FUNDAÇÃO GETULIO VARGAS

ESCOLA DE MATEMÁTICA APLICADA

ARIEL AMADEU EDWARDS TEIXEIRA

Pair Trading in Bovespa With a Quantitative Approach: Cointegration, Ornstein-Uhlenbeck Equation and Kelly Criterion.

Rio de Janeiro

2014

Ariel Amadeu Edwards Teixeira

Pair trading in Bovespa with a quantitative approach: cointegration, Ornstein-Uhlenbeck equation and Kelly criterion.

Dissertação submetida à Escola de Matemática Aplicada como requisito parcial para obtenção do grau de Mestre em Modelagem Matemática da Informação.

Área de concentração: Modelagem e Simulação de Sistemas Complexos

Orientador: Moacyr Alvim Horta Barbosa da Silva

Co-Orientador: Edson Daniel Lopes Gonçalves

Rio de Janeiro

2014

Ficha catalográfica elaborada pela Biblioteca Mario Henrique Simonsen/FGV

Teixeira, Ariel Amadeu Edwards
Pair trading in Bovespa with a quantitative approach : cointegration, Ornstein-Uhlenbeck equation and Kelly criterion / Ariel Amadeu Edwards Teixeira. – 2014.
28 f.

Dissertação (mestrado) – Fundação Getulio Vargas, Escola de Matemática Aplicada.
Orientador: Moacyr Alvim Horta Barbosa da Silva.
Coorientador: Edson Daniel Lopes Gonçalves.
Inclui bibliografia.

1. Cointegração. 2. Operações com pares (Finanças). 3. Ações (Finanças). I. Silva, Moacyr Alvim Horta Barbosa da. II. Gonçalves, Edson Daniel Lopes. III. Fundação Getulio Vargas. Escola de Matemática Aplicada. IV. Título.

CDD – 332.645

ARIEL AMADEU EDWARDS TEIXEIRA

PAIR TRADING IN BOVESPA WITH A QUANTITATIVE APPROACH: COINTEGRATION, ORNSTEIN UHLENBECK EQUATION AND KELLY CRITERION .

Dissertação apresentada ao Curso de Mestrado em Modelagem Matemática da Informação da Escola de Matemática Aplicada da Fundação Getulio Vargas para obtenção do grau de Mestre em Modelagem Matemática da Informação.

Data da defesa: 17/02/2014.

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA

Moacyr Alvim Horta Barbosa da Silva
Orientador (a)

Edson Daniel Lopes Gonçalves

Sabrina Camargo

Rafael Moura Azevedo

Agradecimentos

Agradeço a Deus por tudo, à minha famíla por todo o suporte e carinho, aos meus colegas de curso por terem me ajudado a chegar aqui, e aos meus orientadores por me guiarem no meu trabalho.

"I can calculate the motions of heavenly bodies, but not the madness of the people." - Isaac Newton on the South Sea Bubble

Abstract

Pair trading is an old and well-known technique among traders. In this paper, we discuss an important element not commonly debated in Brazil: the cointegration between pairs, which would guarantee the spread stability. We run the Dickey-Fuller test to check cointegration, and then compare the results with non-cointegrated pairs. We found that the Sharpe ratio of cointegrated pairs is greater than the non-cointegrated. We also use the Ornstein-Uhlenbeck equation in order to calculate the half-life of the pairs. Again, this improves their performance. Last, we use the leverage suggested by Kelly Formula, once again improving the results.

Summary

1. Introduction

"In 44 years of Wall Street experience and study, I have never seen dependable calculations made about common-stock values (…) that went beyond simple arithmetic or the most elementary algebra. Whenever calculus is brought in, (…) you can take it as a warning signal." - *Benjamin Graham*

Traders have known pair trading for decades. Its origin dates back to the 1920s, when the first hedge funds appeared. Although the term hedge fund was coined in the 1950s (the former Forbes magazine reporter Alfred W. Jones is considered the creator of hedge funds [1]), since the 1920s there were investment partnerships as Graham-Newman Corp. headed by Benjamin Graham, that used a variety of different trading strategies, as a hedge fund today [2]. The main difference between a hedge fund and a mutual one is that mutual funds operate in only one market direction, the buy side. Hedge funds, instead, buy and short stocks, therefore hedging its positions. Pair trading (i.e. buying one security against selling other) is a variation in this strategy.

Although calculus and higher algebra have many applications in stocks, it was not until the publishing of Black-Scholes formula in 1973 that the use of mathematics beyond simple arithmetic, as demeaned by Graham, became widespread. Although it was known and used before by Edward Thorpe (who will be discussed later, when we talk about the Kelly Formula), from this event on, mathematical arbitrage trading started to evolve.

As already said, traders were aware of the possibility of hedging by selling shares. For example, if one was bullish on a company, he could buy its stocks and hedge the position by selling an equivalent amount of money in stocks of other company of the same sector. By doing this, he would be protected against market direction. The market could drop, but the stock he bought should only outperform the one he sold, and there would be a guaranteed profit.

In the 1980s, Gerry Bamberger and Nunzio Tartaglia started a variation of this strategy in Morgan Stanley [3]. Now, the investors did not have an idea on whether a stock would outperform the market. His interest was in trading the spread between stocks itself. He would buy the spread if it became too small, or sell it if became too big. The focus here would be stocks issued by the same company, within the same sector, and a plenty of other options as different calendar contracts of commodities, foreign exchanges, etc.

