

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO**

**DEPARTAMENTO DE ECONOMIA**

**MONOGRAFIA FINAL DE CURSO**

**Pairs Trading: Aplicação no Mercado de Ações Brasileiro**

**Joaquim Pedro Sampaio**

**Matricula: 1113017**

**Orientador: Marcelo Cunha Medeiros**

**Rio de Janeiro, Dezembro de 2016**

Declaro que o presente trabalho é de minha autoria e que não recorri para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

---

Joaquim Pedro Sampaio

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do autor.

## SUMARIO

<b>1. Introdução</b> .....	<b>5</b>
<b>2. Revisão Literária</b> .....	<b>8</b>
<b>3. A Abordagem de Cointegração</b> .....	<b>11</b>
<b>3.1 O conceito de Cointegração</b> .....	<b>11</b>
<b>3.2 O Modelo</b> .....	<b>14</b>
<b>4. Método</b> .....	<b>17</b>
<b>4.1 Dados</b> .....	<b>18</b>
<b>4.2 Aplicando a estratégia</b> .....	<b>18</b>
<b>4.3 Calculando retornos</b> .....	<b>19</b>
<b>5. Resultados Empíricos</b> .....	<b>21</b>
<b>6. Conclusão</b> .....	<b>25</b>
<b>7. Referências Bibliográficas</b> .....	<b>26</b>

## **Lista de Figura:**

<b>Figura 1:</b> Exemplo de operação com as acoes LMAE4 e MRVE3	15
<b>Figura 2:</b> Demonstração do mecanismo de janelas de treinamento e operação	18
<b>Figura 3:</b> Histograma do tempo de duração de cada operação	22
<b>Figura 4:</b> Retorno acumulado da estratégia de “Pairs Trading” e da Ibovespa	24
<b>Figura 5:</b> Evolução do preço e z-score de GGBR4 e EGIE3 (2007.1 – 2008.1)	25

## **Lista de Tabelas:**

<b>Tabela 1:</b> Estatísticas descritivas das operações com pares de ações	22
<b>Tabela 2:</b> Estatísticas descritivas dos retornos obtidos com o “Pairs Trading”	23

## **Lista de Equações:**

<b>Equação 1:</b> Especificação do teste Dickey-Fuller	12
<b>Equação 2:</b> Derivando equação final do teste Dickey-Fuller	12
<b>Equação 3:</b> Equação final do teste Dickey-Fuller	12
<b>Equação 4:</b> Primeira etapa do teste de cointegração	13
<b>Equação 5:</b> Resíduo do teste de cointegração	13
<b>Equação 6:</b> Aplicando teste Dickey-Fuller no resíduo	13
<b>Equação 7:</b> Cálculo do desvio entre dois ativos	14
<b>Equação 8:</b> Cálculo do z-score	14
<b>Equação 9:</b> Igualdade entre o capital alocado	16
<b>Equação 10:</b> Índice de Sharpe	21

## 1. Introdução

Arbitragem estatística é uma estratégia de investimento que visa explorar um desalinhamento de preços de mercado entre um ou mais ativos através de um método quantitativo. Em outras palavras, arbitragem estatística consiste em identificar um desalinhamento estatístico das relações de preços de ativos, que são verdadeiras em expectativa, e em seguida realizar um investimento para extrair lucro desse desequilíbrio temporário. A motivação para esse tipo de técnica é pautada na existência de uma relação de longo prazo nos preços de ativos; esperasse encontrar padrões estatísticos observados nos dados históricos. Mesmo que tais padrões existam, a dificuldade com esse tipo de estratégia é identificar e precisar o tempo que os mesmos perdurarão.

“*Pairs Trading*” é um tipo de estratégia de arbitragem estatística que busca explorar desvios de curto prazo de uma relação de longo prazo entre duas ações, sendo uma das estratégias mais comuns de arbitragem estatística na literatura. A ideia principal é identificar um par de ações que se mova conjuntamente ao longo do tempo e tomar uma posição de compra e venda simultaneamente quando os preços divergirem significativamente. Essa estratégia é utilizada no mercado financeiro desde os anos oitenta e continua sendo uma importante ferramenta utilizada por Investidores ao redor do mundo.

Para a estratégia ser bem sucedida, é necessário supor que exista algum mecanismo de reversão de média no preço dos ativos. Se o preço de ações fosse verdadeiramente aleatório, a estratégia não funcionaria. Nesse contexto, a Lei de Preço Único afirma que dois ativos com a mesma estrutura de retorno em todos os estados da natureza, em equilíbrio, deveriam ter o mesmo valor presente (Caldeira & Moura, 2013; Gatev, Goetzmann, & Rouwenhorst, 2006). Portanto, ações com estrutura de retorno similar deveriam possuir uma relação estável de preço ao longo do tempo.

A estratégia pode ser considerada neutra a riscos de mercado, porque ela envolve a compra e a venda de um ativo, do mesmo mercado, ao mesmo tempo. O lucro advindo desta tática é derivado da divergência de preços entre dois ativos, em vez da direção que cada instrumento segue. Portanto, esperamos que os retornos da estratégia não sejam

correlacionados com o retorno de mercado. Isso não significa que a estratégia seja livre de riscos, como qualquer investimento, essa estratégia está exposta a uma variedade de riscos. Mas dado suas características, a estratégia mantém exposição oposta a riscos de mercado na posição comprada e vendida ao longo do tempo, o que neutraliza a exposição a riscos sistêmicos específicos do mercado que estaremos atuando. Essa característica torna a estratégia especialmente atraente porque possibilita obter retornos no mercado de ações em momentos de contração ou estabilidade.

Na literatura, existem diferentes maneiras de selecionar pares. Cointegração e correlação são dois dos conceitos mais utilizados neste contexto. Neste trabalho, uma abordagem de cointegração será utilizada. Procedimento que foi utilizada pioneiramente por Lucas (1997) e Alexander (1999). Cabe neste momento uma breve distinção dos conceitos de correlação e cointegração no contexto de arbitragem estatística, que apesar de serem relacionados, possuem implicações diferentes. Alta correlação entre ativos não necessariamente implica que os ativos são cointegrados. Correlação reflete movimentos em comum entre ativos, mas costuma ser instável ao longo do tempo. Alta correlação por si só não é suficiente para garantir performances de longo prazo. Estratégias de investimento baseadas em correlação frequentemente necessitam calibragem devido à variabilidade na correlação ao longo do tempo<sup>1</sup>. Por outro lado, cointegração mede a relação de longo prazo entre ativos, mesmo que a correlação aparente baixa em alguns períodos. Portanto, estratégias baseadas em cointegração podem ser mais efetivas no longo prazo.

