

Análise do Sentimento Textual dos Relatórios de Desempenho Trimestral das Indústrias Brasileiras

Textual Sentiment Analysis of Quartely Performance Reports of Brazilian Industry

Márcio André Veras Machado

Doutor em Administração pela Universidade de Brasília
Professor Adjunto - Centro de Ciências Sociais Aplicadas - UFPB
Cidade Universitária - Campus I, Castelo Branco – João Pessoa, PB
CEP: 58059-90

E-mail: mavmachado@hotmail.com

Maria Daniella de Oliveira Pereira da Silva

Doutoranda do Programa de Pós-graduação em Ciências Contábeis Multiinstitucional UnB/
UFPB/ UFRN

Professora Assistente do Centro de Ciências Sociais Aplicadas - UFPB
Cidade Universitária - Campus I, Bessa – João Pessoa, PB
CEP: 58059-90

E-mail: mariadaniella75@hotmail.com

Resumo

O objetivo do estudo foi investigar a dinâmica do mercado financeiro diante do desempenho empresarial e do sentimento textual dos relatórios de desempenho trimestral da indústria brasileira. Para isso, foi utilizada como técnica de mensuração do sentimento textual a métrica sugerida por Loughran e McDonald (2011), que faz uso de uma análise das palavras empregadas no processo de comunicação, para capturar o sentimento de otimismo, neutralidade ou pessimismo expresso em um texto. Já o desempenho empresarial foi mensurado mediante as *proxies* de *benchmarks* de desempenho, que foram: ganho inesperado e indicador do desempenho da estimativa dos analistas. O comportamento do mercado, foi avaliado mediante uma *proxy* para retorno anormal, calculada com base no modelo de mercado. Assim, por meio de um estudo de evento, analisou-se o comportamento do mercado no período de divulgação dos relatórios de desempenho trimestral das indústrias brasileiras, referente ao terceiro trimestre de 2014. A pesquisa foi delimitada às indústrias com ações negociadas na BM&FBOVESPA, totalizando uma amostra de 54 empresas. As evidências empíricas sugerem que o mercado reage positivamente ao ganho inesperado e reage negativamente ao sentimento textual dos relatórios trimestrais quando as empresas apresentam um desempenho abaixo das expectativas dos analistas, porém, quando a empresa supera as expectativas dos analistas, apenas o ganho inesperado provoca reações positivas sobre os retornos. Portanto, a dinâmica do mercado é influenciada pelo conteúdo dos relatórios trimestrais apenas em momentos que remetem a uma maior incerteza econômica.

Palavras-chave: Mercado financeiro; Análise de Sentimento; Desempenho empresarial.

Artigo submetido em janeiro de 2017 e aceito em março pela editora Fernanda Sauerbronn, após processo de double blind review.

Versão preliminar apresentada no VI AdCont - 2015

Abstract

This paper aims to investigate the dynamics of the financial market on the business performance and textual sentiment of quarterly performance reports of Brazilian industry. For this, it was used as a measurement technique of metrics textual sentiment suggested by Loughran and McDonald (2011), which makes use of an analysis of the words used in the communication process, to capture the sentiment of optimism, neutrality or pessimism expressed in a text. The business performance was measured by benchmarks proxies, unexpected gain and indicator of the estimate of analysts. The market behaviour was measurement by a proxy for abnormal return, calculated by the market model. Thus, through an event study, it was analysed the market behaviour in the period of disclosure of quarterly performance reports of Brazilian industries, for the third quarter 2014. The research was delimited to industries listed on the BM&FBOVESPA, comprising 54 firms. The empirical evidence suggests that the market reacts positively to the unexpected gain and negatively to the feeling of textual quarterly reports, when companies present a performance below the expectations of analysts, however, when the company beats analysts' expectations, only unexpected gain causes positive reactions about the returns. Therefore, the dynamics of the market is influenced by the content of the quarterly reports only in moments that refer to greater economic uncertainty.

Keywords: Financial market; Sentiment Analysis; Business Performance.

1 Introdução

As investigações empíricas sobre o mercado de capitais, comumente embasadas por informações quantitativas, vêm sofrendo mudanças tanto na área de finanças, quanto na área contábil, com a inclusão de pesquisas que favorecem o uso de informações qualitativas (LI, 2008; HUANG; TEOH; ZHANG, 2014).

Em finanças, muitos estudos passaram a analisar o sentimento dos investidores, o sentimento dos textos e os seus reflexos no mercado de capitais (KEARNEY; LIU, 2014). O sentimento do investidor, segundo Baker e Wurgler (2007), diz respeito à crença dos investidores sobre os fluxos de caixa futuros e sobre os riscos do investimento, que nem sempre são expressos por meio de comportamentos racionais. Já o sentimento textual, diz respeito ao modo como os textos são apresentados, em termos das palavras empregadas, podendo, assim, exprimir um sentimento ou um tom de otimismo, pessimismo ou de neutralidade.

No contexto de estudos sobre sentimento textual, os trabalhos concentram-se em avaliar a dinâmica do preço das ações associada ao tom de relatórios anuais, trimestrais e de anúncios de desempenho, como o de Henry (2008), que analisa os anúncios dos lucros, com o intuito de verificar se os investidores são influenciados pelo tom.

Outros exemplos de estudos dentro desse contexto podem ser citados, como o trabalho de Huang, Teoh e Zhang (2014), no qual foi analisado o poder informacional do tom dos anúncios dos lucros, partindo do pressuposto de que as notícias estão sujeitas à discricionariedade dos gestores; e o de Aguiar (2012), que se concentrou em analisar o comportamento do mercado brasileiro diante do sentimento textual das narrativas utilizadas nos relatórios anuais.

As narrativas contábeis correspondem às palavras empregadas no processo de comunicação da empresa com seus *stakeholders* (JONES, 2001; BEATTIE, 2014). Em geral, essas narrativas não passam por auditorias e, conseqüentemente, ficam sujeitas a manipulações.

Entretanto, é preciso levar em consideração que as informações financeiras quantitativas sozinhas não possibilitam aos investidores uma avaliação completa das circunstâncias econômicas (HUANG; TEOH; ZHANG, 2014). Nesse sentido, as narrativas possibilitam uma complementação das informações apresentadas nas demonstrações financeiras. Ao longo dos últimos 20 anos, as narrativas têm ganhado maior ênfase nos relatórios anuais em países como os Estados Unidos e Reino Unido (JONES, 2011).

No contexto de informações qualitativas que subsidiam o mercado, os relatórios de desempenho trimestral correspondem a uma importante fonte de comunicação da empresa, uma vez que apresentam informações periódicas sobre o desempenho da empresa. De acordo com a Comissão de Valores Mobiliários – CVM, no seu artigo 25, da instrução CVM nº 480/09, são informações obrigatórias, e que devem ser disponibilizadas à CVM na mesma data em que são colocadas à disposição do público.

Diante da importância informacional dos relatórios trimestrais, e considerando que o sentimento textual apresenta potencial para impactar a dinâmica dos preços das ações, o presente estudo visa responder a seguinte indagação: *qual a reação do mercado diante do desempenho empresarial e do tom linguístico apresentado nos relatórios de desempenho trimestral pela indústria brasileira?* Sendo assim, o objetivo deste estudo é verificar como o mercado reage diante do desempenho empresarial e do tom empregado nos relatórios de desempenho trimestral da indústria brasileira.

Tendo em vista que, em momentos de incerteza e de crise, os investidores ficam mais vulneráveis à utilização de informações não financeiras (HENRY, 2008; GARCÍA, 2013), e que o setor industrial brasileiro apresentou, no terceiro trimestre de 2014, um declínio no resultado econômico, chamando a atenção do mercado para uma possível crise no setor, o presente estudo visa compreender como os investidores em um momento de crise se comportam em relação ao sentimento textual dos relatórios de desempenho, verificando se o tom dos relatórios provoca reações no mercado.

O estudo corrobora com a discussão sobre a influência das informações textuais no mercado acionário, na medida em que se propõe a investigar se os investidores seriam influenciados pelo tom do conteúdo apresentado nos relatórios de desempenho, levando em consideração que o comportamento dos investidores é refletido no movimento dos preços. Nesse sentido, a realização da pesquisa é motivada pela necessidade de compreender o comportamento do sentimento textual no mercado acionários brasileiro, pois foram encontrados poucos estudos analisando o contexto brasileiro. Além disso, o estudo é motivado pela necessidade de explorar *big data* relacionados a relatórios de desempenho trimestral para a extração de informações que possam contribuir com a compreensão sobre o processo de tomada de decisão dos investidores.

