Universidade Federal Rural do Semi-Árido
Centro de Ciências Sociais Aplicadas e Humanas
Departamento de Ciências Sociais Aplicadas
Curso de Ciências Contábeis – Mossoró-RN

Trabalho de Conclusão de Curso (2021)

SENTIMENTO TEXTUAL E RETORNOS DE CARTEIRAS DE AÇÕES BRASILEIRAS: UMA ANÁLISE BASEADA EM NOTÍCIAS NO MERCADO NORTE AMERICANO

Mikelline Carla De Lima Melo[1], Kléber Formiga Miranda[2], Lucas Lúcio Godeiro[3]

Resumo: A pesquisa analisou a relação entre retornos de ações de carteiras e índices de sentimento baseado em notícias relativas ao mercado acionário americano, justificada pela influência dos investidores estrangeiros no volume negociado na bolsa brasileira (46%, em média, no ano de 2020). Partiu-se do pressuposto de que esses investidores utilizarão as notícias como parâmetro para tomada de decisão em seus investimentos, já que o Brasil oferece taxas de juros mais altas, caracterizado como poder de atratividade para tal, embora associadas ao seu risco. A análise, com dados mensais, abrange o período de 1999 a 2019. Foram aplicados em pesquisa um dicionário dinâmico de palavras, o qual utilizada uma lista de termos variam com o tempo, seguindo o poder preditivo que essas palavras possuem sobre o retorno do $S\&P_{500}$. Após definido o índice, avaliou-se o impulso-resposta do seu efeito sobre carteiras criadas com base em características e empresas listadas no mercado acionário brasileiro. A análise realizada por meio de vetores autorregressivos (VAR) demonstrou que o índice de notícias, captado pelo dicionário dinâmico, possui influência nos retornos das carteiras baseadas no tamanho e no índice *Market-to-book*. No geral, os resultados sugerem haver um retardamento de 2 a 3 meses na reação dos retornos das ações brasileiras ao choque das notícias do mercado americano. Nessa linha, percebe-se a possibilidade de efeitos ocorridos no mercado americano demorarem a refletir para o mercado brasileiro. Como implicação prática, tem-se que eventos ocorridos no mercado americano podem ser usados como preditores para alterações nos preços das ações do mercado brasileiro até 3 meses após seu acontecimento. A hipótese de que o retorno de carteiras de ações do mercado acionário brasileiro sofre influência de notícias publicadas sobre o mercado americano não foi rejeitada. Contudo, o teste de robustez com índice gerado com dicionário estático apresentou resultados divergentes.

Palavras-chave: Sentimento textual. Retorno das ações. Carteiras. Notícias. Dicionário dinâmico.

1 INTRODUÇÃO

A análise textual em seus mais diversos segmentos da ciência é objeto de discussões e pesquisas, para os mais diversos objetivos (Das, 2014; Frazier et al., 1984; Kos et al., 2017). Em finanças e contabilidade, a análise do tom das notícias vem sendo amplamente discutida, boa parte desta discussão são pesquisas embasadas em relatórios financeiros (Huang et al., 2014; Kearney & Liu, 2014; Loughran & McDonald, 2011), notícias de jornais especializados (Galdi & Gonçalves, 2018; L. S. D. A. Lima & Aragón, 2019) e alguns sobre o tom dos discursos dos CEO, como em conferência de chamada (Kothari et al., 2009; Viana Junior et al., 2019). Enriquecendo, assim o sentido de avaliar o impacto das notícias, bem como sua influência na sensibilidade do preço de ações do mercado financeiro.

Os investimentos realizados em nível global, considera a oportunidade de alocar recursos de locais menos arriscados, em detrimento daqueles mais arriscados. Dessa forma, notícias consideradas relevantes para os investidores podem influenciá-los no sentido de

movimentar seu capital entre países considerando oportunidades de melhores retornos. Considerando o mercado acionário americano como um dos mais maduros e o de dívida como o mais seguro, investidores podem basear-se em notícias americanas para decidir sobre investimentos entre países. Dada a relevância dos investidores estrangeiros no mercado acionário brasileiro (C. P. Silva & Machado, 2020), as notícias relacionadas ao mercado americano, podem influenciar o retorno de ações no mercado brasileiro.

Assim, esta pesquisa se concentra na análise de possíveis oscilações nos preços das ações de carteiras de ações no mercado brasileiro serem afetadas por indicadores de sentimento textual. Com a emergência de técnicas de *machine learning* voltadas a criação de indicadores de sentimento baseados notícias financeiras publicadas em relevantes jornais do mercado americano. Tem-se como objetivo analisar a relação entre indicadores de sentimento textual e o retorno de carteiras compostas por ações negociadas na Brasil, Bolsa, Balcão (B3).

Vale salientar que nesse artigo foram utilizados dois modelos de dicionários para identificação do sentimento do investidor, o modelo textual estático e o dinâmico. Sendo assim, como modelos base do estudo temos o dicionário textual estático de Loughran & McDonald (2011), onde as palavras não variam com o tempo e o dicionário dinâmico de Lima e Godeiro (2020), onde as palavras variam com o tempo. Como as palavras contidas no dicionário dinâmico sempre mudam, entende-se que este pode ser mais fiel a atualidade, independente de em qual momento a análise da pesquisa está acontecendo, devido sua capacidade de identificar palavras dentro do texto e usa-las para derivar novos preditores, do que o dicionário textual estático, que utiliza sempre as mesmas palavras.

A pesquisa avança ao incorporar métricas de sentimento textual, atualmente difundidas nas análises de trabalhos acadêmicos e decisões corporativas em geral (Galdi & Gonçalves, 2018; Huang et al., 2014; Kearney & Liu, 2014; Kothari et al., 2009; L. S. D. A. Lima & Aragón, 2019; Viana Junior et al., 2019). Considerando a possibilidade de notícias serem capazes de afetar as decisões de investidores, há uma ampla discussão a ser enfrentada com a facilidade de acesso à informação por mídias sociais ou em notícias veiculadas via *web*.