Apesar da literatura no tema ser ampla, existem poucas aplicações desta estratégia documentadas academicamente no mercado de ações brasileiro. A motivação do trabalho é montar uma estratégia automatizada de trading que siga regras simples e que possua retornos atraentes e não correlacionados com o mercado. Nossas regras de investimento giraram em torno de dois estágios principais. O primeiro consiste em identificar ações que se movam conjuntamente e o segundo de comprar e vender as ações simultaneamente quando seus preços divergirem além de certo ponto.

---

<sup>1</sup> Segundo Caldeira & Moura (2013)

O restante deste trabalho será organizado da seguinte maneira. O capítulo 2 apresentará uma revisão literária sobre o tema. O capítulo 3 explicará o conceito de cointegração e sua aplicação com maior riqueza de detalhes. O capítulo 4 descreverá o método utilizado. O capítulo 5 apresentará os resultados obtidos simulando a estratégia descrita no capítulo anterior. E finalmente, o capítulo 6 fará uma conclusão do trabalho.

## 2. Revisão Literária

Não é possível precisar a origem exata de estratégias de investimento que visavam arbitrar desequilíbrios na relação de preços entre ativos. Porém, na década de oitenta, Nunzio Tartaglia montou uma equipe de físicos, matemáticos e engenheiros de computação para explorar estratégias de arbitragem no mercado de capitais.<sup>2</sup> O grupo de acadêmicos desenvolveu métodos estatísticos sofisticados para desenvolver programas que executavam estratégias de trading automatizadas. Uma das estratégias utilizada pelo grupo consistia em achar pares de ações que tendessem a fazer um movimento conjunto. Essa estratégia fez grande sucesso e ficou conhecida como “*Pairs Trading*”.

O grupo teve grande sucesso em seu primeiro ano de existência. Porém, após uma série de performances decepcionantes, a congregação foi diluída em 1989. Apesar do sucesso restrito em suas estratégias, o evento despertou o interesse de grandes fundos e investidores, tornando métodos quantitativos de arbitragem estatística objeto de estudo ao redor do mundo.

A partir daí surgiram estudos que tentaram fundamentar a estratégia com explicações estatísticas e comportamentais. Jegadeesh and Titman's (1995) argumentam que os lucros obtidos investindo em pares de ações é parcialmente explicado por reações exageradas do mercado a informações específicas das empresas, em vez da reação de preços a fatores comuns. O próprio Tartaglia utiliza uma explicação psicológica para o *Pairs Trading*. Ele dizia que seres humanos não gostam de investir contra a natureza humana, que aspira comprar ações quando o preço está subindo e não caindo (Gatev, Goetzmann, & Rouwenhorst, 2006).

Outro conceito relevante para explicar o racional por trás de investir em pares de ação é a teoria do preço relativo. A precificação de ativos pode ser vista de maneira absoluta ou relativa. Apreçamento absoluto precifica ativos baseando-se em fundamentos, como o fluxo de caixa descontado, com o fim de achar um nível ‘justo’ para o preço do ativo. Porém o apreçamento relevante para a análise é o relativo. Apreçamento relativo, não envolve precisar o patamar que o preço de um ativo deveria estar, porém, ele denota que

---

<sup>2</sup> Gatev, Goetzmann, & Rouwenhorst, 2006).

dois ativos considerados substitutos deveriam ter o mesmo preço (Gatev, Goetzmann, & Rouwenhorst, 2006). Portanto, deveria existir uma razão estável entre o preço de duas ações parecidas.

A Lei de Preço Único fornece embasamento teórico á esse conceito. Ela afirma que dois ativos com retornos iguais em todos os estados da natureza, em equilíbrio, deveriam ter o mesmo valor presente.<sup>3</sup> Em outras palavras, dois ativos com o mesmo fluxo de caixa esperado deveriam ter o mesmo preço de mercado. Portanto, ações com estrutura de retorno similar deveriam possuir uma relação estável de preço ao longo do tempo. Tal raciocínio fornece base teórica para a ideia básica do *Pairs Trading*, que consiste em explorar divergências em padrões de preços relativos. Hendry & Juselius (2001) utilizam a intuição derivada da Lei do Preço Único para mostrar que divergências de curto prazo no preço relativo podem criar oportunidades de arbitragem dependendo da duração de tais divergências. De acordo com Perlin (2007):

*“The main logic behind the expected profits of pairs trading strategy is: if the correlated movement between the pairs is going to continue in the future then, when the distance between an asset and its pair is higher than a particular threshold value (d), there is a good possibility that such prices are going to converge in the future, and this can be explored for profit purposes.”* (Perlin, 2007, p.7)

Apesar do embasamento teórico descrito ate o momento, existe vasta literatura corroborando a lucratividade do “*Pairs Trading*” empiricamente. Uma das referencias literárias na comprovação de rentabilidade é o trabalho de Gatev et Al (2006). O estudo utiliza dados diários do preços de ações de 1962 a 2002. As regras de investimento utilizadas por esse autor são razoavelmente simples e parecidas com a estratégia que será utilizada nesse trabalho. A grande diferença entre os estudos se deve ao método utilizado para selecionar pares. O método utilizado no trabalho em questão consistia em montar um índice de retorno cumulativo para cada ação. Em seguida os pares eram formados minimizando as distancias ao quadrado entre as series de preços normalizados, além de restringir possíveis pares por setor. O estudo resultou em uma estratégia de investimento com retorno médio anualizado de 11%.

---

<sup>3</sup> Caldeira & Moura (2013) e Gatev, Goetzmann, & Rouwenhorst, (2006).

