

自然言語処理による 景況感ニュース指数の構築と ボラティリティ予測への応用

ごしまけいいち / たかはしひろし / やまだてつや
五島圭一 / 高橋大志 / 山田哲也

要 旨

本研究では、日々配信されている経済ニュースを畳み込みニューラル・ネットワークを用いた自然言語処理によって指数化し、景気動向のナウキャストリングを試みた。そのうえで、構築した景況感ニュース指数（以下、ニュース指数）による資産価格のボラティリティ予測の有効性を検証した。分析の結果、ニュース指数は景気動向を表す経済指標と類似した推移を示しており、特に不確実性が高い時期においてナウキャストリングに有用である可能性が示された。また、ニュース指数はボラティリティ・モデルの予測力を向上させる可能性が示された。さらに、訓練データに利用した景気ウォッチャー調査の景気判断理由集のうち、現状判断を学習したモデルから構築されたニュース指数と先行き判断を学習したモデルから構築されたニュース指数では、予測やナウキャストリングに有用な場面が異なることが示された。

キーワード： 自然言語処理、テキスト・データ、畳み込みニューラル・ネットワーク、景気動向、実現ボラティリティ、ナウキャストリング

.....
本稿の作成に当たっては、石島博教授（中央大学）、山本裕樹氏（野村證券）、渡部敏明教授（一橋大学）ならびに金融研究所スタッフから有益なコメントを頂いた。ここに記して感謝したい。本研究で利用したニュース・データは、高橋大志（慶應義塾大学）より拠出したものである。ただし、本稿に示されている意見は、筆者たち個人に属し、日本銀行の公式見解を示すものではない。また、ありうべき誤りはすべて筆者たち個人に属する。

五島圭一 早稲田大学商学部講師、日本銀行金融研究所
(E-mail: keiichi.goshima@aoni.waseda.jp)

高橋大志 慶應義塾大学大学院経営管理研究科教授
(E-mail: htaka@kbs.keio.ac.jp)

山田哲也 日本銀行金融研究所企画役
(現金融機構局企画役、E-mail: tetsuya.yamada@boj.or.jp)

1. はじめに

本研究では、日々配信されている経済ニュースのテキスト・データを用いて、景気動向のナウキャストリングを試みる。ナウキャストリングとは、将来を予測するのではなく、現時点での景気や消費者物価指数などの各種統計について、公的統計が公表される前に把握することである。具体的な取組みとしては、GDPのナウキャストリング指標であるアトランタ連邦準備銀行の「GDPnow」やニューヨーク連邦準備銀行の「The New York Fed Staff Nowcast」が有名である¹。もちろん、官公庁や中央銀行、企業は古くからそれぞれ独自に取材やサーベイを通じてナウキャストリングすることを試みている。さらに近年、ビッグデータ分析や機械学習をはじめとする技術の発展、計算機の性能向上に伴い、これまで利用されていなかったデータをナウキャストリングに活用しようとする研究が進んでいる²。日本の例でみると、小売業のPOSデータから日次の消費者物価指数の算出を試みた日経CPINow（渡辺・渡辺 [2013]）や、SRI一橋大学消費者購買指数がある³。また、景気動向の面では、山本・松尾 [2016] において、政府・日本銀行の景況感を自然言語処理により月次で指標化するという試みが行われている。

計算言語学や自然言語処理の技術（あるいはそれらの研究成果物）によってテキスト・データを分析し、ファイナンスの諸問題に応用する研究については、2000年代から活発に取り組まれるようになった。これまで、中央銀行の議事録（Hansen and McMahon [2016]、Hansen, McMahon, and Prat [2018]）や企業の決算書（Li [2008]、Loughran and McDonald [2013]）、ニュース記事（Tetlock [2007]、Tetlock, Saar-Tsechansky, and Macskassy [2008]、Garcia [2013]）、インターネットへの投稿内容（Antweiler and Frank [2004]、Chen *et al.* [2014]）など、さまざまなテキスト・データが分析対象となっている。これらの研究では、テキスト・データを定量化して分析に活用することで、公的統計をはじめとする従来の数値データだけでは分析が困難であった新たな知見をファイナンス研究にもたらしている。

こうした研究の流れを受け、近年ではテキスト・データを利用して、経済環境や金融市場の動向を捉えるための新たな指標や手法を開発しようとする研究が行われている。有名な例として、経済政策の不確実性指数（Economic Policy Uncertainty 指数: EPU 指数）がある（Baker, Bloom, and Davis [2016]）。Baker, Bloom, and Davis [2016] は、主要な新聞から事前に定義したキーワードが含まれている記事を数え上げるこ

1 GDPnow: <https://www.frbatlanta.org/cqer/research/gdpnow.aspx> (2019年1月10日)、The New York Fed Staff Nowcast: <https://www.newyorkfed.org/research/policy/nowcast> (2019年1月10日)。

2 近年では、このような非伝統的なデータをオルタナティブ・データと総称することもある。

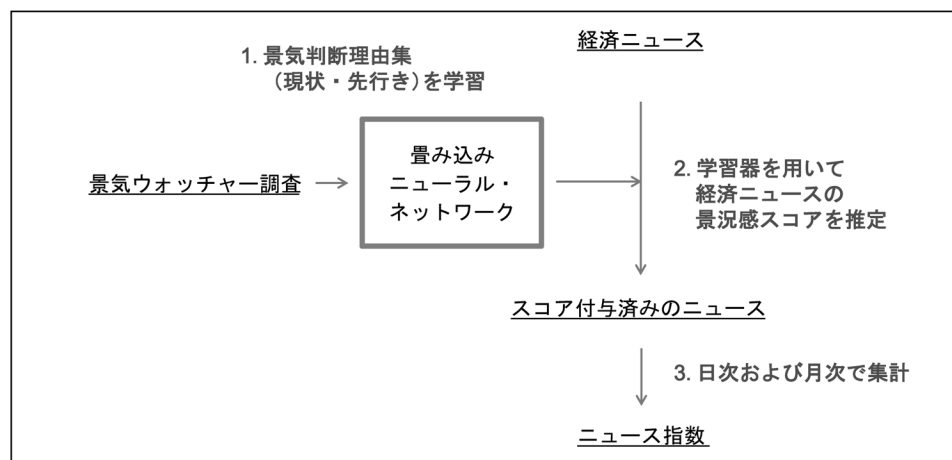
3 SRI一橋大学消費者購買指数: <https://risk.ier.hit-u.ac.jp/Japanese/nei/> (2019年1月10日)。

とで、経済政策の不確実性を表す指標を構築している。例えば、金融政策の不確実性であれば、経済関連のニュース記事であることを特徴付けるキーワード(“economy”あるいは“economic”)が含まれていて、かつ、不確実性を特徴付けるキーワード(“uncertainty”あるいは“uncertain”)が含まれていて、かつ、金融政策に関連するキーワード(“Fed”、“interest rate”、“inflation”など)が含まれている新聞記事を数え上げる。単純なルールではあるものの、EPU 指数は経済政策の不確実性を時系列数値データとして表現していることから、実証分析や経済状況の判断に広く用いられている。また、Loughran and McDonald [2011] は、金融関連のテキストに独自の語彙が多いことを踏まえ、有価証券報告書に相当する米国の Form 10-K のテキスト・データを分析するための専用辞書(Loughran and McDonald Sentiment Word Lists)を開発している。この辞書は、金融関連のテキスト・データ分析に広く利用されている。他に、Manela and Moreila [2017] では、超長期のウォール・ストリート・ジャーナルの記事からテキスト・ベースの不確実性指数である News Implied Volatility (NVIX) を構築している。彼らは、1987 年以降のニュースのテキスト・データを入力、金融市場における代表的な不確実性の指標である VIX を出力として、サポート・ベクター回帰によって学習器を構築し、学習器に過去のニュース記事を読み込ませることで、オプション取引市場が存在せず VIX が算出できない時期(1986 年以前)においても不確実性を表す指標である NVIX を構築できることを示した⁴。そのうえで、1929 年の世界大恐慌や世界大戦などの過去のイベント時における金融市場の不確実性の高まりを分析している。このように、テキスト・データを利用することで、これまで観察が困難であった経済環境や金融市場の動向を定量化して捉えられる可能性が示唆されている。そして、テキスト・データに基づいた新たな指標や、経済・金融関連のテキスト・データを分析するためのツールが活発に開発されている。

本研究は、こうしたテキスト・データ分析の発展に、手法面において、以下 2 つの貢献をなすものである。第 1 に、経済ニュースのテキスト・データから景況感を推定するに当たって、深層学習モデルの 1 つである畳み込みニューラル・ネットワーク(Convolutional Neural Network: CNN)を用いることである。CNN には、文章中の語順や単語の共起関係を考慮できるという強みがある。第 2 に、日次のニュースを指数化することによって、日次での景気動向を計測することである。これら 2 つをあわせ持つ研究は、われわれが知る限り初めてのものである。本研究の主たる分析結果は次のとおりである。はじめに、構築した景況感ニュース指数(以下、ニュース指数)の動きは、景気動向を示す経済指標の動きと類似していることが確認された。次に、リーマン・ショックをはじめとする不確実性の高い時期において、景気動向のナウキャストに資する情報のうち、資産価格が有していな

4 サポート・ベクター回帰とは、分類問題で広く用いられるサポート・ベクター・マシンを回帰問題へと拡張したアルゴリズムであり、 ϵ 不感応関数を利用することでロバストな推定が可能となる。

図1 ニュース指数の構築方法の概略図



いものを同ニュース指数は有している可能性が見いだされた。これを裏付けるように、同ニュース指数は、日経平均先物や長期国債先物のボラティリティの予測精度の向上に資するものであった。さらに、同ニュース指数は、ボラティリティ予測において、ジャンプ項、レバレッジ効果、およびインプライド・ボラティリティとは異なる情報を有していたことが示唆された。

本稿の構成は以下のとおりである。2節ではニュース指数の構築方法を説明し、3節では構築したニュース指数による景気動向のナウキャストに関する検証を行う。4節ではニュース指数によるボラティリティ予測への応用について分析し、5節では得られた結果をまとめるとともに今後の課題を示す。

2. ニュース指数の構築

(1) 枠組み

図1は、ニュース指数の構築方法の概略図である。まず、内閣府が公表している景気ウォッチャー調査の景気判断理由集を訓練データとして、CNNによる教師あり学習を行い、テキスト分類を行う学習器を構築する。次に、訓練した学習器を用いて、経済ニュースの日本語記事を構成する文に対して景況感に関するスコアを付与する。最後に、スコアが付与された文章を月次および日次で集計することによって、ニュース指数を構築する。

表 1 景気判断理由集例

(1) 景気の現状判断

景気の現状判断	判断理由
良くなっている	派遣先において、派遣スタッフの直接雇用への転換事例が増えている。
やや良くなっている	資金ニーズがやや増えてきている。
変わらない	年始は好調に推移したが、月の後半にかけて鈍化してきている。
やや悪くなっている	米国の新大統領就任による影響で、やや悪くなっている。
悪くなっている	当社特有の繁忙期が過ぎたため、受注量が減少している。

(2) 景気の先行き判断

景気の先行き判断	判断理由
良くなる	総決算セールと引っ越しシーズンが重なるため、良くなる。
やや良くなる	3月発売の新機種の予約が好調なので期待したい。
変わらない	既存顧客からの新規案件はなく、今のところ新年度以降も良くなる気配を感じない。
やや悪くなる	海外の政治的な情勢が安定してくれば、景気もやや上向いてくるのではないか。
悪くなる	米国の新大統領就任による影響で悪くなる。

(2) 景気ウォッチャー調査を用いた訓練データ

山本・松尾 [2016] を参考に、景気ウォッチャー調査から訓練データを作成する。景気ウォッチャー調査とは、「地域の景気に関連の深い動きを観察できる立場にある人々の協力を得て、地域ごとの景気動向を的確かつ迅速に把握し、景気動向判断の基礎資料とすることを目的」として、内閣府が毎月調査し公表しているデータである⁵。景気ウォッチャー調査には、景気判断理由集が収録されており、現状および先行きの景気に関して5段階（「良くなっている（良くなる）」・「やや良くなっている（やや良くなる）」・「変わらない」・「やや悪くなっている（やや悪くなる）」・「悪くなっている（悪くなる）」）の判断とその判断理由の文章がペアになっている。景気の現状判断は、「3か月前と比較し現状の景気が上向きか下向きか」という判断基準のもとでの回答であり、景気の先行き判断は、「当月と比較し今後2～3か月先の将来の景気が上向きか下向きか」という判断基準のもとでの回答である。この景気判断理由集を訓練データとして選択した理由は、映画レビューや商品レビューなどのコーパスと比較すると、経済や金融に関する語彙が相対的に豊富であるためである⁶。表1は、景気判断理由集を一部抜粋したものである。

5 景気ウォッチャー調査： https://www5.cao.go.jp/keizai3/watcher/watcher_menu.html (2019年1月10日)。

6 コーパスとは、言語研究に使用するために大量に収集され、電子化された言語資料のことである。

表2 景気判断理由集に収録されている文章数の内訳

景気の現状判断	文章数	景気の先行き判断	文章数
良くなっている	2,077	良くなる	2,248
やや良くなっている	22,882	やや良くなる	26,877
変わらない	49,017	変わらない	57,421
やや悪くなっている	23,345	やや悪くなる	22,671
悪くなっている	6,238	悪くなる	6,014
総計	103,559		115,231

景気の先行き判断に収録されている理由は、景気の現状判断に収録されているものと比較すると、「～とみている」や「～期待する」、「～だろう」などの将来に関して推量する句が入っている傾向があることが特徴である。一方で、現状判断は「～している」などの結果の状態を表す句が入っている傾向がある。この5段階の評価について、「良くなっている（良くなる）」を4、「やや良くなっている（やや良くなる）」を3、「変わらない」を2、「やや悪くなっている（やや悪くなる）」を1、「悪くなっている（悪くなる）」を0と置き換えることで、訓練データを作成する。本研究では、2011年1月から2017年9月までの景気判断理由集を教師あり学習の訓練データとして利用する。文章数はそれぞれ、景気の現状判断が103,559、景気の先行き判断が115,231である。表2は、5段階評価のデータの内訳である。

