日興リサーチセンター
SMBC日興証券グループ

有価証券報告書に含まれるテキスト情報と企業業績の関係について

Research Report
2018 年 3 月

投資工学研究所
佐久間　洋明
田中　良典

要　約

近年、テキストマイニングや人工知能の技術を用い、テキストデータの数値化やテキストデータから投資判断を行う上で重要な文を抽出する手法の研究が盛んに行われている。その中の 1 つに、因果関係文を抽出する研究がある。因果関係文は投資判断を行う上で重要であると考えられるが、抽出された因果関係文が企業業績を実際に反映しているのかを検証した論文は筆者の調べた限り存在しなかった。そこで、本稿では有価証券報告書の「業績等の概要」に含まれる因果関係文を対象として、因果関係文の結果表現に着目して、ポジティブな表現が含まれる場合は正となるように極性を付与し、この極性が企業業績を実際に反映しているかを検証した。その結果、極性と当該期における企業業績の関係は正の相関であり、各極性の企業業績の分布や中央値の推移を見てもその関係は確認された。一方で、極性と翌期の企業業績の間に同様の関係は見られなかった。これらから、有価証券報告書の「業績等の概要」に含まれる因果関係文は、主に当該期の企業業績を反映していることが示唆された。以上より、因果関係文を用いた業績の要因分析など、投資判断に応用する上での因果関係文の有益性がさらに高まったと考えられる。

目次

1. はじめに
2. データの定義及び分析方法
　2.1　分析対象企業、有価証券報告書及び分析期間
　2.2　財務指標の計算方法
　2.3　有価証券報告書の「業績等の概要」に対する極性付与の方法
　2.4　分析方法
3. 結果
　3.1　各極性における企業数の推移
　3.2　極性と当該期の企業業績との関係
　3.3　極性と翌期の企業業績との関係
4. 考察
5. まとめ

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

1. はじめに

投資判断に利用できる基本的なデータの１つとして、財務データやマーケットデータ等の数値データがある。これらの数値データの特徴として、構造化しやすく統計的な解析等に対して、比較的容易かつ大量に扱うことが可能である点が挙げられる。一方、数値以外のデータとして、SNS、新聞記事、決算短信や有価証券報告書等に代表されるテキストデータがある。これらのテキストデータの特徴として、その時点のホットな話題、国内外の重要な出来事、企業の過去の業績や将来の業績の見通しに対する理由や根拠、今後の経営戦略や経営方針に関する情報など、投資判断を行う上で数値データにはない重要な情報を含んでいることが挙げられる。しかしながら、これらのテキストデータは構造化することが難しく、そのほとんどは投資家が直接読まなければ投資判断に利用できない場合が多かった。

近年、この問題に対してテキストマイニングや人工知能の技術を用い、テキストデータの数値化やテキストデータから投資判断を行う上で重要な文を抽出する手法の研究が盛んに行われている。前者では、例えば、あるニュースが株価にとってポジティブに働くかネガティブに働くかでニュースを定量化（極性付与）し、日次の株式リターンとの関係を分析した研究[沖本、平澤 2014][五島、高橋 2016]がある。後者では、例えば、坂地ら[坂地、増山 2011]は、原因や結果を含む文を因果関係文とし、その特徴的な表現と機械学習を用いて、新聞記事から因果関係文を自動的に抽出する手法を提案している。また、坂地ら[坂地他 2015]は、決算短信に含まれる因果関係文から原因・結果表現を抽出する手法を提案している。そして、これらの手法を用いて決算短信に含まれる因果関係文より原因表現及び結果表現を抽出し、ある企業に関連する過去の因果関係や、ある原因（例えば、猛暑）に対する結果（例えば、冷房の需要増）を含む因果関係を記述している企業を検索できるシステム[1]を開発し、公開している。佐藤ら[佐藤他 2017]は、坂地ら[坂地、増山 2011][坂地他 2015]の手法を参考にし、有価証券報告書から因果関係文を抽出する判別モデルを作成し、原因表現及び結果表現の抽出を行っている。因果関係文の抽出以外にも、決算短信から業績要因を含む文を自動的に抽出する手法を提案している研究[酒井他 2015b][2]、決算短信から業績要因を含む業績予測文を抽出する手法を提案している研究[北森他 2017]や、有価証券報告書から企業の課題を含む「課題文」を抽出する手法を提案している研究[小寺他 2018]がある。また、決算短信から抽出した業績要因に対して極性を自動的に付与する手法を提案している研究[酒井他 2015a]もある。

新聞記事、決算短信や有価証券報告書から重要な文として、過去の業績等に対する原因と結果を含む因果関係文を抽出する先行研究はある一方、抽出した因果関係文が実際の企業業績を反映しているのかを検証した論文は筆者の調べた限り存在しなかった。そこで、本稿では有価証券報告書から抽出した因果関係文に対して極性を付与し、企業業績を実際に反映しているかの検証を行った。もし、抽出した因果関係文に含まれる極性が企業業績を反映しているのであれば、その因果関係文には企業業績に関する原因と結果が含まれていることの証左となり、因果関係文を用いた企業業績の要因分析など、投資情報