No Brasil, existe pouca documentação acadêmica comprovando a rentabilidade do *Pairs Trading*. Um dos poucos trabalhos existentes, de João Caldeira e Guilherme Moura (2013), simula uma aplicação da estratégia no mercado de ações brasileiro. Os autores utilizam uma abordagem de cointegração para selecionar candidatos a pares de ações e em seguida selecionam os melhores baseando-se no 'sharpe ratio', calculado com base nos dados obtidos durante a seleção dos candidatos a pares. A estratégia foi testada usando dados de 2005 a 2012, nos quais os últimos quarto anos foram usados para testar a estratégia fora da amostra. Levando em conta custos de transação, uma rentabilidade de 189.29% foi encontrada durante os 4 anos de teste, com média anual igual a 16.38%; fornecendo evidência empírica á lucratividade da estratégia. Em adição, foi encontrada uma baixa correlação entre os retornos da carteira de pares e os retornos do Ibovespa, fornecendo evidência á neutralidade da estratégia em relação ao mercado.

### 3. A Abordagem de Cointegração

A aplicação do conceito de cointegração para alocar ativos foi utilizada pioneiramente por Lucas (1997) e Alexander (1999). Suas principais características, isto é, erro com reversão de média, aumento da estabilidade dos parâmetros e uma melhor utilização das informações incluídas nos preços das ações, permitem um desenho flexível de várias estratégias de negociação, financiadas e autofinanciadas.<sup>4</sup>

Durante a última década, o conceito de cointegração tem sido cada vez mais utilizado em finanças. É uma ferramenta extremamente poderosa que permite modelar séries temporais não estacionárias. Porém, a observação fundamental que justifica a aplicação deste conceito na análise do preço de ações é que um sistema envolvendo preços de ações não estacionárias em nível pode ter um componente de tendência estocástica em comum (Stock & Watson, 1998); salientando uma das principais vantagens em relação ao conceito de correlação; o uso de informações contidas no nível de variáveis financeiras.

Apesar de existir metodologias alternativas na literatura, a abordagem escolhida nesse trabalho será a cointegração. As vantagens de tal abordagem podem ser observadas nos estudos de Alexandre e Dimitriu (2002), Vidyamurthy (2004), Caldeira e Portugal (2011), Caldeira e Moura (2013), entre outros. O livro elaborado por Vidyamurthy (2004) é uma das principais referências na literatura para esse tipo de abordagem. Apesar de não demonstrar resultados empíricos, o autor descreve os principais conceitos teóricos e a forma de implementar o “Pairs Trading” utilizando uma possível relação de cointegração entre as ações. No Brasil, os trabalhos de Caldeira e Moura (2013) são a principal referência no assunto. Cabe nesse momento uma descrição do conceito de cointegração.

#### 3.1 O conceito de Cointegração

---

<sup>4</sup> Caldeira & Moura (2013)

Cointegração é uma ferramenta estatística, onde duas series temporais que são integradas de ordem 1,  $I(1)$ , podem ser combinadas linearmente para produzir uma serie temporal que seja estacionaria,  $I(0)$ . Portanto, a técnica de “Pairs Trading” utilizada nesse trabalho é baseada na presunção que combinações lineares de preços, que sejam  $I(1)$ , tendem a reverter para um equilíbrio de longo prazo, de forma que uma regra de operação possa ser formulada para explorar desvios temporários.

**Definição:** Um vetor  $Y_t$  de series temporais  $n \times 1$  é cointegrado se:

- cada um de seus elementos é não estacionário e
- existe um vetor diferente de zero,  $\beta$ , de forma que  $\beta Y_t$  seja estacionário

A definição acima descreve o conceito de cointegração para um vetor,  $Y_t$ , de series temporais com  $n \times 1$  dimensões, implicando necessariamente que o conceito pode ser aplicado em um sistema com mais de duas variáveis. Porém, neste trabalho, o conceito de cointegração será aplicado em um único par de ações.

O primeiro passo para testar se duas ações são cointegradas é verificar se ambas as series,  $Y_t$  e  $X_t$ , são não-estacionarias. Utilizaremos o teste Dickey-Fuller (DF) para fazer tal verificação. A seguinte especificação do teste foi utilizada:

$$Y_t = \theta + \beta Y_{t-1} + \varepsilon_t \quad (1)$$

Onde  $Y_t$  é o valor da serie temporal  $Y$  no instante  $t$ ,  $\theta$  é uma constante e  $\varepsilon_t$  é um componente de choque aleatório. Se subtrairmos  $Y_{t-1}$  de ambos os lados da equação acima, temos as seguintes equações:

$$Y_t - Y_{t-1} = \theta + (\beta - 1)Y_{t-1} + \varepsilon_t \quad (2)$$

$$\Delta Y_t = \theta + \beta^* Y_{t-1} + \varepsilon_t \quad (3)$$

Onde  $\Delta Y_t = Y_t - Y_{t-1}$  e  $\beta^* = (\beta - 1)$ . Em seguida, testaremos se a serie  $Y_t$  é estacionaria com as seguintes hipóteses:

$H_0: \beta^* = 0$  (a serie é não – estacionaria)

$H_1: \beta^* < 0$  (a serie é estacionaria)

A estatística t do coeficiente  $\beta^*$  é comparada ao valor crítico da distribuição t de Student ajustada para o teste DF referente ao intervalo de confiança de 5%. A necessidade de ajustar a distribuição t de Student decorre de sua utilização no resíduo de uma regressão, alterando suas propriedades. Caso a estatística t seja maior que o valor crítico, a hipótese nula, que a serie é não-estacionaria, não pode ser rejeitada.

Em seguida, para testar a presença de cointegração entre pares de ações, utilizaremos a metodologia desenvolvida por Engle-Granger. Esta metodologia consiste em dois passos fundamentais. Primeiro iremos regressar uma serie de preços  $Y_t$  em seu candidato a par  $X_t$ ; considerando que ambas as series são não-estacionarias. Estimamos a seguinte equação:

$$Y_t = \alpha X_t + \mu_t \quad (4)$$

$$\mu_t = Y_t - \alpha X_t \quad (5)$$

Onde  $\mu_t$  é o resíduo da regressão de  $Y_t$  em  $X_t$ . Em seguida, testamos o resíduo,  $\mu_t$ , por estacionariedade utilizando o mesmo procedimento descrito acima, teste Dicky-Fuller, com os mesmo 5% de intervalo de confiança. A equação tem a seguinte especificação derivada com procedimento análogo a equação (3):

$$\Delta\mu_t = C + \delta\mu_{t-1} + \epsilon_t \quad (6)$$

A opção de manter a constante deve-se ao fato de que nesse ponto não queremos identificar apenas os pares que sejam estacionários com media zero. Pares de ações que sejam estacionários ao redor de uma certo nível são igualmente relevantes. Lembrando que para lucra em uma operação de “Pairs Trading” é necessária a característica de reversão de media, não importando o nível desta mesma.

