \lambda_m \left(F(\widetilde{w'} \cdot [x_t, h_{t-1}] - b_m) \cdot \frac{e^{C_t} - e^{-C_t}}{e^{C_t} + e^{-C_t}} \right) + e_t \tag{28}$$

for $t = 1, \dots, T$, and where $\tilde{x}_t = [1, x_t']'$, the parameters $\lambda_1, \dots, \lambda_M$ represent the weights vector between hidden and output nodes, and $\widetilde{w_1}, \dots, \widetilde{w_M}$ represent the weight vector between the

input and hidden nodes. $F(\cdot)$ are the activation functions in the hidden unit, in this case sigmoid function applied in output gate.

The LSTM network will be used by the reasons explained before and its advantages respect to the traditional RNN model, which are the long-term memory given the pairs spread long-run mean-reverting property and the gradient descent problems solved by LSTM network respect to standard RNN.

4.2.3 LSTM network architecture

In this section will be explained the proposed neural network architecture structure and the parameters used to make the network training and spread predictions.[5]

Network layers and neurons: The proposed network has 3 layers, the input layer containing 17 neurons, one hidden layer with 17 neurons and the output layer with 1 neuron because the output is the spread prediction for next day. The number of neurons in input and hidden layer is equal to the number of explanatory variables in table 3.

Activation functions: The hidden layer is the LSTM layer explained in section 4.2.2, containing the three gates and same activation functions mentioned before, that is sigmoid function in forgot gate, for input gate both sigmoid and hyperbolic tangent functions, and for output gate a sigmoid function applied to input and previous hidden layer data and hyperbolic tangent function applied to cell state values. Output layer has a linear activation function because using an activation function like sigmoid generates values between 0 and 1, and the real spread can take values out of that range (similar case if use hyperbolic tangent function with range between -1 and 1).

Table 3 LSTM network explanatory variables

Number	Variable	Lags (Days)
1	Pair spread cumulative returns	1-5
2	Stock i cumulative returns	1-5
3	Stock j cumulative returns	1-5
4	1-year spread mean (250 days mean)	1
5	1-year spread standard deviation (250 days volatility)	1

---

[5] Network training and prediction total time execution was 8.5 hours. Computer Intel core I3 with windows 10 pro, processor intel Pentium CPU N3540 2.16 GHz, solid D.D. 256 GB and RAM 4GB.

Cost function and optimizer: the function to be minimized by the network is the mean squared error (MSE) defined as follows:

$$MSE = \frac{1}{N} \cdot \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \qquad (29)$$

where $\hat{y}_i$ is the network prediction, y_i is the spread true value and N is the sample size (3267 in this case, see training dataset in table 1). Adam optimizer is used for the training, using the default parameters provided in the original paper published by Kingma and Ba (2015)[25]. O*ther parameters* values can be found in table 4.

Table 4 LSTM architecture parameters

Parameter name	Parameter Value	Description
Batch size	20	Number of samples per gradient update
Epochs	20	Number of times that feedforward − backpropagation procedure is executed during network training
Learning rate	1%	Learning rate used for weights adjustment
Dropout rate	20%	Consisting in randomly setting a fraction rate of input units to 0 at each update during training time, which helps prevent overfitting.

Finally, is important to mention that following Dunis et al. (2006a, 2006b, 2015)[11][9][10] it will be estimated ten networks per pair because the starting point for each network is a set of random weights, affecting the training procedure and making sensitive the outputs respect to the starting point. Therefore, the average results of all ten neural networks was used as the spread prediction in this work.

Looking for biggest accuracy in network forecast, especially in the last years of out-of-sample dataset, is proposed a one-year rolling window for the predictions, allowing the network to retrain itself and adjust their weights using the new information. This way, after network training with in-sample data, next year spread is predicted by the network and then, when this year ends, is used the new available information to retrain the network and fit the weights. This process is repeated for all years in out-of-sample dataset.

4.3 Stocks signal predictions using MLP multi-classification network

After pairs spread predictions, it's proposed a classification problem for all supercointegrated pairs stocks components. This problem main objective is getting each stock trend direction

28

predictions using as inputs the fundamental ratios per stock and training a neural network enable to predict whether a specific stock have bullish, bearish or no trend signal, and finally use the outputs as an additional factor to trade with pairs trading strategies. Therefore, it will be proposed an MLP classification network to make signal predictions per each stock in the supercointegrated pairs. It is used a MLP network because is one of the most common network types used for classification problems in previous literature, like Nanopoulos et al. (2001)[30], who use a MLP to classify time series based on statistical data, or Pham et al. (1994)[31] who proves that MLP network performs better than standard learning vector quantization (LVQ) neural network when are applied to control chart pattern recognition.

4.3.1 Multilayer perceptron (MLP)

The MLP is a supervised learning algorithm that learns by training on a dataset, enable to learn a non-linear function approximator for classification or regression. As mentioned before, in this work will be applied for classification purpose. The MLP basic structure consists of three layers: an input layer (with explanatory variables), hidden layer and output layer (the model estimation), where the input and hidden layer includes a bias node (similar to the intercept in standard regression). Figure 6 shows an example of a single hidden layer MLP network.
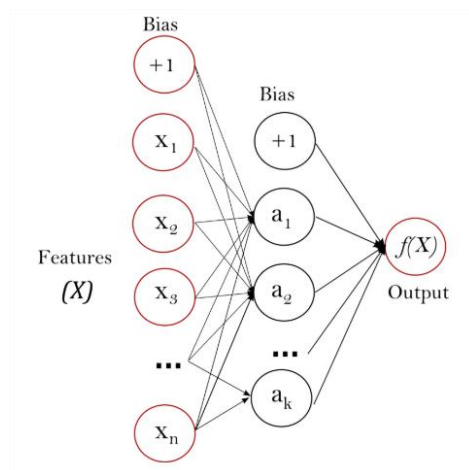
Figure 6 **One hidden layer MLP network**. The red circles in the left represents the $n + 1$ neurons in the input layer. Black circles in the middle represents the $k + 1$ neurons in the hidden layer and red circle in the right is the output layer neuron.
Source: https://scikit-learn.org/stable/modules/neural_networks_supervised.html

The information process made by the network starts with the input nodes containing the explanatory variable values (or features). These values are transmitted to hidden layer nodes as the weighted sum of its inputs and there is applied a non-linear activation function before

passing the information to output layer, when the information is again transmitted as the weighted sum of previous hidden layer output and is applied another activation function. The final output of the network is used to compute a cost (or loss) function to be minimized by well-known error backpropagation method (also called generalized delta rule), whose main idea is after compare both network and desired outputs, implement a gradient descent method for finding the weights that minimize the cost function value, transmitting backwards the errors.[6]

Notice that the training process can be divided in two steps: first, the algorithm has a feedforward advance, and then it is adjusted in a backward way. During feedforward face, the inputs pass through the networks until it arrives the output layer. Once the network output values are obtained, the backward test starts comparing these outputs with target values in order to compute the error and adjust the weights since output layer until it arrives the input layer.

The math behind the MLP network example shown in Figure 6 is the following:

1. The hidden layer receives the input values weighted sum, applies an activation function and gets as output a_i defined as follows:

$$a_i = f_1(w_{hi} \cdot X + b_I) \tag{30}$$

where w_{hi} is the hidden layer weights vector applied to neuron $i \in [1, 2, ..., k]$, X represent the input values vector, b_I is the bias vector received from input layer and $f_1(\cdot)$ represents the activation function.

2. Output layer receives hidden layer outputs, computes a weighted sum with them and applies an activation function (not necessarily the same used in hidden layer) to get the network output $\hat{y}$:

$$\hat{y} = f(X) = f_2(w_o \cdot A + b_h) \tag{31}$$

where w_o represent the output layer weights vector, $A = [a_1, a_2, ..., a_k]$ is a vector containing all hidden layer outputs, b_h is the bias vector received from hidden layer, and $f_2(\cdot)$ is the activation function.

---

[6] For more details about the error backpropagation method, see Rumelhart et al. (1986)[32].

3. The cost function is calculated using the network outputs $\hat{y}$ and target outputs. Let's call the cost function $E(\cdot)$, which represents the overall network error from all $k+1$ neurons from hidden layer. The partial derivative of the errors is calculated using chain rule:

$$\frac{\partial E}{\partial w_l} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w_l} \tag{32}$$

where w_l is the weight from layer $l \in [h, o]$ (hidden or output layer). Notice that in $\frac{\partial \hat{y}}{\partial w_l}$, the network output depends directly of output layer weights w_o, but also depends of hidden layer weights w_h in A. Therefore, we can rewrite equation (32) for hidden layer weights as follows:

$$\frac{\partial E}{\partial w_h} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial A} \cdot \frac{\partial A}{\partial w_h} \tag{33}$$

4. Finally, the weights update is computed using the following formula:

$$\widehat{w_{lj}} = w_{lj} - \alpha \cdot \frac{\partial E}{\partial w_{lj}} \tag{34}$$

where w_{lj} is the old weights vector of layer l neuron j, $\alpha > 0$ the learning rate and $\widehat{w_{lj}}$ is the new weights vector.

5. The process is repeated according to epoch parameter value.

4.3.2 MLP network architecture

In this section will be explained the proposed MLP neural network architecture structure and the parameters used to make the network training and stocks signal predictions. [7]

Network layers and neurons: The proposed MLP network has 4 layers, the input layer containing 20 neurons, two hidden layers with 20 neurons per layer and the output layer with 3 neurons. The number of neurons in input and hidden layers is equal to the number of explanatory variables in table 5.

---

[7] Network training and prediction time execution was 9 hours.

Table 5 MLP network explanatory variables

Number	Variable	Ratio formula	Lags (Days)
1	Price to book value	$\dfrac{P_{i,t}}{BV}$	1-5
2	Market capitalization	$P_{i,t} \cdot N_i$	1-5
3	Enterprise value to EBITDA	$\dfrac{EV}{EBITDA}$	1-5
4	Net debt to enterprise value	$\dfrac{Net\ debt}{EV}$	1-5

Where $P_{i,t}$ is the price of stock i in time t, BV is the asset carrying value on the balance sheet given by the following formula:

$$BV = \frac{Shareholders\ equity - preferred\ equity}{Number\ of\ common\ shares}$$

N_i is the total number of outstanding shares, EV represent the company value (how much worth the company) and its value is given by the addition of market cap and net debt. $EBITDA$ (earnings before interest, tax, depreciation and amortization) is a measure of a company's operating performance and $Net\ debt$ indicates the overall debt situation of a company by netting the value of the liabilities and debts of a company along with its cash and other similar liquid assets.

These fundamental ratios were selected because they have daily frequency, giving more data to train the network than ratios with lower frequency. Also, some of these ratios have been used in previous works about value investing. For example, market capitalization and price to book value are used in the well-known three and five factor models proposed by Fama and French (1993, 2014)[30][31]. Moreover, enterprise value to EBITDA is a good measure to compute the return a company makes respect to the company's enterprise value and helps to determine whether a company is undervalued or overvalued for shares of the same sector, something helpful for the trend prediction we want to do with the MLP network and to compare two shares of the same industry[8]. Finally, net debt to enterprise value give us a measure to determine whether a company is increasing or reducing his debt and its ability to afford its future obligations. Leverage measures are usually a good choice for value investing, e.g., Piotroski

---

[8] Pinto et al. (2019)[28] state EV/EBITDA is clearly a widely used metric in current valuation practice.

(2006)[29] states that companies with growing debt shows a bad signal about its financial risk and stocks with low leverage level are recommended.

The number of neurons in output layer corresponds with the 3 possible classes per stock: bullish, no trend or bearish. The network output are probabilities of each class and the signal is given by the neuron with highest probability, converting this value to one and making the other classes zero. Summarizing, the signals are described as follows:

$$Bullish\ if\ output = \ [1,0,0]$$

$$No\ trend\ if\ output = \ [0,1,0]$$

$$Bearish\ if\ output = \ [0,0,1]$$

It is important to mention that network outputs are long-run (one-year) trend predictions, because supercointegration relationship implies long-run mean-reversion property. Bullish trend happens during training period when spread in $t + 250$ grows more than two standard deviations respect to spread in t, bearish trend occurs when spread in $t + 250$ drop more than two standard deviations respect to the spread in t and network output must be no-trend in other cases. It was selected as target the one-year trend because it was used one-year mean and one-year standard deviation to compute the thresholds.

Activation functions: For both hidden layers are used sigmoid activation functions because is one of the most widely used in practice. Output layer uses the usual choice for multi-class classification problems, the softmax function. The main reason to use this function is that the probability of a class is not independent from the other class probabilities, i.e., one stock cannot be bullish and bearish at same time. Softmax function $\in [0,1]$ computes conditional probabilities while sigmoid function $\in [0,1]$ computes unconditional probabilities, so is used the first one since we only want to predict a single label per sample (for networks that wants to predict more than a single label, sigmoid works better). Softmax function is given by the following formula:

$$Softmax = f(s)_i = \frac{e^{s_i}}{\sum_{j=1}^{C} e^{s_j}} \tag{35}$$

Where i is the ith class, C is the total number of classes and s is the weighted sum of previous hidden layer output.

Cost function and optimizer: the function to be minimized by the network is the categorical cross entropy function defined as follows:

$$CCE = -\sum_{i=1}^{C} t_i \cdot \log(f(s)_i) \tag{36}$$

where $f(s)_i$ is the softmax function in ([35](#)) and t is the target vector. Given that in our problem, target output labels are one-hot, only the positive class C_p keeps its term in the loss, i.e., there is only one element of the target vector t which is not zero $t_i = t_p$. So, rewriting ([36](#)):

$$CCE = -\log\big(f(s)_p\big) \tag{37}$$

Adam optimizer is used for the training, using the default parameters provided in the original paper published by Kingma and Ba (2015)[25]. O*ther parameters* values can be found in [table 6](#).

Table 6 MLP network architecture parameters

Parameter name	Parameter Value	Description
Batch size	1	Number of samples per gradient update
Epochs	20	Number of times that feedforward − backpropagation procedure is executed during network training
Learning rate	1%	Learning rate used for weights adjustment
Dropout rate	20%	Consists in randomly setting a fraction rate of input units to 0 at each update during training time, which helps prevent overfitting.

Like the LSTM network, MLP uses a one-year rolling window to retrain itself and fit the weights to new information (see [section 4.2.3](#)).

4.4 Trading strategy criteria

The trading strategy consists in using the supercointegrated pairs to trade during the out-of-sample period defined in [section 3](#). To take the trading decisions, will be followed the previous literature criteria, which usually compute threshold limits for the spread and when the spread reaches the limit, an operation is executed. Also, will be propose a new alternative to trade based on the fundamental signal of the stocks in the pair, predicted using the MLP neural network and different ways to combine the trading signals and close position criteria to determine which one gives the best results and to prove whether adding fundamental ratios improves the strategy returns.

