

FUCAPE FUNDAÇÃO DE PESQUISA E ENSINO

SÉRGIO ARTHUR DE SOUZA RABELLO

**PAIR TRADING, COINTEGRAÇÃO E COVID-19: uma análise
intrassetorial de retornos sob diferentes frequências de dados**

**VITÓRIA
2022**

SÉRGIO ARTHUR DE SOUZA RABELLO

PAIR TRADING, COINTEGRAÇÃO E COVID-19: uma análise intrassetorial de retornos sob diferentes frequências de dados

Dissertação apresentada ao Programa de Mestrado Profissional em Ciências Contábeis, da Fucape Fundação de Pesquisa e Ensino, como requisito parcial para obtenção do título de Mestre em Contabilidade – Nível Profissional.

Orientadora: Profa. Dra. Neyla Tardin

**VITÓRIA
2022**

SÉRGIO ARTHUR DE SOUZA RABELLO

PAIR TRADING, COINTEGRAÇÃO E COVID-19: uma análise intrassetorial de retornos sob diferentes frequências de dados

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Contábeis da Fucape Fundação de Pesquisa e Ensino, como requisito parcial para obtenção do título de Mestre em Ciências Contábeis.

Aprovada em 29 de agosto de 2022.

COMISSÃO EXAMINADORA

Profa. Dra. NEYLA TARDIN
Fucape Fundação de Pesquisa e Ensino

Prof. Dr. FELIPE STORCH DAMASCENO
Fucape Fundação de Pesquisa e Ensino

Prof. Dr. POLIANO BASTOS DA CRUZ
Fucape Fundação de Pesquisa e Ensino

RESUMO

O estudo tem como objetivo verificar se uma possível deterioração na relação de longo prazo entre ativos, causada pela pandemia do Covid-19, afetou o desempenho médio de estratégias de *pair trading* no mercado brasileiro, e como medidas de gerenciamento de risco e dados intradiários impactam nesse resultado. Com foco na abordagem de cointegração, o trabalho busca refinar os pares de ativos por meio de divisão setorial, utilizar o Valor em Risco (VaR) como critério de gerenciamento de risco e reproduzir diferentes estruturas de custo para todo tipo de investidor. Além disso, o uso da frequência intradiária no auge de volatilidade do mercado de ações em 2020 agrega a literatura de *HFT* (negociações de alta frequência) para trabalhos de cointegração, explorado poucas vezes no Brasil. Estudou-se o período entre janeiro de 2019 e junho de 2022 para dados de fechamento diários e janeiro de 2020 a junho de 2020 para dados de 15 e 5 minutos das ações que compõe a carteira teórica do Índice Bovespa. Para formação dos pares cointegrados foi utilizado o teste de Dickey-Fuller Aumentado com 1% de significância em uma janela contínua de operações ao longo de todo o estudo. Os resultados encontrados não sugerem mudanças no equilíbrio de longo prazo dos ativos ou perda de desempenho da estratégia. Pelo contrário, a divisão setorial proposta resultou em uma taxa de acerto superior a 80% em todo período para os dados diários, o que permitiu mesmo investidores com estruturas maiores de custos obterem retornos superiores ao índice de referência.

Palavras-chave: arbitragem estatística; cointegração; alta frequência.

ABSTRACT

The study aims to verify whether a possible disruption in the long-term relationship between assets caused by the Covid-19 pandemic affected the average performance of pair trading strategies in the Brazilian equit market, and also how risk management measures and intraday data impact this result. Focusing on the co-integration approach, this paper seeks to refine asset pairs through industry division, use Value at Risk (VaR) as a risk management criterion and reproduce different cost structures for all types of investors. In addition, the use of intraday frequency at the height of stock market volatility in 2020 aggregates the HFT (high frequency trading) literature for cointegration work, barely explored in Brazil. The period between January 2019 and June 2022 was studied for daily closing data and January 2020 to June 2020 for 15 and 5 minute data of the stocks that make up the theoretical portfolio of Bovespa Index. To form the cointegrated pairs, the Augmented Dickey-Fuller test was used with 1% significance in a continuous window of operations throughout the study. The results found do not suggest changes in the long-term balance of assets or loss of strategy performance. Contrariwise, the proposed sectoral division resulted in an accuracy rate of more than 80% throughout the period for daily data, which allowed even investors with higher cost structures to obtain excess returns from the benchmark.

Keywords: pair trading; cointegration; high frequency trading.

SUMÁRIO

1. INTRODUÇÃO	6
2. REFERENCIAL TEÓRICO	11
2.1 EFICIÊNCIA DE MERCADO	11
2.2 <i>PAIR TRADING E COINTEGRAÇÃO</i>	13
2.3 NEGOCIAÇÕES DE ALTA FREQUÊNCIA	17
3. METODOLOGIA	20
3.1 BASE DE DADOS	20
3.2 CARACTERIZAÇÃO DA AMOSTRA	21
3.3 HORIZONTE TEMPORAL	23
3.4 GERENCIAMENTO DE RISCO	27
3.5 TESTE DE COINTEGRAÇÃO PARA SELEÇÃO DOS PARES	28
3.6 ENTRADA DOS PARES	29
3.7 CUSTOS TRANSACIONAIS	31
3.8 APURAÇÃO DE RESULTADO E INDICADORES	33
4. RESULTADOS	36
5. CONCLUSÃO	47
REFERÊNCIAS	50
APÊNDICE A – Carteiras teóricas por período e setor	53
APÊNDICE B – Código usado para <i>backtest</i> (python)	55

Capítulo 1

1. INTRODUÇÃO

O conceito de *pair trading*, de acordo com Gatev *et al.* (2006), envolve a análise da correlação entre os retornos de dois ativos marcados a mercado. Primeiro, deve-se encontrar um par de ativos cujos retornos, historicamente, caminham juntos. Segundo, deve-se aguardar o momento em que ocorre uma distorção nessa relação e a diferença entre esses retornos, chamada de *spread*, aumenta. A partir daí, o investidor compra o ativo que desempenhou pior e aposta contra o que foi melhor. Se a relação de equilíbrio desses ativos se mantiver, o *spread* tenderá a diminuir novamente e então o ganho será realizado ao encerrar as posições.

Vidyamurthy (2004) foi um dos primeiros autores a analisar estratégias de *pair trading* com cointegração, modelo econométrico proposto por Engle e Granger (1987). Esse modelo busca estudar relações de equilíbrios de longo prazo entre variáveis, condição essa que sugere o retorno à média da estratégia, visto que, segundo a teoria, no curto prazo podem existir desvios dessa tendência em comum, porém os ativos não podem se divergir muito um do outro e logo tenderão a convergir de volta ao seu equilíbrio. Por essa abordagem se tratar de um modelo econométrico aplicado a séries temporais, eventos que alterem de forma drástica a forma com que os preços evoluem ao longo do tempo, como ocorrido na pandemia do ano de 2020, podem afetar essa relação e comprometer a capacidade de estimar qual ativo está subvalorizado ou

sobrevalorizado em um determinado momento no tempo.

Mazur *et al.* (2021) observou a reação dos preços de ações no mercado americano em março de 2020 e constatou que a mudança de dinâmica da economia, conforme evolução do número de contaminados pelo coronavírus e anúncios de intervenções abruptas feita pelo governo, fizeram com que 90% das ações do S&P1500 (responsável por aproximadamente 90% da capitalização do mercado americano) apresentassem grandes retornos negativos distribuídos assimetricamente, com alguns setores apresentando perdas médias de até 70% de valor de mercado. O autor aponta que tal evento difere de outras crises registradas, visto que grande parte dessas empresas apresentavam ativos valorizados e grande potencial de geração de fluxo de caixa livre, motivos os quais o colapso dos preços não ocorreu devido a uma bolha econômica.

Como o evento do Covid-19 afetou de forma rápida e abrupta os preços dos ativos, além de mudar a dinâmica na economia de todos os países do mundo, as relações de equilíbrio entre ativos e mercados podem ter se comprometido. Portanto, o objetivo desse estudo será verificar se uma possível deterioração na relação de longo prazo entre ativos, causada pela pandemia do Covid-19, afetou o desempenho médio de estratégias de *pair trading* no mercado brasileiro, e como medidas de gerenciamento de risco e dados intradiários impactam nesse resultado.

Estudos como de Rad *et al.* (2016) no mercado americano e Chen *et al.* (2017) no mercado chinês, apontam que o método de cointegração aplicado ao *pair trading* apresenta uma melhor eficácia em período de crises e maior volatilidade, desde que não ocorra quebras estruturais na relação entre as variáveis. Cavalcanti *et al.* (2021) também

constatou, para o mercado brasileiro, que um aumento na volatilidade tem efeito positivo para o retorno da estratégia.

Chen (2018) estudou comovimentos entre os 34 maiores mercados do mundo e observou que mercados emergentes, devido principalmente a fatores idiossincráticos, apresentam em média maior volatilidade do que mercados desenvolvidos (6,93% contra 4,93%). Neste mesmo estudo o Brasil apresentou, em média, a maior volatilidade (11,47%) da amostra durante os 24 anos de dados, o que indica que um choque na dinâmica dos preços pode oscilar mais intensamente neste mercado e afetar assim a relação de longo prazo que algumas variáveis eventualmente possuam.

Tais fatos tornam necessário uma análise do desempenho da estratégia com cointegração durante o período do Covid-19 para buscar verificar mudanças em sua lucratividade, visto que ainda não há publicações relevantes que contemplem esse período para o mercado brasileiro. De forma oportuna, este trabalho busca aproximar a metodologia da prática vivenciada por investidores no mercado brasileiro e ainda utilizará da frequência intradiária de dados para agregar a essa literatura ainda pouco explorada, quando comparada ao uso de dados diários, em *pair trading*.

O estudo contempla dados diários e intradiários (frequência de 5 e 15 minutos) de ações das carteiras do Índice Bovespa de cada período entre janeiro de 2019 e junho de 2022, agrupadas por seus respectivos setores. Foram definidas bandas de entrada, saída da operação para cenário de ganho e um limite máximo de perda tolerável, chamada de *stop loss*, calculada a partir do Valor em Risco (*VaR*), a fim de minimizar o resultado negativo das operações. Todos os dados foram extraídos do terminal MetaTrader 5 e sua

API (*Application Programming Interface*) para *python*, linguagem de programação também utilizada para tratar os dados e executar os testes deste estudo.

A validação da estratégia no período proposto é importante para que gestores e investidores tenham em mãos uma alternativa de diversificação para agregar desempenho aos seus portfólios mesmo em momentos conturbados, tornando-os mais descorrelacionados com o mercado. Além disso, promover o uso da frequência intradiária tende a aumentar a quantidade de operações a serem realizadas, o que de acordo com Perlin (2009), Benos e Sagade (2012) e Pontuschka e Perlin (2015), auxilia na liquidez dos ativos e ajuda a reduzir as distorções do mercado, fatores estes positivos tanto para investidores quanto para os reguladores.

Os resultados encontrados não sugerem mudanças no desempenho da estratégia ou equilíbrio de longo prazo dos ativos apesar do abrupto aumento da volatilidade. O uso do *VaR* como gerenciamento de risco prejudicou o desempenho e aumentou o desvio padrão dos retornos da carteira, enquanto a divisão setorial foi bem-sucedida e obteve para todos os períodos taxas de acertos superiores a 80% nas operações. Não foi possível encontrar pares cointegrados suficientes para analisar aumento de eficácia da estratégia ao utilizar a frequência intradiária de preços no período de maior volatilidade dos dados.

Na seção seguinte serão abordados a evolução dos métodos de *pair trading* até a cointegração, conceitos por trás da estratégia e resultados encontrados em outros estudos. Com os conceitos bem definidos, na seção 3 será apresentada a base de dados juntamente com a metodologia a ser adotada. As seções 4 e 5 serão discutidos os

resultados e a conclusão do estudo.

Capítulo 2

2. REFERENCIAL TEÓRICO

2.1 EFICIÊNCIA DE MERCADO

De acordo a Hipótese de Mercado Eficiente (HME), proposto por Fama (1970), o mercado é considerado eficiente quando os preços refletem totalmente as informações disponíveis, portanto qualquer nova informação é rapidamente incorporada na precificação dos ativos de forma correta e impossibilita assim ganhos anormais, já que os preços são compatíveis com as novas informações. Apesar de eficiência de mercado implicar em preços que refletem as informações disponíveis, ou seja, teoricamente não existem ativos com preços subvalorizados ou sobrevalorizados, isso não significa que não há carteiras de investimentos com desempenhos superiores a outras. Caso haja fatores imprevisíveis e aleatórios em novas informações divulgadas, o retorno para um determinado ativo, após os preços incorporarem as novas informações, pode ser diferente do originalmente esperado. Dessa forma, mesmo em um mercado eficiente, há a possibilidade de investidores apresentarem resultados superiores uns aos outros.

Décadas após o estudo de Fama (1970) sobre a HME, houve o crescimento de uma linha de pesquisa onde autores buscaram desenvolver modelos para explicar o movimento nos preços com base em fatores comportamentais. Estudos como Shefrin e Stratman (1994), Barberis *et al.* (1998), Daniel *et al.* (1998) e Hong e Steain (1999) foram

as principais referências de modelos citados no desenvolvimento da área de finanças comportamentais, área essa que rendeu um Nobel ao economista Richard Thaler, em 2017, ao explicar com sua teoria da contabilidade mental como as pessoas simplificam as tomadas de decisões financeiras e nem sempre a fazem de forma racional.