Por fim, para que as ações sejam consideradas cointegradas, é necessário que o resíduo,  $\mu_t$ , seja estacionário. Portanto, desta vez queremos rejeitar a hipótese nula de que o coeficiente  $\delta$  seja significativamente igual a zero. Rejeitando a hipótese nula ao nível de significância escolhido (5%), podemos concluir que as ações são cointegradas.

### 3.2 O Modelo

Normalmente, algoritmos de negociação de pares possuem dois componentes principais. O primeiro item fundamental dessa metodologia é um algoritmo de seleção de pares que, no nosso caso, é essencialmente baseado em testes de cointegração. O objetivo desta fase é identificar pares cuja combinação linear exibe um componente previsível significativo que não está correlacionado com movimentos subjacentes do mercado como um todo. Com este objetivo, verificamos primeiro se todas as séries são integradas da mesma ordem, I (1). Isto é feito por meio do teste Dickey Fuller (DF). Após ter passado o teste DF, testes de cointegração são realizados em todas as possíveis combinações de pares. Para testar a cointegração, adotamos a abordagem de Engle e Grangers em duas etapas, seguindo o procedimento descrito anteriormente.

Para as relações de cointegração detectadas, a segunda parte do algoritmo cria sinais de negociação com base em regras de decisão de investimento predefinidas. Para implementar a estratégia, precisamos seguir algumas regras de negociação, ou seja, determinar quando abrir e quando fechar uma posição. Primeiro, calculamos o spread entre as ações. O spread é calculado como:

$$\epsilon_t = \log(Y_t^A) - \alpha \log(X_t^B) \quad (7)$$

Onde  $\epsilon_t$  é o valor do spread no tempo  $t$  e  $\alpha$  é o coeficiente de cointegração. Em conformidade, calculamos o que chamaremos de z-score, analogamente a Caldeira e Moura (2013), definido como:

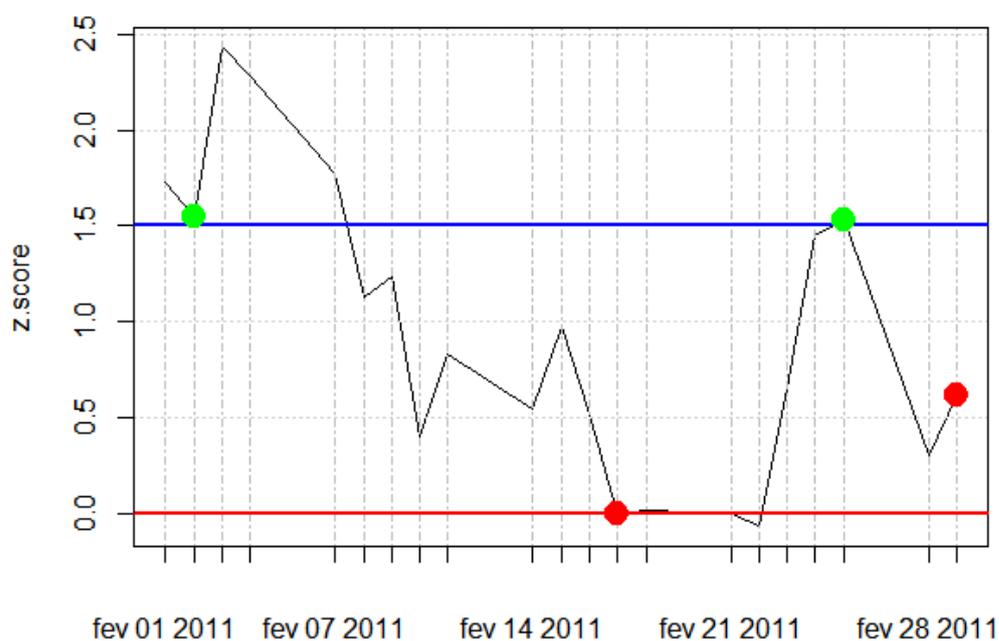
$$Z_t = \frac{\epsilon_t - \mu_\epsilon}{\sigma_\epsilon} \quad (8)$$

Onde  $\mu_\epsilon$  é a média de  $\epsilon_t$  durante o período de estimação da relação de cointegração e  $\sigma_\epsilon$  é o desvio padrão da mesma. O z-score mede a distância da média de longo prazo em unidades de desvio padrão (de longo prazo).

Nossa regra básica será abrir uma posição quando o z-score atingir os 1,5 limiares de desvio padrão de cima ou de baixo. Esta situação implica que os estoques estão mal precificados em termos de seu valor relativo. Se o z-score atingir o limiar de -1,5 desvio padrão, isso significa que a carteira de pares está abaixo do seu valor de equilíbrio de longo prazo. Neste caso, deve-se comprar a carteira, o que significa comprar a ação A e vender ação B. Se o z-score atinge o limite de 1,5 desvio padrão acima, a carteira de pares está sobrevalorizada e deve-se vendê-la a descoberto, o que significa vender a ação A e comprar ação B. A posição é fechada quando o z-score retornar a media ou se a janela encerrar. A razão para abrir negócios somente quando o z-score está longe do equilíbrio é negociar somente quando pensamos que detectamos uma excursão anômala do resíduo de cointegração. Em todos os casos abrir ou fechar uma posição significa comprar e vender as ações simultaneamente. Mais especificamente, os sinais são definidos da seguinte forma:

$$\begin{aligned} \text{Comprar a carteira se } & Z_t < -1,5 \\ \text{Vender a carteira se } & Z_t > 1,5 \\ \text{Fechar a operação comprada se } & Z_t > 0 \\ \text{Fechar a operação vendida se } & Z_t < 0 \end{aligned}$$

**Figura 1:** Exemplo de operação com as ações LMAE4 e MRVE3 em fevereiro de 2011.



Os pontos verdes denotam o momento de abertura das operações, enquanto os pontos vermelhos representam os fechamentos. A primeira operação foi encerrada quando o z-score retornou a média. Já a segunda operação foi encerrada no final da janela. As operações acima obtiveram um retorno de 8%.

