European Research on Management and Business Economics 24 (2018) 160–167

European Research on Management and Business Economics

www.elsevier.es/ermbe

Pairs trading techniques: An empirical contrast

Mario Carrasco Blázquez, Carmen De la Orden De la Cruz, Camilo Prado Román

Rey Juan Carlos University, Department of Business Management, Facultad de Ciencias Jurídicas y Sociales, Paseo Artilleros s/n, Madrid 28933, Spain

ARTICLE INFO

Article history:
Received 1 March 2018
Received in revised form 23 May 2018
Accepted 25 May 2018
Available online 5 July 2018

JEL classification:
C02
C40
C50

Keywords:
Pairs trading
Market neutral
Mean reversion
Co-integration

ABSTRACT

Pairs trading is one of the most commonly used market neutral strategies. Over the last few years, several hedge funds have used different ways to successfully implement this trading strategy. The most extensively used techniques (correlation, distance, stochastic, stochastic differential residual and cointegration) use different methodologies and statistical tools to determine the two key elements of the strategy: pairs selection and the establishment of the long-term relationship between them. The purpose of this paper is to analyze the process of selecting pairs and determining the residual series using each one of the different techniques and comparing the outputs. Results indicate that far from being differentiated systems, relationships exist between the various techniques in terms of pairs selection and residual series creation. However, some techniques are more efficient at creating residual series than others, which then means that these techniques would have the highest probabilities of generating profits. The analysis concludes that cointegration is the most efficient method of structuring a pairs trading strategy.

1. Introduction

Pairs trading, together with statistical arbitrage and risk arbitrage, has been one of the strategies most commonly used by hedge funds since the end of the 1990s (Nicholas, 2004). This type of strategy seeks to obtain profits from inefficiencies existing in the market, irrespective of whether it is a bull, bear or neutral market. Pairs trading consists of the simultaneous opening of long and short positions in two assets with a balance point between them. In this way, the earnings from a long position cover the losses from a short position and vice versa, meaning that the market risk is close to zero, as is the joint beta strategy. Therefore, the key elements that determine the success of a trade consists of determining the balance point between two securities and the point in time that prices move sufficiently away from the balance point to take positions. Pairs trading is not without risks as a miscalculation of these two elements can lead to a failure of the strategy (Opiela, 2004). Securities volatility is an additional risk that needs to be considered, even if there is a high degree of correlation between the securities (Whistler, 2004). Nevertheless, pairs trading can be used to not only generate profits regardless of market trend, but also to balance a portfolio given its market neutral properties. But to optimize trading results, it is necessary to first select the best method to implement a pairs trading strategy. There are five main techniques that can be utilized to execute a pairs trading strategy. These are: correlation, distance, stochastic, stochastic differential residual and co-integration although other authors mention others such as the machine learning and the time-series methods (Krauss, 2017).

These five techniques have been developed and proposed by different authors, however there have been no studies that analyze all of them jointly and under the same conditions. Accordingly, it is necessary to take a general and objective approach to be able to compare and contrast the properties of each one in relation to one another.

To do so, all techniques were compared, by performing an analysis of how pairs are chosen and how the balance point is determined as measured by the residual series. This allowed for the comparison of similarities that helped to determine the most efficient technique in each of the key aspects, given that until now, there have been no detailed comparative studies of the different techniques used to develop a pairs trading strategy. The sole exception is the study by Jurek and Yang (2007) which compared the results obtained from the distance and the stochastic techniques. This, then, is the first comprehensive study in which all five techniques are analyzed for interrelationships and to determine each one's strengths and weaknesses.

The remainder of the paper is organized as follows. The second section describes the theoretical aspects of pairs trading. The third section presents the data used and outlines the empirical techniques. The fourth section shows empirical results and the final section presents conclusions.

https://doi.org/10.1016/j.iedeen.2018.05.002
2444-8834/

2. Theory

The principal challenge faced by financial investors using a pairs trading strategy is to find pairs of assets, be they stock, debentures, futures, currency, etc. with sound and lasting statistical relationships, based on the assumption that the different financial assets are, more or less, related (Arfaoui & Ben Rejeb, 2017). If the relationships are not stable in the medium or long-term, the investment system may incur substantial losses. The evolution of the main techniques includes theoretical components that relate different timeframes over which the relationships between a pair of assets break down. The most widely used include statistical tools, such as distance and correlation, stochastic and mathematical processes such as the stochastic and stochastic differential residual and statistical and econometric processes, such as that of co-integration (Do & Faff, 2010).

