

PAIRS TRADING APLICADO AO MERCADO DE CAPITAIS BRASILEIRO: UMA ABORDAGEM VIA COINTEGRAÇÃO E ORNSTEIN-UHLENBECK

PAIRS TRADING APPLIED TO THE BRAZILIAN CAPITAL MARKETS: AN APPROACH VIA COINTEGRATION AND ORNSTEIN-UHLENBECK

DOI: [HTTP://DX.DOI.ORG/10.13059/RACEF.V15I1.1083](http://dx.doi.org/10.13059/RACEF.V15I1.1083)

Lucas Stumpf Venturini

lucas.stumpf@edu.pucrs.br

Pontifícia Universidade Católica do Rio Grande do Sul

Gustavo Inácio de Moraes

gustavoim@yahoo.com.br

Pontifícia Universidade Católica do Rio Grande do Sul

Data de envio do artigo: 09 de Fevereiro de 2023.

Data de aceite: 07 de Fevereiro de 2024.

Resumo: Este estudo é realizado explorando o fato de que a arbitragem estatística como estratégia de trade obteve grandes avanços na sua metodologia nos últimos 35 anos acompanhando a evolução econométrica. Como consequência, aplicou-se a estratégia de *pairs trading* por cointegração com um filtro de meia-vida, calculado via equação de Ornstein-Uhlenbeck, e através de diferentes testes de raiz unitária no mercado brasileiro de capitais com o objetivo de verificar se a metodologia apresenta lucratividade. Apesar de haver perdas em alguns pares de ações, verificou-se que a estratégia apresenta lucratividade..

Palavras-chave: Mercado de Capitais, *Pairs Trading*, Cointegração, Método de Ornstein-Uhlenbeck, Ibovespa.

Abstract: *This study is carried out by exploring the fact that statistical arbitrage as a trading strategy has made great advances in its methodology in the last 35 years, following the econometric evolution. As a consequence, the pairs trading strategy by cointegration was applied with a half-life filter, calculated via the Ornstein-Uhlenbeck equation, and through different unit root tests in the Brazilian capital market, with the objective of verifying whether the methodology shows profitability. Although there were losses in some stock pairs, it was found that the strategy is profitable.*

Keywords: *Capital Markets, Pairs trading, Cointegração, Ornstein-Uhlenbeck Procedure, Ibovespa.*

1 INTRODUÇÃO

Em meados de 1980 um pequeno grupo de matemáticos, físicos, cientistas da computação e economistas que trabalhavam na Morgan Staneley foram liderados por Nunzio Tartaglia e Garry Bamberger para construir estratégias quantitativas de trade. Dentre essas estratégias, eles foram pioneiros na formação de pares de ativos como forma de trade, essa estratégia ficou conhecida posteriormente como *pairs trading*

(VIDYAMURTHY, 2004). Logo após o pioneirismo do grupo pertencente à Morgan Staneley, alguns fundos de investimentos e bancos adotaram essa abordagem. *Pairs Trading* é a formação de pares com a finalidade de encontrar padrões e ineficiências dentro do mercado financeiro.

Dentro da literatura estatística e financeira a formação de pares é um tipo de arbitragem estatística. Isto significa que se escolhe dois ativos que possuem alguma semelhança no movimento de seus preços e após verifica-se o spread dos ativos, ou seja, algum pequeno desvio de preços no curto prazo de uma relação de equilíbrio de longo prazo entre dois ativos, constituindo ali uma maneira vantajosa de performar sob o mercado. Há diversos modelos que podem ser úteis para descrever a arbitragem estatística dos pares sendo importante destacar alguns como, por exemplo, o método da distância (GATEV; GOETZMANN; ROUWENHORST, 2006), o método por expoente de Hurst (RAMOS-REQUENA; TRINIDAD-SEGOVIA; SÁNCHEZ-GRANERO, 2017), a abordagem via cópulas (LIEW; WU, 2013), o método por cointegração (ENGLE; GRANGER, 1987) e, também, o método por state space models (ELLIOTT; VAN DER HOEK*; MALCOLM, 2005). São diferentes formas de estimações para formar o par e seu spread, porém a idéia que se constitui é a mesma para todos os modelos, isto é, utilizar o spread como forma de trade. É importante ressaltar que a estratégia de *pairs trading* possui um caráter importante no mundo das finanças, pois além de ser uma arma vantajosa de lucro, ela é, também, uma boa forma de hedge, sendo conhecida como, *market neutral strategy* (JACOBS; LEVY; STARER, 1999). Isto significa que quando posicionado num par, o agente estará seguro, pois dado que é uma estratégia de long-short (posição comprada em determinado ativo e vendida em outro ativo) e caso o mercado suba como um todo, irá se obter lucro com a ação comprada e ação vendida irá subir, o que gera uma perda pela parte vendida, porém a posição comprada irá compensar a perda do outro ativo. O contrário também é verdade, caso o mercado caia, a ação vendida irá gerar lucro e a ação comprada ocasionará numa perda, porém será compensada pela ação

vendida.

Uma forma de otimizar os modelos de reversão à média é aplicar o critério de meia-vida na série, isto é, o modelo de Ornstein-Uhlenbeck (UHLENBECK; ORNSTEIN, 1930). O modelo de Ornstein-Uhlenbeck, também conhecido como meia-vida, advém do cálculo estocástico que forma uma equação diferencial estocástica com a finalidade de descrever a velocidade que uma partícula browniana regride à média (SMITH, 2010), podendo ter as suas aplicações em estratégias de trade como visto em Chan (2013), na qual se pode verificar se uma série é reversível a média além de já calcular a meia-vida do processo, sendo uma importante ferramenta de encerramento de trade e de filtragem de pares.

Dado que estratégias de pairs trading já são bem estabelecidas na academia e no ambiente de trading, o objetivo do presente artigo irá verificar se o modelo de cointegração, formado por diferentes testes de raiz unitária, aplicado juntamente com um filtro de meia-vida via processo de Ornstein-Uhlenbeck oferece pares lucrativos. Para isso irá se utilizar algumas das ações com maiores volumes do índice Ibovespa para se testar os possíveis pares durante o período de abril de 2020 até setembro de 2021, sendo 80% como base de treino e o restante como período de teste.

Como estrutura, o artigo na sua seção 2, irá se revisar a literatura com a finalidade de apresentar alguns trabalhos que utilizam a cointegração e outros métodos como forma de arbitragem estatística; na seção 3 irá revisar tópicos de séries temporais apresentando a metodologia do trabalho e as estratégias combinadas como forma de trading; na seção 4 irá se apresentar os resultados dos pares com os seus retornos, retornos anualizados, volatilidade e o índice Sharpe juntamente com o índice Calmar como forma de performance dos pares; finalmente, na seção 5 observações conclusivas serão construídas.

2 FUNDAMENTAÇÃO TEÓRICA

Um dos passos primordiais para

o entendimento sobre o fenômeno de cointegração surgiu com o trabalho de Granger e Newbold (1974). O seu trabalho verificou por meio de simulações de Monte Carlo que, ao possuir uma ou mais séries não estacionárias, isto é, integradas de ordem um, $I(1)$, no modelo de regressão, os resultados parecerão significativos, o que leva ao problema de regressão espúria. Em exemplos de casos onde a regressão espúria é verificada as estatísticas usualmente utilizadas não são mais úteis. Tornando-se necessário novas técnicas para desenvolver análises robustas sobre séries não estacionárias. O caso de regressão espúria posteriormente foi estudado, também, por Phillips (1986) utilizando teoria assintótica de regressão sobre relações de séries aleatórias, os seus resultados explicam algumas das descobertas de Granger e Newbold.

