



UNIVERSIDADE FEDERAL FLUMINENSE
FACULDADE DE ADMINISTRAÇÃO E CIÊNCIAS CONTÁBEIS
Programa de Pós-Graduação em Administração
Mestrado Acadêmico em Administração

**IMPACTO DO BID-ASK SPREAD NA DETERMINAÇÃO DO PREÇO NO MERCADO
DE AÇÕES**

LEONARDO MELLO CAMILO DA SILVA

Niterói

2024

LEONARDO MELLO CAMILO DA SILVA

**IMPACTO DO BID-ASK SPREAD NA DETERMINAÇÃO DO PREÇO NO MERCADO
DE AÇÕES**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Administração da Faculdade de Administração e Ciências Contábeis da Universidade Federal Fluminense, como requisito para obtenção do título de Mestre em Administração.

Orientador: Prof. Dr. Claudio Henrique da
Silveira Barbedo

Niterói/RJ

2024

Resumo

Este estudo investiga o impacto do spread bid-ask na formação de preços de ativos financeiros, com foco no mercado de ações brasileiro. A pesquisa é dividida em duas partes: uma Revisão Narrativa de Literatura (RNL) e um estudo empírico. A RNL explora teorias e abordagens relacionadas ao spread bid-ask, revisando modelos clássicos e contemporâneos.

O estudo empírico aplica métodos reconhecidas para estimar o spread bid-ask no mercado brasileiro, analisando dados de 2017 a 2023. Os resultados indicam que o spread bid-ask é crucial na formação de preços, afetando liquidez e custos de transação. A comparação entre estimadores tem como proposta a identificação da metodologia mais adequada ao mercado brasileiro, oferecendo insights para investidores e reguladores. Este trabalho contribui para a literatura ao preencher lacunas na compreensão do spread bid-ask em mercados emergentes.

Palavras-chave: *Bid-ask spread; Microestrutura de mercado; Liquidez de mercado; Custos de transação.*

SUMÁRIO

1	Introdução	8
1.1	Contexto	8
1.1.1	Relevância e justificativa	8
2	Objetivos	10
3	Referencial teórico	11
3.1	Hipótese de eficiência de mercado	11
3.2	Microestrutura de mercado	12
3.3	Spread de oferta e compra (spread bid-ask)	13
3.4	Modelos de estimação de Spread bid-ask	15
4	Metodologia	17
4.1	Amostra	17
4.2	Modelos econométricos	18
4.3	Cálculo do spread bid-ask	23
4.4	Métricas de Avaliação	24
4.4.1	Métricas de Erro	24
4.4.2	Teste de Similaridade Estatística - Kolmogorov-Smirnov	25
5	Resultados.	26
5.1	Análise dos estimadores	26
5.2	Análise de período de volatilidade	29
5.3	Análise Comparativa com Dados de Referência da Bloomberg	31
5.3.1	Teste de Similaridade Estatística	31
5.3.2	Métricas de Erro	32
5.3.3	Interpretação Integrada	33
5.3.4	Implicações para o Mercado Brasileiro	33
6	Conclusão	35
7	Referências	37

LISTA DE FIGURAS

LISTA DE TABELAS

1	Tabela indicando as principais estatísticas do método EDGE	26
2	Tabela indicando as principais estatísticas do método CS	27
3	Tabela indicando as principais estatísticas do método AR	27
4	Tabela indicando as principais estatísticas do método CS para o ano de 2020	29
5	Tabela indicando as principais estatísticas do método EDGE para o ano de 2020 . . .	30
6	Tabela indicando as principais estatísticas do método AR para o ano de 2020	30
7	Tabela indicando o resultado do teste de Kolmogorov-Smirnov para cada metodologia de estimativa de spread bid-ask	32

LISTA DE EQUAÇÕES

1 Introdução

1.1 Contexto

O spread bid-ask é uma medida fundamental na microestrutura de mercado, representando a diferença entre os preços de compra (bid) e venda (ask) de um ativo. Este spread reflete os custos de transação e a liquidez do mercado, sendo influenciado por fatores como a volatilidade do ativo, a frequência de negociação e a assimetria de informações entre os participantes do mercado. A análise do spread bid-ask é crucial para entender a eficiência do mercado e os custos enfrentados pelos investidores.

No contexto dos mercados emergentes, a B3 (Brasil, Bolsa, Balcão) desempenha um papel central. A B3 é um importante centro de negociação de ativos financeiros na América Latina, ela não só facilita a negociação de ações, mas também de derivativos, títulos de renda fixa e outros instrumentos financeiros. A bolsa tem se destacado por sua infraestrutura tecnológica avançada e por iniciativas voltadas à inclusão financeira e ao aumento da liquidez do mercado. No entanto, os mercados emergentes, como o brasileiro, ainda enfrentam desafios significativos, incluindo menor liquidez e maior volatilidade em comparação com mercados desenvolvidos.

A literatura sobre spread bid-ask tem evoluído significativamente ao longo das últimas décadas, com diversos métodos sendo desenvolvidos para sua estimativa. Métodos tradicionais, como o estimador de Roll (1984), foram complementados por abordagens mais recentes, como os estimadores de Corwin e Schultz (2012) e de Abdi e Rinaldo (2017). Apesar desses avanços, há uma lacuna notável na aplicação dessas metodologias a mercados emergentes. A maioria dos estudos concentra-se em mercados desenvolvidos, deixando uma lacuna na compreensão de como essas metodologias se comportam em contextos de menor liquidez e maior volatilidade, como o mercado brasileiro.

1.1.1 Relevância e justificativa

Medir o spread bid-ask é de extrema relevância para diversos stakeholders no mercado financeiro. Corwin (2012) destaca que, para os investidores, o spread bid-ask é um indicador direto dos custos de transação e da liquidez do mercado, influenciando decisões de compra e venda de ativos.

Reguladores utilizam essa medida para monitorar a eficiência do mercado e identificar possíveis práticas de manipulação de preços. Acadêmicos, por sua vez, empregam o spread bid-ask como uma variável chave em estudos sobre microestrutura de mercado, eficiência de preços e assimetria informacional.

Este trabalho busca preencher lacunas importantes na literatura ao realizar uma análise comparativa de metodologias de estimativa do spread bid-ask aplicadas ao mercado brasileiro. Ripamonti (2016) observa que a ausência de estudos que comparam diferentes metodologias nesse contexto específico representa uma oportunidade para contribuir com novos insights sobre a eficiência e a liquidez do mercado brasileiro. Os resultados deste estudo podem auxiliar investidores na seleção de ações e na avaliação de riscos de liquidez, fornecendo uma base mais sólida para suas decisões de investimento.

A escolha do mercado brasileiro se justifica por suas características únicas como mercado emergente. A B3, com sua diversidade de ativos e participantes, oferece um ambiente ideal para testar a robustez das metodologias de estimativa do spread bid-ask. Além disso, Ripamonti (2016) sugere que os resultados obtidos podem ser generalizados para outros mercados emergentes com características semelhantes, ampliando o impacto e a relevância do estudo. Dessa forma, este trabalho não só contribui para a literatura acadêmica, mas também oferece implicações práticas para investidores e reguladores no contexto de mercados emergentes

2 Objetivos

Considerando a relevância do spread bid-ask na formação de preços de ativos financeiros e seu impacto na liquidez e nos custos de transação, este estudo tem como objetivo geral comparar, no mercado acionário brasileiro, estimativas de spread bid-ask obtidas pelas abordagens de Corwin & Schultz (2012), Abdi & Ranaldo (2017) e EDGE (Ardia, Guidotti & Kroencke, 2024), tendo como referência a cotação de spread divulgada pela Bloomberg.

A análise comparativa das metodologias selecionadas abrangerá tanto períodos de relativa estabilidade quanto momentos de volatilidade acentuada, com especial atenção ao período da pandemia de COVID-19 em 2020. Esta abordagem permitirá avaliar não apenas a precisão geral dos estimadores, mas também sua robustez e capacidade de capturar as variações nos custos de transação em condições extremas de mercado, fornecendo insights valiosos sobre a microestrutura do mercado brasileiro em diferentes contextos econômicos.