De Wet (2004) aponta que conforme a volatilidade em um mercado cresce, há maior incerteza acerca das informações disponíveis e se torna difícil ou custoso para os investidores distinguirem o que é bom ou ruim, o que leva a um aumento na assimetria de informação existente. Para o autor, o potencial comportamento do investidor sob maior assimetria de informação pode impedir ou retardar até crescimento econômico em países emergentes.

Barberis e Thaler (2003) fizeram um estudo a fim de revisar os principais pensamentos acerca de finanças comportamentais e ineficiência de mercado. No estudo, são enfatizados diferentes vieses psicológicos que levam os investidores a irracionalidade, vieses esses associados as expectativas e preferências dos investidores, restrições como o efeito de atenção dividida, e limites à arbitragem, onde riscos, regulações e custos relacionados as operações contribuem para existência de anomalias por períodos maiores que o esperado.

Thaler (2016) traz diversas referências de preferência de indivíduos, como escolhas intertemporais e preferência por cooperação em casos como caridade, e crenças que violam as expectativas racionais, como excesso de confiança e extrapolação excessiva, para mostrar que humanos são mais “gentis, educados e irracionais” do que os *Homo economicus*, ser racional que sempre toma decisões financeiras com base na

razão. Tais crenças retomam Bondt e Thaler (1985) onde investidores sobrestimam informações recentes e atribuem menos importância as antigas, baseada em seus julgamentos. Tal fator, chamado *overreaction*, poderia levar a um excesso de otimismo nas notícias boas e um extremo pessimismo nas notícias negativas, o que cria um desvio temporário dos valores intrínsecos dos preços de ações.

2.2 PAIR TRADING E COINTEGRAÇÃO

Segundo Alexander e Dimitriu (2005), estratégias de arbitragem permitem ao investidor obter ganhos ao se aproveitar de distorções momentâneas entre dois mercados distintos, incorrendo para isso baixo ou até mesmo nenhum risco. Para Friedman (1953), sempre que um ativo tem um desvio temporário do seu valor justo, uma oportunidade atrativa de investimento aparece e, com isso, investidores racionais imediatamente pegarão essa oportunidade e com isso corrigir esse desvio.

Para Ehrman (2006), *pair trading* é um tipo de operação de viés não direcional e valores relativos, ou seja, o sucesso da estratégia independe da direção do mercado e se baseia em um desvio no diferencial de preços histórico entre os ativos. Quando o desvio é observado, compra-se o ativo subvalorizado e aposta contra o sobrevalorizado em relação a esse diferencial.

Gatev *et al.* (2006) foram os primeiros a utilizarem uma abordagem de reversão a média para estratégias de *pair trading* em um estudo publicado em 1999 (Gatev *et al.*, 1999). O trabalho compreendeu os períodos entre 1962 e 1997 no mercado acionário americano e obteve um retorno anualizado bruto de 11% com baixa correlação com o

mercado de ações. Após os custos de transações, o retorno anualizado ficou entre 1,76% e 3,38%, o que confirmou a viabilidade da estratégia no mercado aplicado. Na ocasião, utilizou-se parâmetros definidos com base em consulta com operadores de mercado e analistas gráficos, que apesar de não contemplar medidas de gerenciamento de risco (como um limite de perda), popularizou a abordagem de entradas e saídas nas operações com base em uma métrica de desvio padrão.

A partir da metodologia de Gatev *et al.* (2006), os autores Do e Faff (2010 e 2012) elaboraram um dos estudos mais notados de *pair trading* com abordagem no método da distância. A ideia foi expandir o período de análise original até 2009 e refinar o critério de seleção de pares, visto que o critério original tinha como base a correlação dos ativos, o que tornou crescente as relações espúrias e, conseqüentemente, o número de pares que não convergiam. Tal problema fez com que a estratégia original, após os custos de transação, se tornasse inviável. Dessa forma, Do e Faff (2010) separaram os pares por indústria e montaram uma classificação dos melhores pares com base em critérios estatísticos como soma dos desvios quadrados, o que tornou a estratégia viável novamente.

A reversão a média proposta no trabalho de Gatev *et al.* (1999) abriu a porta para que outros autores desenvolvessem novas metodologias na área de *pair trading*. Como correlação não é uma condição suficiente para um retorno a média, Vidyamurthy (2004) propôs aprofundar a estratégia e elaborar regras com base em conceitos estatísticos e econométricos. Tal estudo trouxe a cointegração para o universo de *pair trading*, onde o autor em sua metodologia realizou a seleção de pares com base nessa abordagem e um filtro que leva em consideração a correlação entre os pares, assim como feito também

em Whistler (2004). Huck e Afawubo (2015) mostraram que pares escolhidos com base em cointegração apresentam reversão média mais frequente do que os escolhidos com base no método da distância.

Rad *et al.* (2016) fez o primeiro estudo empírico de larga escala que utilizou cointegração no mercado americano. No estudo, o autor comparou em 52 anos de dados os principais métodos de *pair trading* da literatura: Distância, Cointegração e Copula. Os resultados sugerem que tais métodos na média geram retornos positivos mesmo após os custos de transações e o método de Cointegração se comporta significativamente melhor do que os demais em períodos de crise e maior volatilidade.

Jacobs e Weber (2015) fizeram um estudo de *pair trading* entre o período de 2000 a 2013, onde são comparadas as rentabilidades em 34 países distintos, sendo 14 deles emergentes (Brasil incluso). Os autores constataram que a divergência entre os pares de mercados emergentes tem extensão maior do que em mercados desenvolvidos, o que levaram esses países a apresentarem retornos em média 25% maiores no período. Já Chen *et al.* (2017) elaborou um estudo de *pair trading* utilizando cointegração no mercado de *commodities* chinês entre os anos 2006 e 2016, e obteve resultados positivos robustos que sugerem também eficácia do método para mercados emergentes.

No Brasil, Kabashima (2006) foi pioneiro ao seguir a metodologia proposta por Vidyamurthy (2004) e utilizar a abordagem da cointegração para o mercado brasileiro. O estudo fez a seleção de pares com critério na correlação e teste de cointegração com *Dickey-Fuller* aumentado para ações do Índice Bovespa entre janeiro de 1993 e junho de 2005. Os resultados encontrados foram superiores a taxa básica de juros e não foi

possível observar uma relação entre a significância da cointegração e o resultado do par, ou seja, pares cointegrados com 1% de significância não necessariamente terão desempenhos melhores do que pares com 10% de significância.

Já Caldeira e Moura (2013) realizou um trabalho de cointegração com relevância internacional, como visto em Krauss (2017), em que aplicou um modelo univariado de cointegração para as 50 ações mais líquidas do Índice Bovespa entre janeiro de 2005 e dezembro de 2009. Caldeira e Moura (2013) classificou os 20 melhores pares com base no índice de Sharpe e obteve retornos excedentes significativos estatisticamente mesmo após custos transacionais.

Perlin (2009) introduziu para o do mercado brasileiro a ideia da frequência de observação dos dados para *pair trading*. O autor utilizou a abordagem da distância, entre os anos 2000 e 2006 para o IbrX-100, índice que mede as 100 ações mais líquidas do mercado acionário brasileiro, com frequência de dados mensais, semanais e diários. Os resultados do autor sugerem que quanto maior a frequência de observação, melhores os resultados da estratégia. Pontuschka e Perlin (2015) realizaram uma ampliação do mesmo estudo para dados diários e intradiários, nas frequências de 1, 5, 15, 30 e 60 minutos, comparados a uma carteira passiva de mercado composta por 20 ativos. Os resultados sugeriram indícios de melhora na rentabilidade líquida em relação aos dados diários ao se utilizar as frequências de 1, 5 e 15 minutos. Os autores apontam que quanto mais frequente uma amostra é observada, maior será a probabilidade de observar distorções.

Cavalcanti *et al.* (2021) aplicaram para o mercado brasileiro um estudo sobre os

efeitos da volatilidade nas estratégias de *pair trading* sob a perspectiva da cointegração. O estudo compreende os 90 pares mais líquidos da Bolsa, Brasil, Balcão (B³), entre os períodos de janeiro de 2016 e dezembro de 2018, e sugeriu que carteiras de *pair trading* compostas por ativos de maior volatilidade apresentam menores prazos operacionais e desempenho médio superiores as carteiras cointegradas montadas sem critério de seleção. Tal resultado corrobora com as contribuições de Chen *et al.* (2017) e Rad *et al.* (2016) a respeito dos melhores retornos estarem associados a volatilidade.

2.3 NEGOCIAÇÕES DE ALTA FREQUÊNCIA

As negociações de alta frequência, também conhecidas como *HFT* (*high-frequency trading*), ganharam bastante notoriedade após o evento que ficou conhecido como *Flash Crash* de 2010, ou simplesmente *Flash Crash*, onde o mercado americano teve uma “quebra relâmpago” de trilhões de dólares em sua bolsa de valores e se recuperou minutos depois, com duração de evento estimado de aproximadamente 36 minutos. Segundo Kirilenko *et al.* (2017), a mídia na época se fascinou com a mistura de tecnologia e frequência de atividade do *HFT*, culpadas por expor a fragilidade de mercados automatizados, apesar dos autores constatarem no estudo que o *HFT* não causou o *Flash Crash*, porém contribuiu significativamente para a volatilidade extraordinária ocorrida no evento.

No estudo, Kirilenko *et al.* (2017) apontam que sob situações normais de mercado, o *HFT* contribui apenas um pouco a mudança de preço, sem causar movimentos direcionais e ainda agrega ao volume de negociação. Entretanto, em momentos de

mercado estressado e elevada volatilidade, quando os preços se movem de forma direcional devido algum desequilíbrio no fluxo de ordens, a atividade do *HFT* pode amplificar o movimento direcional e contribuir para o aumento de volatilidade.

Dewan e Mendelson (1998) apontam que com a integração dos mercados e constante evolução dos sistemas computacionais, oportunidades de extrair ganhos sem risco através de distorções entre ativos e mercados, chamadas “arbitragens”, se tornaram escassas, visto que a alta frequência, elevada capacidade de processamento e velocidade de tomada de decisão dos sistemas eliminavam rapidamente a diferença entre os preços. Portanto, tal evolução dos sistemas faz com que alternativas a tradicional arbitragem livre de risco, como a abordagem da cointegração aplicada ao *pair trading*, surjam a fim gerenciar algum controle sobre o comportamento dos ativos.

Miao (2014) relata que com a evolução da área e dos métodos computacionais, estratégias de arbitragem estatística, como o *pair trading*, tem sido utilizada em conjunto do *HFT*. O autor aponta ainda que tal junção fez com que períodos de operações diminuíssem significativamente, como de semanas para minutos, o que aumentou a frequência dos lucros potenciais. Além disso, segundo Benos e Sagade (2012) e Pontuschka e Perlin (2015) como estratégias de *HFT* buscam obter lucros com ineficiências temporárias no mercado, elas podem auxiliar na velocidade com que as informações são repassadas ao preço, o que pode ajudar na estabilização e correção daquele mercado.

Apesar de estudos sobre *pair trading* apontarem que a estratégia de cointegração se beneficia em cenários de aumento de volatilidade, a pandemia do Covid-19 trouxe

oscilações extremas, mudou a dinâmica na economia de todos os países do mundo e se diferenciou de quaisquer outras crises econômicas, conforme Mazur *et al.* (2021). Além disso, Kirilenko *et al.* (2017) apontou que o *HFT* em momentos de grande estresse agrava movimentos direcionais, o que pode colaborar no distúrbio da relação de equilíbrio entre ativos e mercados. Caso esse cenário afete o equilíbrio dos pares, o benefício da volatilidade pode não ser observado como em Rad *et al.* (2016), Chen *et al.* (2017) e Cavalcanti *et al.* (2021). Além disso, a ausência de estudos relacionados que aproximem a metodologia da prática e explorem medidas de gerenciamento de risco no período da pandemia deixam uma lacuna sobre sua efetividade em mitigar efeito adverso de eventos de grande magnitude, enquanto também não é claro se o uso da maior frequência de dados, como sugerido por Perlin (2009) e Pontuschka e Perlin (2015), de fato tem benefício nesses períodos, em especial se combinado a medidas de gerenciamento de risco. Tais fatos levam a formulação das seguintes hipóteses de pesquisa:

H1: o aumento volatilidade após o início da pandemia agregou desempenho a carteira cointegrada de *pair trading*;

H2: parâmetros de gerenciamento de risco agregam desempenho ao portfólio cointegrado de *pair trading* após o início da pandemia;

H3: dados de frequência intradiária apresentam desempenhos superiores no período de maior volatilidade da pandemia.

Capítulo 3

3. METODOLOGIA

Alexander e Dimitiu (2005) mostraram que os potenciais lucros de uma estratégia de arbitragem dependem da existência de equilíbrio de longo prazo entre o preço dos ativos, choques de curto prazo que afetem o spread dessa relação e a convergência de volta para o equilíbrio anterior. Como este é um trabalho que busca aproximar a metodologia com a prática vivenciada por investidores, após refinar quais pares podem ser utilizados, conforme sugerido por Do e Faff (2010 e 2012), será utilizado o método proposto por Engle e Granger (1987), devido sua simplicidade e eficácia, para selecionar pares cointegrados e observar desvios na relação de equilíbrio, a fim de buscar oportunidades de entradas nas operações. Nessa mesma linha de implementação no dia a dia dos investidores, para realizar a estratégia será utilizada a abordagem cash neutral, em que o financeiro da venda da ação sobrevalorizada é o mesmo da compra da ação subvalorizada, não envolvendo caixa na operação.