No geral, os resultados demonstram diferença entre análises realizadas com dicionários de palavras estático e dinâmico. Ao se analisar o índice de sentimento baseado na mudança de palavras com o tempo (dinâmico), observou-se um *delay* de até 3 meses no efeito do retorno das carteiras de ações brasileiras. Embora o estudo não tenha avançado para as motivações desse retardo, algumas explicações podem decorrer do aproveitamento do otimismo no mercado americano.

Além dessa introdução, esta pesquisa possui uma seção dedicada à revisão da literatura na qual se destacam aspectos relevantes para a relação pretendida na pesquisa. Em seguida apresenta-se a metodologia, contendo os delineamentos metodológicos da pesquisa. A quarta seção trata da análise de resultados e a quinta das considerações finais. Ao final são apresentadas as referências.

2 SENTIMENTO TEXTUAL E RETORNO DAS AÇÕES

A disponibilidade de informações por diversos canais permitiu a todas as áreas usarem de técnicas de *big data*, ciência de dados e aprendizado de máquina para avançar nos seus respectivos conhecimentos. Esse contexto permitiu, sobretudo, utilizar de técnicas de leitura estruturada e não estruturada de textos de notícias e documentos oficiais, consolidando a era do *big text* (Das, 2014). Com base nas informações disponíveis os investidores passam a contar com mais um instrumento para avançar na análise dos fundamentos das empresas e incorporar novos atributos aos preços das ações (Tetlock et al., 2008). A análise textual e o uso de *big data* passa a exercer um papel fundamental na forma como os diversos *stakeholders* processam e interpretam os dados, resultando em informação relevante para influenciar o processo de avaliação das empresas (Loughran & McDonald, 2016).

O discurso de gestores, por exemplo, está diretamente atrelado ao valor das ações. O sentimento de otimismo transmitido em conferências acabam gerando associação positiva com os retornos anormais (Viana Junior et al., 2019). Na mesma linha, divulgações de grandes corporações, de analistas especializados e de imprensa de negócios afetam diretamente as medidas de risco da empresa. Divulgações negativas resultam no aumento do custo de capital, enquanto as positivas as reduzem, apresentando em ambos os vetores, volatilidade no retorno das ações (Kothari et al., 2009). Outro exemplo de efeito de notícias sobre as ações das empresas é a análise de notícias relacionadas ao acidente da Samarco em que constatou-se o aumento do risco sistemático (beta) da ação da Vale (M. D. de O. P. da Silva & Machado, 2019).

Ao analisar a sensibilidade do valor agregado pelo mercado financeiro, relacionado às informações que são divulgadas, é possível verificar o potencial de influência das notícias sobre o preço. Nessa perspectiva, Lima & Aragón (2019) verificaram que o sentimento de notícias financeiras coletadas do The New York Times e do The Wall Street, em um período de 1980 a 2017, possuem poder preditivo sobre a variação de todas as taxas de câmbio analisadas. Esses achados reforçam a influência de notícias coletadas em canais voltados ao setor financeiro na macroeconomia. Na mesma linha, essas notícias podem afetar alterações nos preços das ações de mercados de capitais de países diversos.

Outros achados revelam que o pessimismo exagerado da mídia provoca queda nos preços das ações disponíveis no mercado, precedendo a um alto volume de negociações, podendo ser seguida de uma possível reversão, diante de verificações nos fundamentos que contradigam ou relevem tais afirmações (Tetlock, 2007). Um exemplo ocorreu com o preço de grandes montadoras, após a divulgação do escândalo de fraude de emissões da Volkswagen, quando foi reduzido em torno de 20% (Griffin & Lont, 2016).

Consoante a isso, as palavras negativas e de incerteza que são divulgadas pela mídia financeira especializada, têm influência negativa com a rentabilidade e positiva com a volatilidade dos índices de mercado (Galdi & Gonçalves, 2018). Reforçando, portanto, que a sensibilidade, por parte dos investidores, às notícias e informações disponibilizadas pelas empresas de capital aberto provoca influências no movimento de indicadores gerais do mercado financeiro de ações, de maneira clara, criando-se outras notícias em decorrência disso.

Observa-se, por conseguinte, que o volume e o tom das notícias veiculadas na mídia influenciam o mercado acionário de diversas formas, pontuando-os como fatores importantes. Deste modo, quando as divulgações realizadas pela mídia apresentam indícios de risco sistêmico no mercado ou reflexos de uma maior exposição a riscos, atrelado a um período de incertezas o beta das ações pode ser influenciado (Bianconi et al., 2015; M. D. de O. P. da Silva & Machado, 2019).

Nesse contexto, chegamos à possibilidade de compreender de qual forma as notícias podem impactar em diversas economias mundiais. Entender como os riscos políticos de determinado países, bem como os globais afetam o retorno das ações e a economia é fundamental para se obter êxito em investimentos (Pástor & Veronesi, 2013; Smales, 2014). Compreende-se que determinados setores podem ser mais sensíveis a notícias políticas e econômicas do que outros. Assim como, países parceiros comerciais de indústrias dependentes do comércio apresentam maior volatilidade quanto a riscos políticos existente entre eles (Boutchkova et al., 2012). Deduz-se, portanto, que diante da atual conjuntura, em que as informações e transações trafegam rapidamente entre continentes, grandes níveis de incerteza política e econômica podem reverberar por diversos países com base em informações e notícias de qualquer lugar.

Quando as informações de incerteza são provenientes de países desenvolvidos, a incerteza tende a se alastrar para todo o mundo. Afinal, os países desenvolvidos normalmente são o *benchmark* de locais para investimento seguro. Ao comparar dados do Reino Unido, EUA, França, Alemanha, Itália e Canadá demonstrou-se que os EUA e Reino Unido são os grandes responsáveis pelas repercussões da incerteza, enquanto os demais são apenas receptores dos choques dessa incerteza política (Klößner & Sekkel, 2014). Assim,

informações do Reino Unido e dos EUA são relevantes para outros países, convergindo com a proposta desta pesquisa de avaliar efeitos de notícias sobre o mercado americano em carteira de ações do mercado de capitais brasileiro.