3 Referencial teórico

3.1 Hipótese de eficiência de mercado

De acordo com Fama (1970) a Hipótese de eficiência de mercado propõe que os investidores do mercado irão fazer a melhor alocação de ativos possível com todas as informações relevantes que terão acesso, e os preços dos ativos irá refletir todas as informações disponíveis para os investidores, que se basearão na teoria econômica para realizar tais cálculos. Em Finanças, a HEM (Hipótese da Eficiência de Mercado) afirma que os mercados financeiros são “eficientes em relação à informação”. Ou seja, um agente não consegue alcançar consistentemente retornos superiores à média do mercado, considerando as informações publicamente disponíveis no momento em que o investimento é feito (Fama, 1965; Malkiel, 2007; Jensen, 1978).

Existem três versões principais da hipótese: a fraca, a semi-forte e a forte. A hipótese fraca considera que os preços negociados para os bens (por exemplo, ações, obrigações ou propriedade) refletem toda a informação histórica disponível publicamente. Em seu trabalho seminal, Fama (1965) demonstra que os preços em mercados eficientes seguem um passeio aleatório, refletindo todas as informações históricas disponíveis, o que implica que os preços futuros não podem ser previstos com base em padrões passados.

A hipótese semi-forte afirma que os preços refletem todas as informações publicamente disponíveis e que os preços mudam instantaneamente para refletir as novas informações públicas. Esse conceito é extensivamente discutido no trabalho de Fama (1970), onde ele categoriza a eficiência de mercado em formas fraca, semi-forte e forte. A forma semi-forte da HEM sugere que nem mesmo a análise fundamentalista pode proporcionar vantagens, pois todas as informações disponíveis publicamente já estão incorporadas nos preços dos ativos. Estudos empíricos, como os de Malkiel (2007), corroboram essa visão ao demonstrar que as tentativas de prever movimentos de preços com base em análises fundamentais geralmente não resultam em retornos acima da média.

Embora a hipótese de eficiência de mercado proponha que os preços dos ativos refletem todas as informações disponíveis, o spread bid-ask é um fator crucial que pode introduzir ineficiências no mercado. Custos de transação elevados, informação assimétrica, baixa liquidez e problemas de microestrutura do mercado são elementos que podem ampliar o spread e impedir que os preços se

ajustem rapidamente às novas informações. Isso sugere que, na prática, os mercados podem não ser tão eficientes quanto a teoria propõe, e os investidores precisam considerar esses fatores ao tomar decisões de investimento.

3.2 Microestrutura de mercado

A microestrutura de mercado é um campo de estudo que analisa os mecanismos e processos pelos quais os preços dos ativos são formados nos mercados financeiros. Este campo abrange uma variedade de tópicos, incluindo a organização e funcionamento dos mercados, o comportamento dos participantes do mercado, e os custos de transação. Segundo Harris (2003), a microestrutura de mercado examina como a estrutura do mercado e os mecanismos de negociação afetam a formação de preços, a liquidez e a eficiência do mercado.

Compreender a microestrutura é fundamental para analisar como as informações são incorporadas nos preços e como os custos de transação influenciam as decisões de investimento. Os mercados financeiros podem ser organizados de várias maneiras, incluindo mercados de leilão e mercados de dealer. Em mercados de leilão, como as bolsas de valores, os preços são estabelecidos através de leilões competitivos, onde compradores e vendedores submetem ordens que são combinadas para formar o preço de mercado. Nos mercados de dealer, como o mercado de câmbio, os dealers fornecem liquidez ao comprar e vender ativos de seu próprio inventário.

Harris (2003) destaca a importância do livro de ordens, que registra todas as ordens de compra e venda pendentes e desempenha um papel crucial na determinação dos preços de mercado. A liquidez do mercado, definida por como a facilidade com que um ativo pode ser comprado ou vendido sem causar uma mudança significativa em seu preço, é fortemente influenciada pela microestrutura do mercado (Harris, 2003). Chordia, et. al (2000) mostram que a estrutura do mercado e a presença de intermediários financeiros desempenham um papel vital na determinação da liquidez. A presença de market makers, por exemplo, pode aumentar a liquidez ao fornecer cotações firmes de compra e venda e ao absorver temporariamente desequilíbrios entre oferta e demanda. Além disso, a profundidade do livro de ordens e a quantidade de informações disponíveis aos participantes do mercado também afetam a liquidez.

Os custos de transação são um componente crucial da microestrutura de mercado e têm um

impacto significativo na eficiência do mercado. Stoll (1989) e Glosten e Milgrom (1985) identificam três principais componentes dos custos de transação: custos de processamento das ordens, custos de assimetria de informação e custos de inventário. Os custos de processamento referem-se aos custos operacionais incorridos pelos intermediários ao executar ordens de compra e venda. Os custos de assimetria de informação surgem da incerteza sobre a informação que outros participantes do mercado possuem. Já os custos de inventário são os custos associados à manutenção de um estoque de ativos para facilitar a negociação. A eficiência do mercado é afetada por esses custos, pois eles influenciam a capacidade dos preços refletirem todas as informações disponíveis.

As inovações tecnológicas, particularmente o trading eletrônico e os algoritmos de negociação, têm transformado a microestrutura dos mercados financeiros. Hendershott, Jones e Menkveld (2011) investigam o impacto do trading algorítmico na liquidez do mercado e concluem que ele melhora a liquidez ao aumentar a frequência e a velocidade das transações. A introdução de plataformas de negociação eletrônica reduziu os custos de transação e aumentou a transparência do mercado, permitindo uma melhor incorporação de informações nos preços. No entanto, essas inovações também trazem desafios, como a necessidade de regulamentos adequados para mitigar os riscos associados ao trading de alta frequência e para garantir a estabilidade do mercado.

3.3 Spread de oferta e compra (spread bid-ask)

Custos de inventário ou de estoque, segundo De Jong e Rindi (2009), existem apenas em mercados quote-driven, pois nesses mercados os formadores têm a obrigação de fornecer preços de compra e preços de venda ininterruptamente ao mercado. Os custos relacionados à assimetria de informação e de transação ou de processamento de ordens existem em qualquer mercado. O mercado de ações da B3 é order-driven e nesse caso não se identifica o custo de estoque.

Nos mercados quote-driven, também conhecidos como mercados de dealer, um intermediário (dealer) participa de todas as transações. Nesse tipo de mercado, os traders negociam diretamente com o dealer, que fornece cotações dos preços de compra e venda, assumindo o lado oposto de cada transação e, assim, proporcionando liquidez ao mercado. Esses mercados são comuns no comércio de moedas estrangeiras e títulos no mercado de balcão (OTC).

Por outro lado, nos mercados order-driven, os compradores e vendedores negociam direta-

mente entre si, sem a intervenção de um dealer, utilizando um conjunto de regras de negociação estabelecidas. Essas regras incluem precedência de ordens, que determina como as ordens de compra e venda são casadas, e regras de precificação de transações. Os mercados order-driven são prevalentes em bolsas de futuros, ações e opções, bem como em redes eletrônicas de comunicação para negociação de títulos e moedas. Em suma, a principal diferença reside no papel do intermediário: enquanto nos mercados quote-driven o dealer facilita todas as transações fornecendo cotações, nos mercados order-driven, as transações são realizadas diretamente entre os traders com base em regras de mercado (Peat, 2009).

No contexto do mercado de ações no Brasil, a B3 adota um modelo de mercado order-driven, mas algumas ações contam com a presença de formadores de mercado. Os formadores de mercado, ou market makers, são contratados pelas empresas para aumentar a liquidez de suas ações. A liquidez refere-se à facilidade com que um ativo pode ser comprado ou vendido sem causar uma mudança significativa em seu preço (Harris, 2003). A presença de um formador de mercado pode reduzir o spread de compra e venda, o que diminui os custos de transação para os investidores e aumenta a atratividade da ação. Empresas contratam formadores de mercado para aumentar a liquidez de suas ações, melhorar a precificação e atrair mais investidores, especialmente em períodos de baixa negociação ou para novos papéis que estão sendo introduzidos no mercado.

A assimetria de informação, por sua vez, ocorre quando diferentes participantes do mercado têm diferentes níveis de acesso à informação relevante para a precificação dos ativos. Nos mercados order-driven, a assimetria de informação pode ser mitigada pela transparência do livro de ordens e pelas regras claras de negociação. No entanto, a assimetria ainda pode existir, especialmente em mercados menos líquidos ou em ações com menor cobertura de analistas. Em mercados quote-driven, a assimetria de informação pode ser mais pronunciada, pois os dealers podem ter acesso a informações privilegiadas que não estão disponíveis para todos os participantes do mercado.