3.1 BASE DE DADOS

Um clássico problema apontado nos estudos sobre *pair trading*, segundo Gatev et al. (2006), Do e Faff (2012) e Jacobs e Weber (2015) é a diferença entre o preço de compra e venda disponíveis no livro de oferta ao montar a estratégia, chamado de *bid-ask spread*. Se essa diferença for grande, corre o risco de o investidor entrar na estratégia

em parâmetros que não configurem a entrada desejada, o que poderá prejudicar seu desempenho ao longo do tempo por se tornar algo não mais embasado estatisticamente. Para atenuar esse efeito, foram selecionados apenas ações que compõem o Índice Bovespa, principal índice de ações brasileiro composto por 92 ativos que representam cerca de 80% dos números de negócios e volume financeiro do mercado de capitais brasileiro.

Os dados foram coletados entre o início do primeiro semestre de 2019 até o final do primeiro semestre de 2022, a fim de observar o comportamento dos retornos em torno do início da pandemia. A escolha da frequência diária e intradiária para os preços se deve ao intuitivo fato de que quanto menor o tempo observado, maiores as chances de serem encontradas distorções e ineficiências no mercado (Perlin, 2009).

Para tratar eventos corporativos como dividendos, juros sobre capital próprio, desdobramento, agrupamento, entre outros, a plataforma utilizada para extrair os preços de fechamento, que nesse estudo foi o MetaTrader 5, disponibiliza as cotações já ajustadas, a fim de não haver saltos inexistentes na relação de preços entre as duas variáveis e promover com isso quebras estruturais. O servidor utilizado no MetaTrader 5 foi da corretora XP Investimentos, que é a responsável por fornecer todos os dados extraídos da plataforma para realização do estudo.

3.2 CARACTERIZAÇÃO DA AMOSTRA

A fim de evitar problemas relacionados ao uso de informações ainda não disponíveis na época do teste, conhecido como *look-ahead bias*, foi utilizada a carteira

teórica do Índice Bovespa dos respectivos anos que o estudo ocorreu. Como os ativos Smiles (SMLS3), Unidas (LCAM3) e NotreDame Intermédica (GNDI3) foram incorporadas por outras empresas e deixaram de ser negociadas em bolsa, eles foram excluídos da análise, visto que o ajuste dos preços por eventos corporativos futuros que ocorreram em outros ativos pode não refletir a base comparativa da época de análise.

A fim de obter uma melhor relação de risco e retorno para o portfólio e otimizar a quantidade de pares a serem observados, conforme apontado por Do e Faff (2010 e 2012), este estudo refinou o universo de análise ao agrupar empresas por setores e assim mitigar riscos específicos entre setores distintos. Como o objetivo do teste de cointegração é encontrar, respeitado sua significância estatística, ativos cujos preços possuem um equilíbrio no longo prazo, é natural evitar pares com estruturas de riscos muito diferentes. O número máximo de pares é obtido por meio de arranjo matemático, visto de que regredir o ativo X contra Y terá um resíduo diferente do que regredir Y em X. A quantidade de possíveis arranjos por setor é exposta na Tabela 1:

TABELA 1: NÚMERO MÁXIMO DE PARES COINTEGRADOS POR SETOR

Setores definidos	Número de ativos	Número máximo de pares
Aéreas	2	2
Alimentos e Bebidas	5	20
Bens Industriais	5	20
Construção Civil	4	12
Educação	2	2
Exploração de imóveis	5	20
Financeiro	14	182
Mineração e Siderurgia	7	42
Papel de Celulose	2	2
Petróleo, gás e biocombustíveis	7	42
Saúde	6	30
Tecnologia da Informação	3	6
Telecomunicações	2	2
Utilidade Pública	12	132

Varejo	12	132
Total	88	646

Fonte: Autor

Os ativos Braskem (BRKM5) e Localiza (RENT3) foram excluídos da divisão setorial por representarem um tipo de negócio muito específico ou legislações que não afetam em mesma intensidade outras empresas do índice. A divisão setorial fez com que o arranjo dos 88 ativos tenha formado no máximo 646 possíveis pares cointegrados, contra 7.656 caso não houvesse a divisão. As carteiras de ativos a serem analisados foram separadas por semestre e sua relação, junto de sua divisão setorial, está disponível no apêndice A deste trabalho. O universo completo dos ativos estudados é encontrado na Tabela 2:

TABELA 2: ATIVOS ESTUDADOS

ABEV3	BRML3	ECOR3	HYPE3	MULT3	SOMA3
ALPA4	CASH3	EGIE3	IGTI11	NTCO3	SULA11
AMER3	CCRO3	ELET3	IGTI3	PCAR3	SUZB3
ASAI3	CIEL3	ELET6	IRBR3	PETR3	TAE11
AZUL4	CMIG4	EMBR3	ITSA4	PETR4	TIMS3
B3SA3	CMIN3	ENBR3	ITUB4	PETZ3	TOTS3
BBAS3	COGN3	ENEV3	JBSS3	POSI3	UGPA3
BBDC3	CPFE3	ENGI11	JHSF3	PRI03	USIM5
BBDC4	CPLE6	EQTL3	KLBN11	QUAL3	VALE3
BBSE3	CRFB3	EZTC3	LOGG3	RADL3	VBBR3
BEEF3	CSAN3	FLRY3	LREN3	RAIL3	VIIA3
BPAC11	CSNA3	GGBR4	LWSA3	RDOR3	VIVT3
BPAN4	CVCB3	GOAU4	MGLU3	RRRP3	WEGE3
BRAP4	CYRE3	GOLL4	MRFG3	SANB11	YDUQ3
BRFS3	DXCO3	HAPV3	MRVE3	SBSP3	-

Fonte: Autor

3.3 HORIZONTE TEMPORAL

Estudos de *pair trading* que utilizam cointegração comumente dividem a

metodologia em dois horizontes temporais bem definidos para dados diários: o período de formação (ou período de treino), composto por 12 meses, e período de *trading*, composto por 4 ou 6 meses. O primeiro período, chamado de formação, consiste em separar os potenciais pares a serem utilizados no segundo período, chamado período de *trading*. Todas as possíveis combinações de pares do período de formação são submetidas a testes de cointegração, a fim de encontrar quais séries possuem a potencial característica para reverter a média.

O principal problema dessa metodologia é que a partir do momento que as séries passam para o segundo período, a condição de cointegração não é mais testada até que se finalize o período de *trading*, além de que ao final desse período os pares são finalizados para que outra janela com novos potenciais pares se inicie. Na prática, um investidor não definirá janelas de *trading* e sempre terá preferência por checar a existência de cointegração antes de entrar em um par. Por este motivo, este trabalho adotou o mesmo período de 12 meses para checar a existência de cointegração entre pares de frequência diária, porém todos os possíveis pares tiveram sua condição de cointegração checada diariamente, a fim de abraçar possíveis novas relações e evitar de estudar pares que não possuem mais condições estatísticas de serem operados.

A proposta pode ser observada na Figura 1, antes de uma possível entrada ocorrer em um determinado par, observou-se os 12 meses anteriores a fim de avaliar a existência de cointegração, o que cria uma janela móvel de observação com o decorrer dos dias.

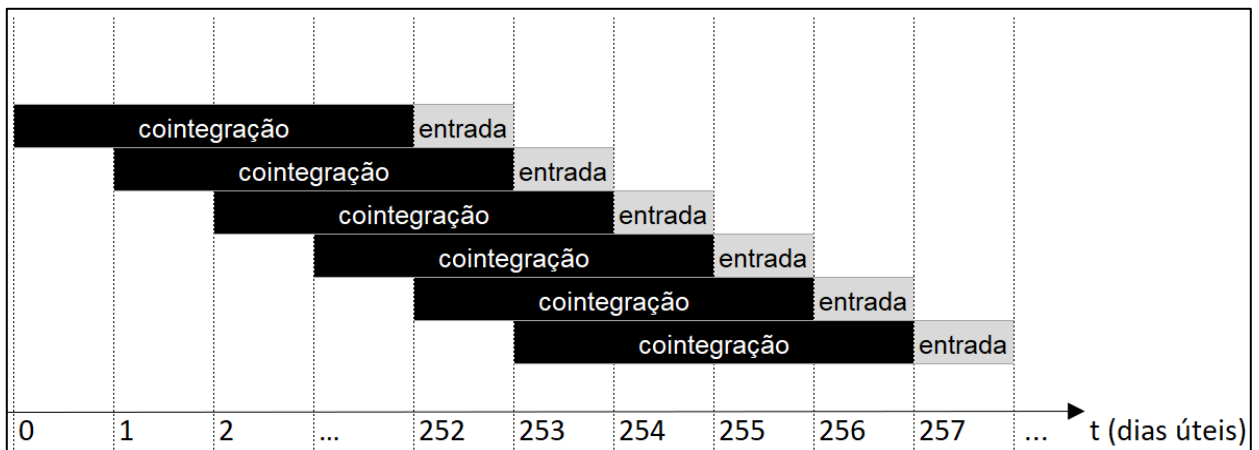


Figura 1: Análise dinâmica de pares cointegrados
 Fonte: Autor

Para frequência intradiária foi adotado o período de formação de 10 semanas (50 dias úteis) como proposto por Ernest Chan (2009) e as verificações ocorreram com a mesma janela móvel que os dados diários. Outro ponto importante para constatar é o período em que os dados intradiários foram utilizados, visto que a abordagem de maior frequência se propõe a estudar o período de volatilidade mais extrema da pandemia. Para chegar a esse período, primeiramente foram extraídos na B³ os dados de volatilidade do Índice Bovespa entre 2008 e 2019, maior janela disponível antes da pandemia. A volatilidade média observada nesse período foi de aproximadamente 23,88%. Em seguida, os dados de volatilidade entre 2009 e 2020 foram extraídos a fim de avaliar qual foi a volatilidade mensal encontrada neste último ano analisado. As volatilidades mensais podem ser observadas na figura 2:

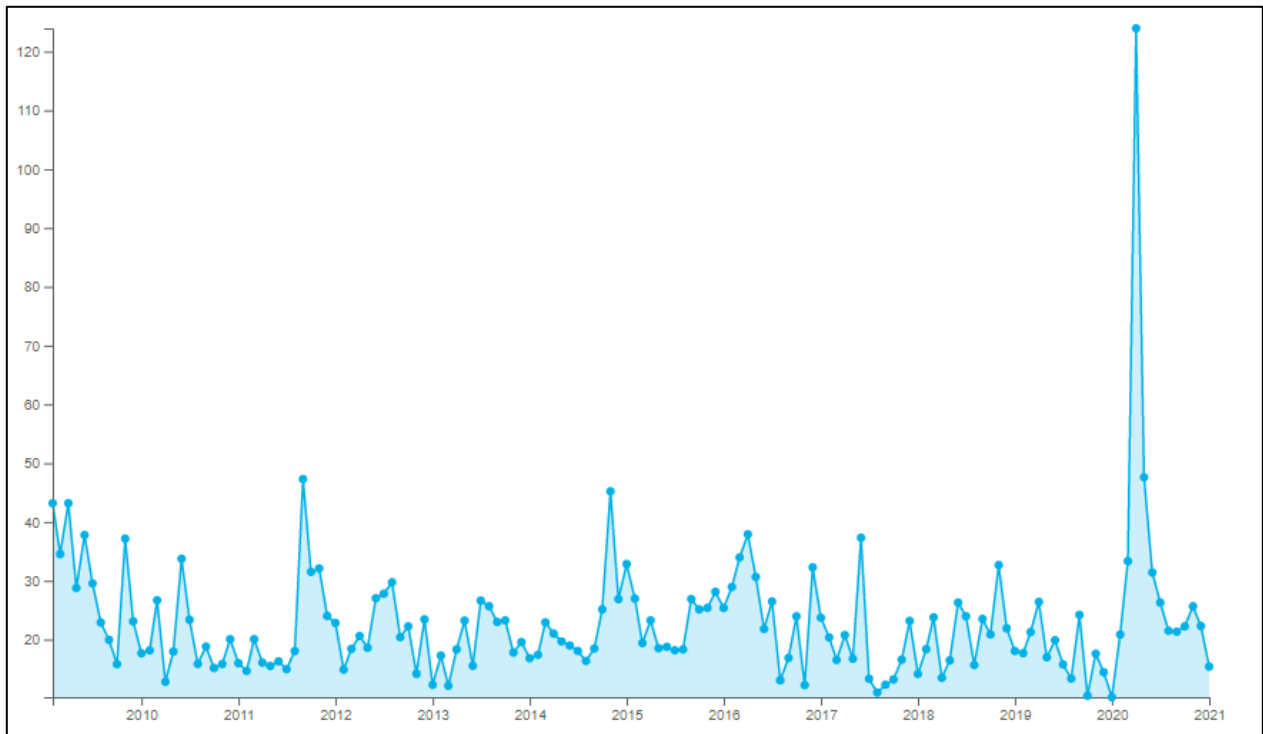


Figura 2: volatilidade mensal Índice Bovespa
 Fonte: Bolsa, Brasil, Balcão [B³]

Os valores mensais de 2020 foram extraídos da figura 2 e podem ser observados na Tabela 3:

TABELA 3: VOLATILIDADE IBOVESPA 2020

Mês	Volatilidade (%)
Janeiro	20,85
Fevereiro	33,32
Março	123,94
Abril	47,59
Maiο	31,41
Junho	26,3
Julho	21,5
Agosto	21,33
Setembro	22,24
Outubro	25,65
Novembro	22,3
Dezembro	15,4

Fonte: Autor

O mês de março destoou dos demais e foi ao menos cinco vezes a média história

observada anteriormente. Em seguida, buscou-se os períodos sequenciais em torno de março onde a volatilidade foi também superior a histórica, que são fevereiro, abril, maio e junho. A fim de tornar os resultados comparáveis com outros trabalhos, o mês de janeiro foi incluído para compor o semestre.

