



Instituto Superior de Economia e Gestão

UNIVERSIDADE TÉCNICA DE LISBOA

DESDE 1911

MASTER IN FINANCE

MASTER'S FINAL ASSIGNMENT DISSERTATION

**PAIRS TRADING PERFORMANCE AND IMPLICATIONS APPLIED TO THE
PORTUGUESE MARKET**

FRANCISCO CASTELO BRANCO

SEPTEMBER - 2012



Instituto Superior de Economia e Gestão

UNIVERSIDADE TÉCNICA DE LISBOA

DESDE 1911

MASTER IN FINANCE

MASTER'S FINAL ASSIGNMENT DISSERTATION

**PAIRS TRADING PERFORMANCE AND IMPLICATIONS APPLIED TO THE
PORTUGUESE MARKET**

FRANCISCO CASTELO BRANCO

ORIENTATION:

RAQUEL M. GASPAR

SEPTEMBER - 2012

Pairs Trading Performance and Implications Applied to the Portuguese Market

Francisco Castelo Branco†
Technical University of Lisbon

September 2012

Abstract

Pairs trading is a statistical arbitrage strategy that gains popularity on Wall Street in the mid-1980's. Such concept is actually very simple to understand: select two assets that had a closely past behavior and when that path is disturbed, take a long/short position in the hope that in the future, history repeats itself. Thus, the primary aim of this final master thesis is to explore and analyses the effectiveness of implementing such strategy in the Portuguese stock market during the period comprised between 2002 and 2012. Since Pairs Trading strategy only uses past stock information and the Efficient Market Hypothesis in its weak-form, states that stock prices already contain all historical data available on the market and therefore, there is no way of using that information in order to produce excess risk adjusted returns, the second objective of the thesis is to confirm that Pairs Trading do in fact explore market inefficiencies to generate statistical excess returns over the market, and present new evidence about the informational efficiency of the Portuguese stock market. This research verifies which value of the thresholds and what method maximises the strategy profitability. It also verifies the importance of that threshold and necessity of an accurate selection of the pairs on the strategy performance.

JEL: C22; C32; G11; G12; G14;

Keywords: pairs trading, statistical arbitrage, quantitative strategy, market efficiency, cointegration, Portuguese stock market.

*Technical University of Lisbon - School of Economics and Management (ISEG), Student of Masters in Finance. I am grateful to Raquel M. Gaspar for the helpful comments and for the Brainstorming's. I am indebted to my family for all the aid and to my friends Ana Arriaga e Cunha, Diogo Belchior and Diogo Jalles for the help and assistance which have benefited the outcome of the thesis. True and sincere thanks to Afonso Eça for the outstanding assistance and support.

Pairs Trading Performance and Implications Applied to the Portuguese Market

Francisco CASTELO BRANCO ¹
Technical University of Lisbon

September 27, 2012

¹Technical University of Lisbon-School of Economics and Management (ISEG), Student of Masters in Finance. I am grateful to Raquel M. Gaspar for the helpful comments and for the Brainstormings. I am indebted to my family for all the aid and to my friends Ana Arriaga e Cunha, Diogo Belchior and Diogo Jalles for the help and assistance which have benefited the outcome of the thesis. A true and sincere thanks to Afonso Eça for the outstanding assistance and support.

Contents

1	Introduction	4
2	Literature Review	7
3	Data	13
4	Methodology	15
4.1	Pairs Formation	15
4.1.1	Square Distance Approach	15
4.1.2	Cointegration Approach	17
4.2	Returns Computation	18
4.3	Transaction Costs	20
4.3.1	Explicit Costs	20
4.3.2	Implicit Costs	20
4.4	Assessing Pairs Trading Performance	21
4.5	Efficiency Tests	22
5	Results	23
5.1	Square Distance Approach	23
5.2	Pairs Trading under different d	26
5.3	Assessing Pairs Trading Performance	27
5.4	Efficiency Tests	27
5.5	Cointegration Method	29
5.6	Updated Data - The first 3 Quarters of 2012	31
6	Conclusions and Implications	33
	References	35
7	Appendix	38

List of Figures

4.1	Pairs at the Selection Period	16
4.2	Pairs at the Trading Period	17
5.1	Volatility	23
5.2	Pairs Trading Portfolio versus Market Performance	24
5.3	Pairs Trading Portfolio Annual Returns	25
5.4	Pairs Trading and the Issue of Luck Analysis	28
5.5	PSI-20 and the Issue of Luck Analysis	28
5.6	Pairs Trading Portfolio Annual Returns	30
5.7	2012 Cumulative Return	31

List of Tables

3.1	List of Financial Stocks used in this Research	13
7.1	Stock's Pairs Over the Time	38
7.2	Pairs Trading Returns per year	39
7.3	Pairs Trading and Market Returns	39
7.4	Apha and Beta for Pairs Trading	40
7.5	Pairs Trading Summary Statistics	40
7.6	d Value Sensitivity Analysis	41
7.7	Summary Statistics for Pairs Trading under different d	41
7.8	Random Trigger Test	42
7.9	Random Trigger OLS Analysis	42
7.10	Random Pair Test	43
7.11	Random Pair OLS Analysis	43
7.12	Run Test	44
7.13	Autocorrelation Test	45
7.14	Cointegration method Analysis	46
7.15	Pairs Trading Strategy Vs. Psi-20	46

1. Introduction

In the financial markets there have been, for a very long time, numerous investment strategies that tried to make profit out of the market inefficiencies and anomalies. A cluster of such strategies, known as quantitative analysis or statistical arbitrage, tries to take advantage of short term detour from the “fair” asset price to generate statistical excess returns through sophisticated techniques and effective algorithms. Pairs Trading is a strategy enclosed in this type of statistical arbitrage and widely practice by hedge funds. The two major attractive characteristics of the strategy are its neutrality to the market, in sense that it generates a zero beta over it, and its effectiveness on exploiting the market inefficiencies to generate excess returns.

The efficient market hypothesis (EMH), in its weak form, postulates that the past available information of an asset is entirely incorporated on the asset value. This means that by only taking into account historical data, there is not any potential on predicting the future behavior of an asset price. The consequence of such hypothesis is that there is not any logical trading rule, based only on past information, that could generate a significant positive excess return over the market. Considering all of that, it is interesting to assess if in practice it is possible to generate systematically positive excess returns by conducting one statistical arbitrage strategy as it is Pairs Trading.

Pairs Trading is framed with quantitative strategies presenting as the main glamour its low risk due to the expected non correlation to the market behavior, and its high positive performance owing to the consistent returns. When money is at stake in the financial markets, “Human beings do not like to trade against human nature, which wants to buy stocks after they go up not down¹”. By contrast, in Pairs Trading and in the majority of statistical arbitrage strategies, the idea is to sell assets that are overvalued and buy the ones that are undervalued, when analysed and compared with its “fair” value. Obvious the “fair” value is not an easy thing to determine and due that reason, Pairs Trading does not base the strategy on assessing the “fair” value of an asset, but else, on the relative pricing. The idea of relative pricing is sustained upon the notion that assets with similar characteristics should present, more or less, the same value. Thus being, when the value of two similar assets present different prices, it is possible to assume that, one or both, are overvalued or undervalued relative to its “fair” value.

¹Nunzio Tartaglia in Hansell, S. (1989), “Inside Morgan Stanley’s Black Box”.

Pairs Trading might be a concept of widespread understanding, yet very complex to implement. The process is based on identifying, with high accuracy, assets that had similar historical behavior and trade them when the correlation deviates, shorting the higher value asset and buying the lower value asset. By exploring such spreads they create a valid possibility of making profits when, eventually, the assets revert to moving together. The two most important phases in order to achieve success in this strategy are the one of identifying the pairs and the one of implementing an efficient trading algorithm.

There are some typical aspects and questions regarding Pairs Trading implementation that are enlightened in this Thesis. For instance, which assets should be designated to be pairs? What distance deviation between the pairs is good enough in order to initiate the trade? When is that the trade is closed? And what if it never closes? In order to go deeper than the usual questions, it is important to conclude with proper proof that Pairs Trading is not an issue of luck, rather an efficient strategy.

From academic literature it is possible to define four distinctive methods for implementing Pairs Trading. (i) Gatev et al. (2006), introduces the minimum distance method based on the distance relation of two normalised assets; (ii) Nath (2006) and Elliott and Malcolm (2005) outline a stochastic approach to a spread by modeling its behavior of mean reversion; (iii) Combined Forecast approach which is given by Huck (2009,2010) and does not assume any equilibrium model and; (iv) Finally, Vidyamurthy (2004) outlines the cointegration approach, that tries to parametrise the strategy.

This Thesis covers both the Minimum Distance and the Cointegration methods. Although the methods are used in some literature, there exist specifications that are not always obvious. For that reason, the undertaken methods follow the original works whenever it is possible but also take certain assumptions without undermining the essence of the methods. Such essence is the same in both methods: the selection of the pairs is done using only historical data; and the trigger of the trade as well as the positions to take are always based upon a spread process behavior.

The purpose of this Thesis is to analyse Pairs Trading under the concepts of Minimum Distance and Cointegration methods and derive the dynamic implications of it on the balance of price series, this is, the evaluation of the mean reversion on the pairs strategy. Thus, after implementing both strategies, the aim of the Thesis is to assess the profitability and implications of Pairs Trading applied to the Portuguese index market (PSI-20).

Such study have a high component of motivation since it is the first work that turns possible to verify the behavior of Pairs Trading on a much less liquid market as it is the Portuguese one. There exist other aspects of interest on studying Pairs Trading applied implications to the Portuguese

market. First, Pairs Trading is relatively known in the trading sphere but it did not have the same attention nor the same rigorous analysis when compared with momentum or contrarian investment strategies. Second, since it is a market neutral and self-financing strategy, it is interesting to observe if it is possible to obtain better results than the market and verify which of these results are due to the behavior of the market and what part is due only to implementation of the strategy. Finally, since Weak-Form efficiency is based on the premise that past stock prices can not be used to generate alpha i.e. informational efficiency states that no one can get an alpha² different from zero using any past information, this thesis also provides evidence on the informational efficiency on Portuguese stock market.

This study is organised as follows. Section 2 introduces the Literature Review. Section 3 and 4 take into account, respectively, the Data and the Methodology used for the empirical study of the strategy. The discussion of the empirical results is given in Section 5, followed by the Conclusions and Implications on Section 6.

²Alpha is considered to represent the excess return of a strategy regarding the return of the market.

2. Literature Review

Any strategy is considered neutral to market returns when their strategy returns are not correlated with market returns. A market neutral strategy follows its own path regardless the market movements. If a strategy is uncorrelated with the market, in the CAPM¹ context, it means that β is equal to zero. It is intuitive that market neutral strategies are able to eliminate the systematic risk². Pairs Trading, as said before, gained popularity among the hedge funds due to the ability of profiting through the anomalies in the financial markets but also due to its independence from the market actions.

Jacobs et al. (1993) affirms that in order to eliminate the exposure to the market risk, a portfolio should use a long-short strategy to become a zero beta portfolio and achieve returns only by its alpha. Such strategy may be called a market neutral strategy. This statement is also sustained by Fung and Hsieh (1996) where it is defended that a strategy that generates returns independently from the market returns may be called as neutral. Long-short market neutral strategies, as Alexander and Dimitriou (2002) states, displays no correlation with the market, lower volatility and, at most of the time, presented better Sharpe Ratios than the market index.

In order to construct a market neutral strategy it is imperative to have a zero β portfolio. Vidya-murthy, G.(2004) demonstrates that it is possible to do that by holding both a long and a short position on different stocks. Consider that r_A as the portfolio A returns, r_B as the portfolio B returns and r as the ratio³:

$$r_A = \beta_A * R(M) + \theta_A \quad (2.1)$$

$$r_B = \beta_B * R(M) + \theta_B \quad (2.2)$$

So as to initiate the trading strategy, it is establish the r units of A as being the portfolio to short and one unit of B as the one to buy,

$$r_{AB} = -r * r_A + r_B \quad (2.3)$$

$$r_{AB} = -r * (\beta_A * R(M) + \theta_A) + \beta_B * R(M) + \theta_B \quad (2.4)$$

¹Capital Asset Pricing Model, firstly proposed by William T. Sharpe (1964), Jack Treynor (1961, 1962), John Lintner (1965) and Jan Mossin (1966).

²Market Risk

³ β_B/β_A

rearranging the terms,

$$r_{AB} = (-r * \beta_a + \beta_b) * R(M) + (-r * \theta_A + \theta_B) \quad (2.5)$$

knowing that r represents β_B/β_A the combined portfolio β is equal to:

$$\beta_{AB} = -r * \beta_A + \beta_B = 0 \quad (2.6)$$

as we wanted to prove.