The correlation technique used by Wong (2010) chooses pairs of stocks according to the correlation coefficient existing between them and determines the residual series by means of a ratio of prices. This system has been used by different authors such as Ehrman (2006). With a given pair of stocks A and B, a ratio between their prices is used to generate the residual series:

$$\frac{P_t^A}{P_t^B} \quad (1)$$

The distance technique was developed and revised by Gatev, Goetzmann, and Rouwenhorst (2006). Their work is considered by many authors as the best-known work of pairs trading (Smith & Xu, 2017). It determines the pairs to be used by means of the distance between them, such distance being the total sum of squares of the difference between the standardized prices of both assets. The residual series is determined by the difference in standardized prices. Based on their work, Do and Faff (2010) developed a profitable pairs trading strategy.

The distance between each pair of assets is determined by:

$$D = \sum_{i=1}^{n} (P_{Ai} - P_{Bi})^2 \quad (2)$$

where D is the distance between both assets, P_{Ai} is the standardized price of asset A at moment i and P_{Bi} is the standardized price of asset B at moment i.

The standardized price of an asset is determined by:

$$P_{it} = \frac{P_t - E(P_i)}{\sigma_i} \quad (3)$$

where P_{it} is the standardized price of asset i at moment t, P_t is the price of the asset at moment t, $E(P_i)$ is the mean or expected value of asset i and σ_i is the volatility or standard deviation of asset i.

The residual series is determined by the standardized price difference between both assets:

$$P_{Ai} - P_{Bi} \quad (4)$$

The stochastic technique is based on the Hornstein–Uhlenbeck process and assumes a stationary residual series according to:

$$dx_t = k(\theta - x_t)dt + \sigma dW_t \quad (5)$$

where k indicates the speed at which the residual series converges on its mean.

This was later used by Kanamura, Rachev, and Fabozzi (2008) and Jurek and Yang (2007) who, in addition, claimed to have achieved better results than those obtained by Gatev et al. (2006) based on a simulation of data. The residual series is established with the difference in price logarithms using the formula:

$$\log(P_t^A) - \log(P_t^B) \quad (6)$$

The stochastic differential residual technique developed by Do, Faff, and Hamza (2006) uses the CAPM and APT theoretical models to determine the balance between the assets. The residual series is obtained with:

$$G_t = G(p_t^A, p_t^B, U_t) = R_t^A - R_t^B - \Gamma r_t^m \quad (7)$$

where R_t^A is the profitability of asset A at moment t, R_t^B is the profitability of asset B at moment t, Γ is the difference between the betas of both assets and r^m is the profitability of the reference market, measured by the index.

Finally, the co-integration technique used by Vidyamurthy (2004) establishes the relationship between assets based on the concept of co-integration developed by Engle and Granger (1987). The cointegration equation is determined by formula:

$$y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t \quad (8)$$

And the residual series can be written as the equation:

$$D_r = y_t - \alpha_0 - \alpha_1 x_t - \varepsilon_t \quad (9)$$

where α_1 is the cointegration coefficient.

3. Data and methodology

3.1. Data and sample set used

Stocks within the US financial sector, specifically, the SP500 bank subgroup, were the sample set chosen for this study. This subgroup was chosen based on the recent economic crisis that severely punished the international banking industry, partly because of its high exposure to the mortgage and real estate markets. In addition, the banking sector is one of the most representative sectors of the economy, reaching almost 7% of worldwide GNP (World Bank, 2014). The total number of potential pairs analyzed in the study was determined by the formula:

$$P = \frac{n^2 - n}{2} \quad (10)$$

where P is the number of possible resulting pairs and n is the number of stocks to be analyzed.

The sample period over which each technique was analyzed extended from January 1, 2008 to December 31, 2013. This time frame was chosen in order to more strongly contrast a market neutral investment strategy such as pairs trading, due to the major stock market fluctuations from the beginning of the economic crisis at the end of 2007 in which even the most advanced statistical tools of risk measure, as VaR, had shown to be inefficient and weak (Muela, Martin, & Sanz, 2017). Pairs trading is one of the most suitable strategies under current financial conditions that feature an unstable political environment and a high degree of uncertainty in the financial markets and the global economy (Ehrman, 2006).

3.2. Methodology used by each technique for pairs selection

Correlation: to determine the pair of stocks to be used, the correlation coefficient of all the pairs studied was calculated and the pair showing the highest degree of correlation was chosen.

Distance: to determine the pair of stocks to be used, the standardized prices of each pair of stocks and the distance between them were calculated according to formula (2) and (3). The pair of stocks chosen was the one that had the least distance between them.

Stochastic: to determine the pair of stocks to be used, the pair with the highest mean reversion was chosen; this being the pair with the highest k coefficient.