

Pairs Trading: Um Estudo Quantitativo no Mercado Financeiro Brasileiro

Hugo P. Villanova, Marcelo R. Bettarello, Antonio T. A. Paiva,
João P. C. Schmitt e Rosa F. N. Zhu

Mentor: Luciano B. França

Resumo

O presente estudo analisa a aplicação da estratégia de *pairs trading* no mercado brasileiro. Essa estratégia de arbitragem estatística busca identificar pares de ativos com uma relação histórica de equilíbrio, permitindo operações de compra e venda quando essa relação se desvia e tende a retornar à média. Para selecionar os pares, foi utilizado o Teste de Engle-Granger para verificar a cointegração entre os ativos, o Teste de Dickey-Fuller Aumentado para avaliar a estacionariedade das séries temporais e a distância euclidiana como critério de filtragem. A análise foi conduzida no setor de utilidade pública, com dados diários de preços de 26 ativos ao longo de um período de seis anos. A eficácia da estratégia foi avaliada por meio de um backtesting nos períodos de negociação, considerando diferentes combinações de critérios para a abertura e o fechamento de posições.

Palavras-chave: *pairs trading*, cointegração, estacionariedade, arbitragem, finanças quantitativas.

1 Introdução

Pairs trading é uma estratégia de arbitragem estatística que explora relações de cointegração entre dois ativos financeiros, assumindo que desvios temporários em seus preços tenderão a se corrigir ao longo do tempo. O método consiste em identificar pares de ativos historicamente cointegrados, abrindo simultaneamente uma posição comprada no

ativo subvalorizado e uma posição vendida no ativo sobrevalorizado quando a diferença entre seus preços se afasta de um valor médio esperado. Essa abordagem é baseada na premissa de reversão à média, um fenômeno amplamente estudado em mercados financeiros. A estratégia ganhou notoriedade na década de 1980, quando o matemático e trader Nunzio Tartaglia, liderando uma equipe quantitativa no Morgan Stanley, desenvolveu modelos estatísticos para identificar oportunidades de arbitragem entre pares de ações altamente correlacionadas. Desde então, *pairs trading* tem sido amplamente estudado e aplicado por fundos de hedge, gestores quantitativos e acadêmicos, evoluindo com o desenvolvimento de novos modelos estatísticos e computacionais.

A estratégia de *pairs trading* tem sido amplamente estudada na literatura acadêmica, com diferentes abordagens para a seleção de pares e definição de pontos de entrada e saída das operações. Os primeiros estudos sobre essa estratégia concentraram-se na metodologia baseada na distância entre preços, conforme proposto por Gatev, Goetzmann e Rouwenhorst (2006), que analisaram pares de ações no mercado norte-americano entre 1962 e 2002, obtendo retornos médios anuais de 11%. No entanto, a eficácia dessa abordagem tem sido contestada ao longo dos anos, pois a simples correlação histórica entre ativos pode não ser suficiente para garantir a persistência de oportunidades de arbitragem (DO; FAFF, 2010). Em contrapartida, abordagens mais avançadas, como a cointegração, têm se mostrado mais robustas para identificar pares com relações estatísticas mais estáveis ao longo do tempo (VIDYAMURTHY, 2004).

A aplicação da cointegração em *pairs trading* foi amplamente explorada em mercados emergentes, incluindo o Brasil. Caldeira e Moura (2013) demonstraram que a seleção de pares com base em testes de cointegração gera retornos superiores no mercado de ações brasileiro, alcançando um Sharpe Ratio de 1,34 e retornos anuais de 16,38%. De maneira similar, Huck e Afawubo (2015) compararam diferentes metodologias de seleção de pares e concluíram que a cointegração supera o método da distância em termos de rentabilidade e estabilidade dos retornos. Esses resultados reforçam a importância de utilizar modelos estatísticos mais sofisticados para evitar pares que apresentem apenas correlações espúrias e não uma relação de longo prazo fundamentada em forças econômicas comuns.

Além dos métodos de seleção de pares, a literatura também investigou o impacto de fatores exógenos na eficácia do *pairs trading*. Engelberg, Gao e Jagannathan (2009) analisaram o papel das notícias idiossincráticas na divergência de preços entre ativos cointegrados, concluindo que eventos específicos de empresas podem gerar desvios duradouros, tornando a arbitragem ineficaz em determinados momentos. Da mesma forma, estudos demonstraram que parte dos retornos de *pairs trading* pode ser explicada pelo momentum

de pares, sugerindo que algumas operações lucrativas derivam de efeitos de tendência e não exclusivamente da reversão à média. Essas descobertas indicam que os modelos tradicionais podem precisar de ajustes para incorporar fatores de risco adicionais e melhorar a precisão na identificação de oportunidades (CHEN et al., 2019).

Diante dessas limitações, estudos mais recentes buscaram aprimorar os modelos tradicionais de *pairs trading* com novas abordagens quantitativas. Avellaneda e Lee (2010) propuseram o uso da Análise de Componentes Principais (PCA) para capturar fatores latentes no comportamento dos preços e otimizar a seleção de pares. Em outra linha, Rad, Low e Faff (2016) compararam a eficácia de métodos baseados em cópulas estatísticas, mostrando que essa abordagem permite modelar melhor as relações não lineares entre ativos e oferece maior robustez em períodos de alta volatilidade. Esses estudos ressaltam a importância da evolução dos modelos matemáticos para aumentar a eficácia da estratégia e reduzir riscos operacionais.

A persistência da lucratividade do *pairs trading* também tem sido objeto de debate. Do e Faff (2010) e Do e Faff (2012) mostraram que, ao longo do tempo, os retornos da estratégia vêm diminuindo devido ao aumento da concorrência e da eficiência do mercado. No entanto, há evidências de que a estratégia ainda pode ser lucrativa em momentos de crise financeira. Do e Faff (2010) identificaram que o *pairs trading* apresentou desempenho superior durante períodos de turbulência, como a crise de 2008, sugerindo que a estratégia pode servir como um hedge contra eventos de mercado extremos. Esse fenômeno está alinhado com as descobertas de Rad, Low e Faff (2016), que apontam que métodos alternativos, como cópulas e modelos de regressão não linear, podem aumentar a resiliência da estratégia em condições adversas.

Embora a literatura mostre avanços significativos na modelagem e implementação de *pairs trading*, desafios como custos de transação e restrições de short selling ainda limitam a aplicabilidade prática da estratégia. Do e Faff (2012) analisaram o impacto dessas restrições e concluíram que, embora *pairs trading* ainda seja lucrativo, os retornos são significativamente reduzidos quando custos reais são considerados. Além disso, Elliott, Hoek e Malcolm (2005) desenvolveram um modelo baseado em cadeias de Markov Gaussianas para descrever a dinâmica do spread entre pares de ativos, permitindo uma melhor calibração dos pontos de entrada e saída da estratégia. Esse tipo de modelagem quantitativa, juntamente com abordagens mais recentes como as exploradas por Krauss (2017), que analisam o desempenho de diferentes metodologias de *pairs trading* incluindo machine learning, reforça a necessidade de adaptação constante da estratégia.

Embora a literatura demonstre a viabilidade da estratégia de *pairs trading* em diversos

mercados e setores, sua aplicação no Brasil ainda é pouco explorada. Assim, este estudo tem como objetivo testar e analisar a efetividade dessa estratégia no mercado acionário brasileiro, avaliando a persistência das relações de cointegração entre os ativos e sua capacidade de gerar retornos ajustados ao risco. Em particular, a pesquisa será focada no setor de utilidade pública, que se caracteriza por ativos regulados e historicamente menos expostos à volatilidade extrema, permitindo uma análise mais robusta da estratégia em um ambiente de menor incerteza. Este trabalho será desenvolvido no âmbito de um projeto semestral da Liga de Finanças Quantitativas (*FGV Quant*) em parceria com a *Avantgarde Asset Management*, integrando técnicas estatísticas para otimizar a seleção de pares e aprimorar o controle de riscos operacionais.