In 1990s, a discussion about pair trading arose. The strategy had become popular and disseminated among traders, and the profit in some pairs vanished. Traders started losing big positions in a type of trade that was supposed to be safe. The problem was buying or selling a spread that simply did not stop trending, becoming larger and larger, or always smaller.

In this context, Carol Alexander publishes “Cointegration and asset allocation: A new active hedge fund strategy” [4]. The main message of her paper was whether to buy or sell a spread; traders had to check if the spread itself was stationary. Otherwise, nothing would guarantee that the spread would return to its mean. Economists had known the tool to make this test since the 1980s, when Engel and Granger began studying spurious regressions. One had to make cointegration test to evaluate whether the spread between two non-stationary processes was stationary.

2. AR Model, stationary and non-stationary variables

One of the most frequent approaches used in pair trading is using correlation, instead of checking wether the spread between the securities is stable and won’t collapse. As an example, Petrobras has two different stocks issued in Bovespa, PETR3 (common) and PETR4 (preferred). They have an almost perfect correlation (98%) and by the standard approach they are regarded as a perfect pair to trade. However they do not cointegrate. That means the spread between them is not stable and not a reliable one to trade agains deviations from the mean, although both of them represent the same company and the same fundamentals.

Now let us take a different pair, VIVT4 and BRTO4. They represent stocks that belong to the same sector, telecommunications. Though, they are different companies with different fundamentals, and they have little correlation between them (-0.08%). They are not a very promising pair by the usual approach using correlation. However they cointegrate and hence they form a pair more reliable than PETR3-PETR4.So in order to check if a pair of stocks is good candidate for trading one should check if they cointegrate. To get there, we will talk about autoregressive models, stationary and non-stationary variables.

According to Wooldridge [5] and Gouriroux [6], an autoregressive model is a random process in which future values are estimated through weighted past values.So, the time series y_t follows an autoregressive process of order 1 (AR1), if and only if it can be written as

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t$$

which around $\alpha/(1 - \rho)$, is stationary. By stationary, we mean the series has a constant mean, and its covariance (and therefore, its variance) and autocorrelation do not depend on time.

Set $E(y_t) = E(y_{t-1})$. Hence,

$E(y_t) = \alpha + \rho E(y_t) + 0$ and $E(y_t) = \alpha / (1 - \rho))$. Therefore, if $\rho = 1$, the series mean will not exist. Also note that if, $|\rho| > 1$, the series will diverge, and will not have a constant mean. Hence, we need $|\rho| < 1$ to yield a stationary series.

Besides,

$$Cov(y\ ,y_{t+h}) = Cov(\alpha + \sum_{l=0}^{\infty} \rho^l \varepsilon_{t-h}, \sum_{k=0}^{\infty} \rho^k\ \varepsilon_{t-h-k}) =$$

$$\sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \rho^{l+k} Cov(\varepsilon_{t-l,}\, \varepsilon_{t-h-k}) = \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \rho^{l+k} Var(\,\varepsilon_{t-h-k}) = \sigma^2 \sum_{k=0}^{\infty} \rho^{h+2k} = \frac{\rho^h \sigma^2}{1-\rho^2}$$

Therefore, the autocorrelations are power functions of the autoregressive coefficient:

$$\gamma(t,h) = cov(y_t, y_{t-h}) / var\ (y_t) = \frac{Cov(y_t, y_{t-h})}{Var(y_t)} = \frac{\sigma^2}{1-\sigma^2} \rho^{|h|} \left[\frac{\sigma^2}{1-\sigma^2}\right] = \rho^{|h|}.$$

Hence, $E(y_t)$ and $Var(y_t)$ are constant, and $Cov(y_t, y_{t-h})$ do not depend on time.

A non-stationary series is said to be integrated of order *n*, where *n* is the number of times the series needs to be differenced to yield a stationary process. A stationary series is integrated of order 0, I(0), and a I(1) series (integrated of order 1) is a non-stationary series that needs to be differenced one time to become stationary.

3. Cointegration

"If you find her [the drunk], the dog is unlikely to be far away." -*Michael Murray*

Cointegration can be easily explained by Murray's example [7] about a drunk and her dog. Both the dog and the drunk follow random walks, and the dog is not on a leash, which enforce a fixed distance between them. However, the dog does not want to get too far from the drunk, or else she will lock him out of home. In the same way, the drunk does not want miss the dog so he will not wake her up in the middle of the night with his barking to get in. Hence, despite the non-stationarity of the paths, the distance between the two paths is stationary, and the walks of the woman and her dog are said to be cointegrated.

A good intro is given by Carol Alexander in "Cointegration and asset allocation: A new active hedge fund strategy" [4]. Traditionally, correlation has been used in order to seek candidates to pair trading. As Alexander showed, this is a mistake, because it does not say anything about the stability of the spread (which is the most important element in this kind of strategy). In Brazil, whether in the financial market or even among academics, correlation is used in determining pairs, instead of cointegration. This is what guarantees the spread stability and that the pair will be profitable. For instance, Alves [8] finds an arbitrage between GOUA4 and GGBR4 as in the classic case of 'siamese twin' companies. However, they do not cointegrate, in the time span we study, and do not show a reasonable performance.