O trabalho encontra-se dividido em cinco seções, considerando esta seção de introdução. A seção seguinte expõe uma revisão da literatura sobre os aspectos conceituais e empíricos referentes ao sentimento textual no mercado de capitais. Na seção três, são evidenciados os procedimentos metodológicos utilizados. Na seção quatro, são apresentados os resultados da pesquisa. E, por fim, na seção cinco, a conclusão da pesquisa.

2 Sentimento Textual e o Mercado de Capitais

As narrativas contábeis correspondem a uma importante fonte de informação sobre as empresas, visto que possibilitam a apresentação de informações incrementais relativas ao desempenho da organização, que nem sempre estão refletidas nas demonstrações financeiras (RUTHERFORD, 2005; JONES, 2011; BEATTIE, 2014). No entanto, as narrativas estão sujeitas à discricionariedade dos gestores, que escolhem o conteúdo e formas de relatar as

informações qualitativas, visto que as narrativas não estão sujeitas a normas, nem são passíveis de auditorias (JONES, 2001; CLATWORTHY; JONES, 2003, HENRY, 2008).

Assim, mesmo com a possibilidade de sofrerem manipulações, diversos estudos têm procurado investigar o reflexo da composição textual das narrativas para os indivíduos e para o mercado (CLATWORTHY; JONES, 2003; LOUGHRAN; MCDONALD, 2011; LARCKER; ZAKOZYUKINA, 2012; HUANG; TEOH; ZHANG, 2014).

Os estudos que tratam sobre narrativas, se diferem em relação a duas características (HUANG, TEOH, ZHANG, 2014), quais sejam: o (a) tipo de informação textual analisada; e o (b) tipo de análise linguística: legibilidade, viés de auto referência e sentimento do texto.

Em relação à característica da informação analisada, nas pesquisas de Li (2008) e Aguiar (2012), foram utilizados textos dos relatórios anuais. Já Clatworthy e Jones (2003) usaram como fonte de informação as mensagens dos presidentes das empresas de capital aberto, enquanto, Henry (2008), Demers e Vega (2011) e Huang, Teoh e Zhang (2014) fizeram uso das notícias sobre o desempenho trimestral.

Dentre os tipos de análise linguística, a legibilidade corresponde à complexidade de compreensão de um texto. Nessa perspectiva, Li (2008) procurou analisar a relação existente entre a legibilidade dos relatórios anuais junto ao desempenho empresarial, constatando que as empresas com baixo desempenho costumavam apresentar textos mais complexos.

Quanto ao viés de auto referência, as análises linguísticas fazem referência às investigações da presença de termos usados intencionalmente para distorcer ou manipular a realidade. Larcker e Zakolyukina (2012) utilizaram esse tipo de análise para investigar o discurso dos executivos em uma conferência de apresentação dos resultados trimestrais. No geral, os resultados obtidos indicaram a ausência de manipulações na apresentação das informações.

Já o sentimento ou tom textual analisa as palavras utilizadas, as quais, no contexto geral, podem exprimir um tom otimista/positivo, um tom pessimista/negativo ou um tom neutro, conforme a escolha dos termos utilizados no processo de comunicação.

No âmbito financeiro, as pesquisas que tratam sobre o sentimento textual procuram investigar os efeitos que o tom linguístico provoca no mercado financeiro. Nesse contexto, Henry (2008) investigou a divulgação do lucro, com intuito de diagnosticar se os investidores eram influenciados pelas narrativas apresentadas nos relatórios de desempenho e pelos *benchmarks* de lucro. Desse modo, analisou o setor de telecomunicações e a indústria de computadores, no período de 1998 a 2002, por meio de um estudo de evento. Com base em uma lista de palavras positivas e negativas propostas, e identificou que os retornos eram elevados à medida em que o tom apresentava-se mais positivo.

Demers e Vega (2011) analisaram o sentimento dos textos de divulgação dos resultados trimestrais na imprensa e o seu impacto no preço das ações. O estudo foi realizado com 20.000 anúncios de lucros, referentes ao período de janeiro de 1998 a julho de 2006. Para o cálculo do tom, foram utilizadas três metodologias linguísticas, respaldadas por dicionários: (1) dicionário de palavras proposto por Loughran e McDonald (2011) para área de finanças; (2) *Diction*; (3) *General Inquirer*, dicionário de termos proposto pela área de psicologia da Harvard. Desse modo, o tom dos anúncios foi obtido por meio de uma análise fatorial dos resultados encontrados nas metodologias. Como resultado principal, os autores observaram que os tons no anúncio dos lucros, mesmo com divulgação voluntária, influenciavam a dinâmica do preço das ações.

Huang, Teoh e Zhang (2014) procuram investigar o tom dos anúncios em relação ao seu poder informacional para investidores, levando em consideração à discricionariedade linguística. Assim, fazem uma análise de conteúdo dos textos, utilizando o dicionário de palavras proposto por Loughran e McDonald (2011). O tom dos anúncios foi determinado por um conjunto de fundamentos econômicos e por incentivos administrativos. Dessa forma, os

autores admitem a coexistência de divulgações verdadeiras e estratégicas nos anúncios, decompondo o tom em duas partes: (a) tom não discricionário – formando por componentes oriundos dos fundamentos econômicos; e (b) tom discricionário – formando por declarações provenientes dos incentivos administrativos (informações privadas e gerenciamento dos gestores). Desse modo, o tom discricionário seria uma *proxy* para o tom anormal dos anúncios. Considerando que o mercado é eficiente, e que a precificação é dada em meio às informações disponíveis no mercado, os autores encontraram evidências de gerenciamento do tom dos anúncios e que os investidores eram enganados temporariamente por tons gerenciados.

No Brasil, os reflexos do sentimento textual no mercado financeiro foi a temática dos estudos de Aguiar (2012) e Faustino (2013). Na pesquisa desenvolvida por Aguiar (2012), foi analisada a reação do mercado em relação aos relatórios anuais de 1997 a 2009. Nesse trabalho, foi proposto um glossário em português de palavras que remetiam a um contexto positivo, contexto negativo, contexto de incerteza e verbos modais. Assim, o sentimento textual foi mensurado como base na classificação sugerida no glossário, e não foram encontradas evidências significativas de que o conteúdo dos relatórios da administração influenciava a movimentação do mercado acionário brasileiro. Já o estudo de Faustino (2013) analisou a publicação de notícias e os seus efeitos no mercado financeiro. Os dados sobre notícias foram obtidos no *Lexis Nexis Academic* e a análise do sentimento foi feita com base no dicionário de palavras proposto por Loughran e McDonald (2011). O autor encontrou evidências de que a quantidade de notícias negativas desencadeava um aumento da volatilidade e do volume negociado do Ibovespa.

Diante do contexto apresentado, esta pesquisa se propõe a analisar o comportamento do mercado diante da divulgação dos relatórios de desempenho trimestral da indústria brasileira. Kearney e Liu (2014) sugerem que o sentimento textual tem alto potencial para impactar os retornos e os volumes de negociação. Dessa forma, a hipótese a ser testada no presente estudo é a seguinte:

H_a: O retorno das ações é influenciado pelo desempenho empresarial e pelo tom empregado nos relatórios trimestrais.

Espera-se, portanto, que o desempenho empresarial influencie positivamente o retorno das ações, e que o tom dos relatórios trimestrais influencie negativamente (positivamente) os retornos nos períodos de maior (menor) incerteza econômica.

3 Metodologia

3.1 Amostra e coleta de dados

A pesquisa delimitou-se a investigar o comportamento do mercado acionário no período de divulgação do resultado trimestral, analisando especificamente a reação do mercado quanto ao desempenho empresarial e ao tom ou sentimento textual dos relatórios, que apresentaram o resultado do terceiro trimestre de 2014 das indústrias brasileiras.

A escolha do setor industrial e do período para a realização da pesquisa foi motivada pela importância econômica desse setor para o país, e em função da queda no resultado econômico das indústrias ao longo dos trimestres de 2014, com o terceiro trimestre de 2014 apresentando um dos piores resultados econômicos, em comparação a 2013, que provocou a um recuo de 1,5% do Produto Interno Bruto – PIB (IBGE, 2014), sinalizando, com isso, incertezas para os investidores no setor industrial.