Visto que séries não estacionárias no modelo de regressão podem produzir resultados significativos, o que pode levar ao caso de regressão espúria, (ENGLE; GRANGER, 1987) introduziram a metodologia de cointegração conectando alguns conceitos como médias móveis, termo autoregressivo e correção de erros para sistemas cointegrados, expandindo o trabalho de Granger (1983). No seu trabalho, Engle e Granger desenvolveram um teste de cointegração para analisar se há uma relação entre as séries $I(1)$. Este teste consistia em formar os resíduos da cointegração e verificar se os resíduos são, ou não estacionários, para isso se pode utilizar alguns testes de raiz unitária como o Augmented Dickey-Fuller test (DICKEY; FULLER, 1979) e Phillips-Perron test (PHILLIPS; PERRON, 1988). O trabalho de Engle e Granger supõem que há, no máximo, apenas um vetor de cointegração, já (JOHANSEN, 1988) supõem que há t vetores de cointegração, sendo $0 \leq t < n$. O teste desenvolvido de (JOHANSEN, 1988) é uma expansão do teste de Engle-Granger para o caso multivariado de séries não estacionárias.

Após o desenvolvimento do modelo de cointegração diversos trabalhos analisaram uma relação de equilíbrio de longo prazo entre variáveis econômicas, porém grande parte dos trabalhos eram mais teóricos e possuíam pouca

aplicação no mercado financeiro, conforme visto em Alexander (1999). Um dos trabalhos pioneiros de cointegração no preço de ações foi o trabalho de Cerchi e Havenner (1988). Já para a estratégia de arbitragem estatística, isto é, utilizando o spread como forma de trade foi o trabalho de Girma e Paulson (1999), na qual estudou-se os contratos futuros de petróleo. O trabalho de Herlemont (2003) realiza uma revisão sobre estratégias de trade com cointegração além de apresentar o modelo de Ornstein-Uhlenbeck como processo de stop-loss para operação.

Alexander; Giblin e Weddington (2002) trabalhou uma estratégia para construção de portfólio utilizando métodos de pairs trading como, por exemplo, a cointegração e o método por correlação dos retornos. Avellaneda e Lee (2010) estudou a arbitragem estatística no mercado norte-americano, o seu estudo contribui mostrando que diferentes conjuntos de fatores de risco levam a diferentes resíduos, o que por sua vez leva a diferentes retornos de portfólio, os fatores de riscos utilizados foram o *Principal Component Analysis*, e o método que utiliza ETFs do setor de indústria como fator de risco.

Já Caldeira e Moura (2013) estudaram a relação de pairs trading no mercado brasileiro durante o período de janeiro de 2005 a outubro de 2012 sendo um ano como período de treino e quatro meses como período de teste. Teixeira (2014) utilizou como metodologia na aplicação de pairs trading a abordagem por cointegração juntamente com o processo de Ornstein-Uhlenbeck e o critério de Kelly, os resultados encontrados performaram o mercado.

Hong e Susmel (2003) foram alguns dos pioneiros a implementar o modelo de cointegração no mercado de ações, no seu trabalho aplicou-se o modelo de cointegração no mercado de ADRs da Ásia, os resultados encontrados foram significativos. Lin; McRae e Gulati (2006) expandiu a estratégia de cointegração desenvolvendo um procedimento que embuti uma condição de lucro mínimo. Taylor e Tonks (1989) e também Mylonidis e Kollias (2010) estudaram o mercado de ações

na europa e encontraram fortes evidências de cointegração. Sipila et alii (2013) teve como objetivo verificar se a abordagem de cointegração para pairs trading possuía retornos significativo em ETFs, sua pesquisa confirmou tal hipótese.

Alguns livros já foram escritos voltados à arbitragem estatística como, por exemplo, (Vlidy- Murthy (2004), Eherman (2006) e Whistler (2004). Todos os livros possuem uma excelente revisão de conteúdo e excelentes exemplos nas aplicações. Confirmando a abordagem como lucrativa. Além de se utilizar o modelo de cointegração como forma de trade, pode-se, também construir uma abordagem de otimização de portfólio na alocação das operações conforme visto em Chiu e Wong (2011). A arbitragem estatística via séries cointegradas, também pode ser usada em dados de alta frequência (MIAO, 2014).

Como há diferentes modelos para lidar com a arbitragem estatística entre duas séries, alguns trabalhos lidaram a seguinte dúvida: o modelo de cointegração provém mais lucros que os outros métodos? Caldas; Caldeira e Mourab (2016) encontraram evidências que o método de cointegração é superior ao método da distância. Gatev; Goetzmann e Rouwenhorst (2006) nos mercados europeus (período de análise de 1973 a 2012) e brasileiro (período de análise de 1995 a 2012), porém nos EUA (período de análise de 1962 a 2012) o método da distância performou melhor que o método por cointegração.

Huck e Afawubo (2015) utilizaram três metodologias para verificar qual é a mais lucrativa, dentre elas:

Método da distância: Neste modelo se escolhe para cada ação o seu par que minimiza a soma dos quadrados da diferença nos preços diários normalizados. Quando a diferença entre os pares aumenta se entra numa posição comprada em um ativo e vendido no outro.

- Método da distância: Neste modelo se escolhe para cada ação o seu par que minimiza a soma dos quadrados da diferença nos preços

diários normalizados. Quando a diferença entre os pares aumenta se entra numa posição comprada em um ativo e vendido no outro.

- A estacionariedade da relação de preços: Neste método após escolher duas ações para especular, o agente toma a razão entre os ativos e realiza um teste de estacionariedade no resultado da razão. Se confirmado a hipótese de estacionariedade na razão de preços pode se utilizar o par para trade.

- A cointegração entre as ações.

Foram escolhidos os preços das ações que compõem o S&P 500 durante o período de Agosto de 2000 a Setembro de 2011 sendo de um a dois anos como período de treino e o restante como período de teste. O modelo de cointegração performou melhor do que os outros métodos.

Clegg (2014) realizou um estudo sobre a persistência da cointegração para pairs trading utilizando ativos do EUA durante o período de 2002 a 2012. No seu estudo foi comparado a persistência da cointegração para os pares de ativos cointegrados em um ano, o seu método se constituiu em realizar diferentes abordagens para verificar a estacionariedade do spread de cada par, utilizando testes como, por exemplo, o de Dickey-Fuller, Phillips-Perron, teste de Johansen com autovalores (JOHANSEN, 1988), teste Johansen de traço (JOHANSEN, 1988), Elliot-Rothenberg- Stock Feasible point optimal test (ELLIOT; ROTHENBERG; STOCK, 1996), Elliot-Rothenberg- Stock Detrended ADF-type test (ELLIOT; ROTHENBERG; STOCK, 1996), expoente de Hurst (TAQQU; TEVEROVSKY; WILLINGER, 1995), Breitung's Variance ratio (BREITUNG, 2002) e Weighted Symmetric Estimator (PANTULA; GONZALEZ-FARIAS; FULLER, 1994). Concluiu-se que, no mercado americano, um par cointegrado durante um ano possui grandes chances de perder a sua relação de longo prazo no próximo ano.

No mercado americano, mais especificamente no SP500, há demonstrações que os resultados permanecem quando da presença de operações de alta frequência, como demonstrado por Enders e Stubinger (2019). Os

retornos obtidos a partir das estratégias testadas foram expressivos, mesmo após deduções de custos de operação, e estáveis ao longo do tempo.

Em um contexto de mercado europeu, Virtaren (2020) aplica a metodologia de Ornstein-Uhlenbeck ao mercado acionário norueguês. A novidade está na aplicação de técnicas de redes neurais, para mais de dois mil e novecentos pares, mas em outro sentido o autor destaca a limitação de utilizar dados diários, sobretudo com os dados de fechamento e abertura para a obtenção dos sinais. Nesse sentido, o presente artigo compartilha a limitação desta pesquisa.