[1] CS(Causal expressions Search system)として Web で公開されている。http://hawk.ci.seikei.ac.jp/CS/
[2] CEES(Causal Expressions Extraction System)として Web で公開されている。http://hawk.ci.seikei.ac.jp/cees/

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

としての因果関係文の有益性が高まると考えられる。

有価証券報告書は、金融商品取引法によりすべての上場企業が提出することを義務付けられている開示資料[3]である。そのため、SNS や新聞記事は大企業や注目を集めている企業などの限られた企業に関する情報が多いのに対して、有価証券報告書は全ての上場企業から公表されているという利点がある。また、決算短信と比較すると、有価証券報告書は速報性では劣るものの、決算短信よりも開示を義務付けられている項目が多く、「企業の概況」、「事業の状況」、「設備の状況」、「提出会社の状況」や「経理の状況」等、投資判断に利用することができると考えられる情報を多く含んでいる。そして、これらの項目の中でも「事業の状況」に記載されている「業績等の概要」は、「企業内容等の開示に関する内閣府令」において、該当項目の記載上の注意として、会計年度における業績及びキャッシュ・フローの状況について前年同期と比較して分析的に記載することを求められている[4]。つまり、該当項目には企業の業績に関する定性的な情報が記載されているため、企業業績との関係を見る上で該当項目より因果関係文を抽出することが最も適切であると考えられる。

本稿では、因果関係文が企業の業績を反映しているかを検証するために、有価証券報告書の「業績等の概要」の項目を分析対象とし、その中の因果関係文に対する極性（ポジティブ、ネガティブ）に着目した。極性以外にも、単語の頻度や共起等を用いることも考えられるが、企業や業種によって使われる単語は多種多様であり、統一的に扱うことが難しいと推測される。一方で、極性であれば、企業業績が前年と比較して良ければ文章もポジティブな表現を用いると推測されるため、統一的に取り扱うことが比較的容易であり、因果関係文が企業業績を反映しているかの検証を行う上で、因果関係文に含まれる極性を用いることは適切であると考えられる。また、本稿では因果関係文の中でも企業業績に関する結果が記載されていると考えられる結果表現に着目し、結果表現に含まれる単語に基づき因果関係文へ極性を付与することにより、「業績等の概要」に対する極性の付与を行った。その後、有価証券報告書の「業績等の概要」の極性が実際の企業業績を反映しているかの検証を行うために、以下の分析を行った。まず始めに、「業績等の概要」の極性がポジティブ（ネガティブ）であれば、当該期の企業業績は前年対比プラス（マイナス）であると仮定を置き、「業績等の概要」の極性と当該期の企業業績（売上高、営業利益、経常利益、当期利益の伸び率）の間に相関があるか分析を行った。次に、比較対象として「業績等の概要」の極性は、翌期の企業業績との間に関係はないと仮定を置き、同様の分析を行った。

2.データの定義及び分析方法

2.1 分析対象企業、有価証券報告書及び分析期間

本稿では、TOPIX1000 採用銘柄（除く金融）の中で 3 月期決算の企業かつ有価証券報告書を取得できた企業を分析対象企業とした。そして、2008 年 3 月期～2016 年 3 月期の有価証券報告書を対象と

[3] 有価証券報告書は金融商品取引法の第二十四条に基づき提出を義務付けられている。EDINET（金融商品取引法に基づく有価証券報告書等の開示書類に関する電子開示システム）から入手可能。http://disclosure.edinet-fsa.go.jp/

[4] 企業内容等の開示に関する内閣府令　第二号様式において定められている。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

し、その決算期に対応する期間を分析期間とした。

2.2 財務指標の計算方法

本稿では、企業業績を表す財務指標として、売上高、営業利益、経常利益、当期利益の伸び率を用いた。t 期の伸び率の計算方法は以下の通りである。

$$\text{財務指標}_{t}\ (\%) = \frac{\left(\text{財務数値}_{t} - \text{財務数値}_{t-1}\right)}{\left|\text{財務数値}_{t-1}\right|} \times 100$$

ただし、t 期と t-1 期が同じ会計基準ではない場合、対象外とした。

2.3 有価証券報告書の「業績等の概要」に対する極性付与の方法

佐藤ら[佐藤他 2017]に従い、有価証券報告書の「業績等の概要」から因果関係文を抽出し、その因果関係文から原因表現及び結果表現を抽出した。次に、抽出した結果表現に含まれる単語に基づき因果関係文へ極性を付与し、その集計結果を有価証券報告書の「業績等の概要」の極性とした。下記にその手順を示す。