(3) 指数化するニュース・データ

ニュース指数を構築するためのニュース・データとして、ロイター・ニュースを用いる。ロイター・ニュースは、世界で最も広く知られたニュース提供会社の1つであるトムソン・ロイター社が配信しているニュースである。ロイター・ニュースは、各団体が発信する1次情報に比べると、記者やアナリストによる情報の取捨選択が行われており、社会や市場に対して相対的に重要な情報が含まれていると考えられる。また、新聞や雑誌のニュースに比べ、イベントからニュース記事配信までのラグが小さく、速報性に優れている。本研究では、「[東京 ○○日 ロイター]」から始まるニュース記事のみを分析対象とし、2003年1月から2015年5月までに配信された1,516,245本の日本語のニュース記事を取り扱う。利用したタグ情報は、ニュース記事の「配信日時」である。配信日時タグを利用することで、日次レベルの指標を構築することが可能となる。ただし、ニュース記事のテキスト情報に注目するため、決算情報や社債の発行要項、テクニカル・データなどの数値情報のみのニュース記事は分析対象外として、ルール・ベースで除外している。そして、

文単位でポジティブな表現とネガティブな表現を集計するために、ニュース記事から句点で終わる文のみを抽出している。また、抽出する際にニュース内容と関係ない文に関してもルール・ベースで除外している⁷。文の総数は 3,168,616 であり、語彙に関しては、延べ語数は 95,986,958、異なり語数は 100,233 であった⁸。

(4) テキスト分類モデルおよび集計方法

本研究では、深層学習モデルの 1 つである CNN によるテキスト分類モデルを用いている。自然言語処理や計算言語学の分野では、深層学習によるアプローチは、テキスト分類問題に対して高い分類精度が報告されている (Socher *et al.* [2013]、Kim [2014]、Lin *et al.* [2017])。なかでも、CNN は文章中の語順や単語の共起関係を考慮できるため、高い分類精度が期待される⁹。本研究で採用した Kim [2014] の CNN に基づくテキスト分類モデルの詳細は、補論 1 を参照されたい。

ニュース指数を構築するために、CNN によって 4 つの学習器を作成する。1 つ目の学習器 (learner 1) は景気判断理由集の現状判断と先行き判断の両方を入力し、5 段階のスコア (0~4) を出力するように学習する。2 つ目の学習器 (learner 2) は景気判断理由集の現状判断を入力に、3 つ目の学習器 (learner 3) は景気判断理由集の先行き判断を入力にして、どちらも 5 段階のスコアを出力するように学習する。4 つ目の学習器 (learner 4) は景気判断理由集の現状判断と先行き判断のどちらに近いかを予測する学習器であり、景気判断理由集の現状判断と先行き判断の両方を入力として、現状判断を 0、先行き判断を 1 としたバイナリ値を出力するように学習する。表 3 は、4 つの学習器をまとめたものである。

ロイター・ニュースの各文に対して、learner 1~learner 3 を用いてスコアを、learner 4 を用いて重みを付与する。learner 1~learner 3 は、5 段階のスコアを連続値として出力する学習器なので、各文に対しても連続値が付与され、4 より大きい値や負の値を取りうる。learner 1、learner 2、learner 3 から付与されたスコアを、それぞれ、 $Score_{both}$ 、 $Score_{current}$ 、 $Score_{future}$ と表す。また、learner 4 からは現状判断と先行き判断のどちらに近いかを表す確率値 W が出力される。learner 4 によって付

7 ニュース内容と関係ない文とは、「(前日比)は週初や休日明けには(前営業日比)を示します。」や「この記事の詳細はこの後送信します。」などの補足に関する文のことを指す。

8 延べ語数とは文章中に出現した単語の総数であり、異なり語数は重複を除いた単語の種類数である。

9 テキスト分類問題に関しては、Bag-of-Words に基づく方法、すなわち、ベクトルの各次元を 1 つの単語に対応付け、文章中の単語の出現回数や出現有無を入力することで、テキストをベクトルとして表現する方法もある。ニューラル・ネットワークに基づくモデルが Bag-of-Words に基づくモデルよりも常に優れているとは限らず、言語や問題に依存する。ただし、本研究では、景気判断理由集に関して、Bag-of-Words に基づく線形サポート・ベクター・マシンよりも CNN の分類精度が高いことを確認している。

表3 ニュース指数の構築に用いる学習器

	訓練データ	出力
learner 1	景気判断理由集 (現状判断・先行き判断)	景況感に関するスコア (0~4)
learner 2	景気判断理由集 (現状判断)	景況感に関するスコア (0~4)
learner 3	景気判断理由集 (先行き判断)	景況感に関するスコア (0~4)
learner 4	景気判断理由集 (現状判断・先行き判断)	バイナリ値 (0 or 1)

与えられる確率値 W は、テキストが先行き判断に近いほど 1 に、現状判断に近いほど 0 に近づく¹⁰。なお、深層学習モデルのパラメータを推定する際に用いられる確率的勾配法では、ドロップアウトや重み行列の初期値がランダムに与えられるため、同じハイパー・パラメータのもとでの学習結果が厳密に一致しないこともある。そこで本研究では、CNN の学習とスコアの付与を 10 回行い、その平均値を利用した。これは、アンサンブル学習の一種に相当する。

続いて、(1) 式~(3) 式によってスコアを月次および日次ごとに集計し、ニュース指数を構築する。

$$News\ Index_{both,t} = \frac{1}{N_t} \sum_{k=1}^{N_t} Score_{both,t,k} \tag{1}$$

$$News\ Index_{current,t} = \frac{\sum_{k=1}^{N_t} Score_{current,t,k} \cdot (1 - W_{t,k})}{\sum_{k=1}^{N_t} (1 - W_{t,k})} \tag{2}$$

$$News\ Index_{future,t} = \frac{\sum_{k=1}^{N_t} Score_{future,t,k} \cdot W_{t,k}}{\sum_{k=1}^{N_t} W_{t,k}} \tag{3}$$

ここで、 $Score_t$ は CNN によって付与されたスコアであり、 W_t は同様に付与された確率値である。添字 t は、各ニュースの文が配信された t 月あるいは t 日を表す。また、 N_t は t 月あるいは t 日に配信されたニュースに含まれる文数を表している。(1) 式は単純平均であり、(2) 式と (3) 式では確率値を文が有するトピック (現状あるいは先行き) 割合とみなして加重平均を取っている。月次の指数に関しては、月初から月末までのスコアを集計したものであり、例えば、1 月の指数であれば、1 月 1 日 0 時から 1 月 31 日 24 時までの間に配信されたニュース記事のスコアを集計している。日次の指数に関しては資産価格との関連性を分析するため、東京証券取引所の営業日の大引け 15 時時点のスコアを集計する。すなわち、ある営業日 ($t-1$) の 15 時から翌営業日 t の 15 時までの間に配信されたニュース記事からスコアを集計して、 t 日のニュース指数を算出する。以下、learner 1 から構築したニュース指数を「 $News\ Index_{both}$ 」、learner 2 と 4 から構築したニュース指数を「 $News\ Index_{current}$ 」、learner 3 と 4 から構築したニュース指数を「 $News\ Index_{future}$ 」と表記する。

10 出力される確率値のうち高い方のクラスを選択すれば、文の所属するクラスが 1 つに決まるが、本研究では出力された確率値をそのまま利用している。

表 4 ニュース指数の記述統計量

(1) ニュース指数（月次）の記述統計量

	<i>News Index_{current}</i>	<i>News Index_{future}</i>	<i>News Index_{both}</i>	文数
平均値	1.75	1.85	1.90	21,265
中央値	1.75	1.85	1.90	20,091
最小値	1.66	1.75	1.79	13,840
最大値	1.82	1.91	1.96	36,900
標準偏差	0.03	0.03	0.04	4,928
観測数	149	149	149	

(2) ニュース指数（日次）の記述統計量

	<i>News Index_{current}</i>	<i>News Index_{future}</i>	<i>News Index_{both}</i>	文数
平均値	1.75	1.85	1.90	1,042
中央値	1.75	1.85	1.91	978
最小値	1.56	1.66	1.68	104
最大値	1.89	1.99	2.04	3,021
標準偏差	0.04	0.05	0.05	321
観測数	3,039	3,039	3,039	

3. ニュース指数の検証

(1) ニュース指数の概観

ここでは、2 節までの手順によって構築したニュース指数を俯瞰する¹¹。まず、月次および日次で構築したニュース指数の記述統計量は表 4 のとおりである。2003 年 1 月から 2015 年 5 月までの東京証券取引所の営業日数は 3,043 営業日であったが、そのうち 4 営業日はニュースが配信されなかったため、観測数は 3,039 となっている¹²。表 4 をみると、ニュース指数の値は景気判断理由集の「変わらない」を示す 2 を常に下回っている。また、ニュース指数の値の範囲が 1.6 程度から 1.9 程度までであり、個々のニュース文に対して付与されたスコアの範囲と比較すると狭い。これは、中立な表現のテキストや景気と無関係なテキストが含まれており、

.....
11 補論 2 では、辞書に基づいて構築したニュース指数を用いて本節の分析を行った結果をまとめている。本研究で構築したニュース指数は、景気動向との連動性やナウキャストへの利用可能性といった点で、相対的に優れていることが確認された。

12 具体的には、2003 年 10 月 16・30・31 日、2004 年 10 月 13 日である。

表5 ニュース指数（月次）と景気動向の変化率との相関係数

	先行指数	一致指数	遅行指数	鉱工業生産指数
<i>News Index_{current}</i>	0.49	0.48	0.36	0.35
<i>News Index_{future}</i>	0.50	0.46	0.41	0.36
<i>News Index_{both}</i>	0.50	0.47	0.40	0.37

日次あるいは月次で集計すると、それらのスコアによってニュース指数の範囲が狭くなることが要因と考えられる。

図2は、表4で得られた平均値と標準偏差を用いて標準化したニュース指数の推移を図示したものである¹³。太い実線が先行き判断から構築したニュース指数（*News Index_{future}*）、実線が現状判断から構築したニュース指数（*News Index_{current}*）、点線が先行き判断と現状判断の両方から構築したニュース指数（*News Index_{both}*）の推移を示したものである。図2から、2008年9月のリーマン・ショックや2011年3月の東日本大震災などのイベント発生時にニュース指数が大きく低下していることが観察できる。

(2) 景気動向ナウキャストの検証

次に、構築したニュース指数が景気動向のナウキャストに利用できるかについて、内閣府が公表している景気動向指数および経済産業省が公表している鉱工業生産指数を対象として検証する。景気動向指数は、「生産、雇用など様々な経済活動での重要かつ景気に敏感に反応する指標の動きを統合することによって、景気の現状把握及び将来予測に資するために作成された指標」であり、通常、速報値が翌々月の上旬に公表されている。ここでは景気動向指数のうち、コンポジット・インデックス（先行指数・一致指数・遅行指数）を利用する。一方、鉱工業生産指数は、「鉱工業製品を生産する国内の事業所における生産、出荷、在庫に係る諸活動、製造工業の設備の稼働状況、各種設備の生産能力の動向を捉え、また、生産の先行き2か月の計画を把握することで、日本の生産活動をいち早く把握」するための指標であり、翌月末に速報が発表され、翌々月中旬に確報が発表されている。これらは月次の指標であるため、ここでは月次で集計したニュース指数を利用する。

まず、構築したニュース指数（月次）と景気動向指数および鉱工業生産指数との関係を概観する。表5は、ニュース指数の水準（月次）と景気動向指数および鉱工業生産指数の変化率との相関係数をまとめたものである。変化率は月次の対数階差（%）を用いている。各ニュース指数と景気動向に関する各指標の変化率との相関