Statistical arbitrage⁴ strategies are created in order to explore the inefficiencies of the market and therefore eliminate the exposure to the market. Minimizing such exposure will traduce on eliminating the β and conferring to the strategy returns non correlation to the market returns. Jarrow et al. (2003) mentioned that statistical arbitrage strategies, by definition, follow four conditions: (i) self-financing or zero initial costs; (ii) positive expected profits; (iii) probability of loss converging to zero and; (iv) time average variance also converging to zero. All the prior conditions automatically imply that such strategies will produce riskless incremental profit and positive Sharpe Ratio. In fact, the notion of statistical arbitrage does not support the idea that market is in an economic equilibrium. As is supported in Jarrow (1988) this economic equilibrium is an essential premise for an efficient market.

The efficient market hypothesis⁵ has been discussed for decades and was developed by Eugene Fama (1965,1970), Cootner (1964), Samuelson (1965), Roberts (1967) among many others in the early 1960s. EMH considers that financial markets are informational efficient and therefore it is not possible to consistently achieve excess returns on the market given the information available, at the time the investment is made. Moreover Jensen (1978) stated that market efficiency with respect to an information set, ϑ , implies that it is impossible to make economic profits⁶ by trading on the basis of ϑ .

Efficient Market Hypothesis, as explained in Fama (1970), claims that a stock price should both reflect the true value of a stock and also all the available information. Those statements imply that there is not any opportunity of arbitrage in a rational economy.

Stocks expectations can be defined by:

$$E(P_{i,t}|\Phi_{t-1}) = [1 + E(r_{i,t})|\Phi_{t-1}] P_{i,t-1} \quad (2.7)$$

⁴More details on the terminologies can be found on Khandani and Lo (2007).

⁵Now on referred as EMH and referring to its weak form.

⁶Understand by economic profits positive risk adjusted returns net of all costs.

where, E is the expected value, $P_{i,t}$ is stock's price in the period t , $r_{i,t}$ represents the return of the stock i during period t , and Φ_{t-1} is the set of information presented on the market at time $t-1$. The left-hand side of the equation exhibits the expected price of stock i given the information available at the time $t-1$. On the right-hand side, the expression denotes the expected return of stocks having an equal risk to stock i .

Now consider that $x_{i,t}$ accounts for the difference between the price of a stock i at time t and the expectation that an investor had at time $t-1$ of that stock for the next period:

$$x_{i,t} = P_{i,t} - E(P_{i,t}|\Phi_{t-1}) \quad (2.8)$$

considering that in an efficient market: $E(x_{i,t}|\Phi_{t-1}) = 0$, it is possible to infer that the information is always incorporated in stock prices and therefore the potential expectation of a stock return, according with EMH, may be represented as: $P_t = E_{t-1}P_t + \varepsilon_t$. Underlying the efficiency market hypothesis, it is opportune to mention that expected stock returns are entirely consistent with randomness in assets returns (Samuelson (1965)).

Informational efficiency holds on the market capacity of replicate, almost instantaneous, the financial reliable information on the asset prices. The *Random Walk* Hypothesis (RWH) is about independent and identically distributed (i.i.d.) of the increases, in which the price formation should be:

$$P_t = \mu + P_{t-1} + \varepsilon_t \varepsilon_t \sim \text{idd}(0, \sigma^2), \quad (2.9)$$

where μ is the expected variation of the price. According to Lucas (1978), the major difference between RWH and EMH is the risk-return trade-off. Although it is not entirely true to affirm that RWH tests can conclude about the market efficiency, those can provide valid information about weak-form efficiency questions.

The Efficient Market Hypothesis treats mainly about the (un)predictability of the share prices and the adjustment of those prices to the new available information. For Campbell, Lo and Mackinlay (1997) the natural logarithm of the prices follow a random walk with increases normally distributed. The first study that lies on this theme is the Bachelier (1900) where it is related that "Past, present and discounted future are reflected on market prices" but with regularity they appear to show non relation to the prices changes. Other important study about the subject was developed by Kendall (1953). The author implies that prices follow a random walk.

Nevertheless, there exists a lot of empirical studies that seems to contradict the EMH. Jegadeesh and Titman (1993), using one trading strategy that only uses historical well and poor performing assets, obtained an excess return of 12% per year. A similar conclusion was found by Lakonishok, Shleifer, and Vishny (1994) when using historical value versus historical glamour. Coval, Hirshleifer and Shumway (2005) found evidence that the investor's persistent abnormal returns were not due to inside information, rejecting therefore the EMH. Another study that finds violations of the EMH is the research of Cahn, Jegadeesh, and Lakonishok (1996) where it is confirmed that a strategy based only on past returns and earning announcements provide future excess returns.

EMH is, after all, only a hypothesis not a proven fact, nor it is its rejection. So many years had passed since the introduction of this hypothesis and there is still no consensus amongst researchers. About the present research, EMH is tested on the principle that: given a market and a set of information about both the market and the assets that belong in it, EMH affirms that those assets prices are at its fair price given that information; obvious the significance of such premise is that there is no way of using that information in order to produce excess risk adjusted returns, meaning that it is not possible to beat the market performance.

Among the quantitative or statistical arbitrage, Pairs Trading is a very popular strategy being used by hedge funds and investment banks but still not widely notorious as momentum and contrarian trading strategies. Such concept is actually very simple to understand: (i) Find two stocks with both historical similar characteristics and correlated paths; (ii) when the path of those stocks is disturbed and the spread between them is higher than "d"⁷ a trading signal is created; (iii) sell the stock with the higher value and buy the other one; (iv) if history repeats itself, and both prices converge, the open positions are closed and profit will be made⁸.

It all begun in the 70's at Princeton-Newport Partners, an investment company founded by the mathematical financier Edward Thorp (Ed Thorp), who was a fervent advocate of markets inefficiencies. Overwhelmed with proofs of such, he decided to verify how inefficient the market was and how it would be possible to exploit profits from that. The answer came from the statistical arbitrage toolbox and the idea was to rank the stocks performance and thus construct a market neutral portfolio by taking long and short positions on the most down and most up stocks, respectively. Ultimately due to the good performance of Princeton-Newport Partners, the idea was left behind. Disconnectedly from

⁷The value of "d" is not a value stipulated by rule but arbitrary decision of the investor.

⁸When the prices converge, according to logic, the shorted stock will reduce the value and the other will increase the value.

the work that was developed by Ed Thorp, in the early 80's at Morgan Stanley⁹, Gerry Bamberger developed another statistical arbitrage scheme similar but at the same time less variable than the Ed Thorp's. Although the initial phase translated into significant profits, it was never implemented as a major strategy on Morgan Stanley. Displeased with the non-recognition of his work, Gerry Bamberger left Morgan Stanley and Nunzio Tartaglia, come up with a team in order to improve and explore the strategy and new possibilities of profit through quantitative arbitrage strategies. Such team was composed by mathematicians, physicists and computer scientists. The result of their study was the Pairs trading. Since the first time that it was applied, such strategy has become very popular in the investment banks and hedge funds sphere.

Pairs Trading is a strategy where the "*human subjectivity had no influence whatsoever in the process of making the decision of buy and sell a stock*" (Perlin, M. (2009)). The first empirical work referring this strategy is attributed to Gatev, Goetzman and Rouwenhorst (1999) where it was documented an analysis of Pairs Trading strategy and how it affects the theory about market efficiency. Few years after Gatev et al. (2006) upgrading the original work, with more recent data, it was found proof of considerable profits uncorrelated to the S&P 500.

Gatev et al. (2006) is the most cited paper regarding Pairs Trading. The identification of the pairs is elaborated by using the distance method¹⁰. Such method was also used by Do and Faff (2009), Perlin (2009), Bolgiün, Kurun and Güven (2009) and Do and Faff (2011) when testing the performance of Pairs Trading. Perlin (2009), when examining Pairs Trading on the Brazilian market concluded that the strategy had a good performance, returns not correlated with the market and the strategy, in fact, took advantage of the market inefficiencies. Bolgiün, Kurun and Güven (2009) affirmed that the strategy produced powerful returns especially during volatile periods. Finally Do and Faff (2009), when extending Gatev et al. (2006) work stated that although profit was declining such fact could be avoided if pairs were calculated more often. More recently Do and Faff (2011) introduced accurate transaction costs and concluded that although the disappearing of high profits, Pairs Trading still presents positive returns and low risk.

However, there are others models to identify the pairs through more rigorous tests and forecasts. The cointegration approach drawn in Vidyamurthy (2004) is an attempt of parametrise Pairs Trading, by exploring the possibility of two time series, with integration of the same order d , could be linearly matched in order of generating a single time series integrated of order $d-b$, with the condition that

⁹A global financial services firm located in New York City and founded in 1935

¹⁰Minimum square distance between normalised price series

$b > 0$.

The notion of cointegration emerged with Engle and Granger (1991) by explaining that unlike correlation, cointegration is a measure of long-term dependencies. Facing the perspective of having to work with non stationary time series it is more thoughtful trying to construct portfolios that can be related to stationary time series. Such approach is highly recurrent among all asset classes and accomplished by cointegration techniques.

Engle and Granger (1991) states that two time series, x_t and y_t , are cointegrated if, and only if, each is $I(1)$ and a linear combination $x_t - \alpha - \beta y_t$, where $\beta_1 \neq 0$ is $I(0)$.

Since the appearance of cointegration views, a lot of academics and researchers uses cointegration to model the dependencies between assets. Tsay (2005) presents numerous examples of cointegration modeling time series with the respective detailed justification. Vidyamurthy (2004) tries to parametrise Pairs Trading strategy by exploring the possibility of cointegrated pairs. Alexander and Dimitriu (2002) show that statistical techniques used for Pairs Trading should present a model for the price behavior over time and that the result would depend on the presence of a long run equilibrium of the spread. Such conclusions are driven by a comprehensive back test through cointegration approach. It is also emphasised that by using this technique on trading strategies it would produce encouraging results. Lucas (1997) also find good results when dealing with the optimal asset allocation in presence of cointegrated series. Galenko, Popova and Popova (2007), using a set of cointegrated financial time series to build a new trading strategy, find out that the strategy significantly outperformed in comparison with the market. Also note that, by using cointegration for trading strategies, it implies the existence of a long term stochastic trend. In general it contradicts the hypothesis that returns follow a random walk. Such idea is supported by Lo and Mackinley (1988).

3. Data

The data consists on the 14 most liquid stocks of the Portuguese stock market, PSI-20. Such list can be viewed in Table 3.1.

Table 3.1: List of Financial Stocks used in this Research

Name	Sector of Activity	Stock Code
Banco Comercial Português	Banking	BCP
Banco Espírito Santo	Banking	BES
Banco Português de Investimento	Banking	BPI
Banif Financial Group	Banking	BNF
Brisa	Transportation Services	BRI
Cimpor	Building Materials and Fixtures	CPR
Cofina	Media Conglomerate	CFN
Energias de Portugal	Electricity	EDP
Jerónimo Martins	Food Retailers and Wholesalers	JMT
Mota-Engil	Heavy Construction	EGL
Portucel	Paper	PTI
Portugal Telecom	Fixed Line Telecommunications	PTC
Semapa	Building Materials and Fixtures	SEM
ZON Multimédia	Broadcasting and Entertainment	ZON

Liquid stock means that in the period of analysis it had at least 90% of valid close prices. This criteria is crucial since the lack of liquidity of the stocks is an enormous risk when implementing Pairs Trading strategy. This risk is characterized by higher costs as *bid-ask spread* and the difficulty of having short positions opened for more than 3 days.

The data, from 2002 until 2012, was obtained through Datastream and adjusted for splits and dividends on a daily frequency. Note that, when choosing the stocks to consider, they were filtered in order to only choose those that were listed during the all period.

Portuguese stock market is less liquid then the American and also composed by less stocks, which could constitute a problem. Since the study lies on the Portuguese market, PSI-20, and as Alexander et al. (2002) shows, efficient long short hedge strategies can be achieved with relatively few stocks, therefore this potential problem is ignored.

Additionally, hedge funds tend to only use pairs that belong to the same activity sector. However, this aspect will not be taken into consideration: stocks from different sectors will be considered pairs

if they obey to the establish criteria.

Finally it is used daily risk-free rates and Psi-20 returns for the analysis of the strategy performance and as a risk-free rate it is used Euribor rates.

4. Methodology

4.1 Pairs Formation

From academic literature is possible to identify 4 methods of pairs selection. The first one is the Square Distance approach implemented on the works of Gatev et al.(1998,2006) and Do and Faff (2003) among others. As it is explained with more detail on the next section, this approach merely searches a statistical relationship within pairs. The second major approach is know as Cointegration approach which incorporates the mean reversion on Pairs Trading structure, simply the most significant statistical relationship in order to succeed. Such approach is introduced on the researches of Vidyamurthy (2004) and Alexander and Dimitriu (2002) inter alia. Those two approaches mentioned above are the two methods that are included and tested in this research. Finally the last two approaches mentioned on the literature are the Stochastic and the Combined Forecast approaches. The first one is included in Elliot et al. (2005) and Do et al. (2006). Such approach models the mean reversion behavior of the spread and is widely defended due to the ability of capturing it on a continuous time model and with parameters easy to estimate. Nevertheless this approach have a fundamental limitation on restricting the long run relationship of two pairs to one of return parity (Do et al. (2006)). The last approach is encouraged by Huck (2009,2010) given that provides more trading chances. Such approach differs from the others since it does not make any reference to an equilibrium method.