2 Dados

2.1 Ações Escolhidas

O presente estudo se inspirou fortemente na metodologia e nos resultados obtidos por Gatev, Goetzmann e Rouwenhorst (2006), que investigaram a estratégia de *pairs trading* no mercado acionário dos Estados Unidos. No artigo, os autores analisaram todas as ações presentes no banco de dados *CRSP daily files* que possuíam negociação contínua ao longo do período avaliado, verificou-se que o setor de utilidade pública foi um dos segmentos que mais se destacou em termos de retorno na estratégia. No nosso caso, para facilitar a condução do estudo e proporcionar uma abordagem mais direcionada, optou-se por restringir a análise ao setor de utilidade pública do mercado acionário brasileiro.

A escolha desse setor se justifica por diversos fatores. Primeiramente, empresas do mesmo setor tendem a apresentar maior correlação e, conseqüentemente, maior probabilidade de cointegração entre os ativos, o que aumenta a efetividade da estratégia de *pairs trading*. Além disso, empresas do setor de utilidade pública compartilham características fundamentais semelhantes, como regulação tarifária e demanda relativamente estável, o que favorece a formação de pares com padrões de longo prazo previsíveis. Outro fator relevante é a liquidez dos ativos, pois a negociação de ações com volume reduzido pode impactar negativamente a execução das ordens e aumentar os custos de transação. Por fim, reduzir a exposição a fatores idiossincráticos é essencial, dado que, conforme discutido por Engelberg, Gao e Jagannathan (2009), choques informacionais específicos podem afetar os retornos de determinados ativos e comprometer a estratégia.

Além de reduzir a complexidade da análise, a restrição a um único setor favorece a aplicabilidade do estudo dentro do contexto acadêmico da *FGV Quant*, permitindo uma investigação mais detalhada e replicável para estudos futuros. O foco no mercado

brasileiro também contribui para preencher lacunas na literatura, que frequentemente se concentra no mercado norte-americano. Os ativos selecionados para compor o estudo foram: *ALUP11*, *AMBP3*, *CEBR3*, *CGAS3*, *CLSC4*, *CMIG4*, *COCE5*, *CPFE3*, *CPLE6*, *CSMG3*, *EGIE3*, *EKTR4*, *ELET3*, *EMAE4*, *ENEV3*, *ENGI11*, *EQPA3*, *EQTL3*, *GEPA4*, *LIGT3*, *NEOE3*, *REDE3*, *SAPR11*, *SBSP3*, *TAE11* e *TRPL4*.

Com esse conjunto de 26 ativos, o número total de pares possíveis a serem avaliados na estratégia é dado pela combinação de dois elementos retirados de um conjunto de 26, sem repetição, conforme a equação abaixo:

$$\text{ParesPossíveis} = \frac{n \times (n - 1)}{q!} = \frac{26 \times 25}{2} = 325 \quad (1)$$

onde n representa o número total de ativos e q é o número de elementos escolhidos por par (no caso, $q = 2$). Assim, a análise considera um total de 325 pares potenciais para a implementação da estratégia.

2.2 Preços das Ações

Os dados utilizados neste estudo foram extraídos por meio da API *Yahoo Finance* utilizando a biblioteca *yfinance* do *Python*. Para garantir a integridade dos dados e evitar distorções causadas por eventos corporativos, utilizamos os preços de fechamento ajustados (*Adjusted Close*). Esses preços já incorporam ajustes por proventos, como pagamento de dividendos e desdobramentos de ações, refletindo a valorização real dos ativos ao longo do tempo. A base de dados cobre o período de 1º de janeiro de 2018 até 31 de dezembro de 2023, proporcionando uma amostra de seis anos para análise.

Após a extração, os dados foram tratados para remover eventuais valores ausentes (*NaN*) a fim de garantir a consistência das séries temporais. Após essa filtragem, verificou-se que todas as ações possuíam o mesmo número de observações ao longo do período analisado. Para padronizar a análise e possibilitar a comparação relativa entre os ativos, os preços foram normalizados para 1 no início do período de formação, conforme a equação abaixo:

$$P_t^{\text{norm}} = \frac{P_t}{P_0} \quad (2)$$

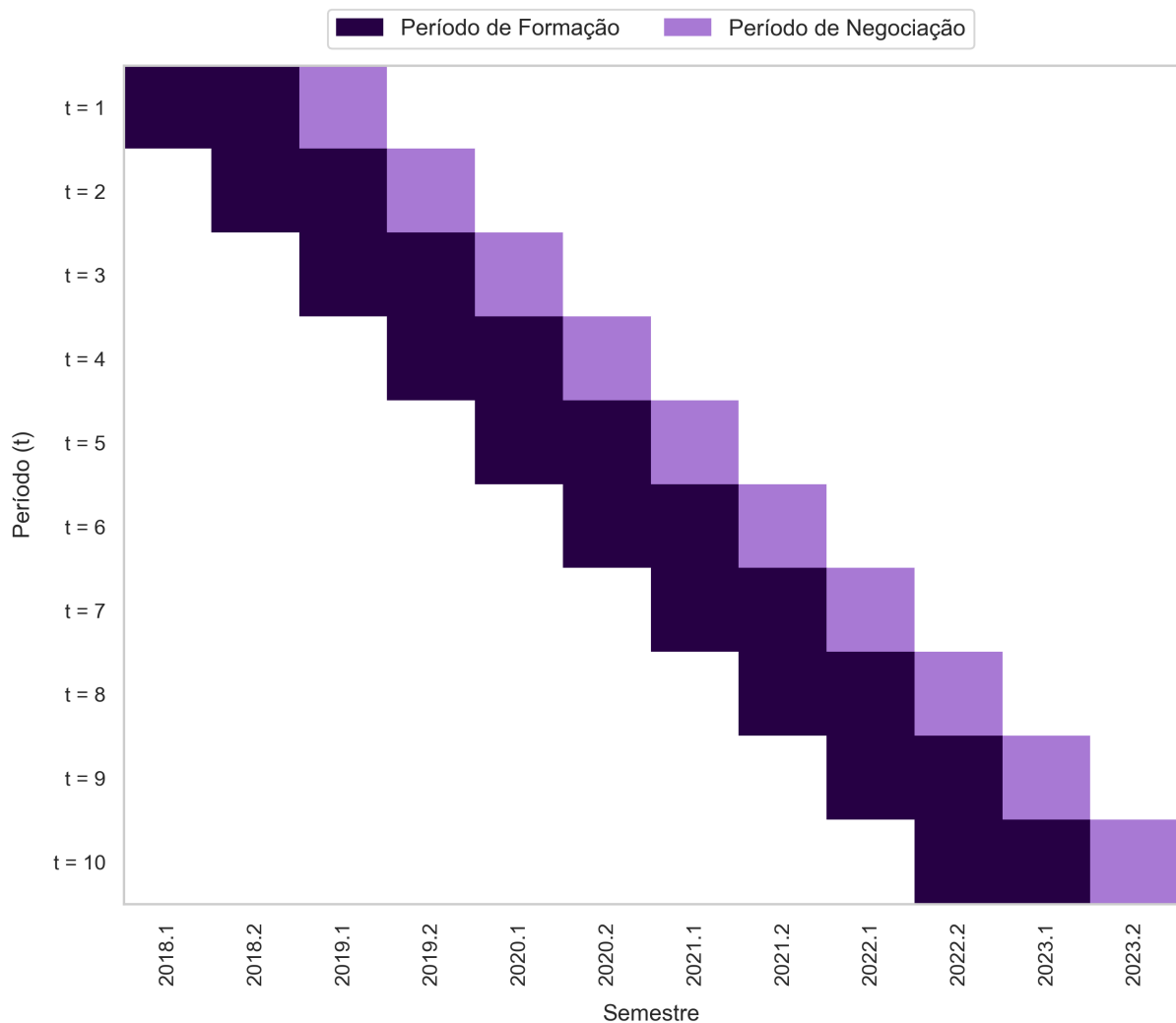
onde P_t representa o preço ajustado do ativo no instante t e P_0 é o preço ajustado na data inicial da série. Essa normalização permite comparar as variações percentuais dos preços dos ativos independentemente de seus valores nominais, facilitando a identificação de pares com movimentações semelhantes ao longo do tempo.

