

テキスト分析を用いた米景気後退確率 の検討[†] —Beige Book の文書データから—

石川 亮

(日本政策投資銀行設備投資研究所)

[†] 2024 年 3 月 19 日に開催された日本政策投資銀行設備投資研究所フリーディスカッションでは、東洋大学中村純一教授、トロント大学野澤良雄助教授、立正大学慶田昌之准教授、日本政策投資銀行産業調査部、同設備投資研究所の方々から有益なコメントを頂いた。これらの方々に感謝申し上げたい。

なお、本稿は、2024 年 3 月 28 日に日本政策投資銀行のウェブサイトで公表した「DBJ Research テキスト分析を用いた米景気後退確率の検討」をもとに、分析手法、先行研究との関係および考察を加筆・再構成したものである。詳細は補論 4 を参照。なお、本稿の内容や見解等はすべて執筆者個人に属し、残された誤りはすべて筆者の責任に帰す。

Assessing U.S. Recession Probabilities via Textual Analysis:
Evidence from Beige Book Documents
Economics Today, Vol.47, No.3, July 2026

Ryo ISHIKAWA
Research Institute of Capital Formation
Development Bank of Japan

要旨

近年、オルタナティブデータを経済分析に活用する事例が増加している。オルタナティブデータの特徴のうち、速報性に注目した分析は数多くみられる。一方で、テキストデータなど、数値以外のデータを分析対象とする「異質性」に着目した分析例は相対的に少ない。多面的な経済分析を行ううえでは、こうした分野の発展が期待される。

特に金融政策分野においては、中央銀行のコミュニケーションを分析することへの関心が高く、経済分析におけるテキスト分析の重要性は増している。そこで本稿は、米連邦準備制度が公表する地区連銀景況報告（Beige Book）をテキスト分析の対象とし、数値データ以外の情報を用いて、米国景気の動向を分析した。

具体的には、センチメント分析で Beige Book が景気動向を捉える要素を含むことを確認し、類似度分析で単語の出現回数に注目すると景気変動を捉えられることを示した。

これらの分析結果を踏まえ、単語の出現回数に注目して景気後退確率を推計するモデルを構築したところ、既存のモデルと比較して、先行性および推計精度の両面で良好な結果を得た。本稿は、テキストデータを用いた景気分析の有用性を示すとともに、中央銀行の景気認識を定量的に把握する方法を提案する。

Keywords： オルタナティブデータ、テキスト分析、機械学習、ランダムフォレスト、Beige Book、景気、中央銀行

目次

1. はじめに	1
(オルタナティブデータの特徴、分析例)	2
(金融政策とテキスト分析)	4
2. データおよび分析方法	6
(1) 使用データ	6
(2) テキスト分析の手法	7
(先行研究)	7
(分析のアプローチ)	7
(データの加工)	7
(センチメント分析)	8
(文書間の類似度を計算)	9
(景気後退確率の試算)	12
3. 結果	14
(センチメント分析)	14
(文書間の類似度)	15
(景気後退確率)	19
4. ディスカッション、結論	25
(補論1) シーザー暗号化による文字化けの復号について	28
(補論2) 見出し語化について	30
(補論3) 過学習の対処について	31
(補論4) テキスト分析に関する近年の研究と本研究の位置づけ	32
参考文献	33

1. はじめに

近年、経済分析において、オルタナティブデータの活用が期待されている。オルタナティブデータとは政府統計を代替するデータとされ、様々な面で政府統計を補完する役割が期待されている。例えば、Banbura & Reichlin (2010)は、日次の原油価格や原材料価格などの高頻度データを活用し、欧州の GDP 統計を公表前に把握する手法を提案するほか、Schmidt & Vosen (2011)は、Google の検索情報を用いて米国の個人消費の予測を行っている。日本でも、白木・松村・松本 (2013)が、景気判断において検索情報の利用可能性の検討を行っている。

もっとも、オルタナティブデータの活用が急速に進んだのは、2020 年の新型コロナウイルス（以下、コロナ）感染症拡大以降とみられる。コロナ禍でオルタナティブデータの活用が進んだ要因の一つとして、経済状況が大きく変化する中、政府統計が公表される前に景気を把握し、政策などの意思決定を行う必要性が高まったことが指摘できる。実際、コロナ禍の初期には、日本銀行が 2020 年 4 月の「経済・物価情勢の展望（展望レポート）」¹において、国内外の高頻度データ（石炭使用量やレストランの客足、繁華街の夜間人口など）を用いて経済分析を行ったほか、内閣府も、2020 年 4 月の「月例経済報告」²において、日次の有効求人数やクレジットカード支出の動向などをもとにして、政府統計の公表を待たずに景気判断を引き下げた。その後も、厳格な活動制限がとられたことなどから、特に飲食・宿泊業などのサービス業を中心に大きな影響があらわれたが、サービス産業の政府統計公表は約 2 ヶ月遅延するため、サービス消費と強く連動する人流データへの関心が高まった。同時期に、Google が「コミュニティ モビリティ レポート」³と呼ばれる各地の人流データの公表を開始したほか（2022 年 10 月 15 日に更新停止）、内閣府も、コロナが地域経済に与える影響を可視化することを目的に、「V-RESAS」⁴のウェブサイト上で高頻度データの公表を行うなど、データソースの整備も進んだため、人流データを用いた景気分析が急速に広まった。

近年は学界でもオルタナティブデータへの関心が高まっており、2023 年秋に開催された 2023 年度日本経済学会では、複数の特別セッションでオルタナティブデータ関連の報告がなされ、多数の経済学者が議論に参加する（表 1-1a,b）など、関心の高さがうかがえた。経済学会では、オルタナティブデータの特徴や有用性が共有される一方、オルタナティブデータの分析結果に既存の経済理論の枠組みをどのように適応させるかといった課題意識もみられた。

本稿では、本章（第 1 章）の後半で、このところ注目が集まるオルタナティブデータの特徴や先行研究を簡単にまとめたうえで、経済分析への活用事例を紹介する。第 2 章では、オルタナティブデータのうち、テキストデータにフォーカスし、分析のアプローチを紹介する。第 3 章では、米国の連邦準備制度 (Fed) が年に 8 回公表する地区連銀経済報告 (Beige Book)

¹ <https://www.boj.or.jp/mopo/outlook/index.htm>

² https://www5.cao.go.jp/keizai3/getsurei/kako_getsurei.html

³ <https://www.google.com/covid19/mobility/>

⁴ <https://v-resas.go.jp/>

のテキスト分析を行い、米経済の分析や景気後退確率の試算等を行う。

表 1-1a 日本経済学会の概要

日程・場所	2023年9月16-17日、関西大学
参加者	国内外の研究者、学生、一部エコノミストなど
主なビッグデータ、AI関連の報告	・ 滞在人口データを用いた取引先企業のリース需要予測(加藤ほか) ・ コロナ禍と年次有給休暇関する一考察(井草ほか) ・ How do Inflation Expectation Responded to News? Evidence from a High Frequency Survey of Consumers(丹後ほか) ・ オルタナティブデータの景気判断および政策分析効果への活用(宇南山ほか) ・ 人流データの活用と進展(重岡ほか)

表 1-1b 日本経済学会でのオルタナティブデータに関する主な意見等

・ オルタナティブデータには、①速報性、②網羅性、③接続可能性(他のデータベースと接続可能)という強みがある
・ オルタナティブデータには、①代表しない(標本に偏りがある)、②正解がない(政府統計との比較が必要)、③目的が違う(統計用に作られていない)という限界がある
・ オルタナティブデータで、政府統計をほぼ完全に捉えることができたとして、数週間の速報性が政策決定に役立つ場面は、金融政策や緊急時などの一部に限られる
・ 人流データの活用においては、経済学的な側面をどう絡めるかに苦労している
・ 統計のバイアスを理解するのに有用である可能性がある
・ オルタナティブデータは、ものによっては高価であり、研究費で購入できないものもある

(オルタナティブデータの特徴、分析例)

近年、技術革新やデジタル化が進む中で、利活用可能なデータが大きく増加した。日常生活においても、インターネット販売でのレコメンドシステムや生成 AI など、ビッグデータを分析、活用している事例は数多く存在する。経済分析においても、ビッグデータの利用は進んでおり、内閣府や日本銀行をはじめ、様々な機関が分析結果を公表している。

経済分析分野においてのビッグデータの活用は、政府統計の補完といった役割が多い。政府統計を代替するデータは、特にオルタナティブデータと呼ばれる。オルタナティブデータの特徴として、①詳細性（細かい）、②速報性（公表が速い）、③異質性（新しいデータを扱う）などがあると考えられる（表 1-2）。

表 1-2 オルタナティブデータの特徴と利用目的、分析例

データの特徴		分析に利用する目的	具体的な分析例
量が多い	細かい（頻度/濃度）	政府統計よりも高頻度なデータや粒度の細かいデータを用いて、経済の状況をより詳細に把握	- 家計簿データを用いた定額給付金の影響分析（内閣府、2023） - クレジットカードデータを用いた小売売上高の予測（BEA、2020）
	公表が速い	政府統計の公表前に経済の状況を把握	- 電力使用量データを用いた生産量予測（水門ほか、2021） - 新聞記事データを用いた景況感分析（SF連銀、2020）
質が異なる	新しいデータを扱う	分析可能なデータの幅を広げ、多面的に経済の状況を把握	- 決算説明会の音声データを用いた景況感分析（橋本、2021） - 衛星画像を用いた新興国のGDP予測（Luis、2022）

以下では、①～③について、特徴や利用目的、先行研究を紹介する。まず、①の詳細性については、政府統計は回答者負担を軽減する制約などからデータの収集に限りがあるが、オルタナティブデータは回答者の負担なく粒度の細かいデータを取得することが可能な場合がある。例えば、家計の収入、支出動向を把握することができる代表的な政府統計として、総務省統計局が公表する家計調査⁵があり、ある程度詳細なデータが公表されているが、家計簿アプリのデータは、家計の収入・支出がリアルタイムでかつ両者が一体的・網羅的に記録され、家計の収入、支出動向をより詳細に分析できる可能性を有する。内閣府（2023）は、家計簿アプリのデータが高頻度、高粒度である特徴を活用し、特別定額給付金が家計消費に与えた影響を定量的に分析した。

続いて、②の速報性については、経済分析において最も利活用が進んでいる分野であり、金融市場での関心も高い。例えば、コロナ禍において、家計の消費行動が大きく変動し、これが財政・金融政策にも強く影響したため、家計の消費行動をいち早く掴むことへの関心が高まった。米国商務省経済分析局（Bureau of Economic Analysis）は、クレジットカードの消費額の動きを公表しており、月次ベースで公表される政府統計（小売売上高など）に先行する指標として、エコノミストや市場参加者から注目されている。もちろん、クレジットカードの決済情報は、政府統計よりも詳細に品目を把握できるため、①の詳細性といった特徴も併せ持つ。他にも、工場の電力使用量のリアルタイムデータを用いて、企業の製造業生産活動を精緻に計測する手法も提案されている（水門・和泉・坂地・島田・松島、2021）。