13 後述の分析では利用しなかったが、スコアの（重み付き）分散、（重み付き）歪度、（重み付き）尖度を営業日ごとに集計した結果を補論3に記載した。

図2 標準化されたニュース指数の推移

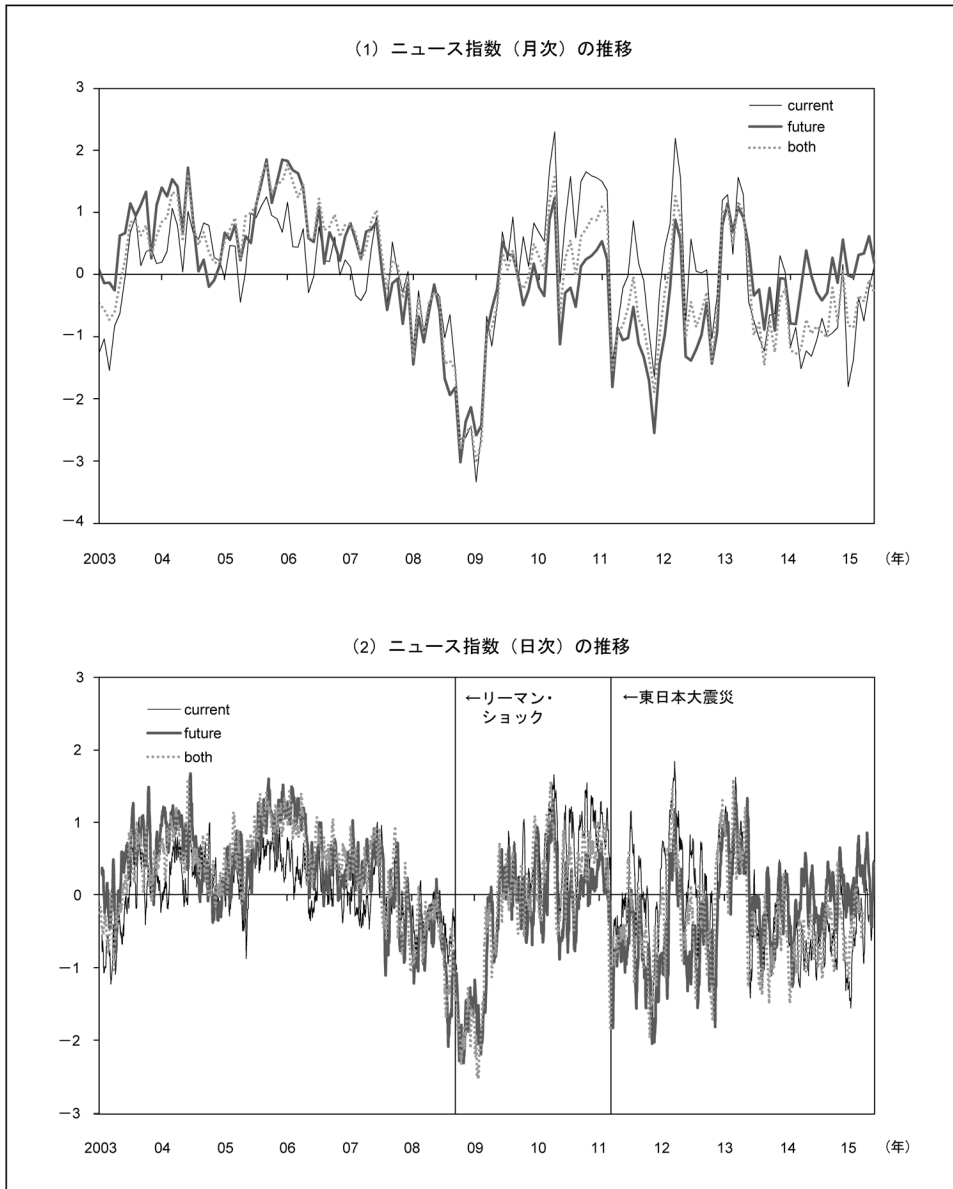
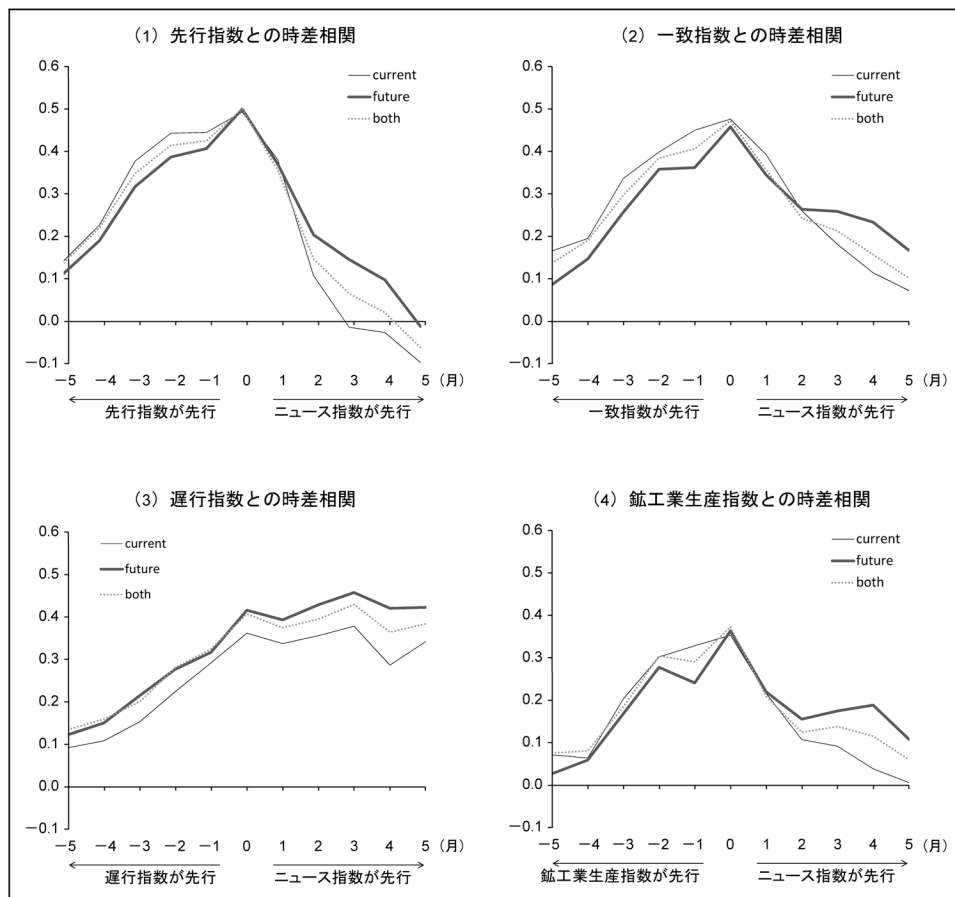


図3 ニュース指数と景気動向変化の先行遅行関係（時差相関）



係数は0.3～0.5程度であり、特に先行指数との相関係数が相対的に大きいことがわかる¹⁴。こうした特徴については、ニュース指数間で大きな差はみられない。

図3は、先行遅行関係を把握するために、ニュース指数と景気動向の変化との時差相関を図示したものである。縦軸が相関係数、横軸はラグを表しており、右方向はニュース指数が先行する方向、左方向は景気動向変化が先行する方向を示している。図3をみると、先行指数、一致指数および鉱工業生産指数との時差相関については、ラグが0のとき最も相関係数が大きくなり、ラグが大きくなるほど相関係数が小さくなる傾向にある。他方で、遅行指数との時差相関に関しては、ニュース指数が先行する方向において、ラグが大きい場合も高い正の相関を維持している傾向が窺える。これは、遅行指数が実際の景気状況から遅行する指標である一方、

14 景気動向の水準との相関係数を調べたところ、遅行指数との相関はみられたものの、先行指数・一致指数・鉱工業生産指数との相関はみられなかった。

ニュース指数のもととなるロイター・ニュースは速報性が高いことを反映していると考えられる。

また、ニュース指数が先行する方向では他の 2 指標 ($News Index_{current}$ 、 $News Index_{both}$) と比較して、 $News Index_{future}$ の相関係数が全体的に大きい傾向があることがわかる。とりわけ、ニュース指数が先行するほど差異が顕著である。一方で、景気動向の変化が先行する方向では、 $News Index_{current}$ と最も高く相関していることがわかる。これらの結果は、各ニュース指数を構築する際に、訓練データとして用いた景気判断理由集の文章の表現の違いに起因している可能性がある。すなわち、表 1 からわかるように、先行き判断は将来に関して推量する句が入っている文が多い一方で、現状判断は現在の景気に関して結果の状態を表す句が多いためだと考えられる。

次に、ニュース指数を景気動向のナウキャストに用いることができるかを検証するため、被説明変数を景気動向を表す指標、説明変数をニュース指数とする回帰分析を行う。ここで、被説明変数には、先行指数、一致指数、鉱工業生産指数、それぞれの変化率を用いる。また、ニュース指数の時系列を標準化した後の値を用いることで、ニュース指数の 1 標準偏差当たりの変化に対する景気動向の変化を考察する。さらに、コントロール変数として、被説明変数のラグ項と、景気動向指数の採用系列のうち、即時に観察可能な金融データである長期国債 (10 年) 新発債流通利回りの月次変化 (ΔJGB) と東証株価指数の月次対数収益率 ($\Delta TOPIX$) を加えた回帰分析も行う。長期国債 (10 年) 新発債流通利回り と東証株価指数のデータは日経 NEEDS より取得した。具体的な推計式は、(4) 式のとおりである。

$$\Delta Business Trend_t = \alpha + \beta_{NI} News Index_t + \beta_X X_t + \epsilon_t. \quad (4)$$

$News Index_t$ 、 $\Delta Business Trend_t$ 、 X_t 、 ϵ_t はそれぞれ、 t 期におけるニュース指数、景気動向関連変数の変化率、コントロール変数、誤差項を表す。ここで重要なパラメータは β_{NI} である。 β_{NI} が有意な値であれば、ナウキャストに用いることができる可能性がある。表 6~8 は推計結果をまとめたものである。

推計結果をみると、コントロール変数を考慮しない推計式とラグ項のみを考慮した推計式では、ニュース指数の回帰係数が有意になっている。しかしながら、資産価格変動をコントロール変数に追加すると、ニュース指数の回帰係数は総じて有意ではなくなる。これは、ニュース指数は景気動向をナウキャストするうえで、資産価格変動と類似した情報を持っている可能性が高いことを示唆している。また、決定係数や赤池情報量規準 (Akaike's Information Criterion: AIC) やベイズ情報量規準 (Bayesian Information Criterion: BIC) といった各種情報量規準の観点で、先行指数のナウキャストに資産価格変動が大きく寄与しているのは、先行指数を構成する指数の中に TOPIX や金利が採用されていることが理由と考えられる。

表6 先行指数のナウキャストイング

先行指数の変化率												
<i>News Index_{current}</i>	0.85*** (2.76)			0.49*** (2.08)		0.63*** (3.41)		0.37** (2.30)		0.22 (1.60)		0.24* (1.71)
<i>News Index_{future}</i>	0.86*** (2.87)			0.53*** (2.79)		0.62*** (3.67)		0.42*** (2.78)		0.07 (0.59)		-0.04 (-0.37)
<i>News Index_{both}</i>	0.86*** (2.83)					0.63*** (3.58)				0.16 (1.23)		
$\Delta L a g_{t-1}$				0.30*** (2.71)		0.30*** (2.97)		0.30*** (2.82)		0.27** (2.54)		0.30*** (3.61)
$\Delta L a g_{t-2}$				0.02 (0.18)		0.04 (0.30)		0.03 (0.24)		0.01 (0.05)		0.06 (0.55)
$\Delta T O P I X$										0.17*** (7.71)		0.18*** (7.71)
$\Delta J G B$										0.17*** (7.73)		0.17*** (7.20)
定数項	0.08 (0.47)	0.08 (0.43)	0.08 (0.44)	0.08 (0.49)	0.04 (0.33)	0.05 (0.38)	0.05 (0.36)	0.05 (0.40)	-0.03 (-0.31)	-0.03 (-0.34)	-0.03 (-0.31)	-0.03 (-0.34)
観測数	149	149	149	149	147	147	147	147	147	147	147	147
修正済み決定係数	0.24	0.24	0.25	0.28	0.31	0.32	0.32	0.34	0.50	0.49	0.49	0.50
AIC	549.51	548.22	548.04	541.09	530.95	529.05	529.58	526.14	486.23	488.78	487.62	488.16
BIC	558.52	557.23	557.06	553.11	545.91	544.00	544.54	544.08	507.17	509.71	508.55	512.08

備考：表は(4)式を最小二乗法によって推計した結果をまとめたものである。括弧内は各回帰係数の t 値を表している。 t 値は Bartlett Kernel と Prewhitening に基づく HAC 推定量によって補正した値である。ラグ次数は2とした。なお、ラグ次数を変えても同様の結果が得られることを確認している。***、**、*はそれぞれ、両側確率1%、5%、10%で回帰係数が有意であることを示している。また、修正済み決定係数、AIC および BIC については最もパフォーマンスがよい値を太字にしている。以下、表7、8とも同じ。

表7 一致指数のナウキャストイング

一致指数の変化率												
<i>News Index_{current}</i>	0.84* (1.95)			0.54 (1.61)		0.54** (2.37)		0.33 (1.64)		0.33 (1.44)		0.27 (1.33)
<i>News Index_{future}</i>	0.81* (1.94)			0.44* (1.82)		0.51*** (2.70)		0.33** (2.04)		0.27 (1.38)		0.14 (0.88)
<i>News Index_{both}</i>	0.84* (1.90)					0.53** (2.57)				0.31 (1.44)		
$\Delta L a g_{t-1}$				0.26** (2.12)		0.29** (2.19)		0.27** (2.20)		0.25** (2.07)		0.28** (2.20)
$\Delta L a g_{t-2}$				0.18 (1.02)		0.18 (1.02)		0.18 (1.01)		0.16 (0.93)		0.19 (1.14)
$\Delta T O P I X$										0.08*** (3.03)		0.08*** (2.81)
$\Delta J G B$										0.08*** (2.86)		0.07** (2.35)
定数項	0.07 (0.36)	0.07 (0.32)	0.07 (0.34)	0.07 (0.36)	0.02 (0.14)	0.03 (0.18)	0.02 (0.16)	0.03 (0.17)	-0.01 (-0.11)	-0.01 (-0.08)	-0.01 (-0.09)	-0.01 (-0.07)
観測数	149	149	149	149	147	147	147	147	147	147	147	147
修正済み決定係数	0.22	0.20	0.22	0.25	0.33	0.33	0.33	0.35	0.36	0.36	0.36	0.36
AIC	559.89	563.14	560.64	555.14	533.42	533.09	533.04	531.36	528.10	530.10	529.01	529.47
BIC	568.90	572.15	569.66	567.15	548.38	548.05	547.99	549.30	549.04	551.03	549.94	553.39

これらの結果は、景気動向に関して資産価格がニュース指数よりも重要な情報を有している可能性を示唆している。

表 8 鉱工業生産指数のナウキャストイング

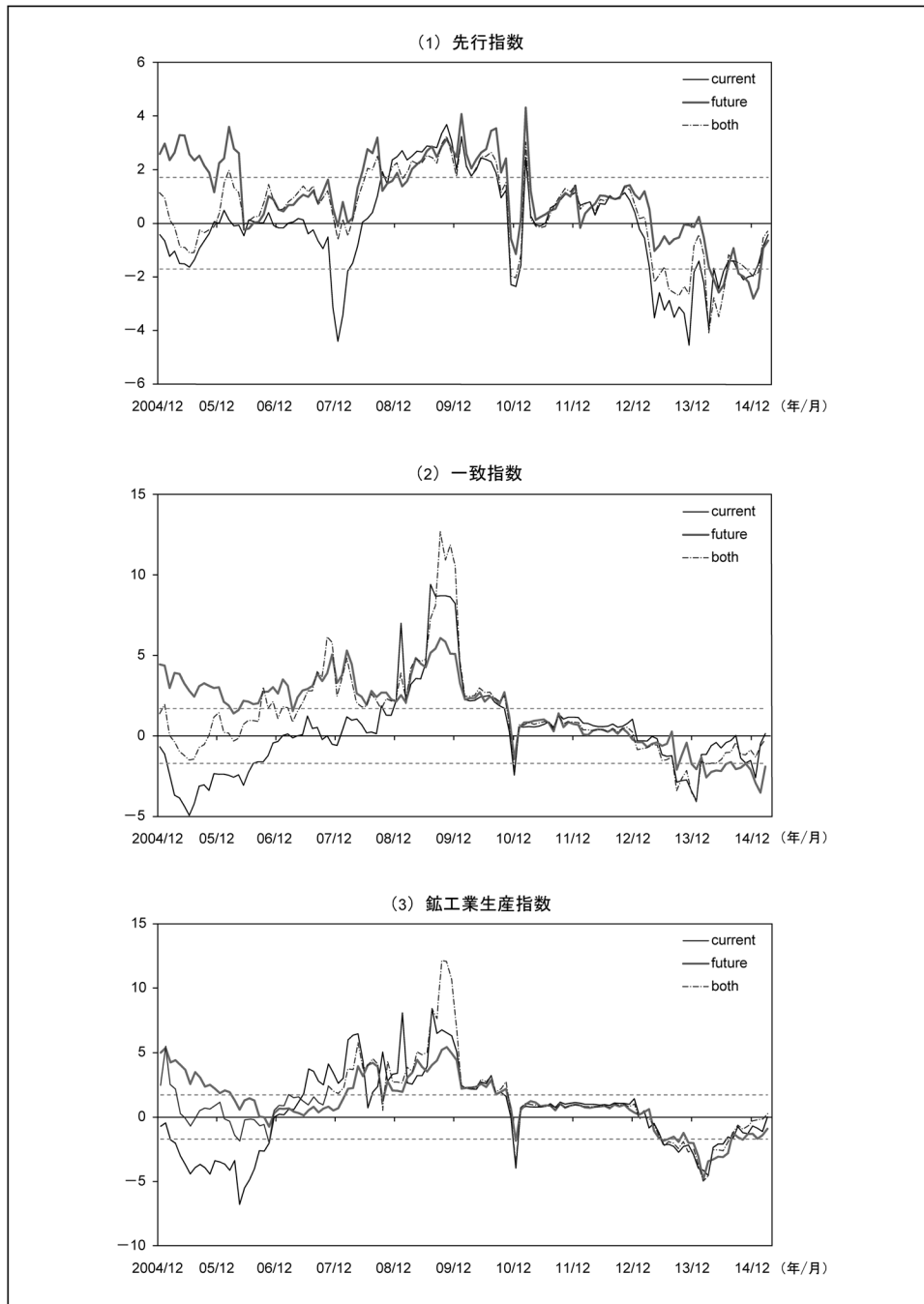
鉱工業生産指数の変化率											
<i>News Index_{current}</i>	0.90*		0.51		0.83**		0.45		0.62		0.41
	(1.89)		(1.31)		(1.99)		(1.38)		(1.39)		(1.20)
<i>News Index_{future}</i>	0.93**		0.59**		0.84**		0.57**		0.66		0.44
	(2.15)		(2.02)		(2.40)		(2.36)		(1.59)		(1.53)
<i>News Index_{both}</i>	0.95**				0.88**				0.69		
	(2.13)				(2.35)				(1.62)		
ΔLag_{t-1}			0.10		0.13		0.11		0.10		0.12
			(0.77)		(0.90)		(0.82)		(0.78)		(0.93)
ΔLag_{t-2}			0.00		0.00		-0.01		-0.01		0.00
			(0.03)		(0.02)		(-0.04)		(-0.10)		(0.00)
$\Delta TOPIX$							0.10**		0.08		0.08
							(2.01)		(1.48)		(1.65)
ΔJGB							-1.34		-1.70		-1.63
							(-0.81)		(-1.04)		(-1.00)
定数項	-0.02	-0.02	-0.02	-0.02	-0.03	-0.02	-0.02	-0.02	-0.07	-0.06	-0.06
	(-0.07)	(-0.07)	(-0.07)	(-0.07)	(-0.15)	(-0.09)	(-0.12)	(-0.13)	(-0.40)	(-0.31)	(-0.35)
観測数	149	149	149	149	147	147	147	147	147	147	147
修正済み決定係数	0.12	0.13	0.13	0.14	0.12	0.13	0.13	0.14	0.14	0.14	0.14
AIC	687.57	686.33	685.13	684.61	682.22	680.37	679.87	679.68	681.23	681.27	680.12
BIC	696.58	695.34	694.14	696.62	697.17	695.33	694.83	697.62	702.16	702.20	701.05