Testes de estratégias com custos de operações proporcionas aos volumes transacionados foram testados por Xing (2022) para seis pares de ativos americanos. Demonstra-se neste contexto que os resultados permanecem consistentes com estratégias lucrativas.

Multiplicando as aplicações, os mercados asiáticos, no caso o chinês, também foram examinados pela técnica. Luo, Lin e Wang (2022) focaram sua análise nos mercados futuros de energia em um período de 23 meses, obtendo um retorno expressivo em condições medianas de mercado. Contudo, constataram que em condições qualificadas como adversas de mercado a técnica não foi capaz de trazer resultados lucrativos.

2.1 Conceitos de Séries Temporais

Para um entendimento mais claro do fenômeno de cointegração, sendo a primeira subseção para o desenvolvimento do fenômeno da estacionariedade e o passeio aleatório, a segunda subseção para o desenvolvimento dos testes de raízes unitárias que serão utilizados, a terceira subseção para o desenvolvimento do modelo de cointegração,

2.1.1 Processos lineares estacionários

Para um entendimento maior do fenômeno

da estacionariedade iremos, primeiramente, desenvolver a definição de sequência aleatória num espaço amostral. Considerando uma sequência de variáveis aleatórias X_n no mesmo espaço amostral Ω , isto é, $(X_n, n = 1, 2, \dots)$. Para todo $n \geq 1$ se tem:

$$(X_1 = a_1, \dots, X_n = a_n) = P(X_1 = a_1) \times P(X_2 = a_2 | X_1 = a_1) \times \dots \times P(X_n = a_n | X_1 = a_1, \dots, X_{n-1} = a_{n-1}) \quad (1)$$

Com α_j representando os estados do processo e o espaço amostral de cada processo é tomado como conjunto dos reais. Uma simplificação importante da equação 1 é quando se possui uma sequência $(X_n, n \geq 1)$ com variáveis aleatórias mutuamente independentes, podendo ser representado como:

$$P(X_1 = a_1, \dots, X_n = a_n) = P(X_1 = a_1) \dots P(X_n = a_n) \quad (2)$$

Logo, se todas as variáveis aleatórias possuírem a mesma distribuição se tem que estas variáveis aleatórias podem ser chamadas de variáveis independentes e identicamente distribuídas. Possuindo-se, então, que o processo X_n é estacionário. Caso o valor esperado do processo X_n , isto é, $E(X_n) = \mu$ e a variância do processo X_n , isto é, $Var(X_n) = \sigma^2$, para todo $n \geq 1$ se tem:

$$\begin{aligned} Cov\{X_n, X_{n+\tau}\} &= \sigma^2, \text{ se } \tau = 0 \\ &0, \text{ se } \tau \neq 0 \end{aligned} \quad (3)$$

Quando se possui correlação, isto é, $\rho_t = 1, t = 0$ e quando $\rho_t = 0$ se tem $t \neq 0$. Neste caso, o processo de variáveis aleatórias passa a ser chamado de processo puramente aleatório ou ruído branco forte.

Portanto, chega-se a conclusão que um processo estocástico X_n é estacionário de segunda ordem se a sua média é constante e sua variância também é constante. Logo,

$$\begin{aligned} E\{X_t\} &= \mu_t \text{ e} \\ E\{X_{t-1}\} &= \mu_t, \text{ portanto} \\ E\{X_t\} &= E\{X_{t-1}\}, \\ E\{X^2\} &= E\{X^2_{t-1}\} \end{aligned} \quad (4)$$

2.1.2 Processos Não Estacionários

Em análise de séries temporais um processo muito importante é o passeio aleatório. Dentro da literatura econométrica financeira este conceito possui uma suma importância, pois como demonstra alguns pesquisadores os preços dos ativos seguem um passeio aleatório, e essa verificação pode levar a conclusão da hipótese de mercados eficientes, pois dado que os preços se movem aleatoriamente é difícil prever exatamente onde o preço estará no futuro. Um exemplo de passeio aleatório segue-se abaixo, considerando uma sequência aleatória, isto é, $\{Y_t, t \geq 1\}$ de variáveis aleatórias, tem-se a sequência

$$X_t = Y_1 + \dots + Y_t \quad (5)$$

A seguir o valor esperado de X_t depende de uma constante na média das variáveis aleatórias, isto é, $E(X_t) = t\mu$ e variância depende também de t , isto é, $Var(X_t) = t\sigma^2$. Nota-se que com o passar do tempo, o processo X_t tenderá ficar próximo de $t\mu$, numa amplitude não estável. Conclui-se que o

processo não possui média constante, portanto não é estacionário. Grande parte dos ativos não são estacionários, pois a média dos preços está sempre se alterando com o passar do tempo, e o mesmo se observa com a variância.

2.1.3 Testes de Raíz Unitária

Verificado a definição de estacionariedade e passeio aleatório, torna-se necessário uma formalização de testes para verificar se uma série é ou não estacionária. Esses testes são conhecidos na literatura como testes de raíz unitária. No presente trabalho irá se trabalhar com quatro testes de raíz unitária, sendo eles: teste de Dickey-Fuller aumentado, teste de Phillips-Perron, teste de Pantula, Gonzales- Farias e Fuller, mais conhecido como teste PGFF.

2.1.4 Teste de Dickey-Fuller Aumentado (ADF)

Um dos testes mais famosos para verificação de estacionariedade em séries temporais é o teste Aumentado de Dickey-Fuller. Introduzido em 1979, o teste de Dickey-Fuller verifica se uma série é estacionária ou não. Para formalizar o teste suponha que X_t siga um processo AR(1):

$$X_t = \phi X_{t-1} + v_t, \text{ com } \{v_t\} \text{ sendo o ruído branco com distribuição } (0, \sigma^2) \quad (6)$$

Então se testa a seguinte hipótese:

$$\begin{aligned} H_0 &: \phi = 1 \text{ (raíz unitária, não é estacionária)} \\ H_1 &: \phi < 1 \text{ (é estacionária)} \end{aligned} \quad (7)$$

Onde $\hat{\phi}$ é estimado via mínimos quadrados e $se(\hat{\phi})$ é a estimação dos mínimos quadrados do erro padrão de $\hat{\phi}$. Logo:

- Se $|\phi| = 1$, então $\sqrt{T}(\hat{\phi} - \phi) \xrightarrow{d} N(0, (1 - \phi^2))$
- Ou, se $\phi < 1$, então $\hat{\phi}$ é superconsistente com taxa de $(1/T)$. T possui a distribuição de Dickey-Fuller. $\phi=1$

2.1.5 Teste de Phillips-Perron (PP)

O teste de Phillips-Perron foi desenvolvido posteriormente ao Dickey-Fuller. A sua diferença advém do fato que o teste PP corrige o problema de correlação serial e heterocedasticidade após as estimações via MQO. Portanto se tem a equação:

$$X_t = \alpha + \phi X_{t-1} + \epsilon_t \quad (8)$$

Da qual utiliza-se um teste não paramétrico para corrigir a presença de correlação serial e heterocedasticidade.

A hipótese do teste é a mesma do Dickey-Fuller, isto é:

$$\begin{aligned} H_0 &: \phi = 1 \text{ (raíz unitária, não é estacionária)} \\ H_1 &: \phi < 1 \text{ (é estacionária)} \end{aligned} \quad (9)$$

2.1.6 Teste de Pantula, Gonzales-Farias E Fuller (PGFF)

O teste PGFF introduzido em 1996 consistiu em estimar um parâmetro simétrico ponderado, ρ_{ws} , para verificar se a série apresenta raíza unitária. O formato da equação é:

$$Y_t = \mu(1 - \rho) + \rho Y_{t-1} + \epsilon_t, t \geq 2 \quad (10)$$

O teste de hipóteses é semelhante com os já mencionados:

$$\begin{aligned} H_0 : \rho &= 1 \text{ (raíz unitária, não é estacionária)} \\ H_1 : \rho &< 1 \text{ (é estacionária)} \end{aligned} \quad (11)$$

2.2 Cointegração

No presente artigo irá se utilizar o procedimento de Engle e Granger para verificar se duas séries são cointegradas. O procedimento para verificar se duas séries são cointegradas é dado a seguir.