3 Metodologia

3.1 Divisão dos Períodos

Seguindo a metodologia proposta por Gatev, Goetzmann e Rouwenhorst (2006), o presente estudo adota uma estrutura de períodos baseada em janelas móveis para a realização da estratégia. Cada ciclo da estratégia é composto por um período de formação de 12 meses, onde os pares são identificados, seguido por um período de negociação de 6 meses, onde as operações são executadas e os retornos são mensurados. Essa configuração resulta em um total de 10 períodos de negociação ao longo do intervalo analisado. A Figura 1 ilustra a estrutura utilizada, destacando a sobreposição dos períodos.

Figura 1: Divisão dos períodos de formação e negociação



Fonte: Elaborado pelos autores.

A aplicação da metodologia ocorre de maneira estruturada ao longo desses períodos. Nos períodos de formação, a estratégia de seleção de pares é aplicada para identificar os ativos que exibem maior potencial de cointegração e retorno. Em seguida, nos períodos de negociação, os pares identificados são negociados conforme regras que serão explicadas posteriormente. Essa divisão clara entre formação e execução proporciona um maior rigor estatístico, garantindo que os dados utilizados para selecionar os pares não sejam contaminados por informações dos períodos de negociação futuros. Além disso, essa abordagem permite avaliar a robustez da estratégia ao longo do tempo, assegurando que os pares escolhidos mantenham suas propriedades estatísticas mesmo após a transição para o período de negociação.

3.2 Teste de Dickey-Fuller Aumentado

Antes de procedermos com a identificação de pares de ativos cointegrados, é necessário garantir que as séries temporais analisadas sejam integradas de ordem 1, ou seja, $I(1)$. No contexto dos preços de ativos financeiros, uma série é dita $I(1)$ quando seus valores em nível (preços) seguem um comportamento não-estacionário, mas suas primeiras diferenças (retornos diários) são estacionárias. Isso significa que os preços ajustados dos ativos podem exibir tendências estocásticas ao longo do tempo, mas as variações diárias não.

O Teste de Dickey-Fuller Aumentado (ADF) é utilizado para verificar a presença de uma raiz unitária em uma série temporal, indicando sua não-estacionariedade (ENDERS, 2008). O teste é aplicado tanto sobre os preços ajustados dos ativos quanto sobre suas primeiras diferenças (retornos diários) para determinar se as séries são de fato $I(1)$. A formulação geral do ADF pode ser expressa matematicamente da seguinte forma para os preços dos ativos:

$$\Delta P_t = \alpha + \beta t + \gamma P_{t-1} + \sum_{i=1}^p \delta_i \Delta P_{t-i} + \varepsilon_t \quad (3)$$

Já para os retornos diários, ou seja, as primeiras diferenças da série de preços, o teste é aplicado na seguinte forma:

$$\Delta^2 P_t = \alpha + \beta t + \gamma \Delta P_{t-1} + \sum_{i=1}^p \delta_i \Delta^2 P_{t-i} + \varepsilon_t \quad (4)$$

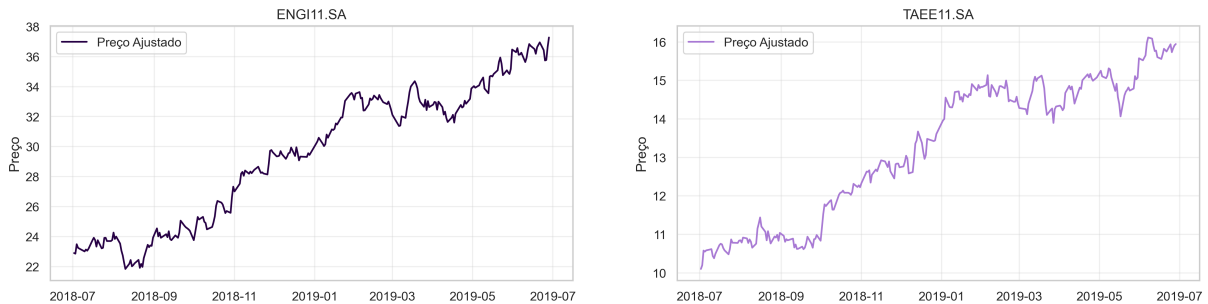
onde P_t representa os preços ajustados dos ativos, α é o termo constante, β representa uma possível tendência temporal, γ é o coeficiente associado à presença de raiz unitária, δ_i são coeficientes de defasagem para capturar autocorrelação, e ε_t é um termo de erro branco.

As hipóteses do teste de Dickey-Fuller Aumentado são formuladas da seguinte maneira: a hipótese nula (H_0) estabelece que a série possui raiz unitária ($\gamma = 0$), indicando que é não-estacionária, enquanto a hipótese alternativa (H_1) postula que a série não possui raiz unitária ($\gamma < 0$), caracterizando uma série estacionária. Para rejeitarmos a hipótese nula e concluirmos que a série é estacionária, utilizamos o p-valor do teste a um nível de significância de 5%. Caso o p-valor seja inferior a 0.05, rejeitamos H_0 e consideramos a série como estacionária, enquanto que, se o p-valor for superior a esse limite, não podemos rejeitar a presença de raiz unitária na série analisada.

3.2.1 Exemplo 1

Para fins ilustrativos, na Figura 2 analisamos os preços de fechamento ajustado dos ativos ENGI11.SA e TAEE11.SA durante o segundo período de formação do estudo, que compreende de 1º de julho de 2018 a 30 de junho de 2019. Ao aplicar o Teste de Dickey-Fuller Aumentado (ADF) nessas séries, obtivemos os seguintes resultados: para a ENGI11.SA, a estatística-teste foi de -0.35 com um p-valor de 0.92; para a TAEE11.SA, a estatística-teste foi de -1.05 com um p-valor de 0.74. Como os p-valores são elevados, não há evidências suficientes para rejeitarmos as hipóteses nulas de que essas séries possuem raiz unitária, indicando que ambas são não-estacionárias. Esse comportamento também é evidente pela representação gráfica das séries.