Considerando o efeito *spill-over*, em que um evento acontece em decorrência de outro, verificou-se que o mercado de ações americano exerce influência nos mercados de capitais dos BRICS. Assim, os BRICS monitoram de perto as notícias os EUA e por conseguinte suas possíveis implicações no mercado, principalmente em períodos de instabilidade econômica (Dakhlaoui & Aloui, 2016). Sob o ponto de vista da incerteza política, é possível verificar indicativos de que a previsão de modelos macroeconômicos podem ser viabilizados com índices de incerteza como o *Economic Policy Uncertainty* (EPU), contribuindo para análises com uso de variáveis não observáveis (Abid, 2020).

Assim, considerando o efeito das notícias em todo o ambiente macroeconômico de um país, esta pesquisa avaliará a relação entre um indicador de sentimento textual, baseado em notícias sobre o mercado americano, e o retorno de carteiras de ações no mercado brasileiro, sob a perspectiva de que o otimismo das notícias revelado pelo índice afetará positivamente os retornos das ações brasileiras. Tem-se, portanto, a hipótese de pesquisa (H_p):

H_p: O retorno de carteiras de ações do mercado acionário brasileiro sofre influência de notícias publicadas sobre o mercado americano.

3 METODOLOGIA

A coleta de dados foi realizada em três plataformas, quais sejam a plataforma da Economatica® e os sites dos jornais The Wall Street Journal (WSJ) e The New York Times (NYT). O período de coleta dos dados foi de 1999 a 2019, em frequência mensal. As informações coletadas no Economatica foram o retorno das firmas e informações utilizadas para caracterização das firmas conforme seu tamanho (valor de mercado) e índice *market-to-book*, além do volume diário de negociação para filtro por liquidez.

A Tabela 1 demonstra o número médio mensal de empresas por ano constante na amostra. Onde verificamos que no decorrer dos anos o número de empresas cresce, aumentando a viabilidade da amostra e de análise dos dados, visando um resultado de pesquisa ainda mais esclarecedor.

Tabela 1 Quantidade de empresas da amostra por ano

1999	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
21	23	34	46	47	47	56	57	62	76	101
2010	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	
144	152	149	151	157	151	143	135	130	139	

A coleta de notícias escritas em língua inglesa, nos sites dos jornais WSJ e NYT, foi realizada de forma manual, enquadradas nos critérios dispostos no Quadro 1.

Quadro 1 Critérios para coleta de notícias no WSJ e NYT

Seção	**Termos para filtro**
• Notícias Econômicas • Notícias do Mercado Financeiro • Notícias de Commodities	S&P 500 index, Dow Jones index, Bond yield, oil price, stock index, stock market, DJIA, SPX, equity market, Dow Jones Industrial Average, S&P 500, Standard & Poor's 500-stock index, U.S. treasury rates, U.S. treasury yield, stock price, earnings surprise, earnings surprises, oil prices, Nasdaq Composite or standard & poor's 500 index. 30-year Treasury bond, 10-year Treasury bond, 30-year Treasury, 10-year Treasury, Bond prices, bull market or bear market, bull market or bear market, bullish, bearish, financial markets, financial market, volatile, volatility, Nasdaq

3.1 FORMAÇÃO DAS CARTEIRAS

A formação das 15 carteiras seguiu o critério de classificação baseado no tamanho das empresas e no seu índice *market-to-book*. Inicialmente, calculou-se o logaritmo natural do valor de mercado das empresas para representar o tamanho das firmas e o índice *market-to-book* pela razão entre o valor de mercado e o patrimônio líquido das empresas, sendo:

1) 3 carteiras para tamanho, sendo classificadas como SMALL aquelas empresas com valor de mercado inferior ao valor do percentil 33%, BIG aquelas com valor superior ao percentil 66% e MEDIUM as demais empresas que ficavam entre os dois percentuais citados.
2) 3 carteiras para índice *market-to-book*, sendo classificadas como LOW aquelas empresas com valor de mercado inferior ao valor do percentil 33%, HIGH aquelas com valor superior ao percentil 66% e MEDIUM as demais empresas que ficavam entre os dois percentuais citados.
3) 9 carteiras com a combinação de todas as criadas nos itens 1 e 2, ficando: SMALL-LOW, SMALL-MEDIUM, SMALL-HIGH, MEDIUM-LOW, MEDIUM-MEDIUM, MEDIUM-HIGH, BIG-LOW, BIG-MEDIUM, BIG-HIGH.

As carteiras com qualquer classificação MEDIUM não foram analisadas, com o intuito de evitar que a amostra apresentasse um possível viés na análise dos dados. Assim as análises apresentadas se concentraram nas 8 carteiras cuja classificação se concentrou no extremo das distribuições. Os retornos das 8 carteiras analisadas são plotados na Figura 1.

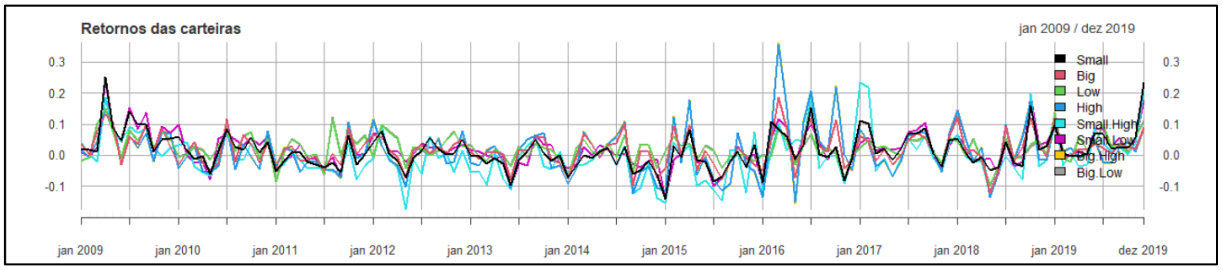

Figura 1. Retorno das Carteiras

3.2 GERAÇÃO DO ÍNDICE DE SENTIMENTO TEXTUAL

O dicionário que gera o índice de sentimento textual utilizado, como já foi citado anteriormente, seguiu a metodologia proposta por Lima e Godeiro (2020), na qual usa o dicionário variante no tempo ao invés dos dicionários fixos, até então utilizados. O primeiro passo para a construção do índice é a coleta das notícias financeiras do WSJ e do NYT. O próximo passo consiste em efetuar-se a contagem de palavras. A partir de matriz *Document Term Matrix* (DTM), que contém a contagem das palavras, usa-se aprendizado de máquina supervisionado, por meio de *Elastic Net*, para classificar as palavras em positivamente e negativamente correlacionadas com o retorno do índice $S\&P_{500}$. Então, calcula-se o índice de sentimento como a proporção das palavras negativamente correlacionadas com o $S\&P_{500}$ em relação ao total de palavras selecionadas. Ou seja, sendo X_{neg} o somatório de palavras negativas que apareceram nas notícias e X_{pos} o número de palavras positivas, o índice de sentimento S foi calculado conforme Equação 1:

