

ニュース指標による株式市場の予測可能性

沖 本 竜 義
平 澤 英 司 CMA

目 次

- | | |
|----------------|-------------|
| 1. はじめに | 4. 評価項目別の分析 |
| 2. タグ情報とニュース指標 | 5. 結論 |
| 3. ニュース理論とその検証 | |

本稿では、QUICK端末で配信されている日経ニュースを定量化してニュース指標を作成し、そのニュース指標を用いた株式市場の予測可能性について検証した。分析の結果、ニュース指標は、翌営業日の株式リターンや出来高に対して有意な説明力を持つこと、先行研究で見られたようなリバウンドが観測されることはないこと、などが確認された。これは、本稿で用いたニュース指標が株価に対して本源的な情報を有していることを示しており、非常に興味深い結果である。

1. はじめに

通常、株式投資は株価・財務指標等の数値情報だけでなく、ニュース等のテキスト情報（注1）も用いて総合的に意思決定することが多い。従って、本来、その意思決定の検証は、数値情報とテキスト情報の両方を用いて行うべきである。しか

しながら、後者を用いるにはテキスト情報を定量化可能な情報に変換する必要がある、これには技術的ならびに費用的にハードルが高いため、今まではそのような検証が十分に行われてこなかった。

今回は、(株)金融工学研究所がニュースから推計したタグを用いて、ニュースというテキスト情報



沖本 竜義（おきもと たつよし）

オーストラリア国立大学クロフォード公共政策大学院准教授。2005年カリフォルニア大学サンディエゴ校大学院経済学研究科博士課程修了（経済学Ph. D）。同年、横浜国立大学大学院国際社会科学部研究科准教授、08年9月、一橋大学大学院国際企業戦略研究科准教授を経て、14年3月より現職。主な著書に『経済・ファイナンスデータの計量時系列分析』（朝倉書店、2010年）がある。



平澤 英司（ひらさわ えいじ）

(株)金融工学研究所 取締役 業務統括。1996年早稲田大学理工学部応用物理学科卒業、05年一橋大学大学院国際企業戦略研究科修士課程修了。96年野村證券入社。ブルームバーグL. P. を経て、06年4月より現職。現職では、モデル構築・メンテナンス、コンサルティング、新商品開発等業務全般の統括を担当。

を定量処理することが可能となったので、テキスト情報と株式市場との関連性、特に予測可能性について分析を行いたいと思う。テキストにはニュース以外にもインターネット上のSNS情報等もあるが、本稿では株式の投資家が優先的に利用していると思われるニュースに限定して分析を行うこととする。

現在、新聞や情報端末を通じて、日々、膨大な株式市場に関するニュースが配信されている。これらのニュースが、株価に対して本源的な情報を持っていることが期待されるが、Roll [1988] がニュースと株式リターンの間にほとんど関係がないことを示した研究に代表されるように、両者に強い関係を見出している研究は少ない。例えば、Tetlock [2007] は、ウォールストリートジャーナル（以下、WSJ）のコラムからニュース指標を作成し、ニュース指標が株式市場のセンチメントを表す指標となるが、株価に対して本源的な情報を持っていないことを確認している。

ニュースが日本の株式市場に与える影響に関しては、Griffin *et al.* [2011] が、ニュースが株式市場のボラティリティを上昇させる可能性を示唆している程度で、多くの研究は行われていない。そこで、本稿では、QUICK端末で配信されている日経ニュースを定量化してニュース指標を作成し、その指標が株価に対して本源的な情報を有しているかどうかを検証する。具体的には、Tetlock [2007] と同様に、ニュース指標、株式リターン、出来高からなる3変量VARモデルを用いて、ニュース指標に対して情報理論、センチメント理論、無情報理論という3つの理論の検証を行う。さらに、ニュースを評価項目別に細分化し、どの類のニュースが、株式リターンに関してより大きな影

響を持つかを明らかにすることを試みる。

本稿の構成は以下の通りである。第2章では、本稿で用いたニュースならびにタグ情報の概要と、それを基に作成したニュース指標について述べる。第3章では、Tetlock [2007] による3つのニュース理論を紹介し、ニュース指標、株式リターン、出来高からなる3変量VARモデルに基づいて、ニュース指標に対して3つの理論を検証する。第4章では、ニュースを評価項目別に細分化し、どの類のニュースが、株式リターンに関してより大きな影響を持つかを検証する。最後に、第5章で結論を述べる。