Figura 2: Séries de Preços de Fechamento Ajustado

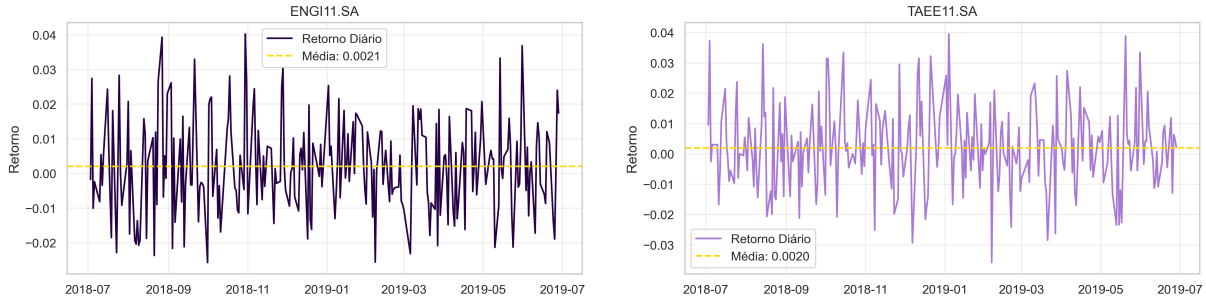


Fonte: Elaborado pelos autores.

Agora, na Figura 3 analisamos os mesmos ativos considerando suas séries de retornos diários. Ao aplicar o Teste de Dickey-Fuller Aumentado nessas séries, obtivemos os seguintes resultados: para a ENGI11.SA, a estatística-teste foi de -15.28 com um p-valor de 4.61×10^{-28} ; para a TAEE11.SA, a estatística-teste foi de -12.47 com um p-valor de 3.25×10^{-23} . Como os p-valores são extremamente baixos, temos evidências suficientes para rejeitarmos as hipóteses nulas de que essas séries possuem raiz unitária, indicando

que ambas são estacionárias. Essa conclusão é corroborada pela Figura 3, onde observamos que os retornos flutuam ao redor de uma média constante sem tendência evidente, confirmando um comportamento estacionário.

Figura 3: Séries de Retornos Diários



Fonte: Elaborado pelos autores.

3.3 Teste de Engle-Granger

O teste de Engle-Granger tem como objetivo verificar se duas séries temporais $P_{1,t}$ e $P_{2,t}$, ambas integradas de ordem 1 ($I(1)$), possuem uma combinação linear que seja estacionária. Essa relação indica que, apesar das séries de preços serem não-estacionárias individualmente, pode existir uma relação linear entre elas que exibe um comportamento estacionário. Caso essa relação de longo prazo exista, dizemos que as séries são cointegradas, o que sugere que seus movimentos estão ligados ao longo do tempo, sendo um elemento essencial para a estratégia de *pairs trading*.

O teste é composto por duas etapas. Na primeira etapa, realizamos uma regressão linear simples entre as duas séries $I(1)$ de preços dos ativos, estimando a seguinte equação:

$$P_{1,t} = \alpha + \beta P_{2,t} + \varepsilon_t \quad (5)$$

onde $P_{1,t}$ e $P_{2,t}$ são os preços dos ativos, α é o intercepto da regressão, β representa o coeficiente de relação entre os ativos, e ε_t são os resíduos da regressão, também chamados de *spread*.

A segunda etapa do teste consiste em aplicar o Teste de Dickey-Fuller Aumentado (ADF) sobre os resíduos ε_t da regressão, de forma a verificar se essa combinação linear é estacionária. A equação do teste ADF sobre os resíduos pode ser expressa como:

$$\Delta\varepsilon_t = \gamma\varepsilon_{t-1} + \sum_{i=1}^p \delta_i \Delta\varepsilon_{t-i} + \eta_t \quad (6)$$

onde γ é o coeficiente a ser testado, δ_i são coeficientes das defasagens dos resíduos para capturar autocorrelação, e η_t é um termo de ruído branco.

As hipóteses do teste de Engle-Granger seguem a mesma lógica do teste ADF, mas aplicadas aos resíduos ε_t . A hipótese nula (H_0) postula que os resíduos possuem raiz unitária, ou seja, são não-estacionários, o que indica que as séries $P_{1,t}$ e $P_{2,t}$ não são cointegradas. Já a hipótese alternativa (H_1) sugere que os resíduos são estacionários, o que implica que $P_{1,t}$ e $P_{2,t}$ são cointegradas e compartilham uma relação de longo prazo. Se rejeitarmos H_0 , concluímos que existe uma relação estatisticamente significativa entre os ativos, permitindo seu uso na estratégia de *pairs trading*.

3.3.1 Exemplo 2

Para facilitar o entendimento do teste de Engle-Granger, utilizamos os mesmos dois ativos do exemplo anterior, ENGI11.SA e TAEE11.SA, durante o segundo período de formação (2018-07-01 até 2019-06-30). A primeira etapa do teste consiste em estimar uma regressão linear simples entre as séries de preços ajustados desses ativos, conforme ilustrado a seguir:

$$P_{\text{ENGI11},t} = \alpha + \beta P_{\text{TAEE11},t} + \varepsilon_t \quad (7)$$

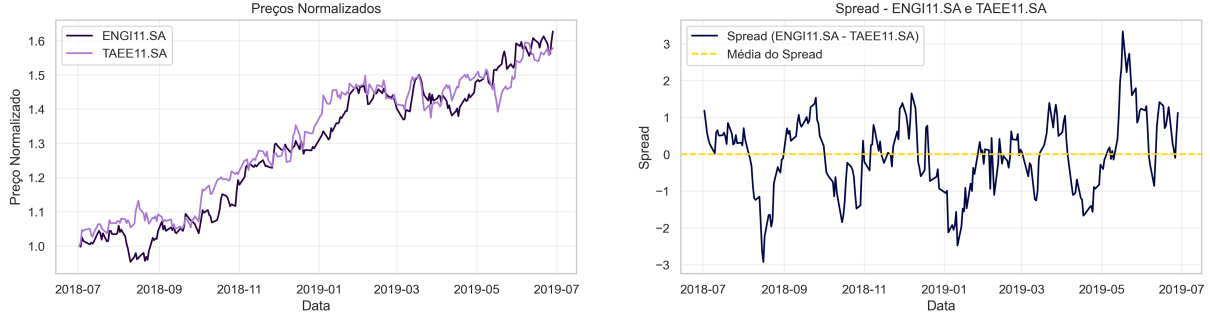
$$P_{\text{ENGI11},t} = -3.2168 + 2.3871 P_{\text{TAEE11},t} + \varepsilon_t \quad (8)$$

$$\varepsilon_t = P_{\text{ENGI11},t} + 3.2168 - 2.3871 P_{\text{TAEE11},t} \quad (9)$$

Os coeficientes α e β foram estimados a partir dos dados do período, e a série de resíduos ε_t , conhecida como *spread*, foi extraída para ser utilizada na segunda etapa do teste.

Na segunda etapa, aplicamos o Teste de Dickey-Fuller Aumentado (ADF) sobre o spread estimado. Os resultados do teste foram os seguintes: estatística-teste de -3.56 e um p-valor de 0.027 . A um nível de significância de 5% , esses resultados fornecem evidências estatísticas suficientes para rejeitarmos a hipótese nula de presença de raiz unitária na série do spread. Isso indica que a série de resíduos ε_t é estacionária, sugerindo a existência de uma relação de cointegração entre $P_{\text{ENGI11},t}$ e $P_{\text{TAEE11},t}$, ou seja, apesar dos preços individuais serem não-estacionários, existe uma relação linear entre eles que se mantém no longo prazo. A Figura 4 ilustra essa relação, onde no gráfico à esquerda apresentamos os preços normalizados das séries e no gráfico à direita o spread estimado, que visualmente aparenta ser estacionário.

Figura 4: Preços Normalizados e Spread Estimado



Fonte: Elaborado pelos autores.