① 有価証券報告書の「業績等の概要」から手がかり表現を含む文を抽出し、それらの文について判別モデルを用いて因果関係があるかないか判別処理を行い、因果関係文を抽出した。その後、抽出された因果関係文から構文パターンによって原因・結果表現を抽出する手法を用いて、原因表現及び結果表現を抽出した。（詳細は佐藤ら[佐藤他 2017]を参照）

② 因果関係文に含まれる結果表現に着目し、結果表現にポジティブ（ネガティブ）な単語が含まれている因果関係文はポジティブ（ネガティブ）と極性を付与した。ただし、結果表現にポジティブな単語とネガティブな単語の両方が含まれる場合は、より文の後ろに現れる単語の極性を採用した。また、1 つの因果関係文に複数の因果関係が含まれる場合も、より後ろに現れる因果関係の結果表現の極性を用いた。なお、ポジティブな単語とネガティブな単語は計 14 個（増益、減益等）を選定し、また前に現れる単語によって極性が変わる単語については、極性を判定する辞書を作成して極性の付与を行った。増加を例にとると、売上の増加はポジティブ、コストの増加はネガティブとした。因果関係文への極性付与の例を図表 1 に示す。

③ 各決算期の各社有価証券報告書について、「業績等の概要」に含まれるポジティブな因果関係文とネガティブな因果関係文の数を集計し、ポジティブの文数からネガティブの文数を引いた。そして、その差が正の場合はポジティブ、負の場合はネガティブ、 0 の場合はニュートラルの極性を付与し、その有価証券報告書の「業績等の概要」の極性とした。以下、有価証券報告書の「業績等の概要」の極性を、「極性」と呼ぶ。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表 1　因果関係文への極性付与の例

因果関係文	極性
営業利益は、コスト削減の効果などにより**増益**となった。	ポジティブ
減損損失の計上などにより**減益**となりました。	ネガティブ

注：因果関係文の点線は原因表現、実線は結果表現、太文字は結果表現に含まれる極性を付与する単語を示す。
(出所)各社有価証券報告書より日興リサーチセンター作成

2.4 分析方法

極性と当該期の企業業績の関係について、以下の分析を行った。まず全分析期間において、極性別に当該期における各財務指標の中央値を求めた。次に、極性と当該期における各財務指標の順位相関係数を計算した。続いて、極性別に各財務指標の分布を調べた。最後に、各期の各極性における各財務指標の中央値の推移について調べた。また、極性と翌期の企業業績の関係についても、同様の分析を行った。

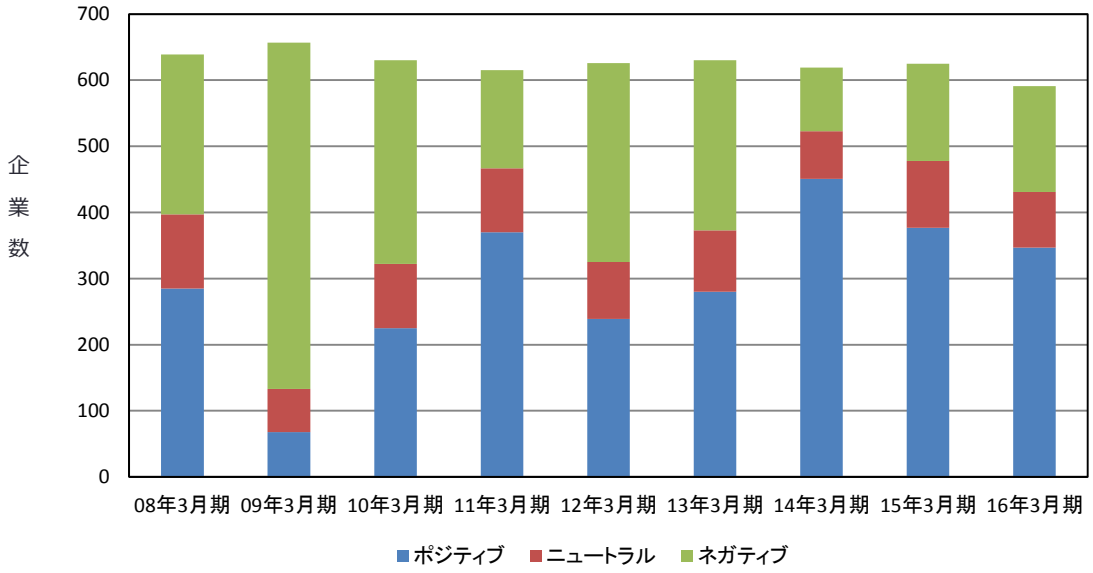
3.結果

3.1 各極性における企業数の推移

分析期間における企業数は各期において 600 社程度であった（図表 2）。各極性における企業数の推移を見ると、09 年 3 月期にネガティブな極性の企業が前期の 242 社から 524 社へ約 2.2 倍に増加した後、10 年、11 年 3 月期はポジティブな極性の企業が年々増加し、ネガティブな極性の企業が減少した。また、12 年 3 月期にネガティブな極性の企業が前期の 148 社から 301 社へ約 2 倍に増加した後、13 年、14 年 3 月期はポジティブな極性の企業が年々増加し、ネガティブな極性の企業が減少した。そして、直近はポジティブな極性の企業が徐々に減少し、ネガティブな極性の企業が増加する傾向にあった。また、09 年、10 年、12 年 3 月期以外、ポジティブな極性の企業数はネガティブな極性の企業数を上回っていた。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表 2　極性別企業数の推移