Consider two nonstationary processes A and B. If there is a linear combination of them which is a stationary process, the two pairs are said cointegrated. If we take stock A and a stock B, processes integrated of order 1, i.e. nonstationy processes, hence:

$$A_t = \beta\ B_t + \varepsilon_t \rightarrow \varepsilon_t = A_t - \beta\ B_t.$$

A_t and B_t are cointegrated if and only if ε_t is stationary.

4. Unit Root Hypothesis Testing - Dickey Fuller Tests

Suppose the model

$y_t = \rho y_{t-1} + \varepsilon_t$,

is fit to an I(1) series. According to Greene [9], the conventional measures will tend to hide the true value of ρ; the sample estimate is biased downward, and by dint of the very small true sampling variance, the conventional *t* test will tend, incorrectly, to reject the hypothesis of $\rho = 1$. The practical solution to this problem devised by Dickey and Fuller was to derive, through Monte Carlo methods, an appropriate set of critical values for testing the hypothesis that ρ equals one in an AR(1) regression when there truly is a unit root. One of their general results is that the test may be carried out using a conventional *t* statistic, but the critical values for the test must be revised; the standard *t* table is inappropriate. The random walk with drift to be analyzed in the DF test is

$$y_t = \mu(1-\rho) + \rho y_{t-1} + \varepsilon_t .$$

4.1 The Augmented Dickey-Fuller Test

In this work, we use Augmented Dickey-Fuller Test in order to check cointegration. According to Pfaff [10], we impose a known cointegrated vector, so all the uncertainty associated with estimating the cointegrated vector is eliminated, when we use the ADF test. In the other hand, we are estimating the cointegrated vector and so the uncertainty associated with this estimation is incorporated in the test when we use the Johansen test framework, which is another cointegration test.

We know the cointegrated vectors from theory, so it makes sense to impose it in our pair trading examples. Hence, the ADF resulting tests will have more power, i.e. ability to reject the null when the alternative is true, than the Johansen tests. Both tests have no-cointegration as the null hypothesis. The ADF test has a higher power than the Johansen test in this context, which explains our preference to find cointegration with the ADF test.

The Dickey-Fuller test described above assume that the disturbance in the model are white noise, i.e., $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) = \sigma^2$. An extension which will accommodate some form of serial correlation is the augmented Dickey-Fuller test. The ADF test is the same as above, carried out in the context of the model

$$y_t = \mu + \beta t + \rho y_{t-1} + \rho_1 \Delta y_{t-1} + \ldots + \rho_p \Delta y_{t-p} + \varepsilon_t.$$

The random walk with drift, which is of our interest, has $\beta = 0$. Another advantage of this formulation is that it accommodates higher-order autoregressive processes in ε_t.

4.2 Pairs Trading Model

A market neutral long/short strategy is the investment strategy we aim at implementing. This implies that we will try to find shares within the same sector, or which are components of an index. By simultaneously taking both a long and short position the beta of the pair tends to zero, where beta is the correlation between the stocks and the market. The hard part of this strategy is to find market neutral positions that will deviate in returns. The starting point of this strategy is that stocks that have historically had the same trading patterns (i.e. constant price ratio) will have so in the future as well. A trading opportunity is created if there is a deviation from the historical mean, which can be exploited. When the price relationship is restored, gains are earned.

If a price series of a security is stationary, it would be a great candidate for a mean-reversion strategy. However, most stock prices are not stationary. However, one can find a pair of stocks such that if you long one and short the other, the market value of the pair is stationary. If this is the case, the two time series cointegrate.

If a price series is stationary, then a mean-reverting strategy is guaranteed to be profitable, as long as the stationarity persists into the future. However, the converse is not true.

So, suppose we are trading a pair spread, long A and short B. How many units of B should we sell short for each unit of A that we buy? The hedge ratio is that number. Assume the spread is $S_t = A_t - \beta B_t$. Hence, β is the hedge ratio. Restating A and B relationship

$A_t = \alpha + \beta B_t + \varepsilon_t$ (19)

Spread is $S_t = A_t - \beta B_t = \alpha + \varepsilon_t$

Therefore, hedge ratio is the β coefficient from linear regression.In summary, we need to find two stocks prices which cointegrate, find the hedge ration between the two stocks, use mean reversion in the ratio of the prices, (correlation is not key), earn gains when the historical price relationship is restored, and let resources invested in risk-free interest rate.

Let us take the pair AMBV4-CRUZ3 as an example. It shows a hedge ratio of 2.724. That means that for every 1,000 shares of AMBV4 we are long, we have to be short 2,724 shares of CRUZ3.