Suponha duas séries integradas de ordem um:

$$\begin{aligned} X_{1t} &\sim I(1) \\ X_{2t} &\sim I(1) \end{aligned} \quad (12)$$

No passo 2, realiza-se uma regressão via MQO para encontrar os parâmetros μ e γ :

$$X_{1t} = \mu + \gamma X_{2t} + \epsilon_t \quad (13)$$

No passo 3, verifica-se os resíduos da regressão. Espera-se que a sua combinação linear resulte num processo estacionário, portanto se utiliza algum teste de raiz unitária para verificar se é estacionário:

$$\begin{aligned} H_0 : \epsilon_t &\sim I(1); P \text{ ar não cointegrado} \\ H_1 : \epsilon_t &\sim I(0); P \text{ ar cointegrado} \end{aligned} \quad (14)$$

Para aplicar esta formulação em *pairs trading* utiliza-se os procedimentos abaixo:

Suponha duas ações não estacionárias

$$\begin{aligned} \log(ação_{1t}) &= \gamma x_t + \omega_1 t \\ \log(ação_{2t}) &= x_t + \omega_2 t \end{aligned} \quad (15)$$

sendo x_t uma tendência estocástica em comum definida como passeio aleatório:

$$x_t = x_{t-1} + \omega_t \quad (16)$$

Na qual os parâmetros ω_1 , ω_2 e ω_t são os resíduos i.i.d.

O coeficiente γ é estimado via MQO formando o spread da série:

$$\log(ação_{1t}) = \mu + \gamma \log(ação_{2t}) + \epsilon_t, \quad (17)$$

$$spread = \log(ação_1t) - \gamma \log(ação_2t) = \mu + e_t \quad (18)$$

Como foi visto, os resíduos de um modelo cointegrado são estacionários, logo o spread também deverá ser estacionário.

Observa-se que o Spread gira em torno da média. Usando-se o teste de raiz unitária verifica-se que o spread é estacionário.

3 MATERIAIS E MÉTODOS

3.1 Processo de Ornstein-Uhlenbeck.

O processo de Ornstein-Uhlenbeck é um processo estocástico que advém da física mais precisamente do campo de equações diferenciais estocásticas. O modelo é estacionário e contínuo em probabilidade, além disso é um processo que descreve algumas características de procedimentos que se distanciam da média, ou seja, processos de reversão à média. Em pairs trading a aplicação se dá devido ao cálculo de velocidade do parâmetro de reversão, isto é, a velocidade de reversão à média, sendo conhecido muitas vezes como o cálculo da meias-vida.

Considerando a variável X_t como sendo o spread do par de ações, sendo ele estacionário, a equação de Ornstein-Uhlenbeck pode ser escrita como

$$dX_t = \lambda(\mu - X_t)dt + dW \quad (19)$$

Onde:

- λ é a velocidade de reversão,
- μ é a média de longo prazo dos preços,
- dW é um ruído branco.

Integra-se, então, do tempo t_0 até o instante t_1 , da qual é o preço esperado em determinado instante sendo representado por X_{t+1} .

A estimação do parâmetro λ pode ser realizada através de uma regressão linear ou pelo processo de máxima verossimilhança (SMITH, 2010). Um exemplo da interpretação do resultado do meia-vida é verificar o valor do resultado e utilizá-lo como dias, por exemplo, se $H = 3,89$ se considera que o processo demora 4 dias para retornar à média.

Como destacado por Glover e Peskir (2023) o resultado de reversão à média pode ser identificado no tempo, através da possibilidade de se diagnosticar um resultado falso-positivo precoce. Assim, bastaria identificar em subamostras que os sinais da constante e os sinais do coeficiente da variável como sendo significativos e em sentidos opostos.

3.2 Estratégia de Trade.

Conforme já visto o processo de estimação para cointegração, utilizará esta subseção para explorar como irá se utilizar a metodologia para trade.

Os procedimentos realizados até agora são:

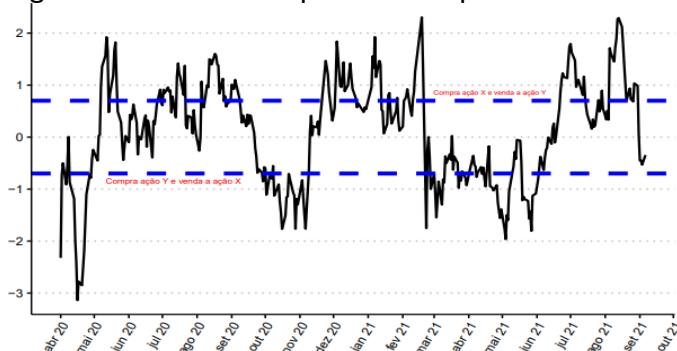
- Escolher duas ações e definir qual é a variável independente e qual é a variável dependente.

- Realizar a regressão linear encontrando os parâmetros μ e γ .
- Se constrói o spread e se testa a estacionariedade da série.
- Confirmada a hipótese de estacionariedade, utiliza-se o spread para fazer as operações.

Acrescenta-se, agora, um passo posterior aos já mencionados, após a confirmação de estacionariedade no spread se irá transformar o mesmo, utilizando o método de normalização Z-Score.

Após realizado a transformação, realiza-se os pontos de entradas e os pontos de saída do trade. Estes sinais determinarão a ordem das operações nos pares. Quando o spread cruza o sinal de entrada com sinal negativo se deve entrar long na ação dependente, Y_t , e entrar short na ação independente, X_t , encerrando-se as operações quando ocorre o retorno à média. Porém quando spread cruza o sinal positivo de entrada, entra-se short na ação dependente, Y_t , e se entra long na ação independente, X_t . No presente trabalho irá se utilizar o sinal de entrada, após a transformação Z-Score, como sendo 0,7. Escolha-se esse valor difentemente dos outros trabalhos, como, por exemplo, Caldeira e Moura (2013) que escolheram o valor de 2 como entrada, pois decide-se verificar mais operações. A figura 1 mostra o spread da série juntamente com os sinais de entrada.

Figura 1 - Sinal de compra e venda por ativo



Fonte: Elaboração Própria.

A proporção das posições compradas e vendidas é dado pelo coeficiente γ . Ou seja, quando $\gamma = 0,4$ irá entrar vendido numa posição de 40% a menos ações. Por exemplo, no par usado de exemplo até agora, PETR4-CMIG4, o coeficiente γ estimado é 0,81, portanto cada vez que se entrar long comprando, por exemplo, 1000 ações deverá entrar na posição short com 810 ações.

Logo, pode-se construir o retorno do portfólio sendo:

$$\text{Retornos} = \ln\left(\frac{P_t^l}{P_{t-1}^l}\right) - \gamma \ln\left(\frac{P_t^s}{P_{t-1}^s}\right) \quad (20)$$

Onde:

- P_t^l - preço no tempo t da posição comprada;
- P_{t-1}^l - preço no tempo t-1 da posição comprada;
- P_t^s - preço no tempo t da posição vendida;
- P_{t-1}^s - preço no tempo t-1 da posição vendida;

Yan, Chiu e Wong (2022) identificaram o potencial de se utilizar um mecanismo de correção de erros para aprimorar o modelo, independentemente da ordem da defesagem, capturando assim efeitos defasados na ordem de cointegração. Assim, modelos de pairs trading podem ser aprimorados, e manterem-se lucrativos, em condições de captura de processos de memória das séries.

3.3 Métricas de performance

As métricas utilizadas para avaliar a performance das operações seguem a seguir.