2. タグ情報とニュース指標

本稿で利用したニュースデータは、2007年7月1日から11年12月31日までにQUICK端末で配信された日経ニュースの809,667件であり、ニュースで扱われた企業数は3,535社にのぼる。ニュースデータは、(株)日本経済新聞社と(株)QUICKから拠出頂いた。当ニュースは、証券市場に参加している機関投資家、個人投資家にリアルタイムで読まれており、市場への情報伝達力が高いニュースといえることができる。

このニュースデータに対して、定量処理を可能にするために、(株)金融工学研究所が付与したタグ情報を利用した。主に利用したタグ情報は、対象ニュースの主要企業名（コードを含む）・事由付きポジティブ・ネガティブ（ポジ・ネガ）情報である（注2）。事由付きポジ・ネガ情報は、図表1のとおりに定義されており、ニュースの見出しと本文から同社定義の基準で推計されている。図表2は、ポジ・ネガ情報の分布である。市況情

（注1） 本稿では表現方法がテキストで、そのままだとパソコンなどの機械で処理を行いにくいタイプの情報のことをテキスト情報と呼ぶこととする。

図表1 事由付きポジ・ネガ情報

評価項目	事由	ポジティブ / ネガティブ
証券会社アナリスト推奨	評価上げ 評価下げ	ポジティブ ネガティブ
コーポレートアクション	自社株買い、増配等 無配・減配等	ポジティブ ネガティブ
利益	上方修正・上昇等 下方修正・低下等	ポジティブ ネガティブ
売上	売上増・堅調等 売上減・低迷等	ポジティブ ネガティブ
マーケットコメント	続伸・反発等 続落・反落等	ポジティブ ネガティブ
信用格付	格上げ 格下げ	ポジティブ ネガティブ
製品・サービス	新サービス リコール	ポジティブ ネガティブ
コーポレートガバナンス	不祥事・偽装等	ネガティブ
マクロ指標	日銀短観プラス等 機械受注減等	ポジティブ ネガティブ

(図表注) ポジティブ・ネガティブの評価は各ニュースで言及されている主要企業に紐付けられている。一つのニュースに複数の事由が該当する場合は、主要企業にポジ・ネガ評価が複数発生したと処理している。

(出所) 筆者作成、以下同じ

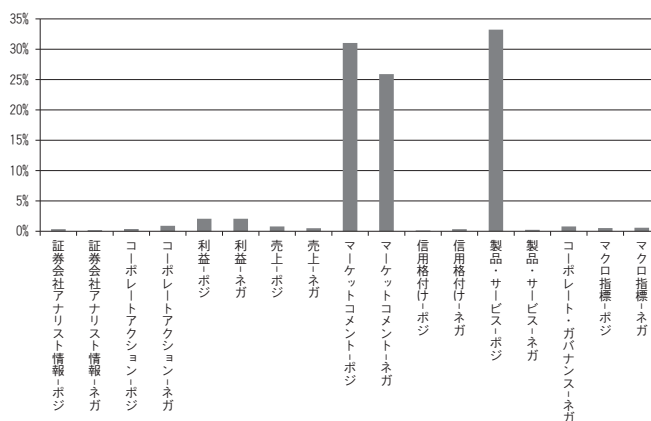
報を伝えるマーケットコメントや新商品・新サービスの情報を伝える製品・サービス（ポジティブ）のニュースが大部分を占める分布となっている。

ここで、データの制約について説明する。今回のニュースデータからは、時・分・秒単位のタイムスタンプ情報を得ることが出来なかった。引け後（15時以降）を翌日のニュースとして処理する考え方もあるが、本稿では24時までのニュースを当日のニュースとして扱って日次分析を行う（注3）。ニュース指標を構築する際は、銘柄ごとにポジ・ネガ情報が付与されたニュース数を日次でカウントし、Tetlock, *et al.* [2008] を参考に、次の2つの指標を算出した（注4）。

$$NI1 = \frac{P-N}{P+N+1}, NI2 = \frac{-N}{P+N+1} \quad (1)$$

ここで、 P はポジティブ情報が付与されたポジテ

図表2 ポジ・ネガ情報の分布



(図表注) 図表2の事由付きポジ・ネガ情報の分布を表している。銘柄の値動き・理由等を伝えるマーケットコメント、新商品・新サービスの案内をする製品・サービスのニュースが多くを占めている。件数は少ないものの株価に対するインパクトが強い項目もある。