4.1.1 Square Distance Approach

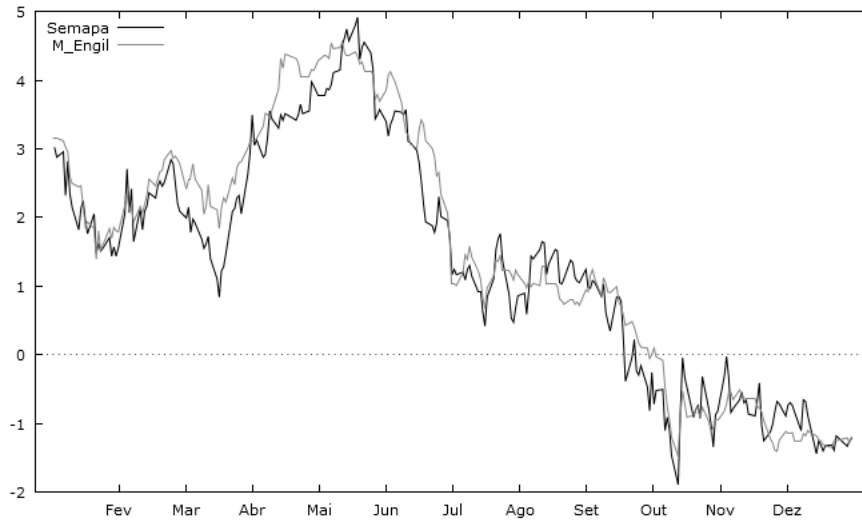
It is essential that the price series pass through a normalisation procedure, $P^n = \frac{P_i - E(P_i)}{\sigma_i}$, so that they are brought to the same standard unit in order to permit a quantitatively fair choice of the pairs. This procedure is very important in order to avoid the impossibility of setting a distance trigger¹ on the pair. After this procedure, using the minimum distance rule, presented on Equation(4.1), each stock respective pair that minimises the square difference between the normalised prices is searched.

$$\text{Min}(\sum_1^t P_i^n - \sum_1^t P_j^n)^2, i = 1, 2, \dots, n \wedge j = 1, 2, \dots, n \wedge i \neq j \quad (4.1)$$

¹Meaning the impossibility of knowing when to open or close the trade of the pairs.

As Gatev et al. (2006) points, pairs traders normally search for two stocks whose prices followed the same patterns. This suggests that square distance approach is the most similar to how traders find their pairs. That said, the minimum distance approach will be used to match stocks with their pairs as shown in Figure 4.1.

Figure 4.1: Pairs at the Selection Period



Afterwards is important to set the distance, d , that will trigger the trade. d value, as referred before, is arbitrary to the investor will. Nevertheless its important to take into account some considerations: (i) if the value of d is very high probably no trade will take place (ii) if the value of d is very low, there will be a significant number of transactions increasing the transaction costs. Hereupon the value of d will be 2 standard deviation², meaning that every time the absolute distance³ is higher than 2 a trade will take place⁴. Although this value of 2 standard deviations is used as standard, it is also interesting to test other values to unleash the trade. Having in mind all the considerations, this research verifies such ideas by exploring the results of exercising the strategy with different d .

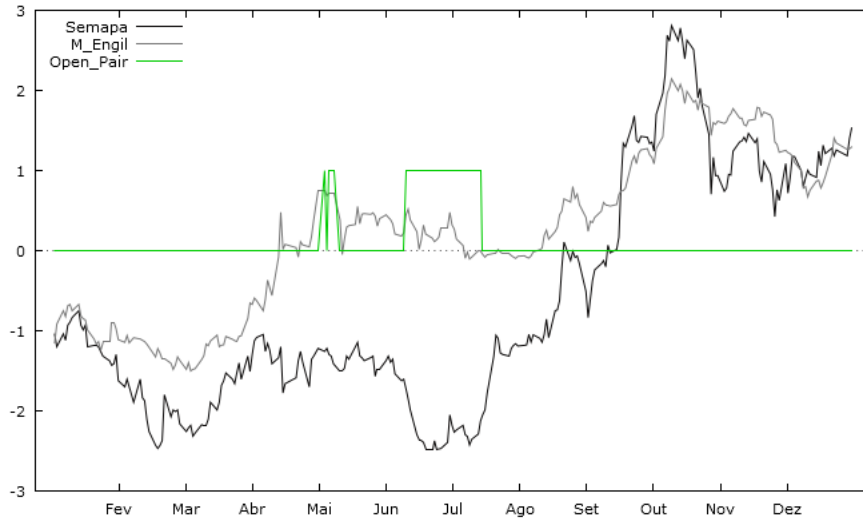
Note that also in order to be a self-financing and market neutral strategy, there will always be a ratio, P_b/P_a , to confer the same weight to both bought and sold assets. The Figure 4.2 shows the behavior of one pair during trading period and the respective trades that are opened.

²This value is the most common on the literature and therefore, before analysing which value maximises the strategy, this 2 standard deviation value is used as default.

³By absolute distance it is meant the spread.

⁴The price that will be consider on the trade will be the last price of the day.

Figure 4.2: Pairs at the Trading Period



4.1.2 Cointegration Approach

The main challenge in this type of strategies is to identify the stocks that tend to move together over time and therefore make potential pairs. It is known on advance that if an asset price series is stationary it will respond with success to a mean reverting strategy. But we also know that asset price series are not stationary. Nevertheless we can build a stationary series with two cointegrated asset time series. Cointegration theory is definitely an innovation in theoretical econometrics. Unlike correlation, it does not refer to the co-movements in returns but to co-movements in raw asset prices.

Let y_t and x_t be two nonstationary time series. If in a certain moment of y , the series $y_t - \gamma x_t$ is stationary we can say that y and x are cointegrated. The idea behind cointegration is that y and x have a long-run equilibrium that when disturbed from the long-run mean it is expected that, one or both time series, adjust and return to the equilibrium.

Given a time series that is not stationary and only stationary after differing n times can be called "integrated of order n ", denoted $I(n)$. Consider a random walk, $y_t = \alpha + y_{t-1} + \varepsilon_t$, is a $I(1)$ where α denotes the drift and the error process ε_t is i.i.d. stationary process. Said that, x and y are cointegrated if $x, y \sim I(1)$ but there exists α such that $z = x - \alpha y \sim I(0)$. Cointegration is a statistical relation of two time series that are integrated on the same order, n , and can be linear combined to produce on single series. In the specific case of a Pairs Trading strategy, the series are combined to produce a series $I(0)$.

By definition a x_t series is I(1) if the first difference is stationary, I(0). Considering $x_{1t}, x_{2t}, \dots, x_{kt}$ a time series I(1); if there are $\beta_1, \beta_2, \dots, \beta_k \neq 0$ such that $\beta_1 x_{1t}, \beta_2 x_{2t}, \dots, \beta_k x_{kt}$ results in one series I(0), than $x_{1t}, x_{2t}, \dots, x_{kt}$ are said to be cointegrated⁵.

Assuming a long-short position in two time series that are not cointegrated it is possible that the value of the pair is stationary, meaning that they are cointegrated. If the time series is stationary, one strategy based on mean reverting is profitable, at least until the stationary status is not interrupted.

Applying the cointegration idea on the Pairs Trading analysis in order to detect potential pairs: (I) Using the Augmented Dikey Fuller test, we verify if all asset time series are I(1); (II) After confirming the nonstationary status of the time series is used a Engle-Grager or a Johansen test on all possible combinations to verify which pairs are cointegrated; (III) Having the pairs it is formulated a linear combination between the two time series to obtain the β of the regression.

In order to start the strategy it is also necessary do calculate the spread between the time series,

$$P_{xt} - \beta P_{yt} = Z_t \quad (4.2)$$

where β is the cointegration coefficient and Z_t is the value of the spread over time. The spread is mean reverting giving that both time series are cointegrated. Having already the spread it is done a normalisation procedure by its historical mean and standard deviation. Once again it is indispensable to settle the rule to trigger the trade: sell (buy) the spread every time it deviates above (below) two standard deviation of the mean and close the trade when the spread reverts to less than 1 standard deviation of the mean.

4.2 Returns Computation

In order to conclude if the strategy created value and be able to take conclusions about Pairs Trading it is necessary to calculate the total return of the strategy. The following equation represents the raw return of the strategy, R_{raw} .

$$R_{raw} = \sum_{t=1}^T \sum_{i=1}^n \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) P_{i,t}^{LS} \quad (4.3)$$

where,

⁵For more detailed information see Hamilton(1994), Hendy and Jeselius (2001) and Enders (2002).

T = Number of observations in the period;

n = Number o stocks considered in the period and;

P^{LS} = Dummy variable that takes the value of 1 for a long position and -1 for a short position.

Such equation means that each time the strategy opens a position, the raw return return is $R_{i,t} = \ln(P_{i,t}/P_{i,t-1})$, $i \in 1, \dots, n$, controlled by the dummy P^{LS} that confers the position on the market.

Although the results of the whole strategy are considered as being before transaction costs, it seems convenient to include them on the formula:

$$RTc = \ln\left(\frac{1-c}{1+c}\right) \left(\sum_{t=1}^T \sum_{i=1}^n Tc_{i,t} \right) \quad (4.4)$$

where,

c = transaction cost per trade and;

Tc = Dummy variable that takes the value of 1 if a transaction is made and 0 otherwise.

joining both equations,

$$R = \ln(P_{i,t}/P_{i,t-1}) + \ln\left(\frac{1-c}{1+c}\right) \Leftrightarrow \quad (4.5)$$

$$\Leftrightarrow R = \ln\left(\frac{P_{i,t}(1-c)}{P_{i,t-1}(1+c)}\right) \Leftrightarrow \quad (4.6)$$

$$\Leftrightarrow R = R_{i,t} + R_{Tc} \quad (4.7)$$

But as it was explained before, Pairs Trading strategy always performs a transaction with a long and a short position in two different assets, the stock and its pair. Said that, next equation incorporates this long-short position on the returns computation:

$$R_T^{LS} = \sum_{t=1}^T \left[\sum_{i=1}^n \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) P_{i,t}^{LS} + \sum_{j=1}^n \ln\left(\frac{P_{j,t}}{P_{j,t-1}}\right) P_{j,t}^{LS} \right] * (d) + \ln\left(\frac{1-c}{1+c}\right) \left(\sum_{t=1}^T \sum_{i,j=1}^n Tc_{i,j,t} \right), \quad \begin{cases} i \neq j; \\ \text{if } P_{i,t}^{LS} = 1 \Rightarrow P_{j,t}^{LS} = -1; \\ \text{if } P_{i,t}^{LS} = -1 \Rightarrow P_{j,t}^{LS} = 1. \end{cases} \quad (4.8)$$

$$\text{where, } d = \begin{cases} 1 & \text{if } \text{Absolute}(P_{i_{t-1}}^n - P_{j_{t-1}}^n)^2 > 2\sigma \\ 0 & \text{otherwise} \end{cases}$$

Note that it refers to the case where it is used the Mini-

imum Distance Method. If it is performed the Cointegration Method, $d = \begin{cases} 1 & \text{if } \text{Absolute}(P_{i_{t-1}} - \beta P_{j_{t-1}}) > 2 \\ 0 & \text{otherwise.} \end{cases}$

4.3 Transaction Costs

When working with strategies that involve trading stocks it is important to have in mind the impact coming from the transaction costs. Such term can be viewed as a necessary loss while pursuing the potential gains. Nevertheless it is a topic that concerns each type of investor at a different level. Once Pairs Trading is based on trading stocks (both long and short positions), this concern might have impact on the strategy result depending on who is trading.

Although transaction costs are the *cost of carrying out a transaction by means of an exchange on the open market* their value highly differ whether it is a small investor or it is an investment bank. For instance for a single small investor the transactions costs will have a great impact on the results since the trading scale will not compensate the costs. But on the other hand, the impact of transaction costs for a hedge fund will not be so critical as for the small investor. Hedge funds place huge amounts of trading orders and as consequence the transaction costs have a lower impact on the final result of their trades. Finally, it is important to highlight that for the investment banks the impact of the explicit transaction costs might be null. This consideration is sustained by the fact that banks are indeed the market makers and therefore banks will not pay commissions or fees for something that is done by themselves. In this present thesis, due to the difficulty of accounting for correct transaction costs, the research pursues a theoretical point of view, assuming 0 transaction costs, although it may not waive away much of the reality given the fact explained above.

There are explicit costs and implicit costs associated with this kind of strategies.