$$S = \frac{X_{neg}}{1 + X_{pos} + X_{neg}} \tag{1}$$

Desta maneira, se o valor do índice for de 0,6, indica que 60% da informação contida nas notícias de determinado mês é de conotação negativa. Cabe ressaltar que a vantagem do dicionário variante no tempo em relação ao fixo consiste na capacidade do mesmo em captar palavras de acordo com o contexto econômico e financeiro do momento. Por exemplo, a palavra coronavírus, que não está presente nos dicionários fixos, pode indicar um sentimento negativo. Essa palavra pode ser selecionada pelo dicionário variante no tempo.

A Figura 2 apresenta o gráfico das séries temporais dos índices de sentimento usados na pesquisa. Onde é possível observar que o dicionário dinâmico mantém uma variação no índice de sentimento estável no tempo, enquanto que o dicionário estático apresenta uma variação claramente irregular.

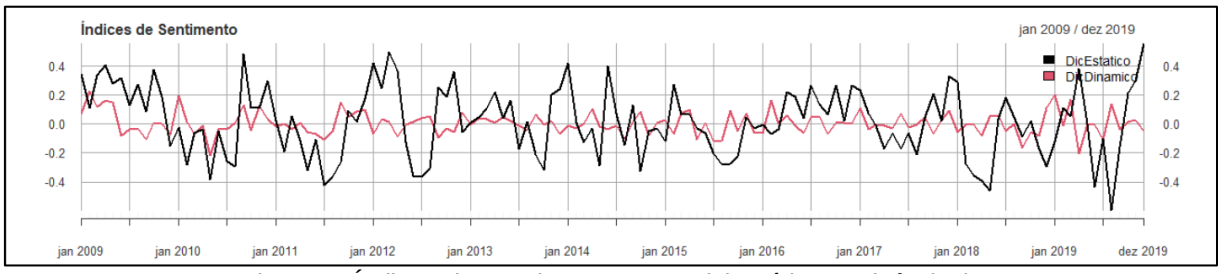

Figura 2. Índices de Sentimento Textual (Estático e Dinâmico)

3.3 ANÁLISE DA HIPÓTESE DE PESQUISA

Para análise da hipótese de pesquisa utilizou-se o modelo de vetores autorregressivos (VAR), cuja estimação pressupões a estacionariedade das séries. Como a relação pretendida pela hipótese de pesquisa envolve o retorno que já possui estacionariedade, por construção, avaliou-se que a estacionariedade do índice de sentimento por meio da sua primeira diferença, dada a rejeição da hipótese nula do teste de Zivot-Andrews, a qual determina a presença de raiz unitária na série analisada. Em seguida, foram avaliados os *lags* (defasagens) ótimas para análise da estimação do VAR. O critério para escolha do *lag* ótimo obedeceu ao valor mínimo do critério de *Schwarz* – SC(n), dentre os 10 primeiros *lags* testados.

Logo após realizou-se o teste de Granger com o intuito de verificar a ordem das causas das séries analisadas. Como pela hipótese de pesquisa e, pela lógica, os retornos futuros devem ser afetados por notícias pretéritas e não o inverso, pois não seria razoável imaginar notícias presentes serem impactadas por retornos futuros desconhecidos, o teste de Granger permitiu identificar se, em alguma relação o retorno causaria sentimento. Nesses casos, a relação foi desconsiderada como desfavorável à hipótese de pesquisa. O critério para não rejeição da hipótese de pesquisa foi realizado para cada carteira, seguindo a linha de que todas deveriam ter influência das notícias norte americanas, sob a justificativa de que a movimentação de capitais para o mercado acionário brasileiro se daria para todos os tipos de ações, deste modo todas foram analisadas igualitariamente, novamente evitando que houvesse algum viés na análise da amostra.

4 ANÁLISE DE RESULTADOS

A análise da hipótese de pesquisa se concentra em duas variáveis relativas ao índice de sentimento textual e retorno de carteiras. A Tabela 2 apresenta as estatísticas descritivas de todas as variáveis utilizadas nos modelos VAR da pesquisa.

Tabela 3 Estatísticas descritivas

Variáveis	Posição		Distribuição				
	Média	Desvio	Mínimo	Q1	Mediana	Q3	Max
Painel 1 – Índice de Sentimento							
D1sentDinam	-0.001	0.091	-0.314	-0.056	-0.001	0.053	0.288
Painel 2 – Retornos das carteiras							
SMALL	0.018	0.084	-0.262	-0.022	0.011	0.059	0.741
BIG	0.025	0.073	-0.283	-0.020	0.023	0.072	0.295
LOW	0.029	0.067	-0.256	-0.008	0.027	0.066	0.297
HIGH	0.015	0.089	-0.317	-0.040	0.011	0.068	0.367
SMALL-HIGH	-0.002	0.091	-0.378	-0.051	-0.004	0.043	0.277
SMALL-LOW	0.028	0.112	-0.208	-0.017	0.017	0.068	1.457
BIG-HIGH	0.015	0.090	-0.309	-0.041	0.012	0.070	0.392
BIG-LOW	0.029	0.068	-0.262	-0.008	0.027	0.066	0.296

Nota. D1sentDinam: primeira diferença do índice de sentimento textual dinâmico (L. R. Lima & Godeiro, 2020); Small: carteira formada por empresas com capitalização inferior ao percentil 33%; Big: carteira formada por empresas com capitalização superior ao percentil 66%; Low: carteira formada por empresas com índice *Market-to-Book* inferior ao percentil 33%; High: carteira formada por empresas com índice *Market-to-Book* superior ao percentil 66%; Small-High: carteira formada por empresas classificadas como Small e High no mesmo mês; Small-Low: carteira formada por empresas classificadas como Small e Low no mesmo mês; Big-High: carteira formada por empresas classificadas

como Big e High no mesmo mês; Big-Low: carteira formada por empresas classificadas como Big e Low no mesmo mês; Q1: primeiro quartil; Q3: terceiro quartil.

