

論 文

テキスト情報と機械学習を用いた景気動向分析*

新谷 元嗣**

＜要旨＞

本稿では、テキスト情報を利用して、政府統計よりも速報性の高い景気動向指数を作成する方法と日本経済への応用例を概観する。分析手法に関しては、重要な単語の出現頻度に着目する辞書アプローチと、自然言語処理のモデルをテキストデータから学習する機械学習アプローチの2つに分類して整理する。辞書アプローチの中では、古典的なセンチメント分析が、特に計算や経済学的な解釈の容易性の観点から、現在でも十分有用性が高いと考えられる。ただし、その指数の作成過程では、マクロ経済ドメインに特化した極性辞書の利用やテキストデータの慎重な前処理作業が不可欠である。一方で、純粋な予測精度向上の観点からは、文脈を含めたテキスト情報を有効に反映できる機械学習アプローチが望ましい。今後は、新しい言語モデルの景気動向分析への応用が益々増加することが予想される。同時に言語モデルは近年急速に進化し大規模化しているため、モデルが変更された場合の過去系列の遡及推計や指数の継続性は重要な課題である。

JEL Classification Codes : C53、C55、E17

Keywords : 自然言語処理、センチメント分析、ナウキャスト

* 本稿の作成にあたり五島圭一、福永健介、松林洋一の各氏から有益なコメントを頂いた。記して謝意を表したい。

** 新谷 元嗣：東京大学大学院経済学研究科教授。

Business Cycle Analysis Using Textual Information and Machine Learning

By Mototsugu SHINTANI

Abstract

This paper provides an overview of techniques for utilizing textual information to compute business cycle indices available prior to official government statistics, along with their applications to the Japanese economy. In the review, we classify the methods into two categories: the lexicon approach, which focuses on term frequencies of predefined keywords, and the machine learning approach, which utilizes the text data to learn language models. In the lexicon approach, classical sentiment analysis remains highly valuable due to its low computational cost and economic interpretation. However, when computing the index, utilizing a domain-specific polarity dictionary for macroeconomic analysis, and conducting careful preprocessing of text data are essential. On the other hand, in terms of forecast accuracy, machine learning approach that can effectively capture textual information, including context, is preferable. We can anticipate a rising trend in the utilization of new language models for conducting business cycle analysis. At the same time, given the rapid evolution of large language models in recent years, ensuring the continuity of reporting the same index, as well as retrospective estimation when models are replaced, is important.

JEL Classification Codes: C53, C55, E17

Keywords: Natural Language Processing, Nowcast, Sentiment Analysis

1. はじめに

適切な経済政策の運営には、足もとの景気動向を迅速かつ正確に把握することが望ましい。ところが、一国のマクロ経済活動の最も重要な指標である GDP（国内総生産）は、3か月に1度しか観測されない四半期データであり、一次速報値の公表までに約6週間を要する。より速報性を重視した内閣府の景気動向指数の一致指数については、月次データである鉱工業生産指数、有効求人倍率、全産業の営業利益、小売業や卸売業の商業販売額等、合計10個の一致系列を統合することで計算される。また2022年から内閣府で参考公表が始まった景気を把握する新しい指数（一致指数）では、第3次産業活動指数、建設出来高、ソフトウェア投資やサービス輸出など、多くの非製造業関連の月次データが追加されている。ただしこれらの現行指数や新指数についても、対象とする月から1か月以上公表が遅れてしまう。

そこで GDP や景気動向指数の採用系列のような伝統的な数値データに代わり、テキスト情報、位置情報、決済情報、衛星画像情報を含むオルタナティブデータを利用することによって高頻度かつ高精度に足もとの経済活動を把握するナウキャストの試みが近年進められている。特に、新型コロナの感染拡大局面では、位置情報のデータを用いた生産活動のナウキャスト（王他, 2021）やクレジットカード決済情報のデータを用いた消費活動のナウキャスト（ナウキャスト社の JCB 消費 NOW 等）が注目を集めた。本稿では、オルタナティブデータ的一种であるテキストデータを取りあげ、景気動向のナウキャストや将来予測のためのテキスト情報を用いた分析手法の整理を行う¹。

2. 自然言語処理と経済分析

テキスト情報をどのように利用すれば、政府統計よりも速報性の高い景気動向指数を構築することができるだろうか。指数の作成方法を整理するにあたり、経済活動に関して重要な単語の出現頻度に着目する古典的な方法と、自然言語処理のモデルをテキストデータから学習する方法の2つに分類すると便利である。本稿では便宜的に前者を「辞書 (lexicon) アプローチ」、後者を「機械学習 (machine learning) アプローチ」と呼ぶことにする。辞書アプローチでは、あらかじめ設定されたキーワードの出現頻度から、そのキーワードに関連する経済活動の注目度が評価される。このアプローチは計算手続きが単純であり、経済学的な解釈も比較的容易なことが利点としてあげられる。同時に単語の出現頻度のみでは、文脈や否定表現が考慮されない等の問題があり、利用可能な情報がすべて反映されない欠点も持っている。一方、機械学習アプローチではテキスト情報を何らかの手法で数値ベクトル形式に変換したデータを入力とし、分析対象である出力との関係を学習することによ

¹ なお金融市場や金融政策の分野でもテキスト分析の先行研究が多数存在するが、本稿では主に景気動向に関する応用に焦点を当てるため、一部の研究の紹介以外は割愛している。

り、経済活動状況を把握する。このアプローチでは予測に関して高い精度が期待できる利点がある。欠点としては、複雑な言語モデルの学習では計算負荷が大きくなることや、学習されたモデルが基本的にブラックボックスであり経済学的な解釈が困難であることがあげられる。

2.1 辞書アプローチ

伝統的なデータでは迅速に観測することが困難であったマクロ経済活動の側面を、テキスト情報を用いて定量化する研究は、当初は英語圏の経済分析を中心に進められてきた。以下ではまず辞書アプローチの応用例をとりあげる。辞書アプローチでは関心のある事象に直結するキーワードをあらかじめ選択する必要がある。たとえば、Carroll (2003) では、New York Times 紙と Washington Post 紙の新聞一面における「inflation」で始まる単語を含むインフレ関連記事数から家計のインフレ予想へ与える影響を分析している。