3.3.1 Índice Sharpe Anualizado

Considerando o trabalho seminal de (SHARPE, 1994) irá se utilizar a sua métrica para verificar a razão do retorno e da volatilidade de cada operação, a interpretação simplória deste índice é de quanto maior o valor, melhor a performance. O índice Sharpe anualizado é calculado a partir de:

$$\text{Índice Sharpe Anualizado} = \frac{\sqrt{(n)} (1 + Ra)^{escala} - 1}{\sqrt{escala} \sqrt{\sigma}} \quad (21)$$

Onde:

- R_p é o retorno de cada operação e
- R_f é a taxa livre de risco.

3.4 Amostragem

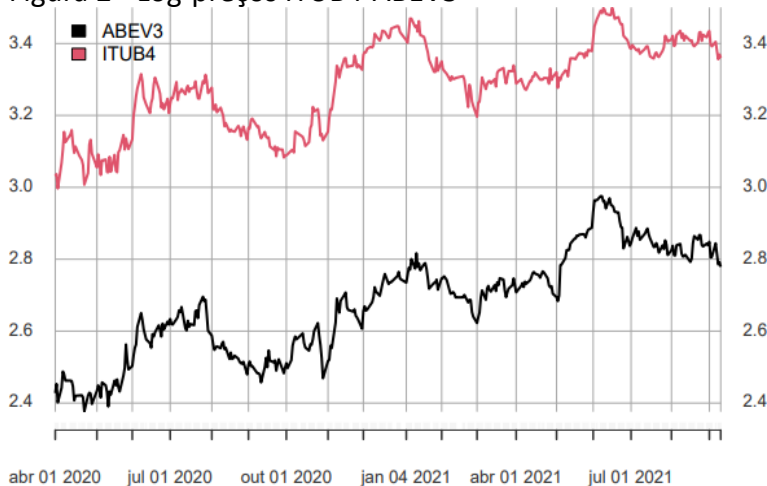
Foram escolhidas 43 ações mais o ETF da bolsa brasileira conhecido como BOVA11 com periodicidade diária. As 43 ações escolhidas pertencem ao índice B3, foi escolhido tais ações durante o período de 2020/04 a 2021/09 devido o seu grau de negociações do período possuindo maiores volumes se comparados as Small Caps. Para a realização de trades se relaxou a hipótese de que cada par deveria pertencer ao mesmo setor, portanto agora se pode ter pares de setores completamente diferentes. Isso significa que o mais importante é apenas encontrar pares com significância estatística, em outras palavras, deixar os dados falarem por si. Além desses fatores se decidiu trabalhar apenas com as ações que possuem um γ maior que zero e o μ maior que zero, também. Essas escolhas foram feitas, pois quando se possui um γ negativo se trabalha com operações na mesma direção, isto é, long-long ou short-short. Além disso, adiciona-se um filtro de meia-vida nos pares, ou seja, escolhe-se os pares que possuem um retorno à média de no máximo 15 dias. A base de dados foi extraído pelo pacote quantmod (RYAN et al., 2020) do software R, da qual busca os dados pela API do Yahoo Finance (DATA, 2012).

4 RESULTADOS E DISCUSSÃO

Antes de já apresentar os resultados de todos os pares, primeiro irá se exemplificar de forma prática todos os procedimentos desenvolvidos até agora. No exemplo se utilizará o par ITUB4-ABEV3.

Primeiramente irá se verificar o gráfico do log preços.

Figura 2 - Log-preços ITUB4-ABEV3



Fonte: Elaboração Própria a partir de dados do Yahoo Finance

Visualmente as ações parecem caminhar juntas indicando uma possível cointegração entre os ativos. Lembrando que para verificar uma relação de cointegração no spread se deve, primeiramente, encontrar os coeficientes μ e γ através da regressão linear. Conforme já estabelecido se utilizou 80% do período como in-sample e o restante como out-of-sample. A equação da regressão linear é:

$$\log(ITUB4) = \mu + \gamma \log(ABEV3) + \epsilon_t \quad (23)$$

Os coeficientes estimados da equação 23 se encontram na Tabela 1.

Tabela 1 - Coeficientes estimados do par

Variável	Valor do Coeficiente (Desvio Padrão)
ABEV3	0,815 *** (0,024)
Constante	1,108*** (0,064)
Observações	286
R ² Ajustado	0,799
Erro Residual Padrão	0,049
Estatística F	1.134,07

Fonte: Resultados da Pesquisa

Nota:*p<0.1; **p<0.05; ***p<0.01

Logo o spread possui o formato de:

$$Spread = \log(ITUB4) - 0,815\log(ABEV3) \quad (24)$$

Um passo importante na análise é verificar o gráfico da ação independente juntamente com o formato do spread com o coeficiente μ_{xx} , a figura 3 demonstra a evolução de ambos. A interpretação deste gráfico é verificar quando ambos se distanciam se tendo observações de quando ocorre as operações:

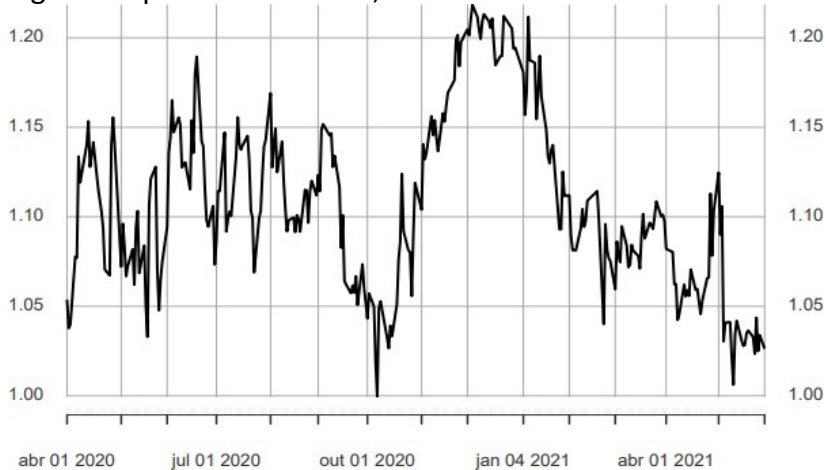
Figura 3 - Log-preço ITUB4 e equação da regressão linear



Fonte: Elaboração Própria

O próximo passo é verificar o gráfico do spread da qual se encontra na figura 4.

Figura 4: Spread = ITUB4 – 0,815*ABEV3



Fonte: Elaboração Própria.

Visualmente o spread aparenta ser estacionário, portanto se deve testar estatisticamente a estacionariedade do mesmo. Os resultados se encontram na Tabela 2.

Tabela 2 - Resultados dos testes de raiz unitária

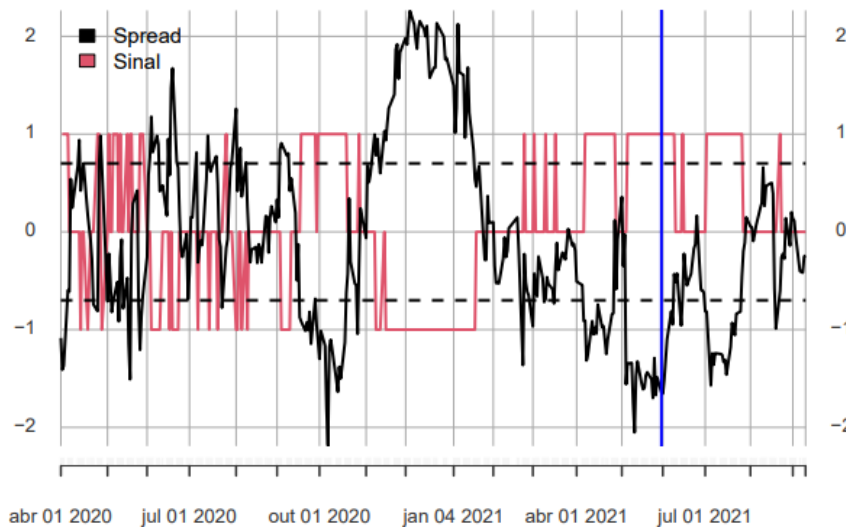
Nome do teste	Estatística do teste	P-Valor	Hipótese Alternativa
Augmented Dickey-Fuller Test	-1,98	0,58	estacionário
Phillips-Perron Unit Root Test	-20,54	0,06	estacionário
Pantula, Gonzales-Farias and Fuller Unit Root Test	0,91	0,01	estacionário

Fonte: Resultados da Pesquisa

Observa-se que pelos testes de raiz unitária ADF e PP se aceita a hipótese nula de não estacionariedade, porém pelo teste PGFF se rejeita a hipótese nula aceitando a hipótese alternativa de estacionariedade.