De acordo com a Tabela 2, o índice de sentimento, medido pela sua primeira diferença, capta variações negativas (pessimismo das notícias) até a metade dos dados observado (Mediana = -0,001) que coincide com o mesmo valor de seu ponto médio. O desvio padrão de 0,091 demonstra dispersão no índice calculado em todo o período. Como os valores correspondem a um período relativamente longo (1999-2019), esse desvio da média é, de certa forma, esperado, dadas as nuances econômicas acontecidas nos períodos, incluindo crises.

As carteiras analisadas apresentam retornos que variam de -0.002 (SMALL-HIGH) a 0.029 (LOW e BIG-LOW). Os desvios-padrão dos retornos das carteiras representam, no mínimo, 2 vezes os valores das respectivas médias, corroborando a noção de dispersão entre os valores dos retornos das carteiras, tornando-as diferentes. Essa diferença também é visualizada nas diferenças entre os quartis e valores de máximo e mínimo. A carteira que merece destaque é a SMALL-HIGH por ser a única que apresenta valores negativos de retorno em até metade de seus valores (mediana = -0.004). A carteira SMALL é a que apresenta maior retorno (0.741), justificado pelos baixos preços de empresas com baixa capitalização, já que uma variação pequena em valores pequenos pode se refletir em percentuais expressivos.

Portanto, mesmo com algumas divergências de valores médios, o comportamento não é muito discrepante entre as carteiras. Desse modo, espera-se que o efeito das notícias do mercado americano seja similar (mesma direção) para todas. Para complementar a análise de similaridade entre as carteiras, apresenta-se, na Tabela 3, as correlações entre os retornos de todas as carteiras analisadas.

Tabela 4 Correlação entre os retornos das carteiras

Carteiras	SMALL	BIG	LOW	HIGH	SMALL HIGH	SMALL LOW	BIG HIGH
BIG	0.640						
LOW	0.575	**0.895**					
HIGH	0.683	**0.806**	0.648				
SMALL-HIGH	0.726	0.600	0.530	0.651			
SMALL-LOW	**0.883**	*0.481*	*0.454*	0.514	*0.437*		
BIG-HIGH	0.666	**0.802**	0.646	**0.999**	0.619	0.508	
BIG-LOW	0.528	**0.887**	**0.998**	0.631	0.513	*0.395*	0.629

Nota. Todas as correlações possuem significância estatística a 1%; Em negrito destaca-se as correlações superiores a 80% e em itálico as correlações inferiores a 50%.

As correlações apresentadas na Tabela 3 demonstram alta associação entre os retornos das carteiras analisadas com destaque para aquelas com grande capitalização (BIG), dado o valor da maioria de seus coeficientes serem superior a 80%. A maior associação (0.999) foi verificada entre a carteira de ações com alto índice *market-to-book* (HIGH) com aquelas de alta capitalização e alto índices *market-to-book* (BIG-HIGH). Essas associações convergem com a noção de os retornos dessas carteiras possuem movimentos similares em seus preços.

4.1 ANÁLISE DA HIPÓTESE DE PESQUISA (H_P)

Para análise da Hipótese de Pesquisa serão apresentados os valores correspondentes ao impulso-resposta dos 3 primeiros meses e o teste de Granger. Para não rejeição de H_p, espera-se que o teste de Granger revele que apenas o índice de sentimento gera retorno, mas o retorno não gera índice de sentimento. Adicionalmente, os valores apresentados não devem conter o valor zero entres os seus limites inferior (LI) e superior (LS). A Tabela 4 demonstra os resultados para análise de H_p em relação à carteira de empresas com baixa capitalização (SMALL).

Tabela 5 Efeito do índice de notícias nas carteiras classificadas pelo tamanho

Painel 1 – Carteira com baixa capitalização (SMALL)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	0.009	-0.001	0.018
2	**0.014**	**0.006**	**0.021**	0.002	-0.005	0.011
3	0.003	-0.005	0.011	0.004	-0.006	0.013
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.011				Retorno não causa Sentimento - pvalor: 0.164		
Painel 2 – Carteira com alta capitalização (BIG)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	0.003	-0.006	0.014
2	**0.021**	**0.012**	**0.030**	0.005	-0.004	0.013
3	0.006	-0.001	0.013	0.002	-0.007	0.012
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.000				Retorno não causa Sentimento - pvalor: 0.171		

Os resultados da Tabela 4 demonstram que H_p não é rejeitada quando se analisa carteira de empresas classificadas pela capitalização. Empresas com baixa capitalização sofrem um efeito positivo (Painel 1: 0.014) das notícias do mercado americano 2 meses após o indicador de sentimento apontar para otimismo nas notícias sobre o mercado americano. A H_p também não é rejeitada quando se analisa carteira de ações com alta capitalização. Assim como a carteira SMALL, a carteira BIG sofre influência positiva (Painel 2: 0.021) do otimismo de notícias americanas com defasagem de 2 meses. Em ambas as carteiras o teste de Granger convergiu com a noção de que o sentimento gera retorno e retorno não gera sentimento das notícias.