4.3.1 Explicit Costs

On this section there are two types of transaction costs: commissions and short selling costs. Commissions are the fee paid to an agent or a trading company for the services of facilitating transactions between buyers and sellers. Short selling costs comes from the procedure of selling an asset not owned by the investor with the intention of have profit later when the asset is bought back at a lower price. The investor does not own the asset, therefore he has to borrow it and that has a cost.

4.3.2 Implicit Costs

The implicit costs are the bid-ask spread. This cost is undoubtedly the most difficult to measure and has a great presence on this specific strategy. There are three considerations to make about this theme in order to explain exactly what is this cost about. Primary the bid-ask spread is the quantity by which the ask price is exceeding the bid price. Such definition is in actual fact the difference between the highest and lowest price that the market is willing to pay for buying and selling an specific asset, respectively. Secondly it is important

to have in mind that if a stock is quoted on the market at 50 euros and at that time is placed a buy order to the market, if the stocks tick size is 1 euros, the stock was in reality bought at 51 euros. Here it is described a cost of 1 euros that when working with stocks quotes it is not considered and therefore it has impact on the strategy results. Finally, consider that the stock and its pair, at a precise moment, diverge from the path by two standard deviations and that trigger the trade. The one that is bought is the one that was losing value and the one that is shorted is the one gaining value. The result of this finding is that it is more likely that a stock is bought at the quoted bid price when it should be at an ask price and the other is shorted at the quoted ask price when it should have been shorted at a bid. This inconsistencies also generate a bid-ask spread that affects the strategy performance.

4.4 Assessing Pairs Trading Performance

The idea behind this section is to verify if Pairs Trading has any science or it is all due to an issue of luck. Such analysis is performed in two different ways: (i) attribute random pairs and (ii) simulate random entries to simulate the values that d takes over the time.

In the first test, attributing randomly the pairs, it is possible to test the importance of realising, for instance, the square distance method and then verify whether the trading performance is in fact due to the strategy or is just an issue of luck. In the second test, there are simulated randomly entries and retrieved the median to trigger the trade. Both analysis have a critical importance to verify the part of Pairs Trading strategy return that is not due to mean reversion.

Testing the strategies mean reversion, the market neutrality and verifying the annual returns are the major conclusions of this section. At the random pairs strategy there are used the actual stock prices vectors and therefore the trade triggers when its normalised prices separates for more than two standard deviation unit. Here the test relies on the importance of calculating the suitable pairs in order to explore the market inefficiencies. At the random trigger strategy, the pairs are the most suitable, calculated by the minimum square distance method, but the trade do not bases upon the deviation of the long-term relation of the normalised pairs. Here the test lies on the importance of waiting for the divergence of the pairs behavior in order to explore the mean reversion possible profit.

In brief the goal of such tests is to verify if Pairs Trading performance is due to luck or the strategy is able to explore the market inefficiencies.

4.5 Efficiency Tests

According to the efficient market hypothesis (EMH), in its weak-form, it is impossible to forecast future prices or returns based on historical information. Thus, there should not be possible to find a situation where exists a linear dependence of future returns in relation to past performances.

In order to provide evidence on the informational efficiency of the Portuguese stock market, this Thesis have two different approaches: The Run test and the Autocorrelation test.

Also known as Wald Wolfowitz test⁶, the Runs Test is a nonparametric test addressed to the analysis of the randomness hypothesis. A Run + (−) is a sequence of positive (negative) returns. The null hypothesis is that the two elements + and − are independently drawn from the same distribution.

Such test is realised through the comparison of the obtained number of runs in the series of returns with the expected number of runs, $E(R)$. Under the null hypothesis, the number of runs in the sequence is a random variable with a conditional distribution determined by the observation of n_+ positive values and n_- negative values. The Run Test can be defined as: H_0 : the sequence was generated randomly; H_1 : the sequence was not generated randomly and; with a test statistic:

$$Z = (R - E(R)) / \sqrt{\text{var}(r)}, \quad (4.9)$$

approximately a normal distribution, computed as: $\mu = \frac{2n_+n_-}{N} + 1$ and $\sigma^2 = \frac{(\mu-1)(\mu-2)}{N-1}$.

The main question that this test answers is if this series was generated from a random process, this is, if the series follow the random walk model.

The Autocorrelation test, sometimes called as *lagged correlation* or *serial correlation*, examines the relationship between random variables within past periods and current periods, based on different ranges of lags. Such test may be used to test the dependency or not of a time series but also to identify a suitable model if that series is not generated randomly. The statistic test Q , known as *Box Pierce* is given by $Q_k = n(n+2) \sum_{k=1}^n \frac{\widehat{\rho}_k^2}{n-k}$, where: H_0 : The data is independently distributed, meaning that, if there is any observed correlations in the data it would only result from randomness; H_1 : The data is not independently distributed and; $\widehat{\rho}_k$ is the series autocorrelation at lag k with h lags to test, such that:

$$\widehat{\rho}_k = \frac{\sum_{k=1}^{n-i} (r_k - \widehat{\mu})(r_{k+i} - \widehat{\mu})}{\sum_{k=1}^n (r_k - \widehat{\mu})} \quad (4.10)$$

Note that the metric Q_k follows an asymptotic distribution when the process returns are i.i.d.. That being, when the RWH is verified the expected value of Q_k should exceed $E(\chi_k^2) = k$.

⁶The name comes from its developers Abraham Wald and Jacob Wolfowitz.

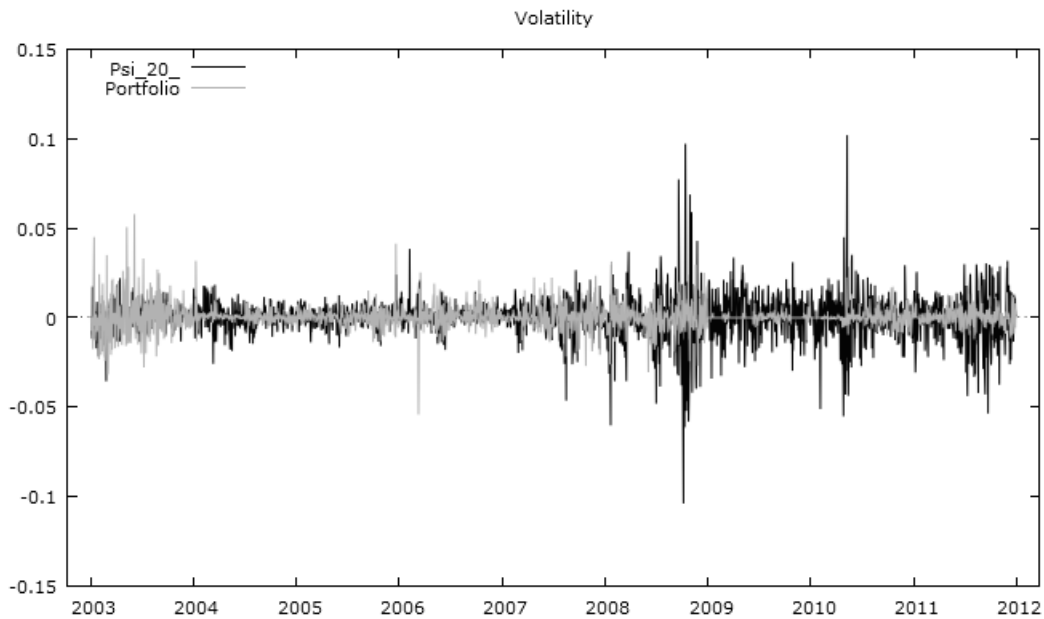
5. Results

5.1 Square Distance Approach

Table 7.1, presented on the next page, reveals the stocks and its pairs over the observation period. The stock that over time was the most profitable was *BES*¹ earning more than 330%, representing about 37% per year. On the opposite side stands the stock *Mota Engil* with losses on the order of -47%, meaning -5.24% per year. Another aspect that is possible to observe is that the stock *BES* had the same pair four times during the period while *BCP*², *Cimpor*, *Cofina*, *Portucel* and *Zon* never held the same pair for more than one year.

The strategy performed encouragingly when compared to the market performance. Figure 5.1 and 5.2 replicate the volatility during the period and the accumulated returns, respectively.

Figure 5.1: Volatility



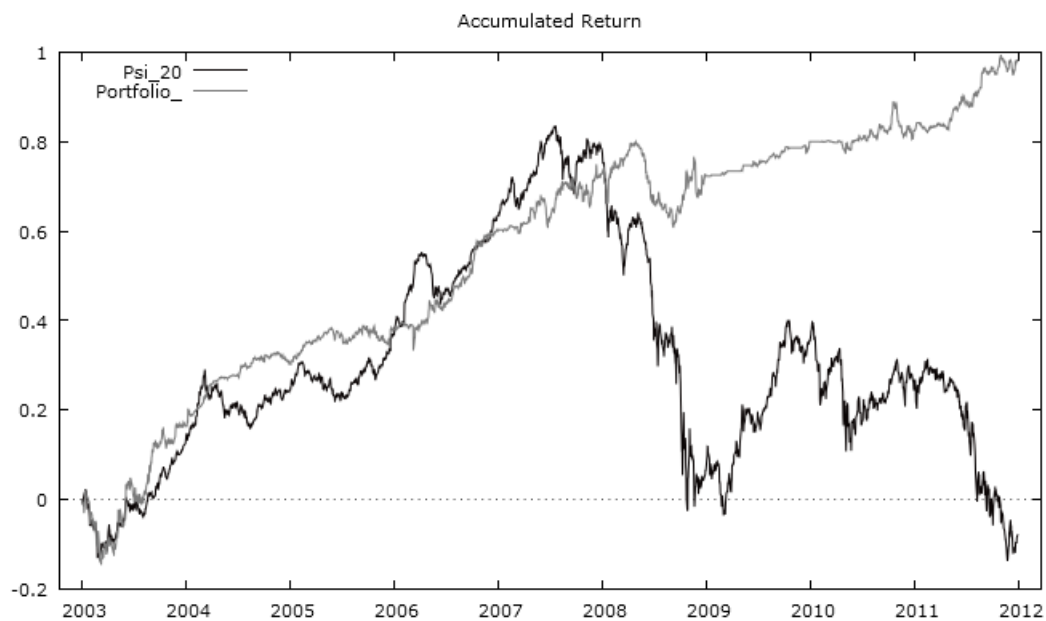
(i) Pairs Trading annual Volatility=8.40% (ii) Portuguese index Market annual Volatility=15.84%

From Figure 5.1 it is possible to observe that, as expected from a strategy like Pairs Trading, the volatility is very low, lower than the market and also with a time average variance converging to zero. It is also possible

¹Banco Espírito Santo

²Banco Comercial Português

Figure 5.2: Pairs Trading Portfolio versus Market Performance



(i) Pairs Trading annual Return=10.91% (ii) Portuguese index Market annual Return=-0.88%

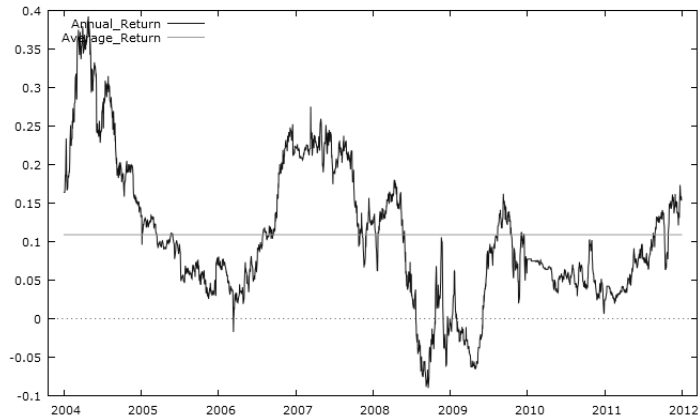
to enhance that the lowest observed volatility was on the year of 2009, year after the world crisis. Through Figure 5.2 it is possible to verify that for about 75% of the study temporal horizon, Pairs Trading strategy outperformed the benchmark³ finishing the analyses period with a profit of 98.2%. Once again, Pairs Trading distinguish form the Portuguese market during and after 2008, when it continued to achieve returns as if there had been no crisis whatsoever unlike the Portuguese index Market that accumulated huge losses. In brief, the both annual return and the annual volatility were, respectively, 10.91% and 8.4% for Pairs Trading strategy and for the benchmark -0.88% and 15.84%.

For a best view of the performance, Figure 5.6 exhibits the annual returns of the strategy over the analysis period. The contribute of presenting 5.6 is to have a clear view of the annual strategy outcome. As it is possible to observe, only approximately 8% of the all possible annual investments, within the observation period, presented a negative performance. This means that, for an investment with a maturity of one year, 92% of the days would generate positive returns.

The strategy returns distributed by year and month is represented on Table 7.2. It is possible to verify that the best year was 2006 with a total of 22.29% and the best month October with profits of 24.41%. One of the reasons why it was the best year of the whole period is probably due to the fact that it was the year where occurred most transactions, meaning that the mean reversion had more effect. By contrast the worst

³Benchmark refers to the Portuguese index market, PSI-20.