新聞記事のテキスト情報以外でも、検索サイト Google で提供される検索件数を表す Google トレンドがナウキャストに有用であることを Choi and Varian (2009) が指摘した。彼らは実証例として米国の自動車販売と住宅販売に関連したカテゴリーに属する単語の検索数を用いている。似た試みとしては McLaren and Shanbhogue (2011) が、失業者数のキーワードとして、「jobs」「Jobseeker's Allowance」「unemployment benefit」等、住宅価格のキーワードとして「house prices」「buy house」「sell house」等を設定し、やはり検索件数から英国経済動向のナウキャストを行った。また Antenucci et al. (2014) は Twitter 社 (現 X 社) のツイートを用いて、「lost job」や「unemployment」等の失職や求職のシグナルを含む n グラム (n -gram、 n 個の隣接する単語) の出現頻度から「ミシガン大学ソーシャルメディア失職指数 (The University of Michigan Social Media Job Loss Index)」を作成し、2012 年のハリケーン・サンディや 2013 年の政府閉鎖等の米国の大きなイベント時のナウキャストに有益であったことを示している²。

複数の単語の出現頻度を同時に分析に反映させる辞書アプローチとしては、 n グラム以外にも、共起語 (文中で同時に出現する傾向のある単語) や、複数の単語の出現を条件とする論理型 (ブーリアン型) の指数も採用されている。論理型指数の応用例としては Baker et al. (2016) の「経済政策不確実性 (EPU) 指数」がよく知られている。彼らは、米国の主要新聞 10 紙 (USA Today 紙や Wall Street Journal (WSJ) 紙等) を用いて、「経済」「政策」「不確実性」の 3 つのカテゴリーの単語リストの単語が少なくとも 1 語ずつ含まれる記事を探し出し、この条件を満たす記事の数が新聞全体の記事の数に占める割合を月次で集計した指数を提案した。マクロ経済活動を予測する上で不確実性の度合いの重要性は理論的にも示されており、現在もウェブサイト上でリアルタイムで更新されている EPU 指数は研究者や民間の分析で活用されている。

² Antenucci et al. (2014) の指数はリアルタイムで週次データが公開されていたが、2017 年に更新が停止された。

辞書アプローチの中でおそらく最も経済への応用例の多い手法が、センチメント分析である。特に株価形成に関する市場センチメント分析については古くから多くの研究が行われてきた。代表的な研究としては、Tetlock (2007) が WSJ 紙の市場最新情報コラムからセンチメント指数（感情スコア）を計算し、株価上昇率や出来高との相互依存関係を調べている。センチメント分析では、文に含まれる単語の感情極性値を集計することによって、その文の内容の評価が行われる。各単語の持つ意味のポジティブ/ネガティブの度合いを数値化した感情極性値は、感情極性辞書で参照される。Tetlock (2007) の分析では、一般的な単語の感情極性辞書である General Inquirer (Stone et al., 1966) の Harvard IV-4 Psychosocial Dictionary (以下、GI 辞書) が用いられている。一方、Loughran and McDonald (2011) は、GI 辞書に含まれる一般的なネガティブ極性値と、金融（ファイナンス）のドメインで用いられる単語のネガティブ極性値に乖離があることを指摘したうえで、米国上場企業の有価証券報告書に相当する 10-K で用いられる単語から辞書を構築し、金融市場分析で広く利用されるようになった (以下、LM 辞書)。ただし、マクロ経済活動と金融（ファイナンス）のドメインは必ずしも同一ではないため、Shapiro et al. (2022) は景気動向のニュースセンチメントを把握する目的で新聞の経済関連記事で用いられる単語をもとにした新しい辞書を構築した (以下、SSW 辞書)。SSW 辞書を用いて米国主要新聞 24 紙から計算された日次のニュースセンチメント指数はサンフランシスコ連銀のウェブサイトで「Daily News Sentiment Index」としてリアルタイムで公開されている。Buckman et al. (2020) は Daily News Sentiment Index が新型コロナ感染症発生時の 2020 年 1 月時点で他の指数に先駆けて落ち込みが観測できている点を指摘し、ナウキャストの有効性を議論している。また、株価形成に関するセンチメント分析では、2021 年に S&P Dow Jones Indices 社が Twitter 社（現 X 社）の協力のもと S&P 500 の構成銘柄の中からポジティブなツイートのあった上位 200 銘柄にもとづく「S&P 500 Twitter Sentiment Index」のリアルタイムな提供を開始した。さらに、Barbaglia et al. (2023) では米国主要新聞 6 紙について「経済」、「金融部門」、「インフレ」、「製造業」、「金融政策」、「失業」の 6 つのカテゴリーの属する単語リストを用いて記事のテーマを分類し 6 つのセンチメント指数を計算し、他の数値データと組み合わせることによって、主要なマクロ経済変数の予測精度が改善される結果を示している。

次に辞書アプローチの日本の景気動向分析への応用例を紹介しよう。検索件数を用いたナウキャストの試みとして 2013 年にポータルサイトの Yahoo!JAPAN が、景気動向指数の一致指数と相関の高い頻度で検索される 206 語のキーワードの検索数の加重平均から「Yahoo!JAPAN 景気指数」を構築し公表前の景気動向指数に対して先行性があることを示した。Google トレンドの日本経済への応用可能性についても白木他 (2013) でサービス消費のナウキャストを用いて議論されている。

EPU 指数については Baker et al. (2016) によって日本の値も計算されている。日本の EPU 指数の計算には、1988 年から 2013 年の朝日新聞と読売新聞の邦文記事が用いられた。論文付録の解説によれば、「経済」カテゴリーの単語は「経済」のみが、「不確実性」カテゴ

リーの単語には「不透明」と「不確実性」が採用されている。一方で、「政策」カテゴリーの単語リストに含まれるものは、「政策」、「税」、「歳出（あるいは公共事業費）」、「公共投資（あるいは国費）」、「規制」、「日本銀行（あるいは日銀）」、「財政（あるいは赤字）」である。Baker et al. (2016) による日本の EPU 指数は Arbatli Saxegaard et al. (2022) によって、日本経済新聞、毎日新聞を加えた 4 紙に拡張され、「日本の政策不確実性指数」として公開されている³。また Tetlock (2007) で採用された市場センチメント分析の手法は沖本・平澤 (2014) が日本のデータに応用している。さらに Shapiro et al. (2022) の SSW 辞書のようにマクロ経済のドメインに特化した辞書を用いて、日本経済新聞の記事から計算したニュースセンチメント指数は「日経・UTEcon 日次景気指数」の名称で、日本経済新聞社と東京大学エコノミックコンサルティング株式会社（以下、UTEcon 社）から日リアルタイムで提供されている。なお日経・UTEcon 日次景気指数とその計算に用いられた辞書については後述する。