Por fim, os primeiros e últimos 15 minutos do pregão regular foram descartados independente da frequência intradiária observada, enquanto todos os dados de *after market* foram descartados por representar um período de liquidez reduzida e regras diferentes de negociação como bandas de volatilidade e preço. Assim o período de estudo intradiário contemplou o primeiro semestre de 2020, enquanto o estudo diário de janeiro de 2019 até o final de junho de 2022.

3.4 GERENCIAMENTO DE RISCO

Ao buscar uma medida de gerenciamento de risco amplamente utilizada e de simples implementação para investidores, o *VaR* não paramétrico não necessita de pressupostos quanto a forma de distribuição dos retornos e seu uso como limite de perda para um par não é encontrado na literatura de *pair trading*. Como o período de estudo deste trabalho contempla períodos de alta volatilidade, o uso do *VaR* pode auxiliar o investidor a manter a posição tempo suficiente até que ocorra a reversão esperada do equilíbrio dos pares.

Abad et al. (2014) aponta que o método de simulação histórica do *VaR* é amplamente o mais utilizado entre os modelos não paramétricos. A partir do número de dias adotados, calcula-se os retornos históricos e os ordena em ordem crescente, do

menor retorno par ao maior retorno. Em seguida, adota-se um intervalo de confiança para separar os piores retornos da amostra. Os valores encontrados são o *VaR* histórico estimado. Para exemplificar o método, se um *VaR* histórico com 99% de confiança resultar em -5%, isso significa que há 1% de chance do ativo se desvalorizar ao menos em -5%, portanto cria-se uma referência de valor para gerenciar o risco do portfólio.

Para o estudo, foi adotado um *VaR* de 1% com simulação histórica de 12 meses, tempo máximo que asseguramos a existência de cointegração para um par, visto que esta é a janela temporal em que a cointegração é testada. A escolha do período é importante pois, ao olhar esta janela, garante-se que o retorno do par se dá em condições de existência de equilíbrio, confirmada pela cointegração. Já a significância, buscou-se adotar um valor que não fosse muito restritivo, visto que a alta volatilidade pode levar o retorno até o *VaR* com frequência, entretanto em momento de mensurar a eficácia da estratégia, múltiplos valores de *VaR* são comparados.

3.5 TESTE DE COINTEGRAÇÃO PARA SELEÇÃO DOS PARES

Iniciado o período de formação, deve-se aplicar o teste de cointegração de Engle e Granger (1987), a fim de avaliar se as séries são cointegradas. O procedimento é composto pelas seguintes etapas:

- Determina-se a ordem de integração dos pares envolvidos através do Teste de Dickey-Fuller Aumentado, também conhecido como *ADF* (*Augmented Dickey Fuller*);

- Se as variáveis forem integradas de primeira ordem, isto é $I(1)$, estima-se uma regressão linear simples da relação de longo prazo e obtêm-se os resíduos estimados $\hat{\alpha}_t$;
- Checar se os resíduos possuem raiz unitária pelo ADF.

Rejeitar a hipótese nula do teste *ADF* indica que os resíduos não possuem raiz unitária e, assim, as variáveis são cointegradas. Como os resíduos são estimados, devemos utilizar a tabela disponibilizada no trabalho de Engle e Granger (1987) para encontrar os valores apropriados. Para assegurar maior rigor estatístico, este estudo adota o nível de 1% de significância para os testes de *ADF*.

3.6 ENTRADA DOS PARES

Ao iniciar a análise, novos dados são introduzidos de forma dinâmica para que os resíduos dos pares sejam calculados a cada observação pelo *ADF*, a fim de encontrar a relação de cointegração, e desvios dos retornos dos ativos em relação à média de longo prazo. O cálculo do desvio é feito pelo escore padrão, referido na literatura por *z-score* (Z_t), calculado pela fórmula:

$$z_t = \frac{x - \bar{x}}{DP_{amostra}} \quad (1)$$

As variáveis são definidas a seguir:

x = o valor do dado;

\bar{x} = média da amostra;

DP é o desvio padrão;

Assim como utilizado nos trabalhos de Caldeira e Moura (2013), Huck e Afawubo (2015), Rad *et al.* (2016), Chen *et al.* (2017) e Cavalcanti *et al.* (2021), o escore padrão utilizou dois desvios como um dos critérios de entrada na operação:

$$\begin{cases} \text{Se } Z_t > 2, \text{ compra o ativo X e vende o ativo Y} & (a) \\ \text{Se } Z_t < -2, \text{ compra o ativo Y e vende o ativo X} & (b) \end{cases}$$

Vidyamurthy (2004) e Whistler (2004) sugerem como filtro de pares uma análise da correlação daqueles a serem utilizados, portanto se o escore padrão satisfazer as condições determinadas acima, observa-se se a correlação do par está contida no quartil superior de todos os pares disponíveis naquele momento. Caso ambas as condições estejam satisfeitas, inicia-se uma operação e são aguardados um dos três cenários distintos para o encerramento:

- Há uma reversão a média do *spread* e ocorre o lucro se $Z_t < 0,5$ em (a) ou $Z_t > -0,5$ em (b);
- O retorno do par no dia cai além do seu valor em risco (*VaR*) determinado. Nesse caso a operação é encerrada independente do resultado e o par deve aguardar o escore padrão retornar para baixo de 2 desvios e permanecer assim por ao menos um dia;
- Caso o período de análise finalize, a operação é encerrada e o resultado apurado independente do resultado.

3.7 CUSTOS TRANSACIONAIS

Os custos de transação em uma estratégia de *pair trading* estão divididos entre os custos da corretora para emitir as compras e vendas (conhecido como corretagem), as taxas da bolsa para efetuar as transações (liquidação, emolumentos e seus respectivos tributos), e os custos de aluguéis que incidem sobre as posições vendidas. A fim de apurar os resultados mais próximos da realidade de diferentes investidores, a corretagem (único custo referente a corretora) será dividida para três tipos de investidores: o que paga o custo de “tabela bovespa”, muito utilizado ao utilizar robôs de *trading* e mesa de operação das corretoras e dos bancos, o que tem 50% de desconto de corretagem, comum de ocorrer visto que uma operação envolve dois ativos, e o investidor que investe a corretagem zero. O investidor que opera pela “Tabela Bovespa” paga 0,5% de corretagem, o que somado as taxas da B³, resultará em uma taxa aproximada de 0,61% sobre o volume financeiro operado.

A fim de simplificar alguns custos relacionados ao contrato de aluguel da ação, tais como taxa de registro, renovação, entre outras, foi criada a taxa de intermediação, cobrada uma única vez e proporcional ao volume da ponta vendida da operação. Esse custo é de 0,25% sobre o volume vendido, valor este relevante que será considerado no estudo.

Outro custo envolvido é o aluguel da ação vendida, custo este de valor proporcional ao número de dias na operação e com taxa a depender da procura pelo aluguel deste ativo. Como as ações do Índice Bovespa possuem boa liquidez, essas taxas não costumam ser elevadas. Para o estudo, foi coletado do site da B³ (Brasil, Bolsa,

Balcão) todas as taxas negociadas entre os ativos do índice no pregão de 21/01/2022. Em seguida, a fim de não distorcer a amostra, retirou-se as taxas superiores a 10% ao ano, visto que para ativos de maior liquidez tal taxa representa algum evento específico, e não a média negociada ao longo do tempo. Dessa forma, a taxa obtida e, conseqüentemente, adotada para o estudo, foi de aproximadamente 1,43% ao ano. Para obter a taxa proporcional ao período, o custo é dado por:

$$Tx = Q * P * \left\{ \left[\left(1 + \frac{i}{100} \right)^{\frac{n}{252}} \right] - 1 \right\} \quad (2)$$

As variáveis são definidas a seguir:

Tx = Taxa de liquidação do contrato (em reais)

Q = Quantidade de ativos

P = Cotação do ativo no início do contrato de aluguel

i = Taxa de empréstimo ao ano

n = Número de dias úteis com contrato em aberto

A fim de tornar o custo do aluguel mais próximo da realidade praticada no mercado, após o encerramento da operação são necessários mais 2 dias úteis para liquidação da posição e, conseqüentemente, do contrato de aluguel. Portanto, o número de dias com contrato em aberto será composto pelo tempo que a operação ficou em aberto somado ao tempo necessário para que o contrato seja liquidado.

Por fim, acrescentou-se o *bid-ask spread*, que apesar de ser um custo implícito e de influência menor em ativos mais líquidos, impacta no resultado da estratégia. Pontuschka e Perlin (2015) apontam que o valor adotado costumeiramente na literatura

é de 0,05%, portanto este foi o valor final adotado uma única vez na entrada e outra na saída. A relação final dos custos é encontrada na Tabela 4:

TABELA 4: ESTRUTURAS DE CUSTO POR INVESTIDOR

Custo de corretagem	Custo compra	Custo venda
0%	0,135%	0,385% + aluguel
50%	0,385%	0,635% + aluguel
100%	0,64%	0,89% + aluguel

Fonte: Autor

3.8 APURAÇÃO DE RESULTADO E INDICADORES

O estudo utiliza da estratégia *cash neutral*, onde o volume financeiro da venda será tão próximo quanto o volume de compra, de modo que não há desembolso de capital visto que o volume de venda financia o volume de compra. Entretanto, a bolsa brasileira exige do investidor ativos em garantia, tais como títulos públicos de renda fixa, ouro ativo financeiro, dólar, ações de companhia aberta admitida à negociação na B3, ETF (*Exchange Traded Fund*), entre outros. Logo, será considerado que há um volume financeiro invariável, suficiente para cobrir a garantia de todas as operações, aplicado como garantia. Apesar de ser possível mensurar o desempenho da estratégia somada a uma taxa livre de risco, visto que a abordagem *cash neutral* não precisa de desembolso financeiro, ou seja, o capital pode estar alocado em uma renda fixa, este estudo focou em utilizar apenas o desempenho das estratégias de *pair trading*.

O retorno de um par é apurado com base na diferença entre os preços compra e venda de cada um dos ativos, com incidência dos custos no preço médio das compras e vendas, apurados da seguinte forma:

$$\text{Entrada da operação} \begin{cases} \text{Compra X: } P_x = \text{Preço compra} * (1 + \text{custos}) & (c) \\ \text{Venda Y: } P_y = \text{Preço venda} * (1 - \text{custos}) & (d) \end{cases}$$

$$\text{Saida da operação} \begin{cases} \text{Venda X: } P_x = \text{Preço venda} * (1 - \text{custos}) & (e) \\ \text{Compra Y: } P_y = \text{Preço compra} * (1 + \text{custos}) & (f) \end{cases}$$

Como as operações são feitas com base no mesmo volume financeiro, sem reinvestimento do resultado da operação, o resultado diário é apurado como a soma ponderada do retorno de todas as operações abertas e fechadas:

$$R_{n,t} = \sum_{n=1}^N w_{n,t-1} * r_{n,t} \quad (3)$$

As variáveis são definidas a seguir:

n = par operado

t = dia analisado

N = quantidade de pares abertos e fechados no dia

w = peso do volume financeiro do par no portfólio

r = retorno do par

Os indicadores utilizados para comparar o desempenho das estratégias serão o índice de informação (*information ratio*), assim como utilizado por Caldeira e Moura (2013) e Pontuschka e Perlin (2015), com a volatilidade calculada com base na mesma metodologia da B3:

$$IR = \frac{\widehat{R}_t}{\widehat{\sigma}} \quad (4)$$

$$\hat{\sigma} = \sqrt{252} * DP(\ln R_t) \quad (5)$$

Capítulo 4

4. RESULTADOS

Foram analisados os dados diários das ações que compõem a carteira teórica do Índice Bovespa entre janeiro de 2019 e junho de 2022. Os períodos de formação para cálculo da cointegração foram de 12 meses para os retornos, calculados diariamente de forma contínua ao longo do tempo. Após agrupar os ativos por setor e encontrar os pares cointegrados por meio do procedimento Engle e Granger (1987), foram separados apenas os que apresentaram 1% de significância no teste de *Dickey-Fuller* aumentado e, antes de entrar no par, avaliou-se a relação de correlação entre eles. Todos os testes e operações foram realizados por meio de um código próprio elaborado na linguagem de programação *python* juntamente com suporte de uma API da plataforma de *trading* MetaTrader 5.

Ao todo 687 operações foram executadas, entre diferentes níveis de *VaR*, e sem limitação de pares em abertos simultaneamente, a fim de observar, na média, a efetividade da estratégia durante todo horizonte de estudo. Observou-se o total de 88 pares distribuídos em 15 setores, além do Índice Bovespa, as estatísticas do retorno estão descritas na Tabela 5:

TABELA 5: ESTATÍSTICA DESCRITIVA DA ESTRATÉGIA COM 1% ADF

Estatística	Resultado - 1% <i>VaR</i>	Resultado - sem <i>VaR</i>	Ibovespa
Observações	765	765	765
Média	0,5486	0,4800	0,1664
Desvio Padrão	0,5466	0,4863	0,1318

Menor Valor	-0,1768	-0,1329	-0,3015
Maior Valor	1,8963	1,8659	0,4369
Q1	0,0178	0,0327	0,0812
Q2	0,5468	0,3638	0,1629
Q3	0,9899	0,8475	0,2655
Teste Normalidade (KS)	0,4382	0,4674	0,4314
Normalidade (p-valor)	< 0,0001	< 0,0001	< 0,0001

Fonte: Autor

Nota-se que no período o Ibovespa obteve média e desvio padrão inferiores as carteiras com e sem gerenciamento de risco, apesar de um menor valor mais baixo que os demais. Curiosamente, a média e desvio padrão da carteira com *VaR* apresentaram maiores valores do que a sem *VaR*. Os retornos brutos em todas as amostras não apresentam distribuição normal pelo teste de Kolmogorov-Smirnov.