また、これまではデータとして利活用ができない、もしくは難しかったものを分析の対象とする取り組みも進んでおり、本稿ではこれを③の異質性と分類する。政府統計は、基本的

⁵ <https://www.stat.go.jp/data/kakei/>

に数値データであり、経済分析においては、数値データを統計分析することが主流である。ただし、最近では、数値データ以外のデータも分析対象として利用が進んでいる。数値データ以外のデータとは、テキストデータや、音声データ、光データなどであり、それぞれ具体的な研究例を紹介する。これらの中で比較的活用事例が多くみられるものは、テキストデータである。例えば、サンフランシスコ連邦準備銀行（SF 連銀）は、日々の新聞記事のデータから、景気のセンチメントを分析し、Daily News Sentiment Index として公表している（図 1-1）⁶。なお、この指数は、日次で公表されるため、①の詳細性（高頻度）や②の速報性といった側面も持つ。また、テキストデータ以外では、決算説明会の音声データを用いた景況感分析（橋本，2021）のほか、衛星画像の光データを用いた研究（Martinez, 2022）などがある。Martinez (2021)は、衛星画像の光データを用いて、新興国の GDP 予測を行い、権威主義国家の経済成長率は政府の公式発表よりも弱い可能性を指摘し、英 Economist 誌に取り上げられるなど話題を呼んだ。

以上にみるように、オルタナティブデータの経済分析での利活用が様々な場面で進んできた。また、実務面での関心の高さから、特に速報性を持つオルタナティブデータの分析が盛んに行われている印象がある。ただし、経済を多面的に把握するという意味では、これまで分析があまりなされていなかった観点からの分析も重要である。そこで本稿は、テキスト分析に焦点を当て、経済分析分野では取り組み事例が少ない分析を行い、分析の幅を広げ、新たな示唆を得ることを目指す。

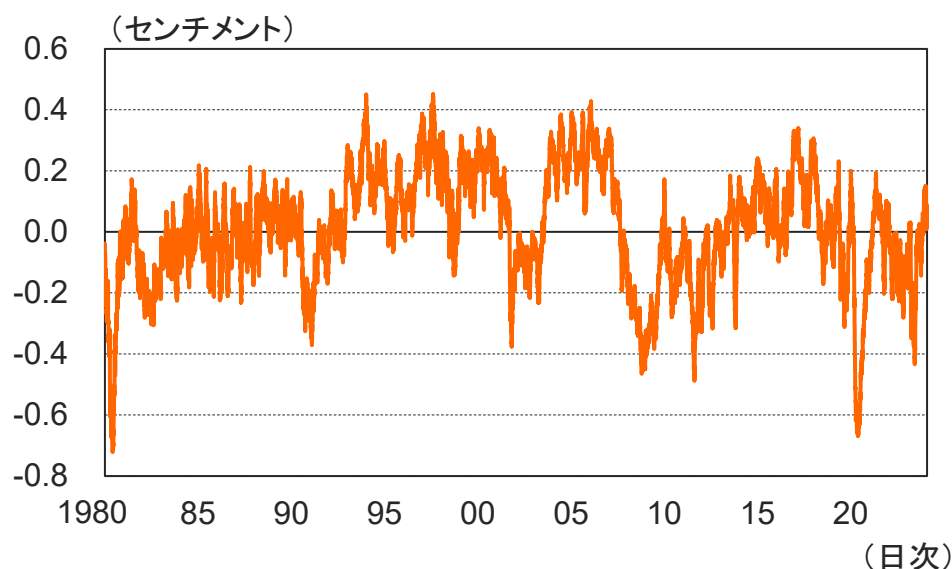


図 1-1 Daily News Sentiment Index

（金融政策とテキスト分析）

本稿では、テキスト分析の対象として、米国の連邦準備制度が公表する地区連銀景況報告

⁶ 詳細は Buckman, Shelby R., et al.(2020)、データは <https://www.frbsf.org/research-and-insights/data-and-indicators/daily-news-sentiment-index/>

（詳細は第2章）を用いるが、景気の分析にあたって中央銀行が公表する文書を用いる意義を整理する。先述の通り、経済分析においてオルタナティブデータを活用する動きは活発化しており、中央銀行も金融政策の判断にあたって様々なオルタナティブデータを分析しているが、2000年代に入ってから以降、各国中銀が非伝統的金融政策にシフトし、中銀のコミュニケーションへの関心が高まる中、中銀の公表するテキスト資料自体を分析する必要性も強まっている。実際、日本銀行の白川総裁と黒田総裁の定例記者会見文書に対して、統計的自然言語処理のトピックモデルを適用することで意味論的分析を行い、日銀のコミュニケーションの変化について明らかにした研究（Keida & Takeda, 2019）もある。本稿は、一義的には、テキスト分析を用いて質の高い米景気後退確率を推計する方法の提案を目的としているが、分析データが中央銀行の公式文書であるため、中銀の景気認識を定量的に捉える試みと考えることもできる。

2. データおよび分析方法

(1) 使用データ

本稿では、テキスト分析の対象として、米国の連邦準備制度が公表する地区連銀景況報告（Beige Book）⁷を用いる。Beige Book は、連邦公開市場委員会（FOMC）の基礎資料として用いられ、米国の中央銀行が政策決定を行ううえでの判断材料となる。公表は、FOMC の 2 週間前の水曜日とされており、FOMC の開催にあわせて年に 8 回公表される（2003 年のみ 7 回公表）。Beige Book は、1970 年から公表されており、1996 年 10 月分以降は連邦準備制度理事会（FRB）のホームページ⁸、それ以前はミネアポリス連銀のホームページ⁹で公表されている。

Beige Book の構成は何度か変更されているが、内容は、17 の地区と米国全体について、総合の景況感のほか、雇用や物価事情などを文書形式で記載したものとなっている。具体的には、表 2-1 のような記載となっている。なお、単語数は、24 年 1 月公表分で 15,932 語となっており、各回とも概ね同水準となっている。

表 2-1 Beige Book の内容(23 年 11 月)

全体として経済活動は減速した<総括> (前回報告(10/18公表)では横ばい)
今後6-12カ月の経済の見通しは低下した<総括> (前回報告では横ばい、もしくはわずかに軟化)
地域別にみると、経済は、4地域が緩やかに成長し、2地域がほぼ横ばい、6地域がわずかに縮小した<総括>
節約のため、高価な商品避けたり、ルームシェアを行う動きがみられた<カンザスシティ>
消費者は一層価格に敏感になっている<フィラデルフィア>
建設業では、高金利のために遅延、キャンセルとなった案件がある<セントルイス>
需要の減少などにより価格支配力が弱まった<クリーブランド>

本稿では、FRB のホームページで入手可能な、1996 年 10 月分から 2023 年 11 月分までのデータを用いた。基本的に PDF でダウンロードしたものを用いているが、2003 年 1 月分から 2008 年 1 月分までは HTML 版のみの公表であるため、PDF 化して利用した。

⁷ <https://www.federalreserve.gov/monetarypolicy/publications/beige-book-default.htm>

⁸ <https://www.federalreserve.gov/monetarypolicy/beige-book-archive.htm>

⁹ <https://www.minneapolisfed.org/region-and-community/regional-economic-indicators/beige-book-archive>

（２） テキスト分析の手法

（先行研究）

文書に対して統計的手法を用いて分析を行う試みは 2000 年初頭から行われており、以降、多変量解析を用いて言葉や文章を分類しようとする試みが広がった（Renz & Franke (2003) や樋口(2004)など）。小木(2015)によれば、「テキストマイニング」に関する論文数や書籍数などは、2000 年初以降、増加基調にある。テキスト分析を経済分野に応用する動きも前述の Daily News Sentiment Index などのように広がっており、Beige Book を用いたテキスト分析も米国を中心に行われている。Armesto, Hernández-Murillo, Owyang & Piger (2009)は、Beige Book の出現単語を楽観的（Optimism）と悲観的（Pessimism）にテキスト分析の手法を用いて分類（後述のセンチメント分析）し、出現頻度の差をスコアとして計算し、これと GDP や雇用統計の関係を分析したほか、Sadique, Shibley, et al.(2013)も同様の手法を用いて、Beige Book のセンチメントと景気の間関係を考察している。コロナ禍以降では、2022 年にセントルイス連銀が Beige Book における特定の単語の出現数を用いて、雇用やインフレ¹⁰、サプライチェーン圧力¹¹を測る指数を作成し、これが消費者物価などの既存の政府統計等と似た動きをすることを示した。

（分析のアプローチ）

本稿では、前処理と 3 つの段階に分けて分析を行った。前処理では、Beige Book の文書を、テキスト分析しやすいように加工した。第 1 段階では、Sadique, In, Veeraraghavan & Wachtel (2013)などの先行研究と同様に、出現単語のセンチメント分析を行った。第 2 段階では、Beige Book の文書毎に、文書と文書の類似度を計算した。第 3 段階では、文書の特徴を用い、機械学習の手法によって景気後退確率を試算するモデルを作成した。

（データの加工）

前処理において、Beige Book の文書を、テキスト分析しやすいように加工する。文書は全て pdf 形式であるため、統計ソフト R（以下、R）と、いくつかのパッケージを用いて、pdf から文字データを抽出した。この際、Beige Book の一部の資料について、pdf 上で閲覧する上では問題が生じないものの、データ上で文字化けが起きている資料が存在するため、復号処理を行っている（詳細は、補論 1 を参照）。また、R とそのパッケージ（分析ツール）の問題により、全ての単語を正確に抽出できていないわけではないが、分析上大きな問題はないと判断した。

¹⁰ 詳細は Gascon & Werner (2022)

¹¹ 詳細は Kliesen & Werner (2022)

続いて、抽出した文字データを、語を構成する最小の意味のある単位に分割する（形態素解析）。日本語のテキストでは、形態素解析において、分かち書きの処理（図 2-1）が必要であるものの、本稿で扱う文章は全て英語であるため、分かち書きの処理は行っていない。

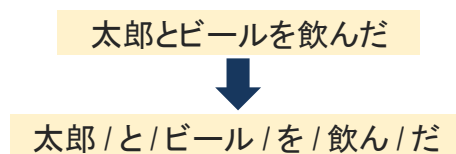


図 2-1 分かち書きの例

他方、英語は、主として語形変化によって、性、数・格などの文法的関係を示す言語（屈折語）であるとされる（保田、2006）。そのため、出現単語について、語形変化を統一する見出し語化を行った。見出し語化の手法はいくつかあるが、本稿では、単語を全て原形に戻す処理を行った。例えば、'drive', 'drove', 'driven', 'drives', 'driving' という単語は、全て 'drive' に置換する（詳細は補論 2 を参照）。