Adicionalmente, iremos incluir um critério para selecionar os pares de ação que serão utilizados no portfólio. Dentre todos os pares que possuem uma relação de cointegração em uma determinada janela, apenas os 20 pares que tiverem o menor p-valor no teste de cointegração serão selecionados para teste fora da amostra. Desta forma, estaremos selecionado os pares que são considerados cointegrados com o maior nível de significância.

Assim que a distancia entre as ações se afasta de sua média de longo prazo, pode-se apostar que o spread retornará à sua média, porém não sabemos se vamos ganhar mais em posições compradas ou vendidas. Uma vez que uma operação é iniciado, a carteira não é balanceada. Portanto, após a abertura de uma posição, mesmo quando os preços se movem e a posição deixa de ser neutra, a carteira não é ajustada. Só temos dois tipos de transações que são admitidas pela metodologia da estratégia; mover para uma nova posição, ou a liquidação total de uma posição previamente aberta.

A estratégia de investimento que visamos implementar é neutra ao mercado, portanto, manteremos uma posição longa e uma posição curta, ambas com o mesmo valor em moeda local. Essa abordagem elimina a exposição líquida ao mercado de ações, portanto, os retornos fornecidos não devem ser afetados pela direção do mercado. Adicionalmente, a estratégia aqui adotada pretende possuir um investimento líquido de zero. Faremos isso alocando um montante igual no ativo A e B, como podemos ver na equação (9), eliminando a necessidade de desembolso financeiro. Isso pode ser feito através do empréstimo de um numero de ações do ativo B, imediatamente vendendo essas ações e investindo o montante em  $\theta$  ações de A.

$$\theta P_t^A = P_t^B \quad (9)$$

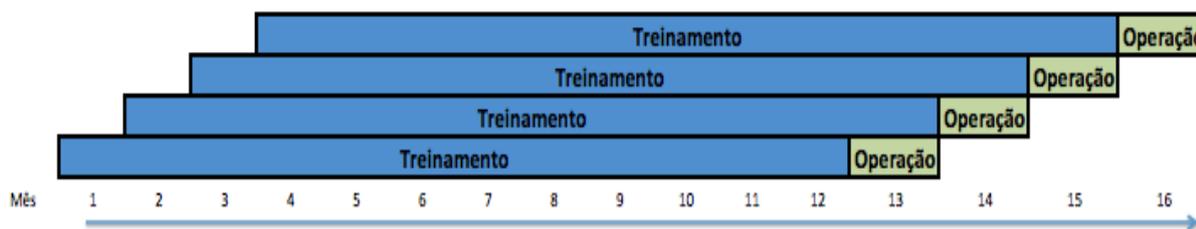
As principais limitações desse método decorrem da dificuldade de estimar o vetor de cointegração e de estimativas incorretas do desvio esperado, podendo gerar falsos sinais

de operação. Além disso, o procedimento de teste de cointegração em duas etapas gera resultados sensíveis à ordem das variáveis, portanto, os resíduos podem ter diferentes propriedades estatísticas. Além disso, se as duas séries não forem cointegradas, a "regressão de cointegração " leva a estimativas espúrias (Lim & Martin, 1995), tornando a análise de reversão à média nos resíduos não confiável.

#### **4. Método**

Nossa implementação de “*Pairs Trading*” envolve dois estágios simples. O primeiro estágio envolve a escolha de pares utilizando um período de 12 meses. Em seguida, a estratégia de investimento é aplicada aos pares selecionados por 1 mês. Esses passos serão repetidos sucessivamente até que toda janela de dados disponível seja utilizada.

**Figura 2:** Demonstração do mecanismo de janelas de treinamento e operação.



## 4.1 Dados

Os dados utilizados nesse trabalho foram obtidos através do Software Economatica. Foram colhidos os preços de fechamento e volume diário de todas as ações transacionadas na IBOVESPA entre os anos de 2006 e 2015; corrigindo para “*stock splits*” e dividendos para evitar falsos sinais de negociação (Broussard & Vaihekoski 2012). Apenas as 50 ações com maior volume durante o período de formação serão candidatas a pares. Como as ações mais transacionadas variam ao longo de cada período de formação, as ações na amostra também variam, livrando a amostra do viés de sobrevivência. Adicionalmente, utilizar as ações com maior liquidez minimizam problemas relacionados a grandes custos operacionais e a dificuldade no aluguel de ações. É comum na literatura limitar a escolha de pares por setor, por exemplo Dunis et al. (2010). Porém, neste trabalho não adotaremos esta restrição. De modo que pares de companhias pertencentes a diferentes setores possam ser operados livremente, contanto que eles satisfaçam os critérios de seleção.

## 4.2 Aplicando a estratégia

O primeiro estagio da estratégia envolve a formação de pares de ações. As 50 ações com maior volume durante os 12 meses do período de formação serão as candidatas iniciais. Em seguida, é aplicado um teste de estacionariedade, utilizando o teste Dickey-Fuller com parametrização descrita no terceiro capítulo. Apenas as ações consideradas não-estacionárias serão selecionadas para a próxima etapa. Na prática, aproximadamente 95% da ações não apresentaram evidências de estacionariedade.

Após verificar a presença de raiz unitária, todas as ações serão testadas entre si para verificar a presença de cointegração utilizando metodologia desenvolvida por Engle-Granger e descrita em detalhe no terceiro capítulo. A equação (4) é estimada e, em seguida, o resíduo é testado por raiz unitária através do teste DF com valores críticos alterados por se tratar de uma variável estimada. Caso o resíduo seja estacionário, existe uma relação de cointegração entre  $X_t$  e  $Y_t$ . Os pares de ações que possuem uma relação de cointegração serão ranqueados pelo p-valor do teste de estacionariedade e apenas os pares com os 20 menores p-valor serão testados fora da amostra. Esse procedimento é repetido para cada janela de formação de pares.