(注2) 同社によると、タグの推計精度は主要企業名、ポジ・ネガ情報ともに99%超であるが、モデルによる推計を行っているので予期しない書式・内容のニュースが相対的に増えた場合などに精度が落ちる可能性がある。

(注3) 本稿で利用したデータとは別に2013年からサンプル週（4週）を定めて引け後から24時までのニュースの割合を求めたところ、全ニュースに対する割合が約3割となった。

イブニュースの数であり、 N はネガティブ情報が付与されたネガティブニュースの数である。つまり、 $NI1$ はポジティブニュースとネガティブニュースの数の差を、ニュース数の合計で割り、基準化したものであり、 $NI2$ はポジ・ネガ情報が付与されたニュースに占めるネガティブニュースの割合を計算したものである（注5）。

ニュースに含まれる数値情報を用いて、 N と P に重みをかける考え方もあるが、業種などの銘柄属性によって最適な重みが異なる可能性もあり、別途議論を重ねる必要があると考える。したがって、本稿ではニュース件数が数値情報と関係なく株式市場の予測に有意な情報かどうかを検証することを目的とし、単にニュースの数を扱っていることに注意されたい。

分析に用いたその他の株価指標、市場売買高等の株式関連データは、(株)日本経済新聞デジタルメディアのNEEDS-Financial QUEST 2.0から取得した。

3. ニュース理論とその検証

Tetlock [2007] はニュースが持つ情報に対して、次のような3つの理論を考えている。まず、1つ目は情報理論であり、これはニュースが株価に対して本源的な情報を保有するという理論である。この場合、ニュースは将来の株価に対して、有意な影響をもち、その影響は恒久的に消滅しないこととなる。2つ目は、センチメント理論であり、これはニュースが株価に対して本源的な情報を保

有しないが、マーケットのセンチメントに影響を与えるというものである。この場合、ニュースは将来の株価に関して、短期的には有意な影響を持つが、長期的にはその影響は消滅することとなる。3つ目の理論が無情報理論であり、これはニュースが株価やセンチメントに対して何の影響も持たないという理論である。この場合、ニュースは将来の株価に対して、短期的にも長期的にも何の影響も与えないこととなる。

Tetlock [2007] はWSJのコラムに関して、これら3つの理論を検証している。具体的には、ニュース指標、株式リターン、出来高からなる3変量VARモデルを考え、ニュース指標が将来の株式リターンに対して有意な影響を持つかどうか、その影響が時間とともに消滅するかどうかなどを検証した。その結果、WSJのコラムから作成した指標が、短期的には株価に有意に正の影響を与えるが、長期的にはリバウンドが生じ、有意な影響を持たないことを指摘している。つまり、この結果はセンチメント理論と整合的なものであり、WSJのコラムには株価に対する本源的な情報は含まれていないが、市場のセンチメントに関する情報を含んでいる可能性を示唆している。

本章では、前章で述べたニュース指標を用いて、Tetlock [2007] と同様の分析を行い、3つのニュース理論を検証する。具体的には、ニュース指標 (NI_t)、株式リターン (Tpx_t)、出来高 (Vol_t) からなる3変量VARモデルを考え、ニュース指標が株価に対する本源的な情報を含んでいる可能性があるかどうかを検証する。株式リターンとして

（注4） このため、企業名が付与されていないニュースは分析の対象外とした。

（注5） Tetlock, et al. [2008] では、ポジティブニュースの説明力が弱いということで、 $NI2$ と同様の指標に基づいた結果だけが報告されている。Tetlock, et al. [2008]の指標と本稿で用いているニュース指標の違いは、分母が0になることを防ぐために、分母に1を加えていることと、 $NI1$ と $NI2$ を使用した際の結果の符号が同一となるように、 $NI2$ にマイナスを掛けていることである。

は、TOPIXの日次対数収益率（％）を利用し、出来高としては、東証1部全銘柄の出来高の合計の対数を計算したものを利用した。また、ニュース指標は、平均が0、標準偏差が1となるように標準化した。Tetlock [2007] と同様に、VARのラグ次数は5次、つまり各変数の限界的な影響は1週間以内に消滅すると仮定した。この仮定の下では、例えば、株式リターンに関する回帰モデルは次のように表される（注6）。

$$Tpx_t = \alpha_1 + \sum_{j=1}^5 \beta_{1j} Tpx_{t-j} + \sum_{j=1}^5 \gamma_{1j} NI_{t-j} + \sum_{j=1}^5 \delta_{1j} Vol_{t-j} + \varepsilon_{1t} \quad (2)$$