A pesquisa restringiu-se a análise das empresas que apresentavam ações negociadas na BM&FBOVESPA e que eram integrantes do setor industrial, conforme a classificação setorial da *Thomson Reuters*. No entanto, foram analisadas apenas as empresas que apresentaram dados concomitantes às seguintes bases de dados: Comissão de Valores Mobiliários - CVM
Machado, M. A. V.; Silva, M. D. O. P

(Relatórios de desempenho trimestral); Economatica (dados contábeis); e *Thomson Reuters* (previsão dos analistas, período de coleta 19/12/2014). Desse modo, a amostra inicial do estudo limitou-se a análise de 69 indústrias, das seguintes atividades: bens industriais (6); calçados (2); combustíveis (2); componentes elétricos (1); construção civil e operações (23); cosméticos (1) ferro e aço (6); indústria alimentícia (5); indústria farmacêutica (4); indústria química (1); óleo e gás (4); papel (3); pesca e agricultura (3); siderúrgica (1); têxtil (1); tabaco (1); e vestuário (5).

As indústrias selecionadas foram submetidas à estimação do retorno dos seus ativos em torno da data de divulgação do relatório do terceiro trimestre de 2014. Com isso, foram eliminadas da amostra 15 indústrias que não atenderam aos pressupostos de estimação do modelo de mercado, conforme resultados apresentados nas Tabelas 2 e 3. Portanto, o estudo foi realizado com uma amostra de 54 indústrias.

3.2 Variáveis

3.2.1 Retorno Anormal

Para investigar o retorno anormal do mercado, no período de divulgação do resultado trimestral à CVM, foram utilizados os procedimentos de estudo de evento, que apresenta uma série de aplicações em contabilidade e finanças, podendo ser utilizado para investigar o comportamento do mercado mediante fusões, aquisições e anúncios de lucros, por exemplo (MACKINLAY, 1997).

A implementação de um estudo de eventos está atrelada à realização de procedimentos específicos, que, segundo Mackinlay (1997), podem ser definidos como: (1) definição do evento; (2) seleção da amostra; (3) medição do retorno anormal; (4) estimação e realização de testes; (5) interpretação e conclusão.

A definição do evento consiste em delimitar o que será estudado e o período a ser analisado. Dessa forma, que o período de análise é formado pela data do evento (t_0) e pelas datas em torno do evento (t_{-x} a t_{+x}). Assim, o fruto dessa junção de datas no período análise é chamado de janela de evento (MACKINLAY, 1997).

No presente estudo, o evento consistiu na data de apresentação do relatório de desempenho trimestral de cada empresa à CVM. Como período de análise, foi utilizada uma janela de evento de -3 a +3 dias em torno da data do evento. Já a janela estimação considerou o intervalo de -48 a -8 dias que antecederam a data do evento. Considerou-se um intervalo de -5 dias entre a janela de evento e janela de estimação, para que os coeficientes não sofressem as influências oriundas do evento, conforme recomenda MacKinlay (1997). A escolha dos intervalos temporais, representada na Figura 1, seguiu a média de dias utilizados em estudos anteriores desta natureza (DEMERS; VEGA, 2011; KEARNEY; LIU, 2014).

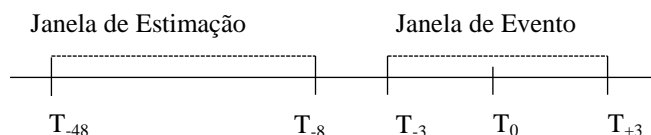


Figura 1 - Linha do tempo do estudo de evento

Uma vez definido o evento, a etapa seguinte consistiu na delimitação da amostra de empresas a serem analisadas, tomando como base critérios de exclusão relacionados à ausência de dados sobre a previsão de analistas, relatórios de desempenho trimestral e dados

contábeis. Neste trabalho, a amostra foi delimitada às empresas que integravam o setor industrial brasileiro.

A avaliação do impacto de um evento, geralmente, ocorre mediante a avaliação do retorno anormal. Mackinlay (1997) define o retorno anormal como a diferença entre o retorno observado dentro da janela de eventos menos o retorno esperado ou retorno normal, que corresponde ao retorno considerando que o evento não ocorra. Assim, o retorno anormal para uma empresa i na data t será obtido por meio da Equação 1:

$$AR_{it} = R_{it} - E(R_{it}|x_t) \quad (1)$$

Em que: AR_{it} – retorno anormal; R_{it} retorno efetivo ou retorno observado; $E(R_{it}|x_t)$ corresponde ao retorno esperado, para cada período de tempo t , condicionado a um conjunto de informações x_t .

A mensuração do retorno anormal pode ser obtida por meio de distintos modelos, como o CAPM (*Capital Asset Pricing Model*) e o modelo de mercado, o qual foi adotado em pesquisas anteriores desta natureza (HENRY, 2008; LOUGHRAN; MCDONALD, 2011; AGUIAR, 2012; HUANG; TEOH; ZHANG, 2014).

No presente estudo, utilizou-se o modelo de mercado, o qual relaciona o retorno do ativo com o retorno da carteira de mercado (Equação 2):

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \quad (2)$$

Em que: R_{it} - corresponde ao retorno do ativo i na data t ; α_i - intercepto; β_i - coeficiente angular para o ativo i , e R_{mt} corresponde ao retorno da carteira de mercado na data t ; ε_{it} – erro aleatório do ativo i na data t (com distribuição normal, média zero e variância constante).

A estimação dos parâmetros do modelo (α_i e β_i) se deu pelo método dos mínimos quadrados ordinários, levando em consideração os pressupostos de homocedasticidade, linearidade, normalidade e ausência de auto correlação. Na análise das regressões foi levado em consideração a significância dos parâmetros, por meio do teste *t-Student*, e a significância do conjunto de coeficientes, por meio do teste *F*. Adotou-se como *proxy* para a carteira de mercado, o retorno do IBOVESPA, considerando para o cálculo dos retornos dos ativos e do mercado, a capitalização contínua, conforme Equação 3:

$$R_t = \ln \frac{p_t}{p_{t-1}} \quad (3)$$

Em que: p_t - preço de fechamento do ativo no período t e p_{t-1} - preço de fechamento do ativo no período $t - 1$.

O período considerado para a estimação dos parâmetros foi de - 48 a -8 dias anteriores ao pregão da data do evento. Assim, a janela de estimação foi formada por 40 observações diárias de cada um dos ativos analisados, levando em consideração a data de divulgação do relatório de cada empresa, visto que os relatórios não costumam ser apresentados na mesma data, conforme pode ser observado na Tabela 2. Após a estimação dos parâmetros do modelo com as observações da janela de estimação, conforme modelo de mercado, foram obtidos os retornos anormais da janela de evento.

Diante disso, os retornos anormais obtidos na janela de evento, foram acumulados, para serem passíveis de interpretação. Assim, a junção dos retornos anormais formou o *Cumulative Abnormal Return* (CAR), conforme Equação 4:

$$CAR_i(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} AR_{i\tau} \quad (4)$$

Em que: CAR_i é o retorno acumulado do ativo i ; τ_1 primeiro dia da janela de evento; e τ_2 último dia da janela de evento. Portanto, o CAR_i foi utilizado como variável dependente no estudo.

Nesta pesquisa, também foi investigado o impacto da divulgação do relatório trimestral no retorno. Para isso, foi realizada uma agregação do retorno acumulado dos ativos na janela de evento (Equação 5), conforme recomenda Mackinlay (1997):

$$CAAR(\tau_1, \tau_2) = \sum_{i=1}^n \frac{CAR(\tau_1, \tau_2)}{n} \quad (5)$$

Em que: $CAAR$ é o retorno anormal agregado de uma janela de eventos (τ_1, τ_2) , estimado mediante uma média aritmética dos retornos acumulados dos eventos. Para testar a hipótese nula de que o $CAAR$ é estatisticamente igual a zero, foi utilizado o teste não paramétrico de Cowan (1992); não foram utilizados testes paramétricos em função da distribuição do retorno anormal agregado não apresentar distribuição normal.

O teste não paramétrico de Cowan (1992), conhecido como teste de sinais generalizados, examina se o número de ativos (w) com retornos anormais acumulados positivos na janela de eventos excede o valor esperado do parâmetro (\hat{p}), caso não tivesse ocorrido o evento, sendo apurado pela janela de estimação. A estatística do teste (Z_G) é obtida pela Equação 6:

$$Z_G = \frac{w - n\hat{p}}{[n\hat{p}(1-\hat{p})]^{\frac{1}{2}}} \quad (6)$$

Considerando que o parâmetro (\hat{p}) possui uma distribuição binomial, ele foi estimado conforme Equação 7:

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n \frac{1}{100} \sum_{t=-1}^{-100} S_{i,t} \quad (7)$$

Em que: \hat{p} é a expectativa baseada na fração de retornos positivos anormais no período de estimação de 100 dias. $S_{i,t}$ é a classificação do retorno anormal, admitindo $S_{i,t} = 1$, quando o retorno anormal (AR) do ativo i , no tempo t , for maior que zero; $S_{i,t} = 0$, quando o retorno anormal for igual ou menor que zero.