Is important to mention that all criteria of trading involving thresholds will be applied to both real pair spread value and LSTM network spread predictions. The purpose of use LSTM

predictions is to test whether trading with neural network predictions improves results from real spread value, which it's usually used in pairs trading.

4.4.1 Thresholds

In this work, will be used the two most widely threshold limits used in previous literature about pairs trading, the fixed volatility threshold used by Gatev et al. (1999)[18] and others like Do et al. (2009)[6], and the conditional volatility threshold used by Figuerola-Ferretti et al. (2017)[17].

4.4.1.1 Fixed volatility threshold

This threshold consists on compute a constant volatility and define the upper and lower bounds as two unconditional standard deviations from the mean. As will be used around 6.5 years in the out of sample period to trade, this fixed threshold will be a little different from previous literature works because in general the past pairs trading researches do not used an out of sample period as long as the one used in this work. Like a rolling window is used to re-train the neural networks as explained in section 4.2.3, the unconditional volatility will be updated with the rolling window, getting at the beginning of each year an unconditional volatility equal to last year spread cumulative returns volatility, and using the result as the following year constant volatility to compute the thresholds. The fixed thresholds can be defined as follows:

$$Upper\ fix\ threshold = \ \mu_t + 2 \cdot \sigma$$

$$lower\ fix\ threshold = \ \mu_t - 2 \cdot \sigma$$

Where μ_t represent a 1-year spread mean and σ is the unconditional standard deviation from previous year. The reason to use a dynamic mean μ_t is explained later in this work. Figure 7 shows for the supercointegrated pair "BBY.N (Best Buy) & FL.N (Foot Locker)" the fixed unconditional volatility, where at the end of each year, is recalculated and used as following year constant volatility.
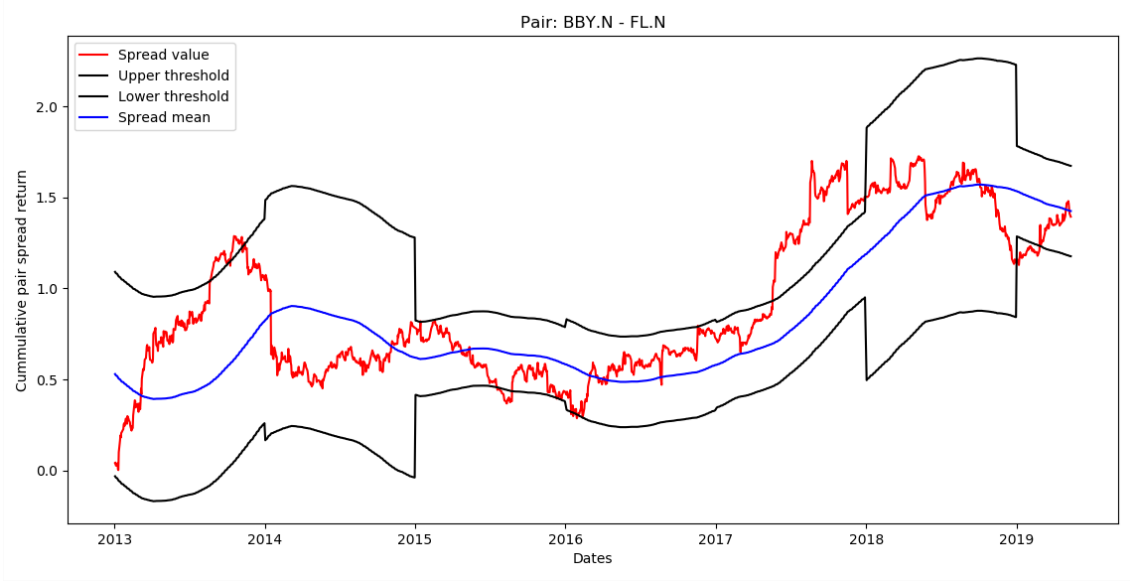
Figure 7 **Fixed threshold for pair BBY.N & FL.N:** The blue line represents the last year spread mean, the red line represents the spread cumulative returns and black lines represents upper and lower thresholds. Notice that the thresholds moves parallel to the mean and his level change when a new year begins.

4.4.1.2 Conditional volatility threshold

As Figuerola-Ferretti et al. (2017)[17] do, it will be used a dynamic threshold based on conditional volatility computed as the standard deviation using 1-year rolling-window, and the upper and lower bounds will be defined as two standard deviation from the mean:

$$Upper\ dynamic\ threshold = \ \mu_t + 2 \cdot \sigma_t$$

$$Lower\ dynamic\ threshold = \ \mu_t - 2 \cdot \sigma_t$$

Where μ_t and σ_t represent 1-year mean and standard deviation respectively. Figure 8 shows for the supercointegrated pair "BBY.N (Best Buy) & FL.N (Foot Locker)" the dynamic conditional volatility. This way, as stated by Roa (2018)[1], we capture the time-varying nature of spread time series. Notice that difference between fixed and dynamic threshold is the volatility update frequency. With fixed method, volatility update occurs at the end of each year, while dynamic threshold volatility has daily update.
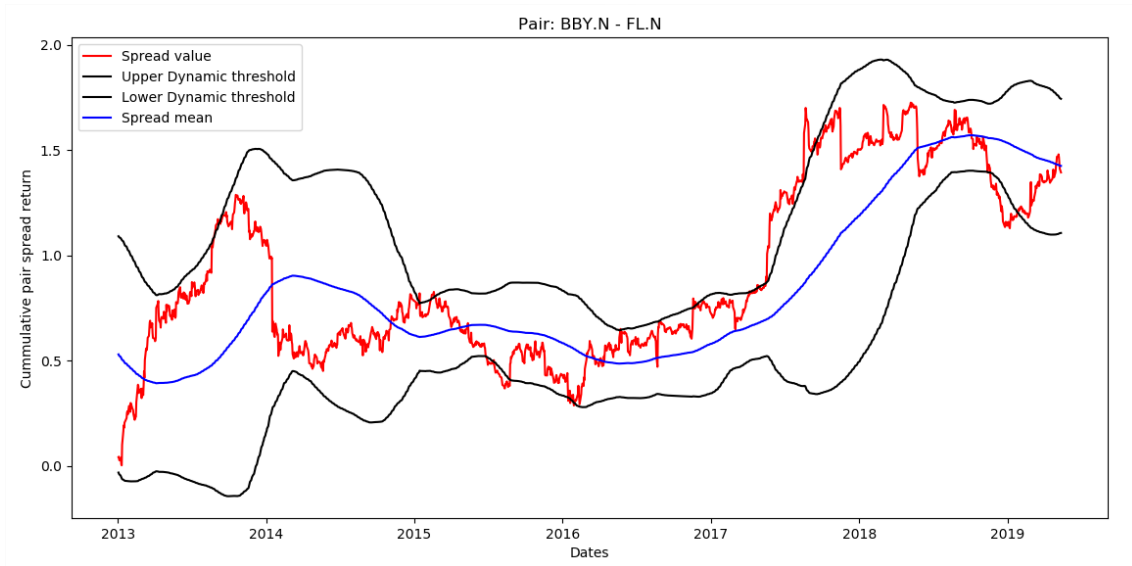
Figure 8 **Dynamic threshold for pair BBY.N & FL.N**: The blue line represents the last year spread mean, the red line represents the spread cumulative returns and black lines represents upper and lower dynamic thresholds. Notice that after high volatility periods, the distance between the bands increases and in moments of low volatility the bands get closer to spread mean.

4.4.2 Trade criteria

It will be proposed three alternatives to trade, using fixed and conditional volatility thresholds explained in section 4.4.1, using stocks trend prediction with the MLP network explained in section 4.3.2 and combining both threshold and trend predictions methods.

4.4.2.1 Trading with thresholds signals

First trade criterion is the widely used threshold trade method. Traditional methods propose trading when the spread reaches one of the two standard deviation thresholds from zero and closing position when spread mean-reverts. This method works well with short period of out-of-sample data, usually 1 or 2 years in previous literature. However, in this work we are working with 6.5 years of out-of-sample data to trade, generating that in most cases the cumulative spread returns takes great values given by trends during this period, which implies cases when the spread gets over the threshold and never returns inside the thresholds. Figure 9 shows an example applied to supercointegrated pair "CLX.N (The Clorox Company) & PG.N (The Procter & Gamble)" where the spread overpass one threshold and stays out of volatility limits.
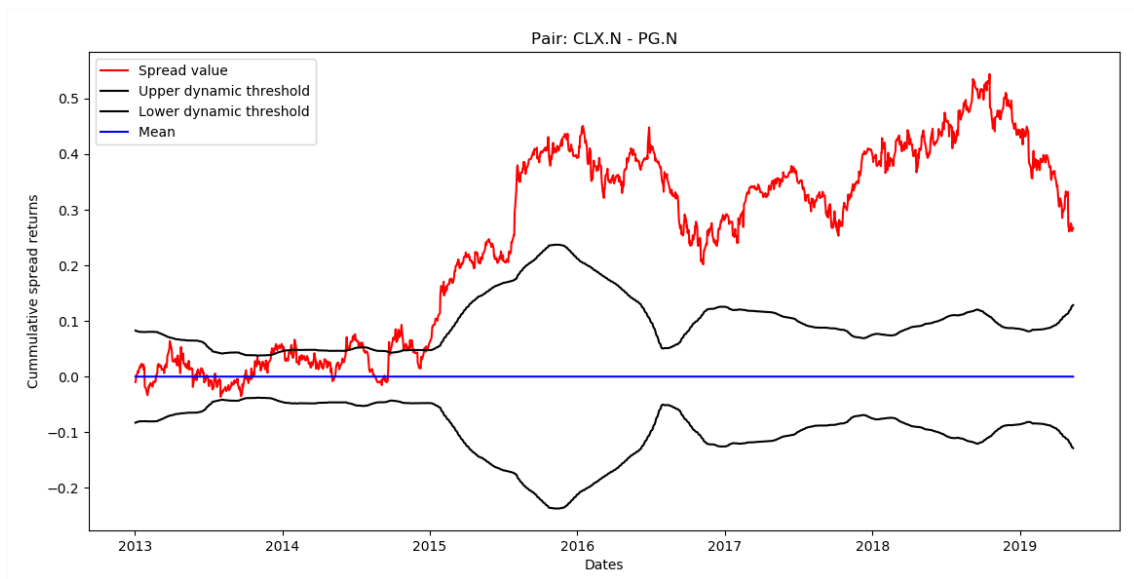
Figure 9 **Dynamic threshold assuming zero-mean for pair CLX.N & PG.N**: Blue line represents zero-mean label, black lines represents the upper and lower conditional volatility thresholds and red line is the cumulative spread returns during out-of-sample period. Notice that after a bullish trend in 2015, the spread value breaks upper threshold and didn't come back inside the thresholds.

Therefore, a dynamic mean is used to compute the spread confidence bands in each moment of time t. This way, it is avoided the problem because the mean value reflects the trend followed by the spread. Figure 7 and Figure 8 shows examples of fixed and conditional volatility thresholds moving around a dynamic mean. Will be used a 1 year mean for this purpose.

Hence, trade will be triggered in the following scenarios:

1. Traditional method: When spread value reaches lower (upper) threshold, a long (short) position in the spread is taken.
2. Alternative method: When spread value is below (above) the lower (upper) threshold and breaks this threshold from down (up) to up (down).

With the alternative method, we want to reduce bad trades avoiding them when the spread reaches the lower (upper) band by first time, since can be in a bearish (bullish) trend that still does not ends. When the spread breaks the lower (upper) threshold from bottom up (top to bottom), is in its long-run mean-reverting process, the main assumption in most of pairs trading approaches, including this one given the supercointegration relationship. Figure 10 shows an example of both scenarios to open a position using thresholds for supercointegrated pair "BBY.N (Best Buy) & FL.N (Foot Locker)".
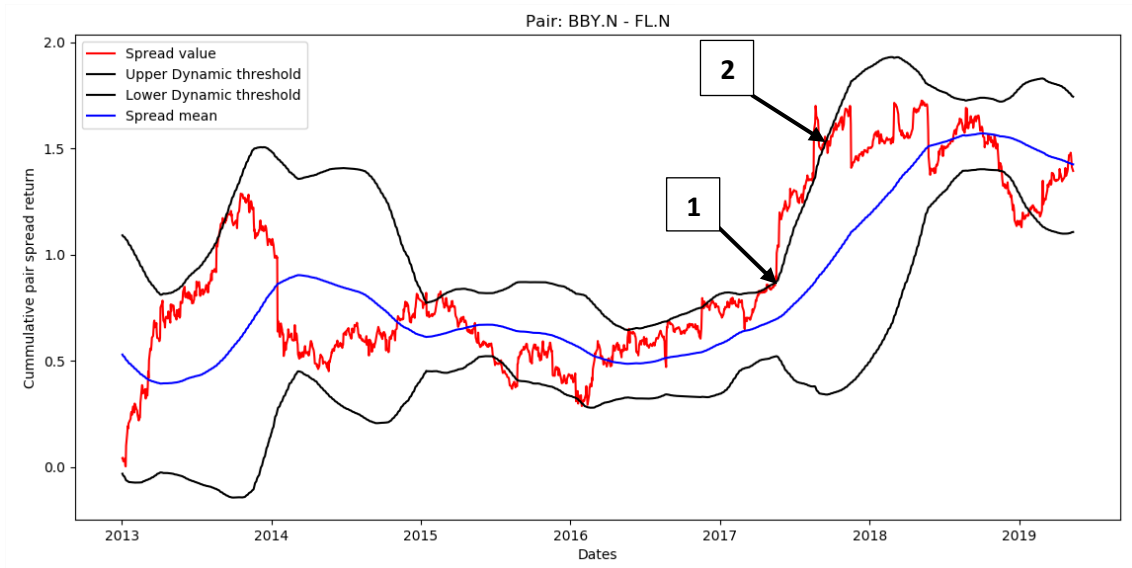
Figure 10 **Graphical illustration of threshold trade criteria**: Black square containing number 1 indicate short position in the spread when its value reaches by first time the upper threshold. Black square containing number 2 indicate short position in the spread when its value is breaking upper threshold from top to bottom.

4.4.2.2 Trading with MLP network predictions

As mentioned in section 4.3.2, the MLP network give us an output vector with three elements in each time period t, indicating whether the stock is bullish, bearish or has no trend. Using these outputs, trades will be triggered in the following cases:

1. When at least one of the pair stocks components gives a bullish or bearish signal and the other one shows a different trade signal (i.e. will not be trades when both stocks shows same trading signal).