3.4 Distância Euclidiana

Com a finalidade de filtrar os pares de ativos que apresentaram evidências de cointegração, selecionamos os cinco pares com as menores distâncias euclidianas. A distância euclidiana é uma métrica que quantifica a separação entre dois pontos em um espaço, sendo utilizada neste estudo para medir a diferença acumulada entre as séries de preços normalizadas dos pares testados como cointegrados em cada período de formação. Quanto menor essa métrica, mais próximas as trajetórias dos preços ao longo do tempo, o que sugere um movimento conjunto mais forte. A formulação matemática da distância euclidiana entre dois ativos $P_{1,t}$ e $P_{2,t}$ é dada por:

$$D(P_{1,t}, P_{2,t}) = \sqrt{\sum_{t=1}^T (P_{1,t} - P_{2,t})^2} \quad (10)$$

O objetivo dessa filtragem é priorizar pares de ativos que apresentem variações próximas ao longo do tempo, tornando as relações lineares entre eles mais previsíveis e consistentes. A seleção dos pares com menor distância euclidiana aumenta a confiabilidade de que os *spreads* entre os preços desses ativos tenderão a reverter à média, característica essencial para a identificação de oportunidades de arbitragem estatística. Dessa forma, essa métrica permite que a estratégia de *pairs trading* se concentre em pares com maior potencial de convergência, reduzindo o risco de falhas na identificação de relações espúrias entre os preços.

3.5 Cálculo dos Z-scores

Após selecionarmos os cinco pares cointegrados que, possivelmente, manteriam um comportamento adequado para a estratégia, é necessário definir uma metodologia para

operar esses pares durante o período de negociação. Para isso, utilizamos o *z-score* do *spread* entre os ativos, que nos fornece uma forma padronizada de monitorar a relação entre os preços ao longo do tempo e tomar decisões de entrada e saída em operações. O conceito do *z-score* e seus benefícios serão melhor explicados posteriormente.

Antes de calcularmos os *z-scores*, devemos primeiro estimar os parâmetros da relação linear entre os ativos, como feito na primeira etapa do teste de Engle-Granger. Dessa forma, realizamos a seguinte regressão linear entre os preços dos ativos durante o período de formação:

$$P_{1,t} = \alpha_{form} + \beta_{form}P_{2,t} + \varepsilon_t \quad (11)$$

Os coeficientes α_{form} e β_{form} estimados no período de formação são salvos e utilizados como parâmetros fixos para o cálculo do *spread* entre os ativos durante o respectivo período de negociação. Assim, o *spread* em um instante t do período de negociação (S_t) é calculado como:

$$S_t = P_{1,t} - (\alpha_{form} + \beta_{form}P_{2,t}) \quad (12)$$

Agora que temos uma forma de calcular o *spread* entre os ativos, podemos calcular o *z-score*, que é uma métrica padronizada para medir o quão distante o *spread* atual está da sua média histórica. O *z-score* nos permite identificar desvios significativos e criar regras objetivas de entrada e saída de operações, aumentando a previsibilidade da estratégia. A média e o desvio padrão utilizados nesse cálculo são aqueles obtidos durante o período de formação dos pares, garantindo que a estratégia se baseie em informações históricas e evitando a contaminação por dados futuros. A equação para o cálculo do *z-score* (z_t) é dada por:

$$z_t = \frac{S_t - \mu_{Sform}}{\sigma_{Sform}} \quad (13)$$

onde S_t é o *spread* calculado no instante t , μ_{Sform} é a média do *spread* no período de formação e σ_{Sform} é o desvio padrão do *spread* no período de formação. Com essa métrica, conseguimos definir regras claras para a execução das operações, que serão explicadas na próxima seção.

3.5.1 Exemplo 3

Assim como nos exemplos anteriores, utilizamos os ativos ENGI11.SA e TAEE11.SA no mesmo segundo período de formação (2018-07-01 até 2019-06-30) para ilustrar a aplicação do cálculo do *z-score*. Conforme realizado na primeira etapa do Exemplo 2, estimamos uma regressão linear simples entre as séries de preços ajustados dos ativos, obtendo os coeficientes $\alpha = -3.2168$ e $\beta = 2.3871$, como mostrado na Equação (8). Com

esses parâmetros fixados durante o período de formação, podemos calcular o *spread* ao longo do período de negociação (2019-07-01 até 2019-12-31), conforme a seguinte equação:

$$S_t = P_{\text{ENGI11},t} - (-3.2168 + 2.3871P_{\text{TAE11},t}) \quad (14)$$

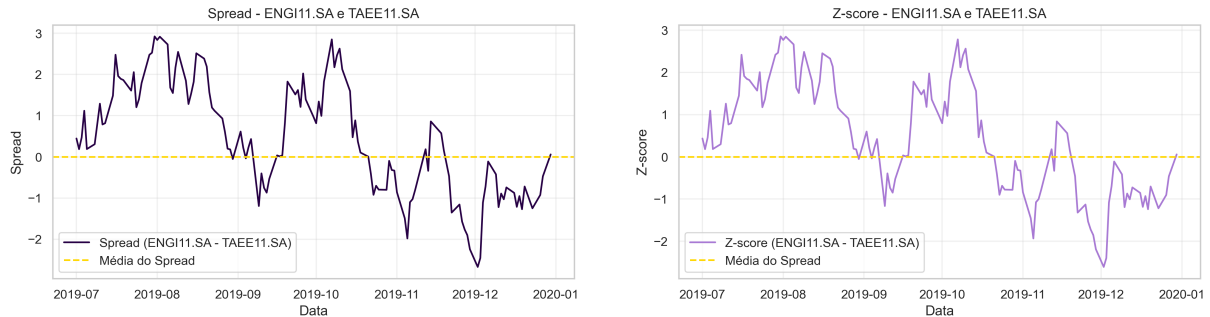
onde S_t representa o *spread* no instante t durante o período de negociação, $P_{\text{ENGI11},t}$ e $P_{\text{TAE11},t}$ são os preços ajustados dos ativos no instante t , e os coeficientes α e β foram estimados no período de formação. A Figura 5 apresenta uma representação gráfica do *spread* ao longo do período de negociação, ilustrado à esquerda da imagem.

Com o *spread* do período de negociação em mãos, podemos calcular o *z-score* do *spread* utilizando a média e o desvio padrão do *spread* estimados no período de formação. Para esse par de ativos, obtemos uma média de -4.3826×10^{-15} (valor extremamente próximo de zero) e um desvio padrão de 1.0257 (valor próximo de 1). Isso faz com que o gráfico do *z-score* do *spread*, apresentado à direita na Figura 5, seja bastante semelhante ao gráfico da esquerda. No entanto, vale ressaltar que essa característica nem sempre se mantém para todos os pares analisados. O cálculo do *z-score* para esse par específico é realizado conforme a seguinte equação:

$$z_t = \frac{S_t - (-4.3826 \times 10^{-15})}{1.0257} \quad (15)$$

onde z_t representa o *z-score* do *spread* no instante t , S_t é o *spread* calculado anteriormente em (14), -4.3826×10^{-15} e 1.0257 são a média e o desvio padrão, respectivamente, do *spread* no período de formação.

Figura 5: Spread e Z-score de Negociação



Fonte: Elaborado pelos autores.

3.6 Critérios para Ordens

A definição dos critérios para abertura e fechamento de ordens é uma das partes mais importantes da estratégia de *pairs trading*. A forma como essas regras são estabelecidas influencia diretamente a frequência das operações e a rentabilidade da estratégia. De modo geral, quanto maior for o valor do gatilho para entrada em uma ordem, menor será o número de aberturas de ordens em um período de negociação. Estudos como Gatev, Goetzmann e Rouwenhorst (2006) e Do e Faff (2010) adotam uma abordagem clássica, utilizando um limite de entrada de ± 2 desvios padrão do z -score do *spread* para abrir uma posição e um limite de saída de 0 desvios padrão para o fechamento da posição. Como destacado por Huck (2013), os resultados de uma estratégia de *pairs trading* são altamente sensíveis à escolha desses critérios, tornando essencial a análise de diferentes configurações para avaliar sua eficácia.