3.2.2 Sentimento textual dos relatórios

As narrativas dos relatórios do desempenho trimestral foram analisadas com base em *sentiment analysis*, que consiste em um método de linguagem natural, o qual possibilita que a opinião expressa em um texto seja capturada mediante a análise das palavras utilizadas (LOUGHRAN; MCDONALD, 2011).

Assim, a mensuração do sentimento textual dos relatórios foi obtida em três etapas: (a) os textos foram preparados para a leitura automatizada; (b) as narrativas foram submetidas à técnica do *vector space model* para o cálculo dos pesos das palavras; (c) calculou-se o tom linguístico dos relatórios, com base no peso das palavras.

Sentiment analysis é um método que pode ser operacionalizado por meio de diferentes técnicas, dentre elas a técnica mais comum é o *vector space model*, já utilizada em diversos estudos, como Loughran e McDonald (2011) e Aguiar (2012).

A técnica do *vector space model* reconhece as palavras como vetores e faz uso da frequência das palavras e de uma prévia categorização delas (por exemplo, palavras positivas *Análise do Sentimento Textual dos Relatórios de Desempenho Trimestral das Indústrias Brasileiras*

e negativas), para determinar se um texto apresenta um tom pessimista, otimista ou neutro. Assim, com base na frequência da palavra é calculado o seu peso em relação ao documento, no qual a mesma está inserida ou em relação a um conjunto de documentos analisados.

Neste trabalho, utilizou-se como critério de categorização das palavras para mensuração do tom dos relatórios, a lista de termos positivos e negativos, propostos por Aguiar (2012). Para a mensuração da média da palavra adotou-se uma ponderação de pesos, para reduzir a discrepância das palavras que apresentam uma alta frequência em relação às palavras com baixa frequência, com o intuito de minimizar os *outliers* (LOUGHRAN; MCDONALD, 2011; AGUIAR, 2012).

Desse modo, o cálculo da média ponderada da palavra ou do seu peso em um texto foi baseado na técnica do *vector space model*, obtido mediante o modelo proposto por Loughran e McDonald (2011), conforme Equação 8:

$$w_{i,j} \begin{cases} \frac{(1+\log(tf_{i,j}))}{(1+\log(a_j))} \log \frac{N}{df_i} & \text{se } tf_{i,j} \geq 1 \\ 0 & \text{caso contrário} \end{cases} \quad (8)$$

Em que: $w_{i,j}$ - média ponderada da palavra i no documento j ; $tf_{i,j}$ - total de ocorrências de uma palavra i em um documento j ; a_j - média de palavras contadas em um documento j ; N - total de documentos na amostra; e df_i - total de documentos com ao menos uma ocorrência da palavra i .

A operacionalização do *sentiment analysis* pode ser dada de forma manual ou automatizada, fazendo uso de programação computacional. No presente trabalho, a análise do sentimento textual foi obtida mediante um algoritmo desenvolvido no software R.

Desse modo, após a coleta dos relatórios trimestrais das indústrias selecionadas, na CVM, os documentos foram preparados para a leitura automatizada. Para isso, foram removidos dos documentos todos os gráficos, tabelas, quadros e figuras.

Mediante a mensuração do peso das palavras positivas e negativas encontradas nos relatórios de trimestrais das indústrias, foi calculado o tom dos relatórios de cada uma das empresas analisadas, por meio do índice do tom, o mesmo utilizado nos estudos de Henry (2008) e Loughran e McDonald (2011):

$$TOM_e = \frac{\sum \text{Peso das Palavras Positivas} - \sum \text{Peso das Palavras Negativas}}{\sum \text{Peso Palavras Positivas} + \sum \text{Peso Palavras Negativas}} \quad (9)$$

Desse modo, o tom da empresa (TOM_e) será otimista quanto mais próximo o índice encontrar-se de 1 (um); apresentará um tom pessimista quanto mais próximo o índice estiver de -1 (menos um); e será neutro quando o seu resultado do índice for 0 (zero). Assim, os relatórios do desempenho trimestral de cada uma das empresas estudadas foram transformados em dados quantitativos, mediante o cálculo do tom dos relatórios.

3.2.3 Desempenho empresarial

Estudos empíricos tem ressaltado que os gestores levam em consideração os *benchmarks* de desempenho para a apresentação dos cenários empresariais nos seus relatórios de desempenho (GRAHAM; HARVEY; RAJGOPAL, 2005; HENRY, 2008; HUANG; TEOH; ZHANG, 2014).

Nesse sentido, Graham *et al* (2005) realizaram uma *survey* com executivos, visando investigar os motivos que levavam os mesmos a atenderem aos *benchmarks* de desempenho/lucro. Os autores diagnosticaram que os executivos procuravam alcançar os *benchmarks*, em função da possibilidade de apresentarem perspectivas de crescimento futuro

para os investidores. Diante disso, afirmaram que eram realizados sacrifícios financeiros de modo a atender as essas previsões. Desse modo, os autores levantaram como *proxies* para os *benchmarks* de desempenho: (1) previsão dos analistas; e (2) ganho inesperado.

A previsão dos analistas corresponde à expectativa do lucro por ação que, em geral, são previsões trimestrais e anuais. Neste estudo, utilizou-se a previsão trimestral dos analistas, com o intuito de analisar se o valor do lucro por ação excedia ou não a previsão dos analistas. Assim, mediante a comparação entre o lucro por ação e a previsão dos analistas, foi construída uma variável dicotômica para avaliar o desempenho da estimativa dos analistas, nomeada de MBE, na qual foi atribuído 1 para as situações em que o lucro por ação era igual ou maior ao valor previsto pelos analistas, e 0 para as situações em que o lucro por ação era inferior ao valor previsto pelos analistas.

Quanto ao ganho inesperado, considerou-se como o resultado excedente à previsão dos analistas, mensurado mediante uma comparação entre lucro real por ação e o lucro previsto por ação (previsão dos analistas), semelhante à estimação de Henry (2008). Desse modo, o ganho inesperado foi obtido conforme Equação 10:

$$GI = \frac{EPS_r - EPS_p}{SP_t} \quad (10)$$

Em que: GI - representa o ganho inesperado; EPS_r - lucro por ação real no período; EPS_p - corresponde ao lucro por ação previsto para o período; SP_t - é preço da ação no início do trimestre. Assim, a variável apresenta o ganho que excedeu a o lucro por ação previsto para o período.

Portanto, o desempenho empresarial foi avaliado mediante as *proxies benchmarks* de desempenho: indicador do desempenho da estimativa dos analistas (MBE) e ganho inesperado (GI).

3.3 Modelo Econométrico

Realizou-se uma adaptação das metodologias empregadas por Henry (2008) e Huang; Teoh e Zhang (2014) para analisar a relação existente entre o mercado acionário, desempenho empresarial e o tom dos relatórios de desempenho trimestral. Para isso, utilizou-se um modelo de regressão linear múltipla, conforme Equação 11:

$$CAR_{i,t} = \alpha + \beta_1 (1 - MBE_{i,t}) * GI_{i,t} + \beta_2 (1 - MBE_{i,t}) * TOM_{i,t} + \beta_3 (MBE_{i,t}) * GI_{i,t} + \beta_4 TOM_{i,t} + \beta_5 Ln_AT_{i,t} + \varepsilon_i \quad (11)$$

Em que: $CAR_{i,t}$ - variável dependente, representada pelo retorno anormal acumulado na janela de evento da empresa i , no semestre t ; $GI_{i,t}$ - ganho inesperado; $MBE_{i,t}$ - variável dicotômica que classifica a previsão dos analistas em relação ao lucro por ação real da empresa i no trimestre t ; $TOM_{i,t}$ - variável quantitativa referente ao tom do relatório de desempenho do trimestre; $Ln_AT_{i,t}$ - logaritmo natural do ativo total das empresas, *proxy* para tamanho das empresas e ε_i - termo de erro.