Além disso, os custos de transação referem-se aos custos associados à realização de uma transação econômica, que incluem os custos de negociação e de contratação (Demsetz, 1968). Esses custos incluem taxas pagas à bolsa, comissões de corretagem e custos relacionados ao tempo e esforço necessários para processar as ordens. Em mercados order-driven como a B3, esses custos podem ser reduzidos pela automação e eficiência das plataformas de negociação eletrônica, que

facilitam a correspondência rápida e precisa das ordens de compra e venda.

Em suma, a presença de formadores de mercado em ações na B3 visa principalmente aumentar a liquidez, reduzir o spread de compra e venda e melhorar a eficiência do mercado. Ao contratar formadores de mercado, as empresas buscam criar um ambiente de negociação mais ativo e atraente para os investidores, promovendo assim uma melhor precificação e maior volume de negociações. A compreensão dos diferentes custos associados ao spread de compra e venda, bem como das diferenças entre mercados quote-driven e order-driven, é essencial para avaliar o impacto da microestrutura de mercado na formação de preços dos ativos financeiros.

3.4 Modelos de estimação de Spread bid-ask

A mensuração do spread bid-ask pode ser realizada por meio de dados de alta frequência, que registram de forma contínua as cotações e negociações ao longo do dia, ou através de dados de baixa frequência, os quais utilizam informações agregadas, como os preços de fechamento, máxima e mínima diários. Em contextos onde os dados de alta frequência são escassos ou de difícil acesso, os métodos de estimação baseados em dados agregados tornam-se fundamentais para a avaliação dos custos de transação e da liquidez dos ativos.

Neste trabalho, serão utilizadas três metodologias de estimação de baixa frequência que vêm contribuindo de maneira inovadora para aprimorar a medição do spread bid-ask. A abordagem de Corwin e Schultz (2012) introduz uma forma simplificada de inferir o spread a partir das oscilações diárias, enquanto a proposta de Abdi e Rinaldo (2017) aprimora esse método ao reduzir incertezas inerentes às estimativas com dados diários. Por fim, a metodologia de Ardia, Guidotti e Kroencke (2024) integra informações adicionais, como o preço de abertura, proporcionando uma estimativa mais robusta e precisa.

O modelo de Corwin e Schultz (2012) foi um dos primeiros a propor uma metodologia baseada apenas em preços diários de máxima e mínima para estimar o spread bid-ask. A ideia central do modelo é que o intervalo entre os preços mais altos e mais baixos de um dia reflete tanto a volatilidade do ativo quanto a diferença entre os preços de compra e venda. Assim, os autores derivam um estimador baseado na comparação entre a soma dos intervalos de dois dias consecutivos e o intervalo total de dois dias. Esse modelo trouxe uma abordagem simples e aplicável a mercados

onde não há disponibilidade de dados de cotações intradiárias

Abdi e Ranaldo (2017) expandiram essa metodologia ao incluir informações do preço de fechamento, além dos preços de máxima e mínima diários. O modelo desenvolvido pelos autores tem como premissa a ideia de que o preço de fechamento contém informações adicionais sobre o comportamento do spread bid-ask. Utilizando uma formulação matemática que combina essas três variáveis, o estimador proposto por Abdi e Ranaldo mostrou-se mais preciso para ativos com menor liquidez, superando modelos anteriores em termos de correlação com medidas de spread obtidas a partir de dados de alta frequência

Mais recentemente, Ardia, Guidotti e Kroencke (2024) propuseram um modelo ainda mais abrangente, que incorpora informações dos preços de abertura, máxima, mínima e fechamento. Diferentemente dos modelos anteriores, que ignoram o preço de abertura, essa nova abordagem busca reduzir vieses na estimação ao aproveitar um conjunto mais completo de informações de preços. Além disso, os autores desenvolveram um estimador eficiente que minimiza a variância da estimativa, tornando-se uma alternativa superior para mensurar spreads em mercados com diferentes níveis de liquidez. O modelo de Ardia et al. (2024) representa um avanço na literatura ao abordar problemas de viés e eficiência estatística na estimação do spread bid-ask

Dessa forma, a evolução das metodologias de estimação do spread bid-ask reflete a necessidade de desenvolver modelos que equilibrem simplicidade e precisão, especialmente em mercados onde dados de alta frequência não estão disponíveis. A aplicação desses modelos ao mercado brasileiro permite uma avaliação mais detalhada da liquidez dos ativos e dos custos de transação, contribuindo para uma melhor compreensão das dinâmicas de precificação no mercado acionário nacional. Além disso, tais estimadores são fundamentais para o desenvolvimento de programas computacionais que viabilizem a estimação contínua do spread bid-ask, permitindo uma análise mais eficiente do comportamento do mercado.

4 Metodologia

A metodologia deste estudo foi desenvolvida para permitir uma análise rigorosa e comparativa das diferentes metodologias de estimativa do spread bid-ask no mercado de ações brasileiro. A seguir, são detalhadas as etapas e procedimentos adotados para a coleta e análise dos dados, bem como a aplicação das metodologias selecionadas.

4.1 Amostra

A base de dados deste estudo foi composta a partir das informações de preços diários das ações listadas na B3, aplicando critérios rigorosos de seleção para garantir a representatividade e a qualidade dos dados utilizados. O critério principal para inclusão das ações na amostra foi o índice de liquidez, calculado conforme a metodologia descrita por Andrade et al. (2009). Foram consideradas apenas as ações com índice de liquidez superior a 0,1, calculado com base em uma janela de seis meses de dados, assegurando a representatividade de ativos com maior frequência de negociação.

A amostra final é composta por 138 ações, distribuídas entre os diferentes segmentos de governança corporativa da B3. Entre elas, 96 pertencem ao Novo Mercado, 16 ao Nível 2, 18 ao Nível 1 e 4 ao segmento Tradicional. Essa segmentação reflete distintos níveis de práticas de governança corporativa, permitindo uma análise abrangente sobre como essas características impactam os spreads bid-ask e a liquidez dos ativos.

A definição da amostra seguiu abordagens amplamente utilizadas na literatura, como a adotada por Andrade et al. (2009), que empregaram critérios de liquidez para selecionar empresas listadas. Empresas do setor financeiro foram excluídas devido às particularidades de suas demonstrações contábeis, que poderiam introduzir vieses na análise. Além disso, foram descartadas companhias com dados incompletos ou inconsistentes, assegurando que a amostra final apresentasse informações confiáveis para a investigação empírica.

O processo de filtragem das ações visou garantir a robustez dos dados utilizados, selecionando ativos que refletissem, de forma mais precisa, o valor de mercado real. A inclusão de ações com maior liquidez não apenas melhora a representatividade da amostra, mas também favorece a

precisão das estimativas econométricas realizadas no estudo, reduzindo potenciais vieses causados por falta de negociação.

4.2 Modelos econométricos

Este estudo aplicou três metodologias amplamente reconhecidas para a estimativa do spread bid-ask: CS, AR e EDGE utilizando dados obtidos diretamente da B3. Essas abordagens foram selecionadas devido à sua robustez teórica, ampla aceitação na literatura e capacidade de oferecer insights em diferentes contextos de mercado.

O estimador de spread bid-ask desenvolvido por Corwin e Schultz (2012) utiliza os preços diários de máxima e mínima das ações para calcular o spread de forma simples e eficiente. A metodologia parte de duas premissas principais. Primeiramente, assume-se que os preços máximos tendem a ser resultados de ordens de compra, enquanto os mínimos são influenciados por ordens de venda. Dessa forma, a relação entre os preços de máxima e mínima reflete tanto a volatilidade fundamental do ativo quanto os custos associados ao spread bid-ask. Em segundo lugar, considera-se que a componente de volatilidade aumenta proporcionalmente com o intervalo de negociação, enquanto a parte do spread permanece constante.