一方で、全ての期間において資産価格のみが景気動向のナウキャストイングに安定的に役立つかは定かではない。サンプル期間には、リーマン・ショックや東日本大震災といった不確実性が大いに高まった時期が含まれる。こうした時期において、景気動向のナウキャストイングには、金融資産価格だけでこと足りるのだろうか。この問いに答えるために、ウィンドウを 24 ヶ月としたローリング推計を行い、ニュース指数の回帰係数における t 値の推移を観察することで、統計的な有意性の時間を通じた変化を確認する。図 4 は、その結果である。縦軸は t 値、横軸は時間、点線は t 分布の 90% 点を表している。

分析の結果、特定の期間においては、資産価格変動をコントロール変数に入れた場合でもニュース指数が有意となることを確認できる。具体的には、2008 年前後の結果をみると、先行指数、一致指数、鉱工業生産指数のいずれの結果においてもニュース指数の回帰係数は有意となっている。2008 年前後の時期は、リーマン・ショックなど、市場の不確実性が高まった時期と重なっており、こうした期間において、ニュース指数は一定の説明力を確保している。これらの結果は、不確実性が高い時期において、景気動向のナウキャストイングに資する情報のうち、資産価格が有していないものをニュース指数が有していたと理解できる。

また、時間の概念がないデータである景気判断理由集から構築した学習器に、ロイター・ニュースを配信順に入力すると、景気動向指数や鉱工業生産指数と似た挙動を示す指数を構築できることも興味深い点であるが、これは、ロイター・ニュースの特徴と関連していると考えられる。ロイター・ニュースは、ビジネスマンや投資家向けのニュースであることから、日々の経済・金融に関する重要な情報が記述

図4 ローリング推計による β_{NI} の t 値の推移



されている。すなわち、ニュース・テキストとして形作られる過程で経済・金融に関する重要な情報が取捨選択されている。そのため、ニュース記事中の景況感に関する文章表現を集計できれば、景気動向を捉えられるものと考えられる。本研究の分析結果は、景気に関する文章表現が豊富である景気ウォッチャー調査の景気判断理由集を訓練データとして利用することで、景気動向を反映したニュース指数を構築できることを示唆している。

4. ボラティリティの予測力

3節において、不確実性の高い時期においては、景気動向ナウキャストイングに資する情報のうち、資産価格が有していないものをニュース指数が有していたことが示唆された。この特別な情報は、金融資産にとって新たなショックになり、後の金融資産価格のボラティリティを左右する可能性がある。そこで、本節では、ニュース指数がボラティリティ予測に有用であるか考察する。資産価格の主たる変動要因がニュースであるとの立場から、QGARCH (Quadratic Generalized Autoregressive Conditionally Heteroskedastic) モデルによってニュースが流入した時のボラティリティ・フィードバック効果を分析した研究 (Campbell and Hentschel [1992]) やニュース・インパクト曲線によってニュースがボラティリティに与える影響を分析した研究 (Engle and Ng [1993]) など、これまで多くのボラティリティ分析への応用がなされている。そこで本節では、ボラティリティとの関連性を通じて、ニュース指数が持つ情報の分析を行う。

まず、ニュース指数と主要な金融変数との関連性を概観する。表9は、ニュース指数と資産価格変動およびボラティリティとの相関関係をまとめたものである¹⁵。東証株価指数 (TOPIX)、日経平均株価、日経平均先物、外国為替 (ドル/円) についてはそれらの日次対数収益率との相関係数を、新発 10 年物国債利回りについてはその日次差分との相関係数を、日経平均ヒストリカル・ボラティリティ (以下、日経平均 HV)、日経平均ボラティリティ・インデックス (以下、日経平均 VI)、VXJ (Volatility Index Japan) はそのままの値との相関係数を、実現ボラティリティ (Realized Volatility: RV) についてはその正の平方根を取った値との相関係数を、それぞれ算出している¹⁶。RV の定義および算出方法は後述する。欠落している 4 営業日分のニュース指数に関しては、それぞれの前営業日のニュース指数を用いるこ

15 2014 年 4 月 14 日の新発 10 年物国債利回りは取引が成立せず、値が付かなかったため前営業日の値で補間している。ドル/円のデータは 17 時時点での取引価格を用いているため、ニュース指数との関連性を観察する際には、17 時で集計したニュース指数を用いている。

16 平方根を利用した理由は脚注 17 を参照されたい。

表 9 ニュース指数と金融変数との関連性

(1) ニュース指数と資産価格変動との相関係数

	TOPIX	日経平均株価	日経平均先物	外国為替 (ドル円)	新発 10 年物国債利回り
<i>News Index_{current}</i>	0.25	0.24	0.23	0.18	0.11
<i>News Index_{future}</i>	0.23	0.22	0.21	0.16	0.08
<i>News Index_{both}</i>	0.26	0.25	0.24	0.18	0.12

(2) ニュース指数とボラティリティとの相関係数

	日経平均 HV	日経平均 VI	VXJ	$RV_{\text{日経平均先物}}^{1/2}$	$RV_{\text{長期国債先物}}^{1/2}$
<i>News Index_{current}</i>	-0.29	-0.37	-0.36	-0.32	-0.10
<i>News Index_{future}</i>	-0.32	-0.43	-0.43	-0.33	-0.08
<i>News Index_{both}</i>	-0.36	-0.48	-0.47	-0.38	-0.06

備考：東証株価指数 (TOPIX)、日経平均株価、日経平均先物、日経平均 HV、日経平均 VI、新発 10 年物国債利回りは日経 NEEDS より、VXJ (Volatility Index Japan) は大阪大学数理・データ科学教育研究センターのウェブページより、外国為替 (ドル/円) は日本銀行のウェブページより、各日次データをそれぞれ取得した。日経平均先物と長期国債先物の RV ($RV_{\text{日経平均先物}}$ 、 $RV_{\text{長期国債先物}}$) は日経メディアマーケティング株式会社の NEEDS ティックデータより取得した高頻度データから算出した。

とで補間している。

ニュース指数と資産価格変動との相関係数は 0.1~0.3 の範囲であり、その水準は大きくない。株式に関しては若干の相関がみられる一方、特に国債に関しては相関がほとんどみられない。株式市場のボラティリティとの相関係数は、-0.3~-0.5 の範囲であり、逆相関となっている。また、日経平均 HV や RV に比べて、日経平均 VI や VXJ といったインプライド・ボラティリティとはより強い負の相関が観察される。一方で、国債市場のボラティリティは株式と比べると相関がみられない。

次に、資産価格のボラティリティの予測分析を通じて、ニュース指数が持つ情報の分析を行う。本研究では、RV と、その変動を表すモデルの 1 つである Corsi [2009] の HAR (Heterogeneous Autoregressive) モデルを用いて検証を行う。t 日における RV である RV_t は、(5) 式のように t 日において日中の n_t 個の対数収益率 $\{r_1, r_2, r_3, \dots, r_{n_t}\}$ が観測された時、それらの二乗和として定義される¹⁷。

$$RV_t := \sum_{i=1}^{n_t} r_i^2. \quad (5)$$

日中の対数収益率の観測数 n_t が十分に大きい時、 RV_t は累積ボラティリティの一致

17 RV に関する詳細な解説は、Andersen and Bollerslev [2003] や渡部 [2007] を参照されたい。また、文献によっては (5) 式の RV_t を実現分散 (Realized Variance)、そしてその平方根を実現ボラティリティと表記することもあるが、本稿では (5) 式の RV_t を実現ボラティリティと表記する。

表 10 取引時間の推移

資産	期間	取引時間
日経平均先物	2003/01/01–2011/02/10	9:00–11:00、12:30–15:10
	2011/02/14–2015/05/29	9:00–15:10
長期国債先物	2005/01/04–2011/11/18	9:00–11:00、12:30–15:00
	2011/11/21–2015/05/29	8:45–11:00、12:30–15:00

推定量となる。本研究では、2003年1月から2015年5月までの日経平均先物および2005年1月から2015年5月までの長期国債先物の高頻度データを用いて、それぞれ RV_t を算出する。ここでは、流動性の高い期近限月のデータを用いる。ただし、流動性の観点から最終取引日においては次の限月のデータを利用する。

本研究では、Liu, Patton, and Sheppard [2015] を踏まえ、レギュラー・セッションにおける5分間隔の約定価格データを用いた標準的なRVを採用した¹⁸。例えば、取引時間が9時から11時30分と12時30分から15時までの場合、9:00、9:05、9:10、…、11:25、12:30、12:35、12:40、…、14:55の各時刻において最初に取引が行われた際の約定価格と、前場の終値（11:30）と後場の終値（15:00）を用いる。表10は、本研究で取り扱った高頻度データの取引時間をまとめたものである¹⁹。

なお、RVは日中のデータから算出されるため、夜間や昼休みなどの取引がない時間はリターンが算出できず、推定値が過小評価される。そこで、分析期間において、昼休みと夜間を除いて算出したRV（以下、 RV_t^* ）の平均値と日次リターン（ R_t ）の標本分散が等しくなるように、(6)式によってRVを補正する（Hansen and Lunde [2005]）。

$$RV_t = cRV_t^*, \quad c = \frac{\sum_{t=1}^T (R_t - \bar{R})^2}{\sum_{t=1}^T RV_t^*}, \quad \bar{R} = \sum_{t=1}^T R_t. \quad (6)$$

そして、 t_1 から t_2 までのRVの平均値（ $RV_{t_1:t_2}$ ）を、

$$RV_{t_1:t_2} = \frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} RV_t, \quad (7)$$

.....
 18 RVを推定する際、市場のマイクロ構造に起因する観測誤差であるマーケット・マイクロストラクチャー・ノイズの影響を受けることが知られている。この問題に対処するために、サブサンプル法やカーネル法、自己相関調整法などのさまざまな補正方法や推定方法が提案されている。こうした中、Liu, Patton, and Sheppard [2015]では、31種類の資産価格の高頻度データを用いて400種類のRVの推定精度を比較したところ、5分間隔の約定価格データを用いた日中の収益率の二乗和という標準的なRVで、十分にマーケット・マイクロストラクチャー・ノイズを軽減でき、さらに、他の推定方法に必ずしも劣っていないことを報告している。

19 この期間、前場のみの取引営業日（大発会や大納会）を含んでいる。

と表記したとき、Corsi [2009] の HAR モデルは、(8) 式のように表される。

$$RV_{t:t+h-1} = \alpha + \beta_d RV_{t-1} + \beta_w RV_{t-5:t-1} + \beta_m RV_{t-22:t-1} + \epsilon_t. \quad (8)$$

ここで、 RV_{t-1} 、 $RV_{t-5:t-1}$ 、 $RV_{t-22:t-1}$ は、それぞれ前営業日 ($t-1$ 時点) の日次、週次、月次の RV を表しており、 ϵ_t は正規ホワイト・ノイズに従う誤差項を表している。すなわち、HAR モデルは、将来の RV が過去の日次、週次、月次の RV に依存するモデルとなる。RV の分布は右方向に歪んでおり、対数正規分布に近いことが知られている。そのため、HAR モデルに関する研究では、RV を対数値である $\ln RV$ に代替して正規分布に近似させることが多く、本研究においても同様に対応している²⁰。そして、(8) 式の HAR モデルにおいて、日次、週次、月次の RV をコントロールしたうえで、ニュース指数を追加して推定する。つまり、(9) 式を用いて、ニュース指数が持つ情報を分析する。

$$\begin{aligned} \ln RV_{t:t+h-1} = & \alpha + \beta_d \ln RV_{t-1} + \beta_w \ln RV_{t-5:t-1} + \beta_m \ln RV_{t-22:t-1} \\ & + \beta_{NI} News Index_t + \epsilon_t. \end{aligned} \quad (9)$$