Na equação 11, o coeficiente β_1 capta o efeito do ganho inesperado sobre o retorno anormal nos momentos em que a empresa não superou às expectativas dos analistas, que foi tratado no estudo como um período de desempenho desfavorável. O coeficiente β_2 capta os efeitos do tom dos relatórios nos momentos de desempenho desfavorável, enquanto o coeficiente β_3 apresenta o efeito do ganho inesperado nos momentos de desempenho favorável. O coeficiente β_4 apresenta o efeito do tom dos relatórios sobre o retorno anormal

Análise do Sentimento Textual dos Relatórios de Desempenho Trimestral das Indústrias Brasileiras

quando a empresa superou as expectativas dos analistas. Por fim, o coeficiente β_5 capta os efeitos provocado pelo tamanho das indústrias.

Os parâmetros da regressão linear foram estimados por meio dos mínimos quadrados ordinários. Dessa forma, foram respeitados os pressupostos de homocedasticidade, ausência de autocorrelação, normalidade dos erros e linearidade (BROOKS, 2008). Para analisar a relação existente entre o comportamento do mercado em relação ao desempenho empresarial e ao tom dos relatórios, utilizou-se o teste *t-Student*, para verificar a significância dos parâmetros, o teste *F*, para verificar a significância das variáveis em conjunto, e o R^2 ajustado, para analisar o poder explicativo do modelo.

A variável dependente $CAR_{i,t}$ corresponde ao retorno anormal acumulado, obtido mediante as Equações 1, 2 e 4. Essa variável é uma *proxy* para a reação do mercado em relação à divulgação do resultado trimestral.

O ganho inesperado $GI_{i,t}$ captura o ganho excedente do lucro por ação em relação ao lucro previsto pelos analistas. Desse modo, espera-se que o ganho excedente provoque reações positivas no mercado nos momentos em que a empresa supera ou não as expectativas dos analistas (HENRY, 2008).

Já a variável $MBE_{i,t}$ é uma *dummy* utilizada no estudo como um indicador para o desempenho apresentado pela organização em relação às expectativas dos analistas, classificado no estudo como desempenho favorável ou desfavorável. Assim, a variável serviu como base para a segregação de indústrias que apresentam uma maior incerteza econômica para os investidores.

O sentimento ou tom textual $TOM_{i,t}$ obtido por meio de uma ponderação do peso das palavras encontradas nos relatórios, conforme Equações 8 e 9, representa o posicionamento da empresa na divulgação do resultado em termos de otimismo, neutralidade ou pessimismo. Partindo do pressuposto de que os investidores estão mais vulneráveis às informações textuais em momentos que remetem a uma maior incerteza econômica, espera-se que o tom provoque reações negativas nos momentos de desempenho desfavorável (TETLOCK, 2007; GARCÍA, 2013), e que em momentos de desempenho favorável o tom provoque reações positivas (HENRY, 2008; GARCÍA, 2013)

A variável logaritmo natural do ativo total $Ln_AT_{i,t}$ foi inclusa no modelo para controlar o tamanho das empresas e evitar possíveis efeitos que o tamanho possa provocar sobre o retorno dos ativos, pois foi observado em estudos anteriores, que nos períodos de anúncios de resultados, os retornos apresentavam relação com o tamanho das empresas (BALL; KOTHARI, 1991; HENRY, 2008).

No Quadro 1, apresenta-se um resumo das variáveis e da relação esperada com a variável dependente.

Quadro 1 - Sinais esperados

Variável	Relação esperada com CAR Desempenho Favorável	Relação esperada com CAR Desempenho Desfavorável
Ganho Inesperado - <i>GI</i>	+	+
Tom dos relatórios - <i>TOM</i>	+	-

4 Análise dos Resultados

4.1 Estatística descritiva

A Tabela 1 apresenta a estatística descritiva das variáveis utilizadas no estudo. Verifica-se que, nos dados gerais, a variável dependente retornos anormais acumulados (CAR) não apresenta uma grande dispersão, pois tem um desvio padrão relativamente baixo ($dp=0.077724$) e, em média, apresentou retornos positivos (média= 0.006599).

Em relação às variáveis independentes, constata-se que o ganho inesperado – GI apresentou uma média negativa de -0.09499 e mediana -0.035900, o que significa dizer que as indústrias analisadas, em média, não apresentaram lucros que superassem às expectativas dos analistas. Quanto ao sentimento dos textos analisados, observa-se que prevaleceu um tom de neutralidade nos relatórios trimestrais, em função da média (0.0467623) e mediana (0.020301) serem próximas de zero. A variável usada para controlar os efeitos do tamanho das indústrias (Ln_AT) apresentou uma média de 20.61 e mediana de 22.33.

Ao analisar os dados conforme o desempenho apresentado no trimestre, constata-se que, dentre as indústrias que superaram às expectativas dos analistas, existe uma predominância de retornos anormais positivos (media=0.037079 e mediana=0.018074), o ganho inesperado foi positivo (media=0.108873 e mediana= 0.076740) e os comunicados sobre o desempenho do trimestral apresentam um tom de otimismo (média= 0.178071 e mediana=0.205783), o que, de certa forma, era esperado, em virtude da empresa ter alcançado resultados positivos.

Já na amostra de indústrias que não superaram às expectativas dos analistas, verifica-se que o retorno anormal acumulado, em média, foi negativo (média=-0.005124 e mediana=-0.007662), assim como, o ganho inesperado (média=-0.173405 e mediana=-0.075560). Ademais, o sentimento que prevaleceu nos comunicados do desempenho alcançado no trimestre foi de pessimismo (média=-0.008619 e mediana= -0.084446), estando associado ao fato da empresa não ter alcançado resultados positivos.

Tabela 1 - Estatística descritiva das variáveis

<i>Painel A: Dados gerais</i>						
Variável	N	Mínimo	Média	Mediana	Máximo	Desvio Padrão
CAR	54	-0.188364	0.006599	-0.000355	0.260861	0.077724
GI	54	-1.654.665	-0.094994	-0.035900	0.423190	0.283993
TOM	54	-0.718672	0.043240	0.020301	1.000.000	0.375548
Ln_AT	54	20.387	22.611	22.335	26.370	1.261
<i>Painel B: Desempenho favorável</i>						
CAR	15	-0.084114	0.037079	0.018074	0.260861	0.086798
GI	15	0.004760	0.108873	0.076740	0.423190	0.111224
TOM	15	-0.389970	0.178071	0.205783	0.967469	0.298717
<i>Painel C: Desempenho desfavorável</i>						
CAR	39	-0.188364	-0.005124	-0.007662	0.160783	0.071709
GI	39	-1.654.665	-0.173405	-0.075560	-0.000710	0.291916
TOM	39	-0.718672	-0.008619	-0.084446	1.000.000	0.393908

Fonte: dados da pesquisa

4.2 Retorno Anormal

Com base na janela de estimação do estudo de evento, foram estimados os coeficientes alfa e beta da regressão para cada uma das 69 indústrias. Os resultados encontrados nas regressões e as datas do evento analisado em cada empresa são apresentadas na Tabela 2.

Considerando um nível de significância de 5%, foram eliminadas da amostra as empresas cujo histórico de ações não apresentou o coeficiente beta significativo, bem como as empresas nas quais o *p-valor* da estatística F não foi significativo a 5%, conforme evidências apresentadas na Tabela 3.

Desse modo, foram eliminadas da amostra as empresas que apresentas as ações com as seguintes siglas: SUZB; BRKM; WEGE; MRFG; LEVE; SLCE; MAGG; KEPL; PFRM; EMBR; HRTP; FESA; AGRO; INEP.

As cinquenta e cinco empresas restantes na amostra forma submetidas à avaliação dos pressupostos de normalidade, homocedasticidade, linearidade e ausência de auto correlação. Os resultados encontram-se na Tabela 3.

Ao ser testado o pressuposto da ausência de auto correlação, conforme teste de *Durbin-Watson*, considerando uma significância de 5%, foram encontradas evidências de auto correlação em: BISA; ETER; JHSF; LPSB; PDGR; TUPY. Desse modo, as empresas foram submetidas à correção por meio da matriz robusta de *Newey-West*.