Matematicamente, o estimador é definido pela seguinte expressão:

$$S = \frac{2(e^\alpha - 1)}{1 + e^\alpha},$$

onde S representa o spread bid-ask estimado e e é a base do logaritmo natural. O parâmetro α é calculado por:

$$\alpha = \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}},$$

sendo β e γ definidos a partir dos logaritmos naturais das razões entre os preços máximos, H , e mínimos, L , observados:

$$\beta = E \left[\sum_{j=0}^1 \left(\ln \frac{H_{t+j}}{L_{t+j}} \right)^2 \right],$$

$$\gamma = E \left[\left(\ln \frac{H_{t,t+1}}{L_{t,t+1}} \right)^2 \right],$$

onde $H_{t,t+1} = \max(H_t, H_{t+1})$ e $L_{t,t+1} = \min(L_t, L_{t+1})$. Nesse contexto, β captura as

variações diárias entre os preços máximos e mínimos, enquanto γ reflete a dinâmica entre dois dias consecutivos. Essas variáveis permitem separar a volatilidade fundamental do ativo da componente relacionada ao spread bid-ask.

A partir dessas definições, torna-se evidente que o método permite isolar o impacto do spread bid-ask dos movimentos puramente voláteis do preço. A implementação do estimador é direta, dispensando iterações complexas, o que facilita sua aplicação prática em grandes conjuntos de dados. Corwin e Schultz (2012) validaram o estimador por meio de simulações que replicaram condições realistas de mercado, demonstrando que a correlação entre os valores estimados pelo método e os spreads reais é elevada, em torno de 0,9. Em comparação com o estimador de covariância de Roll (1984), o desvio padrão das estimativas de Corwin e Schultz é cerca de metade, indicando maior estabilidade nos resultados.

O método foi aplicado a ações de 11 países, incluindo o Brasil, evidenciando sua eficácia em medir os custos de transação em diferentes contextos de liquidez e volatilidade. Essa metodologia tem sido amplamente utilizada em estudos sobre microestrutura de mercado. Por exemplo, foi empregada para analisar a relação entre a qualidade das informações financeiras e a assimetria informacional, com resultados que destacaram sua eficiência em comparação a outros indicadores. Embora algumas pesquisas apontem limitações no uso do método como medida de liquidez para prever retornos, ele continua sendo uma ferramenta confiável para estimar custos de transação e compreender a dinâmica dos mercados financeiros.

Portanto, o estimador de Corwin e Schultz (2012) oferece uma solução prática e robusta para mensurar o spread bid-ask com base em dados diários, sendo amplamente adotado em estudos de microestrutura, especialmente em mercados onde dados intradiários são limitados. Sua simplicidade e eficácia tornam-no uma ferramenta valiosa para analisar a liquidez e a eficiência do mercado financeiro brasileiro.

Já o estimador de Abdi e Rinaldo (2017) trouxe uma metodologia nova para o cálculo do spread, sendo desenvolvida para estimar o spread bid-ask utilizando preços diários de fechamento, máxima e mínima das ações. A abordagem se baseia na decomposição do spread em componentes de volatilidade e custos de transação, permitindo uma análise detalhada dos fatores que influenciam a formação de preços nos mercados financeiros. Esse estimador é fundamentado em três etapas

principais. Primeiro, utiliza-se o preço médio entre a máxima e a mínima (mid-range) como um proxy para o preço eficiente, que é definido como $MR_t = \frac{H_t + L_t}{2}$, supondo que o spread bid-ask se cancela nesse cálculo. Em seguida, calcula-se a distância quadrática entre o preço de fechamento e o proxy de preço eficiente derivado da média das médias de dois dias consecutivos, com essa distância sendo representada como: $D_t = (C_t - MR_t)^2$. Por fim, essa distância é ajustada para remover a variância do preço eficiente, σ_D^2 , isolando assim o componente relacionado ao spread bid-ask. Essa abordagem, além de simplificar cálculos, não requer ajustes *ad hoc* para períodos sem negociação, como finais de semana e feriados, o que a torna menos suscetível a vieses comuns em outros modelos.

No final, a fórmula do estimador é dada por:

$$S = 2\sqrt{\sigma_D^2},$$

onde S é o spread bid-ask estimado, e σ_D^2 é a variância do termo D_t , ajustada com base nos preços de dois dias consecutivos. Essa formulação permite capturar de maneira mais precisa os custos de transação associados às negociações

Adicionalmente, o modelo de Abdi e Rinaldo (2017) se destaca em cenários de baixa liquidez. Ao contrário do estimador de Corwin-Schultz (2012), que apresenta viés descendente em ativos menos líquidos, esse método AR foi dito como mais estável e robusto em condições de mercado desafiadoras, o tornando uma ferramenta valiosa para analisar a eficiência e a liquidez em mercados emergentes e em desenvolvimento, onde a disponibilidade de dados intradiários é limitada. A relevância desse estimador é evidente ao compará-lo com outros modelos, como o de Corwin-Schultz. Enquanto o estimador de Corwin-Schultz depende de uma estrutura não-linear para decompor a volatilidade do preço e o spread bid-ask, o modelo de Abdi e Rinaldo (2017) utiliza uma solução em forma fechada, simplificando o processo de estimação e reduzindo a necessidade de premissas adicionais, como a igualdade de chances de preços de alta e baixa serem iniciados por compradores ou vendedores.

Além disso, o método é amplamente aplicável, tendo sido validado em Abdi e Rinaldo (2017) em uma extensa amostra de ações nos EUA desde 1926, demonstrando alto grau de correlação com spreads efetivos medidos por dados intradiários (TAQ). Isso confirma sua eficácia como

uma ferramenta para mensurar custos de transação e liquidez em longos períodos históricos, contribuindo para a análise de risco de liquidez sistemática e de commonality em liquidez.

O método EDGE, desenvolvido por Ardia, Guidotti e Kroencke (2024), surge como uma proposta robusta para estimar o spread bid-ask de maneira eficiente, utilizando preços diários de abertura, máxima, mínima e fechamento (OHLC). Este método aborda uma limitação crítica de estimadores anteriores, como os de Corwin e Schultz (2012) e Abdi e Ranaldo (2017): o viés introduzido pela suposição de preços observados continuamente, que não reflete a realidade de mercados com baixa frequência de negociações.

O EDGE parte de três premissas centrais, compartilhadas com outros métodos: (1) os retornos fundamentais não são autocorrelacionados; (2) os retornos fundamentais são independentes das flutuações causadas pelo spread bid-ask; e (3) as flutuações do spread não possuem média e são não correlacionadas. No entanto, ele se diferencia ao relaxar a suposição de que preços são continuamente observados. Em mercados reais, a quantidade de negociações por intervalo de tempo é finita, o que resulta em um viés de subestimação dos spreads, especialmente em ativos com baixa liquidez. O EDGE corrige essa limitação introduzindo termos analíticos que ajustam o viés decorrente da observação discreta de preços.

Matematicamente, o spread bid-ask estimado pelo EDGE é definido como:

$$S = f(O_t, H_t, L_t, C_t),$$

onde S é o spread estimado, e a função f combina as estatísticas derivadas dos preços OHLC. Em essência, o método utiliza a amplitude intradiária normalizada para capturar a dinâmica entre o preço eficiente e o spread bid-ask. A amplitude é calculada como:

$$A_t = \ln \frac{H_t}{L_t},$$

e o ajuste para a observação discreta de preços é dado por:

$$D_t = \ln \frac{C_t}{O_t},$$

onde A_t reflete a volatilidade e D_t corrige o viés causado pela não observação contínua dos preços. O spread estimado é então ajustado por um termo de correção baseado na variação observada em múltiplos dias consecutivos, representado como:

$$S = \alpha \cdot A_t + \beta \cdot D_t,$$

em que os coeficientes α e β são calibrados para garantir que o estimador seja imparcial e apresente a menor variância possível. Esses coeficientes são derivados empiricamente, considerando simulações e validações em diferentes mercados.

De acordo com Ardia, Guidotti e Kroencke (2024), uma das características mais inovadoras do EDGE é a combinação otimizada de estimadores derivados de diferentes combinações de preços (como abertura e fechamento, máxima e mínima). Isso é feito para minimizar a variância de estimação, resultando em um estimador eficiente que se destaca tanto em cenários de spreads baixos quanto elevados. Segundo o artigo, enquanto estimadores como o de Corwin e Schultz (2012) são mais precisos para spreads pequenos, e o de Abdi e Ranaldo (2017) apresenta menor viés para spreads grandes, o EDGE combina o melhor desses cenários, garantindo estimativas com menor variância e viés uniforme.