3つのニュース理論を検証する上で重要なパラメータが(2)における γ_1 である。もし、情報理論が正しいのであれば γ_{11} は正で有意になり、残りの γ_1 の成分は負で有意になることはないはずである。一方、センチメント理論が正しいのであれば γ_{11} は正で有意になり、 γ_{1j} ($2 \leq j \leq 5$)はいずれかの j において有意に負の値をとるはずである。最後に、無情報理論が正しいのであれば、 γ_1 のいずれの成分も有意にならないはずである。

(2)を推定し、 γ_1 の推定値をNewey-West標準誤差とともにまとめたのが、図表3である（注7）。図表から分かるように、いずれのニュース指標においても、 γ_{11} は有意水準5%で有意に正となっ

図表3 γ_1 の推定値

	NI1		NI2	
γ_{11}	0.167**	(0.078)	0.194**	(0.084)
γ_{12}	-0.021	(0.068)	-0.016	(0.071)
γ_{13}	0.087	(0.073)	0.098	(0.076)
γ_{14}	-0.032	(0.070)	-0.043	(0.071)
γ_{15}	0.043	0.061	0.044	(0.065)
Adj R ²	0.011		0.014	

（図表注） 括弧内はNewey-West標準誤差を表す。*、**、***は、それぞれ10%、5%、1%の有意水準で有意であることを表す。以下同じ。

ている。つまり、ニュース指標は翌営業日の株式リターンに対して、有意な正の影響を持ち、無情報理論が成立していないことを示唆している。また、それ以外の γ_1 は、いずれのニュース指標においても、すべて有意とはなっていない。この結果は、2営業日目以降のいずれかの γ_1 が負で有意になり、リバウンドが観測されたTetlock [2007]の結果とは対照的であり、ニュース指標は株価に対して恒久的な影響を持っていることが示唆されている。つまり、センチメント理論ではなく、情報理論が成立している可能性が高いことを示しているのである。

より具体的にNI1(NI2)の結果を解釈すると、ある日のニュース指標が平均値から1標準偏差大きいとすると、ニュース指標が平均値であったときと比較して、翌営業日の株式リターンは0.167% (0.194%) 有意に大きくなる傾向があることを示している。これは、経済的にも非常に大きな値であり、ニュース指標が株価に対して本源

（注6） 前章で述べたデータ制約のため、引け後（15時以降）を翌日として同時点での依存構造も考慮した構造VARモデルではないことに注意されたい。また、本稿では1日のニュースの件数を情報として用いており、ニュースの件数は1日を経過しないと、観測されないものであるため、(2)のモデルはその意味では自然なモデルとなっている。タイムスタンプを利用し、観測期間をより短期間とした、より精緻な分析については今後の重要な課題の1つである。

（注7） ニュースは株式市場に基づくニュースも多く、 Tpx とNI1の相関による多重共線性の可能性には注意が必要である。実際、両者の相関は0.6程度と高いものになっているが、(2)を推定した結果の中心化分散拡大係数はいずれの変数に関しても4以下となっており、多重共線性が疑われる10よりも、かなり小さくなっているため、多重共線性が問題となる程度にはなっていない。

的な情報を保有している可能性を強く示唆しているということが出来る。また、紙幅の都合上、結果の詳細は載せないが、 β_1 の推定値はいずれの次数においても有意となっておらず、ニュース指数の有意性が際立つものとなっている。

次に、本稿のニュース指標においては、センチメント理論が不適当であることを市場がニュースに与える影響という観点から確認する。そのために、3変量VARモデルにおけるニュース指標の回帰モデル、つまり

$$NI_t = \alpha_2 + \sum_{j=1}^5 \beta_{2j} Tpx_{t-j} + \sum_{j=1}^5 \gamma_{2j} NI_{t-j} + \sum_{j=1}^5 \delta_{2j} Vol_{t-j} + \varepsilon_{2t} \quad (3)$$