Neste estudo, testamos seis combinações diferentes de limites para entrada e saída de ordens, permitindo uma avaliação mais abrangente do impacto dessas regras na estratégia. Especificamente, utilizamos quatro valores distintos para o limite de entrada: ± 1.0 , ± 1.5 , ± 1.75 e ± 2.0 desvios padrão, e dois valores distintos para o limite de saída: 0.0 e ± 0.5 desvios padrão. Essas combinações permitem verificar se limiares mais agressivos ou mais conservadores impactam a estabilidade e a rentabilidade da estratégia, fornecendo informações valiosas sobre a melhor configuração dos critérios.

3.6.1 Exemplo 4

Dando continuidade ao mesmo par de ativos dos exemplos anteriores, ENGI11.SA e TAEE11.SA, agora demonstramos o funcionamento da abertura e fechamento de ordens com base no z -score, utilizando um limite de entrada de ± 2.0 desvios padrão e um limite de saída de ± 0.5 desvios padrão. A interpretação dessa métrica segue a seguinte lógica: sempre que o z -score atinge valores significativamente altos ($z_t > 2.0$), o *spread* está acima da média, indicando uma oportunidade para entrar vendido (*Short*) no ativo 1, que está relativamente mais caro, e entrar comprado (*Long*) no ativo 2, que está relativamente mais barato. Da mesma forma, se o z -score atinge valores significativamente baixos ($z_t < -2.0$), isso indica que o *spread* está muito abaixo da média, sugerindo a posição inversa: entrar comprado (*Long*) no ativo 1 e entrar vendido (*Short*) no ativo 2.

Após a abertura de uma posição, o critério de fechamento ocorre quando o z -score retorna para próximo de 0.5 ou -0.5 , cruzando esses valores em direção à média. Esse comportamento indica que o *spread* está se revertendo para seu valor esperado, sugerindo o fechamento da posição aberta. A Figura 6 ilustra esse processo para os ativos ENGI11.SA

e TAE11.SA durante o segundo período de negociação (2019-07-01 até 2019-12-31), onde podemos observar duas ocorrências de abertura de ordens para z -scores acima de 2.0, uma abaixo de -2.0 e seus respectivos fechamentos quando retorna para próximo de 0.5.

Figura 6: Ordens de Entrada e Saída



Fonte: Elaborado pelos autores.

3.7 Custos de Transação

Muitos estudos passados sobre *pairs trading* acabam por não computar explicitamente os custos associados à execução das operações, como comissões, impacto de mercado e outros custos de transação, ao calcular os retornos das estratégias. No entanto, Do e Faff (2012) realizou uma análise detalhada sobre o impacto desses custos, verificando se eles poderiam comprometer significativamente a rentabilidade das estratégias de *pairs trading*. O estudo concluiu que, embora os lucros tenham se tornado mais modestos ao levar em consideração os custos de transação, a maioria das estratégias continuou apresentando retornos positivos, demonstrando que esses custos não inviabilizam completamente a estratégia, mas precisam ser levados em conta para obter uma visão mais realista de sua viabilidade.

Existem diversas formas de quantificar os custos de transação, que podem variar de país para país, setor para setor e até mesmo de ativo para ativo. No nosso estudo, optamos por uma abordagem simplificada, assumindo um custo fixo de 0.1% sobre o valor negociado sempre que uma ordem de abertura ou fechamento de uma posição fosse realizada. Essa escolha foi baseada na necessidade de equilibrar precisão e eficiência, dado o tempo disponível para a condução do estudo. Apesar de ser uma simplificação, esse custo representa uma estimativa razoável para mensurar o impacto das despesas operacionais na rentabilidade da estratégia.

3.8 Cálculo dos Retornos

A contabilização dos retornos da estratégia foi realizada considerando que se trata de uma estratégia beta neutro, ou seja, sempre que nenhuma operação estivesse ativa para um par, 100% do patrimônio era alocado no CDI. Dessa forma, o retorno do par no instante t , quando não havia posições abertas, era dado por:

$$R_{\text{par},t} = R_{\text{CDI},t} \quad (16)$$

Quando uma operação Long estava ativa, o ativo 1 (subvalorizado) era operado como comprado e o ativo 2 (sobrevalorizado) como vendido. Assim, o retorno do par era calculado como:

$$R_{\text{par},t} = R_{\text{ação1},t} - R_{\text{ação2},t} \quad (17)$$

De forma análoga, quando uma operação Short estava ativa, o ativo 1 (sobrevalorizado) era vendido e o ativo 2 (subvalorizado) era comprado, resultando na seguinte equação:

$$R_{\text{par},t} = R_{\text{ação2},t} - R_{\text{ação1},t} \quad (18)$$

Além disso, sempre que uma operação era aberta ou fechada no dia t , o custo de transação de 0.1% era subtraído do retorno diário, de modo que:

$$R_{\text{par},t} = R_{\text{par},t} - 0.001, \quad \text{se houver abertura ou fechamento de posição no dia } t. \quad (19)$$

O retorno de cada período de negociação foi calculado somando os retornos diários de todos os cinco pares escolhidos no respectivo período de formação e dividindo o resultado pelo número de pares, obtendo assim um retorno médio para o período de negociação:

$$R_{\text{período},t} = \frac{1}{5} \sum_{i=1}^5 R_{\text{par}_i,t} \quad (20)$$

Finalmente, o retorno total da estratégia ao longo dos 10 períodos de negociação foi calculado considerando o reinvestimento dos lucros ao final de cada período. Dessa forma, ao invés de simplesmente somarmos os retornos médios de cada período, adotamos a seguinte estrutura de composição de retornos:

$$R_{\text{total}} = \prod_{j=1}^{10} (1 + R_{\text{período},j}) - 1 \quad (21)$$

Essa abordagem assegura que os retornos obtidos em cada período sejam acumulados e reinvestidos para os períodos seguintes, refletindo uma estratégia mais realista de crescimento de capital ao longo do tempo. Em outras palavras, dentro de cada período de negociação, os retornos dos cinco pares selecionados são somados e um retorno médio $R_{\text{período},j}$ é obtido. No entanto, o ajuste do capital ocorre apenas ao final de cada período, garantindo que os lucros ou perdas sejam incorporados ao portfólio antes do início do próximo ciclo.

4 Resultados

O objetivo desta seção é apresentar e comparar os retornos obtidos pelas diferentes configurações de critérios de ordens testadas em nossa estratégia de *pairs trading*. Como observado, os resultados gerais não foram muito animadores, com a maioria das configurações apresentando retornos inferiores ao CDI, que utilizamos como *benchmark*. A escolha do CDI como referência se deve ao fato de que nossa estratégia é baseada em arbitragem estatística com beta neutro, ou seja, buscamos capturar ineficiências de mercado sem exposição direcional ao risco sistêmico. Dessa forma, um retorno superior ao CDI seria desejável, pois indicaria que nossa abordagem gerou valor adicional em relação ao investimento livre de risco. A Tabela 1 apresenta um resumo dos retornos obtidos por cada configuração de entrada e saída testada.