2.2 機械学習アプローチ

機械学習アプローチでは基本的に次のような手続きを経て、景気動向のナウキャストを行う。まずテキスト情報を数値ベクトル形式のデータに変換する。次に、マクロ経済活動とテキスト情報の関係を記述する予測モデルを過去のデータを用いて推定（学習）する。最後に、直近のテキスト情報を入力としたモデルの出力から足もとの経済状況を把握する。一般的にテキスト情報は膨大であり、数値化されたデータはビッグデータである。説明変数の数が非常に大きい状況で回帰分析を用いると過適合の問題が生じるため、テキストデータを用いた予測モデルの推定には機械学習の手法が採用されることが多い⁴。その結果得られた予測モデルの予測値をテキスト情報にもとづく景気の指数として利用することが可能である。

機械学習には様々な手法が存在するが、分析対象である出力があらかじめ与えられた教師あり学習と、出力の情報がない教師なし学習に大別することができる⁵。Blei et al. (2003) の LDA (latent Dirichlet allocation) に代表されるトピックモデルによる潜在的なトピック分類は、教師なし学習の一種である。トピックモデルでは抽出された潜在的なトピックに経済学的な解釈を付与しやすいという利点があり経済分析で比較的良好に用いられる。たとえば中央銀行は景気や物価、金利、失業率等の様々な経済要因を考慮しながら金融政策を行っているが、中央銀行の議事録に対してトピックモデルを適用し、特に中央銀行が重視す

³ 同指数は独立行政法人経済産業研究所 (RIETI) のウェブサイトにおいて月次の頻度で更新されている。似た試みとして、日本銀行の公式文書に含まれる「物価」カテゴリーと「懸念」カテゴリーの単語から金融政策スタンスを予測する「日銀テキストインデックス」を 2015 年にクレディ・スイス証券が開発したが、数か月で更新が停止されている。

⁴ モデルの過適合が生じるとサンプル内予測に比較してサンプル外予測の精度が著しく低下することで、ナウキャストの実践が困難となる。

⁵ たとえば、通常の回帰分析は被説明変数を出力、説明変数を入力とする教師あり学習である。一方、主成分分析は具体的な予測対象の被説明変数を必要としないため、教師なし学習とみなすことができる。

るトピックを抽出することによって金融政策の要因分析を行うことができる (Hansen and McMahon, 2016)。トピックモデルを景気動向の分析に応用した例としては Larsen and Thorsrud (2019) や Thorsrud (2020) があり、前者はノルウェーの新聞記事から抽出された様々なトピックが、産出、消費、生産性、資産価格等のマクロ経済変数に予測力を持つことを示しており、後者は同じトピックを因子モデルと組み合わせることで日次の GDP 予測系列である「Newsy Coincident Index」を計算している。第 1 段階でトピックモデルからトピック系列を抽出し、第 2 段階で分析対象の予測モデルを推定する 2 段階の手続きは、経済学的解釈の観点からは望ましいかもしれない。このことは辞書アプローチを用いた指数を変数とする予測モデルの推定にもあてはまるだろう。一方、解釈よりもナウキャストの予測精度の向上を追究する場合は、分析対象の変数を出力として、テキスト情報を入力とするモデルの教師あり学習がより効率的だと考えられる。教師あり学習を用いた株式市場分析への応用としては、たとえば Yahoo! ファイナンス等の株式掲示板の投稿内容を利用した Antweiler and Frank (2004) の研究がある。彼らはまず 1000 の投稿文の入力に対し、人手によって付与された {買い、中立、売り} のラベルを出力とした訓練データに対し、ナイーブベイズやサポートベクトルマシン等の機械学習によって予測モデルを構築した。次に得られた予測モデルを用いて残りの 150 万の投稿文に対して新たにラベルを付与している。景気動向分析への応用については、たとえば Rambaccussing and Kwiatkowski (2020) が英国の新聞 13 紙から無作為抽出された 1568 の記事に人手でポジティブ/ネガティブの極性値を付与した訓練データを、ナイーブベイズ、サポートベクトルマシン、ニューラルネットワークで学習し、残りの記事から GDP、インフレ率、失業に関する指数を構築している。教師あり学習はこのような離散値を予測する分類問題だけではなく、連続値の回帰問題への応用も可能である。たとえば Manela and Moreira (2017) の研究では、株式オプションから逆算される市場の不確実性指標である VIX を出力とし、WSJ 紙のニュース記事を入力とするモデルをサポートベクトル回帰によって学習し、VIX が存在しない 1986 年以前についての過去のニュース記事を入力に用いた不確実性指数である「News Implied Volatility (NVIX)」を提案している。景気動向についても、GDP、インフレ率、失業率等の連続値をとる主要なマクロ経済変数を出力とし、数値化されたテキスト情報を入力とする予測モデルを学習すればよい。たとえば Kalamara et al. (2022) は、英国の主要 3 紙の新聞ニュース記事を入力とし、ラッソ、リッジ回帰、エラスティックネット、サポートベクトル回帰、ランダムフォレスト、ニューラルネットワーク等、様々な機械学習の手法を用いて主要なマクロ経済変数の予測結果を示している。

次に日本の景気動向把握に関する機械学習アプローチの応用例を紹介する。和泉他 (2007) は国際金融情報センターが毎週発行している市場レポート 156 週分 (1992–1994 年) に対し人手で「景気」「物価」「雇用」等の 14 種類のカテゴリ分類と (円高方向を正とする) 経済動向 {0, ± 1 , ± 2 } のラベル付与を行い、その訓練データについて予測モデルを決定木で学習した。さらに 1995 年の記事を入力とした予測モデルを使って為替レート変

動要因の分析に応用した。このように分析対象とテキスト情報を結びつける訓練データは人手で作成される場合も多いが、内閣府が公表している景気ウォッチャー調査では回答者（景気ウォッチャー）による「良い、やや良い、どちらとも言えない、やや悪い、悪い」の5段階の景気スコアと景気判断理由集のテキスト情報が直接ひもづけられている。特に景気ウォッチャー調査の景気判断理由集は、掲示板の投稿やツイート等の SNS データと比較して内容が景気動向判断に特化され、かつ誤字や回答者間の表記ゆれが公表前に修正されてノイズが少ないことから、景気動向のテキスト分析では頻繁に用いられる。たとえば、景気ウォッチャー調査のテキスト情報を用いて岡崎・敦賀（2015）は、センチメント指数と共起語による分析、敦賀・岡崎（2017）は、トピック分析やナイーブベイズ、サポートベクトルマシン、ニューラルネットワーク等の機械学習による景気動向の分類評価を行っている。