Figure 5.3: Pairs Trading Portfolio Annual Returns



The figure presents the one year return of the investment. For instance, the return presented on the May first of 2006 refers to an one year investment that took place on the May first of 2005.

period was 2008 and June, with losses of -0.57% and -6.2%, respectively. One of the reasons why in 2008 Pairs Trading did not have profits may be given to the world crisis that affected all stock markets to fell down, disturbing the path of the pairs and resulting on flaws of the mean reversion.

The Table 7.3 summarises the returns before transaction costs from Pairs Trading strategy using the square distance approach and the performance of the Portuguese benchmark. As it is possible to conclude, Pairs Trading outperformed the market achieving an annual return of approximately 11%. Such performance results from constant high positive returns within the years of observation, outperforming more than 100% above the market. As it is possible to conclude from the Table 5.3, the results obtained from Pairs Trading strategy were not driven by a higher risk. Once again, when comparing to the benchmark, Pairs Trading Sharpe Ratio was considerable higher than that of Portuguese market, Psi-20, obtaining a value of 1,43 versus 0,09 , respectively. While the annual return is an relevant metric when verifying the strategy performance because it displays the total amount of profit we may expect from a future investment it does not refers to the possible risk⁴ on assessing such profit. That said the best way on assessing the trading performance will be the Sharpe Ratio. Such metric is the excess return of a strategy by one unit of risk. Since both strategies⁵ have the same time horizon there is no need of annualizing the Sharpe Ratio. The consequence of Pairs Trading Sharpe Ratio being bigger than the market's (meaning that this strategy gives more return for unity of risk) is the contradiction of the efficient market hypothesis.

Table 7.4 shows the statistic inherent to Pairs Trading strategy. Beta, as well known as systematic risk, is, with the exception of the first year, always quite close to zero and not statistically significant at 1%. On the other hand, the alpha, often considered to represent the excess return of the strategy regarding the return of

⁴It is known that the higher the risk the greater will be the profit required.

⁵The Pairs Trading strategy and the benchmark.

the market, although not always statistically significant when looking year by year, it is considerable robust when analysing the whole period.

The results obtained support the idea that Pairs Trading is a market neutral strategy or in another words that the returns do not follow market behavior. The risk-adjusted measure, alpha, is higher than zero representing an annual value of about 11% which indicates that the investment annual return exceeded that from investments with a similar risk. Once again, the EMH is violated since that for being an efficient market, the expect value o alpha shall be zero. Table 7.5 exhibit the Pairs Trading summary statistic and reports that its returns do not follow a normal distribution as it is expected since time series are characterised by excess of Kurtosis. The presented high Kurtosis (10,7535) reflects that the distribution of the returns is leptokurtic. If the returns were normally distributed the value of Kurtosis should be 3. Also note that the strategy is positive skewed, meaning that it has more positive returns than negative ones. In order to be normally distributed it should present a value of Skewness equal to zero.

5.2 Pairs Trading under different d

As explained before the standard strategy is based on a value of d equal to two standard deviation. In order to verify which is the best value for d it was conducted a sensitive analysis. Table 7.6 summarises the results obtained from the different strategies. Although it appears that the lower the value of d the best the strategy outperforms such might not be entirely true. As an investor or an investment fund the result of a strategy should not be only led by the final return but by the risk-adjusted return and in this case the best value of d still prevails two standard deviations.

Another important note that might have impact on the performance of the strategy regrading the value of d is the number of transactions that will affect the final performance. The lower the value of d the higher will be the transaction costs. By analysing the table it is possible to conclude that in terms of total return the value of d that maximises the strategy performance is "0.5". Such value also performed the highest and lowest daily return. By contrast d equal to "0.5" had the most amount of days with negative returns. Nevertheless, since the metric in the present research that is considered to be the most adequate to evaluate performances is the Sharpe Ratio, a d of two standard deviations is the value that exploits the most the strategy potential since it maximises the excess return per unit of risk. Once again, there are provided, on Table 7.7⁶, the summary statistic of each strategy.

⁶The Table is displayed at the end of this chapter.

5.3 Assessing Pairs Trading Performance

When analysing this kind of strategies that explore market inefficiencies in order to achieve excess returns, it is prudent to verify if it did in fact explore market inefficiencies or it was just a question of luck.

Table 7.8 provides information confirming that Pairs Trading strategy is not an issue of luck. From the Random Trigger test, the annual returns were about -0.02% and achieved a Sharpe Ratio of -0.15. Such results indicate that, although it performed better than the benchmark, it did not present both excessive and positive returns. One reason that possibly explains that outcome is the inability of capturing temporal variations in returns.

In order to reject once more the idea that Pairs Trading strategy performance might be an issue of luck, Table 7.9 may present further proof. As evidenced this Random Trigger strategy followed the market, rejecting both the null hypothesis of β equal to zero with significance of 1% and the view that by taking a long-short position it would lead to a market neutral strategy. It is also evident that this Random Trigger strategy does not violate the EMH: an α equal to zero implies that the investment returns were consistent with the risk undertaken and therefore the market was efficient.

The other test for assessing if Pairs Trading strategy was an issue of luck, as reported before, is the Random Pairs strategy.

When the pairs are chosen randomly, the strategy result is a disaster, recording losses of -8,17% per year and a Sharpe Ratio of -0,19.

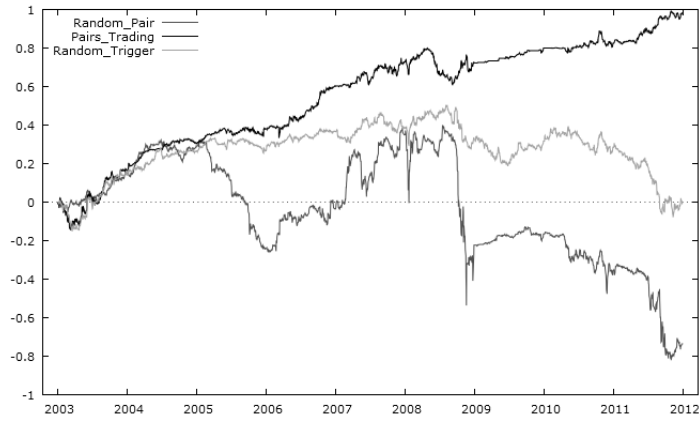
Table 7.11 presents the results of this specific strategy regarding the market behavior. The value of a β equal to 0,025 with a p-value of 0,33 do not drive to the rejection of the null hypothesis⁷. The measure of the risk adjusted return, α , is equal to -0,0003 and, once again, with a p-Value of 0.30 the null hypothesis can not be rejected. This is possible to verify that it does not follow the market behavior and also it does present a negative excess return. In brief, by attributing randomly the pairs, the strategy underperformed when comparing both to the Pairs Trading strategy and the market, implying that although it proved to be a market neutral strategy their mean reversion did not perform, conducting to continuous losses specially since the end of 2008. Figures 5.4 and 5.5 demonstrates the performance of these issue of luck strategies, one comparing with the Pairs Trading strategy and the other with the market.

5.4 Efficiency Tests

Pairs Trading bases their possibility of making positive excess returns on the inefficiencies of the market. Although it is not a core content of this research, testing the Portuguese Market efficiency, with random walk

⁷Null hypothesis: $\beta=0$.

Figure 5.4: Pairs Trading and the Issue of Luck Analysis



The figure presents the comparison between Pairs Trading strategy and the two types of strategies that try to verify if the mean reversion of Pairs Trading is just an issue of luck.

Figure 5.5: PSI-20 and the Issue of Luck Analysis



The figure presents the comparison between the two types of strategies, that tries to verify if the mean reversion of Pairs Trading is just an issue of luck, and the Portuguese Index Market, PSI-20.

hypothesis tests, can provide more explanations about how did the strategy performed so well.

The importance of conducting the Efficient Market Hypothesis emerged due to the substantial empirical implications of the Random Walk tests. One of main focus of the RWH studies relates to the distribution shape of prices variations. The hypothesis affirms that the daily returns are random variables Independent and identically distributed (I.I.D) rejecting therefore Fama (1965), that concludes that the best distribution of the daily returns is the non-normal. Nevertheless many economists have been reluctant on accepting those results, mainly due to the variety of statistical techniques available on processing normal variables.

Observing the Autocorrelation Test, the results are conclusive. Testing the autocorrelation for 30 lagged time periods it is found evidence of dependency over the historical time series, indicating therefore that the market was not efficient by rejecting the random walk hypothesis.

Each of the 14 stocks included for analysis, on average, on all the 30 time lags, the absence of autocorrelation is rejected on about 47,143% 65,000% 76,429% of the lags with, respectively, 1%, 5% and 10% of significance.

The Run Test, did not present such strong evidence of rejecting the efficiency of the market when testing with respect to the mean return. In all the 14 stocks, only 3, 5 and 7 stocks rejected the independence and randomness of the returns with a respective level of significance of 1\$, 5% and 10%. Hereby, it is not clear to affirm that on a overall basis, the null hypothesis may be rejected.

In brief, the Autocorrelation test presents evidence that the market did not followed a random walk, hence, it was not weak form efficient. On the Run test, although not possible to affirm the same, it is clear that is presented some inefficiencies. Both tests indicates that a strategy based on past information could have outperformed the market and that Pairs Trading performance is consistent with these findings.

5.5 Cointegration Method

Such as the results of the Minimum Distance method (MDm), the Cointegration method outperformed the market performance. The strategy produced an annual return of 13% from 2004 until 2011 that represents a total return of 106%.

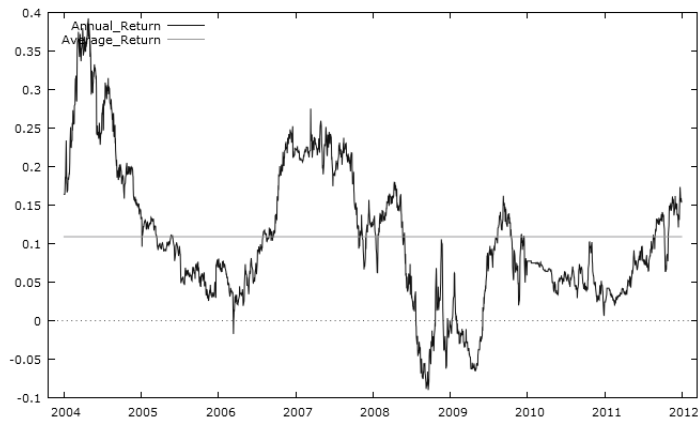
Although it presents an higher annual return than the strategy done by the MDm, the risk-adjusted return metric, Sharpe Ratio, is lower, presenting an annual value of 1,31 versus 1,43. It is not only the Sharpe Ratio that is worst but also the percentage of annual investments that result on a positive performances, as it is possible to observe in Figure 5.6. Through the Cointegration method, the percentage of negative annual investment returns increases from 9,75%, in the MDm,⁸ to 25,18%.

In brief, Pairs Trading conducted by the Cointegration method also produced positive excess returns. In general, this method had a behaviour similar to the MDm although it presented a major difference. In this

⁸Note that for the Cointegration Method the time horizon is different since that for selecting the pairs it is used 2 historical years.

case, the strategy presented a $\beta=0,135$ with a t-Statistic of 6.991 that rejects the null hypothesis with 1% of significance. Hence, the strategy was not market neutral.

Figure 5.6: Pairs Trading Portfolio Annual Returns



The figure presents the one year return of the investment. For instance, the return presented on the May first of 2006 refers to an one year investment that took place on the May first of 2005.

7.14 details some of the information that can be used to analyse the performance of the Pairs Trading under the Cointegration method. First of all, the volatility of this method indicates that it would be preferable to choose the Minimum Distance method because here the volatility exceeded the 10% ending up with more 3% that the other method. Second, and as a result of the high volatility, the Sharpe Ratio, as said before, was worst in the Cointegration method, although better than the market.

The information on the maximum and minimum return, is a little ambiguous. The concern on introducing this type of information is to observe how the methods behaved e terms of return, this is, to see how one days return could affect the total return of the strategy. While Minimum Distance method is more conserved in that point, the Cointegration method had a maximum return of 18% that in any other case could be the minimum return. Said that, although presenting the highest daily return, the Cointegration method should not be preferred over the other just based on this information.

Finally, the % of days with positive and null returns indicates us two things: First that the Cointegration method had more than 50% of the returns positive, contributing for the high total return. The second note that it is possible to find here is that, Minimum Distance method, having the most % of days with null returns, did not trade as intensive as the Cointegration method, meaning that the strategy behaved more likely as Pairs Trading is suppose to and by that it is meant that, probably, this method lost the stationary of its pairs along the time.