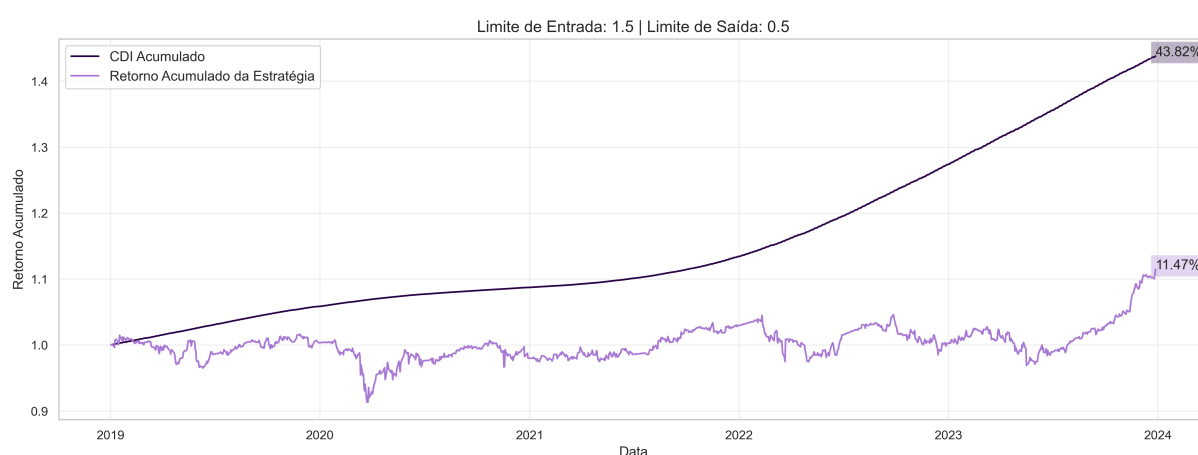
Tabela 1: Comparação dos Retornos das Estratégias Testadas

| Configuração (Entrada, Saída) | Retorno Total (5 anos) | Retorno Anualizado |
|-------------------------------|------------------------|--------------------|
| CDI (<i>Benchmark</i>) | 43.82% | 7.54% |
| (1.0, 0.0) | -6.63% | -1.36% |
| (1.5, 0.0) | 2.85% | 0.56% |
| (1.75, 0.0) | 3.41% | 0.67% |
| (2.0, 0.0) | 5.89% | 1.15% |
| (1.0, 0.5) | -14.66% | -3.13% |
| (1.5, 0.5) | 11.47% | 2.20% |
| (1.75, 0.5) | 5.42% | 1.06% |
| (2.0, 0.5) | 9.37% | 1.81% |

Ao analisarmos os resultados, observamos uma grande variação no desempenho das diferentes configurações. A melhor configuração testada foi com entrada em ± 1.5 desvios padrão e saída em ± 0.5 , apresentando um retorno total de 11.47% e um retorno anualizado

de 2.20%. Estratégias com limites de entrada mais altos, como ± 2.0 , também tiveram retornos positivos, mas menores, enquanto limites mais baixos, como ± 1.0 , resultaram em desempenho negativo. Além disso, o critério de saída influenciou os resultados, com a maioria das configurações melhorando quando o limite de fechamento foi ajustado para ± 0.5 . A Figura 7 apresenta o desempenho da melhor estratégia testada (1.5, 0.5) em comparação ao CDI, evidenciando a diferença de resultados entre os limites analisados.

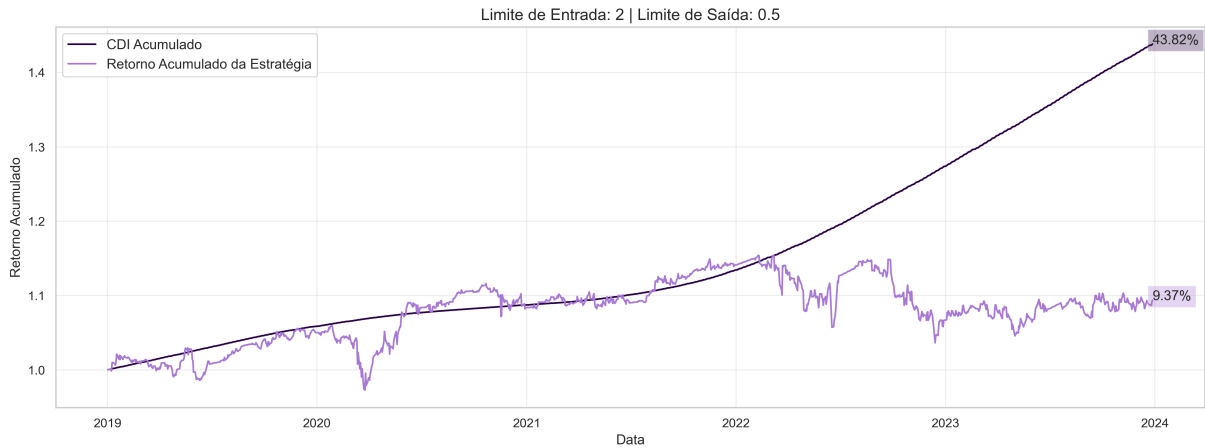
Figura 7: Desempenho da Melhor Estratégia Testada



Fonte: Elaborado pelos autores.

Além da configuração com entrada em ± 1.5 e saída em ± 0.5 , a segunda melhor estratégia foi aquela com entrada em ± 2.0 e saída em ± 0.5 , que obteve um retorno total de 9.37% e um retorno anualizado de 1.81%. Curiosamente, até o início de 2022, essa estratégia apresentou retornos muito próximos ao CDI, e em determinados momentos até superou o benchmark. No entanto, a partir de 2022, observou-se uma forte divergência nos retornos, com a estratégia apresentando um desempenho significativamente inferior ao CDI. Isso sugere que fatores macroeconômicos ou mudanças nas características do mercado podem ter influenciado negativamente a eficácia da abordagem. A Figura 8 ilustra essa trajetória, destacando a diferença na evolução dos retornos ao longo do tempo.

Figura 8: Desempenho da Segunda Melhor Estratégia Testada



Fonte: Elaborado pelos autores.

Na próxima seção, discutiremos os possíveis erros cometidos, limitações da abordagem utilizada e sugestões de melhorias para futuras implementações. Além disso, incluiremos conselhos fornecidos pelo nosso mentor sobre como tornar a estratégia mais robusta para diferentes condições de mercado.

5 Conclusão

O presente estudo implementou e avaliou uma estratégia de *pairs trading* baseada em arbitragem estatística, buscando identificar pares de ativos cointegrados e explorar oportunidades de reversão à média. No entanto, os resultados indicaram que, na forma como foi aplicada, a estratégia não conseguiu superar o CDI, que serviu como *benchmark*. Apesar de algumas configurações terem apresentado retornos positivos, o desempenho geral ficou aquém das expectativas, destacando a necessidade de ajustes na abordagem. Nesta seção, possíveis explicações para esse resultado serão apresentadas, bem como sugestões de melhorias para futuras implementações.

Um dos fatores que pode ter influenciado o desempenho da estratégia foi a seleção dos ativos e o número de pares analisados. Diferentes estudos, como Gatev, Goetzmann e Rouwenhorst (2006), trabalharam com um número significativamente maior de ativos e setores, ampliando a capacidade da estratégia de encontrar pares estatisticamente robustos. Além disso, Do e Faff (2012) testaram 29 metodologias diferentes para seleção de pares, evidenciando o impacto da escolha dos critérios na *performance*. No nosso estudo, utilizamos apenas 26 ativos e limitamos a estratégia a 5 pares por período de formação,

o que pode ter restringido as oportunidades disponíveis.

Outro ponto relevante é a definição fixa dos limites de entrada e saída das operações. Utilizamos limiares constantes de ± 1.5 e ± 2.0 para entrada e 0.0 ou 0.5 para saída. No entanto, como sugerido pelo nosso mentor, estratégias mais dinâmicas poderiam melhorar o desempenho, ajustando os limites de entrada e saída com base na volatilidade do mercado, por exemplo. Essa adaptação permitiria que os limites se ajustassem às condições de mercado, evitando a entrada excessiva em momentos de alta volatilidade e aproveitando melhor oportunidades em momentos de baixa volatilidade.