景気ウォッチャー調査の景気判断理由集で学習した予測モデルの入力として、内閣府の「月例経済報告」と日本銀行の「金融経済月報」のテキスト情報を用いることで、景気センチメントの月次指数を計算した試みが山本他（2015）や山本・松尾（2016）による「野村 AI 景況感指数」である。予測モデルの学習には文脈が考慮できるリカレントニューラルネットワークの一種である LSTM が採用されている。さらに同じ予測モデルを用いて野村證券と経済産業省が共同で開発した「SNS×AI 景況感指数」と「SNS×AI 鉱工業生産予測指数」では月次のレポートの代わりにツイートを入力とすることでより高頻度の更新が可能となった⁶。饗場・山本（2018）では後者を用いることで鉱工業生産指数のナウキャストの精度が高まるという結果が示されている。

他にも景気ウォッチャー調査を訓練データとしてモデルを学習し、様々なテキスト情報の入力データによって指数を構築する試みは数多く存在する。五島他（2019）は畳み込みニューラルネットワークにより学習されたモデルの入力にロイターニュース記事を用いた日次のニュース指数、風戸他（2019）は同じモデルに日本銀行総裁記者会見要旨を用いた中央銀行の景気判断指数を計算している。また日本経済新聞記事を入力に用いて日次のニュース指数を計算とした例としては、モデルの学習に Facebook 社が開発した fastText を用いた Goshima et al.（2021）による「ニュース先行指数（News-based Leading Indicator）」や Google 社が開発した BERT（Bidirectional Encoder Representation from Transformers）を用いた Seki et al.（2022）による「S-APIR 指数」が存在する。

一方、Manela and Moreira（2017）型の機械学習アプローチによる不確実性指数については Goshima et al.（2023）が政府債務の不確実性は国債価格に反映され则认为、日本国債の VIX（JGB-VIX）の変動を日本経済新聞の記事を用いて学習した上で、過去の新聞記事から政策の不確実性指数を遡って財政政策の不確実性指数を計算している。また Oh and Shintani（2022）では毎日新聞のニュース記事データと伝統的なマクロ経済データを入力とし、足もとの GDP を出力として予測モデルを学習し、GDP のナウキャストにおけるテキ

⁶ 両指数は野村證券と経済産業省のサイトで公表されていたが、2021 年をもって更新を終了している。

スト情報の有用性を示している。

3. 景気単語極性辞書とセンチメント指数

3.1 景気動向分析と極性辞書の比較

既に議論したようにセンチメント分析では分析対象のドメインに適した極性辞書を使用することが望ましい。五島他(2022)は日本経済新聞(朝刊・夕刊・電子版)の約510万記事をコーパスに用いて、マクロ経済のドメインに特化した「景気単語極性辞書」を作成した。単語分割、品詞推定および基本形の推定には形態素解析エンジンのMeCabとシステム辞書のmecab-ipadic-NEologd(佐藤他, 2017)の組み合わせを用い、考慮する単語は名詞、動詞、形容詞のみとしている。まず「景気」、「経済」及びその合成語と各単語の共起回数を算出し、その上位の単語から総単語数が3,000語となるように候補単語リストを作成する。その際に「リーマンショック」や「新型コロナウイルス」などの特定のイベント固有の単語は候補単語リストから除外されている。次にマクロ経済分析の専門家により景気動向の観点で各単語から連想される意味について、{ポジティブ、ニュートラル、ネガティブ}の3段階のラベル付与を人手で行った。特に景気の見方は個人の経歴や立場によって異なる可能性があり、特定の業種の意見を反映しないように、政府・中央銀行エコノミスト、民間金融機関エコノミスト、報道機関エコノミストの3つの異なる職種の専門家にラベル付与を依頼している。多数決の結果からニュートラルを除外した単語と、機械学習による追加収録単語を合わせ、最終的な景気単語極性辞書の収録単語数は874語となった。収録された品詞の内訳は名詞が567語、動詞が198語、形容詞が109語となっている。

一方、ドメイン横断型の一般的な単語の極性辞書としては、産業技術総合研究所高村研究室(高村他, 2006)の「単語感情極性対応表」や、東北大学乾研究室(小林他, 2005)の「日本語評価極性辞書」が広く用いられている。またLM辞書のように金融(ファイナンス)分野に関するドメイン特化型の極性辞書としては東京大学和泉研究室(Ito et al., 2018)の「金融専門極性辞書」が知られている⁷。景気単語極性辞書を含めた4つの極性辞書の収録単語のサンプル10語と総単語数の比較が表1に示されている。表からは単語感情極性対応表と金融専門極性辞書の収録単語数が、日本語評価極性辞書(用言編)や景気単語極性辞書に比べて多いことがわかる。この収録単語数の違いは日本語評価極性辞書や景気単語極性辞書の極性値が基本的に人手によるラベル付与であるのに対して、単語感情極性対応表と金融専門極性辞書では学習されたモデルに基づいて極性値が計算されていることが反映されている。またラベル付与の方法の違いから、日本語評価極性辞書や景気単語極性辞書の極性値が{ポジティブ、ネガティブ}の2値であるのに対し、単語感情極性対応表と