（センチメント分析）

第 1 段階では、Beige Book の出現単語について、センチメント分析を行う。センチメント分析とは、テキストの内容に対する評価を、構成する単語をもとに判定する手法のことであり、テキストを計量的に分析する手法の一つである（石田、2020）。センチメント分析は、計算が比較的容易であることなどから比較的先行事例も多い。本稿では、Sadique et al.(2013) に倣い、あらかじめ単語に感情（positive, negative）を紐づけた「辞書」を用いて、Beige Book の出現単語を positive と negative に分類し、出現頻度の差を文書のセンチメントとした。分類は、例えば表 2-2 のように行われる。

表 2-2 センチメントの分類

Positive	頻度	Negative	頻度
good	5	bad	2
increase	3	decrease	3
happy	4	sad	1
（合計）	12	（合計）	6

文書のセンチメント（Sentiment）は、以下の式で計算される。ただし、Positive は、文書における前向きなコメントの数で、Negative は、後向きなコメントの数を表す。

$$Sentiment = \frac{Positive - Negative}{Positive + Negative}$$

本稿の「辞書」は、Liu, Bing が開発した「Bing」(詳細は Liu (2022)) と、Nielsen (2011) が開発した「AFINN」の 2 種類を用いた。「AFINN」は、Positive 度合いと Negative 度合いをそれぞれ 5 段階のスコア (計 10 段階) で判定できるが、本稿では単純平均により Positive と Negative の 2 段階としてセンチメント分析を計算した。なお、スコアに基づく加重平均の計算も行ったが、結果に大きな差はみられなかった。

センチメント分析は、計算が比較的容易であるほか、解釈もしやすいため、先行例が多いものの、分析対象の単語が分類に用いる辞書に依存してしまう、単語毎の重要度が一律に決まってしまうなどの課題もあるため、以降では文書全体の情報を用いた分析を行う。

(文書間の類似度を計算)

第 2 段階では、Beige Book の公表毎に、文書と文書の類似度を計算する。文書の類似度の計算は、テキスト分野では主要な手法の一つで、身近なところでは検索エンジンの処理で用いられているほか、最近では生成 AI などにも応用されている。経済分析での先行研究は、日本銀行が発表する総裁定例記者会見の文書の分類を試みた分析 (Keida & Takeda, 2017) などがあるが、前述のセンチメント分析と比較すると分析例は少ない。文書の類似度は、文書を一定のルールで数値化 (ベクトル化) し、それらの数値 (ベクトル) の近さを計算することで求められる。例えば、bag of words の手法では、文書の出現単語数に注目し、出現頻度が似ている文書同士の類似度は高いとする。以下に bag of words の具体例を示す。

表 2-3 では、①「暖かい日が続く」、②「今日の東京は暖かい」、③「銀行に行った」、という 3 つの文章を、形態素単位に分解し、出現頻度をみている。①と②は、「暖かい」という共通の単語がある一方、①と③は共通の単語がなく、①は、③よりも②に似ていると考えられる。

表 2-3 bag of words

	暖かい	日	が	続く	今日の	東京	は	銀行	に	行く	た
①	1	1	1	1	0	0	0	0	0	0	0
②	1	0	0	0	1	1	1	0	0	0	0
③	0	0	0	0	0	0	0	1	1	1	1

類似度合いを定量的に計算する手法としては、難波 (2020) によれば、Jaccard 係数、Dice 係数、Simpson 係数といった集合間の類似度尺度を用いる方法のほか、コサイン類似度の値を計算することが基本的な手法とされる。本稿では、分析例の多いコサイン類似度を用いる。コサイン類似度は、文章をベクトル表記することにより計算できるが、文章は、出現単語を

ベクトルの次元、出現頻度をベクトルの次元毎の大きさとするすることで、ベクトル表示することができる。例えば、表 2-3 では、出現単語が 12 種類であるため、①～③の文章を 12 次元のベクトルで表すことができる。コサイン類似度とは、文章をベクトル化し、そのコサイン値を計算したものであり、ベクトル $\vec{a}$ と $\vec{b}$ のコサイン値は、内積 $\vec{a} \cdot \vec{b}$ を用いて以下の式で表される。ただし、 $|\vec{a}|$ は $\vec{a}$ の大きさを表す。

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|}$$

表 2-4 について具体的に計算すると、①と②は 0.22、①と③は、0 となるため、①と②の方が、①と③よりも類似度が高いことがわかる。

bag of words をやや発展させた方法もある。出現頻度を文書全体の単語数で除した TF (term frequency) 値と、出現単語のレア度を示す IDF (inverse document frequency) も考慮した TF・IDF 値があり、いずれも単語の重要度を測る手法である。これらについて、具体例を用いて簡単に紹介する。表 2-4a は、文章 X、Y、Z について、「暖かい」「東京」「銀行」の出現回数を記載している。

表 2-4a 文書、単語の出現頻度

	暖かい	東京	銀行
X	7	4	0
Y	8	5	1
Z	0	0	10

これらについて、まず、TF 値の計算を行う。例えば、X の「暖かい」という単語の TF 値は、

$$\begin{aligned}
 TF(\text{暖かい}, X) &= \frac{X \text{ での「暖かい」の出現頻度}}{X \text{ の単語数}} \\
 &= \frac{7}{(7 + 4 + 0)} \\
 &= 0.64
 \end{aligned}$$

となる。これらを各文章、各単語について計算すると、表 2-4b のようになる。

表 2-4b TF 値

	暖かい	東京	銀行
X	0.64	0.36	0
Y	0.57	0.36	0.07
Z	0	0	1.00

続いて、IDF 値は、単語 t に対して、次式で計算される。

$$IDF(t) = \log \frac{\text{文書数}}{t \text{ が含まれる文書数}} + 1$$

式の通り、特定の文書のみに出現する単語ほど値が大きくなる。「暖かい」「東京」「銀行」について、IDF 値を求めると、表 2-4c のようになる。「銀行」という単語は文章 Z にしか含まれないため、単語のレア度を示す IDF 値が高くなっていることがわかる。

表 2-4c IDF 値

暖かい	東京	銀行
1.41	1.00	2.10

最後に、各単語の $TF \cdot IDF$ 値は、

$$TF \cdot IDF = TF \text{ 値} \times IDF \text{ 値}$$

で求めることができる。これを各文書、各単語について計算すると、表 2-4d のようになる。

表 2-4d TF・IDF 値

	暖かい	東京	銀行
X	0.89	0.36	0
Y	0.80	0.36	0.15
Z	0	0	2.10

本稿では、TF 値と TF-IDF 値の 2 種類を、Beige Book の文書毎に、全ての単語に対して計算した。そして、これらを要素とするベクトルを作成し、前述のコサイン類似度を求め、文書同士の類似度を計算した。

（景気後退確率の試算）

第 3 段階では、文書の特徴を用い、機械学習の手法によって景気後退確率を試算する。分類モデルは、大規模データに対して、非線形の関係进行分析できるランダムフォレストを用いた。ランダムフォレストは、Breiman (2001)によって提示された機械学習の手法であり、同論文は 12 万回（24 年 1 月時点）を超える引用がなされるなど、機械学習分野において主要な手法の一つとなっている。ランダムフォレストは、図 2-2 のように、複数の決定木を用いて森を構成して識別などを行う機械学習手法である（波部，2016）。

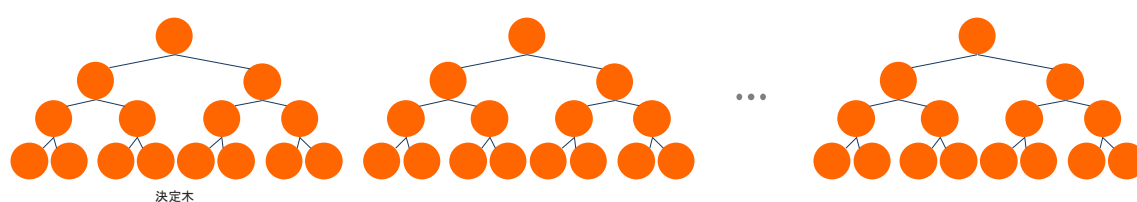


図 2-2 ランダムフォレスト

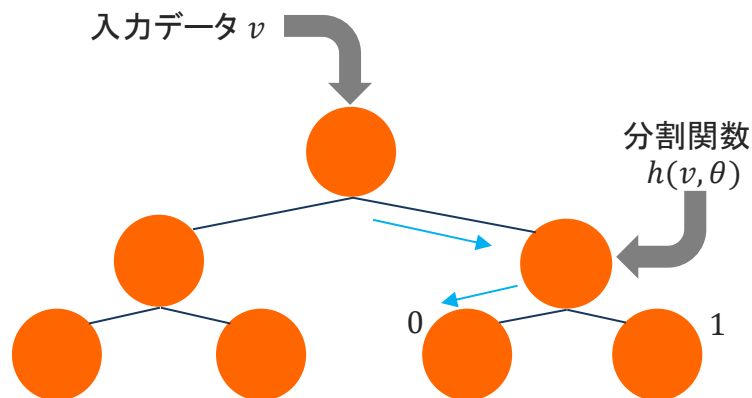


図 2-3 決定木

決定木は、図 2-3 に示すように、結節点において、どちらに分類されるかを分割関数によって決定される構造を持っている。本稿では、(2)で求めた文書の TF 値のベクトルを説明変数として、各文書の公表時が景気後退であるかどうかを判定する分類タスクを行う（TF・IDF 値を用いない理由は後述）。分類タスクとは、例えば銀行の顧客データを用いて、顧客が一定期間内に破産するかどうかの二項分類を行うような分析のことである。ランダムフォレストを用いた先行研究は数多くあり、例えば、金（2007）はランダムフォレストにより文章の書き手を特定する手法を提案しているほか、内閣府（2018）もランダムフォレストを用いて転職前後の賃金変化と関係の深い変数の特定を行っている。

本分析の被説明変数となる景気後退判定データは、全米経済研究所（NBER）¹²が公表するものを用いる。頻度については、Beige Book は 1 ヶ月半に一度公表され、NBER の景気後退判定は月毎に行われているものの、景気の主要指標である GDP は四半期毎の公表であるため、景気後退判定の推計は四半期毎に行った。さらに、使用する単語は単語の出現回数によってサンプリングを行っている。対象期間は、96 年第 4 四半期から 23 年第 4 四半期までとし、96 年第 4 四半期から 19 年第 4 四半期までを訓練（train）データ、20 年第 1 四半期から 23 年第 4 四半期までを検証（test）データとしている。ただし、過学習が行われていないかを確認するため、train データと test データをランダムに抽出して予測精度に差異がないかどうかの検証も行った（詳細は、補論 2 を参照）。また、ランダムフォレストのパラメータ（木の数など）は、グリッドサーチと呼ばれる機械学習の手法を用いて適切なものを設定した。