Além disso, a filtragem dos pares cointegrados foi feita com base exclusivamente na distância euclidiana, o que pode ter limitado a capacidade da estratégia de identificar pares mais robustos. Estudos como Huck e Afawubo (2015) destacam a importância da escolha do critério de seleção dos pares na eficácia da estratégia. Uma possível melhoria seria testar abordagens complementares, como o *Half-Life* da reversão à média, que mede a rapidez com que um ativo retorna ao seu valor médio, ajudando a selecionar pares com maior previsibilidade de reversão.

Outro fator a ser considerado é o tamanho dos períodos de formação e negociação. No estudo, utilizamos uma janela fixa de 12 meses para formação e 6 meses para negociação, um modelo tradicional na literatura. No entanto, Huck e Afawubo (2015) indicam que diferentes tamanhos de janela podem impactar significativamente os retornos. Uma abordagem mais sofisticada seria desenvolver um modelo que automatizasse a escolha dos períodos, otimizando o tamanho da janela para maximizar a efetividade da estratégia em diferentes condições de mercado.

A forma como os pares foram ponderados no portfólio também pode ter impactado os resultados. Atualmente, a estratégia atribui pesos iguais a todos os pares selecionados, sem levar em consideração seu nível de risco ou estabilidade da reversão à média. Como sugerido pelo nosso mentor, métodos como *Kelly Criterion* ou otimizações baseadas em minimização do *VaR* poderiam ser testados para alocar capital de forma mais eficiente, potencialmente reduzindo perdas e aumentando a robustez da estratégia.

Adicionalmente, é importante reconhecer que os últimos anos foram desafiadores para superar o CDI, mesmo para gestores profissionais. O cenário macroeconômico recente, caracterizado por mudanças bruscas nas taxas de juros e incertezas no mercado financeiro, pode ter impactado negativamente as oportunidades de arbitragem estatística. Isso reforça a necessidade de modelos mais adaptáveis às mudanças de regime de mercado, permitindo ajustes conforme as condições macroeconômicas variam ao longo do tempo.

Para aprimoramentos futuros, o ex-presidente da liga, Guilherme Regueira, atualmente

Equity Volatility Trading Analyst na Goldman Sachs, sugeriu algumas melhorias que podem aumentar a análise da estratégia e ampliar as chances de superação do CDI. Dentre essas melhorias, destaca-se a inclusão do *Sharpe Ratio* para avaliar o retorno ajustado ao risco, além do *Sortino Ratio*, que permite uma análise mais refinada da distribuição dos retornos, embora apresente limitações interpretativas quando os valores são negativos (nosso caso atual). Outra sugestão foi a inclusão do IBOV como referência adicional. Apesar de a estratégia ser neutra ao mercado, a comparação com o principal índice da bolsa brasileira pode fornecer um contexto mais amplo sobre o desempenho da abordagem. Tanto Guilherme quanto nosso mentor ressaltam ainda a necessidade de incluir um número maior de ativos no universo de seleção. A limitação a apenas 26 ativos e um único setor pode restringir as oportunidades de formação de pares cointegrados, sendo recomendável a ampliação do universo de ativos analisados para aumentar a robustez dos resultados.

Por fim, destacamos que este foi um projeto semestral desenvolvido dentro da Liga de Finanças Quantitativas, com tempo limitado para explorar todas as possibilidades de refinamento da estratégia. Apesar dessas limitações, o estudo proporcionou *insights* valiosos sobre os desafios do *pairs trading* e abriu caminhos para futuras pesquisas e melhorias. Estamos abertos a sugestões externas e acreditamos que, com os ajustes adequados, essa estratégia pode ser aprimorada para capturar oportunidades mais consistentes.

Agradecimentos

Gostaríamos de expressar nosso profundo agradecimento a todos os membros do grupo que contribuíram para a realização deste estudo, cujo comprometimento e dedicação foram fundamentais para o desenvolvimento do projeto. Agradecemos também ao nosso mentor, Luciano França, pelo suporte técnico e pelas valiosas orientações ao longo do semestre, que nos ajudaram a aprimorar a estratégia e a condução da análise. Nosso reconhecimento se estende à *Avantgarde Asset Management*, pelo incentivo à pesquisa aplicada em finanças quantitativas e pelo apoio institucional ao projeto.

Registramos nossa gratidão ao Diretor de Projetos do semestre, Diego Natel, pelo suporte na organização e estruturação das iniciativas dentro da entidade. Agradecemos também a todos os membros da Liga de Finanças Quantitativas, que, direta ou indiretamente, contribuíram com ideias, discussões e sugestões ao longo do processo. Por fim, um agradecimento especial aos ex-presidentes da entidade, Guilherme Regueira e Guilherme Rodrigues, cujas contribuições foram essenciais para definir aprimoramentos futuros da estratégia, reforçando a busca pela excelência em pesquisas quantitativas dentro da liga.

Referências

- AVELLANEDA, M.; LEE, J. H. Statistical arbitrage in the us equities market. *Quantitative Finance*, Taylor & Francis, v. 10, n. 7, p. 761–782, 2010.
- CALDEIRA, J. F.; MOURA, G. V. Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. *Revista Brasileira de Finanças*, Sociedade Brasileira de Finanças, v. 11, n. 1, p. 49–80, 2013.
- CHEN, H. et al. Empirical investigation of an equity pairs trading strategy. *Management Science*, INFORMS, v. 65, n. 1, p. 370–389, 2019.
- DO, B.; FAFF, R. Does simple pairs trading still work? *Financial Analysts Journal*, Taylor & Francis, v. 66, n. 4, p. 83–95, 2010.
- DO, B.; FAFF, R. Are pairs trading profits robust to trading costs? *Journal of Financial Research*, Wiley Online Library, v. 35, n. 2, p. 261–287, 2012.
- ELLIOTT, R. J.; HOEK, J. V. D.; MALCOLM, W. P. Pairs trading. *Quantitative Finance*, Taylor & Francis, v. 5, n. 3, p. 271–276, 2005.
- ENDERS, W. *Applied econometric time series*. [S.l.]: John Wiley & Sons, 2008.
- ENGELBERG, J.; GAO, P.; JAGANNATHAN, R. An anatomy of pairs trading: The role of idiosyncratic news, common information and liquidity. In: *Third Singapore International Conference on Finance*. [S.l.: s.n.], 2009.
- GATEV, E.; GOETZMANN, W. N.; ROUWENHORST, K. G. Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, Oxford University Press, v. 19, n. 3, p. 797–827, 2006.
- HUCK, N. The high sensitivity of pairs trading returns. *Applied Economics Letters*, Taylor & Francis, v. 20, n. 14, p. 1301–1304, 2013.
- HUCK, N.; AFAWUBO, K. Pairs trading and selection methods: Is cointegration superior? *Applied Economics*, Taylor & Francis, v. 47, n. 6, p. 599–613, 2015.
- KRAUSS, C. Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, Wiley Online Library, v. 31, n. 2, p. 513–545, 2017.
- RAD, H.; LOW, R. K. Y.; FAFF, R. The profitability of pairs trading strategies: Distance, cointegration and copula methods. *Quantitative Finance*, Taylor & Francis, v. 16, n. 10, p. 1541–1558, 2016.
- VIDYAMURTHY, G. *Pairs Trading: Quantitative Methods and Analysis*. [S.l.]: John Wiley & Sons, 2004. v. 217.