⁷ 五島・高橋(2017)も金融(ファイナンス)のドメインに特化した辞書を提案している。

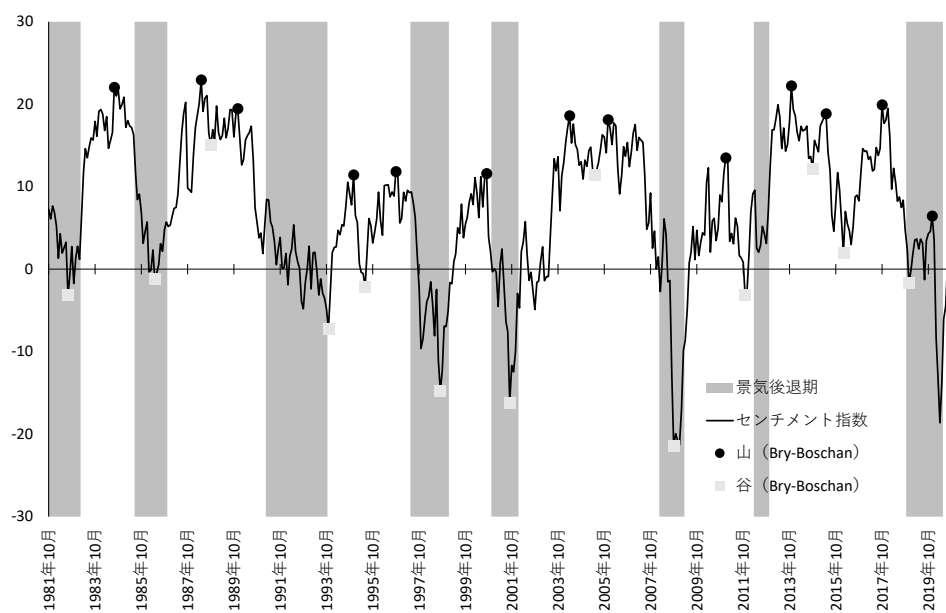
金融専門極性辞書の極性値は正負の連続値で与えられている⁸。

表 1：日本語極性辞書の比較

	単語感情極性対応表 (高村他, 2006)	日本語評価極性辞書 (小林他, 2005)	金融専門極性辞書 (Ito et al., 2018)	景気単語極性辞書 (五島他, 2022)
ポジティブ 単語例	優れる 良い 喜ぶ 褒める めでたい	バツグン 馬力がある 高名です 持ち歩ける ハイパワーだ	増配 営業増益 黒字転換 増益 復配	好況 金融緩和 上向く 活気づく 底堅い
ネガティブ 単語例	ない 酷い 病氣 死ぬ 悪い	野暮だ 手狭だ 冒す 御機嫌斜めの 凄絶の	赤字 苦戦 減益 減配 営業減益	貸し渋り 信用収縮 落ち込む 買い控える 重い
総単語数	55125 語	5280 語	19630 語	874 語

景気単語極性辞書と日本経済新聞朝刊記事データから五島他（2022）によって計算された景気のセンチメント指数が図 1 に示されている。

図 1：景気単語極性辞書によるセンチメント指数



出典：五島他（2022）

⁸ このため表 1 の単語サンプルは単語感情極性対応表と金融専門極性辞書については極性値が最大値と最小値に近い単語それぞれ 5 語が示されており、日本語評価極性辞書や景気単語極性辞書についてはポジティブとネガティブの単語それぞれ 5 語が無作為抽出されている。

具体的には次式について当日の重みを 2/31 とする指数移動平均 (exponential moving average) として計算されている。

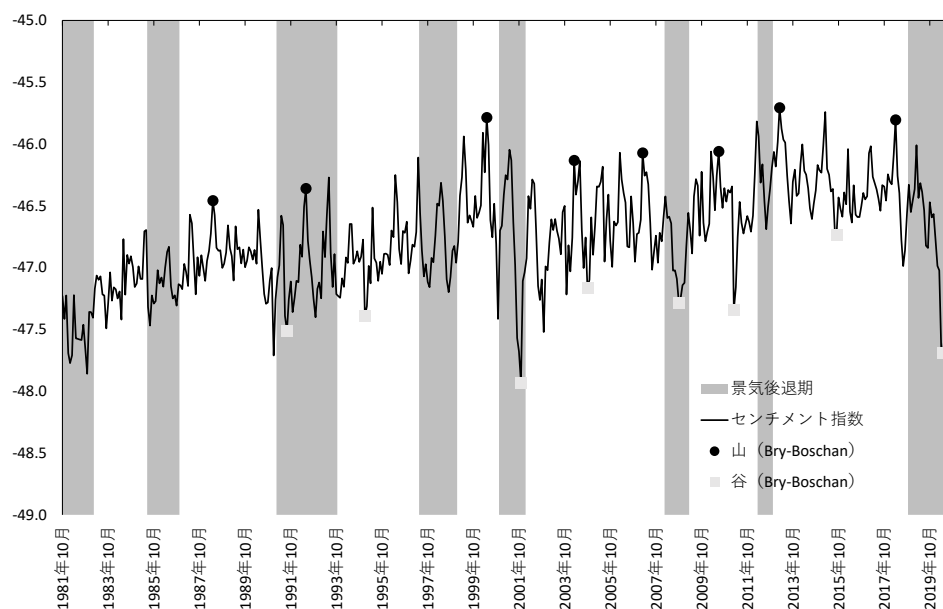
$$\frac{Positive\ Words_t - Negative\ Words_t}{Positive\ Words_t + Negative\ Words_t} \times 100$$

ただし、 $Positive\ Words_t$ と $Negative\ Words_t$ はそれぞれ、 t 日の記事中に出現した (重複を許した) ポジティブな単語数とネガティブな単語数である。同じ図には内閣府の景気基準日付による景気後退期局面が灰色の領域で、Bry-Boschan 法から計算された系列の山谷のタイミングがそれぞれ黒丸と四角で示されている。図 1 からはセンチメント指数の山谷は景気基準日付の景気循環より多いものの、谷のタイミングは概ね景気基準日付と一致していることがわかる。また政府が認定した景気循環よりも早めに転換点が検出されており、先行性が確認できる。

景気単語極性辞書によるセンチメント指数との比較の目的で、五島他 (2022) はドメイン横断型の極性辞書である単語感情極性値対応表を使ったセンチメント指数も計算している。なお、単語感情極性値対応表に収録された 55125 語のうち 0.72% にあたる 397 語が景気単語極性辞書と重複している。また単語感情極性値対応表で極性値が連続値で与えられているため、センチメント指数は次式の指数移動平均として計算されている。

$$\frac{Labels_t}{Words_t} \times 100$$

図 2：単語感情極性値対応表によるセンチメント指数

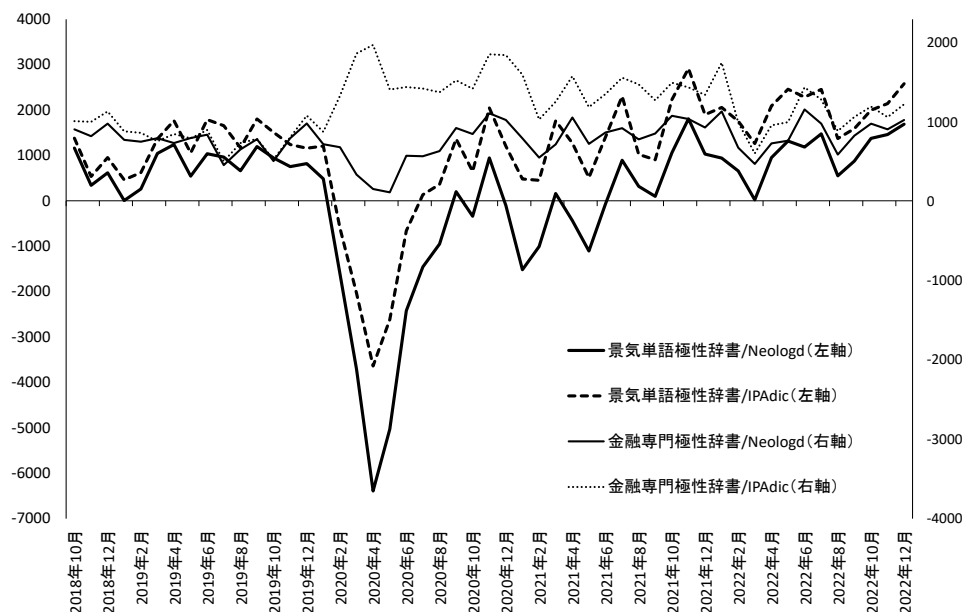


出典：五島他 (2022)