5.6 Updated Data - The first 3 Quarters of 2012

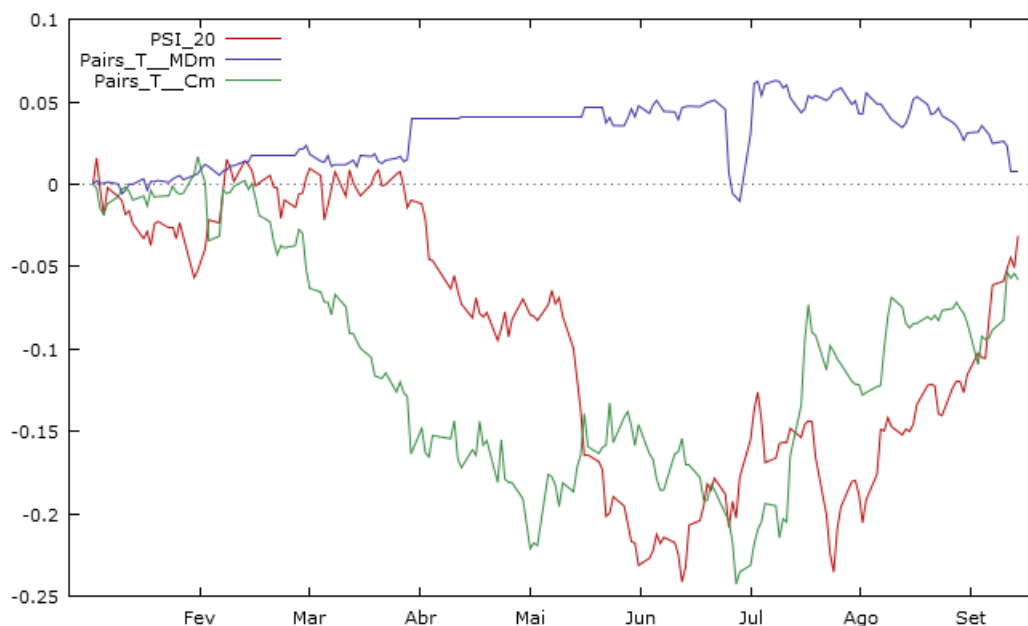
In order to update the analysis of the present thesis, the data from the year 2012 were withdrawn until the 14 of September.

The evolution of the Portuguese economy in 2012 is marked by the continued application of economic assistance and international financial help conducted in 2011. This request has become unavoidable due to the progressive deterioration of access to international funding markets. The framework of the Portuguese economy began a situation in which uncertainty has been dominant. This has penalised the corporate investment cycle operating in Portugal and attracting new investment projects.

As it is possible to observe from 5.7 the behavior from the market produced a return of -3,18% in this first 9 months. This achievement was only possible by the sharp increase after August where PSI-20 raised 15%.

Observing the Pairs Trading performance on both methods, the results are not only different but also not appealing. By the Minimum Distance method, the strategy had a stable behavior but never exceeded the 6,30% of accumulated return. The final 9 months return was of 0,81%. The Cointegration method did not lead to a positive return. With a behaviour very similar to the market, the strategy resulted on losses of 5,77% after the 9 months and presenting only 40% of positive daily returns. This value contrasts with the MDm, where this value is exceeds the 65%.

Figure 5.7: 2012 Cumulative Return



(i) Pairs Trading Return by Mdm=0,81% (ii) Pairs Trading Return by Cm=-5,77% (iii) Portuguese index Market Return =-3,18%

Regarding 7.15 the conclusion is that in this three Quarters of 2012 the strategy was not able to beat the market in terms of Sharpe Ratio. As it is possible to observe, the values are all very close and negative. In terms of volatility, the Minimum Distance method produced the best outcome, presenting a volatility of 9%, approximately half of what Cointegration method and the Market registered. Note that, May was the worst month performance of the market (-15,19%), the second best with the Cointegration method (7,52%) and neutral to the Minimum Distance method (0,00%).

6. Conclusions and Implications

The present research proposes the implementation of a strategy based on quantitative analysis or statistical arbitrage, widely known among the hedge funds and the investment banks, known as Pairs Trading. Such proposal is applied to the Portuguese Index Market (PSI-20) and is conducted by two different approaches, the Minimum Distance method (MDm) and the Cointegration method, in order to verify which explores better the characteristics of the mean reversion of the pairs.

The main goal of implementing the strategy is to compare both the performance of Pairs Trading against the PSI-20 return and add more information about the efficiency of the Market. Although those two points are the major objectives, there are also presented in this research additional information about the characteristics of the strategy. For this purpose there are realised an analysis considering different values for the trigger parameter d and also two different variants for assessing the Pairs Trading performance.

The first notorious conclusion is that for both methods the strategy outperformed the Market achieving an annual return of 10,91% and 13,26%, with the MDm and Cointegration method respectively, versus the -0,88% achieved by the Market. Overall, for the respective analysis period, the MDm ended the year of 2011 with a total return of about 98% which is very close to the 106% obtained with the Cointegration method. This outstanding performances are more notorious when compared to the total return of -15% accomplished by the Market.

The MDm performance proved to be non related to the Markets, presenting a zero β by rejecting the null hypothesis which means that Pairs Trading may be called a Market Neutral strategy. In this analysis stage it is also relevant to refer that the strategy achieved an alpha, this is, an excess return of the strategy regarding the return of the market, of about 11% a year, which indicates that the investment annual return exceeded that from investments with a similar risk. Here it is possible to conclude that the EMH is violated since that in order to be considered a weak-form efficient market, the expected value of alpha, in this type of strategy, should be zero. The Sharpe Ratio, excess return of a strategy by one unit of risk, was considerably higher than that of Portuguese Market, Psi-20, obtaining a value of 1,43 (MDm) versus 0,09, respectively.

On the other hand, the Cointegration method, presented a β statistically significant, equal to 0,135. This results imply that by elaborating the Pairs Trading strategy, in the Portuguese Market, with such approach, ultimately, the strategy can not be defined as a Market Neutral Strategy. Although this major difference from the result obtained with the MDm, the alpha presented a value of 12,5%, supporting once again the idea that the market present inefficiencies on its weak-form. In terms of risk-adjusted return, the Cointegration method presented lower values than the MDm. Thus, the value of the Sharpe Ratio was of 1,31.

From the analysis of the value of the trigger rule, the results were consistent with the literature concluding as the most profitable a d equal to 2 standard deviations. When assessing the performance of Pairs Trading strategy, it is possible to conclude that the results that were obtained were not due to an issue of luck but else to the proficiency of the strategy.

When analysing the first 9 months of 2012 it is possible to conclude that the Minimum Distance method is, in fact, the best approach to conduct Pairs Trading strategy. The fact that 65% of the daily returns were positive and in the Cointegration method there was only 40%, indicates that in times of crisis the mean reversion was not profitable because the stationary does not persist.

Another conclusion that is possible to retrieve is that in trouble times the financial markets tend to become more efficient. In this case, and based on the MDm, both 23% of the days had not recorded any transaction and also 9 out of the 14 stocks did not traded in the all period. This last statement indicates that there was not any inefficiency that could be seized by Pairs Trading.

The overall results confirms both that: it is possible to achieve excess returns by using past information; that there is still space for profiting by using trading strategies that are able to explore the market inefficiencies; periods of crisis and low volatility tend to low the strategy's performance; the adoption of the Minimum Distance method turns the strategy less riskier and not that less profitable and finally; that all the variables that were used in order to perform the strategy are, without a fact, crucial to the strategy.

It is also crucial to address one limitation of this work, that in some cases, may lead to different conclusions. Such limitation is that the performance results are all before transaction costs. For further investigation, it would be interesting if the Pairs Trading strategy could include accurate transactions costs. It would also be of great interest to study the other two most cited Pairs Trading methods¹ and the respectively performance and implications to the Portuguese Index market (PSI-20). Additional, instead of defining a trigger rule, thresholds d , it could be done by, for example, the Markov switching model or the Kalman filter. Such analysis could lead to a new path on the optimization of Pairs Trading implementation.

¹Meaning the two methods of selecting the pairs that are not presented in this research: the Stochastic and the Combined Forecast approaches

Bibliography

- Alexander, C. and A. Dimitriu (2002). The cointegration alpha: Enhanced index tracking and long-short equity market neutral strategies. *Discussion Paper 2002-08, ISMA Centre Discussion Papers in Finance Series.*
- Chan, Louis K.C., J. N. and J. Lakonishok (1996). Momentum strategies. *The Journal of Finance* 51(5).
- Cootner, P. H. (1964). The random character of stock market prices, the mit press, cambridge, massachusetts.
- Coval, Joshua D., H. D. A. and T. Shumway (2005). Can individual investors beat the market? *Working Paper No. 04-025.*
- Do, B. and R. Faff (2009). Does simple pairs trading still work? *Working Paper.*
- Do, B. and R. Faff (2011). Are pairs trading profits robust to trading costs? *Working Paper.*
- Elliott, Robert J., V. D. H. J. and W. P. Malcolm (2005). Pairs trading. *Quantitative Finance* 5(3), 271–276.
- Enders, W. (1988, August). Arima and cointegration tests of ppp under fixed and flexible exchange rate regimes. *The Review of Economics and Statistics* 70(3), 504–08.
- Engle, R. F. and C. W. J. Granger (1987, March). Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55(2), 251–76.
- Fama, E. (1991, December). Efficient capital markets: Ii. *Journal of Finance* 46(5), 1575–617.
- Fung, W. and D. Hsieh (1997). Empirical characteristics of dynamic trading strategies: the case of hedge funds. *Review of Financial Studies* 10(2), 275–302.
- Galenko, A., E. P. I. P. (2011). Trading in the presence of cointegration. *Journal of Alternative Investments.*
- Gatev, E., W. N. Goetzmann, and K. G. Rouwenhorst (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies* 19(3), 797–827.
- Jacobs, B. I. and K. N. Levy (1993). Long/short equity investing. *The Journal of Portfolio Management.*
- Jarrow, R. (1988). Finance theory. *Prentice-Hall.*
- Jegadeesh, N. and S. Titman (1993, March). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48(1), 65–91.

- Juselius, K. and D. F. Hendry (2000, December). Explaining cointegration analysis: Part ii. Discussion Papers 00-20, University of Copenhagen. Department of Economics.
- Kaan Evren Bolgiün, E. K. and S. Güven (2009). Dynamic pairs trading strategy for the companies listed in the istanbul stock exchange. *Working Paper*.
- Khandani, A. E. and A. W. Lo (2008, November). What happened to the quants in august 2007?: Evidence from factors and transactions data. Working Paper 14465, National Bureau of Economic Research.
- Lakonishok, J., A. S. and R. W. Vishny (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance* 49(5), 1541–1578.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47(1), 13–37.
- Lo, A. and A. MacKinlay (1988). Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies* 1(1), 41–66.
- Lucas, A. (1997). Strategic and tactical asset allocation and the effect of long-run equilibrium relations. Serie Research Memoranda 0042, VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics.
- Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica* 34(4), 768–783.
- Murray, M. P. (1994). A drunk and her dog: An illustration of cointegration and error correction. *The American Statistician* 48(1), 37–39.
- Perlin, M. (2007, December). Evaluation of pairs trading strategy at the brazilian financial market. *Journal of Derivatives & Hedge Funds* 15(8308), 122–136.
- Roberts, H. (1961). Market value, time, and risk. *Unpublished manuscript*..
- Rouwenhorst, K. G. (1998). International momentum strategies. *The Journal of Finance* 53(1), 267–284.
- Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* 6.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19(3), 425–442.
- Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. *The Journal of Finance* 52(1), 35–55.

Steve Hogan, Robert Jarrow, M. T. and M. Warachka (2004). Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics* 73(3), 525 – 565.

Treynor, J. L. (1961). Market value, time, and risk. *Unpublished manuscript.*

Treynor, J. L. (1962). Toward a theory of market value of risky assets. *Unpublished manuscript.*

Tsay, R. S. (2005). *Analysis of financial time series*. Wiley.