(出所)各社有価証券報告書より日興リサーチセンター作成

3.2 極性と当該期の企業業績との関係

図表 3 は、全分析期間での各極性における当該期の各財務指標の中央値を示している。中央値はすべての財務指標においてポジティブの伸び率が最も高く、次にニュートラルが高く、ネガティブが最も低かった。各財務指標におけるポジティブとネガティブの中央値の差を比較すると、売上高伸び率の差は 7.5%であり、一方その他 3 つの利益系指標（営業利益、経常利益、当期利益伸び率）は差が 30%以上であった。また、極性と各財務指標の順位相関係数は 0.34～0.39 であり、正の相関であった。

図表 3　各極性における当該期の各財務指標の中央値

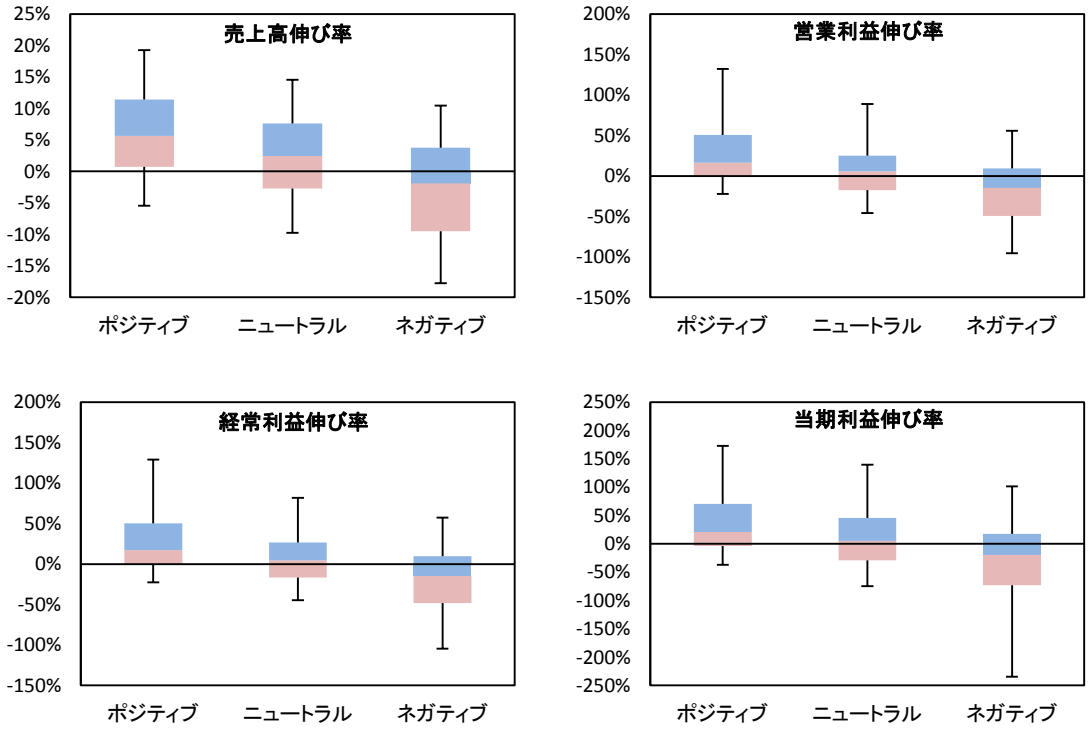
	売上高 伸び率	営業利益 伸び率	経常利益 伸び率	当期利益 伸び率
ポジティブ	5.65%	16.27%	17.10%	20.77%
ニュートラル	2.44%	5.55%	4.61%	5.26%
ネガティブ	−1.89%	−14.74%	−14.79%	−19.91%

(出所)各社有価証券報告書及び QUICK より日興リサーチセンター作成

図表 4 は、当該期の各財務指標における極性別箱ひげ図を示している。すべての財務指標で箱ひげ図は、ポジティブが最も上方に位置しており、その次にニュートラル、ネガティブが最も下方に位置していた。また、各極性における分布の偏りは、どの指標においてもポジティブはプラス方向に、ネガティブはマイナス方向に裾が長い傾向が見られ、利益系指標においてニュートラルはプラス方向に裾が長い傾向が見られた。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表４　当該期における各財務指標の極性別箱ひげ図