ただし、 $Labels_t$ と $Words_t$ はそれぞれ、 t 日の記事中に出現した（重複を許した）単語の極性値の総和と単語の数を表す。単語感情極性値対応表によるセンチメント指数は図2に示されている。図2からは一般的な単語の極性辞書による感情極性値対応表によるセンチメント指数については特に前半の期間において、山と谷のタイミングが景気基準日付とは必ずしも一致しない傾向があることがわかる。

以上2つのセンチメント指数の挙動の差は極性辞書のドメインの差が明確に反映されているといえる。ただし指数の計算に用いたテキストデータが、前者については辞書作成に用いたコーパスと同一であることも結果に有利に影響している可能性も考えられる。そこで景気単語極性辞書の作成に用いられた日本経済新聞以外のテキストデータを用いてセンチメント指数を計算した場合の頑健性を確認しよう。この目的のために新型コロナウイルスの感染拡大前から2022年12月までの毎日新聞記事を利用する。また辞書作成時にはマクロ経済に関連する複合名詞を収録できるように MeCab のシステム辞書には mecab-ipadic-NEologd が使用された。これはたとえば「国土交通省は2001年に設置されました。」という文が標準のシステム辞書である IPAdic を用いた場合、「国土/交通省/は/2001/年/に/設置/さ/れ/まし/た/。」と分かち書きされるのに対して、mecab-ipadic-NEologd を用いた場合、「国土交通省/は/2001 年/に/設置/さ/れ/まし/た/。」と分かち書きされるようになるためである。このため毎日新聞記事データの形態素解析では両者のシステム辞書を用いた場合の頑健性についても確認する。さらに比較対象として、Yahoo!ファイナンスの掲示板投稿データをコーパスとして作成された金融専門極性辞書によるセンチメント指数も計算する。2つの極性辞書と2つのシステム辞書を用いて計算されたセンチメント指数が図3に示され

図3：新型コロナウイルスの感染拡大期のセンチメント指数



出典：筆者作成

ている。なお2値の極性値が収録された景気単語極性辞書と連続値の極性値が収録された金融専門極性辞書によるセンチメント指数の比較と要因分析を容易にするためにここでは単純に t 月の新聞記事中に出現した単語の極性値の総和の月次系列が比較されている。まず景気単語極性辞書によるセンチメント指数（ここでは1か月間に出現したポジティブな単語数とネガティブな単語数の差）は直近の景気の谷にあたる2020年5月の谷の前後で大きく落ち込んでいることがわかる。システム辞書がどちらを用いた場合でも頑健な結果となっている。一方で金融専門極性辞書によるセンチメント指数はシステム辞書がmecab-ipadic-NEologdである場合には落ち込みが見られるものの、システム辞書がIPAdicの場合は谷周辺で逆に増加している。この差はたとえば「感染拡大」という単語がIPAdicによる分かち書きだと「感染」「拡大」に分解されてしまい、ポジティブな単語として「拡大」が多く出現していることが反映された結果である。

以上の結果の簡単な考察からは、以下の点が示唆されるだろう。まず景気動向のセンチメント分析を行うにあたり、ドメイン横断型ではなくドメイン特化型の極性辞書を選択することの重要性である。次に、文章を分かち書きにして、品詞や基本形を推定する形態素解析の手法の違いから最終的な結果が左右されてしまう可能性である。このことから、特定の意味を持たない記号や単語であるストップワードの除去等も含め、テキストデータの前処理段階においても慎重な作業と検討が必要であることがわかる。最後に、景気把握を目的とするセンチメント指数の計算では、収録単語数が比較的小さい極性辞書でも頑健で良好な結果が得られることである。つまり迅速な景気判断の目的には、極性値の曖昧な多くの単語を分析に含めるよりも、特に重要な少数のキーワードに着目することが有利であるかもしれない。文脈を考慮しないセンチメント指数の予測精度は、洗練された言語モデルよりも劣るだろう。しかしモデルの更新及びモデルの学習のコストが高い現状では、単純な計算で頑健な結果が得られるセンチメント指数も十分有用であると考えられる。

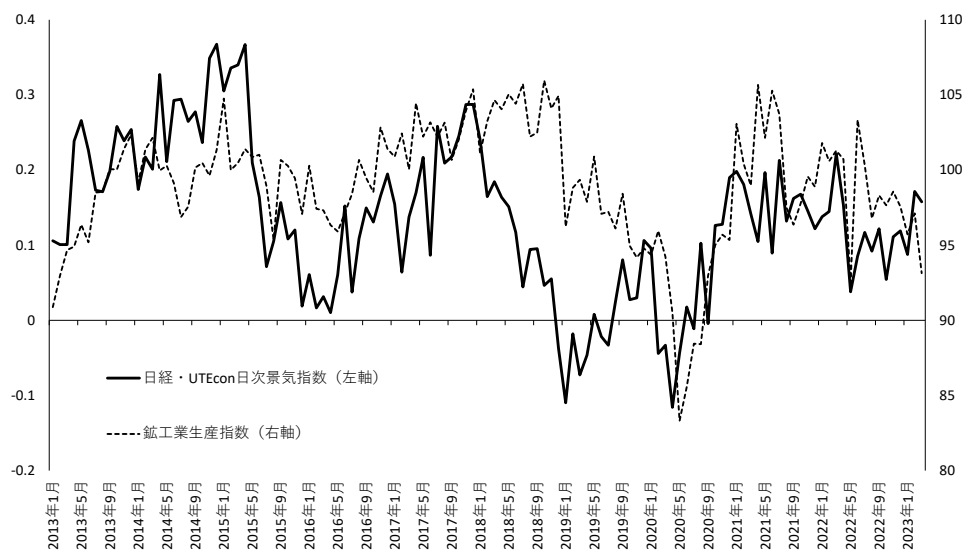
3.2 日経・UTEcon 日次景気指数

景気単語極性辞書を用いて、日本経済新聞の朝刊記事から計算されたセンチメント指数は「日経・UTEcon 日次景気指数」として、2022年3月より日本経済新聞社とUTEcon社から提供が開始された。この指数は日次データとして、日本経済新聞の休刊日を除き毎朝午前8時に更新されている。また2023年8月からは、もとのセンチメント指数を14業種に分解した業種別指数と国内外9地域に分解した地域別指数についても計算されている。業種別指数は特定の業種における景況感、また地域別指数は国内の景気動向における外生的な海外要因の寄与度の把握の目的で利用できる。