Comparando com métodos anteriores, como o de Abdi e Ranaldo, o EDGE generaliza sua abordagem ao considerar informações adicionais de preços de abertura e fechamento. Enquanto o estimador de Abdi e Ranaldo utiliza médias de preços máximos, mínimos e de fechamento para isolar o componente de volatilidade do spread, o EDGE aproveita todas as informações de OHLC e introduz ajustes analíticos para corrigir observações discretas. Além disso, o EDGE não requer ajustes adicionais para períodos sem negociação, como finais de semana e feriados, assim como o estimador de Abdi e Ranaldo.

Sendo assim, o EDGE é uma contribuição relevante para a literatura de microestrutura de mercado, oferecendo uma ferramenta confiável e eficiente para a estimação de custos de transação. Ele não apenas melhora a precisão das estimativas em relação a métodos tradicionais, mas também amplia sua aplicabilidade a diversos contextos de mercado, como períodos históricos e mercados emergentes.

Os três estimadores discutidos – Corwin-Schultz (CS), Abdi-Ranaldo (AR) e EDGE – foram escolhidos para este estudo devido à sua relevância teórica, robustez metodológica e capacidade de oferecer insights em diferentes contextos de mercado. Além disso, suas características complementares permitem uma análise comparativa detalhada de suas performances no mercado brasileiro,

que é classificado como emergente.

A aplicação dos métodos será realizada em três períodos distintos: pré-pandemia, durante a pandemia e no pós-pandemia. O período pré-pandemia é caracterizado por uma maior estabilidade econômica e condições de mercado mais previsíveis, oferecendo uma base para avaliar o desempenho dos estimadores em um cenário mais controlado. Durante a pandemia, o mercado foi marcado por alta volatilidade e incertezas, o que representa um contexto de maior desafio para os estimadores, especialmente em ativos menos líquidos. Por fim, o período pós-pandemia reflete um mercado em processo de readaptação, com uma recuperação gradual da liquidez e uma nova estabilidade, permitindo verificar a consistência dos métodos em cenários intermediários.

Com isso, este estudo busca avaliar como cada método se adapta às características específicas do mercado brasileiro e dos diferentes contextos temporais, contribuindo para a literatura de microestrutura ao explorar os pontos fortes e limitações de cada estimador em mercados emergentes e desafiadores.

4.3 Cálculo do spread bid-ask

As três metodologias selecionadas foram: EDGE (Ardia, Guidotti e Kroencke, 2024), CS (Corwin e Schultz, 2012) e AR (Abdi e Ranaldo, 2017). O uso de uma janela móvel de 21 dias, representada pelo parâmetro $width = 21$, foi adotado para garantir consistência metodológica e facilitar a comparação com estudos anteriores, como os de Ardia, Guidotti e Kroencke (2024), que seguem o mesmo princípio. Essa configuração considera 21 dias como equivalente a aproximadamente um mês de negociações, com base na prática padrão em mercados financeiros, onde a média de dias úteis de negociação em um mês é de cerca de 21.

Além disso, no estudo de Ardia, Guidotti e Kroencke (2024), a aplicação do estimador é baseada em períodos de 21 dias úteis para representar um mês de negociações. Essa abordagem é consistente com os trabalhos de Corwin e Schultz (2012) e Abdi e Ranaldo (2017), que também utilizam essa mesma configuração temporal em suas análises e simulações. O objetivo é criar uma unidade temporal padronizada que permita avaliar variações no spread bid-ask de maneira comparável ao longo de diferentes períodos e contextos de negociação.

Essa escolha também facilita a análise do impacto de diferentes frequências de negociação,

uma vez que uma janela de 21 dias captura tanto oscilações diárias quanto tendências mais amplas em cenários de alta e baixa liquidez. Portanto, a adoção desse parâmetro não só mantém a consistência com a literatura, mas também reflete práticas amplamente aceitas no campo de microestrutura de mercado, garantindo a robustez e a validade dos resultados obtidos no contexto do mercado brasileiro.

O cálculo foi realizado sobre as séries de preços de abertura, máxima, mínima, fechamento (OHLC) e volumes diários obtidos diretamente da B3. As metodologias foram aplicadas sem modificações adicionais, seguindo rigorosamente os algoritmos descritos nos trabalhos originais, de forma a garantir a comparabilidade com os benchmarks teóricos.

4.4 Métricas de Avaliação

As estimativas obtidas pelas três metodologias foram comparadas com os valores de referência fornecidos pela Bloomberg, considerados o padrão-ouro para a estimativa do spread bid-ask. Para realizar essa análise, os dados da Bloomberg foram importados e integrados à base processada, permitindo a construção de comparativos entre as estimativas calculadas e os valores reportados. Essa integração foi necessária para garantir uma avaliação precisa da performance de cada metodologia em relação a uma referência amplamente reconhecida no mercado financeiro. Foram utilizadas métricas estatísticas para medir a proximidade das estimativas com os valores de referência, como o erro quadrático médio (MSE), que avalia a magnitude do erro entre as estimativas, e o coeficiente de correlação de Pearson, que mede a força e a direção da associação entre os valores estimados e os da Bloomberg. Esses indicadores forneceram uma base objetiva para identificar as condições em que os estimadores apresentaram melhor ou pior desempenho, destacando suas capacidades e limitações em diferentes contextos de negociação.

4.4.1 Métricas de Erro

Para quantificar a precisão das estimativas geradas por cada metodologia em relação aos valores oficiais da Bloomberg, foram utilizadas as seguintes métricas de erro:

4.4.1.1 Erro Quadrático Médio (RMSE) O RMSE é uma medida amplamente utilizada para avaliar a magnitude dos erros de estimação, calculada pela raiz quadrada da média dos quadrados

das diferenças entre valores estimados e valores reais:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

onde y_i representa o valor oficial do spread (Bloomberg) e \hat{y}_i representa o valor estimado pela metodologia em análise. O RMSE atribui maior peso a erros maiores devido à operação de elevação ao quadrado, tornando-o particularmente sensível a outliers. Esta métrica foi escolhida como principal indicador de proximidade entre as estimativas e os valores reais, sendo que valores menores de RMSE indicam maior precisão da metodologia.

4.4.1.2 Erro Absoluto Médio (MAE) Como métrica complementar, foi utilizado o MAE, calculado pela média dos valores absolutos das diferenças entre valores estimados e valores reais:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

O MAE atribui peso igual a todos os erros, independentemente de sua magnitude, oferecendo uma perspectiva alternativa sobre a precisão das estimativas.

4.4.2 Teste de Similaridade Estatística - Kolmogorov-Smirnov

Para avaliar a similaridade estatística entre as distribuições das estimativas e dos valores oficiais, foi aplicado o teste de Kolmogorov-Smirnov (KS). Este teste não paramétrico avalia a hipótese nula de que duas amostras provêm da mesma distribuição. A estatística D do teste KS é calculada como:

$$D = \sup_x |F_1(x) - F_2(x)|$$

onde $F_1(x)$ e $F_2(x)$ são as funções de distribuição acumulada empíricas para as duas amostras sendo comparadas, e \sup_x representa o supremo (máximo) das diferenças entre as duas distribuições. A interpretação dos resultados do teste KS é baseada no p-valor:

- p-valor < 0,05: Rejeita-se a hipótese nula, indicando diferença estatisticamente significativa entre as distribuições.
- p-valor ≥ 0,05: Não se rejeita a hipótese nula, sugerindo que não há evidência estatística suficiente para afirmar que as distribuições são diferentes.

5 Resultados

O estimador EDGE tem sua construção baseada na combinação otimizada de diferentes estimadores, minimizando a variância da estimativa final. Essa abordagem faz com que o método seja eficaz tanto em cenários de spreads estreitos quanto amplos, além de se adaptar automaticamente a variações na volatilidade do mercado. Dessa forma, a metodologia EDGE se mostra particularmente superior nos períodos de alta volatilidade, nos quais outras metodologias apresentam desempenho inconsistente.

A seguir, são apresentadas as análises comparativas entre os diferentes estimadores, com o objetivo de avaliar suas performances no contexto do mercado acionário brasileiro e identificar qual metodologia oferece maior precisão na mensuração do spread bid-ask.