2. When both stocks show opposite signals (i.e. one stock shows long position and the other one short position).

4.4.2.3 Trade combining thresholds and fundamental signals

This is an extension of the trading criteria explained in 4.4.2.1 and 4.4.2.2. A trade is executed in the following scenarios:

1. When spread value overcome the threshold, the signal prediction from MLP network of both stocks shows different trading signal and at least one of them shows same signal with both MLP and the spread position to take. This means that whether the threshold method indicates long position in the spread, i.e., long position in stock i and short

39

position in stock j, the MLP signal prediction must be long position for stock i or short position for stock j.

Notice that this is an extension of traditional method described in section 4.4.2.1 adding the MLP predictions to trade decision.

2. When spread value is below (above) the lower (upper) threshold and breaks this threshold from down (up) to up (down), the signal prediction from MLP network of both stocks shows different trading signal and at least one of them shows same signal with both MLP and the spread position to take. This means that whether the threshold method indicates long position in the spread, i.e., long position in stock i and short position in stock j, the MLP signal prediction of stock i shows long position or the signal of stock j shows short position.

Notice that this is an extension of alternative method described in section 4.4.2.1 adding the MLP predictions to trade decision.

Finally, is important to mention that the number of stocks to trade depends of both cointegration and investment amount parameters, where the investment amount per stock is 10,000 USD.[9] Rewriting the definition of the spread in equation (3):

$$S_{i,j} = \epsilon_t = p_{t,i} - \beta \cdot p_{t,j}$$

The long position in the spread consists in buying N_i unit of asset i, and selling N_j units of asset j, where N_j depends of β (the cointegration parameter computed with the OLS made in the supercointegration analysis described in section 4.1 with in-sample data and assumed constant over all out of sample period). Short position in the spread consists in selling N_i unit of asset i and buying N_j units of asset j.

N_i and N_j are defined as follows:

$$N_i = int\left(\frac{X}{OP_{i,t-1}}\right) \quad , \quad N_j = int\left(\frac{X}{OP_{j,t-1}} \cdot \beta\right)$$

where X represent the investment amount per stock, $OP_{i,t-1}$ and $OP_{j,t-1}$ are the opening prices of stock i and j in $t-1$ respectively and β the cointegration parameter as explained before. In theory, each pair trade is beta neutral, however N_i and N_j only can take integer values, so this

---

[9] This parameter is arbitrary, the strategy can be tested with different investment amount limits according to investor risk appetite and wealth.

makes than the position beta will not be equal to zero in most cases. Also, all trades are executed using stocks opening prices next day from the trading signal, this way will be used realistic prices to trade because all signals are triggered or computed using closing prices as inputs, so we need this data before investing.

Notice that N_j increase whether $\beta > 1$, in these cases the initial investment in asset j can be greater than the invested in asset i. Also is important to mention that cannot be opened more than one position at same time for a specific pair.

4.4.3 Close position criteria

When a position is opened by some of the criteria explained in section 4.4.2, there are three alternatives to close position, described as follows:

1. *Mean reversion criterion*: a position is closed when the spread value reaches the dynamic mean value (take profit criterion) or whether the spread overcome again the threshold during two consecutive periods (stop loss criterion). Figure 11 shows examples when a position is closed by taking profit or stop loss. This method is applied only for strategies that includes threshold method in section 4.4.2.1 to trade, except for traditional method from referents, which it does not include stop loss, only mean-reversion criterion.

2. *MLP signal criterion*: a position is closed whether both stocks signal predictions are different respect to the taken position in them when the trade was opened. This method is applied only to strategies that includes MLP signals predictions in section 4.4.2.2 to trade.

3. *Mean reversion and MLP signal combination:* a position is closed when 1 or 2 occurs. Applied only to trades opened based on combined method explained in section 4.4.2.3.
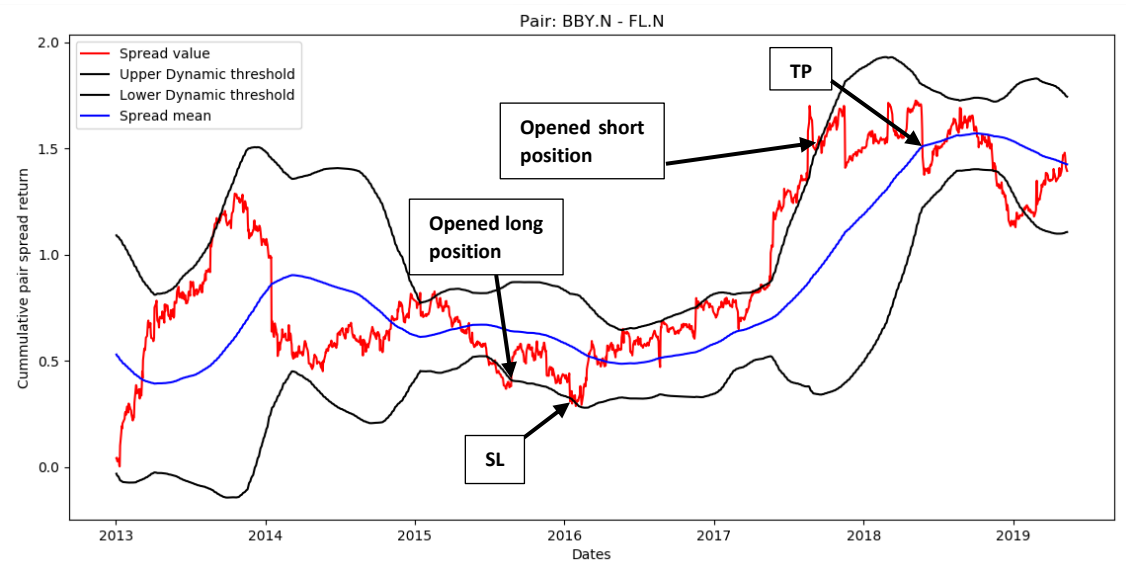
Figure 11 **Graphical illustration of mean reversion criterion**: Black square containing TP (take profit) indicates a short position in the spread closed when its value reaches his mean. Black square containing SL (stop loss) indicates a long position in the spread closed when its value overcome the lower threshold during two consecutive periods.

Summarizing, table 7 shows all possible combination between trade criteria and close position criteria. Remember that traditional and alternative methods use both fixed and conditional thresholds and are applied to both real spread and LSTM prediction values.

Table 7 Strategies resume

Description	Methodology	Threshold signal	MLP signal	Combined threshold and MLP signals
Trade trigger scenarios	Traditional method	X		X
	Alternative Method	X		X
	Alternative Method with stop loss	X		X*
	MLP at least one stock signal		X	X
	MLP both stocks signal		X	
Close criteria	Mean reversion	X		X
	MLP close criterion		X	X
	Mean reversion and MLP close criterion			X

* Only applied to mean-reversion close criterion strategy

42

4.5 Trading costs and returns

One of the main objectives of this study is to show transparency in the strategy results and whether strategy is profitable after discount real market trading costs from operation returns. Therefore, the roundtrip costs applied in this work are commissions applied by real brokerage firms like "Renta 4" to long positions and "IG group" to short positions trading CFD's. Hence, the total roundtrip cost per trade is compute as follows (let's assume a spread long position):

- Open position costs:

$$Long\ pos\ open\ = \max(15\ USD;\ 0.15\% \cdot N_i \cdot OP_{i_1})$$

$$Short\ pos\ open = \max(15\ USD;\ 0.02\ USD \cdot N_j)$$

$$Total\ open\ costs = \ Long\ pos\ open + Short\ pos\ open \qquad (38)$$

- Close position costs:

$$Long\ pos\ close = \max(15\ USD;\ 0.15\% \cdot N_i \cdot OP_{i_2})$$

$$Short\ pos\ close = \max(15\ USD;\ 0.02\ USD \cdot N_j)$$

$$Total\ close\ costs = \ Long\ pos\ close + Short\ pos\ close \qquad (39)$$

Where OP_{i_1} represent stock i open price the day when the trade is opened, OP_{i_2} is the stock i open price the day when the position is closed, N_i is the number of stocks for long position and N_j is the number of stocks for short position. Also, a short position implies the payment of borrowing interests to the real owner of the stock, equal to 2.5% (annual rate).

$$Short\ interests = 2.5\% \cdot \frac{\#days}{365} \cdot N_j \cdot OP_{j_1} \qquad (40)$$

Where $\#days$ represent the number of days elapsed until the position is closed and OP_{j_1} is the opening price of stock j when the position is opened. Finally, the total trading costs is given by the sum of (38), (39) and (40).

$$Total\ trading\ costs = Total\ open\ costs + Total\ close\ costs + Short\ interests \qquad (41)$$

On the other hand, return per trade is defined as follows:

$$r_n = \frac{W_n}{l_n + s_n} \qquad (42)$$

where l_n and s_n represents the long and short position investments in pair trade n respectively, both computed as number of stocks multiplied by stock price when trade was opened. W_n is the net trade income given by

$$W_n = w_{l,n} + w_{s,n} - C_n \tag{43}$$

with $w_{l,n}$ and $w_{s,n}$ the net long and short positions income respectively, and C_n the total trading costs provided by equation ([41](#)) for trade n. Notice that it was not assume a zero (or close to zero in this case given by including beta) net investment provided by compensation between long and short position. In theory, when short position is taken the investor receive money, but in practice the broker requires a margin account as guarantee, this implies that investor must leave cash in the bank until trade is closed. It was assumed that the cash required by the broker is the short position investment s_n and its addition with long position investment l_n represent the total investment amount in trade n.

5. Results

This section discusses the results obtained applying the methodology discussed in chapter 4.

5.1 Pair selection

As discussed in section 4.1, the pairs are selected based on their supercointegration relationship. S&P 500 index has 505 companies from 63 different industries, however only 13 industries have supercointegrated pairs. Number of pairs per industry can be found in table 8.

Table 8 Breakdown of the number of selected pairs per industry

INDUSTRY NAME	# stocks	# of possible pairs	# Supercointegrated pairs
Aerospace & Defense	12	66	1
Banks	19	171	2
Capital Markets	21	210	2
Chemicals	15	105	2
Electric Utilities	14	91	2
Equity Real Estate Investment Trusts (REITs)	31	465	2
Food Products	13	78	1
Health Care Equipment & Supplies	19	171	1
Household Products	5	10	1
Life Sciences Tools & Services	7	21	2
Machinery	15	105	1
Pharmaceuticals	10	45	1
Specialty Retail	15	105	3
TOTAL			**21**

Where *# stocks* represent the number of stocks per industry, *# of possible* pairs are the maximum number of possible combinations of pairs that can be constructed per industry and the *# supercointegrated pairs column* represents the pairs that pass the supercointegration analysis described in section 4.1, i.e., pairs who approved both Johannsen and Engle-granger test.

However, before running the MLP neural network are dropped supercointegrated pairs with stocks without complete fundamental ratios historical data[10]. Finally, 15 pairs from 12 different industries are used to compute trading results.

---

[10] Supercointegrated pairs dropped: DUK.N & XEL.OQ, AMG.N & RJF.N, FRT.N & IRM.N, STI.N & ZION.OQ, MTD.N & TMO.N and TMO.N & WAT.N.

5.2 Strategy performance

Using the 15 supercointegrated pairs, is computed each strategy returns following the methodology explained in section 4. The tables in this section reports the total number of trades, the total returns before trading costs, the average annual returns before costs, the average trading cost per trade (in USD), the total returns after costs and the average annual returns including costs for the different strategies proposed.

5.2.1 Threshold signals

Following the threshold criterion, the strategy returns resume using the real spread value are displayed in table 9, and the results using the LSTM predictions are shown in table 10.

Table 9 Threshold strategy results using real spread values during the entire out-of-sample period.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	97	55.22%	8.62%	146.46	-19.71%	-3.08%
Alternative (Fixed)	97	39.93%	6.23%	132.86	-28.14%	-4.39%
Alternative with stop loss (Fixed)	283	70.00%	10.93%	76.74	-45.02%	-7.03%
Traditional (Dynamic)	106	-10.82%	-1.69%	147.97	-93.63%	-14.62%
Alternative (Dynamic)	105	-22.55%	-3.52%	140.86	-100.67%*	-15.72%
Alternative with stop loss (Dynamic)	322	33.11%	5.17%	82.24	-106.63%*	-16.65%

* For trading strategy execution, it was supposed that investor has the necessary initial investment amount to execute each trade. If the strategy initial investment value were accumulative, the net total return must be -100%. It was used the real cumulative trade returns for comparative purpose, and it is used for all strategies from now on.

Table 9 shows a better performance using fixed threshold bands respect to the dynamic (conditional volatility) threshold with the real pairs spread values to compute the dynamic mean and standard deviation. Even though there are positive annual returns without trading costs, none of the strategies gives positive returns after costs. Notice that the stop loss, that should reduce the losses generate unexpectedly worst returns than the alternative methods without stop loss.

Table 10 Threshold strategy results using LSTM predictions during the entire out-of-sample period.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	81	-25.09%	-3.92%	190.25	-106.82%	-16.68%
Alternative (Fixed)	79	-50.67%	-7.91%	158.79	-116.97%	-18.27%
Alternative with stop loss (Fixed)	171	-17.04%	-2.66%	87.96	-96.39%	-15.05%
Traditional (Dynamic)	83	-18.85%	-2.94%	183.56	-99.83%	-15.59%
Alternative (Dynamic)	83	-65.12%	-10.17%	170.72	-140.27%	-21.90%
Alternative with stop loss (Dynamic)	187	-15.23%	-2.38%	102.38	-116.69%	-18.22%

Table 10 shows the results using LSTM predictions instead of real spread value. The results got worst in all cases and none of the strategies indicates positive returns neither before trading costs. The stop loss applied to the network predictions works better than applied to real spread. The results until now exhibit better performance using real spread with fixed threshold, something unexpected given that there are works in previous literature stating that dynamic threshold works better. Also, the traditional method (trade when spread reaches by first time the threshold) outperforms the alternative method proposed when trades are triggered by threshold signals.

5.2.2 Threshold and fundamental signals

Following both thresholds and stocks trend predictions given by the MLP network to trade and the three close position criteria, the strategy returns are summarized below:

> *Closing with mean reversion criterion*

Table 11 Threshold and fundamental signal trades using real spread values during the entire out-of-sample period and closing position when spread reaches his mean.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	26	71.92%	11.23%	130.08	53.55%	8.36%
Alternative (Fixed)	19	14.59%	2.28%	119.00	2.22%	0.35%
Alternative with stop loss (Fixed)	34	26.29%	4.11%	75.95	12.27%	1.92%
Traditional (Dynamic)	28	40.77%	6.37%	134.99	20.39%	3.18%
Alternative (Dynamic)	22	12.08%	1.89%	134.35	-3.85%	-0.60%
Alternative with stop loss (Dynamic)	41	18.77%	2.93%	81.51	1.01%	0.16%

Table 11 illustrates the results with real spread, where traditional method beats again the alternative proposal and fixed threshold continue being the best choice. These results show that combining the standard pairs trading criterion to trade using the spread value and stocks trend predictions based on fundamental ratios improves the pairs trading returns. The number of trades was significantly reduced, showing a better trade trigger with the fundamental's inclusion.