Tabela 6 Efeito do índice de notícias nas carteiras classificadas com base no índice *market-to-book*

Painel 1 – Carteira com baixo índice *market-to-book* (LOW)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	0.002	-0.007	0.016
2	**0.015**	**0.009**	**0.022**	0.003	-0.007	0.012
3	0.004	-0.003	0.010	0.000	-0.008	0.006
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.000				Retorno não causa Sentimento - pvalor: 0.448		
Painel 2 – Carteira com alto índice *market-to-book* (HIGH)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	0.007	-0.002	0.015
2	**0.020**	**0.012**	**0.028**	0.003	-0.008	0.012
3	**0.008**	**0.001**	**0.018**	-0.002	-0.011	0.007
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.000				Retorno não causa Sentimento - pvalor: 0.402		

Tabela 7 Efeito do índice de notícias nas carteiras classificadas como baixa capitalização, desmembradas com base no índice *market-to-book*

Painel 1 – Carteira com baixa capitalização e alto índice *market-to-book* (SMALL-HIGH)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	0.008	-0.002	0.017
2	**0.011**	**0.002**	**0.018**	0.001	-0.005	0.010
3	0.005	-0.004	0.014	0.004	-0.005	0.010
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.088				Retorno não causa Sentimento - pvalor: 0.177		
Painel 2 – Carteira com baixa capitalização e baixo índice *market-to-book* (SMALL-LOW)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	0.006	-0.009	0.012
2	**0.016**	**0.005**	**0.027**	-0.001	-0.010	0.007
3	0.002	-0.009	0.012	0.003	-0.004	0.012
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.061				Retorno causa Sentimento - pvalor: 0.039		

A Hp não foi rejeitada ao separar empresas de baixa capitalização entre aquelas com alto e baixo índice *market-to-book*. Os coeficientes do impulso-resposta de 2 meses após o período de otimismo revelado pelo índice de sentimento textual das notícias do mercado americano foram positivos (Tabela 6 / Painel 1: 0.011; Tabela 6 / Painel 2: 0.016). O teste de Granger demonstrou que o retorno causa sentimento (Painel 2), contudo o intervalo entre o limite inferior (LI) e o superior (LS) revela a probabilidade de conter o valor zero. Quando se avalia se o sentimento causa retorno tem-se resultado marginal com 10% de significância (Tabela 6 / Painéis 1 e 2: 0.088 e 0.061, respectivamente).

A Tabela 7 revela os resultados das carteiras com alta capitalização, divididas com base no índice *market-to-book*.

Tabela 8 Efeito do índice de notícias nas carteiras classificadas como alta capitalização, desmembradas com base no índice *market-to-book*

Painel 1 – Carteira com alta capitalização e alto índice *market-to-book* (BIG-HIGH)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.007**	**0.001**	**0.015**
2	**0.021**	**0.011**	**0.029**	0.003	-0.008	0.012
3	0.008	-0.002	0.018	-0.002	-0.011	0.006
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.000				Retorno não causa Sentimento - pvalor: 0.420		
Painel 2 – Carteira com alta capitalização e baixo índice *market-to-book* (BIG-LOW)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	0.002	-0.009	0.012
2	**0.015**	**0.008**	**0.023**	0.003	-0.007	0.010
3	0.004	-0.002	0.009	0.000	-0.009	0.006
Teste de Granger						
Sentimento causa Retorno - pvalor: 0.000				Retorno não causa Sentimento - pvalor: 0.407		

As carteiras de ações com alta capitalização, desmembrada entre empresas com alto de baixo índice *market-to-book* também apresentaram comportamento similar às demais carteiras analisadas (Tabela 7), não sendo possível rejeitar a H_p quando se analisa o efeito de notícias do mercado norte americano nos retornos dessas carteiras.

Os resultados da Tabela 7 revelam um aumento do retorno das carteiras BIG-HIGH (Painel 1: 0.021) e BIG-LOW (Painel 2: 0.015) 2 meses após o otimismo no mercado americano captado pelo índice de sentimento textual baseado em notícias do WSJ e NYT. Diferente das outras análises, o impulso-resposta demonstrou, também, um efeito positivo do retorno sobre o sentimento (Painel 1: 0.007), contudo o teste de Granger demonstrou que retorno não causa sentimento (Painel 1: pvalor = 0.420).

No geral, todas as análises convergiram para a não rejeição da H_p, demonstrando a utilidade de um índice de sentimento textual dinâmico para modelagens no processo decisório de investimento em ações no mercado acionário brasileiro.

4.2 TESTE DE ROBUSTEZ

Para fins de análise de sensibilidade realizou-se as mesmas estimações anteriores utilizando um índice de sentimento textual estático proposto por Loughran e McDonald (2011). A intenção foi avaliar a robustez de um indicador de sentimento dinâmico (L. R. Lima et al., 2020), frente ao estático. A Tabela 8 apresenta as estatísticas descritivas.

Tabela 9 Estatísticas descritivas do índice de sentimento estático

Variáveis	Posição		Distribuição				
	Média	**Desvio**	**Mínimo**	**Q1**	**Mediana**	**Q3**	**Max**
Índices de Sentimento							
D1sentEstat	-0.001	0.282	-0.849	-0.200	0.015	0.196	0.732

Nota. D1sentEstat: primeira diferença do índice de sentimento textual estático (Loughran & McDonald, 2011); Q1: primeiro quartil; Q3: terceiro quartil.

Diferente do índice dinâmico, a Tabela 8 demonstra que os valores do índice são positivos (otimismo) na metade dos dados (Mediana = 0.015), embora o valor médio seja igual em ambos os índices (-0.001). A seguir são apresentados, na Tabela 9, os resultados do impulso-resposta com índice dinâmico (por questão de espaço, foram apresentados 2 meses).

Tabela 10 Testes de robustez para carteiras individuais baseadas no tamanho e índice *market-to-book*

Painel 1 – Carteira com baixa capitalização (SMALL)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.064**	**0.032**	**0.089**
2	0.005	-0.003	0.012	**0.077**	**0.041**	**0.103**
Painel 2 – Carteira com alta capitalização (BIG)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.054**	**0.028**	**0.076**
2	0.006	-0.002	0.014	**0.063**	**0.036**	**0.086**
Painel 3 – Carteira com baixo índice *market-to-book* (LOW)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.043**	**0.009**	**0.067**
2	0.006	-0.002	0.012	**0.062**	**0.032**	**0.086**
Painel 4 – Carteira com alto índice *market-to-book* (HIGH)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.053**	**0.024**	**0.078**
2	**0.011**	**0.003**	**0.019**	**0.065**	**0.034**	**0.097**

Os resultados demonstram que a H_p seria rejeitada caso utiliza-se um índice de sentimento textual estático para analisar o efeito no retorno de carteiras de ações do mercado brasileiro. Todos os coeficientes do impulso-resposta demonstraram que o retorno gera

sentimento, enquanto apenas a carteira HIGH (Tabela 9 / Painel 4) apresentou efeito do sentimento no seu retorno.