はじめに、全期間のデータ（日経平均先物は 2003 年 1 月 5 日から 2015 年 5 月 29 日までの 3,043 営業日、長期国債先物は 2005 年 1 月 4 日から 2015 年 5 月 29 日までの 2,552 営業日のデータ）を用いて (9) 式を推定し、 β_{NI} が有意であるかを検証する。(9) 式の推計には最小二乗法を利用する。次に、アウト・オブ・サンプルに対する予測力を検証する。500 営業日をウィンドウとしたローリング推計によって、日次 ($h=1$)、週次 ($h=5$)、月次 ($h=22$) について、それぞれ 1 期先の RV を予測する。最後に、Patton [2011] を参考にして、実現値と予測値との予測誤差について、最小二乗誤差 (Mean Squared Error: MSE) および疑似尤度 (Quasi-Likelihood: QLIKE) の 2 つの損失関数と Mincer-Zarnowitz の R^2 (MZ- R^2) の 3 つの評価基準に基づき、アウト・オブ・サンプルに対する予測力を評価する。MSE と QLIKE については、 t 期における損失関数 (L_t^{MSE} 、 L_t^{QLIKE}) を (10) 式のように定義する。

$$L_t^{MSE} := (RV_t - \widehat{RV}_t)^2, \quad L_t^{QLIKE} := \frac{RV_t}{\widehat{RV}_t} - \ln \frac{RV_t}{\widehat{RV}_t} - 1. \quad (10)$$

ここで、 RV_t は実現値、 \widehat{RV}_t は予測値を表している。そのうえで、全期間 T の MSE と QLIKE を、(11) 式のように、予測誤差の平均値によって評価する。

$$MSE = \frac{1}{T} \sum_{t=1}^T L_t^{MSE}, \quad QLIKE = \frac{1}{T} \sum_{t=1}^T L_t^{QLIKE}. \quad (11)$$

.....
20 本研究でも、RV および $RV^{1/2}$ を変数として HAR モデルを最小二乗法で推定したところ、 $\ln RV$ を用いたモデルの方が決定係数やアウト・オブ・サンプルに対する予測力が高いことを確認している。

表 11 日次の RV (RV_t) に対する予測力

	日経平均先物					長期国債先物				
<i>News Index current</i>	-0.77*** (-4.07)					-0.50** (-2.55)				
<i>News Index future</i>	-0.73*** (-3.88)					-0.47** (-2.39)				
<i>News Index both</i>	-0.78*** (-4.33)					0.41* (1.84)				
$\ln RV_{t-1}$	0.41*** (12.68)	0.39*** (12.64)	0.39*** (12.62)	0.39*** (12.46)	0.39*** (12.59)	0.29*** (10.51)	0.29*** (10.52)	0.29*** (10.54)	0.29*** (10.55)	0.29*** (10.58)
$\ln RV_{t-5:t-1}$	0.29*** (6.50)	0.29*** (6.42)	0.29*** (6.31)	0.29*** (6.35)	0.29*** (6.32)	0.38*** (9.32)	0.38*** (9.27)	0.38*** (9.15)	0.38*** (9.11)	0.38*** (9.15)
$\ln RV_{t-22:t-1}$	0.20*** (5.77)	0.19*** (5.57)	0.20*** (5.55)	0.19*** (5.40)	0.19*** (5.50)	0.25*** (6.43)	0.25*** (6.40)	0.26*** (6.42)	0.26*** (6.48)	0.25*** (6.39)
MSE	18.504	18.363	18.399	18.394	18.345	0.007	0.007	0.007	0.007	0.007
QLIKE	0.128	0.124	0.126	0.124	0.124	0.244	0.244	0.242	0.244	0.242
MZ-R ²	0.311	0.315	0.314	0.314	0.316	0.315	0.316	0.317	0.316	0.317

備考：表はそれぞれ、 RV_t 、 RV_{t+4} 、 RV_{t+21} に対する予測力の検証のため、(9) 式を推計した結果をまとめたものである。括弧内は各回帰係数の t 値を表している。 t 値は Bartlett Kernel と Prewhitening に基づく HAC 推定量によって補正した値である。ラグ次数は先行研究に倣い、5 ($h = 1$)、10 ($h = 5$)、44 ($h = 22$) としている。***、**、*はそれぞれ、両側確率 1%、5%、10% で回帰係数が有意であることを示している。QLIKE、MSE、MZ-R² はアウト・オブ・サンプルに対する予測力をまとめたものであり、最もパフォーマンスがよい値を太字にしている。以下、表 12、13 と同じ。

また、MZ-R² は、(12) 式のように、被説明変数に実現値、説明変数に予測値を用いて推定した単回帰式の決定係数として算出する。

$$RV_t = \alpha + \beta \widehat{RV}_t + \epsilon_t. \tag{12}$$

MSE と QLIKE は値が低いほど、MZ-R² は値が高いほどモデルの予測力が高いことを表している。

表 11~13 はそれぞれ、 RV_t 、 RV_{t+4} 、 RV_{t+21} について、(9) 式を推計した結果をまとめたものである。また、比較対象として HAR モデルの推計結果もあわせて記載している。日経平均先物の RV_t および RV_{t+4} については、ニュース指数の回帰係数が概ね有意であることがわかる。すなわち、イン・サンプルに対する検証では、ニュース指数が予測力向上に寄与している可能性が示された。また、回帰係数は負の値となっていることから、景況感の低いニュースが増えたとき、翌営業日あるいは翌週のボラティリティが上昇する傾向にあることを示唆している。一方で、 RV_{t+21} では回帰係数が有意ではなく、予測力の向上に寄与しないことが示唆される。しかしながら、アウト・オブ・サンプルに対する検証では、ニュース指数を追加すると、いずれの評価基準においても予測力は向上する傾向がある。さらに、 RV_t 、 RV_{t+4} 、 RV_{t+21} のいずれにおいても予測力は向上しており、特に、 RV_{t+21} に対する予測力が相対的に大きく向上している。また、長期国債先物の RV に対する

表 12 週次の RV ($RV_{t:t+4}$) に対する予測力

	日経平均先物					長期国債先物				
	<i>News Index_{current}</i>	-0.56** (-2.26)					-0.32 (-1.34)			
<i>News Index_{future}</i>	-0.60** (-2.05)					-0.44 (-1.45)				
<i>News Index_{both}</i>	-0.63** (-2.32)					0.06 (0.16)				
<i>ln RV_{t-1}</i>	0.27*** (6.76)	0.26*** (6.70)	0.26*** (6.78)	0.26*** (6.56)	0.26*** (6.73)	0.21*** (7.91)	0.21*** (7.93)	0.21*** (7.92)	0.21*** (7.94)	0.21*** (7.98)
<i>ln RV_{t-5:t-1}</i>	0.31*** (4.11)	0.31*** (4.10)	0.31*** (4.12)	0.31*** (4.09)	0.31*** (4.08)	0.33*** (4.95)	0.33*** (4.95)	0.33*** (4.91)	0.32*** (4.91)	0.33*** (4.91)
<i>ln RV_{t-22:t-1}</i>	0.26*** (4.04)	0.25*** (3.99)	0.25*** (3.96)	0.25*** (3.91)	0.25*** (3.94)	0.36*** (4.66)	0.36*** (4.67)	0.36*** (4.68)	0.37*** (4.78)	0.36*** (4.67)
MSE	9.315	9.307	9.310	9.345	9.251	0.003	0.003	0.003	0.003	0.003
QLIKE	0.126	0.124	0.123	0.124	0.122	0.209	0.211	0.207	0.209	0.209
MZ-R ²	0.362	0.364	0.362	0.361	0.365	0.439	0.438	0.443	0.439	0.443

表 13 月次の RV ($RV_{t:t+21}$) に対する予測力

	日経平均先物					長期国債先物				
	<i>News Index_{current}</i>	-0.29 (-0.61)					0.08 (0.18)			
<i>News Index_{future}</i>	-0.63 (-1.09)					-0.67 (-1.08)				
<i>News Index_{both}</i>	-0.47 (-0.92)					0.66 (1.04)				
<i>ln RV_{t-1}</i>	0.18*** (4.94)	0.17*** (4.90)	0.17*** (5.24)	0.17*** (4.99)	0.17*** (5.09)	0.13*** (7.90)	0.13*** (7.92)	0.13*** (7.83)	0.13*** (7.94)	0.13*** (7.82)
<i>ln RV_{t-5:t-1}</i>	0.29*** (3.64)	0.28*** (3.64)	0.28*** (3.69)	0.28*** (3.65)	0.28*** (3.68)	0.21** (2.41)	0.21** (2.36)	0.20** (2.31)	0.20** (2.29)	0.20** (2.31)
<i>ln RV_{t-22:t-1}</i>	0.23** (2.07)	0.22** (2.06)	0.22** (2.02)	0.22** (2.00)	0.22** (2.04)	0.50*** (4.65)	0.50*** (4.83)	0.50*** (4.84)	0.51*** (4.99)	0.51*** (4.87)
MSE	5.712	5.765	5.651	5.754	5.650	0.002	0.002	0.002	0.002	0.002
QLIKE	0.147	0.144	0.138	0.139	0.137	0.222	0.222	0.206	0.212	0.210
MZ-R ²	0.360	0.361	0.365	0.361	0.366	0.527	0.527	0.533	0.526	0.538

予測力に関しても日経平均先物の場合と同様、イン・サンプルに対する検証では、どの RV に対してもニュース指数の回帰係数は有意にならないものの、アウト・オブ・サンプルに対する検証では、QLIKE と MZ-R² では $RV_{t:t+21}$ に対する予測力が向上している。そして、日経平均先物および長期国債先物の $RV_{t:t+21}$ に対する予測力は、*News Index_{future}* を追加したときに向上する傾向がある。

これらの結果は、全ての期間ではなく特定の期間においてニュース指数が予測力向上に寄与している可能性を示唆している。そこで、HAR モデルにニュース指数を追加したときに、どの期間に予測力の向上が観察されるかについて、予測

誤差 (MSE、QLIKE) の時系列推移をもとに考察する。ここでは、特に、予測力の向上が観測された $RV_{t,t+21}$ に対する予測力に焦点をあて、ベースラインとなる HAR モデルの予測誤差と、HAR モデルに 2 つのニュース指数 ($News Index_{current}$ と $News Index_{future}$) を追加したモデルの予測誤差との比較を行う。より具体的には、モデル間の予測誤差の差を時点 t まで累積した値 ($diff_MSE$ 、 $diff_QLIKE$) によって予測精度を比較する。 t 期までの $diff_MSE_t$ と $diff_QLIKE_t$ は (13) 式と (14) 式のように定義される。

$$diff_MSE_t := \sum_{j=1}^t (L_j^{MSE,HAR} - L_j^{MSE,HAR+News Index}), \quad (13)$$

$$diff_QLIKE_t := \sum_{j=1}^t (L_j^{QLIKE,HAR} - L_j^{QLIKE,HAR+News Index}). \quad (14)$$

$L_j^{MSE,HAR}$ と $L_j^{QLIKE,HAR}$ は、HAR モデルで予測した際の損失関数であり、 $L_j^{MSE,HAR+News Index}$ と $L_j^{QLIKE,HAR+News Index}$ は、HAR モデルにニュース指数を追加したモデルで予測した際の損失関数である。

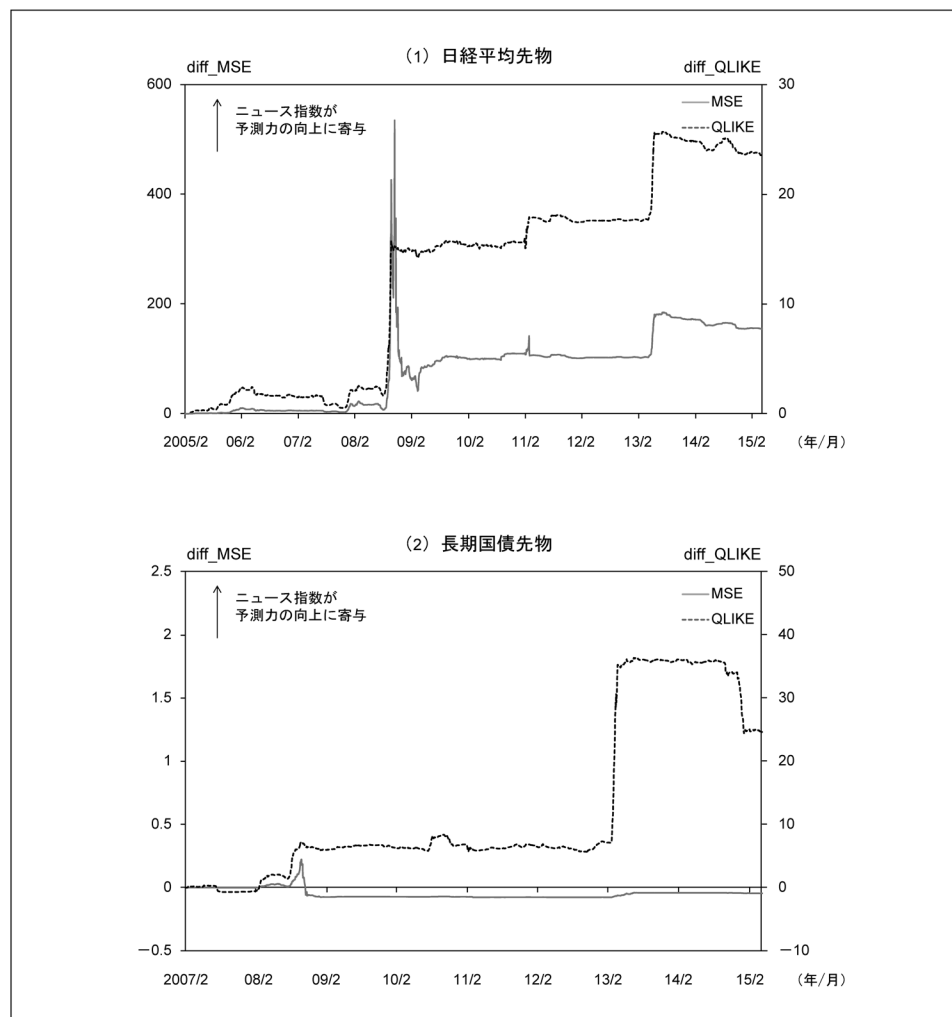
図 5 は、 $diff_MSE$ と $diff_QLIKE$ の推移を図示したものであり、正の方向に大きいほど、ニュース指数が予測力の向上に寄与していることを示している。分析結果から、通常時では、ニュース指数は予測力の向上に寄与しない傾向にあるが、リーマン・ショックや東日本大震災、量的質的金融緩和導入などのイベントが発生した時に、予測力の向上に寄与する傾向にあることがわかる。