Table 12 Threshold and fundamental signal trades using LSTM predictions during the entire out-of-sample period and closing position when the prediction reaches its mean.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	35	105.93%	16.54%	166.38	74.06%	11.56%
Alternative (Fixed)	22	3.28%	0.51%	156.26	-15.46%	-2.41%
Alternative with stop loss (Fixed)	33	-11.22%	-1.75%	91.01	-27.25%	-4.26%
Traditional (Dynamic)	29	131.04%	20.46%	153.50	106.43%	16.62%
Alternative (Dynamic)	20	61.87%	9.66%	138.64	47.26%	7.38%
Alternative with stop loss (Dynamic)	33	56.79%	8.87%	88.39	41.39%	6.46%

LSTM predictions combined with MLP signal predictions outperform standard threshold method too and got better results than using real spread with dynamic threshold. This method shows positive annual returns after including trading costs especially for traditional method with dynamic bands, to determine the reason of this high returns we look at the trades executed and find that one particular trade generates 78.23% of returns which implies that this single trade represent the most part of all strategy returns (equal to 106.43% during the entire out-of-sample period). Removing this trade from results, we got a total return of 28.20% and an annual average return of 4.40% after trading cost. Traditional fixed method also had a trade with extremely high return (78.23%), which is greater than the strategy net total return, i.e., this strategy gave positive returns thanks to that trade. For the alternative methods with dynamic thresholds, even though the trade with high returns represent the 29.77% of total returns which isn't too high, this implies that more than the half of total returns are provided by the for both strategies, giving less robustness to results.

Summarizing, after removing the single trade with high returns, more reliable and robust results are obtained and still LSTM network with MLP signals improved the results from real spread using dynamic thresholds, but the highest annual returns until now are obtained with real spread using traditional method with fixed threshold, with an average annual returns of 8.36%. Details about the trades executed to get this return can be found in appendix 3. Also, the 76.92% of the trades with this method generated benefits, making more robust the results.

> ➢ Closing with MLP signal

Table 13 indicates real spread strategies results resume. The results with this closing criterion for real spread are worse than closing with mean reversion except with alternative method with fixed threshold.

Table 13 Threshold and fundamental signal trades using real spread values during the entire out-of-sample period and closing position when both stocks signal predictions are different respect to the taken position when the trade was opened.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	36	26.92%	4.20%	99.13	7.78%	1.22%
Alternative (Fixed)	18	38.64%	6.03%	119.44	27.26%	4.26%
Traditional (Dynamic)	33	5.16%	0.81%	101.33	-12.65%	-1.98%
Alternative (Dynamic)	25	-1.00%	-0.16%	106.88	-15.12%	-2.36%

Table 14 indicates LSTM predictions strategies results. With this closing criterion, it is observed an important improvement for all strategies respect to mean-reversion closing criterion.

Table 14 Threshold and fundamental signal trades using LSTM predictions during the entire out-of-sample period and closing position when both stocks signal predictions are different respect the taken position when the trade was opened.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	51	182.19%	28.45%	94.12	156.20%	24.39%
Alternative (Fixed)	23	77.59%	12.12%	107.13	64.32%	10.04%
Traditional (Dynamic)	32	184.06%	28.74%	98.50	166.88%	26.06%
Alternative (Dynamic)	20	100.38%	15.67%	106.44	88.99%	13.90%

However, these high returns are affected by one single trade who inflates total returns. Removing that trade from traditional fixed (single trade return = 134.51%), traditional dynamic (single trade return = 134.51%) and alternative dynamic (single trade return = 65.73%), the total returns over the entire out-of-sample period dropped to 21.69%, 32.37% and 23.26%, giving an annual average return of 3.39%, 5.05% and 3.63% respectively.

For the alternative fixed method, it was not removed that trade because his return was 33.31%, which is a high but common return. However, is important to mention that two trades represent almost 60% of the total return of this strategy (equal to 64.32%) during all out-of-sample period. The trades made to get this return can be found in appendix 4. Even though this method generates greater annual returns than the traditional with fixed threshold for the spread, the proportion of trades with benefits is 52.17%, lower than the 76.92% from the method with 8.36% annual returns.

➢ Closing with both mean-reversion and MLP signal criteria

Table 15 Threshold and fundamental signal trades using real spread values during the entire out-of-sample period and closing position when both stocks signal predictions are different respect to the taken position when the trade was opened or when the spread reaches his mean.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	37	6.86%	1.07%	84.20	-9.93%	-1.55%
Alternative (Fixed)	19	15.94%	2.49%	89.17	6.88%	1.07%
Traditional (Dynamic)	33	-20.07%	-3.13%	87.15	-35.45%	-5.54%
Alternative (Dynamic)	25	-26.94%	-4.21%	88.05	-38.67%	-6.04%

Table 15 indicates real spread strategies returns, where the annual returns are worst in all cases than the closing based on MLP signal only. This is a more conservative strategy since the trades can be closed when the spread reaches his mean even if the MLP signals still indicates trend on one stock (or both). Therefore, it was expected that this strategy shows lower returns respect to the previous one.

Table 16 Threshold and fundamental signal trades using LSTM predictions during the entire out-of-sample period and closing position when both stocks signal predictions are different respect to the taken position when the trade was opened or when the prediction reaches his mean.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
Traditional (Fixed)	53	97.75%	15.26%	82.33	74.10%	11.57%
Alternative (Fixed)	24	33.46%	5.23%	80.88	23.05%	3.60%
Traditional (Dynamic)	33	100.53%	15.70%	85.62	85.16%	13.30%
Alternative (Dynamic)	21	38.45%	6.00%	85.97	28.85%	4.51%

Table 16 indicates returns using LSTM predictions, where we got lower returns than previous closing method using MLP signal. Removing one trade with 78.23% return from results of traditional method with dynamic threshold, the annual return drops to 1% and is negative for traditional method with fixed threshold.

Summarizing, for real spread this method is the worst of all, giving positive results only in one case and being this annual return very low. For LSTM predictions, even though only one case shows negative returns (after removing high return trade), the results are lower than previous close criteria. However, with the three closing criteria displayed in this section it has been improved the threshold method results from section 5.2.1, showing that adding fundamental ratios to pairs trading improves the strategy performance.

5.2.3 MLP network predictions criterion

Applying the trade criterion explained in section 4.4.2.2, the strategy performance is the following:

Table 17 MLP trades using MLP signals during the entire out-of-sample period and closing positions when both stocks have different signals respect to the position taken in them when the trade was opened.

TRADING CRITERIA	Trades	Gross total returns	Gross average returns	Average cost (USD)	Net total returns	Net average returns
MLP at least one stock signal	188	297.65%	46.48%	96.41	201.27%	31.43%
MLP both stocks signal	8	24.90%	3.89%	133.91	19.31%	3.01%

The results in table 17 illustrate strategy returns using only the MLP network predictions to open and close positions. Opening trades when both stocks shows opposite signals got positive returns, nevertheless, the number of trades is too low for a strategy applied to more than 6 years' period.

Trading when at least one stock has trend shows impressive returns. Details about the trades can be found in appendix 5. However, there is one trade with 134.51% return which generates an overrate in total returns over the entire out-of-sample period. Removing this trade, total returns of the strategy drops from 201.27% to 66.76% and the annual return drops to 10.42%. Nevertheless, this return is greater than all strategies tested until now and the elevate number of trades gives robustness to the results.

Also, is important to mention that the number of trades with positive returns are 81, which implies that 43.09% of the trades gives benefits. Even though the proportion of positive trades are lower than the traditional method with fixed threshold applied to real spread (76.92%), the profits from good trades compensates the losses from bad trades. In fact, there are a lot of trades with very small loses, this can happen because when the market moves against our position, the network prediction corrects its signal and close the position.

Summarizing, the most profitable strategy was the MLP signal prediction, showing consistent results during out-of-sample period with a high number of trades and an average annual return of 10.42% removing the trade with extreme high return (31.43% including the trade). Top 5 strategies with highest net average returns are shown in table 18.

Table 18 Ranking of strategies with highest net average returns

Trade criterion	Close criterion	Trade trigger	Spread value	Trades	Net total returns	Net average returns
MLP signal	MLP signal	MLP at least one stock signal	N/A	188	201.27%	31.43%
Combined method	MLP signal	Traditional (Dynamic threshold)	LSTM prediction	32	166.88%	26.06%
Combined method	MLP signal	Traditional (Fixed threshold)	LSTM prediction	51	156.20%	24.39%
Combined method	Mean reversion	Traditional (Dynamic threshold)	LSTM prediction	29	106.43%	16.62%
Combined method	MLP signal	Traditional (Dynamic threshold)	LSTM prediction	20	88.99%	13.90%

Also, traditional method outperforms the alternative method proposed in this work, showing better returns when the trade is executed when spread reaches by first time a threshold. Moreover, the returns using threshold approach was negative and were improved when it was included MLP network predictions to trade trigger, indicating positive gross returns in most cases. Table 19 illustrates the top 5 strategies with lowest net average returns, notice that all the strategies in the ranking use threshold trade criterion.

Table 19 Ranking of strategies with lowest net average returns

Trade criterion	Close criterion	Trade trigger	Spread value	Trades	Net total returns	Net average returns
Thresholds	Mean reversion	Alternative (Dynamic)	LSTM prediction	83	-140.27%	-21.90%
Thresholds	Mean reversion	Alternative (Fixed)	LSTM prediction	79	-116.97%	-18.27%
Thresholds	Mean reversion	Alternative with stop loss*	LSTM prediction	187	-116.69%	-18.22%
Thresholds	Mean reversion	Traditional (Fixed)	LSTM prediction	81	-106.82%	-16.68%
Thresholds	Mean reversion	Alternative with stop loss*	Real Spread	322	-106.63%	-16.65%

* Alternative method with dynamic threshold

Finally, even though the stop loss criterion generates less average roundtrip costs, this generates a higher number of trades and that trades returns doesn´t compensate the benefits obtained without this close position criterion.

Furthermore, with all strategies was notice that long positions in the stocks outperforms the short positions performance, giving long positions always positive total returns and short positions negative total returns after costs. This happens because during out-of-sample period, the United States market was in bullish trend. However, if market conditions were different, i.e., market trend was bearish, short positions should perform better than long trades. It can be concluded that pairs trading allows us to compensate part of market risk.

6. Conclusions and recommendations

The main objectives of this work were applying neural network techniques to pairs trading using stocks and prove whether company's fundamental ratios can improve pairs trading strategies performance. It was proposed a LSTM neural network for spread prediction and MLP neural network for multi-classification purpose, giving trend signals and using the outputs as trade trigger in the strategy. It was compared the trading strategies proposed against the standard strategy used in previous pairs trading literature where the authors trade trigger was volatility thresholds, testing in this work two of the most common approaches, the fixed and conditional volatility threshold with both real spread value and LSTM spread predictions.

As result, it was found that standard threshold strategies weren't profitable after roundtrip costs, however, the inclusion of MLP network trend predictions improves the performance of the strategy, giving positive returns in most cases and reducing the number of trades executed respect to the standard threshold strategy, showing that the inclusion of trend signals improves trades trigger criterion. Therefore, it can be concluded that adding fundamental ratios generates a more profitable trading strategy than the traditional threshold method based only in the spread value to trade.

Moreover, the combination of threshold and fundamental signal criteria stated in [section 5.2.2](#) shows better results for the MLP signal close position criterion applied to LSTM spread predictions and using the traditional method with dynamic threshold to trade, obtaining an annual average return of 26.06% and 24.39% for same method with fix threshold. However, the results are inflated by one trade with extreme high return, distorting the strategy performance. Removing the trade with high return from all strategies, the most profitable strategy is the alternative method with fix threshold applied to LSTM predictions and MLP signal close criterion with 10.04% of annual average return, followed by the traditional method with fix threshold applied to real spread value and closing when spread mean-reverts with 8.36% of annual average return.

Finally, the proposal of trade using MLP signals outperforms the other strategies, showing more robustness given by the number of trades and the return after roundtrip costs. This strategy, that trades when at least one stock shows trend signal, had a net average annual return of 31.43% executing 188 trades to get this result. As the previous strategies, was removed a single trade who inflates results to see whether the strategy gives consistent results without this extraordinary trade and the average annual return drops to 10.42%. Even though the returns

had an important reduction after removing that trade, shows better returns respect to the other strategies and is a new way to trade with pairs trading.

Comparing results against previous literature about neural network techniques applied to pairs trading, it was found that it was improved the 11.07% annual return obtained by Van der Have (2017)[35] which did not included trading costs. Further, the strategy stated by Dunis et al. (2015)[10] generates 26.46% of total annual return after roundtrip costs applying GPA (genetic programming algorithm) to corn-ethanol crush spread, showing that the strategy proposed in this work beats that methodology in return terms. However, the leverage trading structure used by Dunis et al. (2015)[10] improves results from 26.46% to 33.92% total annual return after roundtrip costs and gives greater results respect to this work strategy. The difference it's maybe explained by the way it was computed the roundtrip costs, Dunis et al. (2015)[10] didn't specify explicitly how they compute trading costs, but the investigation is built on earlier work carried out by Dunis et al. (2006)[9] who investigate soybean-oil crush spread and where transactions costs are calculated from an average of five bid–ask spreads taken from different times of the trading days, and commission fees are not considered. Also, trading strategy proposed by Dunis et al. (2015)[10] consists on buy and sell one contract of corn and ethanol according to the position to take in the pair. Here it is important to see how strategy returns are calculated, given that if they are assuming short position gives money and they are taking as initial investment amount the difference between long and short position, return would be inflated based on that wrong assumption and it would not be comparable with this work results, where it was assumed short position guarantees as the initial investment required to go ahead with the trade. In any case, it would be interesting the addition of a similar leverage structure to the strategy proposed in this work in future researches.

Also, it is recommended for future researches use a rolling window to compute the cointegration relationship because the supercointegration relation obtained with in-sample data does not implies supercointegration during all out-of-sample period. In addition, can be interesting use a BECC model to estimate a dynamic beta to test the strategy. Moreover, would be interesting to prove other threshold methods with the trading criteria proposed in this work, for example a model able to include the possible asymmetric performance of spread series to compute the thresholds (E-GARCH or GJR models can be useful for this purpose). Also, it can be useful the addition of the Wilcoxon test to determine whether the strategies returns are significantly different from zero at different confidence levels to see whether a trading strategy gives profits just by outlier trades returns. Additionally, it is highly recommended the addition of risk

measures to the strategy and it would be interesting use performance measures (sharpe ratio, Jensen alpha, Treynor ratio, sortino ratio, etc.) to determine which strategy works better.