¹² <https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions>

3. 結果

(センチメント分析)

センチメントは、出現単語を、positive と negative に分類し、その差を取ったものになる。Beige Book に対するセンチメント分析の結果は、図 3-1 の通りである。景気後退期に Beige Book のセンチメントは悪化しており、Sadique et al.(2013)の先行研究と同様に、センチメントと景気変動の連動が確認できる。また複数の辞書で試しても似た動きとなっており、頑健性が示唆される。

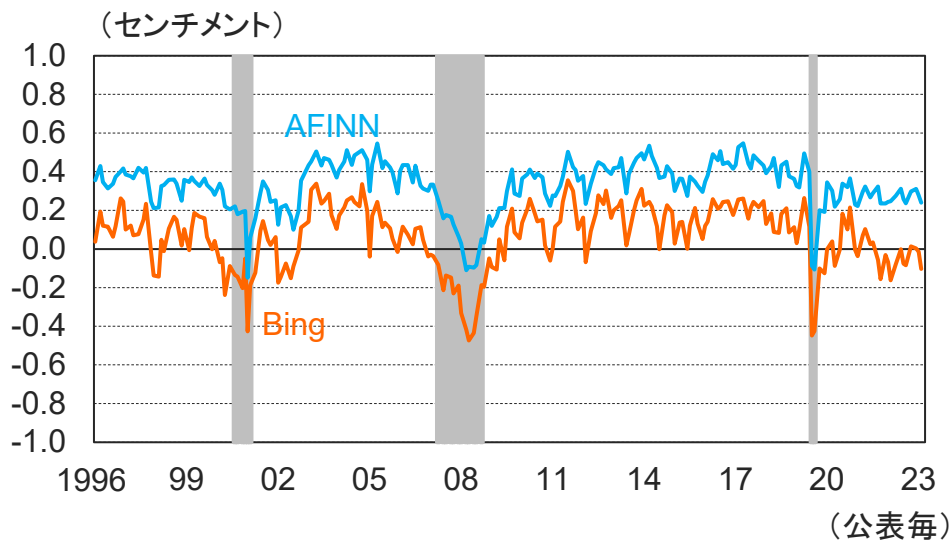


図 3-1 センチメントと景気後退

また、景況感を示す代表的な指数である、ISM 製造業景況指数と比較すると、図 3-2 のようになり、動きがかなり似通っていることがわかる。なお、ISM 製造業景況指数は、米国の景況感を示す指標として最も注目される統計の一つであり、全米供給管理協会 (Institute for Supply Management) が毎月公表している。本稿では、LSEG Datastream からデータを入手し、Beige Book が公表されるタイミングに近いものを抽出してグラフ化した。

センチメント分析の結果から、Beige Book には、景気動向を把握できる要素が含まれていることがわかる。ただし、先行性という意味では ISM 景況指数に劣るため、これだけで景気予測を行う利点がやや薄いほか、第 2 章で指摘した通り、センチメントの計算に利用可能な単語は、使用した辞書に含まれるものに限定されるため、Beige Book の持つ景気予測成分を十分に発揮できていない可能性もある。

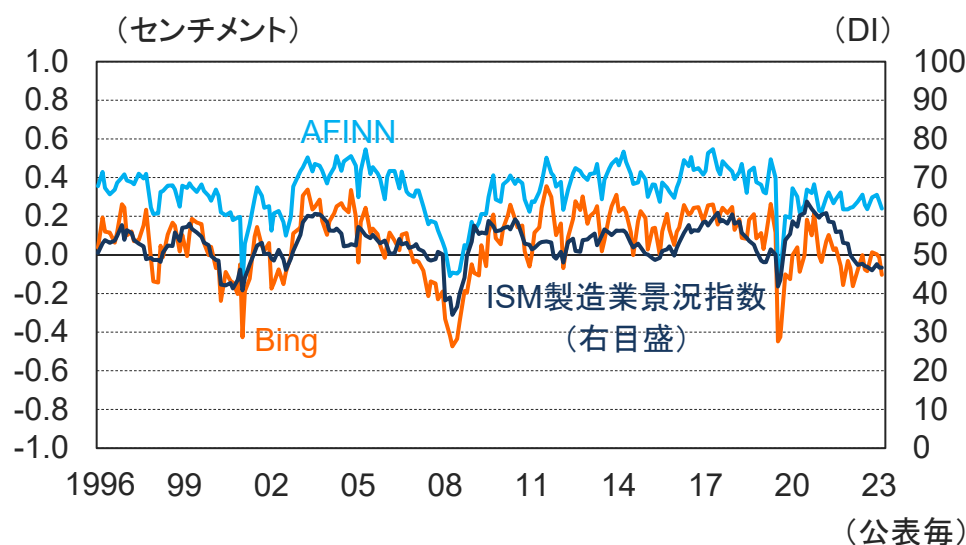


図 3-2 センチメントと ISM 製造業景況指数

(文書間の類似度)

続いて、Beige Book の文書同士の類似度を確認する。類似度は、TF (出現単語の頻度) 値、及び TF・IDF (出現単語の頻度に単語のレア度を考慮したもの) を文書ごとにベクトル表記したのに対して、ベクトルのコサイン類似度をそれぞれ計算する。コサイン類似度を行列表記にした、類似度行列は、図 3-3 の通りである。

縦横の軸は Beige Book が公表された日次を表し、交点の色が類似度の大小を示す。例えば、横軸が 07 年頃、縦軸が 10 年頃の交点は紫色となっており、文書の類似度が相対的に小さいことがわかる。図は 45 度線で対称となっており、45 度線上は同じ文書の比較となるため、類似度が最大となっている。時系列にみると、公表日次が近いものほど黄色、遠いものほど紫色になる傾向、つまり公表日次が近いものは類似度が高い傾向があり、Beige Book の内容や記載ぶりは長期で緩やかに変化していることがわかる。また、図中には縦 (横) 3 本の紫色の線が入っており、ちょうど景気後退局面と重なる。これは景気後退局面に公表された Beige Book と、それ以外の局面に公表されたものとの文章の類似度が小さいことを意味し、景気後退局面には、特徴のある単語の出現頻度が高まることがわかる。このことから、TF 値に注目すれば、景気変動を捉え得る可能性が示唆される。

続いて、単語のレア度を考慮した TF・IDF 値の結果 (図 3-4) をみると、多くの局面で紫色となっており、文書毎に特殊な単語が相応に出現していたことがわかる。

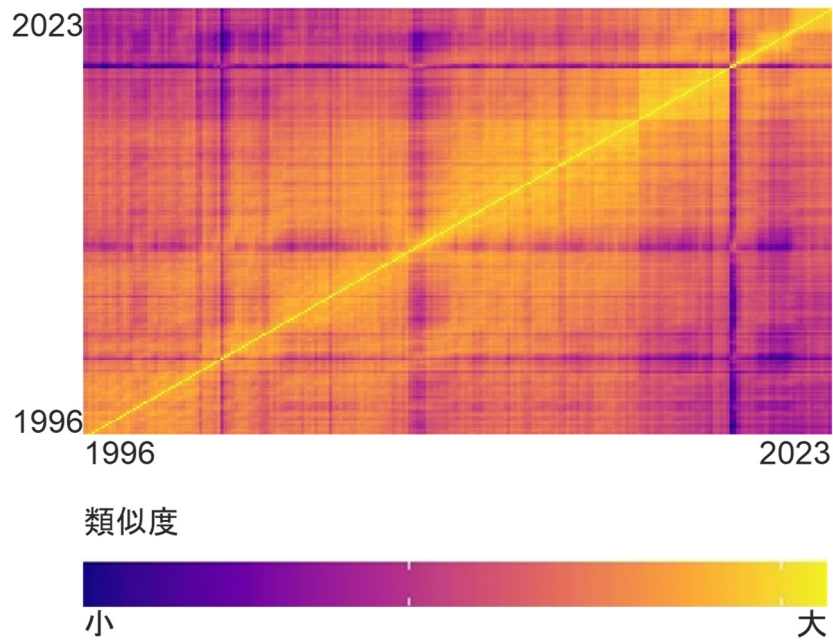


図 3-3 文書の類似度(TF 値)

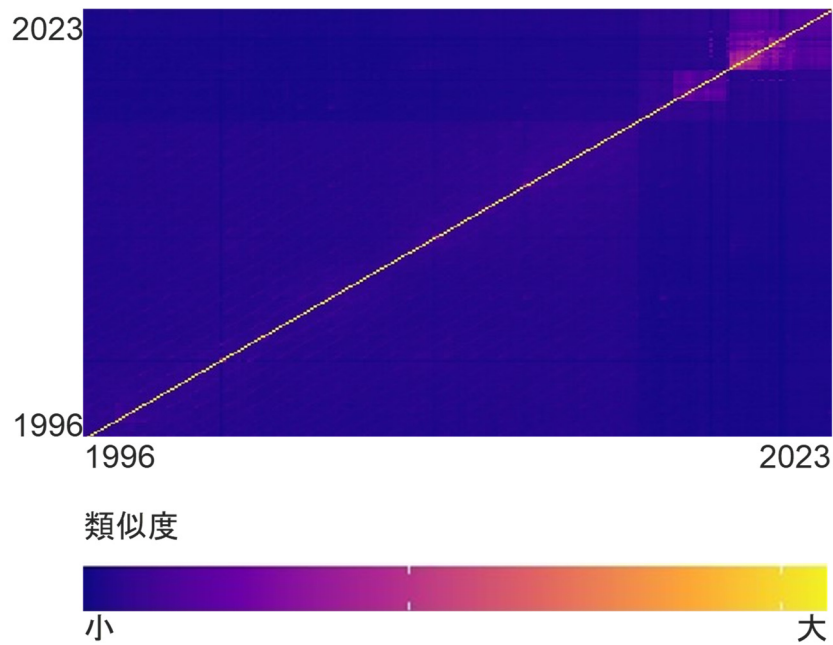


図 3-4 文書の類似度(TF-IDF 値)