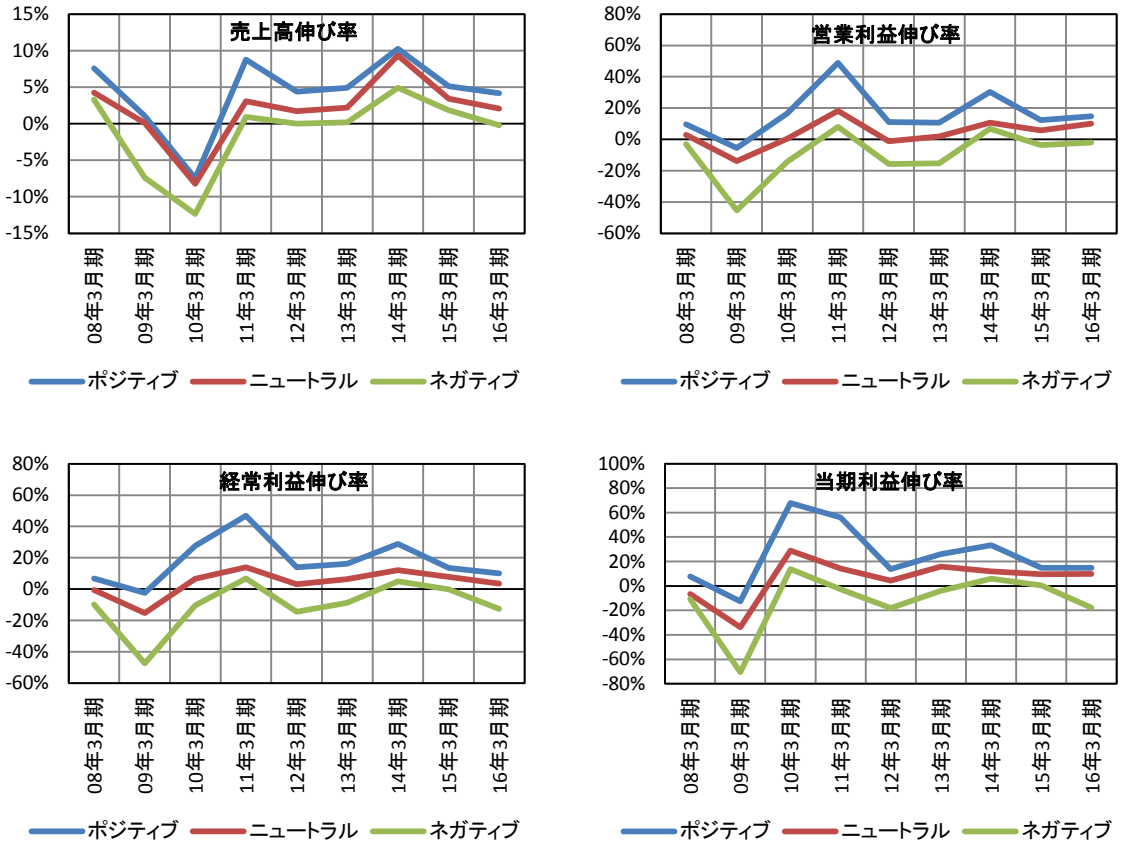


注：箱ひげ図における横線は、下から 10、25 パーセンタイル値、中央値、75、90 パーセンタイル値を示す。
(出所) 各社有価証券報告書及び QUICK より日興リサーチセンター作成

図表 5 は、極性別の当該期における各財務指標の中央値の推移を示している。すべての財務指標の中央値は、ポジティブが最も高く、次にニュートラルが高く、ネガティブが最も低く、その傾向は安定していた。また、グラフの形状を確認すると、分析期間の前半である 08 年～11 年 3 月期において各財務指標は分析期間の後半と比較して大きく動いており、低下した後上昇していた。極性別に見るとネガティブでは低下時の傾きがその他の極性よりも急な傾向にあり、この傾向は利益系の指標においてより顕著であった。一方、ポジティブでは上昇時の傾きがその他の極性よりも急な傾向にあった。12 年 3 月期以降は、分析期間の前半と比較して各財務指標の動きは小さかった。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表 5　当該期における各財務指標の中央値の極性別推移



(出所) 各社有価証券報告書及び QUICK より日興リサーチセンター作成

3.3 極性と翌期の企業業績との関係

図表 6 は、全分析期間での各極性における翌期の各財務指標の中央値を示している。各財務指標におけるポジティブとネガティブの中央値の差をとると、売上高伸び率は 2.7%、営業利益伸び率は-0.8%、経常利益伸び率は-1.5%、当期利益伸び率は-6.7%であった。売上高伸び率では、ポジティブの伸び率が最も高く、ネガティブが最も低かった。一方で、その他 3 つの利益系指標では、図表 3 とは反対にネガティブがポジティブよりも高かった。また、極性と各財務指標の順位相関係数は-0.08～0.14 であり、相関関係は認められなかった。