Furthermore, it is recommended create a program for optimal selection of network architecture, testing different number of layers, neurons, changing learning parameters, etc. to find the best architecture to train both network models. Our network architectures are one of many possible and there is no guarantee that we have selected the best ones, these means that results could be improved with other network parameters.

7. Bibliography

1. Álvarez Roa, A. (2018). *Pairs trading: Optimal threshold strategies*. Interuniversitary program: UPV/EHU - UCM - UV - UCLM.

2. Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning Long-Term Dependencies with Gradient Descent is Difficult. *IEEE Transactions on Neural Networks*.

3. Caldeira, J. F., & Moura, G. V. (2013). Selection of a Portfolio of Pairs Based on Cointegration: A Statistical Arbitrage Strategy (Seleção de uma Carteira de Pares de Ações Usando Cointegração: Uma Estratégia de Arbitragem Estatística). *Revista Brasileira de Finanças*.

4. Cummins, M., & Bucca, A. (2012). Quantitative spread trading on crude oil and refined products markets. *Quantitative Finance*, *12*(12), 1857–1875.

5. Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*.

6. Do, B., & Faff, R. (2009). Does Naïve Pairs Trading Still Work ?*. *Financial Analyst Journal*.

7. Do, B., & Faff, R. (2010). Does simple pairs trading still work? *Financial Analysts Journal*.

8. Do, B., Faff, R., & Hamza, K. (2006). A new approach to modeling and estimation for pairs trading. *Proceedings of 2006 Financial Management Association European Conference*.

9. Dunis, C. L., Laws, J., & Evans, B. (2006). Modelling and trading the soybean-oil crush spread with recurrent and higher order networks: A comparative analysis. *Neural Network World*.

10. Dunis, C. L., Laws, J., Middleton, P. W., & Karathanasopoulos, A. (2015). Trading and hedging the corn/ethanol crush spread using time-varying leverage and nonlinear models. *European Journal of Finance*, *21*(4), 352–375.

11. Dunis, C. L., Laws, J., & Evans, B. (2006). Modelling and trading the gasoline crack spread: A non-linear story. *Derivatives Use, Trading & Regulation*, *12*, 126–145.

12. ELLIOT, ROBERT J.; VAN DER HOEK*, JOHN; MALCOLM, W. P. (2005). Pairs trading. *Quantitative Finance*, *5*(3), 271–276.

13. Elman, J. L. (1990). Finding structure in time. *Cognitive Science*.

14. Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*(2), 251–276.

15. Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*, 1–22.

16. Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3–56.

17. Figuerola-Ferretti, I., Serrano, P., Tang, T., & Vaello-Sebastià, A. (2017). *Supercointegrated*. *SSRN*.

18. Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (1999). *Pairs trading: Performance of a relative-value arbitrage rule*.

19. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*.

20. Huang, C.-F., Hsu, C.-J., Chen, C.-C., Chang, B. R., & Li, C.-A. (2015). An Intelligent Model for Pairs Trading Using Genetic Algorithms. *Computational Intelligence and Neuroscience*.

21. Huck, N. (2009). Pairs selection and outranking: An application to the S&P 100 index. *European Journal of Operational Research*, *196*(2), 819–825. Retrieved from http://dx.doi.org/10.1016/j.ejor.2008.03.025

22. Huck, N. (2010). Pairs trading and outranking: The multi-step-ahead forecasting case. *European Journal of Operational Research*, *207*(3), 1702–1716. Retrieved from http://dx.doi.org/10.1016/j.ejor.2010.06.043

23. Huck, N., & Afawubo, K. (2015). Pairs trading and selection methods: is cointegration superior? *Applied Economics*.

24. Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica: Journal of the Econometric Society*, 1551–1580.

25. Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *arXiv preprint*.

26. Krauss, C. (2015). *Statistical arbitrage pairs trading strategies: Review and outlook*. Nurnberg.

27. Montana, G., & Parrella, F. (2009). Data mining for algorithmic asset management. In *Data Mining for Business Applications* (pp. 283–295).

28. Pinto, J. E., Robinson, T. R., & Stowe, J. D. (2019). Equity valuation: A survey of professional practice. *Review of Financial Economics*, *37(2)*, 219–233.

29. Piotroski, J. D. (2006). Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research*.

30. Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*.

31. Rudy, J., Dunis, C., Giorgioni, G., & Laws, J. (2010). *Statistical Arbitrage and High-Frequency Data with an Application to Eurostoxx 50 Equities*. *SSRN*.

32. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*.

33. Sundermeyer, M., Alkhouli, T., Wuebker, J., & Ney, H. (2014). Translation Modeling with Bidirectional Recurrent Neural Networks Human Language Technology and Pattern Recognition Group. *Emnlp2014*.

34. Thomaidis, N. S., Kondakis, N., & Dounias, G. D. (2006). An intelligent statistical arbitrage trading system. In *Lecture Notes in Computer Science (including subseries Lecture Notes*

in Artificial Intelligence and Lecture Notes in Bioinformatics).

35. Van der Have, R. W. J. (2017). *Pairs Trading Using Machine Learning: An Empirical Study*. Erasmus University Rotterdam.

36. Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*. (J. Wiley, Ed.).

8. Appendix

A.1 Dropped stocks

Table 20 Dropped stocks tickets

TICKETS						
PRGO.N	PFG.OQ	PLD.N	SRE.N	NKTR.OQ	ETFC.OQ	GOOG.OQ
FOX.OQ	KEYS.N	MSCI.N	REGN.OQ	ANTM.N	VRSK.OQ	BKNG.OQ
LIN.N	TTWO.OQ	LYB.N	HCA.N	GOOGL.OQ	AMT.N	STZ.N
AIZ.N	RMD.N	IQV.N	AVGO.OQ	GM.N	PYPL.OQ	ABBV.N
MCO.N	CMG.N	VRTX.OQ	NWSA.OQ	CCI.N	SYF.N	MA.N
MOS.N	CHTR.OQ	BLK.N	CME.OQ	CXO.N	LKQ.OQ	DISCK.OQ
BWA.N	PWR.N	ABMD.OQ	FANG.OQ	A.N	MAR.OQ	OMC.N
VTR.N	PM.N	HII.N	FDX.N	GS.N	MNST.OQ	CRM.N
IVZ.N	UAL.OQ	ZBH.N	PKI.N	MET.N	EW.N	UA.N
STX.OQ	KHC.OQ	GILD.OQ	TWTR.N	WYNN.OQ	LLL.N	TAP.N
VRSN.OQ	LW.N	JWN.N	ANSS.OQ	INFO.OQ	FRC.N	V.N
LB.N	OXY.N	TMK.N	ULTA.OQ	NLSN.N	APH.N	JNPR.N
FB.OQ	FBHS.N	AWK.N	KO.N	CFG.N	ZTS.N	FTNT.OQ
EXPE.OQ	BR.N	PKG.N	WLTW.OQ	DWDP.N	DFS.N	FIS.N
CE.N	TEL.N	FLT.N	ANET.N	HFC.N	APTV.N	DAL.N
COTY.N	KMI.N	IT.N	WCG.N	EIX.N	MPC.N	HBI.N
ARE.N	DVN.N	ALLE.N	GPN.N	VIAB.OQ	DISCA.OQ	DOW.N
UAA.N	MXIM.OQ	FTV.N	RSG.N	NWS.OQ	ICE.N	NTRS.OQ
RHT.N	AAL.OQ	PRU.N	XEC.N	UPS.N	SWKS.OQ	ACN.N
ISRG.OQ	MKC.N	CF.N	HLT.N	TRIP.OQ	EXR.N	XYL.N
TPR.N	SBAC.OQ	ILMN.OQ	NFLX.OQ	EQIX.OQ	IPGP.OQ	HAS.OQ
ADS.N	AMP.N	WU.N	RF.N	EBAY.OQ	WRK.N	NRG.N
CBRE.N	GRMN.OQ	SJM.N	PXD.N	BIIB.OQ	FOXA.OQ	DLR.N
SEE.N	NDAQ.OQ	HPE.N	DG.N	CTSH.OQ	AKAM.OQ	PSX.N
CPRI.N	FFIV.OQ	AAP.N	TDG.N	NCLH.N	TSCO.OQ	FLIR.OQ
MDLZ.OQ	NVDA.OQ	ALGN.OQ	CNC.N	CBOE.Z	FTI.N	

A.2 Supercointegrated pairs

Table 21 Supercointegrated pairs list

Stock 1	Company name	Stock 2	Company name	Industry	Beta	Trace test	Eigen test	Engle-granger p-value
XRAY.OQ	Dentsply Sirona	SYK.N	Stryker	Health Care Equipment & Supplies	1.1626	Approved	Rejected	0.0088
DUK.N	Duke Energy Corporation	XEL.OQ	Xcel Energy	Electric Utilities	0.7779	Approved	Approved	0.0091
EXC.N	Exelon Corporation	PPL.N	PPL Corporation	Electric Utilities	0.8576	Approved	Approved	0.0004
AMG.N	Affiliated Managers Group	RJF.N	Raymond James Financial	Capital Markets	0.7338	Approved	Approved	0.0001
BEN.N	Franklin Resources	RJF.N	Raymond James Financial	Capital Markets	0.9824	Approved	Approved	0.0089
FRT.N	Federal Realty Investment Trust	IRM.N	Iron Mountain	Equity Real Estate Investment Trusts	0.7150	Approved	Approved	0.0007
HCP.N	HCP	O.N	Realty Income	Equity Real Estate Investment Trusts	0.9745	Approved	Approved	0.0004
APD.N	Air Products and Chemicals	ECL.N	Ecolab	Chemicals	0.8629	Approved	Approved	0.0067
IFF.N	International Flavors & Fragrances	PPG.N	PPG Industries	Chemicals	0.9342	Approved	Approved	0.0003
LLY.N	Eli Lilly and Company	PFE.N	Pfizer	Pharmaceuticals	0.8232	Approved	Approved	0.0020
CPB.N	Campbell Soup	TSN.N	Tyson Foods	Food Products	0.7702	Approved	Rejected	0.0071
GD.N	General Dynamics	UTX.N	United Technologies	Aerospace & Defense	0.9815	Approved	Approved	0.0022
BBY.N	Best Buy	FL.N	Foot Locker	Specialty Retail	0.7879	Approved	Approved	0.0011
FL.N	Foot Locker	LOW.N	Lowe's Companies	Specialty Retail	1.1017	Approved	Approved	0.0012
ORLY.OQ	O'Reilly Automotive	TJX.N	The TJX Companies	Specialty Retail	0.5625	Approved	Approved	0.0034
MTB.N	M&T Bank	WFC.N	Wells Fargo & Company	Banks	0.7520	Approved	Approved	0.0053
STI.N	SunTrust Banks	ZION.OQ	Zions Bancorporation	Banks	0.9627	Approved	Approved	0.0030
ITW.N	Illinois Tool Works	SNA.N	Snap-on Incorporated	Machinery	0.9743	Approved	Rejected	0.0018
CLX.N	The Clorox Company	PG.N	The Procter & Gamble	Household Products	0.9951	Approved	Approved	0.0005
MTD.N	Mettler-Toledo International	TMO.N	Thermo Fisher Scientific	Life Sciences Tools & Services	0.8576	Approved	Approved	0.0001
TMO.N	Thermo Fisher Scientific	WAT.N	Waters Corporation	Life Sciences Tools & Services	1.0782	Approved	Approved	0.0002

A.3 Trades details 1

Table 22 Trades for real spread values applying traditional method with fixed threshold and fundamental signals, closing positions with mean-reversion criterion

Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
BEN.N - RJF.N	Long	25/1/2013	3/4/2013	247	40.55	45.87	233	43.76	45.17	4.87%	94.14	12.79%	-3.83%	4.40%	890.12
BEN.N - RJF.N	Short	13/5/2013	21/6/2013	230	42.59	42.86	206	49.41	42.40	7.54%	79.52	0.33%	13.69%	7.14%	1,425.72
BEN.N - RJF.N	Short	21/10/2013	12/11/2013	223	43.84	46.11	204	49.07	47.32	4.35%	71.39	4.87%	3.14%	3.99%	790.24
ITW.N - SNA.N	Short	27/1/2014	7/2/2014	94	100.30	106.19	123	79.41	78.78	3.29%	66.02	5.55%	0.42%	2.94%	565.13
HCP.N - O.N	Long	11/2/2014	31/7/2014	285	34.69	38.03	240	40.98	43.26	2.04%	140.75	9.29%	-6.68%	1.33%	261.66
ORLY.OQ - TJX.N	Short	25/4/2014	9/2/2016	190	29.10	33.77	68	147.51	231.90	-31.18%	369.85	15.49%	-60.60%	-33.56%	-5,222.02
GD.N - UTX.N	Short	5/5/2014	12/3/2015	84	116.10	120.41	90	110.52	132.72	-8.30%	206.65	3.40%	-21.86%	-9.35%	-1,842.61
LLY.N - PFE.N	Short	4/12/2014	13/2/2015	261	31.55	34.74	142	70.98	70.69	4.77%	93.14	9.75%	-0.22%	4.26%	780.63
IFF.N - PPG.N	Long	2/1/2015	13/2/2015	97	101.42	115.34	80	115.58	116.85	6.54%	80.15	13.40%	-1.62%	6.12%	1,168.49
EXC.N - PPL.N	Long	23/9/2015	8/3/2016	328	28.75	33.90	274	31.11	35.67	2.45%	128.24	17.58%	-15.79%	1.74%	311.52
CLX.N - PG.N	Short	30/10/2015	8/3/2016	129	77.08	82.92	81	122.72	123.35	3.53%	120.28	7.26%	-1.41%	2.93%	582.05
MTB.N - WFC.N	Long	11/1/2016	30/3/2016	88	110.18	111.35	148	50.11	48.45	2.04%	87.43	0.75%	2.54%	1.53%	261.21
MTB.N - WFC.N	Short	20/5/2016	8/9/2017	154	48.59	49.58	86	116.57	142.36	-11.80%	285.22	1.64%	-24.67%	-13.43%	-2,350.70
FL.N - LOW.N	Long	16/6/2016	22/8/2016	182	54.17	68.66	143	76.86	78.15	11.76%	98.37	26.41%	-2.27%	11.29%	2,354.34
ITW.N - SNA.N	Short	28/7/2016	13/1/2017	62	155.75	177.93	86	115.47	122.61	3.89%	141.13	13.91%	-7.29%	3.17%	619.99
IFF.N - PPG.N	Short	31/10/2016	5/1/2017	99	93.35	96.50	74	134.89	116.94	8.53%	90.77	3.05%	12.70%	8.06%	1,549.38
CPB.N - TSN.N	Short	2/12/2016	17/2/2017	135	56.86	64.71	176	57.39	58.41	4.95%	95.97	13.42%	-2.43%	4.41%	784.26
ORLY.OQ - TJX.N	Long	12/7/2017	30/11/2017	57	174.97	236.59	163	34.39	37.16	19.65%	103.23	34.86%	-9.27%	18.99%	2,957.60
APD.N - ECL.N	Short	5/1/2018	9/4/2018	63	135.90	140.63	59	168.00	161.23	3.78%	102.77	3.13%	3.30%	3.22%	594.65
BEN.N - RJF.N	Long	6/2/2018	26/10/2018	269	34.67	29.00	105	87.51	73.76	-0.45%	175.17	-16.69%	14.13%	-1.39%	-257.84
HCP.N - O.N	Long	27/2/2018	2/7/2018	447	22.46	25.81	193	50.68	53.75	4.57%	120.65	14.59%	-6.96%	3.96%	784.29
CLX.N - PG.N	Short	1/3/2018	31/10/2018	123	78.35	88.69	76	129.08	146.00	-0.07%	175.59	12.87%	-14.58%	-0.98%	-189.69

Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
GD.N - UTX.N	Short	19/4/2018	30/4/2018	78	124.71	122.53	44	225.55	203.79	4.01%	64.76	-2.06%	9.30%	3.68%	722.64
GD.N - UTX.N	Long	2/8/2018	13/5/2019	50	193.39	170.00	72	133.98	134.03	-6.07%	188.18	-12.40%	-1.68%	-7.05%	-1,361.28
BEN.N - RJF.N	Short	20/11/2018	31/1/2019	125	76.74	80.17	312	31.96	28.79	7.25%	92.13	4.16%	9.30%	6.78%	1,325.66
BBY.N - FL.N	Long	2/1/2019	1/5/2019	193	52.17	74.41	149	52.42	57.23	20.00%	110.51	42.27%	-10.12%	19.38%	3,465.12

A.4 Trades details 2

Table 23 Trades for LSTM predictions applying traditional method with fixed threshold and fundamental signals, closing positions with mean-reversion criterion

Pair	Position	Open date	Close date	# Long stocks	open price	close price	# Short stocks	open price	close price	Gross return	Trade costs	Long return	Short return	Net return	Income (USD)
ITW.N - SNA.N	Long	4/4/2013	6/5/2013	163	61.26	65.59	119	80.87	87.37	-0.35%	75.54	6.76%	-8.50%	-0.73%	-143.25
BBY.N - FL.N	Long	9/4/2013	16/10/2013	402	24.94	39.96	236	33.35	33.09	34.10%	140.83	59.87%	-0.51%	33.31%	5,961.00
BEN.N - RJF.N	Short	3/6/2013	14/11/2013	221	44.26	46.92	204	47.13	48.80	1.27%	136.94	5.70%	-4.66%	0.57%	109.75
BBY.N - FL.N	Short	7/1/2014	3/9/2014	190	41.34	57.66	257	38.14	30.55	28.61%	172.21	39.08%	18.47%	27.64%	4,879.47
HCP.N - O.N	Long	14/1/2014	22/4/2014	283	34.80	37.16	252	38.37	42.09	-1.38%	105.15	6.46%	-10.46%	-1.92%	-375.32
ORLY.OQ - TJX.N	Short	26/8/2014	5/9/2014	189	29.52	30.29	63	156.69	156.85	0.88%	64.73	2.07%	-0.45%	0.46%	70.72
LLY.N - PFE.N	Short	29/12/2014	2/1/2015	261	31.46	31.35	142	69.92	69.74	-0.02%	62.04	-0.72%	-0.07%	-0.36%	-65.19
IFF.N - PPG.N	Long	13/1/2015	19/2/2015	97	102.00	118.27	82	114.17	117.24	6.89%	78.24	15.63%	-3.18%	6.48%	1,248.21
EXC.N - PPL.N	Long	12/10/2015	4/1/2016	323	30.94	27.46	259	33.03	33.93	-7.32%	93.40	-11.55%	-3.47%	-7.82%	-1,450.54
GD.N - UTX.N	Short	23/10/2015	21/12/2015	99	100.46	93.02	69	148.25	138.28	-0.24%	88.03	-7.71%	6.16%	-0.68%	-136.66
IFF.N - PPG.N	Short	1/3/2016	2/3/2016	95	97.55	99.22	96	104.32	105.81	0.08%	60.69	1.39%	-1.73%	-0.23%	-45.08
MTB.N - WFC.N	Short	28/6/2016	3/1/2017	166	45.84	55.67	90	110.38	157.99	-15.12%	148.46	21.05%	-44.33%	-15.97%	-2,801.58
ITW.N - SNA.N	Short	24/10/2016	12/12/2016	61	159.07	176.45	89	113.50	127.02	-0.72%	84.67	10.61%	-12.44%	-1.15%	-227.77
IFF.N - PPG.N	Short	14/11/2016	17/11/2016	96	96.40	96.77	84	118.48	121.19	-1.00%	62.04	0.06%	-2.61%	-1.32%	-254.16
LLY.N - PFE.N	Long	9/12/2016	3/1/2017	149	67.22	74.00	266	31.12	32.69	3.24%	70.07	9.77%	-5.51%	2.86%	522.53
ORLY.OQ - TJX.N	Long	8/1/2018	13/5/2019	38	260.95	360.66	148	38.59	53.58	10.05%	197.37	37.85%	-41.70%	8.78%	1,372.35
APD.N - ECL.N	Short	22/1/2018	23/4/2018	61	139.79	148.22	59	167.64	166.03	3.31%	102.68	5.68%	0.23%	2.75%	506.54
CLX.N - PG.N	Short	9/2/2018	30/7/2018	121	80.69	80.08	78	128.57	132.02	-1.73%	140.37	-1.06%	-3.78%	-2.44%	-483.28
GD.N - UTX.N	Short	12/2/2018	2/8/2018	75	123.67	133.98	48	214.70	193.39	9.17%	144.07	8.01%	8.82%	8.44%	1,652.06
HCP.N - O.N	Long	3/5/2018	17/8/2018	424	23.44	27.07	191	50.72	58.16	0.60%	111.32	15.16%	-15.49%	0.03%	6.76
ITW.N - SNA.N	Short	18/6/2018	23/10/2018	61	159.49	148.04	67	146.89	124.15	4.22%	119.99	-7.49%	14.57%	3.60%	705.14
ITW.N - SNA.N	Long	24/10/2018	14/11/2018	80	125.99	133.91	65	149.11	161.54	-0.88%	71.15	5.98%	-8.75%	-1.24%	-245.50
BEN.N - RJF.N	Short	7/12/2018	13/5/2019	131	77.12	85.35	306	32.48	33.42	3.94%	134.08	10.36%	-3.92%	3.28%	656.41

A.5 Trades details 3

Table 24 Trades using MLP network predictions and at least one stock has trend

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
1	BEN.N - RJF.N	Long	3/1/2013	4/2/2013	256	39.56	42.40	248	40.00	44.90	-2.43%	75.74	6.88%	-12.70%	-2.81%	-563.14
2	APD.N - ECL.N	Short	3/1/2013	16/7/2013	118	74.27	90.83	126	79.47	88.35	4.45%	152.29	21.94%	-12.38%	3.64%	683.51
3	BBY.N - FL.N	Long	3/1/2013	16/10/2013	840	11.60	39.96	239	31.75	33.09	135.65%	198.26	243.95%	-5.97%	134.51%	23,309.17
4	XRAY.OQ - SYK.N	Short	3/1/2013	2/1/2015	209	55.90	94.70	249	40.50	52.22	23.85%	424.64	69.01%	-32.68%	21.90%	4,766.28
5	CPB.N - TSN.N	Short	14/2/2013	4/3/2013	327	23.86	23.10	261	39.00	42.01	-5.75%	67.67	-3.57%	-8.09%	-6.13%	-1,101.80
6	ITW.N - SNA.N	Long	22/2/2013	6/5/2013	159	61.88	65.59	123	78.09	87.37	-2.84%	93.54	5.68%	-12.54%	-3.32%	-645.09
7	HCP.N - O.N	Long	5/3/2013	18/3/2013	224	45.26	44.14	213	46.37	44.50	0.74%	66.30	-2.77%	3.67%	0.41%	81.19
8	IFF.N - PPG.N	Short	15/3/2013	4/4/2013	131	70.99	66.51	132	75.44	74.35	-2.30%	68.87	-6.63%	1.05%	-2.66%	-511.87
9	ORLY.OQ - TJX.N	Long	28/3/2013	4/4/2013	97	102.58	103.32	240	23.39	23.53	0.24%	61.57	0.42%	-1.18%	-0.16%	-24.59
10	HCP.N - O.N	Long	3/4/2013	2/5/2013	217	45.97	48.00	213	45.75	50.68	-3.09%	74.64	4.11%	-11.23%	-3.47%	-684.20
11	IFF.N - PPG.N	Short	24/4/2013	21/6/2013	129	72.50	75.83	130	77.26	77.16	2.28%	88.21	4.27%	-0.45%	1.83%	354.36
12	ORLY.OQ - TJX.N	Long	10/5/2013	5/6/2013	91	109.94	109.49	222	25.39	25.19	0.03%	66.57	-0.71%	0.16%	-0.40%	-62.01
13	BEN.N - RJF.N	Short	13/5/2013	21/5/2013	230	42.59	45.04	206	49.41	50.49	1.71%	64.72	5.44%	-2.52%	1.38%	276.42
14	BEN.N - RJF.N	Short	23/5/2013	14/11/2013	217	43.68	46.92	198	48.77	48.80	3.65%	140.97	7.10%	-1.20%	2.91%	556.68
15	ORLY.OQ - TJX.N	Short	14/6/2013	2/1/2014	224	25.55	31.82	90	111.18	128.42	-0.94%	155.26	24.02%	-16.76%	-1.92%	-302.38
16	IFF.N - PPG.N	Long	24/6/2013	27/6/2014	129	75.46	102.84	123	73.83	100.81	1.14%	223.51	35.93%	-38.61%	-0.05%	-9.41
17	HCP.N - O.N	Short	6/8/2013	10/1/2014	225	42.54	37.80	256	38.97	34.46	0.46%	134.49	-11.46%	10.54%	-0.23%	-45.04
18	APD.N - ECL.N	Short	20/8/2013	5/9/2013	94	91.16	93.42	107	93.67	95.99	-0.19%	67.55	2.13%	-2.85%	-0.56%	-103.41
19	ITW.N - SNA.N	Long	25/9/2013	30/9/2013	131	77.90	76.04	97	100.39	98.64	-0.37%	62.31	-2.68%	1.41%	-0.68%	-136.22
20	CLX.N - PG.N	Short	13/11/2013	14/11/2013	120	82.50	83.50	111	90.63	92.07	-0.20%	60.72	0.91%	-1.89%	-0.50%	-100.56
21	EXC.N - PPL.N	Long	14/11/2013	2/1/2014	360	28.07	27.47	308	28.29	27.94	-0.57%	79.25	-2.44%	0.68%	-0.99%	-186.66
22	ITW.N - SNA.N	Long	15/11/2013	6/12/2013	125	79.72	79.82	91	106.98	105.04	0.96%	69.34	-0.18%	1.41%	0.61%	119.70

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
23	CLX.N - PG.N	Short	15/11/2013	2/1/2014	119	84.17	81.32	108	92.73	92.75	-1.70%	81.29	-3.69%	-0.53%	-2.11%	-422.60
24	ITW.N - SNA.N	Long	10/12/2013	11/12/2013	124	79.77	79.93	91	105.85	105.51	0.26%	60.66	-0.10%	0.00%	-0.05%	-9.88
25	ITW.N - SNA.N	Long	23/12/2013	2/1/2014	121	82.38	83.97	91	107.46	109.28	0.14%	64.26	1.63%	-2.04%	-0.19%	-37.49
26	APD.N - ECL.N	Long	2/1/2014	14/1/2014	96	103.14	100.69	82	103.58	103.31	-1.16%	64.65	-2.68%	-0.15%	-1.51%	-277.71
27	LLY.N - PFE.N	Long	2/1/2014	16/1/2014	195	51.00	53.88	269	30.46	31.16	2.06%	66.37	5.34%	-2.73%	1.69%	306.93
28	BBY.N - FL.N	Short	2/1/2014	3/9/2014	190	41.36	57.66	256	38.89	30.55	29.37%	175.99	39.01%	20.00%	28.38%	5,056.38
29	FL.N - LOW.N	Long	2/1/2014	3/9/2014	241	41.36	57.66	223	49.55	52.70	15.35%	192.99	39.05%	-7.78%	14.43%	3,032.86
30	ORLY.OQ - TJX.N	Long	3/1/2014	13/1/2014	77	128.66	132.92	176	31.64	31.73	2.01%	62.64	3.00%	-0.88%	1.61%	248.66
31	EXC.N - PPL.N	Short	3/1/2014	12/2/2014	306	27.82	28.77	364	26.90	29.07	-2.72%	78.11	3.08%	-8.56%	-3.14%	-575.56
32	ITW.N - SNA.N	Short	3/1/2014	4/1/2016	89	108.73	168.79	119	83.21	90.98	22.58%	408.68	54.85%	-13.09%	20.49%	4,012.03
33	MTB.N - WFC.N	Short	9/1/2014	23/9/2014	165	46.00	52.89	86	116.38	126.51	1.51%	181.34	14.58%	-10.22%	0.48%	84.33
34	HCP.N - O.N	Long	13/1/2014	22/4/2014	290	35.28	37.16	257	38.53	42.09	-1.83%	107.63	5.03%	-10.01%	-2.36%	-476.06
35	APD.N - ECL.N	Long	16/1/2014	30/1/2014	98	102.75	98.44	82	103.65	100.73	-0.98%	65.34	-4.49%	2.40%	-1.34%	-248.11
36	ORLY.OQ - TJX.N	Long	3/2/2014	5/2/2014	76	133.39	133.01	197	28.61	28.00	0.58%	61.14	-0.58%	1.59%	0.19%	30.15
37	CLX.N - PG.N	Short	5/2/2014	11/2/2014	131	75.83	78.00	115	84.41	86.23	0.38%	62.99	2.56%	-2.49%	0.06%	11.98
38	APD.N - ECL.N	Long	19/2/2014	29/12/2014	92	109.28	137.49	83	103.03	106.77	12.28%	191.15	25.48%	-5.47%	11.26%	2,094.28
39	EXC.N - PPL.N	Short	26/2/2014	27/2/2014	290	29.67	29.82	329	30.44	30.33	0.43%	60.69	0.15%	0.05%	0.10%	18.55
40	CPB.N - TSN.N	Short	11/3/2014	9/5/2014	191	40.30	39.13	227	43.85	45.22	-3.03%	88.63	-3.29%	-3.71%	-3.53%	-623.09
41	ORLY.OQ - TJX.N	Short	8/4/2014	5/9/2014	183	29.82	30.29	69	143.03	156.85	-5.67%	130.30	1.01%	-10.68%	-6.52%	-998.79
42	GD.N - UTX.N	Short	5/5/2014	23/7/2014	84	116.10	110.74	90	110.52	120.08	-6.65%	97.47	-4.92%	-9.33%	-7.15%	-1,408.11
43	CPB.N - TSN.N	Short	19/5/2014	29/5/2014	189	40.66	42.26	222	42.39	44.67	-1.19%	64.51	3.54%	-5.75%	-1.57%	-268.27
44	CPB.N - TSN.N	Short	30/5/2014	11/6/2014	182	42.75	36.05	223	45.33	45.60	-7.15%	65.54	-16.06%	-0.95%	-7.52%	-1,345.15
45	HCP.N - O.N	Long	16/6/2014	17/7/2014	265	37.69	38.06	222	43.85	45.14	-0.95%	74.80	0.69%	-3.40%	-1.33%	-262.27