TF・IDF 値では文書間の類似度が大きく低下してしまう要因をもう少し検討する。Beige Book において TF・IDF 値が大きい単語をみると、ハリケーンの単語が多いことがわかる。実際、Beige Book ではハリケーンへの言及が目立ち、それらは、「ハリケーン (hurricane)」という単語ではなく、当時の通称 (2012 年の「サンディ (Sandy)」や 2017 年の「ハービー (Harvey)」など) で記載されることが多い。このような単語は、特定の文書には数多く出現するものの、その他の文書では全くみられないため、IDF 値が高くなりやすい。また、18～19 年はトランプ政権の影響などで「関税 (tariff)」の出現が目立ったほか、コロナ以降は「コロナ (covid)」や「パンデミック (pandemic)」「変異株 (variant)」などの単語が急増した。特に、コロナ関連の語句はコロナ禍で数多く出現したため、コロナ禍の文書間では類似性が多少確認できる。ただし、長期の景気変動を考える上では、TF・IDF 値は文書毎の特殊な単語に注目してしまうため使用が難しく、前述の TF 値を採用した方が良いと考えられる。

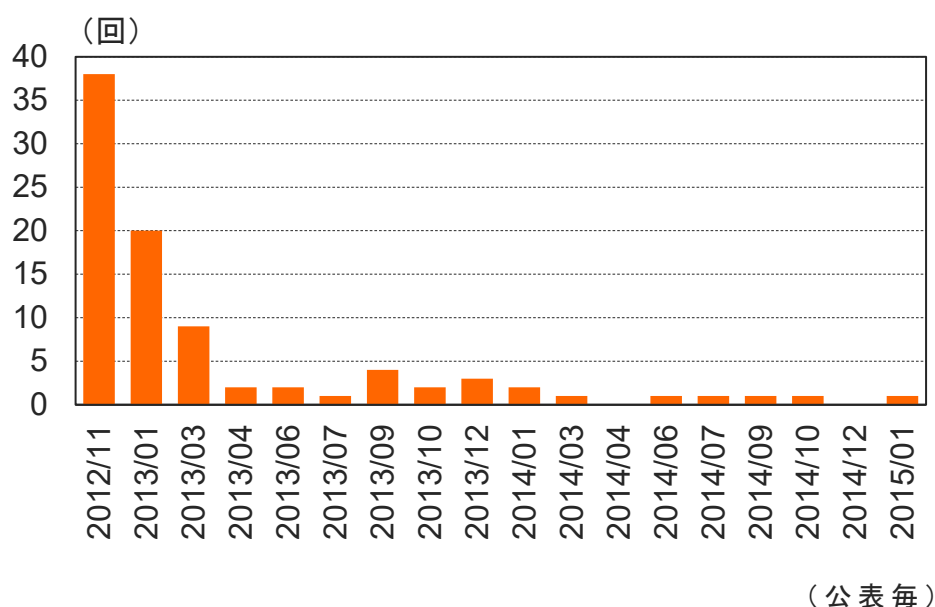


図 3-5a 単語の出現頻度(sandy)

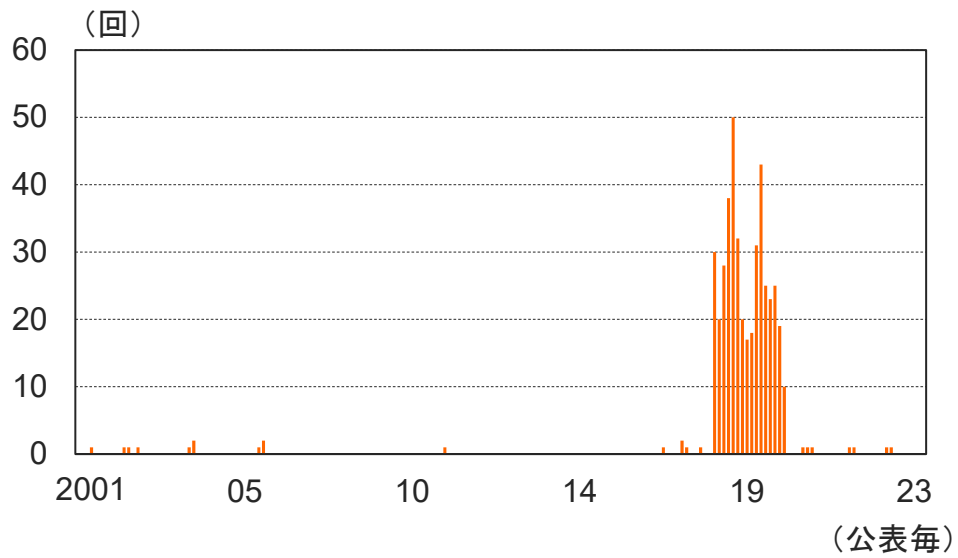


図 3-5b 単語の出現頻度 (tariff)

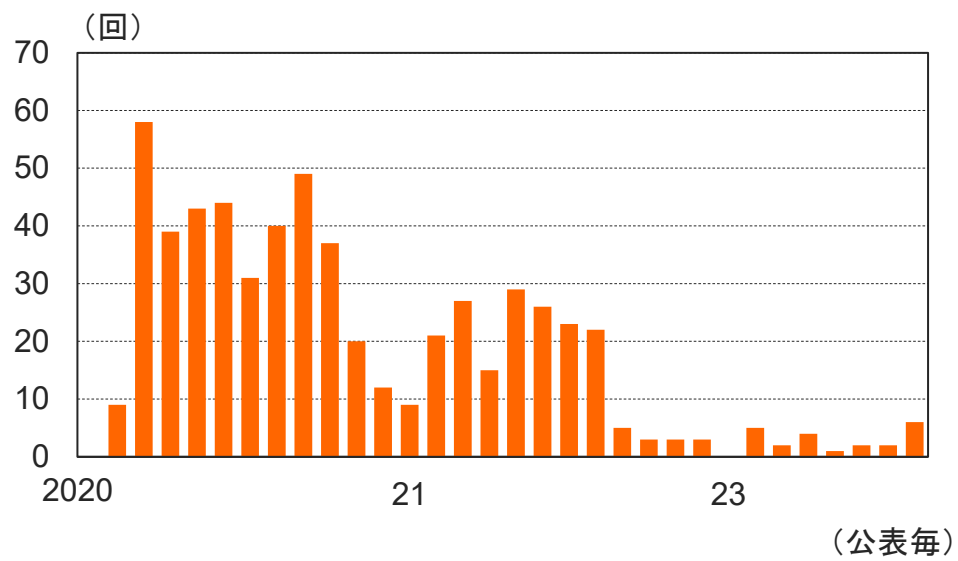


図 3-5c 単語の出現頻度 (covid)

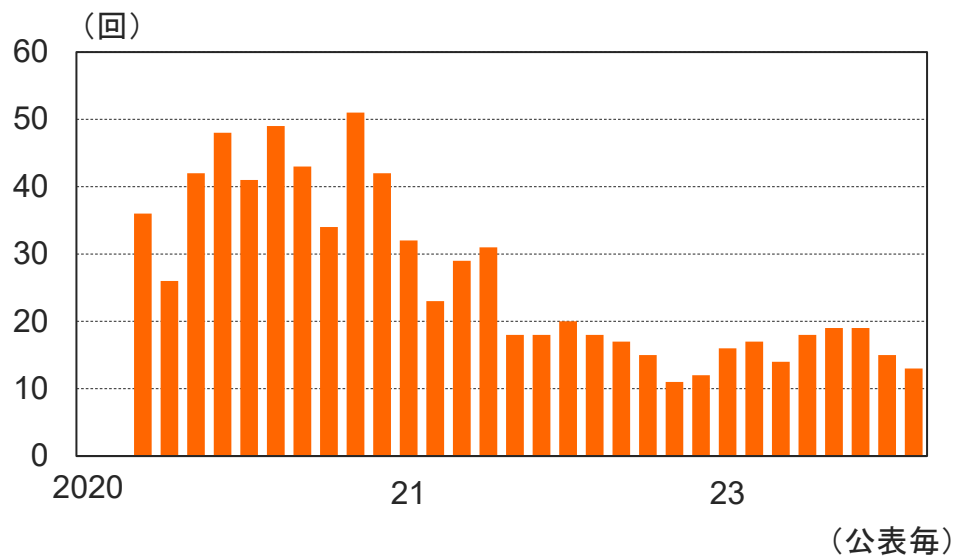


図 3-5d 単語の出現頻度(pandemic)

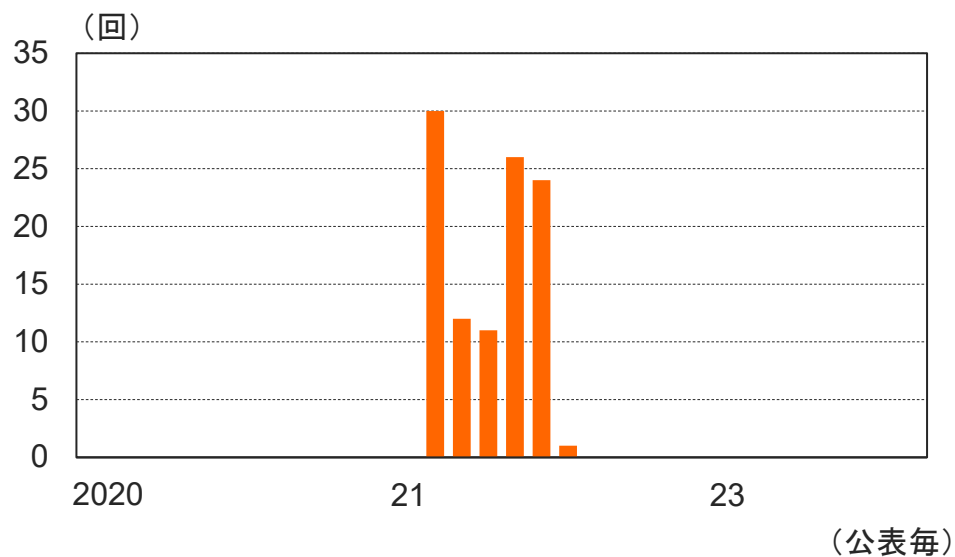


図 3-5e 単語の出現頻度 (variant)

(景気後退確率)

文書の類似度において、出現頻度 (TF 値) でみると、景気後退期には文書間の類似度が大きくなる傾向が確認できたが、レア度も含めたもの (TF・IDF 値) でみると、特殊な単語の影響を大きく受けるため、景気との連動がみられないことが分かった。通常の文書分類においては TF・IDF 値の方を採用する場合も多いものの、本稿では、景気後退確率の試算において TF 値を使用した。特徴量 (説明変数) ベクトルは、出現回数が 100 回以上のものを抽出した。特徴量ベクトルの数は約 530 個になった。

景気後退確率の試算結果は図 3-6 のようになる。なお、ランダムフォレストのパラメータは、グリッドサーチと呼ばれる機械学習の手法により、決定木の数 (ntree) は 500、決定木を作成する際に用いる特徴量 (単語) の数 (mtry) は 160 とした。