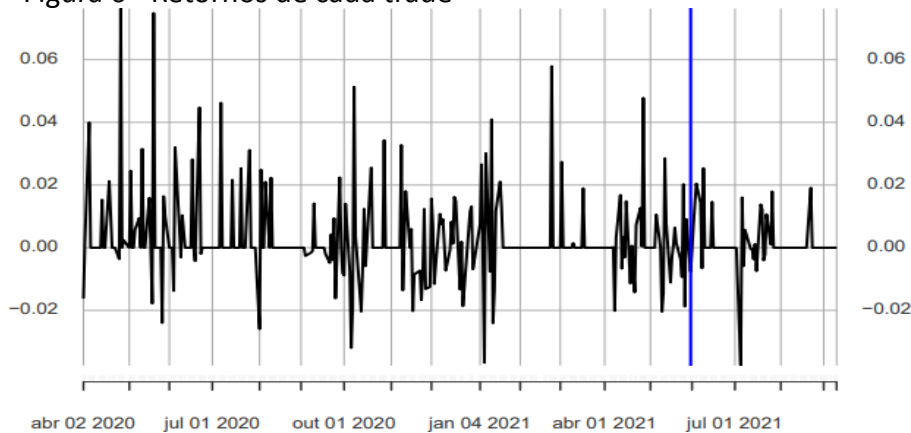
Conforme foi citado na revisão de literatura essa é uma divergência bem comum em pairs trading, neste caso irá se aceitar a hipótese de que o spread é estacionário dado o teste PGFF. Portanto, aceitado a hipótese de estacionariedade se realiza a transformação z-score. Como pode se verificar o spread possui média zero e a cointegração parece persistir para o período de teste. Os sinais de entrada e saída podem ser verificados na figura 5 sendo representados pela linha vermelha. Quando a linha vermelha recebe o valor de +1 significa que se está entrando comprado na ação ITUB4 e vendido na ABEV3, intuitivamente se percebe quando a linha vermelha recebe o valor de -1 significa que se está entrando comprado em ITUB4 e vendido em ABEV3.

Figura 5: Sinais das operações de trade



O número total de sinais com posições comprada em ITUB4 e vendida em ABEV3 foi de 98 operações. Já o número total de sinais com posições vendidas em ITUB4 e comprada em ABEV3 foi de 68 operações. O retorno de cada operação é encontrado na Figura 6.

Figura 6 - Retornos de cada trade



Fonte: Elaboração Própria, 2021.

Para apresentar o P&L (lucro e perda) acumulado é necessário revisar a idéia de reinvestimento das operações:

- Sem reinvestimento: Neste tipo de operação o orçamento inicial é sempre reinvestido. Por exemplo, o orçamento inicial é de 1 real, em todo trade este capital inicial será reinvestido, mesmo se o orçamento em algum momento for de 20 reais.
- Com reinvestimento: Neste tipo de operação todo o capital existente é reinvestido, ganhando a ideia de ser uma operação composta.

A quantidade de lucro acumulado no período de treino foi significativa mesmo sem reinvestimento, o que implica que o teste de raiz unitária PGFF já proveu um par lucrativo ao contrário dos outros testes de raiz unitária.

Observa-se que durante o período de treino o capital acumulado foi expressivo, juntamente o período de teste obteve um bom resultado pelo pouco período.

A Tabela 3 apresenta as métricas de comparação do período de treino e teste. Como pode se observar o índice Sharpe anualizado foi significativo em ambos os períodos.

Tabela 3 - Métricas de performance

Período Treino		Período Teste
Annualized Return	120%	60%
Annualized Std Dev	0,23	0,13
Annualized Sharpe (Rf=0%)	5,26	4,52
Calmar Ratio	16,70	15,88

Fonte: Resultados da Pesquisa

O cálculo da meia-vida foi aplicado ao spread durante o período de treino e foi encontrado uma meia-vida de 7,58 dias.

A partir da seção 4.1, a seguir, inicia-se a descrição dos métodos de tratamento da amostragem.

4.1 Total de pares por teste de raiz unitária

A Tabela 4 mostra o número total de pares por cada teste juntamente com a quantidade de pares com p-valor menor que 0,01.

Tabela 4: -Quantidade de Pares Cointegradas por teste

Teste	Quantidade de Pares	P-Valor < 0.01
ADF Teste	31	10
PP Teste	108	38
PGFF Teste	137	55

Fonte: Resultados da Pesquisa

A Tabela 5 mostra a quantidade de pares em comum pelos testes. Percebe-se que o teste de Phillips-Perron e Pantula, Gonzales-Farias e Fuller teste possuem muitos pares em comum, enquanto que o teste ADF e PGFF possuem uma quantidade de pares cointegrados pequena em comum. O total

de pares que são cointegrados por todos os testes é de apenas 28 pares.

Tabela 5 - Pares em comum por teste

Cointegrado nos 3 testes	Cointegrado pelo ADF e PP	Cointegrado pelo ADF e PGFF	Cointegrado pelo PP e PGFF
28	30	28	100

Fonte: Resultados da Pesquisa

Yu et alii (2023) também apontam a possibilidade de soluções estáticas e dinâmicas apontarem para o mesmo resultado, ou seja, resultados consistentes a partir da técnica de Ornstein-Uhlenbeck e também do processo Média-Variância. Assim, há uma convergência de resultados, mesmo quando um problema de inconsistência temporal se apresenta.

4.2 Retornos nos períodos in-sample e out-of-sample

Os retornos acumulados de todos os pares se encontram na tabela abaixo juntamente com as métricas de performance acumuladas:

Tabela 6 - Retornos Acumulados e Métricas Acumuladas

	Retorno Treino	Retorno Teste	Sharpe Treino	Sharpe Teste	Calmar Treino	Calmar Teste
ADF	3,142%	74%	108,47	9,67	294,68	115,14
PP	10,914%	-16%	353,82	8,65	913,65	189,84
PGFF	13,684%	-18%	443,24	20,88	1117,60	286,94

Fonte: Resultados da Pesquisa

A tabela a seguir apresenta as médias dos retornos anuais, índice Sharpe e índice Calmar tanto no período de treino quanto no período de teste.