4.3 Testing for the mean reversion

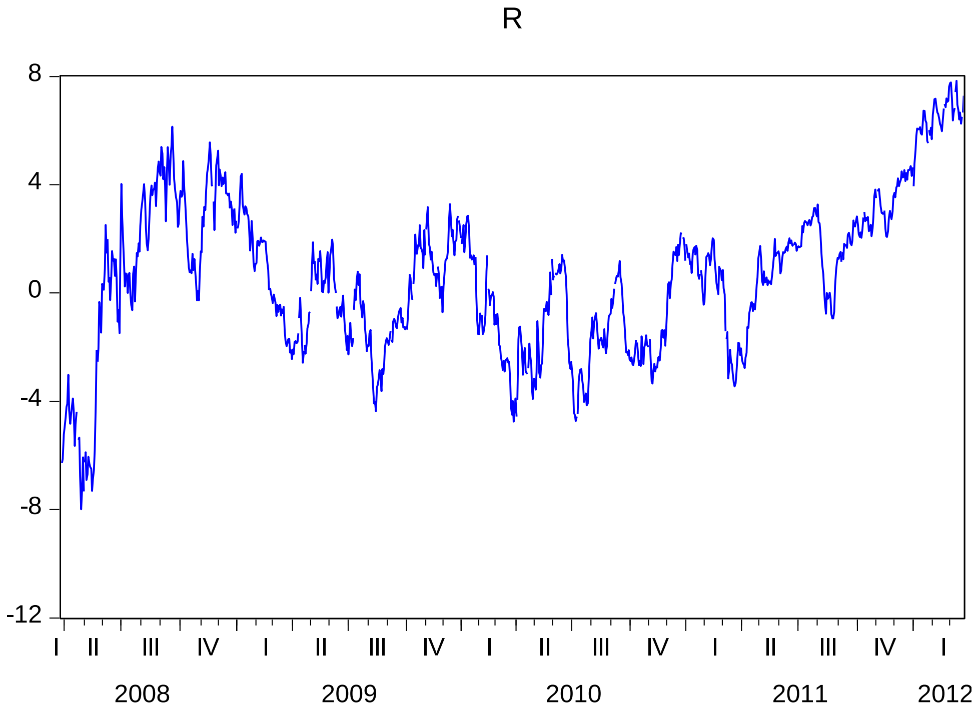
The challenge in this strategy is identifying stocks that tend to move together and therefore make actual cointegrated pairs. Our aim is to identify pairs of stocks with mean-reverting relative prices. To find out if two stocks cointegrate and are mean-reverting, the test conducted is the Dickey-Fuller test of the ratio of the pair.

A Dickey-Fuller test for determining stationarity in the spread $y_t = A_t - \beta B_t$ of share prices A and B

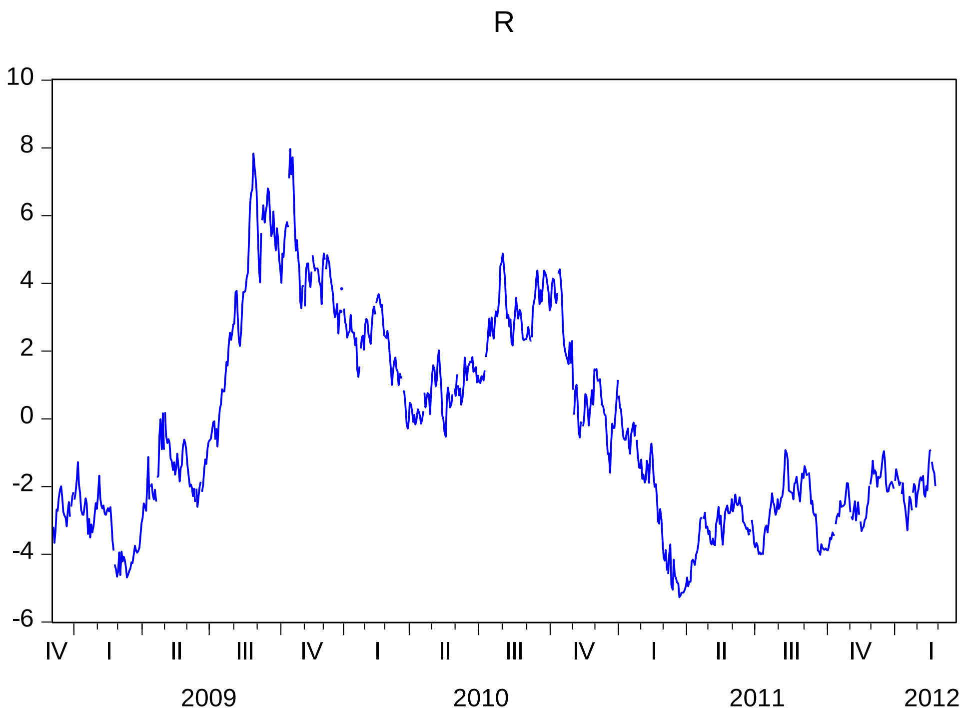
$$\Delta y_t = \mu + \gamma y_{t-1} + \varepsilon_t$$

We are regressing Δy_t on lagged values of y_t. The null hypothesis is that the process is not mean reverting, which means that $\gamma = 0$. The price ratio is following a weak stationary process and is thereby mean-reverting, if the null hypothesis can be rejected on the 99% confidence level.

As an example, we plot the graph of the spread of the cointegrated pair GOAU4-USIM5 generated by EViews.



Now, we plot the graph of the non cointegrated pair CYRE3-BOVA11.



The graphs do not bring us much insight about cointegration. Hence, one must run the ADF Tests in order to check cointegration.

4.4 Screening Pairs

To eliminate the risk we also want to stay sector neutral. This implies that we only want to open a position in a pair that is within the same sector. Due to the different volatility within different sectors, we expect sectors showing high volatility to produce very few pairs, while sectors with low volatility to generate more pairs. Another factor, which influences the number of pairs generated, is the homogeneity of the sector. A sector such as Commercial Services is expected to generate very few pairs, but Financials, on the other hand, should give many trading opportunities. The reason for this is that companies within the Financial sector have more homogenous operations and earnings.