O desempenho da carteira com o *VaR* de 1% é apresentado na Tabela 6 para as múltiplas estruturas de custos:

TABELA 6: DESEMPENHO ACUMULADO (COM GERENCIAMENTO DE RISCO)

Carteira	Retorno Líquido	Desvio Padrão	Índice de Informação	Correlação Ibovespa
Baixo Custo	0,1514	0,3232	0,4683	0,2808
Médio Custo	-0,3849	0,3227	-1,1943	-0,2326
Alto Custo	-0,6751	0,3231	-2,0894	-0,2999
Ibovespa	0,0968	0,2937	0,3296	1

Fonte: Autor

Observa-se que apenas a carteira do investidor que possui a menor estrutura de custos obteve maior índice de informação e desempenhou melhor que o Índice Bovespa, além de apresentar baixa correlação com o índice. Tais resultados são demonstrados na Figura 3:

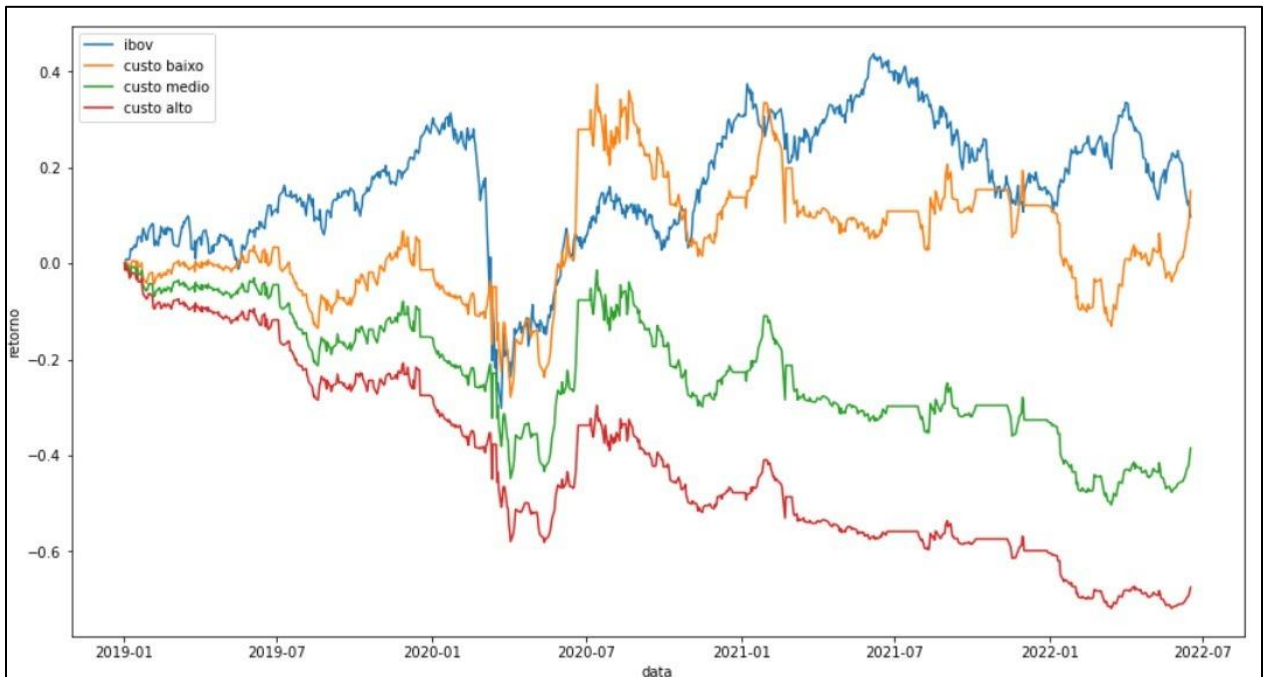


Figura 3: Desempenho das carteiras de diferentes tipos de investidores (com gerenciamento de risco)
Fonte: Autor

Pelo gráfico é exposto a rapidez com que o Índice Bovespa perde valor nos primeiros meses de 2020, enquanto na carteira de *pair trading* não apresenta queda acentuada de mesma magnitude nesse período. A literatura de Kirilenko *et al.* (2017) sobre *HFT* dão indícios sobre a maior intensidade no movimento direcional em momentos de maior estresse. Quanto ao comportamento dos retornos, todos apresentaram um movimento similar de queda e retomada até o início de 2021. Como no movimento de queda acentuada no início de 2020 o *VaR* pode ter interrompido no prejuízo muitas operações abertas, efetuou-se um novo teste sem nenhum parâmetro de gerenciamento de risco, portanto o investidor só saíra da operação quando obtiver lucro:

TABELA 7: DESEMPENHO ACUMULADO (SEM GERENCIAMENTO DE RISCO)

Carteira	Retorno Líquido	Desvio Padrão	Índice de Informação	Correlação Ibovespa
Baixo Custo	1,8659	0,2730	6,8343	0,4546
Médio Custo	0,9784	0,2721	3,5959	0,4676
Alto Custo	0,3568	0,2721	1,3114	0,5240
Ibovespa	0,0968	0,2938	0,3296	1

Fonte: Autor

No cenário sem o gerenciamento de risco, houve um aumento considerável na correlação das carteiras com o Ibovespa, além de redução no desvio padrão e aumento nos índices de informação e retorno líquido. Tal resultado sugere que mesmo um VaR com 1% de significância estatística não foi um bom parâmetro de gerenciamento para o período de maior volatilidade. Rad *et al.* (2016) e Cavalcanti *et al.* (2021) obtiveram retornos positivos também sem utilizar parâmetros de gerenciamento de risco para as operações em andamento do portfólio. O gráfico dos retornos é exposto na Figura 4:

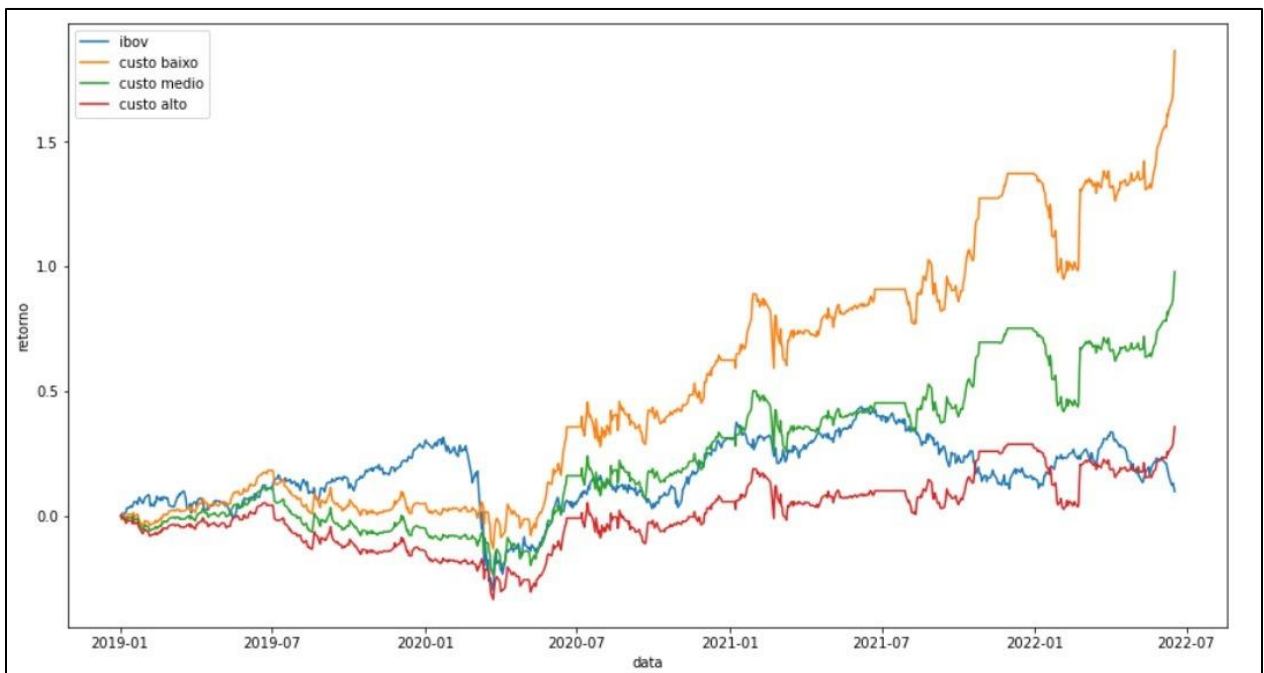


Figura 4: Desempenho das carteiras de diferentes tipos de investidores (sem gerenciamento de risco)
Fonte: Autor

Algo que chama atenção nas Figuras 3 e 4 é o fato do desempenho das carteiras de *pair trading* acumularem retornos positivos apenas após a retomada do Índice Bovespa e não no momento que houve aumento da volatilidade. É curioso notar que durante os meses de março e abril, períodos de volatilidade mais altas de 2020, a carteira oscilou entre o campo neutro ou negativo, com maior sequência de valorização apenas

a partir de maio, mês em que a volatilidade já era decrescente, conforme visto na figura 2 e exposto na Tabela 3. Também é possível observar que em meses de volatilidade menor, não necessariamente há uma piora no desempenho da estratégia, portanto não é possível afirmar que a volatilidade agregou desempenho a carteira para o período pós pandemia. Uma intuição para o aumento do desempenho retoma o estudo de Thaler (2016), em que vieses psicológicos e crenças violam expectativas racionais em momentos de euforia ou aumento de confiança, fatores esses que podem ter desencadeado distorções temporárias entre preços em um ambiente com sequencias retornos positivos no índice e as primeiras notícias relacionadas as vacinas para o vírus.

Como a retirada do *VaR* agregou desempenho a estratégia, buscou-se avaliar diferentes significâncias estatísticas de *VaR* para medir seu impacto na eficácia da estratégia. Para as tabelas seguintes será considerada a estrutura do investidor com menor custo:

TABELA 8: EFICÁCIA ESTRATÉGIA COINTEGRAÇÃO POR SEMESTRE

Período	Número de Operações				Taxa de Acerto				
	<i>VaR</i>	0%	1%	5%	10%	0%	1%	5%	10%
01/2019 - 06/2019		34	38	44	53	91%	84%	73%	60%
07/2019 - 12/2019		11	12	17	18	82%	83%	65%	67%
01/2020 - 06/2020		25	36	39	46	84%	72%	49%	50%
07/2020 - 12/2020		12	14	20	23	83%	71%	55%	52%
01/2021 - 06/2021		17	18	22	24	88%	89%	64%	54%
07/2021 - 12/2021		5	6	8	11	80%	67%	75%	55%
01/2022 - 06/2022		26	33	37	38	85%	73%	62%	55%

Fonte: Autor

Observa-se uma relação direta entre o número de operações realizadas, tamanho do limite de perda e taxa de acerto. Como o *VaR* com 5% e 10% de significância estatística alcança seu limite de perda com mais frequência que 1%, tem-se um menor tempo médio por operação, maior número de operações e, conseqüentemente, maior

nível de custos. A taxa de acerto sem gerenciamento de risco surpreendeu desde 2019, antes do aumento da volatilidade causado pela pandemia, e se manteve em elevados períodos por todo o estudo.

TABELA 9: TEMPO EM OPERAÇÃO - COINTEGRAÇÃO POR SEMESTRE

Período	Tempo médio por operação (d.u.)				Número de stops pelo VaR				
	VaR	0%	1%	5%	10%	0%	1%	5%	10%
01/2019 - 06/2019		23	21	16	10	-	8	28	43
07/2019 - 12/2019		37	23	15	10	-	3	10	15
01/2020 - 06/2020		23	13	7	5	-	20	29	40
07/2020 - 12/2020		20	20	19	14	-	2	9	17
01/2021 - 06/2021		24	21	20	14	-	2	11	17
07/2021 - 12/2021		21	23	11	10	-	1	6	10
01/2022 - 06/2022		21	14	10	9	-	14	24	29

Fonte: Autor

Como o uso do VaR prejudicou o desempenho independente da significância adotada, este mostrou-se ineficaz como parâmetro de gerenciamento de risco. Lin et al. (2006), Puspanimgrum et al. (2010) e Chen *et al.* (2017) tiveram retornos positivos ao utilizar bandas de saída no escore padrão como medida de gerenciamento de risco, entretanto não foi feita uma comparação de como seriam os resultados sem esse critério.