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
46	CPB.N - TSN.N	Long	20/6/2014	23/6/2014	217	46.61	46.40	214	36.07	36.22	-0.44%	60.80	-0.75%	-0.81%	-0.78%	-138.47
47	IFF.N - PPG.N	Short	2/7/2014	6/10/2014	88	104.29	96.51	95	104.91	95.25	1.22%	105.05	-7.78%	8.45%	0.67%	128.45
48	CPB.N - TSN.N	Long	3/7/2014	8/7/2014	216	45.87	45.72	203	39.05	39.08	-0.22%	61.09	-0.63%	-0.47%	-0.56%	-99.58
49	HCP.N - O.N	Long	8/8/2014	14/8/2014	269	37.14	38.06	224	43.54	44.53	0.13%	63.03	2.17%	-2.61%	-0.19%	-37.45
50	ORLY.OQ - TJX.N	Short	8/9/2014	6/11/2014	185	30.35	31.90	63	157.91	176.59	-5.73%	89.30	4.56%	-12.43%	-6.30%	-980.31
51	HCP.N - O.N	Long	18/9/2014	14/10/2014	268	36.88	38.63	226	42.57	42.71	2.24%	72.39	4.43%	-0.76%	1.87%	364.41
52	BBY.N - FL.N	Short	19/9/2014	24/9/2014	137	57.58	56.32	295	34.32	32.78	1.56%	62.08	-2.57%	4.17%	1.22%	219.49
53	FL.N - LOW.N	Long	19/9/2014	24/9/2014	174	57.58	56.32	204	54.42	53.13	0.21%	62.31	-2.49%	2.08%	-0.09%	-18.39
54	MTB.N - WFC.N	Short	24/9/2014	5/11/2014	142	52.22	53.43	79	125.24	121.68	2.62%	80.33	1.91%	2.33%	2.15%	372.73
55	CPB.N - TSN.N	Long	29/9/2014	30/9/2014	235	42.41	42.74	202	37.53	38.93	-1.17%	60.59	0.48%	-4.13%	-1.51%	-265.84
56	CPB.N - TSN.N	Long	1/10/2014	31/10/2014	233	42.51	44.36	197	39.99	39.96	2.46%	72.37	4.04%	-0.46%	2.05%	364.59
57	GD.N - UTX.N	Short	13/10/2014	24/10/2014	98	100.09	103.16	82	122.06	130.27	-1.88%	66.33	2.76%	-7.09%	-2.21%	-438.69
58	IFF.N - PPG.N	Long	27/10/2014	28/10/2014	103	96.15	95.61	95	96.71	97.75	-0.81%	60.63	-0.86%	-1.41%	-1.13%	-215.05
59	IFF.N - PPG.N	Long	3/11/2014	6/11/2014	101	99.51	97.70	95	102.03	102.50	-1.15%	62.07	-2.12%	-0.80%	-1.47%	-290.00
60	IFF.N - PPG.N	Short	12/11/2014	2/1/2015	90	102.62	115.58	100	99.66	101.42	5.16%	83.81	12.29%	-2.30%	4.72%	906.14
61	ORLY.OQ - TJX.N	Short	14/11/2014	1/12/2014	175	31.00	32.80	55	181.75	182.72	1.69%	66.85	5.24%	-0.90%	1.26%	193.93
62	CPB.N - TSN.N	Short	20/11/2014	24/11/2014	180	43.00	42.17	225	44.75	44.97	-1.12%	61.38	-2.32%	-0.80%	-1.46%	-260.28
63	LLY.N - PFE.N	Short	3/12/2014	17/12/2014	261	31.50	30.72	145	70.30	69.56	-0.52%	66.98	-2.84%	0.69%	-0.89%	-163.26
64	ORLY.OQ - TJX.N	Short	11/12/2014	17/12/2014	172	32.32	32.76	52	191.02	190.31	0.73%	62.72	0.84%	0.04%	0.33%	50.74
65	CPB.N - TSN.N	Long	16/12/2014	2/1/2015	228	43.07	44.10	187	40.98	40.11	2.27%	65.86	2.09%	1.66%	1.90%	331.67
66	LLY.N - PFE.N	Short	19/12/2014	2/1/2015	261	31.93	31.35	140	72.19	69.74	1.04%	65.54	-2.18%	3.04%	0.68%	126.08
67	GD.N - UTX.N	Short	26/12/2014	2/1/2015	83	117.52	115.14	70	141.93	138.53	0.21%	62.72	-2.33%	2.07%	-0.11%	-22.26
68	FL.N - LOW.N	Short	2/1/2015	4/1/2016	160	68.99	74.72	176	56.43	64.46	-2.37%	235.91	7.99%	-16.26%	-3.49%	-732.39

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
69	ORLY.OQ - TJX.N	Long	2/1/2015	4/1/2016	51	192.86	248.45	163	34.31	34.86	17.80%	160.52	28.48%	-3.87%	16.76%	2,584.92
70	IFF.N - PPG.N	Long	5/1/2015	19/2/2015	98	100.45	118.27	80	113.46	117.24	7.63%	81.66	17.41%	-3.88%	7.20%	1,361.90
71	LLY.N - PFE.N	Long	5/1/2015	23/2/2015	143	69.73	71.77	262	31.32	34.08	-2.37%	78.94	2.62%	-9.40%	-2.81%	-510.34
72	GD.N - UTX.N	Long	5/1/2015	6/5/2015	72	137.65	138.66	85	114.55	116.02	-0.27%	116.02	0.43%	-2.17%	-0.86%	-168.25
73	CPB.N - TSN.N	Long	7/1/2015	8/1/2015	230	43.69	44.97	197	39.60	40.02	1.19%	61.12	2.63%	-1.45%	0.84%	150.54
74	BBY.N - FL.N	Short	2/4/2015	12/11/2015	125	62.93	62.05	269	36.80	33.79	3.94%	165.77	-1.78%	6.81%	3.00%	533.65
75	IFF.N - PPG.N	Long	10/6/2015	19/6/2015	90	110.27	113.48	81	115.22	117.16	0.68%	64.79	2.61%	-2.06%	0.35%	66.56
76	IFF.N - PPG.N	Long	30/6/2015	16/7/2015	89	110.22	113.05	79	115.69	116.62	0.94%	66.98	2.26%	-1.21%	0.59%	111.42
77	GD.N - UTX.N	Short	10/7/2015	13/7/2015	89	110.38	111.19	69	145.75	147.23	-0.15%	60.69	0.43%	-1.32%	-0.46%	-90.72
78	GD.N - UTX.N	Short	22/7/2015	24/8/2015	94	102.23	88.17	67	146.28	133.00	-2.22%	75.44	-14.07%	8.61%	-2.61%	-507.32
79	IFF.N - PPG.N	Long	27/7/2015	29/7/2015	89	110.51	113.65	88	102.69	107.08	-0.57%	61.41	2.53%	-4.62%	-0.89%	-168.27
80	EXC.N - PPL.N	Long	29/7/2015	4/1/2016	324	31.49	27.46	275	31.15	33.93	-11.03%	124.26	-13.09%	-10.02%	-11.69%	-2,194.48
81	IFF.N - PPG.N	Long	17/8/2015	5/11/2015	88	113.95	116.21	90	104.21	103.72	1.25%	97.00	1.68%	-0.24%	0.75%	145.98
82	LLY.N - PFE.N	Short	4/9/2015	11/9/2015	254	31.46	32.50	121	80.64	80.65	1.48%	62.67	2.93%	-0.35%	1.13%	200.28
83	CLX.N - PG.N	Long	9/9/2015	10/9/2015	90	112.00	109.37	141	70.44	68.10	0.47%	60.80	-2.65%	3.01%	0.16%	32.44
84	CLX.N - PG.N	Long	11/9/2015	17/9/2015	91	108.96	113.71	146	68.00	70.00	0.71%	63.24	4.05%	-3.27%	0.39%	77.01
85	LLY.N - PFE.N	Short	29/9/2015	6/10/2015	259	30.73	33.34	119	80.15	86.39	-0.38%	63.27	8.12%	-8.13%	-0.74%	-129.84
86	GD.N - UTX.N	Short	23/10/2015	21/12/2015	99	100.46	93.02	69	148.25	138.28	-0.24%	88.03	-7.71%	6.16%	-0.68%	-136.66
87	CLX.N - PG.N	Short	30/10/2015	12/11/2015	129	77.08	75.72	81	122.72	122.96	-0.98%	66.13	-2.07%	-0.56%	-1.31%	-261.01
88	IFF.N - PPG.N	Long	9/11/2015	12/11/2015	86	113.40	114.70	91	102.71	100.61	1.59%	61.92	0.84%	1.70%	1.26%	240.98
89	BBY.N - FL.N	Short	24/11/2015	4/12/2015	120	64.66	65.12	332	29.86	31.05	-1.91%	64.75	0.32%	-4.31%	-2.28%	-402.59
90	CLX.N - PG.N	Short	25/11/2015	30/11/2015	131	76.43	75.65	80	124.99	125.51	-0.72%	61.39	-1.32%	-0.73%	-1.03%	-205.17
91	BBY.N - FL.N	Short	7/12/2015	15/12/2015	120	66.10	66.46	322	30.58	29.82	1.62%	64.05	0.17%	2.14%	1.26%	223.76

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
92	CLX.N - PG.N	Short	7/12/2015	18/12/2015	130	77.81	79.60	79	128.70	130.25	0.54%	66.96	2.00%	-1.56%	0.21%	43.29
93	CLX.N - PG.N	Short	22/12/2015	31/12/2015	126	79.21	79.52	77	127.93	127.91	0.20%	64.08	0.09%	-0.33%	-0.12%	-23.48
94	GD.N - UTX.N	Short	22/12/2015	4/1/2016	105	94.31	94.45	72	139.07	135.62	1.32%	64.80	-0.15%	2.13%	1.00%	198.30
95	MTB.N - WFC.N	Long	4/1/2016	2/3/2016	82	118.79	105.38	137	53.05	48.77	-3.02%	79.91	-11.60%	7.38%	-3.49%	-593.17
96	HCP.N - O.N	Short	4/1/2016	3/1/2017	187	51.33	57.71	283	34.23	29.90	12.54%	228.39	12.10%	10.61%	11.36%	2,189.98
97	CPB.N - TSN.N	Short	4/1/2016	3/1/2017	143	52.44	62.48	188	51.88	60.60	-1.18%	228.35	18.75%	-18.84%	-2.50%	-431.99
98	CLX.N - PG.N	Short	11/1/2016	29/1/2016	129	76.17	79.98	79	126.67	126.87	2.40%	69.39	4.69%	-0.55%	2.05%	406.30
99	ITW.N - SNA.N	Long	19/1/2016	28/1/2016	123	81.97	86.78	61	158.00	158.73	2.77%	65.76	5.56%	-0.82%	2.44%	481.34
100	GD.N - UTX.N	Long	22/1/2016	29/1/2016	80	125.10	129.15	115	86.95	87.00	1.59%	63.93	2.93%	-0.39%	1.27%	254.32
101	IFF.N - PPG.N	Short	26/2/2016	29/2/2016	97	97.83	98.01	94	105.50	104.15	0.74%	60.68	-0.13%	0.97%	0.43%	83.68
102	IFF.N - PPG.N	Short	1/3/2016	2/3/2016	95	97.55	99.22	96	104.32	105.81	0.08%	60.69	1.39%	-1.73%	-0.23%	-45.08
103	CLX.N - PG.N	Short	28/4/2016	29/4/2016	124	79.50	79.75	79	124.21	124.36	0.10%	60.67	0.01%	-0.43%	-0.21%	-41.52
104	MTB.N - WFC.N	Long	28/4/2016	4/5/2016	83	119.14	115.31	147	50.47	49.00	-0.59%	62.03	-3.52%	2.48%	-0.95%	-163.83
105	CLX.N - PG.N	Long	2/5/2016	9/6/2016	80	125.25	130.06	124	80.06	82.57	0.37%	79.00	3.53%	-3.62%	-0.03%	-5.44
106	MTB.N - WFC.N	Short	5/5/2016	3/1/2017	153	48.97	55.67	86	114.33	157.99	-15.76%	172.47	13.28%	-39.64%	-16.75%	-2,902.13
107	BBY.N - FL.N	Short	16/6/2016	11/7/2016	143	54.17	57.27	346	29.18	30.97	-0.99%	71.06	5.34%	-6.54%	-1.38%	-247.10
108	FL.N - LOW.N	Long	16/6/2016	11/7/2016	182	54.17	57.27	143	76.86	82.68	-1.29%	72.68	5.41%	-7.95%	-1.63%	-340.74
109	ITW.N - SNA.N	Short	30/6/2016	1/7/2016	63	154.83	158.03	98	102.24	104.31	-0.01%	60.69	1.76%	-2.33%	-0.31%	-61.95
110	CLX.N - PG.N	Short	12/7/2016	1/8/2016	116	85.49	85.42	72	136.73	130.68	2.16%	69.44	-0.38%	4.02%	1.81%	358.04
111	ITW.N - SNA.N	Short	28/7/2016	12/12/2016	62	155.75	176.45	86	115.47	127.02	1.48%	126.03	12.97%	-10.96%	0.84%	164.07
112	CLX.N - PG.N	Short	3/8/2016	11/8/2016	114	86.26	86.80	76	131.87	134.32	-0.63%	64.12	0.32%	-2.20%	-0.95%	-188.76
113	IFF.N - PPG.N	Short	7/9/2016	8/9/2016	87	107.24	106.54	71	139.11	139.26	-0.37%	60.68	-0.97%	-0.42%	-0.69%	-132.23
114	IFF.N - PPG.N	Short	31/10/2016	1/11/2016	99	93.35	93.25	74	134.89	130.90	1.48%	60.68	-0.43%	2.65%	1.17%	224.68