5.1 Análise dos estimadores

Como visto na tabela 1, com as estatísticas do EDGE, na tabela 2, com as estatísticas do CS, e na tabela 3, com as estatísticas do AR, ao analisar os resultados obtidos, observa-se que o estimador EDGE apresentou valores mínimos de spread bid-ask de 2.2×10^{-5} para ativos mais líquidos e $1,6 \times 10^{-5}$ para ativos menos líquidos. Em contraste, o estimador CS retornou valores iguais a zero para ambos os grupos de ativos. Esses achados estão em consonância com o exposto por Ardia, Guidotti e Kroencke (2024), que investigaram o desempenho de diferentes estimadores de spread bid-ask em função da frequência de negociações.

Tabela 1: Tabela indicando as principais estatísticas do método EDGE

	EDGE				
	Média	DesvPad	Mediana	Máximo	Mínimo
Todos	0.008376	0.006727	0.006973	0.153692	0.000016

Fonte: Elaborado pelo autor

Tabela 2: Tabela indicando as principais estatísticas do método CS

	CS				
	Média	DesvPad	Mediana	Máximo	Mínimo
Todos	0.004229	0.003589	0.003466	0.066553	0.000000

Fonte: Elaborado pelo autor

Tabela 3: Tabela indicando as principais estatísticas do método AR

	AR				
	Média	DesvPad	Mediana	Máximo	Mínimo
Todos	0.011402	0.008226	0.009779	0.147883	0.000000

Fonte: Elaborado pelo autor

Segundo Ardia, Guidotti e Kroencke (2024), enquanto todos os estimadores avaliados são não viesados em cenários com alta frequência de negociações (por exemplo, 390 negociações por dia), suas performances divergem significativamente à medida que essa frequência diminui. Especificamente, o estimador CS apresenta um viés negativo acentuado em frequências menores, chegando a estimar spreads nulos em situações com menos de dez negociações diárias. Isso ocorre devido à forte dependência do CS na suposição de que os ativos são negociados continuamente, o que não é válido em mercados com baixa liquidez.

Por outro lado, o EDGE mantém estimativas não viesadas independentemente do número de negociações, graças a um termo de correção analítico que considera a negociação infrequente. Nossos resultados corroboram essa evidência, mostrando que o EDGE não retorna valores nulos mesmo em ativos com baixa frequência de negociação, ao contrário do CS. Dessa forma, o EDGE se apresenta como uma metodologia mais robusta e apropriada para estimar o spread bid-ask em

mercados com ativos menos líquidos.

Ao incluir o estimador de Abdi e Ranaldo (2017) (AR) na análise, conforme apresentado na tabela 3, observa-se que o AR também fornece resultados significativos na estimativa do spread bid-ask no mercado brasileiro. O estimador AR exibiu valores mínimos de spread bid-ask de 0,0 para ativos mais líquidos e $9,2 \times 10^{-5}$ para ativos menos líquidos. Embora apresente valores mínimos iguais a zero em alguns casos, o AR tende a estimar spreads maiores em comparação com o CS e o EDGE.

Os resultados mostram que o estimador AR possui médias de spread bid-ask de 0,011013 para ativos mais líquidos e 0,012382 para ativos menos líquidos, superiores às médias observadas para o EDGE (0,008308 e 0,008709, respectivamente) e significativamente maiores que as do CS (0,004102 e 0,004512). Esse comportamento sugere que o AR captura de maneira mais abrangente os custos de transação, especialmente em mercados com menor liquidez.

Abdi e Ranaldo (2017) destacam que seu estimador é particularmente robusto em condições de baixa liquidez, pois utiliza preços de fechamento, máxima e mínima para estimar o spread bid-ask, sem dependência excessiva da frequência de negociação. Essa característica é essencial no contexto do mercado brasileiro, onde muitos ativos apresentam negociações infrequentes.

Comparando os três estimadores, nota-se que o CS apresenta limitações significativas ao estimar spreads em ativos menos líquidos, frequentemente retornando valores nulos devido à suposição de negociação contínua. O EDGE, ao introduzir termos de correção para a observação discreta de preços, mantém estimativas não viesadas mesmo em ambientes de baixa liquidez. O AR, por sua vez, além de ser robusto em condições de baixa frequência de negociação, tende a estimar spreads mais elevados, possivelmente refletindo de forma mais acurada os custos enfrentados pelos investidores em mercados emergentes.

Esses achados enfatizam a importância de selecionar metodologias adequadas ao perfil de liquidez dos ativos analisados. Enquanto o CS pode ser apropriado para mercados altamente líquidos, o EDGE e o AR demonstram vantagens claras em contextos de menor liquidez, como seria o caso do mercado brasileiro. O fato de o AR estimar spreads maiores pode fornecer insights adicionais sobre a presença de custos de transação elevados e assimetrias informacionais que não são

capturadas pelos outros estimadores.

5.2 Análise de período de volatilidade

Ao comparar os resultados obtidos para todo o período de análise, de 2017 a 2023, com os resultados específicos do ano de 2020, marcado por alta volatilidade devido à pandemia de COVID-19, observamos certas variações nos spreads bid-ask estimados pelas três metodologias: CS, EDGE e AR.

Em 2020, todos os estimadores registraram um aumento significativo nas médias dos spreads bid-ask em comparação com o período total. Esse aumento é consistente com o contexto de alta volatilidade e incerteza do mercado durante a pandemia.

Os dados na tabela 4 mostram que o estimador CS apresentou valores médios de spread bid-ask menores em relação aos estimadores EDGE e AR, tanto no período total quanto em 2020. Contudo, durante o ano de 2020, o CS registrou um aumento na média do spread bid-ask de aproximadamente 31% para ativos mais líquidos, de 0,004102 para 0,005204 e 23% para ativos menos líquidos, de 0,004512 para 0,005569.

Tabela 4: Tabela indicando as principais estatísticas do método CS para o ano de 2020

	CS				
	Média	DesvPad	Mediana	Máximo	Mínimo
Pré-pandemia	0.003874	0.002920	0.003328	0.033160	0.000000
Pandemia	0.005386	0.004826	0.004269	0.051537	0.000000
Pós-pandemia	0.004464	0.003433	0.003793	0.035070	0.000000

Fonte: Elaborado pelo autor

Já o estimador EDGE, na tabela 5, apresentou um aumento mais acentuado nas médias dos spreads bid-ask em 2020. Para ativos mais líquidos, a média passou de 0,008308 no período total para 0,011808 em 2020, um aumento de cerca de 42%. Para ativos menos líquidos, o aumento foi

de aproximadamente 43%, de 0,008709 para 0,012413.

Tabela 5: Tabela indicando as principais estatísticas do método EDGE para o ano de 2020

	EDGE				
	Média	DesvPad	Mediana	Máximo	Mínimo
Pré-pandemia	0.007509	0.004545	0.006791	0.051456	0.000016
Pandemia	0.012225	0.010940	0.008892	0.113239	0.000018
Pós-pandemia	0.008244	0.004891	0.007400	0.072938	0.000071

Fonte: Elaborado pelo autor

O estimador AR, na tabela 6, registrou os maiores valores médios de spread bid-ask em ambos os períodos analisados. Em 2020, a média para ativos mais líquidos aumentou em cerca de 40%, de 0,011013 para 0,015371, enquanto para ativos menos líquidos, o aumento foi de aproximadamente 36%, de 0,012382 para 0,016809.

Tabela 6: Tabela indicando as principais estatísticas do método AR para o ano de 2020

	AR				
	Média	DesvPad	Mediana	Máximo	Mínimo
Pré-pandemia	0.009903	0.005820	0.009098	0.072874	0.000000
Pandemia	0.016079	0.013420	0.012332	0.112654	0.000000
Pós-pandemia	0.011866	0.006788	0.010749	0.074177	0.000046

Fonte: Elaborado pelo autor

Esses resultados podem indicar uma maior sensibilidade dos estimadores EDGE e AR às condições de volatilidade do mercado, capturando de forma mais pronunciada o aumento nos custos

de transação durante períodos de alta volatilidade.