We will also test pairs of shares and the index they compose, BOVA11. We will test the pairs suggested in XP Investimentos, Brazil's greatest retail brokerage firm, which has more than a hundred candidates for pair trading.

5. Optimal Convergence Trading

"The market can stay irrational longer than you can stay solvent." – John Maynard Keynes

One important issue in trading is when a trade should be closed. As in comedy, timing is essential if one wants to achieve profits instead of bankruptcy. If we know a process is mean reverting, we can estimate the period of time it should take to revert to its mean. We can find it by calculating the process half-life [11].

Let define half-life (H) of a variable z as the time for the expected value of z(t) to reach the intermediate (middle) price between the long run mean μ (which the mean reverting process is converging in the long run) and the current value z(0).

The *slowness* of a mean-reversion process is given by half-life. So, there is a direct relation between H and ϴ, and it is an alternative parameter to the reversion speed. This relation is showed below.

Suppose that the variable z, the spread (long market value minus short market value) of a pair of stocks follows a mean-reverting process also known as Ornstein-Uhlenbeck:

$$dz(t) = -\Theta\,(z(t) - \mu)\,dt + dW$$

Where Θ is the speed of reversion, μ is the long-run mean price (or the long-run equilibrium level, which the spread tends to revert), and dW is a white noise [12].

$$E[dz(t)] = -\Theta\,(z(t) - \mu)\,dt$$

The spread half-life is deducted below. Hence, we have the deterministic equation:

$$dz(t) / (z(t) - \mu) = -\Theta\,dt$$

Integrating from x_0 (t_0) to the expected price at the instant t_1, denoted by $z_1(t_1)$, and letting $\Delta t = t_1 - t_0$, we get:

$$\int_{z_0}^{z_1} \frac{d(z(t))}{z(t)-\mu} = \int_{z_0}^{z_1} -\theta dt$$

Hence:

$$\ln\left(\frac{z_1-\mu}{z_0-\mu}\right) = -\Theta\,\Delta t$$

For Δt = half-life H, by definition we have that $(z_1 - \mu) = 0.5\,(z_0 - \mu)$. Substituting:

$$\ln(0.5) = -\Theta H$$

$$-\ln(2) = -\Theta H$$

The half-life is finally given by the equation:

$$H = \frac{\ln(2)}{\theta}$$

For a practical example, take the pair AMBV4-CRUZ3. It has a half-life of 17 days. That means the pair should take 17 days to reverse to its mean, once it deviates away from it. So a trade in this pair should not take more than 17 days. Once that time frame is hit, one should close the trade.

6. Trading rules

The long short report described gives us a large set of pairs that are both market and sector neutral, which can be used to take positions. Since timing is an important issue, this should not be done randomly. We will therefore introduce several trading execution rules.

We need a couple of trading rules to follow, i.e., to clarify when to open and when to close a trade, in order to execute the strategy. When the ratio of two share prices hits the 2 standard deviation, our basic rule will be to open a position and close it when the ratio returns to the mean. In a pair with a spread that is wide and getting wider, however, we do not want to open a position.

By the following procedure, this can be partly avoided: when the price ratio deviates with more than two standard deviations mean calculated in training set, we actually want to open a position. When the ratio breaks the two-standard-deviations limit for the first time, the position is not opened, but rather when it crosses it to revert to the mean again. When the pair is on its way back again, you can say that we have an open position.

A summary would be: open position when the ratio hits the 2 standard deviations band for two consecutive times, close position when the ratio hits the mean, and close position when the trade duration hits the half-life time.

7. Kelly Criterion

"LTCM ran leveraged positions at too-high risk levels. It was not a sustainable business in the longer run (…)" - *Myron Scholes*

In 1998, LTCM collapsed. It was an arbitrage trading hedge fund founded by Nobel laureates Robert Merton, Myron Scholes, and former vice-chairman and head of bond trading at Salomon Brothers, John Merriwether. This hedge fund had $5 billion and $125 billion in debt, a 1:25 leverage. One of the most important issues in investing is how much to allocate in each trade, whether you are an Economics Noble prizewinner or a seasoned trader. If an investor takes much leverage in a trade, he can end bankrupt. In the other hand, he may allocate less money than he could and obtain a poor return. In order to know the optimal leverage, an investor should use Kelly criterion in a trade.

Kelly criterion was described by J. L. Kelly Jr in 1956. Its practical use was to gamble in baseball games. However, it got famous when used by another mathematician, Edward Thorp, who used it to make bets in blackjack and poker. Indeed, the formula worked so well that Thorpe not only made a lot of money (and got himself kicked out of casinos), but also got famous by releasing a book “Beat the dealer”, which became a best-seller. The strategy known as *counting cards* became prohibited in casinos.

After that, Thorp turned his attention to the stock market. He managed a hedge fund and yielded a track record of annualized 20 percent rate of return averaged over 28.5 years. In money terms, it means that $10,000 invested with Thorp would achieve the astonishing amount of $1,805,783 in this period. Thorp also released a book, “Beat the market”, that described his strategies and became a best seller. Among the strategies was the exploration of arbitrages between stocks and its options, an arbitrage he was using before the publishing of Black-Scholes formula, and the use of Kelly formula, the optimal amount of money to be allocated in each bet in the market [3].