業種別（あるいは地域別）の指数は次のような手続きに従って、機械学習で記事を分類し計算されている。まず訓練データとして、日本経済新聞社で提供されている分類タグ付き記事データを利用する。業種に関しては第2次産業7分類（「資源・エネルギー」、「素材」、「機械」、「エレクトロニクス」、「輸送機器」、「食品」、「生活・医薬」）及び第3次産業7分

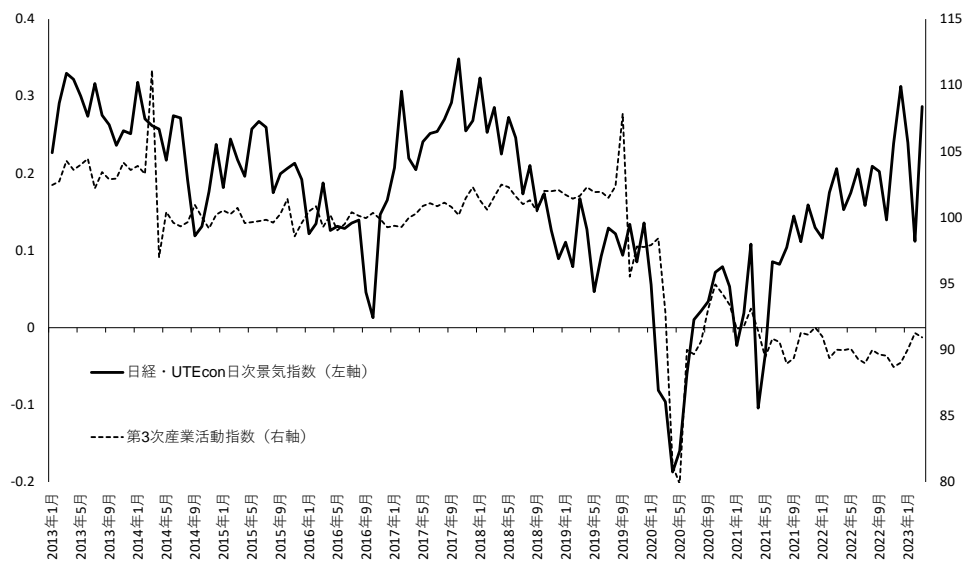
類（「建設・不動産」、「小売・飲食」、「金融」、「運輸」、「情報・通信・ネット」、「サービス」、「その他」）の計 14 種が考慮されている。1 つの記事が複数の業種を取り上げる場合もあるため、1 つの入力に対して複数のタグの出力とすることを許容するマルチラベル分類問題として最初の 256 形態素を入力として BERT（日本語金融 BERT）で学習した。学習したモデルを各記事のタグの予測に用いて、予測結果の和が 1 となるように確率分布に変換する。次に確率と当該記事のポジティブ/ネガティブ単語数の積を計算し、同一の業種タグが付与された記事の値を集計して、業種別の指数として計算される。

図 4：日経・UTEcon 日次景気指数（エレクトロニクス）と鉱工業生産指数
（電子部品・デバイス工業／電気・情報通信機械工業）



出典：UTEcon 社

図 5: 日経・UTEcon 日次景気指数（小売・飲食）と第 3 次産業活動指数
（卸売業、小売業/生活娯楽関連サービス 飲食店飲食サービス）



出典：UTEcon 社

図4では第2次産業の業種別指数の例として日経・UTEcon 日次景気指数のエレクトロニクスとその業種分類と整合的になるように調整したカテゴリーの鉱工業生産指数を示している。一方、図5では第3次産業の業種別指数の例として日経・UTEcon 日次景気指数の小売・飲食とその業種分類と整合的になるように調整したカテゴリーの第3次産業活動指数を示している。2つの図からは業種別指数が関連する鉱工業生産指数や第3次産業活動指数と連動している様子が確認できる。

4. おわりに

本稿では景気動向を把握するためのテキスト情報の利用の現状を概観した。分析手法に関しては便宜的に辞書アプローチと機械学習アプローチに分けて整理を行っている。辞書アプローチの中では、自然言語処理の手法としては古典的なセンチメント分析が特に計算や解釈の容易性の観点から今でも経済分野では盛んに利用されていることを指摘した。また景気のセンチメント分析では、ドメイン横断型ではなくドメイン特化型の極性辞書を選択することが重要であると考えられる。しかし、辞書アプローチにおいても、極性辞書や業種別センチメント指数の作成過程で機械学習の手法が利用されていることを考慮すれば、広義にはすべて機械学習アプローチであるという見方も可能である。今後については、BERT や ChatGPT に代表される大規模言語モデル (LLM) の景気動向分析への応用がますます進んでいくことが予想される。またテキスト情報を用いた日本の景況感指数は過去に何度も提案されてきたが、継続して更新されているものは必ずしも多くない。LLM 等の最新言語モデルは進化の速度も速く近年益々複雑化されておりモデルが変更された場合の過去系列の遡及推計も含めて、指数の継続性は大きな課題であるといえよう。

参考文献

- 饗場行洋・山本裕樹 (2018) 「データサイエンスと新しい金融工学」, 『財界観測』, 第 81 巻, 第 2 号, 30-41 頁.
- 和泉潔・松井宏樹・松尾豊 (2007) 「人工市場とテキストマイニングの融合による市場分析」, 『人工知能学会論文誌』, 第 22 巻, 第 4 号, 397-404 頁.
- 王悠介・須合智広・高橋耕史・松村浩 (2021) 「位置情報データによる経済活動のナウキャスティング」, ワーキングペーパーシリーズ No.21-J-2, 日本銀行.
- 岡崎陽介・敦賀智裕 (2015) 「ビッグデータを用いた経済・物価分析について—研究事例のサーベイと景気ウォッチャー調査のテキスト分析の試み—」, 調査論文, 日本銀行.
- 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用」, 『証券アナリストジャーナル』, 第 52 巻, 第 4 号, 67-75 頁.