という回帰モデルに注目する。もし、センチメント理論が正しいのであれば、過去の株式市場のパフォーマンスはニュース指標に対して正の影響を持つはずであるので、 β_2 は正で有意になる傾向があるはずである。

(3)を推定し、 β_2 の推定値をNewey-West標準誤差とともにまとめたのが、**図表4**である。**図表**からわかるように、いずれのニュース指標においても、 β_{21} から β_{24} は有意水準5%で有意となっていない。また、 β_{25} は有意となっているが、その符号は負である。つまり、過去の株式市場のパフォーマンスがニュース指標に対して、正の影響を持っている統計的証拠は見受けられない。この結

果も、 β_{21} が有意に正となっていたTetlock [2007]の結果とは対照的であり、センチメント理論とは非整合的なものとなっている。

最後に、ニュース指標が出来高に与える影響から3つの理論を検証する。DeLong *et al.* [1990]やCampbell *et al.* [1993]の理論に基づくと、センチメントにショックがあった場合、ノイズトレーダーが合理的トレーダーに対して取引を行うことになる。その結果、センチメントを平均0に標準化した場合、センチメントの絶対値が大きくなると、出来高は上昇することになる。この考察に基づき、Tetlock [2007]はVARモデルにおける出来高の回帰式に、ニュース指標の絶対値を含めた次のような回帰モデルを考えている。

$$Vol_t = \alpha_3 + \sum_{j=1}^5 \beta_{3j} Tpx_{t-j} + \sum_{j=1}^5 \gamma_{3j} NI_{t-j} + \sum_{j=1}^5 \psi_{3j} |NI_{t-j}| + \sum_{j=1}^5 \delta_{3j} Vol_{t-j} + \varepsilon_{3t} \quad (4)$$

センチメント理論が正しいければ、 ψ_{31} は正で有意になるはずである。(4)を推定し、 γ_3 と ψ_3 の推定値をNewey-West標準誤差とともにまとめたのが、**図表5**である。**図表**から分かるように、いずれのニュース指標に対しても、 ψ_{31} は有意とはなっていない。それに対して、 γ_{31} はいずれのニュース指標に対しても、正で有意となっている。つまり、ポジティブニュースは出来高を増大させ、ネガティブニュースは出来高を減少させる傾向があることが分かる。より具体的には、 $NI1$ ($NI2$)が1標準偏差上昇すると、翌営業日の出来高は平均的に1.5%(1.3%)程度上昇することが示唆されている。この結果は、ニュース指標が翌営業日の出来高に関しても、影響力を有していることを示しており、情報理論と整合的なものとなっている。

以上の結果をまとめると、本稿で利用したニュ

図表4 β_2 の推定値

	NI1		NI2	
β_{21}	-0.021	(0.020)	-0.029	(0.021)
β_{22}	-0.028	(0.022)	-0.035	(0.022)
β_{23}	-0.026	(0.018)	-0.031	(0.018)
β_{24}	0.024	(0.018)	0.025	(0.018)
β_{25}	-0.038**	(0.018)	-0.042**	(0.019)
Adj R ²	0.095		0.105	

図表5 γ_3 と ψ_3 の推定値

	NI1		NI2	
γ_{31}	0.015***	(0.006)	0.013**	(0.006)
γ_{32}	-0.005	(0.006)	-0.003	(0.006)
γ_{33}	0.001	(0.006)	0.004	(0.006)
γ_{34}	-0.002	(0.005)	-0.002	(0.005)
γ_{35}	-0.006	(0.006)	-0.007	(0.006)
ψ_{31}	-0.007	(0.008)	-0.007	(0.008)
ψ_{32}	-0.010	(0.007)	-0.008	(0.007)
ψ_{33}	-0.002	(0.007)	-0.002	(0.008)
ψ_{34}	0.013*	(0.007)	0.013*	(0.007)
ψ_{35}	-0.012	(0.007)	-0.012*	(0.007)
Adj R ²	0.547		0.547	

ース指標に対しては、センチメント理論や無情報理論よりも情報理論がより整合的であるということが出来る。これは、Tetlock [2007] の結果とは対照的であり、本稿のニュース指標は株価に対して本源的な情報を含んでいる可能性が示唆されたのである。このような結果が得られたひとつの理由は、Tetlock [2007] のニュース指標はWSJのコラムを基に作成されたのに対して、本稿のニュース指標は、QUICK端末で配信されている日経ニュースを基に作成されていることが挙げられる。WSJのコラムは、市場動向全体に関して記述されており、集約された側面が強い情報であるとともに、読み物的な意味合いが強いこともあり、一時的に市場のセンチメントに影響を与えることがあっても、株価に影響を与える本源的な情報が含まれていることは少ないであろう。それに対して、本稿で分析した日経ニュースは、主に個別銘柄のポジティブニュースやネガティブニュースという、より詳細な情報であり、株価に影響を与える本源的な情報が含まれている可能性が高いこと