Tabela 2 - Data do Evento e Coeficientes de Regressão

Ações	Evento	Coeficientes		Teste T (p-valor)		Teste F (p-valor)	Ações	Evento	Coeficientes		Teste T (p-valor)		Teste F (p-valor)
		Alfa	Beta	Alfa	Beta				Alfa	Beta	Alfa	Beta	
AMAR	06/11/2014	-0.002	0.279	0.461	0.030	0.030	MRVE	12/11/2014	0.000	0.879	0.934	0.000	0.000
ARZZ	29/10/2014	-0.004	0.493	0.364	0.012	0.012	MULT	30/10/2014	-0.002	0.985	0.383	0.000	0.000
BBRK	19/11/2014	-0.004	0.658	0.230	0.000	0.000	MYPK	05/11/2014	-0.001	0.649	0.726	0.000	0.000
BISA	14/11/2014	0.001	0.123	0.221	0.003	0.003	NATU	24/10/2014	-0.002	0.654	0.459	0.000	0.000
BPHA	12/11/2014	-0.002	0.462	0.693	0.019	0.019	PDGR	30/10/2014	-0.005	1.373	0.206	0.000	0.000
BRF	30/09/2014	0.000	0.501	0.795	0.002	0.002	PMAM	31/10/2014	0.000	0.440	0.995	0.001	0.008
BRML	04/11/2014	-0.001	1.356	0.677	0.000	0.000	POMO	10/10/2014	0.003	1.019	0.425	0.000	0.000
BRPR	03/11/2014	-0.002	1.132	0.531	0.000	0.000	QGEP	12/11/2014	0.000	0.478	0.942	0.000	0.000
CRUZ	21/10/2014	0.000	0.631	0.962	0.000	0.000	RAPT	10/11/2014	-0.002	0.728	0.531	0.000	0.000
CSAN	05/11/2014	-0.002	0.995	0.446	0.000	0.000	RSID	14/11/2014	-0.007	1.201	0.053	0.000	0.000
CSNA	14/11/2014	-0.003	0.597	0.444	0.002	0.002	SMTO	10/11/2014	-0.002	0.681	0.406	0.000	0.000
CYRE	13/11/2014	-0.002	0.904	0.333	0.000	0.000	SSBR	05/11/2014	0.001	0.388	0.784	0.004	0.004
DIRR	13/11/2014	-0.001	0.796	0.758	0.000	0.000	TCSA	29/10/2014	-0.003	1.095	0.479	0.000	0.000
DTEX	28/10/2014	0.000	0.798	0.830	0.000	1.949	TECN	30/10/2014	-0.004	0.587	0.309	0.004	0.004
ETER	14/11/2014	-0.003	0.459	0.181	0.022	0.002	Tupy	06/11/2014	-0.001	0.336	0.626	0.006	0.006
EUCA	30/09/2014	-0.003	0.512	0.453	0.019	0.019	UGPA	05/11/2014	0.001	0.825	0.773	0.000	0.000
EVEN	13/11/2014	-0.003	1.113	0.216	0.000	0.000	USIM	28/11/2014	-0.008	0.760	0.133	0.001	0.001
EZTC	10/11/2014	-0.003	0.924	0.384	0.000	0.000	VAGR	30/10/2014	-0.014	0.454	0.000	0.003	0.002
FIBR	22/10/2014	0.001	-0.36	0.717	0.000	0.013	VALE	30/10/2014	0.000	0.281	0.925	0.029	0.029
GFSA	07/11/2014	-0.002	1.074	0.644	0.000	0.000	SUZB	30/10/2014	0.003	-0.17	0.414	0.182	0.181
GGBR	12/11/2014	-0.004	0.293	0.257	0.030	0.030	BRKM	06/11/2014	0.005	0.169	0.033	0.070	0.070
GOAU	12/11/2014	-0.002	0.304	0.557	0.016	0.016	WEGE	29/10/2014	0.002	-0.03	0.376	0.582	0.582
GRND	24/10/2014	0.003	0.535	0.406	0.002	0.002	BEEF	07/11/2014	-0.004	0.346	0.247	0.014	0.014
GUAR	10/11/2014	-0.002	0.521	0.376	0.000	0.000	MRFG	19/12/2014	0.002	0.004	0.723	0.781	0.781
HBOR	13/11/2014	0.001	0.774	0.867	0.000	0.000	LEVE	10/11/2014	0.002	0.156	0.485	0.212	0.212
HGTX	22/10/2014	-0.001	0.726	0.719	0.000	0.000	SLCE	12/11/2014	-0.003	0.089	0.578	0.646	0.646
HYPES	24/10/2014	0.000	0.586	0.956	0.000	0.000	MAGG	13/11/2014	-0.005	0.312	0.292	0.105	0.105

IGTA	03/11/2014	0.000	1.071	0.957	0.000	0.000	KEPL	13/11/2014	0.000	0.233	0.952	0.189	0.189
JBSS	12/11/2014	0.006	0.986	0.093	0.000	0.000	PFRM	19/12/2014	-0.004	0.002	0.377	0.845	0.845
JHSF	13/11/2014	-0.006	0.611	0.018	0.000	0.000	EMBR	06/11/2014	0.000	-0.07	0.885	0.565	0.565
KLBN	28/10/2014	0.000	0.360	0.996	0.003	0.003	HRTF	13/11/2014	-0.008	0.165	0.043	0.255	0.255
LLIS	05/11/2014	0.001	0.606	0.808	0.004	0.004	FESA	11/11/2014	-0.002	0.238	0.671	0.136	0.135
LPSB	14/11/2014	-0.006	0.458	0.154	0.000	0.000	AGRO	06/11/2014	-0.006	0.107	0.225	0.610	0.610
LREN	23/10/2014	0.001	0.650	0.776	0.000	0.000	INEP	18/11/2014	-0.014	-0.32	0.523	0.714	0.714
MDIA	27/10/2014	0.001	0.231	0.644	0.004	0.004							

Fonte: Dados da Pesquisa

Conforme pode ser observado na Tabela 3, ao ser testado o pressuposto da normalidade, conforme teste de *Jarque-bera*, foram levantados indícios de ausência de normalidade em: ETER; GGBR; JBSS; LLIS; LREN; MDIA; PMAM; QGEP. Entretanto, foi levado em consideração o teorema do limite central, no qual o pressuposto pode ser relaxado, ao considerar que ao aumentar o tamanho da amostra a distribuição converge à normal (BROOKS, 2008). Assim, as empresas que não apresentaram distribuição normal não foram eliminadas da amostra.

Tabela 3 - Resultados dos testes – pressupostos MQO

Empresa	DW	Jarque-bera (p-valor)	White (p-valor)	Ramsey (p-valor)	Empresa	DW	Jarque-bera (p-valor)	White (p-valor)	Ramsey (p-valor)
AMAR	1.639	0.007	0.077	0.180	IGTA	2.284	0.535	0.869	0.761
ARZZ	2.325	0.900	0.479	0.554	JBSS	2.555	0.042	0.264	0.912
BBRK	2.474	0.662	0.841	0.859	JHSF	2.699	0.190	0.815	0.969
BISA	2.865	0.328	0.786	0.800	KLBN	2.178	0.656	0.171	0.738
BPHA	2.337	0.669	0.784	0.103	LLIS	1.723	0.000	0.903	0.907
BRF	2.442	0.833	0.542	0.051	LPSB	1.047	0.118	0.442	0.442
BRML	1.873	0.653	0.173	0.101	LREN	2.297	0.015	0.621	0.863
BRPR	2.096	0.000	0.717	0.473	MDIA	2.139	0.045	0.289	0.734
CRUZ	1.675	0.979	0.931	0.461	MRVE	2.403	0.066	0.839	0.668
CSAN	1.501	0.986	0.204	0.115	MULT	2.324	0.919	0.444	0.409
CSNA	1.792	0.859	0.769	0.680	MYPK	2.496	0.856	0.142	0.172
CYRE	1.688	0.807	0.993	0.400	NATU	1.869	0.739	0.918	0.378
DIRR	2.156	0.902	0.128	0.769	PDGR	1.236	0.371	0.028	0.080
DTEX	1.949	0.749	0.623	0.104	PMAM	1.470	0.000	0.444	0.629
ETER	2.615	0.000	0.650	0.209	POMO	2.429	0.183	0.833	0.311
EUCA	2.522	0.752	0.833	0.527	QGEP	1.608	0.000	0.713	0.991
EVEN	2.267	0.652	0.004	0.986	RAPT	1.336	0.068	0.804	0.757
EZTC	2.191	0.041	0.788	0.514	RSID	2.111	0.341	0.827	0.194
FIBR	2.279	0.341	0.785	0.993	SMTO	2.158	0.441	0.607	0.198
GFSA	2.407	0.668	0.404	0.445	SSBR3	1.875	0.283	0.336	0.830
GGBR	1.892	0.002	0.748	0.800	TCSA	2.167	0.908	0.910	0.161
GOAU	1.783	0.060	0.769	0.818	TECN	1.731	0.395	0.808	0.733
GRND	2.416	0.352	0.791	0.114	Tupy	1.006	0.442	0.499	0.464
GUAR	2.493	0.581	0.674	0.651	UGPA	2.038	0.752	0.019	0.360

HBOR	2.494	0.537	0.457	0.633	USIM	1.830	0.548	0.693	0.924
HGTX	1.839	0.672	0.324	0.658	VAGR	1.528	0.098	0.412	0.091
HYPES	1.928	0.589	0.498	0.938	VALE	1.610	0.866	0.255	0.611

Fonte: Dados da pesquisa

O pressuposto da homocedasticidade foi testado conforme *teste de white*, ao nível de 5%. Verificou-se a presença de heterocedasticidade em: BEEF; AMAR; EVEN; PDGR; UGPA. Essas empresas foram submetidas a uma correção por meio da matriz robusta de *white*. Entretanto, após a correção, o beta da empresa BEEF não apresentou significância no nível desejado e a empresa foi eliminada no estudo.