Vidyamurthy, G. (2004). Pairs trading. *John Wiley & Sons Inc.*

7. Appendix

Table 7.1: Stock's Pairs Over the Time

Stock	2003	Return	2004	Return	2005	Return
BPI	BCP	13.95	PT	0.00	PT	-8.70
BCP	EDP	8.554	Brisa	10.18	BES	-6.02
BES	Cofina	32.78	Zon	3.74	BCP	38.65
Banif	PT	5.42	EDP	4.01	M.Engil	-0.692
Brisa	Semapa	0.00	Semapa	12.35	Cofina	4.98
Cimpor	Semapa	7.88	JTM	1.33	BCP	15.34
EDP	BCP	12.64	Portucel	5.30	Portucel	15.82
Cofina	BES	4.07	Semapa	0.86	Brisa	1.36
JTM	PT	1.15	Cimpor	0.47	Brisa	-3.38
M.Engil	BPI	0.53	Semapa	5.02	cofina	-6.94
Portucel	EDP	0.00	BPI	1.11	EDP	9.89
PT	EDP	0.00	Cimpor	7.23	Zon	3.95
Semapa	BCP	0.00	Brisa	7.15	Cofina	-12.28
Zon	Cofina	27.92	Portucel	25.79	PT	5.98
Stock	2006	Return	2007	Return	2008	Return
BPI	M.Engil	5.774	Banif	0.00	Portucel	-28.71
BCP	M.Engil	13.593	Zon	9.9941	BES	10.02
BES	EDP	30.028	EDP	8.104	BCP	55.95
Banif	BPI	22.748	BPI	0.00	Portucel	-24.85
Brisa	EDP	6.24	Cimpor	8.784	EDP	2.39
Cimpor	M.Engil	26.437	Brisa	6.247	Portucel	12.57
EDP	Semapa	-4.78	BES	3.026	JTM	-6.35
Cofina	Zon	7.866	BPI	7.909	Banif	-35.24
JTM	Banif	6.536	EDP	-4.998	EDP	-7.87
M.Engil	BPI	2.68	Brisa	4.124	Banif	-54.86
Portucel	Semapa	6.024	BPI	0.00	Semapa	0.06
PT	Zon	5.377	Portucel	18.753	EDP	25.78
Semapa	Portucel	21.58	Portucel	-10.804	Portucel	0.33
Zon	Cofina	5.98	BCP	36.183	Portucel	46.80
Stock	2009	Return	2010	Return	2011	Return
BPI	BCP	0.00	M.Engil	0.00	M.Engil	0.00
BCP	BPI	0.00	BES	0.00	M.Engil	0.00
BES	BPI	0.00	BPI	0.00	BPI	0.00
Banif	Zon	0.00	Portucel	-12.75	Cimpor	2.49
Brisa	Zon	7.75	Semapa	5.92	Zon	1.70
Cimpor	M.Engil	15.03	Cofina	0.00	Brisa	11.13
EDP	Banif	0.00	JTM	0.00	Brisa	12.30
Cofina	M.Engil	0.00	JTM	4.88	M.Engil	3.73
JTM	Portucel	10.46	Cofina	-3.95	Portucel	30.11
M.Engil	Semapa	7.35	BPI	0.00	BPI	18.49
Portucel	M.Engil	0.00	Cofina	15.47	JTM	-8.77
PT	EDP	0.33	JTM	4.55	Portucel	15.07
Semapa	M.Engil	13.19	Brisa	10.30	Portucel	19.46
Zon	Banif	0.00	Semapa	-7.37	Brisa	1.67

Note: All values are expressed in percentage
The Table presents the stocks and its respective pairs along the temporal period under review.

Table 7.2: Pairs Trading Returns per year

	2003	2004	2005	2006	2007	2008	2009	2010	2011	Total
Jan	-3,05%	3,22%	1,64%	1,34%	-0,21%	0,28%	0,25%	0,00%	0,59%	4,08%
Feb	-3,93%	1,74%	2,87%	-0,59%	0,64%	1,07%	0,51%	0,02%	0,02%	2,37%
Mar	-4,75%	4,38%	0,25%	-0,44%	0,77%	2,72%	0,34%	-0,23%	0,66%	3,71%
Apr	1,19%	1,51%	1,47%	1,22%	3,34%	2,62%	0,00%	-1,46%	-0,66%	9,25%
May	5,75%	0,01%	0,90%	4,32%	3,16%	-4,10%	1,31%	1,65%	4,15%	17,13%
Jun	5,91%	0,59%	-0,82%	0,10%	-6,11%	-7,66%	-0,16%	0,55%	1,40%	-6,20%
Jul	-0,40%	2,35%	-1,46%	3,12%	7,24%	-2,84%	0,96%	2,06%	2,09%	13,13%
Aug	5,41%	1,46%	2,55%	2,37%	0,98%	-0,04%	1,37%	0,74%	4,64%	19,51%
Sept	5,87%	-0,90%	-0,59%	0,94%	-0,56%	-0,75%	0,78%	-1,07%	1,07%	4,80%
Oct	1,03%	1,40%	-0,01%	6,38%	-1,08%	6,53%	0,99%	6,56%	2,60%	24,41%
Nov	3,77%	0,06%	-2,16%	2,32%	2,82%	-2,35%	0,00%	-6,88%	-1,51%	-3,92%
Dec	-0,39%	-1,56%	2,71%	1,18%	1,47%	3,93%	1,36%	0,48%	0,69%	9,89%
Total	16,41%	14,28%	7,39%	22,29%	12,46%	-0,57%	7,73%	2,43%	5,76%	10,91%
σ	13,77%	5,84%	5,84%	7,00%	11,04%	13,35%	1,78%	10,52%	6,29%	8,40%
S.R.	1,02	2,09	0,89	2,76	0,75	-0,40	3,29	0,16	2,28	

Note: Total refers to the annual return, σ represents volatility and S.R. the Sharpe Ratio.

Table 7.3: Pairs Trading and Market Returns

Date	Pairs Trading Return	Market Return	% Time Market Beaten
2003	16,41%	12,65%	74,13%
2004	14,28%	11,90%	89,96%
2005	7,39%	12,58%	100,00%
2006	22,29%	26,18%	14,67%
2007	12,47%	15,07%	1,93%
2008	-0,57%	-71,94%	98,84%
2009	7,73%	28,87%	100,00%
2010	2,44%	-10,92%	100,00%
2011	15,76%	-32,29%	96,14%
Annual Return	10,91%	-0,88%	-
Annual Volatility	8,40%	15,84%	
Sharpe Ratio	1,43	0,09	-

Note: The % Time Market Beaten refers to the time that the accumulated return of Pairs Trading strategy was above the Market return.

Table 7.4: Alpha and Beta for Pairs Trading

Date	Alpha	t-Stat	P-Value	Beta	t-Stat	P-value
2003	13,143%	0,675	0,501	0,281	2,735	0,006
2004	14,329%	2,595	0,009	0,000	0,018	0,985
2005	7,329%	0,937	0,349	0,016	0,288	0,774
2006	20,271%	1,93	0,054	0,092	1,33	0,184
2007	12,055%	1,051	0,294	0,051	0,848	0,397
2008	-3,665%	-0,251	0,801	-0,0427	-1,59	0,11286
2009	2,607%	1,199	0,231	0,0076	1,062	0,289
2010	1,636%	0,233	0,816	0	-0,00,098	0,921
2011	16,642%	2,1643	0,031	0,002	1,085	0,278
		Coefficients	t Stat	P-value		
Annual Alpha (a)		10,63%	3,09	0,00***		
Beta (b)		0,01	0,53	0,60		

Note: 1: Alphas and Betas are obtained from a regression Pairs Trading returns over time against the returns of Portuguese Market Index (Psi-20); 2: Under the null hypothesis that (a) and (b) are equal to zero, respectively; 3: *** indicates significance of 1% and; 4: Annualised values.

Table 7.5: Pairs Trading Summary Statistics

Summary Statistics, using the observations 2003/01/02–2011/12/30 for the variable Pairs Trading (2309 valid observations)

Mean	Median	Minimum	Maximum
10.911*%	0.00%	-5.40%	5.781%
Std. Dev.	C.V.	Skewness	Kurtosis
8.402*%	15.517	0.678	10.754

*Annualised Values

Table 7.6: d Value Sensitivity Analysis

Value of d	2,5	2	1,5	0,5
Date	Sharpe Ratio			
2003	1,844	0,973	0,836	0,975
2004	1,314	1,930	1,672	2,118
2005	0,496	0,751	1,551	2,233
2006	2,547	2,754	1,952	1,673
2007	1,723	0,857	1,512	1,246
2008	-1,829	-0,267	0,615	1,445
2009	-1,619	2,427	1,798	2,269
2010	-0,216	-0,053	-0,909	0,490
2011	1,771	2,026	0,523	-1,681
Annual Return	7,6%	10,9%	11,7%	13,8%
% Negative Days	39,7	42,9	46,5	48,5
Max	5,7%	5,8%	6,1%	7,8%
Min	-3,1%	-5,4%	-5,1%	-6,5%
Sharpe Ratio	0,97	1,43	1,15	1,25

The Table presents the annual Sharpe Ratios for the different given values of d : 0,5;1,5;2;2,5.

Table 7.7: Summary Statistics for Pairs Trading under different d
 Summary Statistics, using the observations 2003/01/02–2011/12/30
 (missing values were skipped)

Variable	Mean	Median	Minimum	Maximum
2,5	0.000294727	0.000000	-0.0304183	0.0565699
2	0.000425475	0.000000	-0.0540467	0.0577791
1,5	0.000455883	4.26344e-005	-0.0515535	0.0615045
0,5	0.000538465	0.000103507	-0.0656937	0.0780942
Variable	Std. Dev.	C.V.	Skewness	Ex. kurtosis
2,5	0.00569059	19.3080	1.45134	14.7135
2	0.00660061	15.5135	0.677752	10.7476
1,5	0.00799186	17.5305	0.266640	7.27350
0,5	0.0112726	20.9347	0.440066	6.49486

Table 7.8: Random Trigger Test

Date	2003	2004	2005	2006	2007	2008	2009	2010	2011
Jan	-2,06%	1,81%	1,82%	4,50%	-3,27%	1,45%	-1,18%	1,72%	-2,34%
Feb	-5,70%	1,94%	3,21%	-0,28%	-1,19%	-0,56%	-1,25%	4,79%	-1,07%
Mar	-5,74%	-0,39%	1,35%	0,45%	1,94%	3,21%	-6,32%	-3,22%	-2,61%
Apr	1,32%	2,45%	-2,85%	2,71%	3,54%	4,42%	1,32%	0,04%	-5,46%
May	4,94%	5,52%	-0,52%	-0,90%	4,88%	-3,38%	-3,63%	0,68%	3,71%
Jun	6,20%	1,46%	2,66%	0,29%	-3,54%	2,55%	-1,23%	-1,26%	-1,07%
Jul	3,11%	-2,22%	-2,18%	1,68%	2,72%	1,67%	4,34%	-3,98%	-7,82%
Ago	4,90%	0,31%	3,20%	-1,90%	3,24%	-2,55%	6,27%	-3,42%	-9,00%
Set	1,66%	1,18%	-2,17%	2,08%	-3,44%	2,22%	-3,17%	-0,73%	-1,85%
Out	3,96%	-0,42%	-1,31%	0,97%	-2,94%	-16,25%	1,41%	10,22%	0,33%
Nov	3,81%	1,37%	-0,95%	0,78%	1,53%	-2,69%	0,15%	-10,05%	-4,71%
Dec	-2,15%	-0,66%	-2,10%	-0,39%	-1,47%	3,81%	2,40%	2,12%	3,13%
Total	14,26%	12,33%	0,14%	9,98%	1,98%	-6,10%	-0,90%	-3,09%	-28,75%
σ	14%	7%	8%	6%	11%	19%	12%	17%	13%
S.R.	0,83	1,52	-0,26	1,17	-0,21	-0,56	-0,18	-0,22	-2,23

Note: Total refers to the annual return, σ represents volatility and S.R. the Sharpe Ratio.

Table 7.9: Random Trigger OLS Analysis

Model 1: OLS, using observations 2003/01/02–2011/12/30 ($T = 2309$)

Dependent variable: Strategy_Random

	Coefficient	Std. Error	t-ratio	p-value
const	9.53638e-007	0.000165019	0.0058	0.9954
Psi_20	0.0471592	0.0140898	3.3471	0.0008
Mean dependent var	-6.60e-07	S.D. dependent var		0.007947
Sum squared resid	0.145056	S.E. of regression		0.007929
R^2	0.004833	Adjusted R^2		0.004401
$F(1, 2307)$	11.20275	P-value(F)		0.000830
Log-likelihood	7893.692	Akaike criterion		-15783.38
Schwarz criterion	-15771.90	Hannan–Quinn		-15779.20
$\hat{\rho}$	0.009638	Durbin–Watson		1.980724

Date	2003	2004	2005	2006	2007	2008	2009	2010	2011
Jan	-0,95%	-0,31%	2,64%	-0,31%	2,04%	-6,93%	0,38%	0,52%	2,19%
Fev	-2,05%	8,17%	-4,56%	6,46%	11,66%	5,82%	0,86%	-2,15%	-0,41%
Mar	1,98%	3,70%	-10,26%	8,40%	14,23%	-4,57%	0,39%	0,91%	-2,03%
Abr	0,39%	2,07%	0,24%	3,97%	1,63%	3,76%	3,21%	-4,67%	-0,98%
Mai	4,00%	1,01%	-1,43%	-3,05%	-16,94%	-2,65%	0,29%	-6,22%	0,21%
Jun	-1,78%	1,17%	-4,97%	-0,15%	4,31%	-1,02%	-0,35%	-1,09%	-2,29%
Jul	2,35%	-2,62%	-7,80%	1,69%	8,00%	7,56%	-0,01%	2,93%	-15,32%
Ago	0,74%	0,04%	-0,23%	4,83%	7,96%	-2,10%	1,65%	-1,88%	7,04%
Set	2,43%	-4,53%	-7,89%	-2,89%	3,80%	-10,14%	2,35%	-6,89%	-29,14%
Out	4,22%	1,08%	-7,49%	-1,58%	-6,65%	-37,06%	-3,30%	9,38%	-3,95%
Nov	0,67%	3,25%	-8,18%	-1,41%	0,40%	-19,68%	0,07%	-17,16%	4,46%
Dez	3,21%	0,99%	-3,41%	4,63%	8,01%	9,74%	0,27%	7,42%	2,10%
Total	15,22%	14,02%	-53,33%	20,58%	38,46%	-57,26%	5,81%	-18,91%	-38,12%
σ	7%	11%	14%	13%	29%	45%	6%	24%	34%
S.R.	1,75	1,09	-3,96	1,32	1,17	-1,37	0,81	-0,82	-1,16

Note: Total refers to the annual return, σ represents volatility and S.R. the Sharpe Ratio.