Ao se analisar a combinação de carteiras com capitalização (tamanho) e índice *market-to-book* obteve-se resultados similares, conforme Tabela 10.

Tabela 11 Testes de robustez para carteiras combinadas baseadas no tamanho e índice *market-to-book*

Painel 1 – Carteira com baixa capitalização e alto índice *market-to-book* (SMALL-HIGH)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.054**	**0.020**	**0.082**
2	0.006	-0.002	0.016	**0.063**	**0.032**	**0.090**
Painel 2 – Carteira com baixa capitalização e baixo índice *market-to-book* (SMALL-LOW)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.066**	**0.031**	**0.088**
2	0.006	-0.004	0.016	**0.065**	**0.024**	**0.100**
Painel 3 – Carteira com alta capitalização e alto índice *market-to-book* (BIG-HIGH)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.052**	**0.022**	**0.071**
2	**0.011**	**0.004**	**0.020**	**0.064**	**0.025**	**0.088**
Painel 4 – Carteira com alta capitalização e baixo índice *market-to-book* (BIG-LOW)						
Meses	**Impulso resposta Sentimento-Retorno**			**Impulso resposta Retorno-Sentimento**		
	Valor	**LI**	**LS**	**Valor**	**LI**	**LS**
1	0.000	0.000	0.000	**0.039**	**0.002**	**0.064**
2	**0.006**	**0.001**	**0.011**	**0.059**	**0.029**	**0.086**

Mesmo com a modificação da estratégia de carteiras, combinando as características adotadas na pesquisa, foi possível rejeitar a H_p. Os resultados da Tabela 10 indicam, conforme impulso-resposta que o retorno gera sentimento já no primeiro mês. Já o sentimento gera retorno no segundo mês nas carteiras BIG-HIGH e BIG-LOW (Tabela 10 / Painéis 3 e 4). Na análise individual essa relação foi apresentada apenas na carteira HIGH (Tabela 9 / Painel 4).

Assim, o teste de robustez indicou que a utilização de um índice de sentimento textual baseado em notícias de forma dinâmica com desempenho superior ao dicionário estático quando se busca prever o retorno de carteiras de ações do mercado brasileiro.

5 CONSIDERAÇÕES FINAIS

O estudo analisou a relação entre indicadores de sentimento textual e o retorno de carteiras compostas por ações negociadas na B3, sob o pressuposto de que as notícias possuem a propriedade de causar oscilação no sentimento do mercado, sendo precificada nas ações. Para fins desta pesquisa, foram utilizadas notícias financeiras do mercado americano, coletadas manualmente no The New York Times e The Wall Street Journal. O fundamento subjacente é de que investidores estrangeiros, que são influentes no mercado brasileiro, analisam a opção de alocar seus recursos em mercados seguros (americano) antes de direcioná-los para mercados arriscados (brasileiro). Assim, quando as notícias indicam otimismo no mercado americano, os recursos tenderão a convergir para esse mercado e, quando em pessimismo, a tendência é a busca por mercados arriscados.

A hipótese de que o retorno de carteiras de ações do mercado acionário brasileiro sofre influência de notícias publicadas sobre o mercado americano não foi rejeitada. Essa conclusão foi obtida ao verificar o efeito positivo do índice textual dinâmico no retorno das carteiras analisadas. No geral, um choque no índice textual gera um efeito no retorno das carteiras 2 meses após, chegando a 3 meses em poucos casos. Os resultados seguem a noção de que ocorrência de *delay* no efeito das notícias no mercado brasileiro. Ressalta-se que todas as carteiras analisadas tiveram comportamento similar frente às notícias. Isso denota o efeito macro das notícias sobre o mercado.

Em alguns casos, os resultados sugeriram que índice de sentimento foi gerado pelo retorno. Esse fenômeno seria um contrassenso, pois não se esperaria que retornos pretéritos influenciassem o índice textual, considerando a construção do índice ser dinâmica, modificando-se de acordo com índice $S\&P_{500}$. Nos casos em que o retorno gerou sentimento o teste de Granger sinalizou de forma divergente, permitindo a rejeição da Hipótese de Pesquisa. Nos demais casos, o teste de Granger sinalizou pela causa do retorno, dado o sentimento textual.

A implicação prática dos resultados ora obtidos se volta para a ocorrência de um tempo de análise, pelos investidores, para deslocarem seus recursos para o mercado brasileiro. Se um choque no índice textual dinâmico demora 2 meses para fazer efeito no mercado brasileiro, então o investidor aproveita o otimismo no mercado americano para realizar investimento naquele mercado, retornando ao mercado brasileiro, após o período otimista, em busca de retornos maiores, normalmente ofertados por mercados emergentes. Dessa forma, o acompanhamento de notícias sobre o mercado americano pode sugerir, em certa medida, estratégias de investimento no mercado brasileiro.

Para fins de robustez, todas as análises foram realizadas, considerando um índice textual estático. Nesse caso, em que o índice segue um dicionário fixo de palavras, os resultados se demonstraram divergentes daqueles provenientes de análises com dicionário dinâmico. Em todos os casos o teste de Granger indicou que a causa do índice era o retorno, passando a noção de o índice ser decorrente de retornos pretéritos. Esse efeito foi captado já em 1 mês após o choque no retorno. Então o índice reflete o choque do retorno de 1 mês atrás. Esse resultado não era esperado e rejeitaria a Hipótese de Pesquisa, caso a análise fosse realizada com índice textual construído com dicionário estático.