Ao ser testada a linearidade, conforme teste de Ramsey, verificou-se que as empresas não apresentaram problemas de ausência de linearidade. Portanto, após a verificação dos pressupostos dos mínimos quadrados ordinários e devidas correções, restaram 54 empresas para o cálculo do retorno anormal.

O estudo também procurou investigar o efeito da divulgação do relatório trimestral sobre o mercado. Para isso, foi realizada uma agregação dos retornos anormais acumulados das indústrias, conforme Equação 5, e realizado o teste não paramétrico de sinais generalizados. Os resultados encontram-se na Tabela 4.

Tabela 4 - Divulgação do resultado trimestral das indústrias

CAAR							
CAAR (-3,+3) = 0.0067							
Estatística		<i>p-valor</i>		Estatística		<i>p-valor</i>	
Teste de sinais generalizado		0,6573	0,511	0,9885	0,3291	0,5608	0,8734

Fonte: Dados da pesquisa

Conforme pode ser observado nos resultados apresentados na Tabela 4, o mercado não apresenta reações à divulgação do resultado trimestral, visto que não foi encontrada significância estatística em nenhuma das janelas de evento avaliadas. Desse modo, pode-se dizer que, ao nível de significância de 5%, o retorno anormal é estatisticamente igual a zero, ou seja, a divulgação do relatório trimestral não provocou efeitos sobre o mercado acionário.

4.3 Resultado da métrica utilizadas para as palavras

Em relação aos relatórios trimestrais, os mesmos foram submetidos a uma leitura automatizada, na qual foram extraídas as palavras e calculados os seus respectivos pesos, de acordo a métrica apresentada por Loughran e McDonald (2011) e mediante a lista de palavras positivas e negativas propostas por Aguiar (2012).

Tabela 5 - Peso das palavras obtidas com o *vector space model*

Positivas		Negativas	
palavras	peso	palavras	peso
Crescimento	0.7382151	Redução	0.7461972
Superior	0.6203534	Queda	0.6694725
Desenvolvimento	0.6126173	Impacto	0.5539929
Expansão	0.6085769	Inflação	0.5230549
Investimentos	0.5814172	Excluindo	0.5141069

Estável	0.5536995	Endividamento	0.5141069
Evolução	0.5407998	Negativo	0.5141069
Recuperação	0.5338575	Inferior	0.5045413
Consolidado	0.5265398	Negativa	0.4578907
Valorização	0.5105985	Impactando	0.4578907
Contribuição	0.5105985	Impactado	0.4578907
Investimento	0.5018636	Retração	0.4269527
Consolidada	0.4925258	Impostos	0.4084391
Positivo	0.4824957	Negativos	0.4084391
Confiança	0.4716624	Depreciação	0.4084391
Estabilidade	0.4469862	Prejuízo	0.4084391
Otimização	0.416785	Desvalorização	0.4084391
Ganho	0.416785	Impactaram	0.3870667
Consolidadas	0.416785	Baixa	0.3870667
Positivamente	0.3987123	Imposto	0.3870667
Superiores	0.3987123	Complexo	0.3870667

Fonte: Dados da pesquisa

Os resultados apresentados, na Tabela 5, mostram os pesos das 20 palavras positivas e negativas mais citadas nos documentos analisados. A palavra positiva “crescimento”, por exemplo, apresentou nos textos uma frequência de ocorrência de 86 vezes. Entretanto, não foi avaliada de acordo com a sua frequência, mas conforme o peso que representou para todo o conjunto de relatórios avaliados, que, nesse caso recebeu o peso de 0.7382, obtido de acordo com o método do *vector space model*. A mesma interpretação foi realizada para as palavras negativas.

4.4 Resultados da regressão linear

O comportamento dos retornos anormais, obtidos mediante o modelo de mercado, foi avaliado em relação ao desempenho empresarial e ao tom dos relatórios trimestrais. A Tabela 6 apresenta os resultados das variáveis utilizadas no estudo, para explicar o retorno anormal.

Tabela 6 - Resultados da regressão linear entre retornos anormais, desempenho empresarial e tom dos relatórios trimestrais

Painel A				
Variável Explicativa	Coefficiente	Erro padrão ^a	Estatística <i>t</i>	<i>p</i> -valor
Constante	0.366593	0.180730	2.028.397	0.0481
(1-MBE)*TOM	-0.061081	0.028080	-2.175.246	0.0346**
(1-MBE)*GI	0.070255	0.040749	1.724.084	0.0911***
(MBE)*TOM	-0.046911	0.055153	-0.850556	0.3992
(MBE)*GI	0.357112	0.129926	2.748.586	0.0084*
Ln_AT	-0.015923	0.008098	-1.966.316	0.0551***
Painel B				
Descrição	Valor	Descrição	Valor	
R ²	0.335263	Teste F (Estatística)	4.84180	
R ² ajustado	0.26602	Teste F (<i>p</i> -value)	0.001153	
Schwarz	-2.25514	Teste de White (Estatística)	0.870375	

Akaike	-2.47614	Teste de White (<i>p-value</i>)	0.6047
Jarque-Bera (estatística)	0.223096	Durbin-Watson ^c	2.04808
Jarque-Bera (<i>p-value</i>) ^b	0.894449	Número de Observações	54

*significante a 1%, **significante a 5% e ***significante a 10%.

^a Erro padrão estimado sem correções, pois a hipótese de variâncias homocedásticas não foi rejeitada, ao nível de significância de 5%.

^bEvidência de distribuição normal dos resíduos.

^cAusência de auto correlação dos resíduos.

Nota: Ausência da presença de alta multicolinearidade entre as variáveis, uma vez que a estatística do Fator de Inflação da Variância (FIV) oscilou entre 1.04 a 1.247.

Fonte: Dados da pesquisa

Para analisar o impacto do ganho inesperado e do tom dos relatórios sobre o retorno anormal, foi realizada uma segregação das empresas conforme o desempenho empresarial apresentado no trimestre analisado, mediante a inclusão da variável *dummy* MBE. Assim, os comportamentos dos momentos favoráveis para a organização foram capturados, no modelo, pelos coeficientes da interação entre MBE com as variáveis independentes, e os momentos desfavoráveis foram capturados pelos coeficientes da interação entre as variáveis independentes com 1-MBE.

Os resultados apontam que o ganho inesperado (GI) afeta positivamente o retorno anormal nos momentos de desempenho favorável, uma vez que foi positivo e significativo o coeficiente da variável (MBE)*GI ($p=0.0084$). Nos momentos de desempenho desfavorável, o ganho inesperado também afetou positivamente o retorno anormal, em função de ser significativo a variável (1-MBE)*GI ($p=0.0911$). Diante desses resultados, pode-se dizer que a divulgação do ganho inesperado provoca impactos sobre o mercado, independentemente da existência de um desempenho favorável ou não na organização, conforme era esperado.

Ao analisar os resultados dos coeficientes relacionados ao tom dos relatórios, constata-se que o conteúdo linguístico dos relatórios trimestrais influencia os investidores nos momentos de desempenho desfavorável, provocando reações negativas sobre o retorno anormal, visto que o coeficiente da variável (1-MBE)*TOM foi negativo e apresentou significância estatística ($p=0.0346$), conforme era esperado. Porém, nos momentos em que a empresa tem um desempenho que supera a expectativa dos analistas, existem indícios de que o tom dos relatórios não provoca reações no mercado, pois não foi encontrada significância estatísticas na variável (MBE)*TOM ($p=0.3992$).