ここまで、RV の系列相関のみを考慮した HAR モデルで分析を行ってきたが、他にも RV の予測力を高めるような諸要因が報告されており、RV を予測するうえで、それらの要因がニュース指数と同一の効果を反映している可能性がある。そこで、ジャンプ拡散過程のもとでの RV モデルの 1 つである Corsi, Pirino, and Reno [2010] の HAR-TCJ モデルをベースに、レバレッジ効果とインプライド・ボラティリティを加味した Leverage HAR-TCJ-VIX (LHAR-TCJ-VIX) モデルをベースラインとして追加検証を行う。

資産価格変動のジャンプは、市場に流入したニュースと密接な関係があることが広く知られており、ボラティリティを予測するうえで、構築したニュース指数がジャンプ項と同様の働きをする可能性がある。そこで、まず、HAR モデルをジャンプ拡散過程のもとでの RV モデルである Corsi, Pirino, and Reno [2010] の HAR-TCJ モデルに拡張する。

また、株式市場では株価が上がった日の翌日よりも下がった日の翌日の方がボラティリティが上昇する傾向があることが知られており、この非対称性はレバレッジ効果と呼ばれている。前節の結果では日経平均先物に関して、ニュース指数の回帰係数は総じて負の値を取っており、これはレバレッジ効果を反映している可能性が

図5 月次のRV ($RV_{t:t+21}$) に対する予測誤差の比較分析



ある。そこで、Corsi and Reno [2012] に倣い、ベースライン・モデルにレバレッジ効果も加味する。

さらに、Bekaert and Hoerova [2014] では、オプション価格から算出されるインプライド・ボラティリティが将来の市場のボラティリティを反映していることから、HAR モデルにインプライド・ボラティリティを追加することで予測力が向上することを報告している。ニュース指数が将来の不確実性を捉えていたとしても、ボラティリティを予測するうえでは、インプライド・ボラティリティと同じ効果しか持たない可能性もある。そこで、Bekaert and Hoerova [2014] に倣い、VXJ の 2 乗の対数値 ($\ln VXJ_t^2$) をベースラインに追加する。以上を勧案した LHAR-TCJ-VIX モデ

表 14 LHAR-TCJ-VIX モデルをベースラインとした予測力分析

予測値	モデル	MSE	QLIKE	MZ- R^2
RV_t	LHAR-TCJ-VIX	28.406	0.120	0.181
	+ <i>News Index</i> _{current}	26.380	0.119	0.194
	+ <i>News Index</i> _{future}	27.364	0.120	0.187
	+ <i>News Index</i> _{both}	26.941	0.120	0.190
	+ <i>News Index</i> _{current} + <i>News Index</i> _{future}	26.338	0.120	0.194
$RV_{t:t+4}$	LHAR-TCJ-VIX	11.421	0.118	0.336
	+ <i>News Index</i> _{current}	11.301	0.118	0.339
	+ <i>News Index</i> _{future}	10.432	0.117	0.358
	+ <i>News Index</i> _{both}	10.739	0.118	0.352
	+ <i>News Index</i> _{current} + <i>News Index</i> _{future}	10.586	0.117	0.354
$RV_{t:t+21}$	LHAR-TCJ-VIX	6.530	0.136	0.364
	+ <i>News Index</i> _{current}	5.860	0.135	0.393
	+ <i>News Index</i> _{future}	5.414	0.132	0.410
	+ <i>News Index</i> _{both}	5.465	0.133	0.409
	+ <i>News Index</i> _{current} + <i>News Index</i> _{future}	5.393	0.131	0.412

ルの詳細に関しては、補論 4 を参照されたい。

HAR モデルをベースラインとした分析と同様に、ここでは、LHAR-TCJ-VIX モデルにニュース指数を追加した場合に、ニュース指数がボラティリティ予測に寄与するかが焦点となる。追加検証では日経平均先物のみを取り扱う。表 14 は、LHAR-TCJ-VIX モデルにニュース指数を加えたモデルを推定し、アウト・オブ・サンプルに対する予測力の結果をまとめたものである。ニュース指数を追加すると、MSE と QLIKE は小さくなり、MZ- R^2 は大きくなる傾向がある。特に、 $RV_{t:t+21}$ に対する予測力は、 RV_t や $RV_{t:t+4}$ と比べて、より向上していることがわかる。また、HAR モデルと LHAR-TCJ-VIX モデルを比較すると、HAR モデルの方がパフォーマンスがよい傾向があるものの、ニュース指数を追加した場合、LHAR-TCJ-VIX モデルをベースラインとしたモデルの方が $RV_{t:t+21}$ に対する予測力が高まる傾向にあることがわかる。これらの結果は、ボラティリティを予測するうえで、ニュース指数はジャンプ項やレバレッジ効果、インプライド・ボラティリティとは異なる情報を持っており、RV 予測力の向上に寄与する可能性があることを示している。

5. おわりに

本研究では、日々配信されている経済ニュースを CNN を用いた自然言語処理に

よって指数化し、景気動向のナウキャストリングを試みた。構築したニュース指数がナウキャストリングに有用であることを確認したうえで、ニュース指数が日経平均先物および長期国債先物のボラティリティ予測に有用であるか分析を行った。景気動向との関係を分析した結果、構築したニュース指数は景気動向を表す経済指標と類似した推移をしており、特に不確実性の高い時期において、景気動向のナウキャストリングに資する情報のうち、資産価格が有していないものを同ニュース指数が有している可能性が示唆された。また、HARモデルに基づいたボラティリティ予測モデルにニュース指数を加えた場合にボラティリティの予測力が向上するか検討した結果、リーマン・ショックや東日本大震災といった大きなイベントがあった場合に、本研究で構築したニュース指数が予測力の向上に役立つ可能性が示された。さらに、訓練データに利用した景気ウォッチャー調査の景気判断理由集のうち、現状判断から学習したモデルから構築したニュース指数と先行き判断から学習したモデルから構築したニュース指数では、予測やナウキャストリングに有用である場面が異なることが示された。

これらの結果はロイター・ニュースの特徴に依存する可能性もあるため、他の媒体との比較分析は今後の課題である。また、本研究では、ニュース指数の構築の際に景気ウォッチャー調査の景気判断理由集を利用して過去のニュース記事の分類を行っていることから、間接的に将来情報を利用している可能性があることに留意が必要である。トピック・モデルやタグ情報の利用による記事内容のトピック別の分析や、本研究で扱わなかった金融資産の分析への応用が今後の課題として挙げられる。

参考文献

- 坪井祐太・海野裕也・鈴木 潤、『深層学習による自然言語処理』、機械学習プロフェッショナルシリーズ、講談社、2017年
- 東山昌彦・乾 健太郎・松本裕治、「述語の選択選好性に着目した名詞評価極性の獲得」、『言語処理学会第14回年次大会発表論文集』、2008年、584～587頁
- 山本裕樹・松尾 豊、「景気ウォッチャー調査の深層学習を用いた金融レポートの指数化」、『第30回人工知能学会全国大会論文集』、2016年
- 渡部敏明、「Realized Volatility—サーベイと日本株式市場への応用—」、『経済研究』第58巻4号、2007年、352～373頁
- 渡辺広太・渡辺 努、「スキャナーデータを用いた日次物価指数の計測」、CARF Working Paper Series No. CARF-J-094、2013年
- Andersen, Torben G., and Tim Bollerslev, “Modeling and Forecasting Realized Volatility,” *Econometrica*, 71, 2003, pp. 579–625.
- , ———, and Francis X. Diebold, “Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility,” *Review of Economics and Statistics*, 89, 2007, pp. 701–720.
- Antweiler, Werner, and Murray Z. Frank, “Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards,” *Journal of Finance*, 59(3), 2004, pp. 1259–1294.
- Baker, Scott R., Nicholas Bloom, and Steven J. Davis, “Measuring Economic Policy Uncertainty,” *Quarterly Journal of Economics*, 131(4), 2016, pp. 1593–1636.
- Barndorff-Nielsen, Ole E., and Neil Shephard, “Power and Bipower Variation with Stochastic Volatility and Jumps,” *Journal of Financial Econometrics*, 2(1), 2004, pp. 1–37.
- Bekaert, Geert, and Marie Hoerova, “The VIX, the Variance Premium and Stock Market Volatility,” *Journal of Econometrics*, 183(2), 2014, pp. 181–192.
- Campbell, John Y., and Ludger Hentschel, “No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns,” *Journal of Financial Economics*, 31(3), 1992, pp. 281–318.
- Chen, Hailiang, Prabuddha De, Yu Hu, and Byoung-Hyoun Hwang, “Wisdom of Crowds: The Value of Stock Opinions Transmitted through Social Media,” *Review of Financial Studies*, 27(5), 2014, pp. 1367–1403.
- Corsi, Fulvio, “A Simple Approximate Long-Memory Model of Realized Volatility,” *Journal of Financial Econometrics*, 7(2), 2009, pp. 174–196.
- , David Pirino, and Roberto Reno, “Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting,” *Journal of Econometrics*, 159(2), 2010, pp. 276–288.
- , and Roberto Reno, “Discrete-Time Volatility Forecasting with Persistent Lever-

- age Effect and the Link with Continuous-Time Volatility Modeling,” *Journal of Business & Economic Statistics*, 30(3), 2012, pp. 368–380.
- Engle, Robert F., and Victor K. Ng, “Measuring and Testing the Impact of News on Volatility,” *Journal of Finance*, 48(5), 1993, pp. 1749–1778.
- Garcia, Diego, “Sentiment during Recessions,” *Journal of Finance*, 68(3), 2013, pp. 1267–1300.
- Glorot, Xavier, and Yoshua Bengio, “Understanding the Difficulty of Training Deep Feed-forward Neural Networks,” in *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, 2010, pp. 249–256.
- Goodfellow, Ian, Yoshua Bengio, and Aaron Courville, *Deep Learning*, MIT Press, 2016.
- Hansen, Peter R., and Asger Lunde, “A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?” *Journal of Applied Econometrics*, 20(7), 2005, pp. 873–889.
- Hansen, Stephen, and Michael McMahon, “Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication,” *Journal of International Economics*, 99(1), 2016, pp. S114–S133.
- , ———, and Andrea Prat, “Transparency and Deliberation within the FOMC: A Computational Linguistics Approach,” *Quarterly Journal of Economics*, 133(2), 2018, pp. 801–870.
- He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,” in *Proceedings of the 2015 IEEE International Conference on Computer Vision*, 2015, pp. 1026–1034.
- Kim, Yoon, “Convolutional Neural Networks for Sentence Classification,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 2014, pp. 1746–1751.
- Kingma, Diederick P., and Jimmy Ba, “Adam: A Method for Stochastic Optimization,” *arXiv: 1412.6980*, 2014.
- Li, Feng, “Annual Report Readability, Current Earnings, and Earnings Persistence,” *Journal of Accounting and Economics*, 45(2–3), 2008, pp. 221–247.
- Lin, Zhouhan, Minwei Feng, Cícero Nogueira dos Santos, Mo Yu, Bing Xiang, Bowen Zhou, and Yoshua Bengio, “A Structured Self-attentive Sentence Embedding,” *arXiv: 1703.03130*, 2017.
- Liu, Lily Y., Andrew J. Patton, and Kevin Sheppard, “Does Anything Beat 5-minute RV? A Comparison of Realized Measures across Multiple Asset Classes,” *Journal of Econometrics*, 187(1), 2015, pp. 293–311.
- Loughran, Tim, and Bill McDonald, “When Is a Liability Not a Liability? Textual Analysis,

- Dictionaries, and 10-Ks,” *Journal of Finance*, 66(1), 2011, pp. 35–65.
- , and ———, “IPO First-day Returns, Offer Price Revisions, Volatility, and Form S-1 Language,” *Journal of Financial Economics*, 109(2), 2013, pp. 307–326.
- Manela, Asaf, and Alan Moreila, “News Implied Volatility and Disaster Concerns,” *Journal of Financial Economics*, 123(1), 2017, pp. 137–162.
- Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean, “Distributed Representations of Words and Phrases and Their Compositionality,” in *Proceedings of the 26th International Conference on Neural Information Processing Systems*, 2013, pp. 3111–3119.
- Patton, Andrew J., “Volatility Forecast Comparison Using Imperfect Volatility Proxies,” *Journal of Econometrics*, 160(1), 2011, pp. 246–256.
- Socher, Richard, Alex Perelygin, Jean Wu, Jason Chuang, Christopher Manning, Andrew Ng, and Christopher Potts, “Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank,” in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013, pp. 1631–1642.
- Tetlock, Paul C., “Giving Content to Investor Sentiment: The Role of Media in the Stock Market,” *Journal of Finance*, 62(3), 2007, pp. 1139–1168.
- , Maytal Saar-Tsechansky, and Sofus Macskassy, “More Than Words: Quantifying Language to Measure Firms’ Fundamentals,” *Journal of Finance*, 63(3), 2008, pp. 1437–1467.