A Tabela 10 expõe os setores com maior número de operações e sua respectiva taxa de acerto ao longo de todo o estudo:

TABELA 10: TAXA DE ACERTO POR SETOR (SEM GERENCIAMENTO DE RISCO)

Setor	Operações	Taxa de acerto	Ganho médio por operação	Retorno líquido acumulado
Financeiro	49	87,76%	2,75%	134,78%
Utilidade Pública	33	87,88%	2,20%	72,51%
Saúde	11	81,82%	2,11%	23,23%
Petróleo, gás e biocombustíveis	10	80,00%	0,62%	6,16%
Mineração e Siderurgia	7	100,00%	2,76%	19,32%
Varejo	6	83,33%	0,80%	4,79%
Papel e Celulose	3	100,00%	0,38%	1,15%
Alimentos e Bebidas	2	50,00%	0,97%	1,94%
Bens Industriais	2	100,00%	5,20%	10,39%

Construção Civil	2	0,00%	0,56%	1,11%
Exploração de imóveis	2	100,00%	0,76%	1,53%
Telecomunicações	2	100,00%	0,24%	0,48%
Aéreas	1	100,00%	2,46%	2,46%

Fonte: Autor

Enquanto os 15 pares de ativos com maior número de operações são discriminados na Tabela 11:

TABELA 11: ATIVOS MAIS OPERADOS (SEM GERENCIAMENTO DE RISCO)

Par Operado	Operações	Taxa de acerto	Tempo médio por operação	Retorno líquido acumulado
PETR3 x PETR4	9	89%	15	9,00%
EQTL3 x TAEE11	6	100%	21	39,03%
BBAS3 x BBSE3	4	100%	17	21,44%
CMIG4 x TAEE11	4	100%	37	8,43%
B3SA3 x BBDC4	3	100%	14	7,73%
BBDC3 x BBDC4	3	100%	16	2,18%
BBDC3 x BBSE3	3	100%	14	16,50%
BBSE3 x ITSA4	3	100%	13	18,39%
BRAP4 x GGBR4	3	100%	5	20,84%
EGIE3 x SBSP3	3	100%	7	25,33%
FLRY3 x QUAL3	3	100%	10	35,69%
GGBR4 x GOAU4	3	100%	17	-0,79%
KLBN11 x SUZB3	3	100%	6	15,61%
SBSP3 x TAEE11	3	100%	32	7,19%
CPFE3 x EGIE3	3	67%	13	1,95%

Fonte: Autor

Como o rápido aumento da volatilidade em março de 2020 não demonstrou indícios de benefício ou mudanças no comportamento da estratégia para os preços diários nos meses subsequentes, utilizou-se uma maior frequência de dados, com 15 minutos para o intervalo dos preços, a fim de analisar a estratégia e o período sob outra perspectiva. O comportamento do retorno chama atenção a partir de 2020:

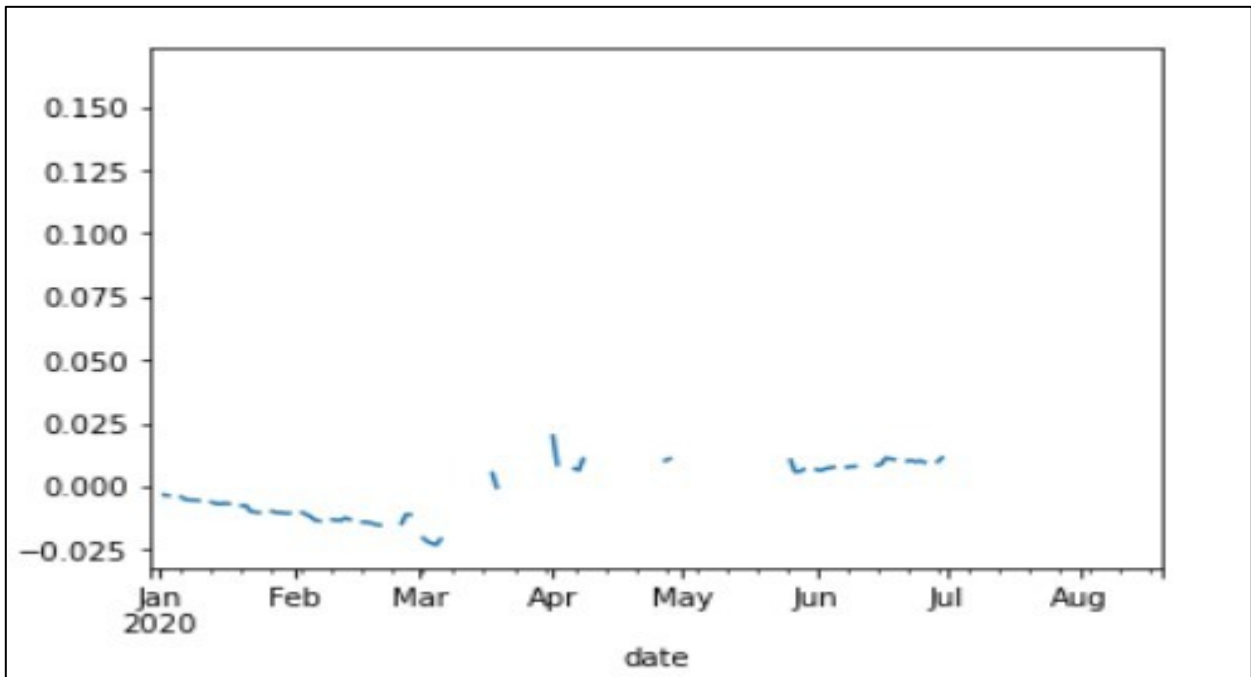


Figura 5: Desempenho intradiário 15 minutos com VaR de 1%
Fonte: Autor

Um comportamento semelhante foi observado após retirar o *VaR*. Por esse motivo, observou-se os retornos da frequência de 5 minutos já sem o *VaR*:

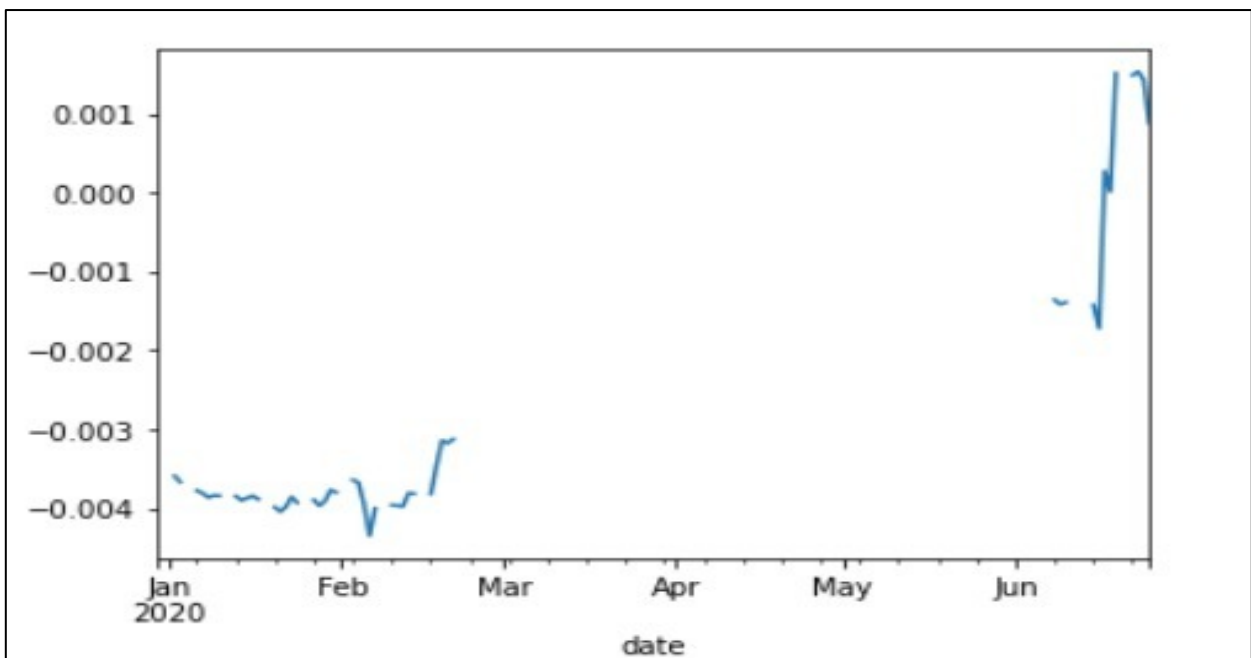


Figura 6: Desempenho intradiário 5 minutos sem VaR
Fonte: Autor

Nota-se que, apesar do aumento da volatilidade e nenhuma diferença perceptível no comportamento dos preços diários, no intradiário tiveram poucas ou nenhuma oportunidade de entrada entre março de junho de 2020. Foi elaborada uma estatística descritiva específica para as duas frequências com apenas os dados que apresentaram retornos:

TABELA 12: ESTATÍSTICA DESCRITIVA DADOS INTRADIÁRIOS

Estatística	15 minutos – com VaR	15 minutos – sem VaR	5 minutos – com VaR	5 minutos – sem VaR
Observações	341	166	191	45
Média	0,0019	0,0073	-0,0010	-0,0005
Desvio Padrão	0,0144	0,0655	0,0047	0,0117
Menor Valor	-0,0414	-0,1210	-0,0150	-0,0216
Maior Valor	0,1513	0,7925	0,0150	-0,0034
Q1	-0,0086	-0,0009	-0,0054	-0,0012
Q2	0,0035	0,0014	-0,0002	0,0013
Q3	0,0070	0,0074	0,0019	0,0609

Fonte: Autor

Os poucos dados encontrados para a frequência intradiária em um momento de volatilidade muito acima da média sugerem que, como a janela de formação de pares é mais curta, as relações de cointegração se comprometeram, tanto que, conforme houve inserção de novos dados, o número de operações volta a aparecer. Dado a impossibilidade de aprofundar os resultados intradiários para este período, utilizou-se os dados diários para março, abril de maio de 2020. O resultado é exposto a seguir:

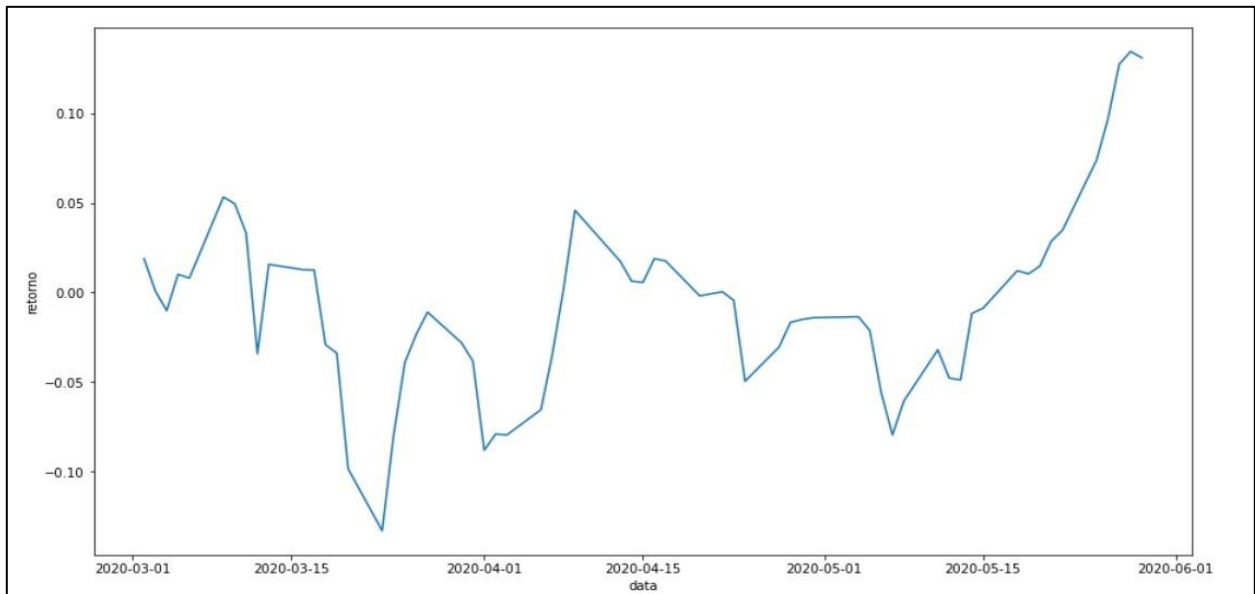


Figura 7: Desempenho dos dados diários da carteira de pair trading no auge da volatilidade
Fonte: Autor

A relação do desempenho por setor é exposta a seguir:

TABELA 13: TAXA DE ACERTO POR SETOR - DADOS DIÁRIOS NO AUGE DA VOLATILIDADE (SEM GERENCIAMENTO DE RISCO)

Setor	Operações	Taxa de acerto	Ganho médio por operação	Retorno líquido acumulado
Mineração e Siderurgia	3	100%	5,33%	16,00%
Saúde	3	67%	1,52%	4,55%
Utilidade Pública	3	100%	2,89%	8,67%
Alimentos e Bebidas	1	0%	2,98%	2,98%
Bens Industriais	1	100%	9,29%	9,29%
Financeiro	1	100%	-3,40%	-3,40%
Papel e Celulose	1	100%	1,81%	1,81%
Varejo	1	100%	1,40%	1,40%

Fonte: Autor

A alta taxa de acerto e o resultado líquido indicam que para os dados diários a relação de longo prazo entre os pares se manteve e a estratégia de cointegração não foi afetada. O desempenho por ativo é exposto na Tabela 14:

TABELA 14: ATIVOS MAIS OPERADOS - DADOS DIÁRIOS NO AUGE DA VOLATILIDADE (SEM GERENCIAMENTO DE RISCO)

Par Operado	Operações	Taxa de acerto	Tempo médio por operação	Retorno líquido acumulado
BRAP4 x GGBR4	3	100%	5	20,84%

EQTL3 x TAE11	2	100%	19	25,87%
HYPE3 x RADL3	2	100%	8	21,50%
AMER3 x MGLU3	1	100%	21	0,50%
BBAS3 x BBDC4	1	100%	31	0,38%
CCRO3 x RAIL3	1	100%	42	18,93%
KLBN11 x SUZB3	1	100%	3	7,29%
SBSP3 x TAE11	1	100%	53	2,89%
BRFS3 x JBSS3	1	0%	83	-10,31%
HYPE3 x QUAL3	1	0%	3	-2,91%

Fonte: Autor

Capítulo 5

5. CONCLUSÃO

O presente estudo aborda as operações de *pair trading* no mercado acionário brasileiro sob a ótica da cointegração no período do Covid-19. Apesar de ser uma estratégia validada na literatura, o choque de volatilidade ocorrido em março de 2020, somado a mudanças nos padrões de consumo e na realidade das empresas por conta dos *lockdowns*, podem afetar modelos econométricos baseados em séries temporais, como o caso da cointegração. Portanto, foi realizada uma simulação histórica dos anos em torno do início da pandemia, em 2020, e observou-se pares de ativos correlacionados com características estatísticas de equilíbrio. Ao ocorrer distorções de magnitude suficiente para afastar os pares do equilíbrio, abriu-se uma posição no par sob a crença de reversão dos pares de volta ao equilíbrio. Frequência de dados intradiários foi utilizada para auxiliar na análise dos retornos em torno de março de 2020, onde a volatilidade superou cinco vezes a média dos últimos onze anos. Este trabalho aproveitou para evidenciar o impacto das diferentes estruturas de custos presentes no mercado brasileiro, tema pouco explorado na literatura local de *pair trading* e bastante relevante por se tratar de uma das operações mais custosas para realizar com ações.