図表 6　各極性における翌期の各財務指標の中央値

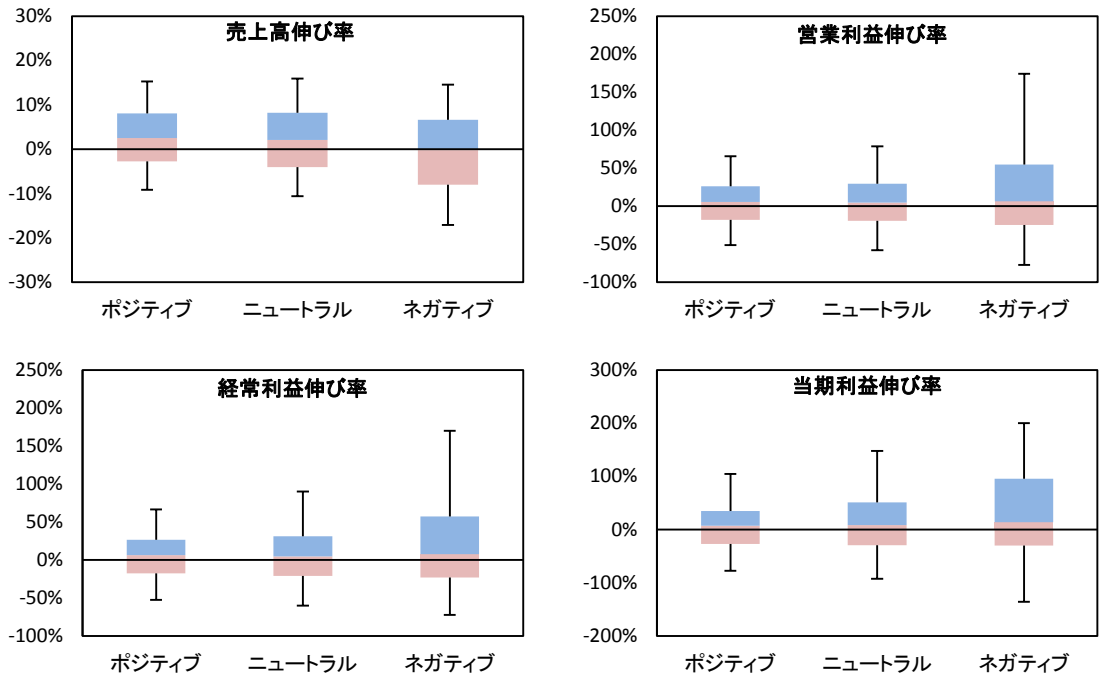
	売上高伸び率	営業利益伸び率	経常利益伸び率	当期利益伸び率
ポジティブ	2.52%	5.63%	6.24%	7.28%
ニュートラル	2.11%	4.95%	4.74%	8.35%
ネガティブ	−0.20%	6.41%	7.78%	14.00%

(出所)各社有価証券報告書及び QUICK より日興リサーチセンター作成

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表 7 は、翌期の各財務指標における極性別箱ひげ図を示している。売上高伸び率の箱ひげ図は、ポジティブがネガティブより上方に位置していたが、その他 3 つの利益系指標ではネガティブがポジティブよりも上方に位置する傾向にあった。また、各極性における分布の偏りを見ると、ネガティブが最も分布の両裾が長く、次にニュートラルが長く、ポジティブが最も短い傾向にあり、特に利益系の指標ではネガティブの裾がプラス方向へより長かった。