が確認されたのである。

4. 評価項目別の分析

第3章で、本稿で用いたニュース指標が、株価に対して本源的な情報を有している可能性が示唆された。そこで、本章では、ニュースを評価項目別に細分化し、どの類のニュースが、株式リターンに関してより大きな影響を持っているかを明らかにすることを試みる。そのために、コーポレートガバナンスを除く各評価項目別に、(1)に基づいて、ニュース指標1 (SNI1) を作成し、次のような回帰モデルを推定した(注8)。

$$Tpx_t = c + \phi Tpx_{t-1} + \sum_j \gamma_j SNI1_{j,t-1} + \varepsilon_t \quad (5)$$

(5)において、全評価項目を用いたものをモデル1として推定し、推定値をまとめたものが図表6の2列目である。結果から分かるように、コーポレートアクションとマーケットコメント(注9)が少なくとも有意水準10%で、正で有意となっている。つまり、コーポレートアクションとマーケットコメントに関して、ポジティブニュースが多かったとすると、翌営業日の株価が有意に上昇する傾向にある。言い換えれば、コーポレートアクションとマーケットコメントに関するニュースに投資家は有用な情報を見出し、それが株価の形成に影響を与えている可能性が高いのである。

マーケットコメントは続伸・反発等の株価の方向性を示す内容であるが、単純なモメンタム効果

(注8) コーポレートガバナンスに関しては、ネガティブニュースしかないため、本章の分析からは除外した。また、紙幅の関係で、ニュース指標1に関する結果だけを報告しているが、ニュース指標2を用いても、ほぼ同様の結果が得られた。

(注9) マーケットコメントは、「◇<東証>大成建が7年ぶり高値 五輪関連銘柄が連日で上昇」のように、記者が選択した銘柄について取引状況を理由付きで配信している場況ニュースのこと。

図表 6 評価項目別の分析

	モデル1	モデル2	モデル3	モデル4	モデル5
C	-0.083	-0.085*	-0.080	-0.081*	-0.086*
φ	-0.063	-0.061	0.006	0.001	-0.064
BR	-0.095**	-0.096**	-0.084*		
CA	0.093*	0.101*		0.109**	
ER	0.011				
SL	0.040				
MC	0.170*	0.170*			0.170*
CR	0.003				
PS	-0.054				
MI	-0.040				
Adj R ²	0.013	0.012	0.002	0.004	0.005

(略語一覧)

BR：証券会社アナリスト推奨

CA：コーポレートアクション

ER：利益

SL：売上

MC：マーケットコメント

CR：信用格付

PS：製品・サービス

MI：マクロ指標

を捉えているわけではないことに注意されたい。実際、(5)式はTOPIXのリターン過去の過去も含んでおり、モメンタム効果はコントロールされており、**図表 6** から分かるように、モメンタム効果は有意となっていない。日本の株式市場において、モメンタム効果が存在しないことは、多くの研究により指摘されており、Fama and French [2012] の最近の研究でも確認されている。したがって、マーケットコメントは、日経（の記者）が全上場銘柄の株価情報から銘柄の選択を行い、その値動きに関して理由を添えることにより、コーポレートアクションと同様に投資家にとっての有用な情報となっていると考えられるのである。

その他の評価項目に関しては、正で有意になっているものはない。証券アナリスト推奨は有意水準5%で、負で有意となっている。つまり、証券アナリスト推奨に関してポジティブニュースが多かったとすると、翌営業日の株価は低下する傾向にあるのである。これは、証券アナリスト推奨のニュースは、同日の間に株価に必要以上に反映されているものが多く、翌営業日にリバウンドが観測されているのではないかと推察される。利益と売上については、ニュースが流れる前にコンセンサスが形成されている場合が多い。ニュースを単に評価しても対コンセンサスで評価したことには