Não foi possível com os resultados do trabalho associar o aumento da volatilidade a um melhor desempenho da carteira de *pair trading* para os dados diários, visto que a alta expressiva do desempenho teve início apenas dois meses após o pico da

volatilidade. Constatou-se também que períodos com volatilidade menor não necessariamente tiveram desempenhos superiores a períodos de maiores volatilidades. Por outro lado, para os meses seguintes a março de 2020, mesmo a estratégia com estrutura mais pesada de custo apresentou um índice de informação na ordem de quatro vezes o desempenhado pelo Ibovespa, além de uma alta taxa de acerto nas operações. Portanto, o resultado segue em linha de estudos de cointegração aplicadas ao *pair trading* e corroboram com Rad *et al.* (2016) no que tange a efetividade da estratégia mesmo em um período de crise.

Apesar das operações na frequência intradiária de dados apresentarem uma redução drástica de pares operados, a curta janela do teste de cointegração realizada (50 dias com dados intradiários) possivelmente afetou o teste a ser analisado, visto que o número de operações cessou quando se observou a frequência de 5 minutos e cresceu com dados diários de outra janela de formação. Para estudos com objetivo de testar melhor desempenho de estratégias, a elaboração de um sistema de análise de cointegração com testes em dados diários mais longos e entradas na estratégia pelo score padrão com dados intradiários tem potencial de gerar resultados interessantes, assim como em Pontuschka e Perlin (2015).

Além disso, os resultados do trabalho sugerem que o uso do *VaR* não parametrizado não foi uma boa escolha como medida de gerenciamento de risco da estratégia, mesmo para períodos de menor volatilidade, enquanto a estratégia analisada de forma contínua no tempo sem limites de perdas teve um resultado superior ao índice de referência. O uso de bandas de saídas com base no score padrão, como feito em Lin *et al.* (2006), Puspanimgrum *et al.* (2010) e Chen *et al.* (2017), aplicados em diferentes

ciclos econômicos de forma intrassetorial e comparados a estratégia sem limites de perdas, é um possível estudo futuro que pode ajudar a entender em quais momentos o gerenciamento de risco aplicado ao *pair trading* pode diminuir o risco e aumentar o índice de informação da carteira. As tabelas de desempenho por setor e ativo deste trabalho podem auxiliar no maior refino da divisão setorial, o que mitigaria ainda mais os riscos e poderiam colaborar para uma maior taxa de acerto e desempenho nesses diferentes períodos.

REFERÊNCIAS

- Abad, P., Benito, S., & López, C. (2014). A comprehensive review of Value at Risk methodologies. *The Spanish Review of Financial Economics*, 12(1), 15-32.
- Alexander, C., & Dimitriu, A. (2005). Indexing and statistical arbitrage. *The Journal of Portfolio Management*, 31(2), 50-63.
- Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1, 1053-1128.
- Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, 49(3), 307-343.
- Benos, E., & Sagade, S. (2012). High-frequency trading behaviour and its impact on market quality: evidence from the UK equity market (No. 469). London: Bank of England.
- Caldeira, J. F. (2013). Arbitragem Estatística, Estratégia Long-Short Pairs Trading, Abordagem com Cointegração Aplicada ao Mercado de Ações Brasileiro. *Revista Economia*, Brasília, 14(1), 521-546.
- Cavalcanti, R. S. G., Santos, J. F. D., Santos, R. R. D., & Cunha, A. G. M. D. (2021). Composição de portfólios por pairs trading com critério de volatilidade no mercado brasileiro. *Revista Contabilidade & Finanças*, 32, 273-284.
- Chan, Ernest P. (2009). *Quantitative Trading*. Wiley Trading. John Wiley & Sons, Inc.
- Chen, D., Cui, J., Gao, Y., & Wu, L. (2017). Pairs trading in Chinese commodity futures markets: an adaptive cointegration approach. *Accounting & Finance*, 57(5), 1237-1264.
- Chen, P. (2018). Understanding international stock market comovements: A comparison of developed and emerging markets. *International Review of Economics & Finance*, 56, 451-464.
- Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance*, 53(6), 1839-1885.
- De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact?. *The Journal of finance*, 40(3), 793-805.
- De Wet, W. D. (2004). The role of asymmetric information on investments in emerging markets. *Economic Modelling*, 21(4), 621-630.
- Dewan, S., & Mendelson, H. (1998). Information technology and time-based competition in financial markets. *Management Science*, 44(5), 595-609.

- Do, B., & Faff, R. (2012). Are pairs trading profits robust to trading costs?. *Journal of Financial Research*, 35(2), 261-287.
- Do, Binh; Faff, Robert. Does simple pairs trading still work?. *Financial Analysts Journal*, v. 66, n. 4, p. 83-95, 2010.
- Ehrman, D. S. (2006). *The handbook of pairs trading: strategies using equities, options, and futures* (Vol. 240). John Wiley & Sons.
- Engle, R., & Granger, C. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251-276.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.
- Friedman, M. (1953). *Essays in positive economics*. University of Chicago press.
- Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (1999). Pairs Trading: Performance of a Relative Arbitrage Rule. National Bureau of Economic Research Inc, NBER Working Papers 7032.
- Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, 19(3), 797-827.
- Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance*, 54(6), 2143-2184.
- Huck, N. (2010). Pairs trading and outranking: The multi-step-ahead forecasting case. *European Journal of Operational Research*, 207(3), 1702-1716.
- Huck, N., & Afawubo, K. (2015). Pairs trading and selection methods: is cointegration superior?. *Applied Economics*, 47(6), 599-613.
- Jacobs, H., & Weber, M. (2015). On the determinants of pairs trading profitability. *Journal of Financial Markets*, 23, 75-97.
- Kabashima, E. M. (2006). *Pairs Trading: Aplicação no mercado acionário brasileiro* (Doctoral dissertation).
- Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72(3), 967-998.
- Krauss, C. (2017). Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, 31(2), 513-545.
- Lin, Y. X., McCrae, M., & Gulati, C. (2006). Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *Advances in Decision Sciences*, 2006.
- Mazur, M., Dang, M., & Vega, M. (2021). COVID-19 and the march 2020 stock market crash. Evidence from S&P1500. *Finance research letters*, 38, 101690.

- Miao, G. J. (2014). High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. *International Journal of Economics and Finance*, 6(3), 96-110.
- Perlin, M. S. (2009). Evaluation of pairs-trading strategy at the Brazilian financial market. *Journal of Derivatives & Hedge Funds*, 15(2), 122-136.
- Pontuschka, M., & Perlin, M. (2015). A estratégia de pares no mercado acionário brasileiro: O Impacto da frequência de dados. *RAM. Revista de Administração Mackenzie*, 16, 188-213.
- Puspaningrum, H., Lin, Y. X., & Gulati, C. M. (2010). Finding the optimal pre-set boundaries for pairs trading strategy based on cointegration technique. *Journal of Statistical Theory and Practice*, 4(3), 391-419.
- Rad, H., Low, R. K. Y., & Faff, R. (2016). The profitability of pairs trading strategies: distance, cointegration and copula methods. *Quantitative Finance*, 16(10), 1541-1558
- Shefrin, H., & Statman, M. (1994). Behavioral capital asset pricing theory. *Journal of financial and quantitative analysis*, 29(3), 323-349.
- Thaler, R. H. (2016). Behavioral economics: Past, present, and future. *American economic review*, 106(7), 1577-1600.
- Vidyamurthy, G. (2004). *Pairs Trading: quantitative methods and analysis* (Vol. 217). John Wiley & Sons.
- Whistler, M. (2004). *Trading pairs: capturing profits and hedging risk with statistical arbitrage strategies* (Vol. 216). John Wiley & Sons.