補論 1. CNN によるテキスト分類

ここでは、本研究で採用した Kim [2014] の CNN に基づくテキスト分類モデルによる景気判断理由集からの学習と、ロイター・ニュースへのスコアおよび確率値の付与について記述する。CNN はニューラル・ネットワーク・モデルの 1 つであり、入力データに対して、フィルタ（もしくはカーネル）と呼ばれる、入力データよりもサイズの小さい重みテンソルを移動（ストライド）させながら積和演算することで、局所的な情報を集約するアルゴリズムである。CNN に関する詳細は、Goodfellow, Bengio, and Courville [2016] や坪井・海野・鈴木 [2017] を参照されたい。Kim [2014] では、「CNN-rand」、「CNN-static」、「CNN-non-static」、「CNN-multichannel」の 4 種類のモデルを考案し、比較実験を行っている²¹。このうち、本研究では「CNN-non-static」を採用した。

まず、景気ウォッチャー調査の景気判断理由集とロイター・ニュースの各文を行列で表現する。例えば、単語 w_a 、単語 w_b 、単語 w_c の順番に出てきた文 S を考える。 S の行列 \mathbf{M}_S は、 w_a 、 w_b 、 w_c の単語ベクトル \mathbf{x}_{w_a} 、 \mathbf{x}_{w_b} 、 \mathbf{x}_{w_c} を文中に出てきた順番に行方向に並べることで表現される。各単語ベクトル \mathbf{x}_w は one-hot ベクトルで表現される。one-hot ベクトルとは、語彙 V 中の各単語に 1 から $|V|$ までの単語番号が割り振られているとき、与えられた文の i 番目に出現した単語の単語番号が n であった場合 n 番目の要素を 1 とし、それ以外の要素は 0 と表現されるベクトルである。このとき、各文に含まれる単語数は異なるので、景気判断理由集およびロイター・ニュースの中の文に含まれる最大の単語数に行列の行数を合わせ、足りない部分はパディングする²²。景気判断理由集とロイター・ニュースの各文の形態素解析による分かち書きには、MeCab を用いている²³。

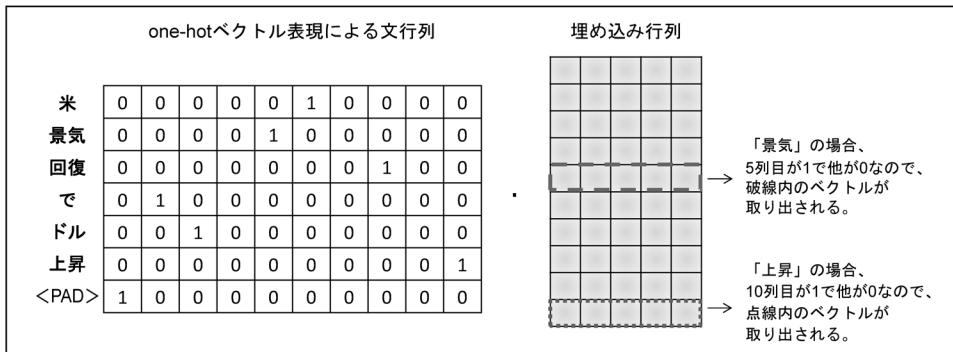
次に、景気判断理由集を入力として、埋め込み層、畳み込み層、最大値プーリング層、全結合層を経て、5 段階のスコアあるいはバイナリ値を出力するモデルを学習する。埋め込み層では、各単語を one-hot ベクトルで表現した文行列を 300 次元の埋め込みベクトルに変換する。埋め込み層における演算は (A-1) 式のように表される。

.....
21 各モデルの違いは埋め込み層にある。埋め込み層の重み行列の初期値に、「CNN-rand」は乱数を、「CNN-non-static」は事前学習した単語の分散表現ベクトルを、それぞれ利用するモデルである。一方で、「CNN-static」は重み行列に事前学習した単語の分散表現ベクトルは利用するが、重み行列を勾配法で更新しないモデルである。そして、「CNN-multichannel」は「CNN-static」と「CNN-non-static」の両方をあわせたモデルである。

22 パディングとは、データを固定長の数字列や行列に変換する際に、不足する部分を同一特殊文字や数字を追加してサイズを合わせることである。

23 MeCab: <https://taku910.github.io/mecab/> (2019 年 1 月 10 日)。システム辞書には IPA 辞書を用いた。加えて、ユーザ辞書として Wikipedia のタイトル・リストを追加している。

図 A-1 埋め込み層における演算例



$$\mathbf{x}_w \mathbf{E} = \bar{\mathbf{x}}_w. \tag{A-1}$$

ここで、 \mathbf{x}_w は単語 w の one-hot ベクトル、 \mathbf{E} は埋め込み行列、 $\bar{\mathbf{x}}_w$ は単語 w の変換後の埋め込みベクトルを表す。図 A-1 は、各単語を 10 次元の one-hot ベクトルで表現した文行列を、埋め込み層によって 5 次元の埋め込みベクトルへと変換する演算例である。図中の <PAD> は、パディング時の特殊文字を表す。

埋め込み層の重み行列の初期値には、日本語版 Wikipedia から獲得した単語の分散表現ベクトルを用いる²⁴。具体的には、あらかじめ日本語版 Wikipedia の本文全文から word2vec を用いて、Wikipedia に出現する単語を 300 次元のベクトルで表現する (Mikolov *et al.* [2013])²⁵。Wikipedia の記事本文の抽出には WikiExtractor を²⁶、単語分割には MeCab を、word2vec のモデルには skip-gram をそれぞれ用いている²⁷。ただし、Wikipedia に出現しない単語の重みの初期値に関しては、区間 $[-0.25, 0.25]$ の連続一様分布からサンプリングした値を用いる。

畳み込み層では、埋め込み層で 300 次元に変換された文行列に対して、フィルタ・サイズの列数を 300 次元に固定した 1 次元の畳み込み演算を行う。このとき、フィルタ・サイズは「 3×300 」、「 4×300 」、「 5×300 」の 3 種類とし、フィルタ数は 100、活性化関数は relu (rectified linear unit) 関数²⁸、パラメータの初期化には He の初期値を用いる (He *et al.* [2015])²⁹。

24 日本語版 Wikipedia: <https://dumps.wikimedia.org/jawiki/> (2019 年 1 月 10 日)。

25 word2vec: <https://code.google.com/archive/p/word2vec/> (2019 年 1 月 10 日)。

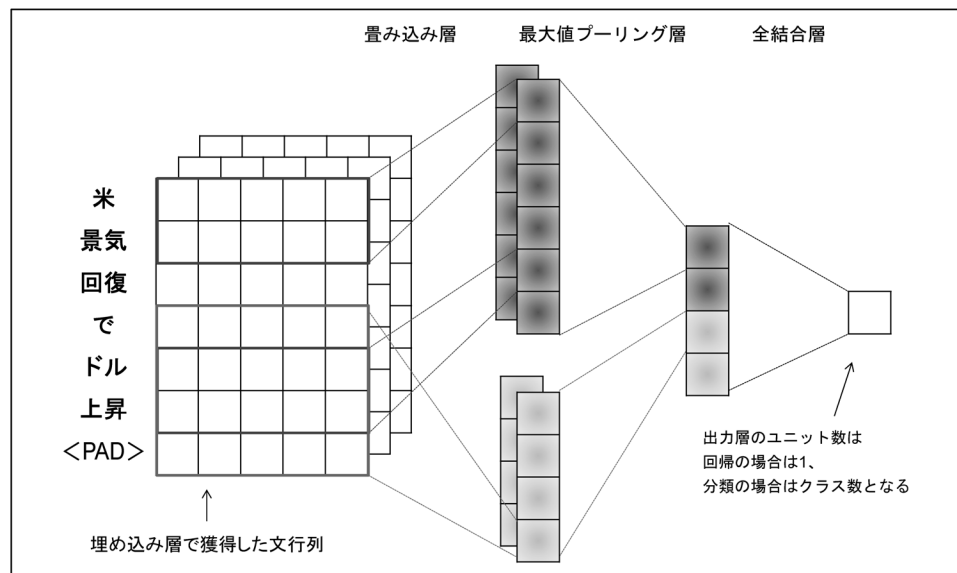
26 WikiExtractor: <https://github.com/attardi/wikiextractor> (2019 年 1 月 10 日)。

27 学習の際のパラメータは、標準設定とした。

28 $\text{relu}(x)$ 関数は $\max(0, x)$ と定義される。

29 He の初期値は、 $w_{he} \sim N\left(0, \sqrt{\frac{2}{fan_{in}}}\right)$ からサンプリングされる初期値であり、 fan_{in} は前層からの入力のユニット数を表す。

図 A-2 CNN に基づくテキスト分類モデルの簡略図



そして、畳み込み層からの出力を、最大値プーリング層への入力とする。最大値プーリングは、与えられた範囲のうち、最大値を取るような演算である。最大値プーリングによって、局所的に素性（単語）が入れ替わっていても出力が不変となる。最大値プーリング層からの出力を連結し、全結合層（出力層）への入力とする。全結合層では、L2 正則化と³⁰、ドロップアウトを採用し³¹、それぞれ正則化パラメータは 3.0、ドロップアウト率は 0.5 とする。そして、その他のパラメータの初期化には Glorot の初期値を用いる（Glorot and Bengio [2010]）³²。活性化関数に関しては、回帰問題のときは恒等関数を、分類問題のときにはソフトマックス関数を用いることができる³³。図 A-2 は、CNN に基づくテキスト分類モデルの簡略図である。

モデル・パラメータの最適化手法には Adam を使用し³⁴、ミニバッチのサイズは

30 L2 正則化とは、損失関数にパラメータの値が大きくなることに罰則を与える正則化項を追加する方法であり、損失関数は $E(\mathbf{w}) = E(\mathbf{w})^* + \frac{\lambda}{2} \|\mathbf{w}\|^2$ と表される。

31 ドロップアウトとは、学習時にノードの一部を 0 とすることで、特定のノードだけを重視した学習を防ぐ方法論であり、学習中の微分計算の度に、無作為にノードを抽出することで、異なる部分構造を評価しパラメータを更新している。

32 Glorot の初期値は、 $w_{glorot} \sim N\left(0, \sqrt{\frac{2}{fan_{in} + fan_{out}}}\right)$ からサンプリングされる初期値であり、 fan_{in} は前層からの入力ユニット数、 fan_{out} は次の層への出力ユニット数を表す。

33 恒等関数： $linear(x) = x$ 、ソフトマックス関数： $softmax(x_k) = \frac{\exp(x_k)}{\sum_{j=1}^K \exp(x_j)}$ 。K はクラス数、k は k 番目のクラスを表す。2 クラス分類の場合はシグモイド関数が利用できる。その際は出力ユニット数を 1 とし、シグモイド関数によって確率値を出力する。

34 Adam は確率的勾配法を拡張した最適化手法の 1 つであり、以下のように勾配の 1 次および 2 次

50、エポック数は25とした。そして、各エポックで訓練データをランダムにシャッフルしたのち、ミニバッチを抽出している。このとき、訓練データのうち、ランダムに抽出した10%を検証データとして、検証データに対する平均二乗誤差あるいは交差エントロピーが最小となるエポックのモデル・パラメータを採用している。これらモデルの実装には、PythonライブラリのKerasを用いた³⁵。

最後に、訓練したモデルに対して行列で表現されたロイター・ニュースの各文を入力することで、スコアおよび確率値を付与する。

モーメントを利用して、モデルのパラメータ w のそれぞれを以下のように更新する (Kingma and Ba [2014])。

$$w_{t+1} = w_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}},$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t},$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \frac{\partial E(w_t)}{\partial w_t}, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) \left(\frac{\partial E(w_t)}{\partial w_t} \right)^2.$$

ここで、 t はエポック、 ϵ はゼロによる除算を防ぐための値、 α は学習率、 β_1 と β_2 はハイパー・パラメータである。

35 Keras: <https://keras.io/> (2019年1月10日)。

補論 2. 辞書に基づくニュース指数との比較

ここでは、主要な代替手法の1つである辞書に基づくニュース指数を用いて、3節の分析を行った結果を示す。辞書によるアプローチは、事前に定義した辞書に含まれる単語の出現頻度をもとに、テキストの内容を判断する方法であり、ファイナンス研究や会計研究、経済研究に広く用いられる。ここでは、東山・乾・松本 [2008] の日本語評価極性辞書（名詞編）を利用して³⁶、(A-2) 式のように辞書に収録されているポジティブな単語の数とネガティブな単語の数をカウントして、ニュース指数を構築する。

$$News\ Index_{dic,t} = \frac{Positive\ Words_t - Negative\ Words_t}{Positive\ Words_t + Negative\ Words_t} \quad (A-2)$$

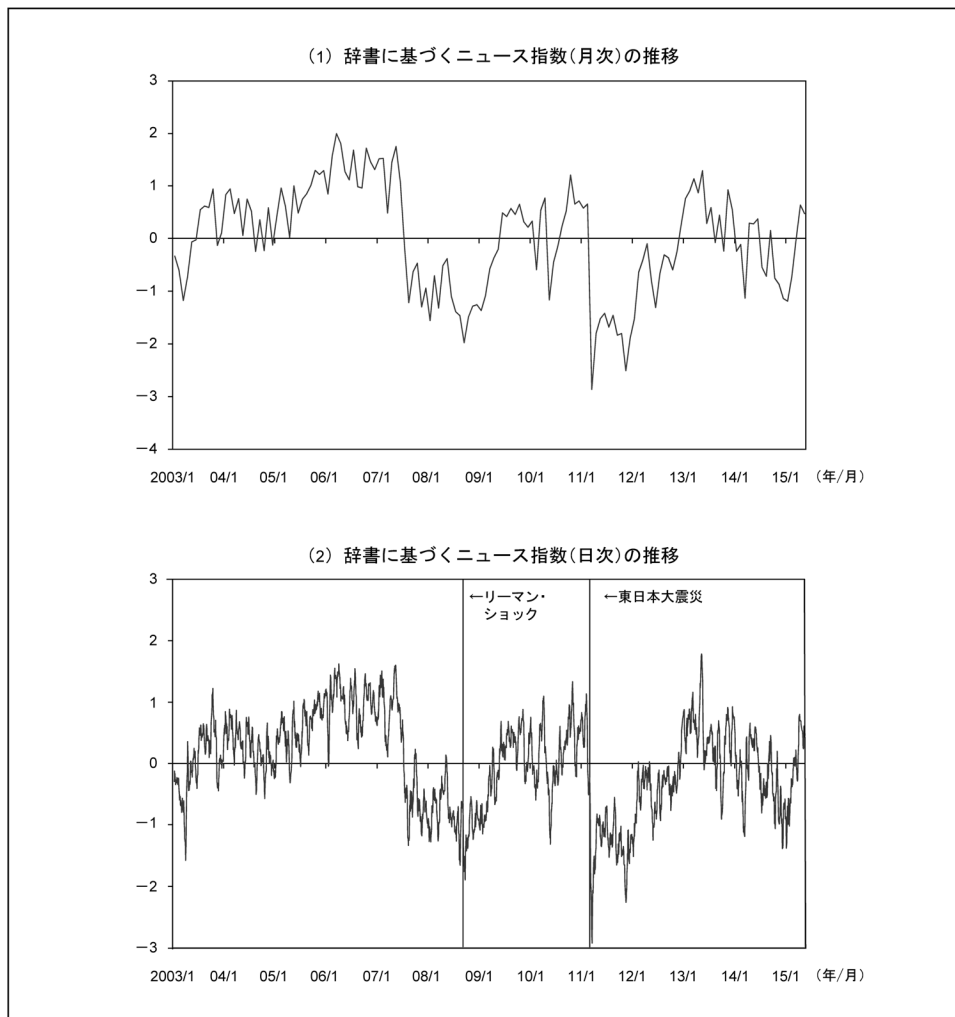
ここで、 $Positive\ Words_t$ と $Negative\ Words_t$ はそれぞれ、 t 月あるいは t 日に配信されたニュース記事の文中に出現したポジティブな単語の数とネガティブな単語の数を表す。MeCab による形態素解析でニュース文を分かち書きする際には、日本語評価極性辞書（名詞編）に収録されている全単語をユーザ辞書として追加している。表 A-1 と図 A-3 はそれぞれ、 $News\ Index_{dic}$ の記述統計量と、その推移を図示したものである。本研究で構築したニュース指数と同様に、リーマン・ショックや東日本大震災などの特徴的なイベントがあった際には大きく下がっていることがわかる。また、表 A-2 は本研究で構築したニュース指数との相関係数をまとめたものである。全体的に相関係数が大きく、特に、 $News\ Index_{dic}$ は、 $News\ Index_{future}$ と最も類似した推移を示している。しかしながら、景気動向の変化率との相関をまとめた表 A-3 をみると、相関係数は本研究で構築したニュース指数と比べて小さいことがわかる。さらに、(4) 式に $News\ Index_{dic,t}$ を追加して推計したところ、 $News\ Index_{dic,t}$ の偏回帰係数は有意にならないことを本研究では確認している。これらの結果は、ニュース・テキスト・データから景気動向を捉えるうえで、景気ウォッチャー調査の判断理由集を訓練データとして利用することが役立つことを示唆している。