Kelly proposed a different criterion for gamblers in his original paper. To maximize the expected value of wealth, the classic gambler would need to invest 100% of his capital in each bet. Kelly maximized the expected value of the utility function, rather than maximizing expected value of capital. Under the assumption that more money can never be worse than less, utility functions are used by economists to value money and are increasing as a function of wealth. We will use the base e logarithm (the natural log), Kelly took the base 2 logarithm of capital as his utility function instead [13].

Let X be a random variable with $P(X = \mu+\sigma) = P(X = \mu-\sigma) = 0.5$. Then $E(X) = \mu$, $Var(X) = \sigma^2$. With initial capital V_0, betting fraction f, the final capital is

$$V(f) = V_0 (1+(1-f)r + fX) = V_0(1+r+f(X-r)),$$

Where, r is the rate of risk free return. Then,

$$g(f) = E(G(f)) = E(\ln V(f)/V_0) = E(\ln(1+r+f(X-r)) = 0.5 \ln(1+r+f(\mu-r+\sigma)) + 0.5 \ln(1+r+f(\mu-r+\sigma)).$$

Now, subdivide the time interval into n equal independent steps. We have n independent X_i, i = 1, …, n, with

$$P(X_i = \mu/n + \sigma n^{-1/2}) = P(X_i = m/n - \sigma n^{-1/2}) = 0.5.$$

Then

$V_n(f)/V_0 = \Pi_{i=1}^{n}$ (1+(1-f)r + fX_i).

Taking E(log(.)) of both sides and expanding the result in a power series leads to

$G(f) = r + f(\mu\text{-}r) \frac{f^2\sigma^2}{2} + O(n^{-1/2})$.

Letting n → ∞, we have

$G_\infty(f) = r + f(\mu\text{-}r) - \frac{f^2\sigma^2}{2}$

Then,

$f^* = (\mu\text{-}r)/\sigma^2$

In practice, as the returns are not Gaussian, it is used half Kelly, i.e., the leverage suggested by Kelly formula divided by 2. For an example, let us take the pair BISA3-RSID3. The leverage suggested by half Kelly would be 2.44. So, one should bet 244% of his equity, in order to maximize the compounded rate of growth

8. Sharpe Ratio

Information ratio is the measure to use when one wants to assess a long only strategy. It is defined as

Information Ratio = $\frac{Average\ of\ Excess\ Returns}{Standard\ Deviation\ of\ Excess\ Returns}$

Where

Excess Returns = Portfolio Returns – Benchmark Returns

The Sharpe ratio is a special case of the information ratio, suitable when one has a money-neutral strategy, so that the benchmark is always the risk-free rate. In practice, most traders use the Sharpe ratio even when they are trading a directional (long or short only) strategy, simply because it facilitates comparison across different strategies [14].

In order to obtain the annualized Sharpe ratio of a pair one should multiply it for $\sqrt{252}$. It should be done because to get the annualized return, one should multiply the return for 252, and multiply the standard deviation for $\sqrt{256}$. Hence: 252 $/ \sqrt{252}$= $\sqrt{252}$.

The Sharpe ratio is important to make comparisons between different strategies with different returns and risks. Suppose we have a strategy A with $E(R_A)$ = 20% and σ_A = 4%, and a strategy B with $E(R_B)$ = 10% and σ = 1%. We have B as the best strategy, because its Sharpe ratio is 10, greater than 5, from A. It means one can leverage B and obtain the same return as A, with a smaller risk.

9. Results

The results were obtained using MatLab and stastistical package Eviews. We use four years of data from Yahoo Finance, from 03/27/2008 to 26/03/2012. We define one year for the training set and the remaining three for the test set. We do not use in our simulations stocks that have less than four years of records, or which have light volume. We do this by avoiding stocks that are not part of Bovespa index. The problem with stocks with light volume is that they have a wide bid-ask spread, and the test results usually get unrealistic, when put in practice.

As we said before, the strategy consists in buying or selling the spread when it gets apart more than two standard deviations from the mean. We find the spread by performing a OLS regression of the stocks on one another. The regression coefficient beta is the spread that we are looking for. Once we have its value, we evaluate the cointegration of the pair.

We do that by performing an Augmented Dickey Fuller test to check the stationarity of the spread. After that, we calculate the half-life of the spread. According to Smith, the function ϴ(z), from the Uhlenbeck-Ornstein equation can be obtained using an OLS regression of the daily change of the spread and the spread itself [12]. We add that to strategy, using now not only a gain stop, but also a time stop. Finally, we use the leverage suggested by the Kelly Formula and obtain the optimal return for the strategy.

When we talk about Sharpe ratios, performances between 0 and 1 are mediocre, results between 1 and 2 are good and greater than 2 are excellent. The results are quite according to the hypothesis we made. Indeed, the average Sharpe ratio of the cointegrated pairs is 0.6 against an average Sharpe ratio of -0.0917 the non-cointegrated pairs. Using the half-life given by the Uhlenbecker-Ornstein equation also improve results. The average Sharpe ratio of the cointegrated pairs improves achieving 2.3, a 283% increase and an excellent performance.