Apos a escolha dos pares, a segunda etapa consiste em aplicar as regras de investimento, descritas no capítulo 3.2, durante o mês que sucede o período de formação. Como a amostra começa em 2006, o período de teste, “out of sample”, se inicia em janeiro de 2007. Desta forma, os retornos apresentados neste trabalho ocorreram entre os anos de 2007 e 2015, com um total de 108 janelas, com duração de um mês.

### **4.3. Calculando retornos**

O cálculo do retorno de operações long-short com pares de ações não é uma tarefa trivial. Como o investimento inicial é igual a zero por construção, não é possível computar retornos sobre um capital investido nulo. Porém, existem duas medidas alternativas de medir o “retorno” de uma operação de pares.

A primeira consiste em adotar a convenção de que o retorno é dado pelo lucro ou perda da operação sobre o valor inicial alocado a cada ponta, no formato de “P&L”. Esta será a métrica utilizada neste trabalho. Usaremos o valor total de R\$1 em cada janela, de forma que o resultado obtido tenha a interpretação de retorno percentual. O lucro obtido em cada operação é apenas reinvestido na mesma janela e par no qual foi obtido. De forma que não reinvestimos o lucro das operações em uma determinada janela nas janelas subsequentes, ou de um par nos outros pares da mesma janela. O retorno total foi calculado simplesmente pela soma dos lucros ou prejuízos obtidos em cada janela. É importante ressaltar que este capital parado poderia ser investido em outros ativos rentáveis, como os títulos do Governo Federal brasileiro, o que tornaria a estratégia mais rentável. Deste modo, os lucros apresentados nesse trabalho são calculados de forma conservadora.

Caldeira e Moura (2013) adotam um desenho “fully invested”, no qual um certo capital é 100% alocado nos pares que possuem transações abertas, e a carteira é constantemente balanceada para manter o capital totalmente investido. Em contraste, neste trabalho, consideramos que  $1/20$  ( $1/\text{numero de pares}$ ) do capital (R\$1) é investido em cada par. Esse valor ficará parado caso nenhuma operação naquele par esteja aberta. Empiricamente, houve casos onde nenhuma operação foi efetuada para um determinado par na janela referente. Nestes casos, o capital alocado permaneceu imóvel. Esse fato corrobora a tese de que os lucros calculados neste trabalho são conservadores.<sup>5</sup>

A segunda forma de calcular o retorno consiste em utilizar o capital alocado para manter a operação como denominador da equação. Como a estratégia é autofinanciada, o custo de manter a operação é igual a zero. Porém, na prática, corretoras cobram um custo operacional e requerem uma quantia de margem; uma forma de limitar a alavancagem da operação. Tais valores poderiam ser utilizados como denominadores na conta do retorno

---

<sup>5</sup> Gatev, Goetzmann, & Rouwenhorst, (2006)

na forma de capital investido. Como não estamos levando em conta custos de operação, essa métrica não será utilizada.

Por fim, utilizaremos o Índice de Sharpe para medir o retorno da estratégia ajustado por sua volatilidade. Os valores foram computadas da seguinte maneira, seguindo procedimento análogo a Caldeira e Moura (2012):

$$SR = \frac{R^A}{\sigma^A} \quad (10)$$

Onde  $R^A$  é o retorno anualizado da estratégia e  $\sigma^A$  é a volatilidade anualizada do retorno.

## 5. Resultados Empíricos

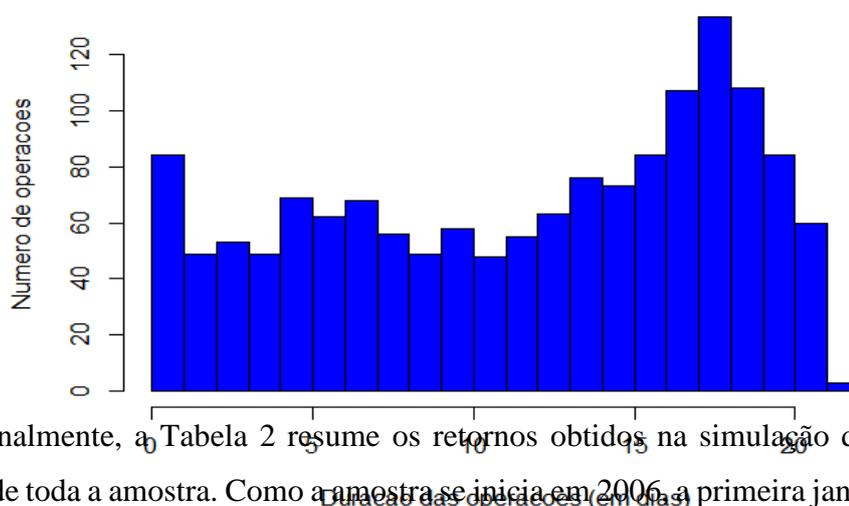
A tabela 1 resume os resultados obtidos nas operações. Somando toda a amostra, 85 ações foram utilizadas para formar 632 pares. Utilizando esse pares, 1595 operações foram

efetuadas com um retorno médio de 0,6% e desvio padrão de 4,8%. Podemos ver que 55% das operações geraram lucro enquanto 45% provocaram prejuízo. Ademais, a media de duração de cada operação ficou em 12 dias, significativamente abaixo do tamanho da janela de operação (22 dias). Porém, podemos ver na Figura 3 que existe uma concentração de operações que perduraram quase todo o período. Indicando que uma janela de operação mais longa poderia ser mais apropriada.