景気後退確率の推計においては、19 年までのデータでモデルを作成し、そのモデルを用いて 20 年以降の推計を行っている。つまり、コロナ禍で特殊な単語 (「コロナ (covid)」や「パンデミック (pandemic)」 「変異株 (variant)」 など) はモデルでは全く利用していない。

結果をみると、20 年初の景気後退局面では、景気後退確率が大きく上昇したほか、22 年末に先行きの景気懸念が強まった瞬間には、小幅に景気後退確率が上昇している。ただし、その後は、幾分か景気後退確率が低下しており、23 年の想定外に堅調な米景気を反映したものになっている。

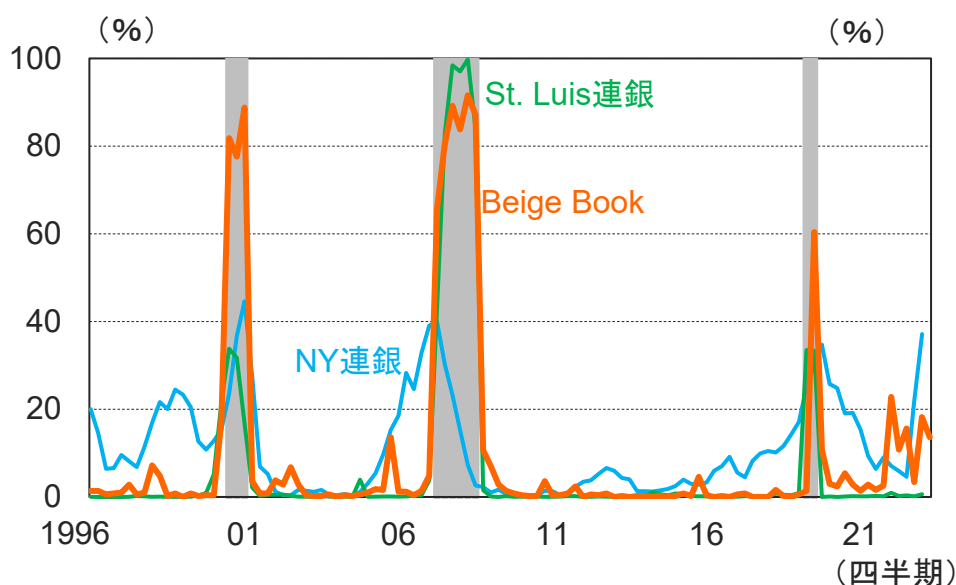


図 3-6 景気後退確率

また、他の景気後退確率モデルと比較すると、セントルイス連銀が公表する¹³、実体経済のデータから推計するモデルに近い動きとなっているが、コロナ禍での判定は本モデルが上回った。また、セントルイス連銀のモデルは、実体経済のデータが揃うまで計算できないため、先行性の面でやや不利でもある。続いて、NY 連銀が長短金利差から推計するモデル¹⁴は、2000 年初頭やリーマンショック前ですら 40%程度にとどまっており、精度にやや難点がある。また、この差を統計のクセと解釈すると、30%を超えると景気後退判定の上で危険水域とみなすことができ、23 年末には、NY 連銀のモデルが 40%近くにまで上昇したため、これをもとに景気後退への懸念を指摘する見方もみられたが、実際の景気は堅調だった。NY 連銀のモデルに採用されている長短金利差について、23 年の状況を簡単にまとめてお

¹³ 詳細は Chauvet & Piger (2008).、データは <https://fred.stlouisfed.org/series/RECPROUSM156N>

¹⁴ https://www.newyorkfed.org/research/capital_markets/ycfaq#/overview

くと、コロナ禍から景気が急回復する中でインフレが深刻化し、FRB が急速な利上げを行ったため、短期金利は急上昇したものの、先行きの景気不安から長期金利は短期に比べれば緩やかな上昇となり、長短金利差は逆転し、その差は 1980 年以来の大きさとなっていた。NY 連銀のモデルは、この差を用いて景気後退確率を計算するため、リーマンショック並みの景気後退確率を示したが、実際の景気は、懸念されたようには悪化せず、むしろ 22 年を上回る成長が実現した。

これらを踏まえると、Beige Book を用いた景気後退確率の結果は、先行性と精度の面でバランスよく実体経済をとらえたものになっていると考えられる（表 3-1）。

表 3-1 景気後退確率モデルの比較

	NY連銀 のモデル	St. Luis連銀 のモデル	今回のモデル
データ頻度	月次	月次	1ヵ月半
時期	月央に12ヵ月先 が更新	月初に2ヵ月前 のものが更新	FOMCの 2週間前
使用データ	長短金利差	実体経済の データ	Beige Bookの テキストデータ
先行性	◎	△	○
精度	△	○	○

また、本モデルにおいて、推定への貢献度が高い単語は図 3-7 の通りである。単語リストをみると、「weaken（弱まる）」「reduce（減らす）」「close（縮小する）」「cut（切る）」など、景気悪化時に増えると示唆される単語が上位に並んでおり、これらの単語の出現頻度が高まると、景気後退確率が上昇すると考えられる（図 3-8）。もっとも、これはあくまで推定への貢献度が高い単語であり、景気後退でない場合に出現する単語も重要と判定されることには注意が必要である。実際、推定への貢献度が高い単語の上位 6 番目に「optimistic（楽観的な）」があるが、図 3-9 の通り、景気後退でない場面での出現頻度が高い。また、「weak（弱い）」は上位 16 番目に位置し、「weaken」と大きな差があることも興味深い。図 3-10 で推移をみると、景気後退期に出現回数が増加する傾向が確認できるものの、「weaken」と比較して、動きが緩やかであることがわかる。つまり、景気を敏感にとらえる単語としては、状態を表す「weak」よりも変化を表す「weaken」の方が適切である可能性がある。

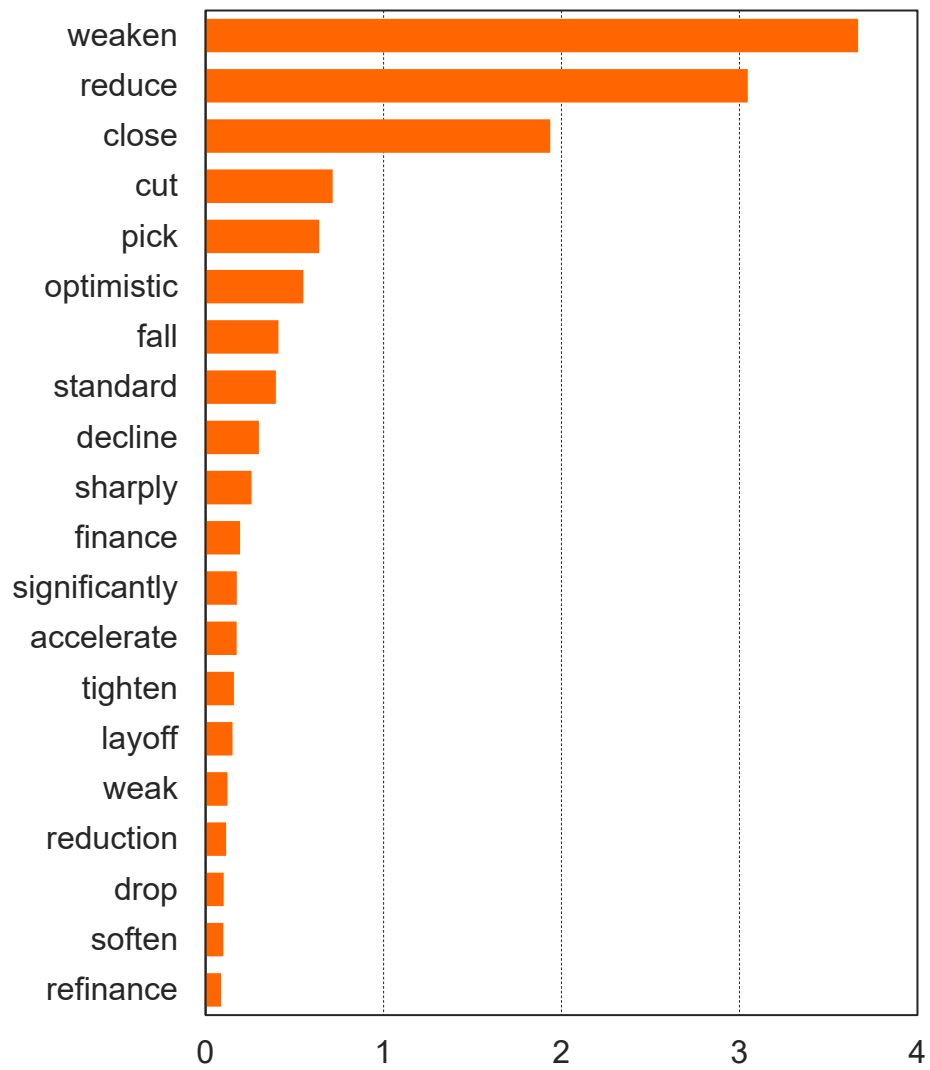


図 3-7 貢献度が高い単語(上位 20 語)