Tabela 7 - Retornos Médio Anual e Métricas de Performance

	Retorno Treino	Retorno Teste	Sharpe Treino	Sharpe Teste	Calmar Treino	Calmar Teste
ADF	101,3%	2,39%	3,499	0,312	9,506	3,714
PP	101,1%	-0,148%	3,276	0,080	8,460	1,758
PGFF	99,88%	-0,13%	3,235	0,152	8.158	2,094

Fonte: Resultados da Pesquisa

Verifica-se que durante o período de treino o teste ADF possuiu um retorno médio anual maior do que os outros testes, além de possuir uma melhor performance pelo Índice Sharpe e o Índice Calmar. A média dos retornos anualizados inferem que os pares cointegrados pelo teste de Dickey-Fuller possuem um retorno médio maior que os outros testes de raiz unitária além de possuir uma

melhor performance no período out-of-sample. É importante verificar como foi o retorno máximo dos pares cointegrados por cada teste no período out-of-sample, a tabela abaixo mostra o maior retorno anualizado do par de cada teste:

Tabela 8: Par com maior retorno por teste de raiz unitária

Teste	Par	Retorno Anualizado	Índice Sharpe	Índice Calmar
ADF	HYPE3.SA-CVCB3.SA	67%	1,48	2,35
PP	PETR4.SA-CMIG4.SA	117%	5,24	16,41
PGFF	WEGE3.SA-LCAM3.SA	122%	3,92	13,73

Fonte: Resultados da Pesquisa

O par com maior retorno pertence a classe de pares cointegrados via teste PGFF. Outro fator importante para verificar é a quantidade de pares com retornos positivos no período out-of-sample e com retornos negativos no período out-of-sample, o resultado se encontra na tabela abaixo:

Tabela 9 - Quantidade de pares com retornos positivos e negativos

Teste	Retornos Anualizados % positivos do total de pares	Retornos Anualizados % negativos do total de pares
Dickey-Fuller	17 / 54,8%	14 / 45,2%
Phillips-Perron	47 / 43,5%	61 / 56,5%
PGFF	60 / 43,8%	77 / 56,2%

Fonte: Resultados da Pesquisa

Verificado a quantidade de pares com retornos positivos e negativos, realiza-se uma média dos retornos anualizados, índice Sharpe e índice Calmar dos pares com retornos positivos.

Tabela 10 - Apenas Retornos no período de teste com valores positivos

	Retorno Treino	Retorno Teste	Sharpe Treino	Sharpe Teste	Calmar Treino	Calmar Teste
ADF	109%	27%	3,71	1,60	9,77	8,19
PP	114%	31%	3,59	1,57	9,44	5,86
PGFF	109%	35%	3,47	1,77	8,87	6,63

Fonte: Resultados da Pesquisa

Na tabela acima verifica-se a média dos retornos anualizados, do índice Sharpe anualizado

e do índice Calmar, somente com valores positivos no período de treino. Observa-se que os pares cointegrados pelo teste PGFF apresentaram um retorno anualizado médio maior dos que os outros testes, além de apresentar um maior índice Sharpe anualizado, porém sem possuir o maior índice Calmar.

Como realizou-se um teste de performance apenas com pares que apresentaram retornos positivos no período de teste, irá se realizar o mesmo para os pares com retornos negativos no período de teste. A tabela abaixo mostra a média de cada performance.

Tabela 11 - Apenas Retornos no período de teste com valores negativos

	Retorno Treino	Retorno Teste	Sharpe Treino	Sharpe Teste	Calmar Treino	Calmar Teste
ADF	93%	-28%	3,25	-1,25	9,18	-1,72
PP	91%	-24%	3,04	-1,07	7,70	-1,41
PGFF	92%	-28%	3,05	-1,11	7,60	-1,44

Fonte: Resultados da Pesquisa

Observa-se que o retorno médio anualizado é praticamente o mesmo para os pares cointegrados via ADF e PGFF, porém o índice Sharpe anualizado médio no período de teste é maior para os pares cointegrados via PP e para os valores médios do índice Calmar se verifica que os pares cointegrados via PP apresentaram melhores resultados.

Para um administrador de ativos, implementar um acompanhamento com sinal de alerta para essas avaliações, procurando otimizar sua carteira, não é difícil computacionalmente. Contudo, é preciso repetir o apelo à cautela quando consideramos as idiosincrasias do mercado de ativos brasileiro, quando se consideram suas características de liquidez e concentração de ativos.

As metodologias, ademais, se prestam a assessorar tanto investidores interessados em posições de hedge como posições que tentem alcançar rentabilidades expressivas, contudo limitadas a períodos curtos, já que os sinais tendem a se dissipar na presença de efeitos paralelos.

5 CONCLUSÕES

Diferentes testes de estacionariedade levam a diferentes pares o que, conseqüentemente, leva a diferentes resultados em estratégias de pairs trading. No presente trabalho foi realizado o processo de formação de pares cointegrados com diferentes testes de raiz unitária a fim de verificar a lucratividade de trade por cointegração. Os resultados encontrados apontam que para o período de análise do estudo, 2020/04-2021/09, sendo 2020/04-2021/06 como período de formação de pares e o restante como período de validação, o teste de raiz unitária que apresentou um retorno médio maior foi o teste de Dickey-Fuller. Além disso foi aplicado o processo de filtragem dos pares dado o resultado do processo meia-vida, da qual é calculado via equação de Ornstein-Uhlenbeck. Observou-se que a estratégia de pairs trading por cointegração apresentou lucratividade para alguns pares.

Observou-se que o teste de Dickey-Fuller apresentou menor número de pares cointegrados se comparados com outros testes vigentes, porém apresentou um retorno médio anual maior que os outros pares além de apresentar métricas de performance melhores. O teste que apresentou uma melhor média dos retornos anuais positivos foi o teste PGFF, portanto a fim de potencializar um maior retorno se deve utilizar o teste PGFF, porém com um risco maior. Por outro lado, o teste que

apresentou uma média dos retornos negativos com menor valor foi o teste de Phillips-Perron, ou seja, a fim de se obter pares com menores riscos na estratégia o teste de Phillips-Perron é o mais indicado.

É importante salientar que retornos anuais negativos em pares de ações é um dos riscos que a estratégia de pairs trading apresenta, este fato acontece, pois dado a estimação do coeficiente γ na regressão linear no período de formação de pares, da qual se obtém determinado valor, no período de validação os regimes podem mudar o que leva, conseqüentemente, o coeficiente γ obter outro valor, essa mudança de regime pode ocorrer devido a fatores econômicos e fundamentalistas, uma possível forma de contornar este problema pode ser através do uso do filtro de Kalman.

Acreditamos, portanto, que as metodologias são úteis para a construção de estratégias de investimento/administração de ativos pois apresentam condições de operações lucrativas, ainda que seus riscos não sejam desprezíveis. Além disso, a estratégia mostrou-se válida em um mercado como o brasileiro, onde a concentração de ativos e uma relativa baixa liquidez é caracterizada frente a outros mercados mundiais.

REFERÊNCIAS

- ALEXANDER, C. Optimal hedging using cointegration. **Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences**, v. 357, n.1758, p. 2039–2058, 1999
- ALEXANDER, C.; GIBLIN, I.; WEDDINGTON, W. Cointegration and asset allocation: A new active hedge fund strategy. **Research in International Business and Finance**, v. 16, n. 5, p. 65–90, 2002.
- AVELLANEDA, M.; LEE, J.-H. Statistical arbitrage in the US equities market. **Quantitative Finance**, v. 10, n. 7, p. 761–782, 2010
- BACON, C. R. **Practical portfolio performance measurement and attribution**. [s.l.] John Wiley & Sons, 2008. v. 546
- BREITUNG, J. Nonparametric tests for unit roots and cointegration. **Journal of econometrics**, v. 108, n. 2, p. 343–363, 2002
- CALDAS, B. B.; CALDEIRAA, J. F.; MOURAB, G. V. **Is Pairs Trading Performance Sensitive to the Methodologies?: A Comparison**. Anais do XLII Encontro nacional de economia [proceedings of the 42nd Brazilian economics meeting]–ANPEC-Associação Nacional dos centros de pós-graduação em economia [Brazilian association of graduate programs in economics]. **Anais...**2016
- CALDEIRA, J.; MOURA, G. V. Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. **Available at SSRN 2196391**, 2013
- CERCHI, M.; HAVENNER, A. Cointegration and stock prices: The random walk on Wall Street revisited. **Journal of Economic Dynamics and Control**, v. 12, n. 2-3, p. 333–346, 1988.
- CHAN, E. **Algorithmic trading: winning strategies and their rationale**. [s.l.] John Wiley & Sons, 2013. v. 625
- CHIU, M. C.; WONG, H. Y. Mean–variance portfolio selection of cointegrated assets. **Journal of Economic Dynamics and Control**, v. 35, n. 8, p. 1369–1385, 2011
- CLEGG, M. On the persistence of cointegration in pairs trading. **Available at SSRN 2491201**, 2014.
- DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American statistical association**, v. 74, n. 366a, p. 427–431, 1979.
- EHRMAN, D. S. **The handbook of pairs trading: strategies using equities, options, and futures**. [s.l.] John Wiley & Sons, 2006. v. 240
- ELLIOT, B.; ROTHENBERG, T.; STOCK, J. Efficient tests of the unit root hypothesis. **Econometrica**, v. 64, n. 8, p. 13–36, 1996
- ELLIOTT, R. J.; VAN DER HOEK*, J.; MALCOLM, W. P. Pairs trading. **Quantitative Finance**, v. 5, n. 3, p. 271–276, 2005