**Tabela 1:** Estatísticas descritivas das operações com pares de ações.

Estatísticas das operações	
# de ações selecionadas	85
# de pares transacionados	632
# de operações	1595
Retorno médio por operação	0,60%
Desvio padrão	4,80%
% de operações ganhadoras	55%
% de operações perdedoras	45%
Pior retorno (por operação)	-19,30%
Melhor retorno	34,20%
Mediana do retorno	0,42%
Duração media das operações	12
Duração mediana das operações	13

**Figura 3:** Histograma do tempo de duração de cada operação.



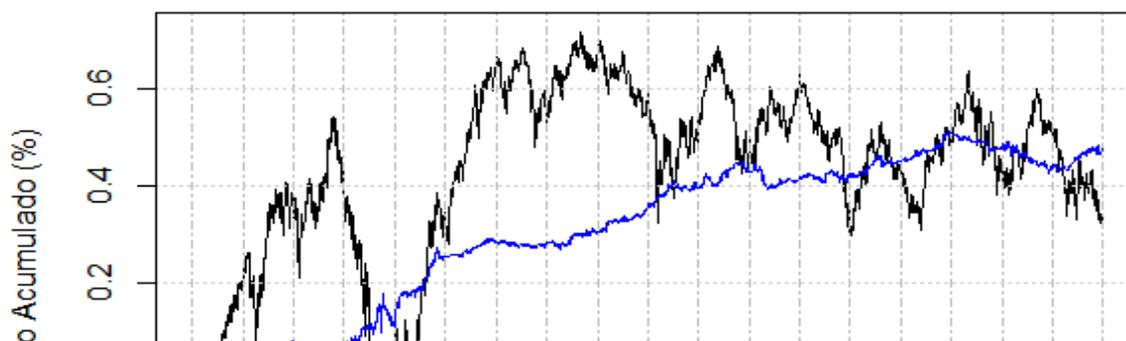
anualizado de 5,15%, resultando em um retorno acumulado de 46,3% em nove anos. Apesar desses retornos serem modestos, principalmente quando comparados as elevadas taxas de juros e inflação no Brasil, vemos que o índice Ibovespa obteve um retorno cumulativo inferior em aproximadamente 15%, tornando a estratégia mais rentável que seu benchmark.

**Tabela 2:** Estatísticas descritivas dos retornos obtidos com a estratégia “Pairs Trading”. Os retornos são referentes ao período que se estende de 2007 a 2015.

Estatísticas do retorno da estratégia “Pairs Trading”	
# de observações dentro da amostra	2473
# de dias na janela de formação	250
# de dias na janela de operação	22
# de janelas	108
# de pares em cada janela	20
Retorno anualizado médio	5,15%
Índice de Sharpe do retorno anualizado	1,07
Retorno Acumulado	46,30%
Retorno Acumulado Índice IBOV (Benchmark)	32,60%
Pior mês	-3,50%
Melhor mês	6,20%
% de dias com retorno positivo	52,50%
% de meses com retorno positivo	63,90%
Skewness	0,84

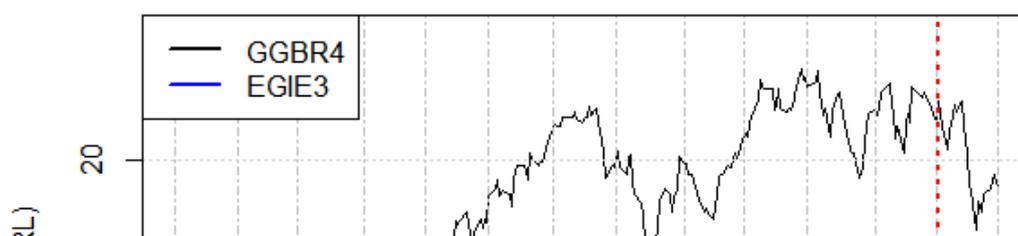
A figura 4 corrobora o sucesso da estratégia em relação ao índice Ibovespa. Podemos ver que enquanto o índice apresentou alta volatilidade em seus retornos acumulados, a estratégia manteve uma volatilidade comparativamente baixa. Tendo como pior e melhor retorno obtido em um mês de operação -3.5% e 6,2%, respectivamente.

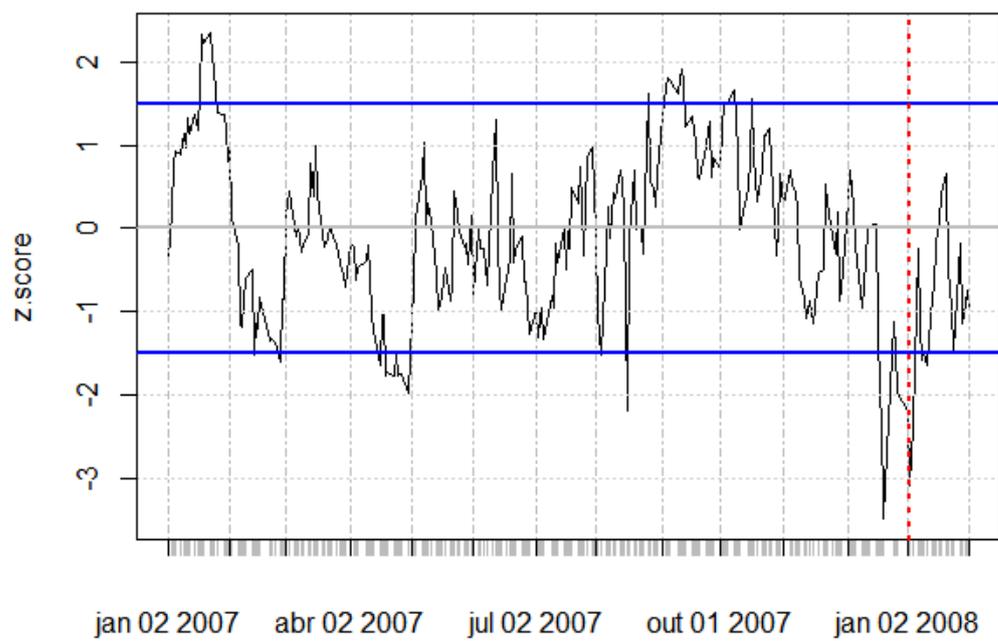
**Figura 4:** Retorno acumulado da estratégia de “Pairs Trading” e da Ibovespa. Período: 2007:01 – 2015:12



Por fim, a Figura 5 ilustra um exemplo de ações que possuíram uma relação de cointegração no período de 2007 a 2008. Pelo gráfico do z-score, podemos ver que essas ações mantiveram uma relação razoavelmente estável e com característica de reversão de media ao longo do período de formação e operação. Porém, tal padrão não foi observado em todos os pares, refletindo um numero significativo de operações, 45%, realizando prejuízos. Uma estimação mais criteriosa da relação de cointegração possivelmente poderia ampliar os lucros apresentados.