APÊNDICE A – CARTEIRAS TEÓRICAS POR PERÍODO E SETOR

1º semestre 2019		2º semestre 2019		1º semestre 2020		2º semestre 2020	
Ativo	Setor	Ativo	Setor	Ativo	Setor	Ativo	Setor
ABEV3	Alimentos e Bebidas	AZUL4	Aéreas	AZUL4	Aéreas	AZUL4	Aéreas
BRFS3	Alimentos e Bebidas	GOLL4	Aéreas	GOLL4	Aéreas	GOLL4	Aéreas
JBSS3	Alimentos e Bebidas	ABEV3	Alimentos e Bebidas	ABEV3	Alimentos e Bebidas	ABEV3	Alimentos e Bebidas
MRFG3	Alimentos e Bebidas	BRFS3	Alimentos e Bebidas	BRFS3	Alimentos e Bebidas	BEEF3	Alimentos e Bebidas
CCRO3	Bens Industriais	JBSS3	Alimentos e Bebidas	JBSS3	Alimentos e Bebidas	BRFS3	Alimentos e Bebidas
ECOR3	Bens Industriais	MRFG3	Alimentos e Bebidas	MRFG3	Alimentos e Bebidas	JBSS3	Alimentos e Bebidas
EMBR3	Bens Industriais	CCRO3	Bens Industriais	CCRO3	Bens Industriais	MRFG3	Alimentos e Bebidas
RAIL3	Bens Industriais	ECOR3	Bens Industriais	ECOR3	Bens Industriais	CCRO3	Bens Industriais
WEGE3	Bens Industriais	EMBR3	Bens Industriais	EMBR3	Bens Industriais	ECOR3	Bens Industriais
CYRE3	Construção Civil	RAIL3	Bens Industriais	RAIL3	Bens Industriais	EMBR3	Bens Industriais
MRVE3	Construção Civil	WEGE3	Bens Industriais	WEGE3	Bens Industriais	RAIL3	Bens Industriais
COGN3	Educação	CYRE3	Construção Civil	CYRE3	Construção Civil	WEGE3	Bens Industriais
YDUQ3	Educação	MRVE3	Construção Civil	JHSF3	Construção Civil	CYRE3	Construção Civil
BRML3	Exploração de imóveis	COGN3	Educação	MRVE3	Construção Civil	JHSF3	Construção Civil
IGT13	Exploração de imóveis	YDUQ3	Educação	COGN3	Educação	MRVE3	Construção Civil
LOGG3	Exploração de imóveis	BRML3	Exploração de imóveis	YDUQ3	Educação	COGN3	Educação
MULT3	Exploração de imóveis	IGT13	Exploração de imóveis	BRML3	Exploração de imóveis	YDUQ3	Educação
B3SA3	Financeiro	MULT3	Exploração de imóveis	IGT13	Exploração de imóveis	BRML3	Exploração de imóveis
BBAS3	Financeiro	B3SA3	Financeiro	MULT3	Exploração de imóveis	IGT13	Exploração de imóveis
BBDC3	Financeiro	BBAS3	Financeiro	B3SA3	Financeiro	MULT3	Exploração de imóveis
BBDC4	Financeiro	BBDC3	Financeiro	BBAS3	Financeiro	B3SA3	Financeiro
BBSE3	Financeiro	BBDC4	Financeiro	BBDC3	Financeiro	BBAS3	Financeiro
CIEL3	Financeiro	BBSE3	Financeiro	BBDC4	Financeiro	BBDC3	Financeiro
ITSA4	Financeiro	CIEL3	Financeiro	BBSE3	Financeiro	BBDC4	Financeiro
ITUB4	Financeiro	IRBR3	Financeiro	BPAC11	Financeiro	BBSE3	Financeiro
SANB11	Financeiro	ITSA4	Financeiro	CIEL3	Financeiro	BPAC11	Financeiro
BRAP4	Mineração e Siderurgia	ITUB4	Financeiro	IRBR3	Financeiro	CIEL3	Financeiro
CSNA3	Mineração e Siderurgia	SANB11	Financeiro	ITSA4	Financeiro	IRBR3	Financeiro
GGBR4	Mineração e Siderurgia	BRAP4	Mineração e Siderurgia	ITUB4	Financeiro	ITSA4	Financeiro
GOAU4	Mineração e Siderurgia	CSNA3	Mineração e Siderurgia	SANB11	Financeiro	ITUB4	Financeiro
USIM5	Mineração e Siderurgia	GGBR4	Mineração e Siderurgia	SULA11	Financeiro	SANB11	Financeiro
VALE3	Mineração e Siderurgia	GOAU4	Mineração e Siderurgia	BRAP4	Mineração e Siderurgia	SULA11	Financeiro
KLBN11	Papel e Celulose	USIM5	Mineração e Siderurgia	CSNA3	Mineração e Siderurgia	BRAP4	Mineração e Siderurgia
SUZB3	Papel e Celulose	VALE3	Mineração e Siderurgia	GGBR4	Mineração e Siderurgia	CSNA3	Mineração e Siderurgia
CSAN3	Petroleo, gás e biocombustíveis	KLBN11	Papel e Celulose	GOAU4	Mineração e Siderurgia	GGBR4	Mineração e Siderurgia
PETR3	Petroleo, gás e biocombustíveis	SUZB3	Papel e Celulose	USIM5	Mineração e Siderurgia	GOAU4	Mineração e Siderurgia
PETRA	Petroleo, gás e biocombustíveis	CSAN3	Petroleo, gás e biocombustíveis	VALE3	Mineração e Siderurgia	USIM5	Mineração e Siderurgia
UGPA3	Petroleo, gás e biocombustíveis	PETR3	Petroleo, gás e biocombustíveis	KLBN11	Papel e Celulose	VALE3	Mineração e Siderurgia
VBBR3	Petroleo, gás e biocombustíveis	PETRA	Petroleo, gás e biocombustíveis	SUZB3	Papel e Celulose	KLBN11	Papel e Celulose
FLRY3	Saúde	UGPA3	Petroleo, gás e biocombustíveis	CSAN3	Petroleo, gás e biocombustíveis	SUZB3	Papel e Celulose
HYPE3	Saúde	VBBR3	Petroleo, gás e biocombustíveis	PETR3	Petroleo, gás e biocombustíveis	CSAN3	Petroleo, gás e biocombustíveis
QUAL3	Saúde	FLRY3	Saúde	PETRA	Petroleo, gás e biocombustíveis	PETR3	Petroleo, gás e biocombustíveis
RADL3	Saúde	HYPE3	Saúde	UGPA3	Petroleo, gás e biocombustíveis	PETRA	Petroleo, gás e biocombustíveis
CMIG4	Utilidade Pública	QUAL3	Saúde	VBBR3	Petroleo, gás e biocombustíveis	UGPA3	Petroleo, gás e biocombustíveis
EGIE3	Utilidade Pública	RADL3	Saúde	FLRY3	Saúde	VBBR3	Petroleo, gás e biocombustíveis
ELET3	Utilidade Pública	CMIG4	Utilidade Pública	HAPV3	Saúde	FLRY3	Saúde
ELET6	Utilidade Pública	EGIE3	Utilidade Pública	HYPE3	Saúde	HAPV3	Saúde
ENBR3	Utilidade Pública	ELET3	Utilidade Pública	QUAL3	Saúde	HYPE3	Saúde
EQTL3	Utilidade Pública	ELET6	Utilidade Pública	RADL3	Saúde	QUAL3	Saúde
SBSP3	Utilidade Pública	ENBR3	Utilidade Pública	CMIG4	Utilidade Pública	RADL3	Saúde
TAEF11	Utilidade Pública	EQTL3	Utilidade Pública	CPLE6	Utilidade Pública	TIMS3	Telecomunicações
AMER3	Varejo	SBSP3	Utilidade Pública	EGIE3	Utilidade Pública	VIVT3	Telecomunicações
CVCB3	Varejo	TAEF11	Utilidade Pública	ELET3	Utilidade Pública	CMIG4	Utilidade Pública
LREN3	Varejo	AMER3	Varejo	ELET6	Utilidade Pública	CPFE3	Utilidade Pública
MGLU3	Varejo	CVCB3	Varejo	ENBR3	Utilidade Pública	CPLE6	Utilidade Pública
NTCO3	Varejo	LREN3	Varejo	ENEV3	Utilidade Pública	EGIE3	Utilidade Pública
VIIA3	Varejo	MGLU3	Varejo	EQTL3	Utilidade Pública	ELET3	Utilidade Pública
		NTCO3	Varejo	SBSP3	Utilidade Pública	ELET6	Utilidade Pública
		VIIA3	Varejo	TAEF11	Utilidade Pública	ENBR3	Utilidade Pública
				AMER3	Varejo	ENEV3	Utilidade Pública
				CRFB3	Varejo	ENGI11	Utilidade Pública
				CVCB3	Varejo	EQTL3	Utilidade Pública
				LREN3	Varejo	SBSP3	Utilidade Pública
				MGLU3	Varejo	TAEF11	Utilidade Pública
				NTCO3	Varejo	AMER3	Varejo
				VIIA3	Varejo	CRFB3	Varejo
						CVCB3	Varejo
						LREN3	Varejo
						MGLU3	Varejo
						NTCO3	Varejo
						PCAR3	Varejo
						VIIA3	Varejo

1º semestre 2021		2º semestre 2021		1º semestre 2022	
Ativo	Setor	Ativo	Setor	Ativo	Setor
AZUL4	Aéreas	AZUL4	Aéreas	AZUL4	Aéreas
GOLL4	Aéreas	GOLL4	Aéreas	GOLL4	Aéreas
ABEV3	Alimentos e Bebidas	ABEV3	Alimentos e Bebidas	ABEV3	Alimentos e Bebidas
BEEF3	Alimentos e Bebidas	BEEF3	Alimentos e Bebidas	BEEF3	Alimentos e Bebidas
BRFS3	Alimentos e Bebidas	BRFS3	Alimentos e Bebidas	BRFS3	Alimentos e Bebidas
JBSS3	Alimentos e Bebidas	JBSS3	Alimentos e Bebidas	JBSS3	Alimentos e Bebidas
MRFG3	Alimentos e Bebidas	MRFG3	Alimentos e Bebidas	MRFG3	Alimentos e Bebidas
CCRO3	Bens Industriais	CCRO3	Bens Industriais	CCRO3	Bens Industriais
ECOR3	Bens Industriais	ECOR3	Bens Industriais	ECOR3	Bens Industriais
EMBR3	Bens Industriais	EMBR3	Bens Industriais	EMBR3	Bens Industriais
RAIL3	Bens Industriais	RAIL3	Bens Industriais	RAIL3	Bens Industriais
WEGE3	Bens Industriais	WEGE3	Bens Industriais	WEGE3	Bens Industriais
CYRE3	Construção Civil	CYRE3	Construção Civil	CYRE3	Construção Civil
EZTC3	Construção Civil	EZTC3	Construção Civil	EZTC3	Construção Civil
JHSF3	Construção Civil	JHSF3	Construção Civil	JHSF3	Construção Civil
MRVE3	Construção Civil	MRVE3	Construção Civil	MRVE3	Construção Civil
COGN3	Educação	COGN3	Educação	COGN3	Educação
YDUQ3	Educação	YDUQ3	Educação	YDUQ3	Educação
BRML3	Exploração de imóveis	BRML3	Exploração de imóveis	BRML3	Exploração de imóveis
IGTI3	Exploração de imóveis	IGTI3	Exploração de imóveis	IGTI11	Exploração de imóveis
MULT3	Exploração de imóveis	MULT3	Exploração de imóveis	MULT3	Exploração de imóveis
B3SA3	Financeiro	B3SA3	Financeiro	B3SA3	Financeiro
BBAS3	Financeiro	BBAS3	Financeiro	BBAS3	Financeiro
BBDC3	Financeiro	BBDC3	Financeiro	BBDC3	Financeiro
BBDC4	Financeiro	BBDC4	Financeiro	BBDC4	Financeiro
BBSE3	Financeiro	BBSE3	Financeiro	BBSE3	Financeiro
BPAC11	Financeiro	BPAC11	Financeiro	BPAC11	Financeiro
CIEL3	Financeiro	CIEL3	Financeiro	BPAN4	Financeiro
IRBR3	Financeiro	IRBR3	Financeiro	CASH3	Financeiro
ITSA4	Financeiro	ITSA4	Financeiro	CIEL3	Financeiro
ITUB4	Financeiro	ITUB4	Financeiro	IRBR3	Financeiro
SANB11	Financeiro	SANB11	Financeiro	ITSA4	Financeiro
SULA11	Financeiro	SULA11	Financeiro	ITUB4	Financeiro
BRAP4	Mineração e Siderurgia	BRAP4	Mineração e Siderurgia	SANB11	Financeiro
CSNA3	Mineração e Siderurgia	CSNA3	Mineração e Siderurgia	SULA11	Financeiro
GGBR4	Mineração e Siderurgia	GGBR4	Mineração e Siderurgia	BRAP4	Mineração e Siderurgia
GOAU4	Mineração e Siderurgia	GOAU4	Mineração e Siderurgia	CMIN3	Mineração e Siderurgia
USIM5	Mineração e Siderurgia	USIM5	Mineração e Siderurgia	CSNA3	Mineração e Siderurgia
VALE3	Mineração e Siderurgia	VALE3	Mineração e Siderurgia	GGBR4	Mineração e Siderurgia
KLBN11	Papel e Celulose	KLBN11	Papel e Celulose	GOAU4	Mineração e Siderurgia
SUZB3	Papel e Celulose	SUZB3	Papel e Celulose	USIM5	Mineração e Siderurgia
CSAN3	Petroleo, gás e biocombustíveis	CSAN3	Petroleo, gás e biocombustíveis	VALE3	Mineração e Siderurgia
PETR3	Petroleo, gás e biocombustíveis	PETR3	Petroleo, gás e biocombustíveis	KLBN11	Papel e Celulose
PETR4	Petroleo, gás e biocombustíveis	PETR4	Petroleo, gás e biocombustíveis	SUZB3	Papel e Celulose
PRIO3	Petroleo, gás e biocombustíveis	PRIO3	Petroleo, gás e biocombustíveis	CSAN3	Petroleo, gás e biocombustíveis
UGPA3	Petroleo, gás e biocombustíveis	UGPA3	Petroleo, gás e biocombustíveis	PETR3	Petroleo, gás e biocombustíveis
VBBR3	Petroleo, gás e biocombustíveis	VBBR3	Petroleo, gás e biocombustíveis	PETR4	Petroleo, gás e biocombustíveis
FLRY3	Saúde	FLRY3	Saúde	PRIO3	Petroleo, gás e biocombustíveis
HAPV3	Saúde	HAPV3	Saúde	RRRP3	Petroleo, gás e biocombustíveis
HYPE3	Saúde	HYPE3	Saúde	UGPA3	Petroleo, gás e biocombustíveis
QUAL3	Saúde	QUAL3	Saúde	VBBR3	Petroleo, gás e biocombustíveis
RADL3	Saúde	RADL3	Saúde	FLRY3	Saúde
TIMS3	Telecomunicações	LWSA3	Tecnologia da Informação	HAPV3	Saúde
VIVT3	Telecomunicações	TOTS3	Tecnologia da Informação	HYPE3	Saúde
CMIG4	Utilidade Pública	TIMS3	Telecomunicações	QUAL3	Saúde
CPFE3	Utilidade Pública	VIVT3	Telecomunicações	RADL3	Saúde
CPLE6	Utilidade Pública	CMIG4	Utilidade Pública	RDOR3	Saúde
EGIE3	Utilidade Pública	CPFE3	Utilidade Pública	LWSA3	Tecnologia da Informação
ELET3	Utilidade Pública	CPLE6	Utilidade Pública	POSI3	Tecnologia da Informação
ELET6	Utilidade Pública	EGIE3	Utilidade Pública	TOTS3	Tecnologia da Informação
ENBR3	Utilidade Pública	ELET3	Utilidade Pública	TIMS3	Telecomunicações
ENEV3	Utilidade Pública	ELET6	Utilidade Pública	VIVT3	Telecomunicações
ENGI11	Utilidade Pública	ENBR3	Utilidade Pública	CMIG4	Utilidade Pública
EQTL3	Utilidade Pública	ENEV3	Utilidade Pública	CPFE3	Utilidade Pública
SBSP3	Utilidade Pública	ENGI11	Utilidade Pública	CPLE6	Utilidade Pública
TAE11	Utilidade Pública	EQTL3	Utilidade Pública	EGIE3	Utilidade Pública
AMER3	Varejo	SBSP3	Utilidade Pública	ELET3	Utilidade Pública
CRFB3	Varejo	TAE11	Utilidade Pública	ELET6	Utilidade Pública
CVCB3	Varejo	AMER3	Varejo	ENBR3	Utilidade Pública
LREN3	Varejo	ASAI3	Varejo	ENEV3	Utilidade Pública
MGLU3	Varejo	CRFB3	Varejo	ENGI11	Utilidade Pública
NTCO3	Varejo	CVCB3	Varejo	EQTL3	Utilidade Pública
PCAR3	Varejo	LREN3	Varejo	SBSP3	Utilidade Pública
VIIA3	Varejo	MGLU3	Varejo	TAE11	Utilidade Pública
		NTCO3	Varejo	ALPA4	Varejo
		PCAR3	Varejo	AMER3	Varejo
		VIIA3	Varejo	ASAI3	Varejo
				CRFB3	Varejo
				CVCB3	Varejo
				DXCO3	Varejo
				LREN3	Varejo
				MGLU3	Varejo
				NTCO3	Varejo
				PCAR3	Varejo
				PETZ3	Varejo
				SOMA3	Varejo
				VIIA3	Varejo

APÊNDICE B – CÓDIGO USADO PARA *BACKTEST* (PYTHON)

Disponível

em

<https://drive.google.com/drive/folders/1PgrHk524sNKv7f85dx0SozIL3e7hQ7D8?usp=sharing>