- 風戸正行・黒崎哲夫・五島圭一（2019）「日本銀行による景気判断のトーン分析」, 金融研究所ディスカッションペーパーシリーズ 2019-J-16, 日本銀行.
- 五島圭一・新谷元嗣・高村大也（2022）「景気単語極性辞書の構築とその応用」, 『自然言語処理』, 第 29 巻, 第 4 号, 1233–1253 頁.
- 五島圭一・高橋大志（2017）「株式価格情報を用いた金融極性辞書の作成」, 『自然言語処理』, 第 24 巻, 第 4 号, 547–577 頁.
- 五島圭一・高橋大志・山田哲也（2019）「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」, 『金融研究』, 第 38 巻, 第 3 号, 1–41 頁.
- 小林のぞみ・乾健太郎・松本裕治・立石健二・福島俊一（2005）「意見抽出のための評価表現の収集」, 『自然言語処理』, 第 12 巻, 第 3 号, 203–222 頁.
- 佐藤敏紀・橋本泰一・奥村学（2017）「単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討」, 言語処理学会第 23 回年次大会発表論文集, 875–878 頁.
- 白木紀行・松村浩平・松本梓（2013）「景気判断における検索データの利用可能性」, 調査論文, 日本銀行.
- 高村大也・乾孝司・奥村学（2006）「スピンモデルによる単語の感情極性抽出」, 『情報処理学会論文誌』, 第 47 巻, 第 2 号, 627–637 頁.
- 敦賀智裕・岡崎陽介（2017）「テキストデータを用いた金融経済分析—サーベイと景気ウォッチャー調査を用いた分析—」, 『統計』, 第 68 巻, 第 4 号, 2–8 頁.
- 山本裕樹・水門善之・木下智夫（2015）「人工知能で政府・日銀の景況感を指数化する」, 調査レポート, 野村證券.
- 山本裕樹・松尾豊（2016）「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」, 2016 年度人工知能学会全国大会（JSAI2016）論文集, 3L3OS16a2 頁.
- Antenucci, Dolan, Michael Cafarella, Margaret C. Levenstein, Christopher Ré, and Matthew D. Shapiro (2014) “Using Social Media to Measure Labor Market Flows,” NBER Working Papers 20010, National Bureau of Economic Research.
- Antweiler, Werner and Murray Z. Frank (2004) “Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards,” *Journal of Finance*, Vol. 59, No. 3, pp. 1259–1294.
- Arbatli Saxegaard, Elif C., Steven J. Davis, Arata Ito, and Naoko Miake (2022) “Policy Uncertainty in Japan,” *Journal of The Japanese and International Economies*, Vol. 64, p. 101192.
- Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2016) “Measuring Economic Policy Uncertainty,” *The Quarterly Journal of Economics*, Vol. 131, No. 4, pp. 1593–1636.
- Barbaglia, Luca, Sergio Consoli, and Sebastiano Manzan (2023) “Forecasting with Economic News,” *Journal of Business & Economic Statistics*, Vol. 41, No. 3, pp. 708–719.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003) “Latent Dirichlet Allocation,” *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022.

- Buckman, Shelby R., Adam Hale Shapiro, Moritz Sudhof, and Daniel J. Wilson (2020) “News Sentiment in the Time of COVID-19,” *FRBSF Economic Letter*, No. 2020-08.
- Carroll, Christopher D. (2003) “Macroeconomic Expectations of Households and Professional Forecasters,” *The Quarterly Journal of Economics*, Vol. 118, No. 1, pp. 269–298.
- Choi, Hyunyoung and Hal Varian (2009) “Predicting the Present with Google Trends,” technical report, Google.
- Goshima, Keiichi, Hiroshi Ishijima, and Mototsugu Shintani (2023) “News Implied Volatility and Aggregate Economic Activity: Evidence from the Japanese Government Bond Market,” *Applied Economics Letters*, forthcoming.
- Goshima, Keiichi, Hiroshi Ishijima, Mototsugu Shintani, and Hiroki Yamamoto (2021) “Forecasting Japanese Inflation with a News-based Leading Indicator of Economic Activities,” *Studies in Nonlinear Dynamics & Econometrics*, Vol. 25, No. 4, pp. 111-133.
- Hansen, Stephen and Michael McMahon (2016) “Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication,” *Journal of International Economics*, Vol. 99, No. S1, pp. S114–S133.
- Ito, Tomoki, Hiroki Sakaji, Kota Tsubouchi, Kiyoshi Izumi, and Tatsuo Yamashita (2018) “Text-Visualizing Neural Network Model: Understanding Online Financial Textual Data,” in *Phung D., Tseng V., Webb G., Ho B., Ganji M., Rashidi L. (eds) Advances in Knowledge Discovery and Data Mining. PAKDD 2018. Lecture Notes in Computer Science, Springer, vol 10939*, pp. 247–259.
- Kalamara, Eleni, Arthur Turrell, Chris Redl, George Kapetanios, and Sujit Kapadia (2022) “Making Text Count: Economic Forecasting Using Newspaper Text,” *Journal of Applied Econometrics*, Vol. 37, No. 5, p. 896–919.
- Larsen, Vegard H. and Leif A. Thorsrud (2019) “The Value of News for Economic Developments,” *Journal of Econometrics*, Vol. 210, No. 1, pp. 203-218.
- Loughran, Tim and Bill McDonald (2011) “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks,” *Journal of Finance*, Vol. 66, No. 1, p. 35–65.
- Manela, Asaf and Alan Moreira (2017) “News Implied Volatility and Disaster Concerns,” *Journal of Financial Economics*, Vol. 123, No. 1, pp. 137–162.
- McLaren, Nick and Rachana Shanbhogue (2011) “Using Internet Search Data as Economic Indicators,” *Bank of England Quarterly Bulletin*, Vol. 2011 Q2, No. 1, pp. 134–140.
- Oh, Yusuke and Mototsugu Shintani (2022) “Nowcasting Japanese GDP Using News Data and Machine Learning,” *paper presented at The 15th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2022)*.
- Rambaccussing, Dooruj and Andrzej Kwiatkowski (2020) “Forecasting with News Sentiment: Evidence with UK Newspapers,” *International Journal of Forecasting*, Vol. 36, pp. 1501–1516.

- Seki, Kazuhiro, Yusuke Ikuta, and Yoichi Matsubayashi (2022) “News-based Business Sentiment and Its Properties as an Economic Index,” *Information Processing and Management*, Vol. 59, p. 102795.
- Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson (2022) “Measuring News Sentiment,” *Journal of Econometrics*, Vol. 228, pp. 221–243.
- Stone, Philip J., Dexter C Dunphy, Marshall S. Smith, and Daniel M. Ogilvie. (1966) *The General Inquirer: A Computer Approach to Content Analysis*: MIT Press.
- Tetlock, Paul C. (2007) “Giving Content to Investor Sentiment: The Role of Media in the Stock Market,” *Journal of Finance*, Vol. 62, No. 3, pp. 1139–1168.
- Thorsrud, Leif Anders (2020) “Words are the New Numbers: A Newsy Coincident Index of the Business Cycle,” *Journal of Business & Economic Statistics*, Vol. 38, No. 2, pp. 393–409.