A análise evidencia que, em períodos de alta volatilidade, como o ano de 2020, os estimadores EDGE e AR tendem a fornecer estimativas de spread bid-ask superiores ao estimador CS. Isso pode sugerir que EDGE e AR capturem de maneira mais eficaz o aumento dos custos de transação e a diminuição da liquidez que tipicamente ocorrem em mercados voláteis.

O estimador CS, por sua vez, embora tenha registrado aumentos nos spreads em 2020, ainda apresenta valores consideravelmente menores, o que pode indicar uma limitação em sua capacidade de refletir plenamente as condições de mercado adversas. Essa limitação pode ser atribuída à suposição implícita no modelo CS de negociação contínua e alta liquidez, que não se sustenta em períodos de turbulência do mercado.

Os estimadores EDGE e AR demonstraram maior robustez em condições de baixa liquidez e alta volatilidade. O EDGE, com seu termo de correção analítico para negociação infrequente, consegue manter estimativas não viesadas mesmo quando a frequência de negociações diminui. O AR também se mostrou eficaz nesse contexto, estimando spreads maiores que refletem os aumentos nos custos de transação enfrentados pelos investidores.

5.3 Análise Comparativa com Dados de Referência da Bloomberg

Para avaliar a eficácia dos estimadores estudados em relação a uma referência de mercado amplamente reconhecida, realizou-se uma análise comparativa entre os spreads bid-ask estimados pelos métodos EDGE, CS e AR e os valores reportados pela Bloomberg. Esta comparação é fundamental para determinar qual metodologia oferece estimativas mais precisas no contexto específico do mercado brasileiro.

5.3.1 Teste de Similaridade Estatística

O teste de Kolmogorov-Smirnov (KS) foi aplicado para verificar se as distribuições dos spreads estimados pelos três métodos são estatisticamente similares à distribuição dos valores de referência da Bloomberg. Este teste não paramétrico avalia a hipótese nula de que duas amostras provêm da mesma distribuição populacional.

Tabela 7: Tabela indicando o resultado do teste de Kolmogorov-Smirnov para cada metodologia de estimativa de spread bid-ask

Estatística D	p-valor	Resultado	Metodologia
0.981907	0.000000	Distribuições diferentes	Corwin-Schultz
0.928711	0.000000	Distribuições diferentes	Abdi e Ranaldo
0.948159	0.000000	Distribuições diferentes	EDGE

Fonte: Elaborado pelo autor

Conforme a tabela 7, os resultados do teste KS revelaram p-valores iguais a zero e estatísticas D próximas a 1 para todas as três metodologias, indicando uma rejeição contundente da hipótese nula. Isso evidencia que as distribuições dos spreads estimados por CS, AR e EDGE são estatisticamente diferentes da distribuição dos spreads de referência da Bloomberg. Estas diferenças não são meramente aleatórias, mas refletem divergências sistemáticas nas estimativas. Este resultado sugere que os três métodos provavelmente capturam diferentes aspectos dos spreads ou possuem vieses estruturais distintos em relação aos valores de referência. A rejeição da hipótese nula para todos os estimadores indica a complexidade inerente à mensuração do spread bid-ask em mercados emergentes, onde fatores específicos podem não ser adequadamente capturados pelos modelos desenvolvidos originalmente para mercados mais líquidos e eficientes.

5.3.2 Métricas de Erro

Para quantificar a precisão das estimativas em relação aos valores de referência, foram calculadas duas métricas de erro: o Erro Quadrático Médio (RMSE) e o Erro Absoluto Médio (MAE). Estas métricas permitem uma avaliação objetiva da magnitude das diferenças entre os valores estimados e os valores de referência. Os resultados das métricas de erro revelaram diferenças significativas entre as metodologias:

O estimador EDGE apresentou desempenho superior, com RMSE de 0,9658148 e MAE de 0,3256090, valores substancialmente menores que os observados para as outras metodologias. Os estimadores CS e AR mostraram desempenho semelhante entre si, com RMSE em torno de

2,74. O estimador AR apresentou um MAE ligeiramente menor (0,4326942) em comparação ao CS (0,4397678).

Estes resultados indicam que, embora nenhuma das metodologias reproduza perfeitamente a distribuição dos spreads Bloomberg (conforme indicado pelo teste KS), o estimador EDGE consegue gerar aproximações pontuais consideravelmente mais precisas. O erro quadrático médio do EDGE é aproximadamente 65% menor que o dos outros estimadores, o que representa uma vantagem significativa em termos de precisão.

5.3.3 Interpretação Integrada

A análise integrada dos resultados do teste KS e das métricas de erro permite uma compreensão mais completa do desempenho relativo dos estimadores no contexto do mercado brasileiro. Apesar de todas as metodologias produzirem distribuições estatisticamente diferentes dos valores de referência, o EDGE demonstrou capacidade superior para aproximar-se dos valores individuais, o que é evidenciado por seu menor RMSE e MAE. Essa característica é particularmente relevante em aplicações práticas, onde a precisão pontual das estimativas é frequentemente mais importante que a reprodução exata da distribuição completa.

A superioridade do EDGE pode ser atribuída à sua abordagem mais abrangente, que incorpora informações dos preços de abertura, máxima, mínima e fechamento, conforme descrito por Ardia, Guidotti e Kroencke (2024). Além disso, o EDGE inclui um termo de correção analítico para a negociação infrequente, característica comum em mercados menos líquidos como o brasileiro. Por outro lado, o desempenho similar dos estimadores CS e AR em termos de métricas de erro, apesar de suas diferentes formulações matemáticas, sugere que ambos podem estar sujeitos a limitações semelhantes quando aplicados ao mercado brasileiro. Isso pode estar relacionado a pressupostos subjacentes a esses modelos que não se aplicam plenamente ao contexto de mercados emergentes.

5.3.4 Implicações para o Mercado Brasileiro

Os resultados desta análise comparativa têm implicações significativas para a escolha de metodologias de estimativa do spread bid-ask no mercado brasileiro:

O estimador EDGE emerge como a opção mais adequada para aplicações práticas, ofere-

cendo estimativas com erro médio absoluto de aproximadamente 0,33 pontos base em relação aos valores de referência da Bloomberg. O fato de todas as metodologias produzirem distribuições estatisticamente diferentes da Bloomberg sugere que fatores específicos do mercado brasileiro, como a concentração de liquidez em determinados ativos e períodos, podem não estar sendo adequadamente capturados pelos modelos existentes. Estes achados corroboram os achados de Ardia, Guidotti e Kroencke (2024) que sugere que o EDGE é mais robusto em cenários de baixa liquidez e mercados emergentes, como mencionado no referencial teórico.

Em suma, enquanto nenhuma das metodologias analisadas reproduz perfeitamente os valores de referência da Bloomberg, o estimador EDGE oferece uma alternativa mais precisa quando os dados de alta frequência não estão disponíveis. Esta conclusão é particularmente relevante para pesquisadores, investidores e reguladores interessados na microestrutura do mercado brasileiro e em seus custos de transação.

6 Conclusão

Este estudo teve como objetivo geral comparar, no mercado acionário brasileiro, estimativas de spread bid-ask obtidas pelas abordagens de Corwin & Schultz (2012), Abdi & Ranaldo (2017) e EDGE (Ardia, Guidotti & Kroencke, 2024), tendo como referência a cotação de spread divulgada pela Bloomberg. A pesquisa desenvolveu-se a partir do reconhecimento da importância fundamental do spread bid-ask como componente da microestrutura de mercado, representando a diferença entre os preços de compra (bid) e venda (ask) de um ativo, refletindo assim os custos de transação e a liquidez do mercado.

A amostra utilizada foi composta por 138 ações listadas na B3, selecionadas com base em critérios de liquidez, onde foram consideradas apenas ações com índice de liquidez superior a 0,1, calculado em uma janela de seis meses. As ações foram distribuídas entre diferentes segmentos de governança corporativa: 96 pertencentes ao Novo Mercado, 16 ao Nível 2, 18 ao Nível 1 e 4 ao segmento Tradicional.

Metodologicamente, o estudo aplicou os três estimadores selecionados aos dados de preços diários (abertura, máxima, mínima, fechamento) e volumes obtidos diretamente da B3, utilizando uma janela móvel de 21 dias para garantir consistência com estudos anteriores. As estimativas foram comparadas com valores de referência da Bloomberg através do teste de Kolmogorov-Smirnov e métricas de erro como RMSE e MAE.