Finally, we use the leverage suggested by Kelly Formula in the training set, apply it in the test set and check the compounded growth rate produced by this leverage. We use 10% as the risk-free rate of return. Indeed, using half-Kelly, we achieve a better compounded growth rate. Although we have a poor result with BRTO4-VIVT4, Kelly criterion instructs one to deleverage in the presence of negative returns, what saves the investor from bankruptcy.

In summary, the results show that using cointegration as the criterion to choose the pairs we will trade leads to better results. Also, they show the use of half-life indicated by the Ornstein-Uhlenbeck equation leads to improvements in performance. Finally, the use of half-Kelly also improves the compounded rate of growth of the pairs. The only exception we have, is when performance turns bad, the use of leverage makes the performance worst.

Table 1: Augmented Dickey-Fuller Tests

These are the Augmented Dickey-Fuller tests values provided by EViews: if the t-statistic is smaller than -2.568627 the pair cointegrates at 90% level of confidence, if the t-statistic is smaller than -2.864925 the pair cointegrates at 95% level of confidence, if the t-statistic is smaller than -3.438268 the pair cointegrates at 99% level of confidence.

Null Hypothesis: R has a unit root

Exogenous: Constant

Lag Length: 0 (Automatic - based on SIC, maxlag=20)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			
Test critical values:	1% level	-3.438268	
	5% level	-2.864925	
	10% level	-2.568627	

Table 2: Sharpe Ratio of Cointegrated Pairs

In the first column we have the pair observed, in the second column we have the Sharpe ratio (the measure we use to evaluate performance) of the pair in the test set, and in the third we have the t-statistic of the ADF test.

Cointegrated-pairs	Sharpe test set	ADF t-statistic
AMBV4 X CRUZ3	1.7692	-4.6341
BISA3 X RSID3	1.2841	-5.2377
BRTO4 X VIVT4	-0.3150	-2.8049
BVMF3 X BOVA11	-	-3.2770
ENBR3 X CESP6	-	-4.1070
GOAU4 X GGBR3	0.5335	-2.6029
GOUA4 X USIM3	0.8910	-3.0026
GOUA4 X USIM5	0.3254	-3.1997
ITSA4 X BOVA11	0.8804	-2.8621
MMXM3 X CSNA3	-	-3.4280
MMXM3 X USIM5	-	-3.2065
OGXP3 X BOVA11	0.0594	-2.6521
TBLE3 X CESP6	-	-3.0943
TRPL4 X CESP6	-	-3.0324
Average Sharpe	***0.6785***	

Table3: Sharpe Ratio of Non-Cointegrated Pairs

In the first column we have the pair observed, in the second column we have the Sharpe of the pair in the test set, and in the third we have the t-statistic of the ADF test. Comparing the average Sharpe ratio with the average one in Table 2, we see that the Sharpe ratio of cointegrated pairs is indeed greater than of non-cointegrated pairs.

Non-cointegrated pairs	Sharpe test set	ADF t-statistic
AMBV4 X BOVA11	-1.4992	1.4326
BBAS3 X BBDC4	-0.4448	-0.8415
BBAS3 X BOVA11	-0.5865	-1.3211
BBDC4 X BOVA11	0.5280	-1.2885
BISA3 X CYRE3	0.5115	-1.9406
BISA3 X EZTC3	-0.8424	1.4840
BISA3 X MRVE3	-0.4400	-1.2553
BISA3 X TCSA3	0.4557	-2.3327
BRFS3 X BOVA11	-0.6612	-0.5455
BRML3 X MULT3	0.0559	-1.3233

BRTO4 X TNLP3	0.1337	-2.4538
CMIG4 X BOVA11	-	-2.4676
CPFE3 X CMIG4	0.1033	-2.3396
CSNA3 X BOVA11	-0.4650	-0.7885
CYRE3 X BOVA11	-	-1.9720
ELET3 X ELET6	0.1028	-1.3222
GETI4 X ELET6	0.3539	-1.8901
GFSA3 X BOVA11	0.0087	-0.7796
GGBR3 x GGBR4	0.0485	-1.1370
GOAU4 X GGBR4	0.0293	-1.0477
GOLL4 X TAMM4	0.9268	-0.0046
HGTX3 X AMAR3	-0.7207	0.2581
HGTX3 X LREN3	-1.2347	0.9063
HYPE3 X BOVA11	0.5137	-1.1830
ITSA4 X BBDC3	0.5581	-1.7785
KLBN4 X FIBR3	0.9946	0.1082
LAME4 X BOVA11	-0.0244	-1.6580
LREN3 X BOVA11	-1.0768	-0.7551
MMXM3 X BOVA11	0.4496	-2.0598
MPXE3 X CESP6	-0.3595	-1.4899
MPXE3 X CMIG4	-0.3702	-0.7046
MPXE3 X ELET3	-0.9278	0.1760
MPXE3 X ELET6	-0.2363	-0.1252
MRFG3 X JBSS3	0.1413	-2.2096
MRVE3 X BOVA11	0.3869	-2.1772
NATU3 X CRUZ3	-0.0994	0.3585
PCAR4 X BOVA11	-0.1328	-0.8818
PDGR3 X BOVA11	0.5489	-2.2187
PETR3 X BOVA11	-0.7814	-0.4704
PETR3 X PETR4	-0.4373	-1.2294
PETR4 X BOVA11	-0.5775	-0.7629
TBLE3 X ELET3	-0.2906	-0.9431
TBLE3 X ELET6	0.4340	-1.6931
TIMP3 X BOVA11	-1.3534	0.9507
TRPL4 X CMIG4	-0.0241	-1.7521
TRPL4 X ELET3	-0.0545	-0.8649
TRPL4 X ELET6	0.9236	-2.0790
TRPL4 X TBLE3	0.7085	-2.4940
USIM5 X BOVA11	-0.2657	-0.8296
VALE3 X BOVA11	0.9784	-1.8082
VALE3 X VALE5	-0.0999	-1.5042
Average Sharpe	***-0.0974***	

Table 4: Half-Life of Pairs

Here we have the half-life of pairs suggested by the OU Equation. We use it as time stop to close trades.