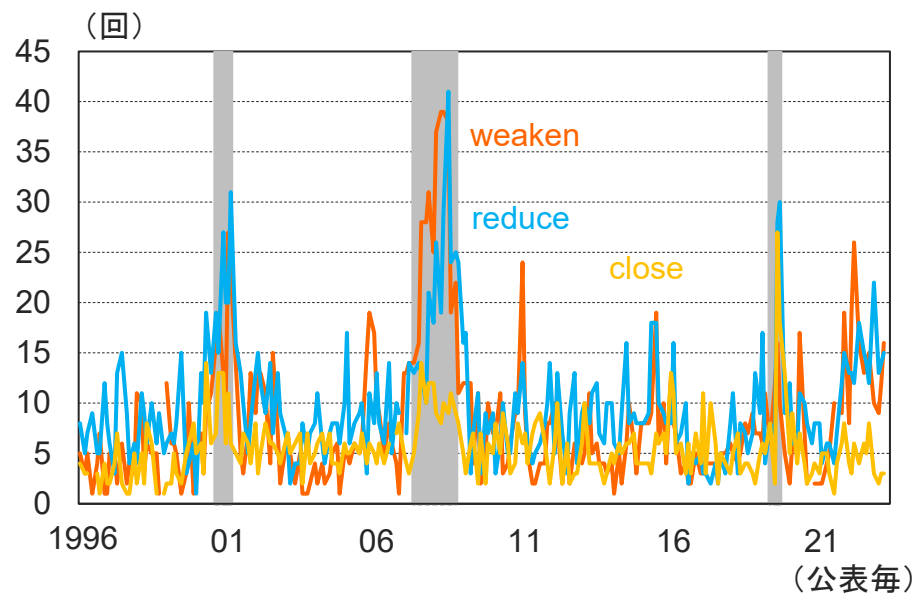


図 3-8 貢献度が高い単語の推移

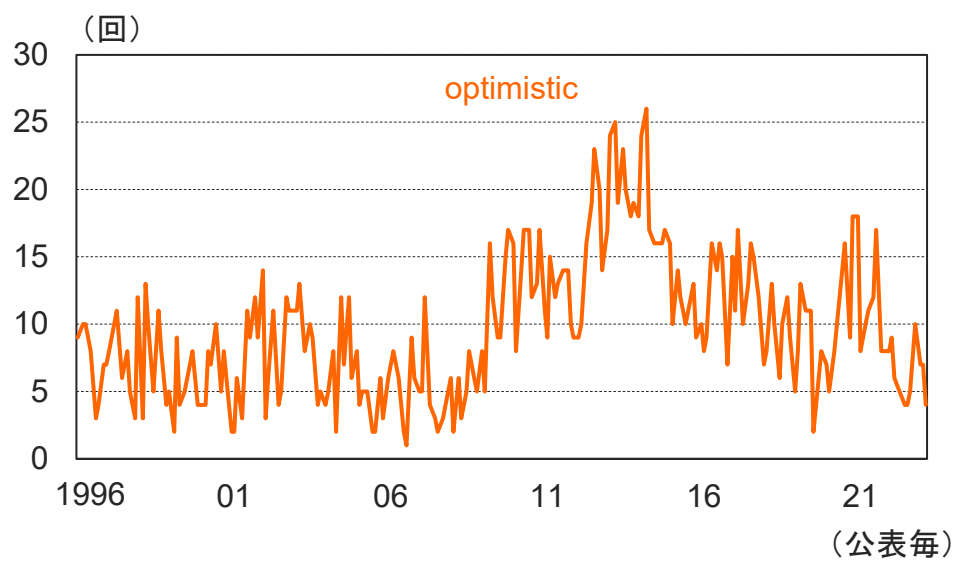


図 3-9 optimistic の推移

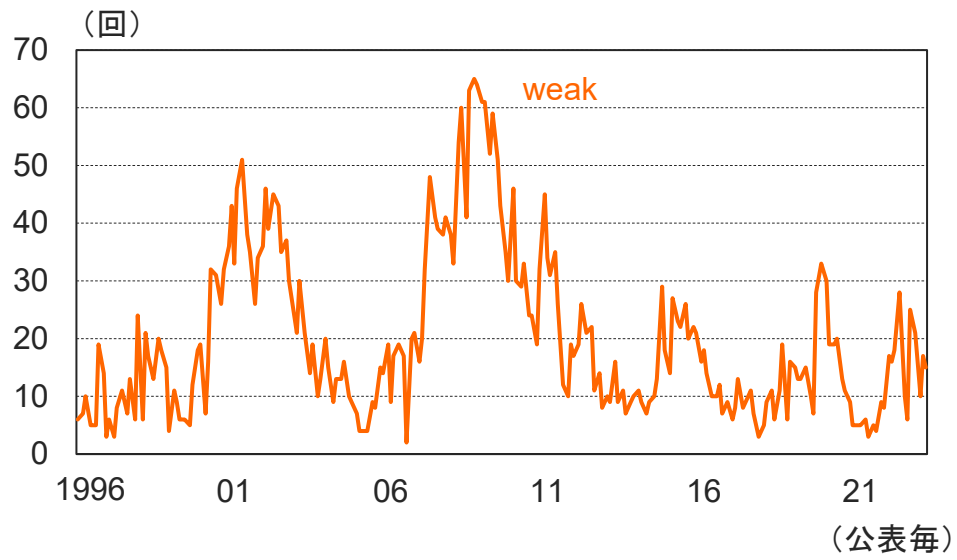


図 3-10 weak の推移

4. ディスカッション、結論

本稿の分析により、コロナ禍以降においても、先行研究と同様に Beige Book のセンチメントは景気と連動することが確認された。さらに、文書間の類似度を、単語の出現頻度をもとに計算したところ、景気後退局面に公表された Beige Book は、そうでない局面のものと文書の類似度が低いという発見が得られた。他方で、単語のレア度も考慮して類似度を計算した場合、いずれの文書間でも類似度が低く、Beige Book には、ハリケーンの通称やコロナ関連語句など、特定の期間にのみ急増する単語が多く存在することがわかった。これらの発見をもとに、単語の出現頻度に注目し、機械学習の手法の一つであるランダムフォレストを用いて景気後退確率を推定するモデルを作成したところ、米国の景気動向を適切に捉えたモデルを得ることができた。既存のモデルと比較すると、先行性と精度の両面で質の高いモデルとなった。また、数値データではなく、文書を用いて景気後退確率の推定を行ったため、景気後退かどうかを判別する際に重要になる単語を確認することもできた。

これらの結果から得られる経済的な示唆は以下の通りである。第一に、景気分析を行う際、オルタナティブデータの活用は有用である。政府統計などの既存の統計には一長一短（例えば景気後退確率について、NY 連銀のモデルは先行性に優れる一方、精度に課題がある）があるが、オルタナティブデータ（本稿ではテキストデータ）を用いることで、既存の統計の弱点をカバーする（本稿の景気後退モデルでは、先行性と精度のバランスが良い）ことが可能となる。そのため、経済分析にあたっては、オルタナティブデータの活用を進め、従前よりも幅広く、またより深い分析を行っていく必要がある。第二に、テキスト分析を用いて景気分析を行う際には、特定期間にのみ多く出現する単語よりも、全文書を通して継続して出現する単語を使用の方が実態に沿った結果を得られるということである。一般のテキスト分析においては、特徴的な単語の発見が必要となることも多いが、景気分析にあたっては、特定の単語の出現数の変化に着目すると、実際の経済に合う結果が得やすい。第三に、経済関連の文書を読む際には、変化を示す単語に注目すべきである。特に、景気が急速に悪化するような場面においては、本稿の分析で示された通り、「弱い」という単語よりも「弱まる」という単語の方がより繊細な動きをしており、経済関連の文書を読む上で注目すべきである。

本稿はテキスト分析により景気の変化を捉えることを主眼としてきたが、分析データである Beige Book は FRB の公式文書であり、中央銀行のコミュニケーションという観点で、いくつかの追加的な示唆が得られる。単語の出現頻度の変化では、トランプ政権下で「関税」の出現回数が急増したほか、コロナやハリケーンなども特定の時期に多く出現することが確認された。これらの単語に対する FRB の関心の高さを窺うことができ、背景としては、関税の引き上げは輸入物価の上昇を通じて FRB の政策目標の一つである消費者物価の直接的な上昇要因となるほか、コロナやハリケーンといった経済ショックは機動的な対応が必要であり、いずれも金融政策において重要な事柄であることが考えられる。続いて、景気後退確率の推計は、FRB の景気認識を定量化したものと考えることもできる。FRB は四半期

に一度「Summary of Economic Projections」¹⁵を公表し、GDPの見通しやリスクバランスを示しているが、景気後退リスクは定量的に公表していない。他方、金融政策の大幅な転換は景気後退時に生じることが多いため、市場での景気後退リスクへの関心は大きく、特に景気減速時にはエコノミストが競って景気後退確率を算出している。本稿の景気後退確率の分析は、FRBが示す定性的な景気認識を、定量的な景気後退確率として示したものとみることができ、金融政策の転換を考える上で有用なものであると考えられる。例えば、2023年は、Beige Bookのテキストによる景気後退確率は10%程度であり、FRBは景気後退リスクが小さいと認識していたと考えられ、大幅な金融政策の転換が生じる可能性も小さかったと考えることができる。また、分析により、景気後退確率の推定において重要視された単語は、「weaken（弱まる）」「reduce（減らす）」「close（縮小する）」などであったが、これらの単語の使用頻度が高まった際には、FRBが景気後退への懸念を強めているとも考えられる。

こうした点を踏まえると、本稿の分析結果は一定の有用性を示す一方で、分析対象や分析手法の観点から、いくつかの課題や今後の発展の方向性も考えられる。

第一に、分析対象文書および分析対象国の拡張である。本稿では米国のBeige Bookを対象に分析を行ったが、日本においても、日本銀行が公表する「地域経済報告（さくらレポート）」など、中央銀行による定性的な景況判断を含む文書が存在している。こうした文書をテキスト分析の対象とすることで、国内景気の把握に資する可能性も考えられる。

第二に、分析対象指標の拡張である。景気後退確率に限らず、GDP成長率やインフレ率、政策金利といった他のマクロ経済指標を推計対象とすることや、雇用や物価、地域別の景況感といった多面的な観点からの分析も考えられる。これにより、中銀のコミュニケーションが経済や金融政策に与える影響を、より幅広く検証することが可能になると考えられる。

第三に、分析に用いるデータの拡張である。例えば、日本政策投資銀行が実施する設備投資アンケートや、有価証券報告書など、企業の投資行動や投資マインドを反映するデータを活用することで、政策当局の景気認識と企業行動との関係について、より踏み込んだ分析が可能になると考えられる。

加えて、近年の生成AIを含む自然言語処理技術の急速な進展も、本稿で用いた分析手法と親和性が高いと考えられる。実際、経済分析に生成AIや大規模言語モデル（LLM）を活用する研究も近年みられている。例えば、Dunn et al. (2024)は、FOMC議事要旨を対象に、生成AIモデルを用いて文書中の文を複数の政策トピックに分類し、人手による分類結果との比較を通じて、中央銀行コミュニケーションの内容把握における生成AIの有用性を検証している。また、高橋ほか (2025)は、LLMをシミュレーション上の経済主体として用いる試行的分析を行い、消費者の支出行動や企業の価格設定行動について、経済理論と整合的な結果が得られる可能性を示している。こうした生成AIを用いた分析は、文脈を踏まえた分類や、表現の揺らぎへの対応に強みを持ち得る点で、本稿の分析手法とは異なる特徴を有し

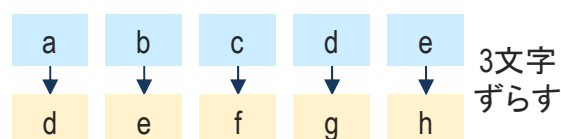
¹⁵ <https://www.federalreserve.gov/monetarypolicy.htm>