Desse modo, foi observado, na amostra analisada, que o tom dos relatórios afeta negativamente o retorno anormal nos momentos de desempenho desfavorável, porém, não provoca reações nos momentos de desempenho favorável. O comportamento encontrado no presente estudo para a variável TOM é condizente com as evidências empíricas levantadas nos estudos de Faustino (2013) e García (2013), quanto aos indícios de que o mercado tende a reagir ao tom das notícias textuais, principalmente nos momentos de maior incerteza econômica. Uma das possíveis explicações atribuídas a esse comportamento está relacionada ao fato de que períodos de incerteza correspondem a momentos nos quais os investidores tendem a apresentar dúvidas sobre os fluxos de caixa futuro e, com isso, ficam mais vulneráveis a utilização de notícias não financeiras para melhor alinhar as suas crenças.

Portanto, as significâncias estatísticas dos coeficientes das variáveis testadas no modelo ratificam a constatação de que o retorno das ações é influenciado pelo desempenho empresarial e pelo tom empregado nos períodos de maior incerteza econômica. Porém, em momentos de menor incerteza, o retorno das ações não sofre a influência do tom dos relatórios.

5 Conclusão

Nas avaliações sobre o desempenho organizacional, as informações financeiras quantitativas configuram-se como uma importante fonte informacional, porém, quando sozinhas, não possibilitam uma avaliação completa das circunstâncias econômicas, visto que, algumas informações que agregam valor para as organizações nem sempre, estão explícitas nas demonstrações financeiras. Desse modo, as informações textuais possibilitam um incremento informacional sobre o desempenho empresarial.

Com o aperfeiçoamento das técnicas de análise de conteúdo baseadas em linguagem de processamento natural, surgiram instrumentos para a extração de atributos linguísticos, como o tom ou sentimento dos textos. No âmbito do mercado acionários, os estudos costumam investigar a dinâmica dos preços das ações relacionada ao tom ou sentimento textual de relatórios, matérias de jornais e dos comentários postados em redes sociais.

No presente estudo, foi examinada a reação do mercado no período de divulgação do resultado do terceiro trimestre de 2014, analisando especificamente o desempenho das indústrias brasileiras e o tom dos relatórios de desempenho trimestral.

O comportamento do mercado foi avaliado mediante o retorno anormal dos ativos, estimado por meio de um estudo de eventos, relacionado ao dia de divulgação do relatório trimestral à CVM. Desse modo, foi analisado o impacto da divulgação dos relatórios sobre o retorno dos ativos e a influência do tom dos relatórios e de *proxies* de *benchmarks* de desempenho sobre o retorno anormal acumulado na janela do evento.

Os resultados levantados, quanto ao impacto do tom dos relatórios sobre os retornos anormal das ações, mostram que, nos momentos em que a empresa apresenta um desempenho acionário inferior ao desempenho previsto pelos analistas, o tom dos relatórios tende a provocar reações negativas no mercado, porém, não são encontrados indícios de reações ao tom dos relatórios nos momentos em que a empresa supera as expectativas dos analistas. Desse modo, os resultados encontrados levam a constatação de que o tom dos relatórios tende a influenciar os investidores nos momentos que remetem a uma incerteza sobre os fluxos de caixa futuro dos ativos.

Já as evidências sobre o desempenho, avaliado mediante a *proxy* ganho inesperado, mostraram que a divulgação de um ganho inesperado provoca impactos positivos sobre o mercado, independentemente da existência ou não de incertezas, corroborando com as evidências levantadas por Henry (2008).

Portanto, os resultados apontam para indícios de que retorno das ações é influenciado pelo desempenho empresarial e pelo tom empregado nos relatórios trimestrais nos períodos de incerteza. Desse modo, pode-se dizer que, na amostra de indústrias analisadas, a dinâmica do mercado sofre a interferência das informações provenientes dos ganhos inesperados e do conteúdo dos relatórios trimestrais, em momentos que remetem a incertezas.

As evidências apresentadas neste estudo encontram-se limitadas à metodologia adotada para a estimação do sentimento textual, que buscou extrair informações de dados textuais mediante a utilização de linguagem de processamento natural, métrica que não tem o intuito de replicar o processo de leitura humana. Além disso, as evidências estão limitadas à amostra de indústrias brasileiras no período referente ao terceiro trimestre de 2014.

Por fim, sugere-se que pesquisas futuras sejam realizadas, utilizando outras métricas de estimação do sentimento textual e com a ampliação da amostra e do período de análise, a fim de verificar a persistência dos resultados sobre a influência do sentimento textual no comportamento no mercado acionário.

Referências

AGUIAR, M. O. **Sentiment Analysis em Relatórios da Administração Divulgados por Firms Brasileiras**. Vitória, 2012. Dissertação. Programa de Pós-Graduação em Administração da Fundação Instituto Capixaba de Pesquisas em Contabilidade, Economia e Finanças – FUCAPE.

BEATTIE, V., Accounting Narratives and the Narrative Turn in Accounting Research: Issues, Theory, Methodology, Methods and a Research Framework, **The British Accounting Review**, 2014.

BAKER, M. P.; WURGLER, J. A. Investor Sentiment in the Stock Market. **Journal of Economic Perspectives**, v. 21, n. 2, 2007, p. 129-151.

BROOKS, C. **Introductory Econometric for Finance**. Cambridge University Press, 2008.

BALL, Ray; KOTHARI, Stephen P. Security returns around earnings announcements. **Accounting Review**, p. 718-738, 1991.

CLATWORTHY, M.; JONES, M. Financial Reporting of Good News and Bad News: Evidence from Accounting Narratives. **Accounting & Business Research**, v. 33 (3), 2003, p. 171-185.

COWAN, A. R.. Nonparametric event study tests. **Review of Quantitative Finance and Accounting**, v. 2, n. 4, p. 343-358, 1992.

COMISSÃO DE VALORES MOBILIÁRIOS. **Instrução CVM nº 480**. Dispõe sobre o registro de emissores de valores mobiliários admitidos à negociação em mercados regulamentados de valores mobiliários. Rio de Janeiro, novembro, 2009.

DEMERS, E; VEGA, C. Linguistic Tone in Earnings Announcements: News Or Noise? **Working paper**, INSEAD, 2011.

FAUSTINO, C. C. R. **O impacto das notícias no mercado financeiro brasileiro**. São Paulo, 2013. Dissertação. Programa de Pós-Graduação em Economia da Escola de Economia da Fundação Getúlio Vargas – FGV.

GRAHAM, J.; HARVEY, C.; RAJGOPAL, S. The economic implications of corporate financial reporting. **Journal of Accounting and Economics**.40, 2005, p. 30–73.

GARCIA, Diego. Sentiment during recessions. **The Journal of Finance**, v. 68, n. 3, 2013, p. 1267-1300.

IBGE. PIB varia 0,1% em relação ao 2º tri de 2014 e chega a R\$ 1,3 trilhão. Disponível em<<http://saladeimprensa.ibge.gov.br/noticias?view=noticia&id=1&idnoticia=2771&busca=1&t=pib-varia-0-1-relacao-2%C2%BA-tri-2014-chega-r-3-trilhao>> Acesso em: 19 de dezembro 2014.

HENRY, E. Are investors influenced by how earnings press releases are written? **Journal of Business Communication**. 45, 2008, p.363–407.

HUANG, X; TEOH, S. H; ZHANG, Y. Tone Management. **The Accounting Review**, v.89, nº3, 2014, p. 1083-1113.

JONES, M. **Creative Accounting, Fraud and International Accounting Scandals**. John Wiley & Sons Ltd, 2011.

KEARNEY, C; LIU, S. Textual sentiment in finance: A survey of methods and models. **International Review of Financial Analysis**, 2014.

LARCKER, D; ZAKOLYUKINA, A. Detecting deceptive discussions in conference calls. **Journal of Accounting Research** 50: 495–540, 2012.

LI, Feng. Annual report readability, current earnings, and earnings persistence. **Journal of Accounting and Economics**, v. 45, 2008, p. 221–247.

LOUGHRAN, T; MCDONALD, B. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. **The Journal of Finance**, v. 66, n. 1, 2011, p. 35-65.

MACKINLAY, A.C. Event studies in economics and finance. **Journal of Economic Literature**, Nashville: American Economic Association, v. 35, n. 1, 1997, p. 13-39.

RUTHERFORD, B. Genre analysis of corporate annual report narratives: A corpus linguistics-based approach. **Journal of Business Communication**, v. 42 (4), 2005, p. 349-378.

TETLOCK, Paul C. Giving content to investor sentiment: The role of media in the stock market. **The Journal of Finance**, v. 62, n. 3, p. 1139-1168, 2007.