**Figura 5:** Evolução do preço das ações e z-score de GGBR4 e EGIE3 de janeiro de 2007 a janeiro de 2008, percorrendo uma janela de formação e operação. A linha vermelha pontilhada indica o fim do período de formação de pares.





## 6. Conclusão

Neste trabalho, aplicamos uma estratégia de arbitragem estatística, conhecida como “*Pairs Trading*”, para ações da Bolsa de Valores de São Paulo. A estratégia foi implementada com base na relação de cointegração entre series de preço e explorando a

característica de reversão de média dos desvios. Testes de cointegração foram aplicados a todas as possíveis combinações de ações selecionadas em cada período, com o fim de identificar pares de ações que compartilham uma relação de equilíbrio de longo prazo. Do total de 1.225 pares possíveis, em média 130 pares possuíram uma relação de cointegração durante o período de treinamento ao nível de significância de 5%. Posteriormente, selecionamos os 20 pares com menor p-valor no teste de cointegração. A partir daí, uma carteira diversificada contendo 20 pares foi selecionada para testes fora da amostra.

O lucro líquido acumulado ao longo de nove anos de testes foi de 46,3%, com média anual de 5,15%. Além disso, os retornos da estratégia implementada apresentaram níveis relativamente baixos de volatilidade e nenhuma relação significativa com o índice Ibovespa, confirmando sua neutralidade ao mercado. Os resultados são atraentes quando comparados ao índice Ibovespa, principalmente se considerarmos que a estratégia não necessita de desembolso financeiro. Porém, os retornos não são tão interessantes quando se leva em consideração as altas taxas de juros e inflação vigentes no período em questão.

Em projetos de pesquisa futuros, o principal refinamento seria levar em conta alguma medida de custo de transação que aproximaria os resultados obtidos neste trabalho a realidade prática. Outra sugestão é testar a estratégia para diferentes janelas de formação e operação; possivelmente utilizando dados de alta frequência. Um método para identificar a estabilidade dos parâmetros de cointegração também seria interessante para tentar aumentar a rentabilidade e mitigar os riscos.

Em conclusão, foi possível obter lucros com a estratégia “*Pairs Trading*” no mercado de ações brasileiro utilizando regras de investimento relativamente simples. Os resultados apresentados reforçam a utilização do conceito de cointegração como ferramenta importante na gestão quantitativa de fundos.

## **7. Referências Bibliográficas**

Alexander, Carol. "Optimal hedging using cointegration." *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 357.1758 (1999): 2039-2058.

Alexander, Carol, and Anca Dimitriu. "The cointegration alpha: Enhanced index tracking and long-short equity market neutral strategies." (2002).

Alonso, Miquel Noguera. *Statistical Arbitrage and Algorithmic Trading: overview and applications*. Diss. Universidad Nacional de Educación a Distancia, 2010.

Andrade, S., Vadim Di Pietro, and M. Seasholes. "Understanding the profitability of pairs trading." Unpublished working paper, UC Berkeley, Northwestern University (2005).

Avellaneda, Marco, and Jeong-Hyun Lee. "Statistical arbitrage in the US equities market." *Quantitative Finance* 10.7 (2010): 761-782.

Broussard, John Paul, and Mika Vaihekoski. "Profitability of pairs trading strategy in an illiquid market with multiple share classes." *Journal of International Financial Markets, Institutions and Money* 22.5 (2012): 1188-1201.

Caldeira, João, and Guilherme V. Moura. "Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy." Available at SSRN 2196391 (2013).

Caldeira, João F., and Marcelo S. Portugal. "Estratégia long-short, neutra ao mercado, e index tracking baseadas em portfólios cointegrados." *Revista Brasileira de Finanças* 8.4 (2010): 469-504.

Dunis, Christian L., et al. "Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities." *Liverpool Business School, Working paper* (2010).

Engle, Robert F.; Granger, Clive W. J. (1987). "Co-integration and error correction: Representation, estimation and testing". *Econometrica*. 55 (2): 251–276. JSTOR 1913236.

Gatev, Evan, William N. Goetzmann, and K. Geert Rouwenhorst. "Pairs trading: Performance of a relative-value arbitrage rule." *Review of Financial Studies* 19.3 (2006): 797-827.

Jegadeesh, Narasimhan, and Sheridan Titman. "Overreaction, delayed reaction, and contrarian profits." *Review of Financial Studies* 8.4 (1995): 973-993.

Khandani, A., and A. Lo. "What Happened to the Quants in August 2007?(Digest Summary)." *Journal of investment management* 5.4 (2007): 29-78.

Krauss, Christopher. "Statistical arbitrage pairs trading strategies: Review and outlook." *Journal of Economic Surveys* (2016).

Lin, Yan-Xia, M. I. C. H. A. E. L. McCRAE, and Chandra Gulati. "Loss protection in pairs trading through minimum profit bounds: A cointegration approach." *Advances in Decision Sciences* 2006 (2006).

Lo, Andrew W., and A. Craig MacKinlay. "Stock market prices do not follow random walks: Evidence from a simple specification test." *Review of financial studies* 1.1 (1988): 41-66.

Lo, Andrew, and Amir Khandani. *What Happened to the Quants in August 2007?.* MIT Working Paper, 2008.

Miao, George J. "High Frequency and Dynamic Pairs Trading Based on Statistical Arbitrage Using a Two-Stage Correlation and Cointegration Approach." *International Journal of Economics and Finance* 6.3 (2014): p96.

Perlin, Marcelo Scherer. "Evaluation of pairs-trading strategy at the Brazilian financial market." *Journal of Derivatives & Hedge Funds* 15.2 (2009): 122-136.

Stock, James H., & Watson, Mark W. 1988. Testing for Common Trends. *Journal of the American Statistical Association*, 83, 1097–1107

Vidyamurthy, Ganapathy. *Pairs Trading: quantitative methods and analysis*. Vol. 217. John Wiley & Sons, 2004.