ENDRES, Sylvia; STÜBINGER, Johannes. A flexible regime switching model with pairs trading application to the S&P 500 high-frequency stock returns. **Quantitative Finance**, v. 19, n. 10, p. 1727-1740, 2019

ENGLE, R. F.; GRANGER, C. W. Co-integration and error correction: representation, estimation, and testing. **Econometrica: journal of the Econometric Society**, p. 251–276, 1987

FAMA, E. F. The behavior of stock-market prices. **The journal of Business**, v. 38, n. 1, p. 34–105, 1965.

GATEV, E.; GOETZMANN, W. N.; ROUWENHORST, K. G. Pairs trading: Performance of a relative-value arbitrage rule. **The Review of Financial Studies**, v. 19, n. 3, p. 797–827, 2006

GIRMA, P. B.; PAULSON, A. S. Risk arbitrage opportunities in petroleum futures spreads. **Journal of Futures Markets**, v. 19, n. 8, p. 931–955, 1999

GLOVER, Kristoffer; PESKIR, Goran. Quickest Detection Problems for Ornstein–Uhlenbeck Processes. **Mathematics of Operations Research**, 2023

GRANGER, C. W. **Co-integrated variables and error-correcting models**. tese de doutorado— [s.l.] Discussion Paper 83-13. Department of Economics, University of California at ..., 1983

GRANGER, C. W.; NEWBOLD, P. Spurious regressions in econometrics. **Journal of econometrics**, v. 2, n. 2, p. 111–120, 1974

HERLEMONT, D. Pairs trading, convergence trading, cointegration. **YATS Finances and Technology**, v. 33, p. 1–31, 2003

HONG, G.; SUSMEL, R. Pairs-trading in the Asian ADR market. **University of Houston, Unpublished Manuscript**, 2003

HUCK, N.; AFAWUBO, K. Pairs trading and selection methods: is cointegration superior? **Applied Economics**, v. 47, n. 6, p. 599–613, 2015

JACOBS, B. I.; LEVY, K. N.; STARER, D. Long-short portfolio management: An integrated approach. **The Journal of Portfolio Management**, v. 25, n. 2, p. 23–32, 1999.

JOHANSEN, S. Statistical analysis of cointegration vectors. **Journal of economic dynamics and control**, v. 12, n. 2-3, p. 231–254, 1988

LIEW, R. Q.; WU, Y. Pairs trading: A copula approach. **Journal of Derivatives & Hedge Funds**, v. 19, n. 1, p. 12–30, 2013

LIN, Y.-X.; MCCRAE, M.; GULATI, C. Loss protection in pairs trading through minimum profit bounds: A cointegration approach. **Advances in Decision Sciences**, v. 2006, 2006

LUO, Jing; LIN, YuCheng; WANG, Sijia. Intraday high-frequency pairs trading strategies for energy futures: evidence from China. **Applied Economics**, p. 1-15, 2022

MIAO, G. 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**, v. 6, n. 3, p. 96–110, 2014
- MURRAY, M. P. A drunk and her dog: An illustration of cointegration and error correction. **The American Statistician**, v. 48, n. 1, p. 37–39, 1994
- MYLONIDIS, N.; KOLLIAS, C. Dynamic European stock market convergence: Evidence from rolling cointegration analysis in the first euro-decade. **Journal of Banking & Finance**, v. 34, n. 9, p. 2056–2064, 2010
- PANTULA, S. G.; GONZALEZ-FARIAS, G.; FULLER, W. A. A comparison of unit-root test criteria. **Journal of Business & Economic Statistics**, v. 12, n. 4, p. 449–459, 1994
- PHILLIPS, P. C. Understanding spurious regressions in econometrics. **Journal of econometrics**, v. 33, n. 3, p. 311–340, 1986
- PHILLIPS, P. C.; PERRON, P. Testing for a unit root in time series regression. **Biometrika**, v. 75, n. 2, p. 335–346, 1988
- RAD, H.; LOW, R. K. Y.; FAFF, R. The profitability of pairs trading strategies: distance, cointegration and copula methods. **Quantitative Finance**, v. 16, n. 10, p. 1541–1558, 2016
- RAMOS-REQUENA, J. P.; TRINIDAD-SEGOVIA, J.; SÁNCHEZ-GRANERO, M. Introducing Hurst exponent in pair trading. **Physica A: statistical mechanics and its applications**, v. 488, p. 39–45, 2017
- RYAN, J. A. et al. **Package ‘quantmod’**, 2020
- SHARPE, W. F. The sharpe ratio. **Journal of portfolio management**, v. 21, n. 1, p. 49–58, 1994.
- SIPILÄ, M.; OTHERS. Algorithmic pairs trading: empirical investigation of exchange traded funds. 2013
- SMITH, W. On the simulation and estimation of the mean-reverting Ornstein-Uhlenbeck process. **Commodities Markets and Modelling**, 2010
- STOCK, J. H.; WATSON, M. W. Testing for common trends. **Journal of the American statistical Association**, v. 83, n. 404, p. 1097–1107, 1988
- TAQQU, M. S.; TEVEROVSKY, V.; WILLINGER, W. Estimators for long-range dependence: an empirical study. **Fractals**, v. 3, n. 04, p. 785–798, 1995
- TAYLOR, M. P.; TONKS, I. The internationalisation of stock markets and the abolition of UK exchange control. **The Review of Economics and Statistics**, p. 332–336, 1989
- TEIXEIRA, A. A. E. **Pair trading in Bovespa with a quantitative approach: cointegration, Ornstein-Uhlenbeck equation and Kelly criterion**. São Paulo: Fundação Getúlio Vargas, Tese de Doutorado, 2014

UHLENBECK, G. E.; ORNSTEIN, L. S. On the theory of the Brownian motion. **Physical review**, v. 36, n. 5, p. 823, 1930

VIDYAMURTHY, G. **Pairs Trading: quantitative methods and analysis**. [s.l.] John Wiley & Sons, 2004. v. 217

VIRTANEN, Anssi. **Pairs trading: an application of pairs selection and outranking in Norwegian stock market**. 2020

WHISTLER, M. **Trading pairs: capturing profits and hedging risk with statistical arbitrage strategies**. [s.l.] John Wiley & Sons, 2004. v. 216

XING, Haipeng. A singular stochastic control approach for optimal pairs trading with proportional transaction costs. **Journal of Risk and Financial Management**, v. 15, n. 4, p. 147, 2022

YAHOO FINANCE. **Yahoo! Finance - Série de Dados Históricos**. Vários Anos. Disponível em: <https://finance.yahoo.com/>

YU, Fenghui et al. Optimal Pairs Trading Strategies: A Stochastic Mean–Variance Approach. **Journal of Optimization Theory and Applications**, v. 196, n. 1, p. 36-55, 2023