Table 7.11: Random Pair OLS Analysis

Model 1: OLS, using observations 2003/01/02–2011/12/30 ($T = 2309$)

Dependent variable: Random_Pair

	Coefficient	Std. Error	t -ratio	p-value
const	-0.000317580	0.000307205	-1.0338	0.3014
PSI_20	0.0254454	0.0262300	0.9701	0.3321
Mean dependent var	-0.000318	S.D. dependent var		0.014762
Sum squared resid	0.502719	S.E. of regression		0.014762
R^2	0.000408	Adjusted R^2		-0.000026
$F(1, 2307)$	0.941073	P-value(F)		0.332104
Log-likelihood	6458.755	Akaike criterion		-12913.51
Schwarz criterion	-12902.02	Hannan–Quinn		-12909.32
$\hat{\rho}$	0.016128	Durbin–Watson		1.967743

Table 7.12: Run Test

Mean	Test Value	Cases < Test Value	Cases \geq Test Value	Total	N ^o of Runs	Z	p-Value
P:BPI	-.000598	1072	1237	2309	1143	-0,276	0,391
P:BCP	-.001114	1069	1240	2309	1105	-1,849	0,032
P:BES	-.000660	964	1345	2309	1093	-1,33	0,092
P:BNF	-.000336	910	1399	2309	1027	-3,344	0,000
P:BRI	-.000320	1088	1221	2309	1236	3,522	0,000
P:CPR	.000213	1247	1062	2309	1170	0,918	0,179
P:ECP	.000186	1253	1056	2309	1206	2,47	0,007
P:CFNA	-.000103	943	1366	2309	1092	-1,066	0,143
P:JMT	.001023	1203	1106	2309	1143	-0,436	0,332
P:EGL	-.000140	1092	1217	2309	1132	-0,84	0,201
P:PTI	.000200	1406	903	2309	1069	-1,386	0,083
P:PTC	-.000098	1100	1209	2309	1152	-0,039	0,484
P:SEM	.000208	1241	1068	2309	1153	0,167	0,434
P:PTM	-.000343	1107	1202	2309	1103	-2,108	0,017

Table 7.13: Autocorrelation Test

Lag	BPI		BCP		BES		BRI		CPR		ACP		ECP	
	AC	Q-stat	AC	Q-stat	AC	Q-stat	AC	Q-stat	AC	Q-stat	AC	Q-stat	AC	Q-stat
1	0.020	1.760	0.107	26.689*	0.155	55.231*	0.020	0.925	0.017	0.698	-0.025	1.404		
2	-0.019	1.876	-0.061	26.832*	-0.030	57.264*	-0.018	1.681	0.000	0.698	-0.029	4.936***		
3	-0.007	1.876	-0.004	26.806*	-0.007	57.377*	-0.036	3.752	0.003	0.728	-0.023	6.117		
4	-0.020	2.804	-0.012	27.235*	0.034	38.405*	-0.036	6.707	0.012	1.057	-0.023	7.341		
5	0.051	8.834	0.110	27.464*	0.034	38.935*	-0.012	7.040	-0.080	15.967**	-0.018	7.492		
6	0.006	8.910	0.010	27.679*	-0.019	39.661*	-0.017	7.683	-0.012	16.317**	-0.010	35.368*		
7	0.036	11.865	0.026	29.285*	-0.003	61.531*	0.052	14.035***	-0.033	18.861*	-0.015	35.864*		
8	-0.011	12.161	-0.027	40.813*	-0.038	64.908*	0.002	14.043***	-0.022	20.007**	-0.011	36.167*		
9	-0.007	12.270	-0.048	34.707*	-0.030	67.010*	0.001	14.044	0.000	20.008**	-0.077	49.991*		
10	-0.003	15.810	-0.059	42.681*	-0.015	47.119*	0.019	14.046	0.009	20.176**	-0.000	49.991*		
11	-0.029	17.669	-0.066	42.776*	-0.026	48.731*	0.009	15.076	-0.030	22.221**	-0.002	50.000*		
12	0.018	17.669	0.018	43.522*	-0.024	68.862*	0.009	15.596	0.017	22.923**	-0.042	54.068*		
13	0.017	18.354	0.020	44.446*	-0.025	52.343*	0.023	16.795	0.005	22.923**	-0.042	54.068*		
14	0.006	18.429	0.022	45.589*	0.014	70.111*	0.033	16.795	0.004	23.011***	-0.059	62.169*		
15	-0.029	20.325	-0.011	46.865*	-0.003	58.798*	0.006	19.450	0.004	23.011***	-0.059	70.188*		
16	0.041	24.298***	0.014	46.864*	0.053	77.611*	0.031	21.644	0.021	24.011***	0.039	73.769*		
17	0.053	30.882**	0.036	49.397*	0.010	71.076*	0.001	21.644	0.002	26.776***	0.036	76.826*		
18	-0.023	32.082**	-0.042	53.547*	0.018	63.138*	0.001	21.647	0.002	26.782***	-0.024	78.119*		
19	-0.016	32.683**	-0.017	54.208*	0.046	82.957*	-0.025	23.058	0.005	26.842***	0.004	78.150*		
20	-0.031	34.947**	-0.035	57.073*	-0.009	67.228*	0.016	23.682	-0.034	27.307***	-0.019	79.006*		
21	-0.015	35.445**	-0.022	58.153*	0.027	67.877*	0.030	25.756	-0.000	30.075***	-0.031	81.225*		
22	-0.011	35.725**	-0.001	58.157*	0.018	92.886*	-0.043	30.150***	-0.000	30.075***	-0.001	81.505*		
23	0.002	35.734**	0.004	58.199*	-0.001	93.666*	0.003	30.321	-0.022	31.167***	-0.005	81.569*		
24	0.004	35.775**	0.047	63.436*	0.001	93.666*	0.005	30.340	-0.012	31.574	-0.055	88.711*		
25	-0.004	37.855**	-0.032	65.854*	-0.007	93.793*	-0.008	30.408	-0.006	31.574	-0.044	93.325*		
26	0.004	37.901**	0.004	65.900*	-0.004	93.825*	0.016	30.559	0.012	32.305	-0.028	95.191*		
27	-0.042	41.991**	-0.037	69.146*	-0.023	95.091*	0.035	33.995	-0.035	32.636	0.026	96.7636*		
28	-0.006	42.085**	-0.015	73.444*	-0.003	95.115*	0.024	35.309	-0.043	39.821***	-0.015	97.284*		
29	-0.007	42.196**	-0.024	75.363*	-0.012	95.439*	-0.024	36.604	0.025	41.304***	0.006	99.259*		
30	-0.056	49.599**	-0.029	78.234*	-0.003	95.460*	-0.003	36.623	0.017	42.001***	0.039	102.895*		
1	0.028	1.760	0.029	17.192*	-0.021	0.988	-0.004	0.035	-0.014	0.427	0.111	28.315*		
2	0.007	5.012***	-0.036	4.959***	-0.036	17.294*	-0.034	2.733	-0.009	0.613	0.013	28.855*		
3	0.009	5.212	-0.014	17.336*	0.001	4.044	-0.045	0.949	-0.012	0.949	0.008	28.830*		
4	0.053	11.804**	-0.033	9.020***	-0.004	17.739*	-0.021	8.326***	-0.004	0.949	-0.012	29.142*		
5	0.011	12.081**	-0.023	10.228***	-0.012	18.384*	-0.021	9.317***	-0.020	1.909	-0.014	29.587*		
6	0.026	13.663**	-0.015	10.753***	0.001	4.428	-0.005	9.379	0.007	2.011	-0.025	31.011*		
7	0.035	16.479**	-0.030	12.862***	0.001	18.663*	-0.049	14.913***	0.038	5.379	-0.012	31.321*		
8	0.017	17.145**	-0.050	18.747**	0.053	12.556	0.005	14.977***	0.033	7.849	-0.019	32.128*		
9	-0.022	18.317**	-0.002	18.759**	-0.063	21.643*	-0.033	17.428**	-0.050	13.564	0.008	32.266*		
10	0.081	33.583*	0.008	18.906**	-0.000	22.943**	-0.006	17.508***	-0.005	13.614	-0.024	33.635*		
11	0.050	39.378*	0.004	18.942***	0.024	22.943**	0.041	21.385***	-0.011	13.890	-0.001	33.637*		
12	0.035	42.272*	0.036	21.929**	0.021	23.953**	0.012	21.800**	0.009	14.075	-0.002	33.650*		
13	0.031	44.487*	0.000	21.930***	0.021	22.985**	-0.047	26.749**	0.015	14.602	-0.023	34.878*		
14	0.080	59.240*	0.009	22.134***	0.029	25.409**	-0.012	27.110**	0.009	14.642	-0.026	36.393*		
15	0.023	60.506*	0.032	24.519***	-0.012	26.205**	0.009	27.286**	-0.004	14.642	-0.021	37.370*		
16	0.021	61.561*	0.004	24.563***	-0.032	28.546**	-0.009	28.272**	0.030	15.391	-0.021	39.581*		
17	0.013	61.975*	-0.025	26.008***	0.054	35.232*	0.021	28.272**	0.018	17.439	0.031	39.581*		
18	0.024	63.336*	-0.080	40.840*	-0.021	36.287*	0.014	28.723**	0.007	17.549	0.010	39.835*		
19	-0.026	64.883*	0.111	41.105*	-0.037	39.541*	-0.028	30.583**	-0.025	19.006	-0.005	39.897*		
20	-0.014	65.107*	-0.026	42.687*	-0.006	40.120*	0.003	30.693**	-0.017	19.687	-0.010	40.108*		
21	0.014	65.597*	-0.005	42.745*	-0.006	40.211*	0.019	32.271***	-0.011	19.980	-0.038	43.488*		
22	-0.005	65.658*	0.045	47.448*	-0.015	40.740*	-0.030	34.437***	0.023	21.197	-0.065	53.240*		
23	-0.046	70.577*	-0.006	47.522*	0.009	36.409**	-0.008	34.606**	0.036	24.671	-0.037	56.404*		
24	0.025	70.592*	0.012	47.876*	-0.005	40.876**	-0.018	35.121***	-0.029	26.608	-0.006	56.578*		
25	0.006	72.151*	-0.005	49.322*	-0.011	36.737***	0.018	42.425**	-0.002	26.619	-0.000	56.578*		
26	0.006	73.230*	-0.036	39.776**	-0.028	44.299**	-0.006	35.851***	0.025	27.078	-0.003	56.598*		
27	-0.021	73.230*	-0.035	42.681**	-0.046	49.281*	-0.038	39.212***	-0.014	28.529	-0.025	58.063*		
28	0.032	75.668*	0.049	65.246*	-0.014	43.125**	0.006	39.309**	-0.053	35.153	-0.005	58.127*		
29	-0.011	75.941*	0.014	65.699*	-0.016	50.322*	0.033	41.793***	-0.047	40.224***	0.039	61.696*		
30	0.030	78.093*	-0.039	69.305*	-0.006	50.416**	0.010	42.008***	0.000	40.225	0.028	63.580*		

Table 7.14: Cointegration method Analysis

	Cointegration method	Minimum Distance method	PSI-20
Volatility	10,35%	7,45%	11,02%
Sharpe Ratio	1,316	1,43	0,73
Max daily Return	18,65%	4,14%	10,20%
Min Daily Return	-6,06%	-5,40%	-10,38%
% of Positive Returns	51,65%	47,71%	52,19%
% of Null Returns	3,213%	10,078%	0,000%

Table 7.15: Pairs Trading Strategy Vs. Psi-20

Month	Cointegration Method	PSI-20	Minimum Distance Method
Jan	0,80%	-4,58%	0,64%
Feb	-5,98%	4,98%	1,53%
Mar	-9,64%	-1,59%	1,83%
Apr	-7,30%	-6,74%	0,08%
May	7,52%	-15,19%	0,00%
Jun	-8,53%	7,62%	-5,12%
Jul	10,95%	-3,32%	6,09%
Aug	1,22%	8,59%	-1,95%
Sep	5,18%	7,06%	-2,29%
Total Return	-5,77%	-3,18%	0,81%
Volatility	22%	24%	9%
S.R.	-0,39	-0,26	-0,24