図表 7　翌期における各財務指標の極性別箱ひげ図

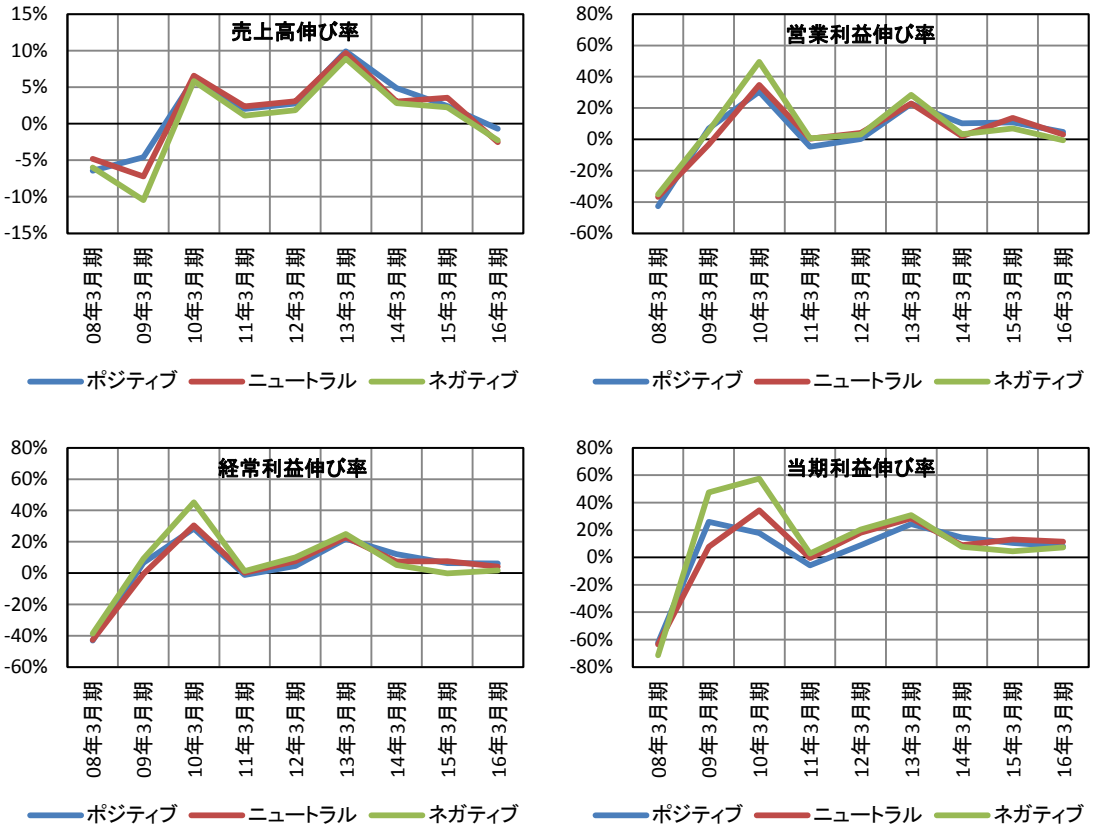


注：箱ひげ図における横線は、下から 10、25 パーセンタイル値、中央値、75、90 パーセンタイル値を示す。
(出所) 各社有価証券報告書及び QUICK より日興リサーチセンター作成

図表 8 は、極性別の翌期における各財務指標の中央値の推移を示している。すべての財務指標において極性間に差は見られず、どの極性も同様の推移をしており、図表 5 のような関係は見られなかった。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表 8　翌期における各財務指標の中央値の極性別推移



(出所) 各社有価証券報告書及び QUICK より日興リサーチセンター作成

4.考察

本稿では、有価証券報告書の「業績等の概要」に含まれる因果関係文が、企業業績を実際に反映しているかの検証を行った。

極性別の企業数の推移を見たところ（図表 2)、リーマン・ショック（2008 年 9 月 15 日）が発生した日を含む 09 年 3 月期は、ネガティブな極性の企業数が急増し、ポジティブな極性の企業数は急減したが、その翌期以降（10 年、11 年 3 月期）はリーマン・ショックからの回復局面にあったため、ポジティブな極性の企業数が年々増加したものと考えられる。また、12 年 3 月期もネガティブな極性の企業数が急増していた。これは、前期末に発生した東日本大震災（2011 年 3 月 11 日）の影響が発生後の日数が僅かである 11 年 3 月期ではなく、12 年 3 月期に現れたためと考えられる。その翌期以降（13 年、14 年 3 月期）はポジティブな極性の企業数が年々増加した。これらの結果は、有価証券報告書の「業績等の概要」がマクロイベント等の影響を受けた企業業績を反映していることを示唆していると考えられる。なお、内閣府より公表されている景気基準日付において、本稿の分析期間は日本経済の景気拡張期に多く属していた（図表 9)。このことは、分析期間を通じてポジティブな極性の企業数が多かった一因であると考えられる。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表 9　2002 年以降の日本の景気循環

循環	谷	山	谷	期間		
				拡張	後退	全循環
第 14 循環	平成 14 年 1 月 (2002 年 1 月)	平成 20 年 2 月 (2008 年 2 月)	平成 21 年 3 月 (2009 年 3 月)	73 ヵ月	13 ヵ月	86 ヵ月
第 15 循環	平成 21 年 3 月 (2009 年 3 月)	平成 24 年 3 月 (2012 年 3 月)	平成 24 年 11 月 (2012 年 11 月)	36 ヵ月	8 ヵ月	44 ヵ月
第 16 循環	平成 24 年 11 月 (2012 年 11 月)					

(出所)内閣府ホームページより日興リサーチセンター作成

次に、極性と企業業績の関係について検証を行った。その結果、全分析期間で見ると、極性と当該期における各財務指標の関係は正の相関であり（図表 3)、極性別の各財務指標の分布を見てもその関係は確認された（図表 4)。また、期毎で見ても各極性と各財務指標の中央値の関係は安定していた（図表 5)。一方、極性と翌期の企業業績の間には相関関係は見られなかった（図表 6、7、8)。これらから、有価証券報告書の「業績等の概要」に含まれる因果関係文は、翌期の企業業績ではなく、当該期の企業業績を反映していることを強く示唆していると考えられる。この結果から、因果関係文を用いた当該期の業績に関する要因分析など、投資判断に応用する上での因果関係文の有益性がさらに高まったと言えるだろう。