ならないので、株価に有意な影響を与えない結果となったと考えられる。信用格付については、性質的には株価にインパクトを与えやすい情報と考えられるが、本稿の分析では、翌営業日の株価に影響を持たない結果となった。この原因の1つとしては、格付け変更（可能性）の情報が事前に十分に株価に反映されていたことが考えられるが、より正確な原因の究明は今後の課題としたい。最後に、製品・サービスとマクロ指標のニュースについては、これらのニュースが、必ずしも株価に対して新しい情報を含んでいるわけではないことが示唆されている。例えば、健康関連の新商品やGDPに関するニュースに絞り込むなどして、ニュースを細分化することにより、株価に有用な情報を抽出できる可能性はあるが、本稿の目的を超えるものであるので、これも今後の課題である。

以上の結果の頑健性を確認するために、説明変数の中から有意なものだけを抽出した重回帰モデル（モデル2）と有意な説明変数だけを用いた単回帰モデル（モデル3からモデル5）を推定し、その推定結果をまとめたものが**図表 6** の4列目から6列目である（注10）。結果から分かるように、モデル2からモデル5の結果は、モデル1の結果と完全に整合的であり、コーポレートアクションとマーケットコメントに関するニュースが、翌営

業日の株価に影響を与えている可能性が高いことが改めて確認できる。

5. 結論

本稿では、QUICK端末で配信されている日経ニュースを基に作成したニュース指標が、将来の株式市場に影響を持つかどうかを検証した。具体的には、ニュースデータに対して、定量処理を可能とするタグ情報を利用し、ニュース指標を作成した上で、そのニュース指標を用いた株式リターンや出来高の予測可能性について分析した。

分析の結果、ニュース指標は、翌日の株式リターンに有意に正の影響を持つこと、先行研究で見られたようなリバウンドが観測されることはないことが確認された。また、ニュース指標が翌日の出来高に対して、有意に正の影響を持つことも確認された。これは、WSJのコラムを基にニュース指標を作成し、同様の分析を行い、ニュース指標が市場センチメントを表すが、株価に対して本源的な情報を持たないとしたTetlock [2007] の結果とは対照的である。つまり、本稿で用いたニュース指標が、市場のセンチメントを表すのではなく、株価に対して本源的な情報を有している可能性を示唆しており、興味深い結果である。

また、ニュースの項目を細分化して分析した結果、コーポレートアクションとマーケットコメントに関するニュースが株価に対して本源的な情報を有していることが示唆された。言い換えれば、コーポレートアクションとマーケットコメントに関するニュースに投資家は有用な情報を見出し、それが株価の形成に影響を与えている可能性が高いことが確認されたのである。

以上より、本稿で用いたニュース指標には株価を予測できる情報が含まれている可能性が示唆され、ニュース情報の活用の有用性を示すことができたといえる。しかしながら、本稿の分析は全てインサンプルの分析であるので、本稿の結果が実際の投資戦略に応用できるかは定かではない。今後は、ニュース指標を用いた投資戦略の構築など、ニュース指標をより実用的な形で活用していくために、更なる研究が行われることを期待したい。

〔参考文献〕

- Campbell, J. Y., S. J. Grossman, and J. Wang [1993] "Trading Volume and Serial Correlation in Stock Returns," *Quarterly Journal of Economics* 108, pp. 905-939.
- DeLong, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann [1990] "Noise Trader Risk in Financial Markets," *Journal of Political Economy* 98, pp. 703-738.
- Fama, E. F., K. R. French [2012] "Size, Value, and Momentum in International Stock Returns," *Journal of Financial Economics* 105, pp. 457-472.
- Griffin, J. M., N. H. Hirschey, and P. J. Kelly [2011] "How Important is the Financial Media in Global Markets?" *Review of Financial Studies* 24, pp. 3941-3992.
- Roll, R. W. [1988] "R-Squared," *Journal of Finance* 43, pp. 541-566.
- Tetlock, P. C. [2007] "Giving Content to Investor Sentiment : The Role of Media in the Stock Market," *Journal of Finance* 62, pp. 1139-1168.
- Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy [2008] "More than Words : Quantifying Language to Measure Firms Fundamentals," *Journal of Finance* 63, pp. 1437-1467.

(この論文は投稿論稿を採用したものです。)

(注10) 紙幅の関係で、結果は載せていないが、有意でない変数を用いた単回帰モデルにおいては、全てのモデルで変数は有意とならなかった。