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
115	LLY.N - PFE.N	Long	3/11/2016	14/11/2016	136	72.81	78.00	265	30.63	32.83	0.68%	64.80	6.82%	-7.60%	0.32%	58.04
116	IFF.N - PPG.N	Short	8/11/2016	17/11/2016	99	93.83	96.77	76	122.00	121.19	1.90%	64.45	2.81%	0.29%	1.55%	288.17
117	XRAY.OQ - SYK.N	Short	9/11/2016	10/11/2016	97	118.55	116.88	163	60.48	61.95	-1.88%	64.93	-1.71%	-2.74%	-2.18%	-466.53
118	IFF.N - PPG.N	Short	18/11/2016	23/11/2016	96	96.23	96.59	82	119.14	119.99	-0.18%	62.01	0.05%	-1.04%	-0.51%	-97.15
119	BEN.N - RJF.N	Short	18/11/2016	1/12/2016	137	72.84	72.77	282	35.86	36.33	-0.71%	65.54	-0.40%	-1.66%	-1.03%	-207.48
120	LLY.N - PFE.N	Long	25/11/2016	3/1/2017	154	68.51	74.00	264	31.68	32.69	3.06%	77.24	7.70%	-3.72%	2.65%	501.58
121	IFF.N - PPG.N	Short	1/12/2016	6/12/2016	94	96.42	95.98	82	121.07	121.04	-0.20%	62.04	-0.79%	-0.30%	-0.53%	-100.94
122	BEN.N - RJF.N	Short	2/12/2016	19/12/2016	135	73.32	71.26	275	36.34	36.40	-1.47%	67.53	-3.11%	-0.53%	-1.81%	-360.81
123	XRAY.OQ - SYK.N	Short	19/12/2016	13/5/2019	97	118.77	184.49	165	60.24	54.23	34.33%	483.28	54.95%	5.56%	32.07%	6,883.21
124	BEN.N - RJF.N	Short	21/12/2016	23/12/2016	136	73.11	71.46	271	37.12	37.10	-1.10%	61.38	-2.56%	-0.26%	-1.40%	-280.79
125	IFF.N - PPG.N	Short	21/12/2016	3/1/2017	97	95.39	95.43	84	118.12	118.45	-0.12%	64.76	-0.28%	-0.63%	-0.46%	-88.60
126	ORLY.OQ - TJX.N	Long	3/1/2017	4/1/2017	35	280.00	280.82	148	37.81	37.91	0.09%	60.38	-0.01%	-0.81%	-0.30%	-46.48
127	ITW.N - SNA.N	Long	3/1/2017	30/1/2017	81	123.73	128.23	56	171.42	181.04	-0.89%	72.45	3.33%	-6.05%	-1.26%	-246.67
128	APD.N - ECL.N	Short	3/1/2017	21/2/2017	73	118.12	122.75	69	144.44	141.01	3.09%	82.53	3.57%	1.85%	2.65%	492.13
129	BBY.N - FL.N	Long	3/1/2017	28/4/2017	230	43.10	52.42	111	71.59	76.77	8.78%	106.63	21.29%	-8.16%	8.19%	1,461.99
130	MTB.N - WFC.N	Long	4/1/2017	2/1/2018	63	156.90	172.48	135	56.20	61.04	1.88%	191.21	9.61%	-10.72%	0.78%	136.93
131	ORLY.OQ - TJX.N	Long	24/1/2017	2/3/2017	37	265.56	270.90	149	37.38	39.12	-0.40%	69.95	1.71%	-5.37%	-0.86%	-131.63
132	ITW.N - SNA.N	Long	2/2/2017	14/2/2017	78	126.60	129.30	53	172.03	172.00	1.12%	65.12	1.83%	-0.37%	0.77%	147.07
133	FL.N - LOW.N	Short	3/4/2017	10/4/2017	134	83.11	82.08	132	74.80	72.44	0.83%	66.58	-1.54%	2.82%	0.51%	106.92
134	FL.N - LOW.N	Short	17/4/2017	18/4/2017	134	81.06	81.80	137	72.07	72.47	0.21%	63.41	0.61%	-0.87%	-0.09%	-19.05
135	FL.N - LOW.N	Short	26/4/2017	30/5/2017	129	84.61	80.77	130	77.08	60.00	8.24%	77.79	-4.83%	21.70%	7.87%	1,647.25
136	BBY.N - FL.N	Long	8/5/2017	9/5/2017	195	51.79	51.89	103	76.87	76.77	0.17%	60.87	-0.11%	-0.26%	-0.17%	-31.07
137	APD.N - ECL.N	Short	12/5/2017	15/8/2017	68	126.63	131.16	68	143.86	146.50	0.70%	103.55	3.23%	-2.59%	0.14%	24.97

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
138	BBY.N - FL.N	Long	18/5/2017	26/5/2017	194	50.85	60.43	110	71.27	59.57	17.77%	65.81	18.51%	15.99%	17.39%	3,079.71
139	FL.N - LOW.N	Short	8/6/2017	14/1/2019	139	78.00	96.36	181	55.29	56.46	11.22%	341.90	23.20%	-5.17%	9.58%	1,998.37
140	ORLY.OQ - TJX.N	Long	12/7/2017	14/7/2017	57	174.97	185.97	163	34.39	35.11	3.27%	61.67	5.98%	-2.66%	2.87%	447.16
141	ITW.N - SNA.N	Long	11/8/2017	15/9/2017	71	138.23	144.97	63	152.49	146.72	4.34%	76.23	4.57%	3.31%	3.94%	765.82
142	APD.N - ECL.N	Short	16/8/2017	17/8/2017	65	131.43	131.51	68	146.91	147.88	-0.33%	60.68	-0.29%	-0.97%	-0.66%	-121.44
143	APD.N - ECL.N	Short	21/8/2017	22/8/2017	66	130.16	130.63	68	145.74	145.84	0.13%	60.68	0.01%	-0.38%	-0.20%	-36.46
144	APD.N - ECL.N	Short	7/9/2017	13/9/2017	64	131.10	129.73	69	144.50	149.01	-2.17%	62.73	-1.40%	-3.45%	-2.51%	-461.60
145	APD.N - ECL.N	Short	28/9/2017	29/9/2017	66	128.95	129.27	66	151.45	151.50	0.10%	60.68	-0.10%	-0.34%	-0.23%	-42.86
146	APD.N - ECL.N	Short	12/10/2017	6/11/2017	64	132.43	131.20	65	152.57	159.48	-2.87%	71.55	-1.28%	-4.95%	-3.26%	-599.42
147	CLX.N - PG.N	Long	2/1/2018	17/1/2018	67	148.79	142.96	107	91.80	90.45	-1.24%	66.73	-4.22%	1.10%	-1.58%	-312.89
148	BEN.N - RJF.N	Short	2/1/2018	6/2/2018	109	89.61	87.51	248	40.28	34.67	5.88%	76.42	-2.65%	13.45%	5.49%	1,084.53
149	APD.N - ECL.N	Short	2/1/2018	23/4/2018	63	134.20	148.22	60	164.88	166.03	4.44%	111.50	10.09%	-1.52%	3.83%	702.76
150	GD.N - UTX.N	Short	2/1/2018	2/8/2018	76	127.90	133.98	49	203.54	193.39	4.87%	160.69	4.44%	3.68%	4.06%	798.74
151	EXC.N - PPL.N	Long	2/1/2018	10/8/2018	253	39.54	43.47	277	31.08	29.00	8.44%	151.72	9.62%	5.30%	7.62%	1,418.73
152	ORLY.OQ - TJX.N	Long	2/1/2018	13/5/2019	41	241.34	360.66	146	38.35	53.58	17.22%	197.94	49.06%	-42.60%	15.94%	2,469.87
153	LLY.N - PFE.N	Short	8/1/2018	31/7/2018	223	36.78	38.25	116	86.69	98.25	-5.55%	157.12	3.63%	-14.60%	-6.41%	-1,170.27
154	CLX.N - PG.N	Long	23/1/2018	26/1/2018	69	143.66	145.00	109	89.55	87.99	1.33%	62.01	0.63%	1.41%	1.02%	200.49
155	CLX.N - PG.N	Short	6/2/2018	30/7/2018	118	80.50	80.08	76	125.62	132.02	-2.81%	138.47	-0.84%	-6.23%	-3.54%	-674.43
156	BEN.N - RJF.N	Long	7/2/2018	16/5/2018	288	36.31	33.52	112	90.78	95.00	-6.19%	108.04	-7.98%	-5.41%	-6.71%	-1,384.90
157	HCP.N - O.N	Long	16/2/2018	17/8/2018	459	22.38	27.07	201	49.33	58.16	1.87%	149.62	20.62%	-19.07%	1.13%	228.26
158	APD.N - ECL.N	Long	24/4/2018	2/5/2018	60	166.56	162.61	58	149.40	145.58	-0.08%	63.56	-2.67%	2.17%	-0.42%	-79.00
159	APD.N - ECL.N	Short	3/5/2018	13/5/2019	59	143.87	178.77	61	161.39	203.80	-2.88%	234.12	23.89%	-28.34%	-4.16%	-762.03
160	ITW.N - SNA.N	Long	16/5/2018	11/6/2018	69	145.97	148.03	65	149.74	157.08	-1.69%	71.76	1.11%	-5.33%	-2.05%	-406.72

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
161	IFF.N - PPG.N	Short	16/5/2018	2/7/2018	90	101.89	102.72	79	123.62	123.72	0.35%	81.40	0.49%	-0.61%	-0.08%	-14.60
162	BBY.N - FL.N	Long	16/5/2018	14/1/2019	128	78.95	56.11	181	44.02	56.46	-28.63%	150.75	-29.23%	-29.77%	-29.47%	-5,325.91
163	BEN.N - RJF.N	Short	17/5/2018	13/5/2019	103	96.31	85.35	298	33.68	33.42	-5.27%	229.80	-11.68%	-1.22%	-6.42%	-1,281.20
164	CPB.N - TSN.N	Short	21/5/2018	18/12/2018	112	68.42	53.50	287	33.77	38.87	-18.06%	156.92	-22.20%	-16.41%	-18.97%	-3,291.66
165	ITW.N - SNA.N	Short	12/6/2018	23/10/2018	62	158.04	148.04	67	149.63	124.15	5.48%	123.86	-6.63%	16.09%	4.86%	963.30
166	CLX.N - PG.N	Long	31/7/2018	20/8/2018	75	134.07	147.73	124	79.85	83.80	2.68%	71.20	9.87%	-5.35%	2.32%	463.50
167	LLY.N - PFE.N	Long	1/8/2018	6/5/2019	101	98.70	115.53	215	39.66	40.99	7.64%	173.47	16.73%	-5.01%	6.71%	1,240.41
168	GD.N - UTX.N	Long	3/8/2018	13/5/2019	51	194.04	170.00	73	134.08	134.03	-6.21%	189.39	-12.69%	-1.59%	-7.17%	-1,411.78
169	HCP.N - O.N	Short	20/8/2018	13/5/2019	167	59.18	67.80	369	27.28	30.15	1.91%	187.47	14.24%	-12.07%	0.97%	193.04
170	EXC.N - PPL.N	Short	27/9/2018	2/10/2018	297	28.45	29.60	234	42.24	43.55	0.19%	62.03	3.69%	-3.43%	-0.15%	-27.02
171	CLX.N - PG.N	Long	15/10/2018	22/10/2018	69	145.29	152.72	126	79.02	86.73	-2.30%	64.25	4.81%	-10.09%	-2.62%	-523.04
172	ITW.N - SNA.N	Long	24/10/2018	14/11/2018	80	125.99	133.91	65	149.11	161.54	-0.88%	71.15	5.98%	-8.75%	-1.24%	-245.50
173	EXC.N - PPL.N	Long	8/11/2018	12/11/2018	224	44.78	44.83	275	31.47	31.51	0.00%	61.29	-0.19%	-0.49%	-0.33%	-61.09
174	ITW.N - SNA.N	Short	15/11/2018	10/12/2018	60	159.85	150.53	74	132.33	129.13	-1.66%	70.06	-6.14%	2.01%	-2.02%	-392.46
175	ITW.N - SNA.N	Long	11/12/2018	16/1/2019	77	133.01	130.11	64	150.74	161.88	-4.71%	76.25	-2.48%	-7.87%	-5.09%	-1,012.51
176	CPB.N - TSN.N	Long	19/12/2018	9/1/2019	257	38.88	33.58	143	52.76	57.34	-11.50%	66.72	-13.93%	-9.17%	-11.88%	-2,083.76
177	EXC.N - PPL.N	Long	3/1/2019	9/1/2019	222	44.18	45.19	303	28.11	29.15	-0.50%	62.38	1.98%	-4.08%	-0.84%	-153.28
178	CPB.N - TSN.N	Short	10/1/2019	13/5/2019	134	56.67	79.00	297	33.96	38.61	9.11%	118.91	39.00%	-14.57%	8.44%	1,492.26
179	ITW.N - SNA.N	Short	17/1/2019	11/2/2019	60	162.44	154.27	76	129.44	136.44	-5.22%	70.78	-5.34%	-5.82%	-5.58%	-1,092.98
180	BBY.N - FL.N	Long	25/1/2019	31/1/2019	174	58.49	58.98	138	57.25	56.50	1.04%	62.82	0.54%	0.90%	0.70%	125.94
181	BBY.N - FL.N	Long	4/2/2019	6/2/2019	169	58.60	59.58	140	54.90	56.67	-0.47%	61.16	1.37%	-3.63%	-0.81%	-143.34
182	FL.N - LOW.N	Short	4/2/2019	6/2/2019	114	97.68	97.92	178	54.90	56.67	-1.38%	64.79	-0.05%	-3.54%	-1.69%	-352.49
183	ITW.N - SNA.N	Long	12/2/2019	13/5/2019	73	137.15	148.96	63	155.62	164.84	1.42%	102.96	8.30%	-6.66%	0.90%	178.31

N°	Pair	Position	Open date	Close date	# Long stocks	Long open price	Long close price	# Short stocks	Short open price	Short close price	Return without costs	Trade costs	Long return	Short return	Return including costs	Income (USD)
184	BBY.N - FL.N	Long	4/3/2019	14/3/2019	143	68.00	68.24	115	64.00	59.68	3.11%	64.03	0.04%	6.29%	2.73%	467.09
185	FL.N - LOW.N	Short	4/3/2019	14/3/2019	104	104.66	101.00	147	64.00	59.68	1.25%	67.24	-3.79%	6.38%	0.92%	187.16
186	BBY.N - FL.N	Long	10/4/2019	13/5/2019	134	74.78	70.26	125	60.22	56.32	-0.67%	71.37	-6.34%	5.93%	-1.08%	-189.55
187	FL.N - LOW.N	Short	10/4/2019	13/5/2019	95	113.69	104.78	158	60.22	56.32	-1.13%	75.54	-8.13%	6.01%	-1.51%	-305.79
188	LLY.N - PFE.N	Long	10/5/2019	13/5/2019	86	115.80	114.85	204	40.62	40.25	-0.03%	60.57	-1.12%	0.54%	-0.37%	-66.79