ところで、極性別に翌期の各財務指標の分布を見ると、3 つの利益系指標において、当該期のそれらの分布とは反対に、ネガティブはポジティブよりも上方に分布する傾向にあり、分布の裾を見るとネガティブはニュートラルやポジティブと比較してプラス方向に裾がより長い傾向にあった（図表 7)。これは、ネガティブな極性の企業の中には、悪材料出尽くし等から翌期に業績が回復する企業が多く含まれている可能性がある。そこで、ネガティブな極性の企業が、翌期にどの極性に推移していたかを調べた。図表 10 は、当期の極性が翌期にどの極性へ推移したかを示している。極性がネガティブの企業は 40%がポジティブへと推移し、46%がネガティブに留まっていた。また、全体としては、ポジティブへの推移や滞留が多かった。これは上述したように、本稿における分析期間は景気拡張期であったことを反映しているのかもしれない。その点を踏まえると、ポジティブな企業よりもネガティブな企業の方がプラス方向の裾が長かった点については、一旦業績が落ち込んだ企業の方が、業績が継続して伸びている企業よりも、利益系指標の伸び率について数値が高く出易いことを示していると考えられる。このことを利用して、ネガティブな極性の企業群から、翌期に業績が急回復する企業を抽出することが出来れば、投資判断の材料として有益な情報になり得るであろう。本稿で得られたネガティブな極性の企業群から、翌期に業績が急回復する企業の抽出を、今後の課題として試みたい。

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

図表 10　当期と翌期の極性の推移の割合

		当期の割合の平均値	翌期		
			ポジティブ	ニュートラル	ネガティブ
当期	ポジティブ	47%	51%	13%	35%
	ニュートラル	14%	48%	15%	37%
	ネガティブ	38%	40%	14%	46%

注：当期の割合の平均値は、分析期間における各期の各極性の割合について平均した値。
　　当期は翌期の前の期を示す。
(出所)各社有価証券報告書より日興リサーチセンター作成

5.まとめ

本稿では、有価証券報告書の「業績等の概要」に含まれる因果関係文が、企業業績を実際に反映しているかを検証した。その結果、主に当該期の企業業績を反映していることが示唆された。これにより、因果関係文を投資判断に用いる有益性はさらに高まったと言えるだろう。今後の課題として、因果関係文の投資判断への応用を試みたい。

＊ 本稿の作成にあたり、東京大学大学院工学系研究科教授和泉潔先生と助教坂地泰紀先生に、多大なるご指導とご助言を頂いた。心から感謝の意を表したい。なお、本稿の内容・意見は全て当社に属する。

以上

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。

参考文献

五島圭一・高橋大志[2016], 「ニュースと株価に関する実証分析―ディープラーニングによるニュース記事の評判分析―」, 証券アナリストジャーナル, Vol.54, No.3, pp.76-86, Mar. 2016.

北森詩織・酒井浩之・坂地泰紀[2017], 「決算短信 PDF からの業績予測文の抽出」, 電子情報通信学会論文誌 (D), Vol.J100-D, No.2, pp150-161, Feb. 2017.

小寺俊哉・佐藤史仁・田中良典[2018], 「テキストマイニングによる有価証券報告書の因果関係文以外の特徴文の抽出」, 日興リサーチレビュー, 2018 年, 1 月号.

沖本竜義・平澤英司 [2014], 「ニュース指標による株式市場の予測可能性」, 証券アナリストジャーナル, Vol.52, No.4, pp.67-75, Apr. 2014.

酒井浩之・小林義和・坂地泰紀[2015a],「企業の決算短信ＰＤＦから抽出した業績要因への極性付与」, 第 15 回 金融情報学研究会, pp.7-12, 2015.

酒井浩之・西沢裕子・松並祥吾・坂地泰紀[2015b], 「企業の決算短信 PDF からの業績要因の抽出」, 人工知能学会論文誌, Vol.30, No.1, pp.172-182, 2015.

坂地泰紀・増山繁 [2011],「新聞記事からの因果関係を含む文の抽出手法」, 電子情報通信学会論文誌 (D), Vol.J94-D, No.8, pp1496-1506, Aug. 2011.

坂地泰紀・酒井浩之・増山繁 [2015],「決算短信 PDF からの原因・結果表現の抽出」, 電子情報通信学会論文誌 (D), Vol.J98-D, No.5, pp811-822, May 2015.

佐藤史仁・佐久間洋明・小寺俊哉[2017], 「テキストマイニングによる有価証券報告書の因果関係文の抽出」, 日興リサーチレビュー, 2017 年, 10 月号.

本資料は、信頼性の高いデータから作成されておりますが、当社はその正確性・確実性に関し、いかなる保証をするものではございません。本資料は、情報提供を目的としており、投資勧誘を目的としたものではございません。証券投資に関する最終判断は、投資家ご自身の判断でなさるようにお願いいたします。本資料の著作権は当社に帰属し、本資料の転用および販売は固く禁じられております。