.....
36 日本語評価極性辞書：<https://www.cl.ecei.tohoku.ac.jp/index.php?Open%20Resources%2FJapanese%20Sentiment%20Polarity%20Dictionary>（2019年1月10日）。

表 A-1 辞書に基づくニュース指数の記述統計量

	ニュース指数 (日次)	ニュース指数 (月次)
平均値	0.32	0.32
中央値	0.32	0.33
最小値	-0.04	0.14
最大値	0.55	0.44
標準偏差	0.09	0.06
観測数	3,039	149

図 A-3 辞書に基づくニュース指数 (標準化後) の推移



備考：辞書に基づくニュース指数（日次）は、標準化されたニュース指数の後方移動平均（10 営業日）によって平滑化した値を描画している。

表 A-2 ニュース指数間の相関係数

	<i>News Index_{current}</i>	<i>News Index_{future}</i>	<i>News Index_{both}</i>
<i>News Index_{dic}</i> (日次)	0.42	0.64	0.64
<i>News Index_{dic}</i> (月次)	0.54	0.80	0.75

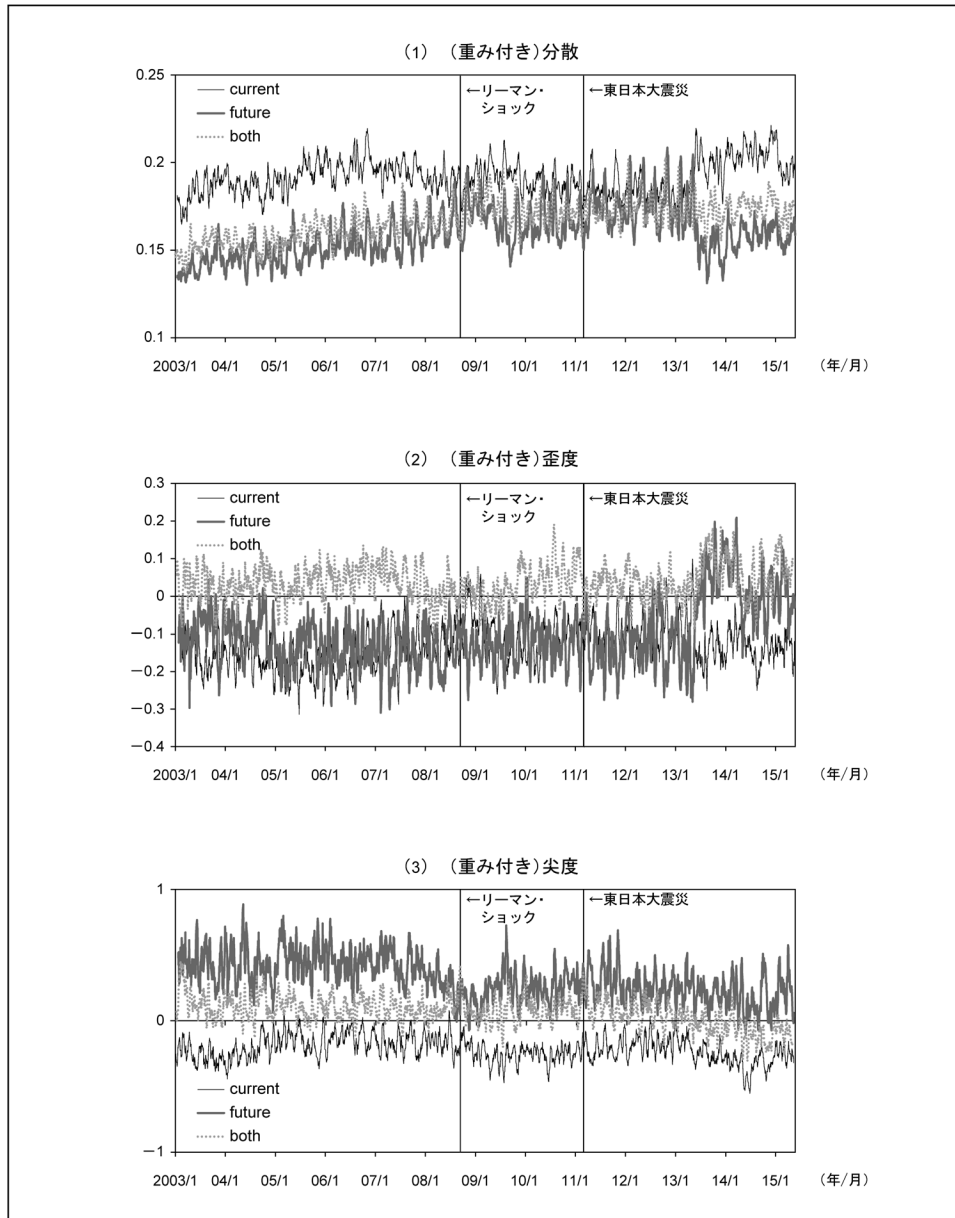
表 A-3 ニュース指数（月次）と景気動向の変化率との相関係数

	先行指数	一致指数	遅行指数	鉱工業生産指数
<i>News Index_{dic}</i>	0.37	0.34	0.28	0.28
<i>News Index_{current}</i>	0.49	0.48	0.36	0.35
<i>News Index_{future}</i>	0.50	0.46	0.41	0.36
<i>News Index_{both}</i>	0.50	0.47	0.40	0.37

補論 3. スコアの分散・歪度・尖度

図 A-4 は、営業日ごとにニュース記事の各文に付与されたスコアの（重み付き）

図 A-4 各営業日ごとのスコアの分散・歪度・尖度



分散、(重み付き)歪度、(重み付き)尖度を集計して推移を図示したものである。視認しやすいように、後方移動平均(10営業日)によって平滑化している。分散と歪度に関しては、2013年以降、推移の傾向が変化していることがわかる。主に先行き判断から算出された分散と歪度が大きく変化していることがわかる。一方で、尖度は分析期間においては、安定した推移を示している。しかしながら、これら高次のモーメントは平均値(*News Index*)に比べ、CNNでの学習によるばらつきが大きくなるため、結果の解釈には注意が必要である。

補論 4. LHAR-TCJ-VIX モデル

ここでは、本研究で取り扱った LHAR-TCJ-VIX モデルについて概説する。LHAR-TCJ-VIX モデルは、ジャンプ拡散過程のもとでの RV モデルの 1 つである Corsi, Pirino, and Reno [2010] の HAR-TCJ モデルをベースに、レバレッジ効果とインプライド・ボラティリティを加味したモデルである。Corsi, Pirino, and Reno [2010] は、資産価格変動がジャンプ拡散過程に従うと仮定して、 RV_t をジャンプ項に起因する変動を取り除いたボラティリティ TC_t とジャンプ項 TJ_t に分解して、本文 (8) 式の HAR モデルを (A-3) 式のように拡張している。

$$RV_{t:t+h-1} = \alpha + \beta_d TC_{t-1} + \beta_w TC_{t-5:t-1} + \beta_m TC_{t-22:t-1} + \beta_{TJ} TJ_{t-1} + \epsilon_t. \quad (\text{A-3})$$

(A-3) 式の TC_t と TJ_t は、TBPV (Threshold Bipower Variation) に基づく変数であり、(A-4) 式と (A-5) 式によって算出される。

$$TJ_t = I_{\{C-Tz > \Phi_\alpha\}} \cdot (RV_t - TBPV_t)^+. \quad (\text{A-4})$$

$$TC_t = RV_t - TJ_t. \quad (\text{A-5})$$

ここで、 $I_{\{\cdot\}}$ は指示関数であり、 $\{\cdot\}$ の条件を満たす時は 1、そうでないときは 0 を返す関数である。また、 $C-Tz$ はジャンプを検出するための統計量であり、ジャンプが発生しないときには、漸近的に標準正規分布 $N(0, 1)$ に従う統計量である。 Φ_α は標準正規分布の累積分布関数の逆関数 (α 分位点) を表しており、本研究ではジャンプを検出するための有意水準を $\alpha = 99.9\%$ ($\Phi_{99.9\%} \approx 3.09$) とした。

Corsi, Pirino, and Reno [2010] では、通常の BPV (Bipower Variation) よりも、TBPV によって推定されたボラティリティの方が真のボラティリティからの乖離がより小さくなるほか、ジャンプの検出力が高いことを明らかにしている³⁷。 t 期の TBPV は、(A-6) 式によって算出される。

$$TBPV_t = \mu_1^{-2} \sum_{j=2}^{\lfloor T/\delta \rfloor} |r_{j-1}| |r_j| I_{\{|r_{j-1}|^2 \leq \vartheta_{j-1}\}} I_{\{|r_j|^2 \leq \vartheta_j\}}. \quad (\text{A-6})$$

ここで、 T は期間、 δ は日中リターンを観測する間隔であり、日中リターンの観測数 N との関係は $N = T/\delta$ となる。本研究では、 T を 1 日、 δ を 5 分としている。また、 $\mu_1 = \sqrt{\frac{2}{\pi}} \approx 0.7979$ である。 ϑ は閾値であり、(A-7) 式によって算出される。

$$\vartheta_j = c_\vartheta^2 \cdot \hat{V}_j. \quad (\text{A-7})$$

.....
37 BPV や BPV に基づく HAR モデルに関しては、Barndorff-Nielsen and Shephard [2004] や Andersen, Bollerslev, and Diebold [2007] を参照されたい。

c_θ は閾値の大きさを決めるパラメータであり、 \hat{V}_j は局所分散 (Local Variance) の推定値である³⁸。Corsi, Pirino, and Reno [2010] では、 $c_\theta \geq 3$ のとき、 c_θ に対する TBPV の感度が小さくなることをシミュレーションによって明らかにしており、本研究では先行研究に倣い $c_\theta = 3$ とした。ジャンプの検出をするための統計量 C-Tz は、(A-8) 式によって算出される。

$$C-Tz = \delta^{-\frac{1}{2}} \frac{(RV_t - C-TBPV_t) \cdot RV_t^{-1}}{\sqrt{(\frac{\pi^2}{4} + \pi - 5) \max\left\{1, \frac{C-TTriPV_t}{(C-TBPV_t)^2}\right\}}} \quad (A-8)$$

$C-TBPV_t$ と $C-TTriPV_t$ はそれぞれ、Corrected Threshold Bipower Variation と Corrected Threshold Tripower Variation の略称であり、(A-9)~(A-11) 式によって算出される。

$$C-TBPV_t = \mu_1^{-2} \sum_{j=2}^{[T/\delta]} Z_1(r_j, \vartheta_j) Z_1(r_{j-1}, \vartheta_{j-1}). \quad (A-9)$$

$$C-TTriPV_t = \mu_{4/3}^{-3} \frac{1}{\delta} \sum_{j=3}^{[T/\delta]} Z_{4/3}(r_j, \vartheta_j) Z_{4/3}(r_{j-1}, \vartheta_{j-1}) Z_{4/3}(r_{j-2}, \vartheta_{j-2}). \quad (A-10)$$

$$Z_\gamma(x, y) = \begin{cases} |x|^\gamma & \text{if } x^2 \leq y, \\ \frac{1}{2N(-c_\theta)\sqrt{\pi}} \left(\frac{2}{c_\theta} y\right)^{\frac{\gamma}{2}} \Gamma\left(\frac{\gamma+1}{2}\right) \cdot \left(\frac{c_\theta^2}{2}\right) & \text{if } x^2 > y. \end{cases} \quad (A-11)$$

ここで、 $N(x)$ は標準正規分布の累積分布関数であり、 $\mu_{4/3} \approx 0.8309$ 、 $c_\theta = 3$ かつ $x^2 > y$ のとき、 $Z_1(x, y) \approx 1.094 \cdot y^{1/2}$ 、 $Z_{4/3}(x, y) \approx 1.129 \cdot y^{2/3}$ となる。

また、Corsi and Reno [2012] では、レバレッジ効果を捉えるために、HAR モデルに (A-12) 式の変数を追加している。

$$R_t^{(h)-} = \min(0, R_t^{(h)}), \quad R_t^{(h)} = \frac{1}{h}(R_t + \dots + R_{t-h+1}). \quad (A-12)$$

ここで、 R_t は日次リターンを表している。過去の日次、週次、月次のレバレッジ効果をそれぞれ取り入れるために、 $R_{t-1}^{(1)-}$ 、 $R_{t-1}^{(5)-}$ 、 $R_{t-1}^{(22)-}$ の3つの変数を追加する。さらに、Bekaert and Hoerova [2014] では、オプション価格から算出されるインプライド・ボラティリティは将来の市場のボラティリティを反映していることから、HAR モデルに VIX を2乗した値を追加することで、モデルの精度が上がることを報告している。そこで、本研究では VXJ の2乗値を追加する。

以上をまとめると、本研究の検証に利用する LHAR-TCJ-VIX モデルは (A-13) 式のように表記される。なお、HAR モデルと同様に、ここでも TC_t 、 TJ_t および VXJ_t^2 を対数変換したモデルを取り扱う。

.....
38 局所分散の算出方法に関しては、Corsi, Pirino, and Reno [2010] の Appendix B を参照されたい。

$$\begin{aligned}\ln RV_{t,t+h-1} &= \alpha + \beta_d \ln TC_{t-1} + \beta_w \ln TC_{t-5:t-1} + \beta_m \ln TC_{t-22:t-1} \\ &\quad + \beta_{TJ} \ln (TJ_{t-1} + 1) + \beta_{Ld} R_{t-1}^{(1)-} + \beta_{Lw} R_{t-1}^{(5)-} + \beta_{Lm} R_{t-1}^{(22)-} \\ &\quad + \beta_{IV} \ln VXJ_t^2 + \epsilon_t.\end{aligned}\tag{A-13}$$