Todas as análises foram realizadas considerando toda a série temporal disponível 1999-2019, não captando possíveis diferenças em períodos de crise, por exemplo, e por se tratar de um período demasiado longo, este pode ter influenciado quanto ao *delay* de 2 meses no impulso resposta. Dessa forma, um avanço para esta pesquisa é a análise de momentos específicos (*bull* ou *bear market*) com adoção de técnicas como estudo de eventos e consequentemente a análise de períodos de tempos mais curtos.

REFERÊNCIAS

ABID, A. (2020). Economic policy uncertainty and exchange rates in emerging markets: Short and long runs evidence. **Finance Research Letters, 37(November 2019), 1–7.** https://doi.org/10.1016/j.frl.2019.101378

BIANCONI, M., HUA, X., & TAN, C. M. (2015). Determinants of systemic risk and information dissemination. International **Review of Economics & Finance, 38, 352–368**. https://doi.org/10.1016/j.iref.2015.03.010

BOUTCHKOVA, M., DOSHI, H., DURNEV, A., & MOLCHANOV, A. (2012). Precarious politics and return volatility. **Review of Financial Studies, 25(4), 1111–1154.** https://doi.org/10.1093/rfs/hhr100

DAKHLAOUI, I., & ALOUI, C. (2016). The interactive relationship between the US economic policy uncertainty and BRIC stock markets. **International Economics, 146, 141–157.** https://doi.org/10.1016/j.inteco.2015.12.002

DAS, S. R. (2014). Text and Context: Language Analytics in Finance. **Foundations and Trends® in Finance, 8(3), 145–261.** https://doi.org/10.1561/0500000045

FRAZIER, K. B., INGRAM, R. W., & TENNYSON, B. M. (1984). A methodology for the analysis of narrative accounting disclosures. **Journal of Accounting Research, 22(1), 318–331.** https://doi.org/10.2307/2490713

GALDI, F. C., & GONÇALVES, A. M. (2018). Pessimismo e incerteza das notícias e o comportamento dos investidores no Brasil. **Revista de Administração de Empresas, 58(2), 130–148.** https://doi.org/10.1590/s0034-759020180203

GRIFFIN, P. A., & LONT, D. H. (2016). Game changer? The impact of the VW emission cheating scandal on the co-integration of large automakerss securities. **SSRN Electronic Journal, 1, 1–38.** https://doi.org/10.2139/ssrn.2838949

HUANG, X., TEOH, S. H., & ZHANG, Y. (2014). Tone Management. **The Accounting Review, 89(3), 1083–1113.** https://doi.org/10.2308/accr-50684

KEARNEY, C., & LIU, S. (2014). Textual sentiment in finance: A survey of methods and models. **International Review of Financial Analysis, 33(Cc), 171–185.** https://doi.org/10.1016/j.irfa.2014.02.006

KLÖßNER, S., & SEKKEL, R. (2014). International spillovers of policy uncertainty. **Economics Letters, 124(3), 508–512.** https://doi.org/10.1016/j.econlet.2014.07.015

KOS, S. R., SCARPIN, J. E., & PINTO, J. S. DE P. (2017). Emoção em textos financeiros: proposta de um dicionário de termos de emoção em contabilidade e finanças na língua portuguesa. **USP International Conference in Accounting, 1–20.**

KOTHARI, S. P., LI, X., & SHORT, J. E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis. **The Accounting Review, 84(5), 1639–1670.** https://doi.org/10.2308/accr.2009.84.5.1639

LIMA, L. R., & GODEIRO, L. L. (2020). Equity premium prediction with diffusion-index, machine learning and text data. **Brazilian Finance Meeting, 20, 1–55.**

LIMA, L. R., GODEIRO, L. L., & MOHSIN, M. (2020). Time-varying dictionary and the predictive power of FED minutes. **Computational Economics, 0123456789.** https://doi.org/10.1007/s10614-020-10039-9

LIMA, L. S. D. A., & ARAGÓN, E. K. DA S. B. (2019). Previsão fora da amostra da taxa de câmbio com sentimento das notícias. **Encontro Nacional de Economia, 1–20.**

LOUGHRAN, T., & MCDONALD, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. **The Journal of Finance, 66(1), 35–65.**

LOUGHRAN, T., & MCDONALD, B. (2016). Textual analysis in accounting and finance: a survey. Journal of Accounting Research, 54(4), 1187–1230. https://doi.org/10.1111/1475-679X.12123

PÁSTOR, Ľ., & VERONESI, P. (2013). Political uncertainty and risk premia. **Journal of Financial Economics,** 110(3), 520–545. https://doi.org/10.1016/j.jfineco.2013.08.007

SILVA, C. P., & MACHADO, M. A. V. (2020). The effect of foreign investment flow on commonality in liquidity on the Brazilian stock market. **Revista Contabilidade & Finanças, 31(84), 425–443.** https://doi.org/10.1590/1808-057x201909530

SILVA, M. D. DE O. P. DA, & MACHADO, M. A. V. (2019). Índice de sentimento textual: uma análise empírica do impacto das notícias sobre risco sistemático. **Revista Contemporânea de Contabilidade, 16(40), 24–42.** https://doi.org/10.5007/2175-8069.2019v16n40p24

SMALES, L. A. (2014). Political uncertainty and financial market uncertainty in an Australian context. **Journal of International Financial Markets, Institutions and Money, 32(1), 415–435.** https://doi.org/10.1016/j.intfin.2014.07.002

TETLOCK, P. C. (2007). Giving content to investor sentiment: the role of media in the stock market. **The Journal of Finance, 62(3), 1139–1168.** https://doi.org/10.1111/j.1540-6261.2007.01232.x

TETLOCK, P. C., SAAR-TSECHANSKY, M., & MACSKASSY, S. (2008). More than words: quantifying language to measure firms' fundamentals. **The Journal of Finance, 63(3), 1437–1467.** https://doi.org/10.1111/j.1540-6261.2008.01362.x

VIANA JUNIOR, D. B. C., CASTRO, L. A., PONTE, V. M. R., & LIMA, M. C. (2019). Do the words matter? an analysis of the linguistic sentiment of earnings conference calls and abnormal stock returns in Brazilian companies. **Advances in Scientific and Applied Accounting, 12(3), 144–158.** https://doi.org/10.14392/ASAA.2019120308.