Os resultados do teste de Kolmogorov-Smirnov indicaram que as distribuições dos spreads estimados por todos os três métodos são estatisticamente diferentes da distribuição dos valores da Bloomberg, o que sugere que nenhum dos estimadores reproduz perfeitamente a complexidade do spread bid-ask no mercado brasileiro. Esta constatação aponta para a existência de características específicas deste mercado emergente que podem não estar sendo adequadamente capturadas pelos modelos desenvolvidos originalmente para mercados mais líquidos e eficientes.

A pesquisa se baseou principalmente nos trabalhos seminais dos autores dos próprios estimadores: Corwin e Schultz (2012), que propuseram uma metodologia baseada em preços diários de máxima e mínima; Abdi e Ranaldo (2017), que expandiram essa abordagem incluindo preços de fechamento; e Ardia, Guidotti e Kroencke (2024), que desenvolveram o estimador EDGE incorpo-

rando informações adicionais dos preços de abertura.

Os principais resultados demonstraram que, embora nenhum dos estimadores reproduza perfeitamente a distribuição dos spreads Bloomberg (conforme indicado pelo teste KS), o estimador EDGE apresentou desempenho significativamente superior, com RMSE de 0,9658148 e MAE de 0,3256090, valores aproximadamente 65% menores que os dos estimadores CS e AR. Durante o período de alta volatilidade de 2020, todos os estimadores registraram aumento nas médias dos spreads, com EDGE e AR demonstrando maior sensibilidade às condições adversas de mercado.

Esta pesquisa contribui para a literatura ao fornecer evidências sobre a eficácia relativa de diferentes metodologias de estimação do spread bid-ask especificamente no contexto de um mercado emergente como o brasileiro. Para estudos futuros, os resultados oferecem uma base para o desenvolvimento de estimadores mais adaptados às especificidades dos mercados emergentes, que tipicamente apresentam menor liquidez e maior volatilidade. A identificação do EDGE como estimador mais preciso também fornece orientação metodológica para pesquisadores interessados em mensurar custos de transação no mercado brasileiro.

Entre as limitações do estudo, destaca-se que os modelos foram originalmente calibrados e testados em mercados desenvolvidos, com estruturas distintas do brasileiro. Os pressupostos subjacentes a esses modelos, como a hipótese de negociação contínua, podem não se aplicar adequadamente ao contexto brasileiro, especialmente em ativos com menor liquidez. Adicionalmente, particularidades estruturais do mercado brasileiro, como a dinâmica da formação de preços na B3 e a concentração de liquidez em determinados horários, podem não ser completamente capturadas pelos modelos analisados.

7 Referências

- ABDI, F.; RANALDO, A. A simple estimation of bid-ask spreads from daily close, high, and low prices. **The review of financial studies**, v. 30, n. 12, p. 4437–4480, 2017.
- ARAÚJO, G. S.; BARBEDO, C. H. DA S.; VICENTE, J. V. M. The adverse selection cost component of the spread of Brazilian stocks. **Emerging markets review**, v. 21, p. 21–41, 2014.
- ARDIA, D.; GUIDOTTI, E.; KROENCKE, T. A. Efficient estimation of bid-ask spreads from open, high, low, and close prices. **SSRN Electronic Journal**, 2023.
- BURTON GORDON MALKIEL. **A random walk down Wall Street : the time- tested strategy for successful investing**. New York: W.W. Norton, 2007.
- CATTIVELLI, L.; PIRINO, D. A SHARP model of bid–ask spread forecasts. **International journal of forecasting**, v. 35, n. 4, p. 1211–1225, 2019.
- CHORDIA, T.; ROLL, R.; SUBRAHMANYAM, A. Commonality in liquidity. **Journal of financial economics**, v. 56, n. 1, p. 3–28, 2000.
- CONROY, R. M.; HARRIS, R. S.; BENET, B. A. The effects of stock splits on bid- ask spreads. **The journal of finance**, v. 45, n. 4, p. 1285–1295, 1990.
- CORREIA, L. F.; AMARAL, H. Determinantes da liquidez de mercado de ações negociadas. **Revista Brasileira de Economia**, v. 68, n. 1, p. 89-110, 2014.
- CORWIN, S. A.; SCHULTZ, P. A simple way to estimate bid-ask spreads from daily high and low prices. **The journal of finance**, v. 67, n. 2, p. 719–760, 2012.
- DE JONG, F.; RINDI, B. **The Microstructure of Financial Markets**. Cambridge: Cambridge University Press, 2009.
- DEMSETZ, H. The cost of transacting. **The Quarterly Journal of Economics**, v. 82, n. 1, p. 33, 1968.
- EASLEY, D.; O’HARA, M. Price, trade size, and information in securities markets. **Journal of financial economics**, v. 19, n. 1, p. 69–90, 1987.

- FAMA, E. F. The behavior of stock-market prices. **The journal of business**, v. 38, n. 1, p. 34, 1965.
- FAMA, E. F. Efficient capital markets: A review of theory and empirical work. **The journal of finance**, v. 25, n. 2, p. 383, 1970.
- FERREIRA, E. et al. Impacto assimétrico do sentimento do investidor na volatilidade do mercado acionário brasileiro. **Revista Brasileira de Gestão de Negócios**, v. 23, n. 2, p. 34-52, 2021.
- GEORGE, T. J.; KAUL, G.; NIMALENDRAN, M. Estimation of the Bid-ask Spreads and its Components: A New Approach. **Review of Financial Studies**, p. 623–656, 1991.
- GLOSTEN, L. R.; MILGROM, P. R. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. **Journal of financial economics**, v. 14, n. 1, p. 71–100, 1985.
- HARRIS, L. **Trading and Exchanges: Market Microstructure for Practitioners**. Oxford: Oxford University Press, 2003.
- HENDERSHOTT, T.; JONES, C. M.; MENKVELD, A. J. Does algorithmic trading improve liquidity? **The journal of finance**, v. 66, n. 1, p. 1–33, 2011.
- HOLDEN, C. W.; JACOBSEN, S. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions: Liquidity measurement problems in fast, competitive markets. **The journal of finance**, v. 69, n. 4, p. 1747–1785, 2014.
- HUANG, R. D.; STOLL, H. R. The components of the bid-ask spread: A general approach. **The review of financial studies**, v. 10, n. 4, p. 995–1034, 1997.
- ICHIMURA, D.; VIDEIRA, R.; RIPAMONTI, A. Asymmetric information and daily stock prices in Brazil. **Estudios gerenciales**, p. 465–472, 2020.
- JENSEN, M. C. Some anomalous evidence regarding market efficiency. **Journal of Financial Economics**, v. 6, n. 2-3, p. 95–101, jun. 1978.
- OLIVEIRA, R. **Como o fator de iliquidez afeta o retorno das ações brasileiras: Estudo Empírico**. São Paulo: Insper, 2018.
- PEAT, M. Market data resources for researchers: The SIRCA data repository. **The Australian**

economic review, v. 42, n. 4, p. 490–495, 2009.

PINTO, P. **Análise dos principais componentes do bid-ask spread de opções sobre ações no mercado brasileiro**. São Paulo: Insper, 2018.

RIPAMONTI, A. Corwin-Schultz bid-ask spread estimator in the Brazilian stock market. **Brazilian Administration Review**, v. 13, n. 1, p. 76–97, 2016.

ROLL, R. A simple implicit measure of the effective bid-ask spread in an efficient market. **The journal of finance**, v. 39, n. 4, p. 1127, 1984.

RYAN, Jeffrey; ULRICH, Joshua. Quantmod: Quantitative Financial Modelling Framework. Disponível em: <https://github.com/joshuaulrich/quantmod>. Acesso em: 24 nov. 2024.

STOIKOV, S. The micro-price: a high-frequency estimator of future prices. **Quantitative finance**, v. 18, n. 12, p. 1959–1966, 2018.

STOLL, H. R. Inferring the components of the bid-ask spread: Theory and empirical tests. **The journal of finance**, v. 44, n. 1, p. 115–134, 1989.

TOLEDO, R. **Determinantes da remuneração do spread de certificados de recebíveis do agronegócio no mercado brasileiro**. São Paulo: Fundação Getúlio Vargas, 2016.