Pair	Half-Life
AMBV4 X CRUZ3	17
BISA3 X RSID3	11
BRTO4 X VIVT4	82
BVMF3 X BOVA11	28
ENBR3 X CESP6	29
GFSA3 X BOVA11	239
GOAU4XGGBR3	19
GOUA4 X USIM3	35
GOUA4 X USIM5	41
ITSA4 X BOVA11	36
MMXM3 X CSNA3	49
MMXM3 X USIM5	36
OGXP3 X BOVA11	54
TBLE3 X CESP6	35
TRPL4 X CESP6	60

Table 5: Sharpe Ratio of Pairs Using Half-Life

Here we have the Sharpe ratio of the pairs using half-life as a time stop. Again, comparing the results with the results in table 2, we see that using half-life improves the performance of the strategy.

Pair	Sharpe
AMBV4 X CRUZ3	6.65
BISA3 X RSID3	1.43
BRTO4 X VIVT4	0.6
BVMF3 X BOVA11	-
ENBR3 X CESP6	-
GFSA3 X BOVA11	0.03
GOAU4XGGBR3	4.6
GOUA4 X USIM3	3.99
GOUA4 X USIM5	2.16
ITSA4 X BOVA11	1.45
MMXM3 X CSNA3	-
MMXM3 X USIM5	-
OGXP3 X BOVA11	-0.2
TBLE3 X CESP6	-
TRPL4 X CESP6	-
Average Sharpe Ratio	***2.3011***

Table 6: Annual Compounded Growth Rate of Unleveraged (*g*) vs. Annual Compounded Growth Rate Using Half-Kelly Leverage (*g)**

Here we have the annual compounded growth rate of the strategy using the leverage suggested by Kelly criterion (g^*) and with no leverage (g). With the exception of BRTO4-VIVT4, this technique brings an improvement in the performance of the pairs.

Pair	***g****	***g***
AMBV4 X CRUZ3	0.6994	0.2587
BISA3 X RSID3	0.2173	0.1606
BRTO4 X VIVT4	-1.8732	-0.2133
GOAU4 X GGBR3	-	0.0542
GOAU4 X USIM3	0.3943	0.3134
GOAU4 X USIM5	-	0.0036
Average	***-0.1406***	**0.0962**

References

[1] Loomis, Carol J. 2012. “Tap dancing to work: Warren Buffett on practically everything, 1966-2012”. *Penguin Books*

[2] Graham, Benjamin. 1973. “The intelligent investor”. *Harper Business Essentials. 4th Edition Revised*

[3] Patterson, Scott. 2010. “The Quants: How a New Breed of Math Whizzes Conquered Wall Street and Nearly Destroyed It.” *Crown Business*

[4] Alexander, C and Dimitriu. A. 2002. “Cointegration-based trading strategies: A new approach to enhanced index tracking and statistical arbitrage”. *Discussion Papers 2002-08, ISMA Centre Discussion Papers in Finance Series*

[5] Wooldridge, Jeffrey. 2012. “Introductory Econometrics: A Modern Approach.” *Cengage Learning. 5th Edition*

[6] Gourieroux, Christian and Iasiak. Joann. 2001. “Financial Econometrics: Problems, Models and Methods.” *Princeton University Press*

[7] Murray, Michael P. 1994. “A Drunk and Her Dog: An Illustration of Cointegration and Error Correction” *The American Statician*

[8] Alves, Rodrigo de Barros. 2006. “Existe arbitragem entre a Metalúrgica Gerdau e a Gerdau S.A.?” *EPGE/FGV*

[9] Greene, William H. 2002. “Econometric Analysis” *Prentice Hall. 5th Edition*

[10] Pfaff, Bernahrd. 2008. “Analysis of Integrated and Cointegrated Series.” *Springer 2nd Edition*

[11] Smith, William. 2010. “On the Simulation and Estimation of the Mean-Reverting Ornstein-Uhlenbeck Process”. *Commodities Markets and Modelling*

[12] Uhlenbeck, George and Leonard Ornstein. 1930. “On the Theory of Brownian Motion.” *Physical Review*

[13] Thorp, Edward. 1997. “The Kelly Criterion in Blackjack, Sports Betting and the Stock Market”. *Handbook of Asset and Liability Management* Engle, R and Granger. C. 1987. “Cointegration and Error-Correction: Representation, Estimation and Testing.” *Econometrica*

[14] Sharpe, William. 1994. “The Sharpe Ratio.” *Journal of Portfolio Management*

[15] Herlemont, Daniel. 2004. “Pairs Trading, Convergence Trading. Cointregration.” *YATS Finance & Technologies*