ている。例えば、Dunn et al. (2024)のようなトピック推定では、文書に含まれる多様な議論を、文単位で柔軟に分類することが可能となる。一方で、分類結果がどのような語句や判断基準に基づいて得られたものかについては、分析者が追加的に確認する必要がある。これに対し、本稿では、前半のセンチメント分析において辞書を用いて Beige Book の記述内容を定量化したほか、後半の景気後退確率の推計では、単語の出現状況を特徴量として用いることで、文書に含まれる情報を機械学習モデルに取り込んだ。このような手法は、分析に用いる語句や特徴量と推計結果との対応関係を確認しやすく、どのような表現が景気後退確率の上昇に寄与しているかを把握しやすいという利点を有している。他方で、辞書の設計や特徴量の選択が分析結果に影響を与え得る点や、新たな表現、文脈依存的な意味、語句の用法の変化を十分に捉えきれない可能性がある点には留意が必要である。

以上を踏まえると、生成 AI を用いたトピック推定やシミュレーション手法と、本稿が採用した分析手法とは、それぞれ異なる強みと制約を有していると考えられる。生成 AI の発展は、本稿で得られた分析結果そのものの有効性を直ちに変えるものではない一方で、テキストデータの前処理、分類候補の抽出、辞書の更新、特徴量設計、推計結果の頑健性確認といった工程を効率化・高度化する可能性がある。こうした観点から、従来型のテキスト分析手法が有する解釈可能性や再現性を維持しつつ、生成 AI を組み合わせた分析フレームの高度化についても、今後の検討課題としたい。

（補論 1）シーザー暗号化による文字化けの復号について

Beige Book の一部の資料について、pdf 上で閲覧する上では問題が生じないものの、データ上で文字化けが起きている資料が存在するため、これの復号を行っている。具体的には、例えば 21 年 4 月公表の Beige Book について、表記上では「and」となっている単語が、データ上では「dgg」となっているものがある（24 年 1 月時点）。この暗号化は、シーザー暗号（Caesar cipher）が行われていると推察される。シーザー暗号とは、補論図 1-1 の通り、文字列に対して、辞書順に数文字ずらしたものである。



補論図 1-1 シーザー暗号

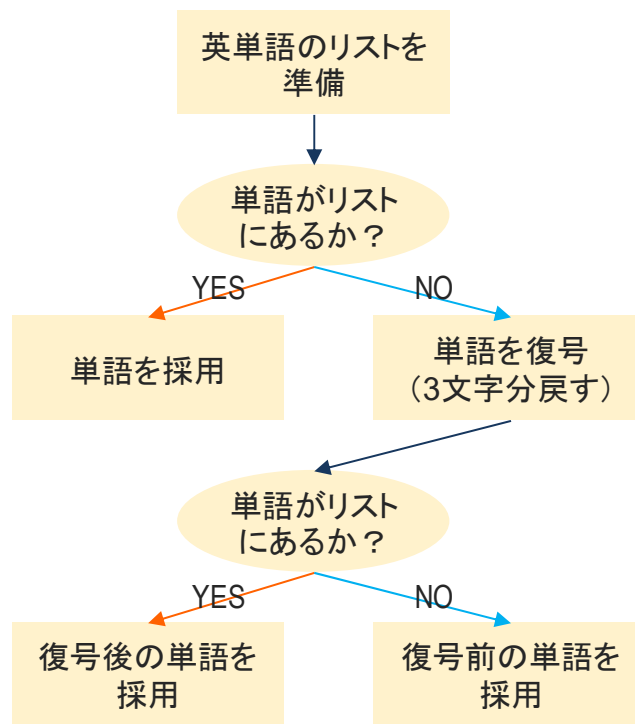
Beige Book においては、偶然の発見であるが、3 文字後ろにずらしたものになっている可能性が高いとみられる。実際、「and」を 3 文字ずらすと、「dgg」となる（補論表 1-1）。

補論表 1-1 シーザー暗号の例

復号前	復号後
dgg	and
sulfh	price
lqfuhdvh	increase
suhylr xv	previous
zlw k	with

このような暗号化が行われた意図は不明であるものの、分析上は、「dgg」を「and」に戻して分析することが望ましい。ただし、同じ「and」でも、暗号化されているものとされていないものが混在し、また、文書によって暗号化された単語が混在するものとししないものが存在する。そのため、一括で文字を置換することは難しく、本稿では、補論図 1-2 のようなステップを踏むことで復号を行った。

なお、復号不要な単語が失われるのを極力防ぐため、フローのように、復号前の単語が一般的な英単語リストになく、復号後の単語が英単語リストにある場合のみ、単語の復号を行った。そのため、英単語リストにないような一般的でない単語はそのまま残すことができるが、復号後の単語が一般的でない場合は暗号のまま残存している。



補論図 1-2 復号のフロー

(補論2) 見出し語化について

英語は、主として語形変化によって、性、数・格などの文法的関係を示す言語とされる。この語形変化について、翻訳などにおいては重要な意味を持つ（現在か過去か、単数形か複数形か、など）ものの、本稿のように多数の単語を扱う場合には、ノイズとなってしまう可能性がある。例えば、bag of words の手法では、出現単語ごとに次元ずつ与えたベクトルを考えるが、語形変化をそのままにしておくと、ベクトルの次元が大きくなりすぎてしまう。景気の局面変化を掴むうえでは現在形か過去形かの情報が有用である可能性はあるものの、分析の安定性の観点から、本稿では見出し語化の処理を行った。

語形変化に対処する方法として主要なものは2つあり、それぞれ語幹化 (Stemming)、見出し語化 (Lemmatizing) と呼ばれる。見出し語化は、単語を原型に戻す処理のことであり、語幹化は、単語の語幹部分で切り、語尾を削除する手法のことである（補論表 2a,2b）。これらの処理には一長一短あるとされ、見出し語化は、原型に戻す処理のため分析が安定する一方、単語を辞書と比較して行うため、処理に時間を要するなどのデメリットがある。他方、語幹化は、単語を機械的に語尾だけ削除するため、高速に処理ができる一方、語幹処理した結果、同一語をすべて処理しきれない、異なる単語を同一とみなしてしまうなど、安定性の面で弱みがある。本稿では、分析の安定性のほか、後々単語を確認しやすいように、見出し語化の処理を行った。

補論表 2a 見出し語化と語幹化の比較(drive)

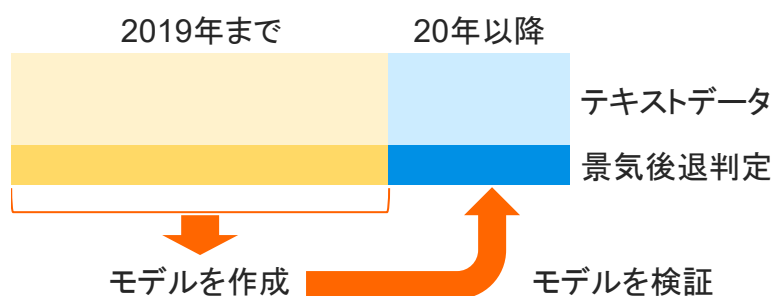
出現語	見出し語化	語幹化
driver	driver	driver
drive	drive	drive
drove	drive	drove
driven	drive	driven
drives	drive	drive
driving	drive	drive

補論表 2b 見出し語化と語幹化の比較(be 動詞)

出現語	見出し語化	語幹化
are	be	ar
am	be	am
being	be	be
been	be	been
be	be	be

（補論 3）過学習の対処について

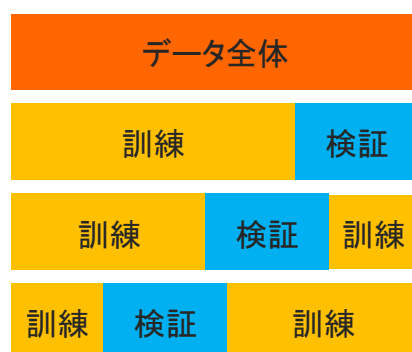
景気後退確率の推定では、大量の説明変数を用いるため、過学習（overfitting）の懸念がある。過学習とは、推定に用いたデータに過度に適合し、モデルの汎用性が失われてしまうことである。そこで、本稿では、主に 3 つの対処を行っている。第一に、サンプリングの工夫である。本稿のモデルでは、補論図 3-1 の通り、96 年第 4 四半期から 19 年第 4 四半期までのデータのみでモデル推定し、20 年第 1 四半期から 23 年第 4 四半期までのデータは検証用に残した。モデルを用いて景気後退確率を計算し、実際の景気後退判定との比較を行った。



補論図 3-1 サンプリング

第二に、モデルの工夫である。ランダムフォレストは特定の関数式を仮定しないため、従来の回帰モデルとは異なり説明変数の選択に制約が非常に少ないとされている（内閣府、2018）。

第三に、交差検証（cross validation）の実施である。交差検証とは、補論図 3-2 のように、訓練データと検証データの分割を複数回行い、それぞれでモデルの推定を行い、モデルの性能を確認することである。本稿では、ランダムに 3 種類の分割を行い、モデルの性能を確認したが、いずれも性能に大きな差はなかった。



補論図 3-2 交差検証

（補論 4）テキスト分析に関する近年の研究と本研究の位置づけ

本稿は、2024 年 3 月 28 日に日本政策投資銀行のウェブサイトで公表した「DBJ Research テキスト分析を用いた米景気後退確率の検討」をもとに、分析手法や先行研究との関係等を加筆・再構成したものである。既出の DBJ Research では、Beige Book のテキスト情報を用いて米国の景気後退確率を推計する分析の概要を示したが、本稿では、オルタナティブデータおよびテキスト分析に関する背景整理、センチメント分析および類似度分析の位置付け、景気後退確率モデルの構築過程、ならびに今後の課題に関する考察を拡充している。

なお、DBJ Research 公表後、Beige Book を用いた景気後退確率の推計に関する近接した研究として、Gascon and Martorana (2024) が公表されている（公表時期は 2024 年 12 月）。同研究は、Beige Book に含まれる定性情報を用いて、米国景気循環の把握を試みたものである。具体的には、Beige Book のテキストからセンチメント指数を作成するとともに、政治的事象や自然災害など、通常のセンチメント指標を歪め得る各種ショックを捉える指数を作成し、これらを説明変数とするプロビットモデルにより景気後退確率を推計している。Gascon and Martorana (2024) は、こうしたモデルから得られるリアルタイムの景気後退確率が NBER の景気後退期と密接に対応し、景気循環の転換点を把握するうえで有用であることを示している。

同研究は、Beige Book に含まれる定性情報が景気循環の把握に有用であることを示している点、同情報を用いて景気後退確率の推計を行っている点で、本稿と問題意識を共有している。一方で、分析アプローチには相違がある。Gascon and Martorana (2024) は、Beige Book から作成したセンチメント指数およびショック指数を用いてプロビットモデルにより景気後退確率を推計しているのに対し、本稿は、センチメント分析は行いつつも、景気後退確率の推定にあたっては、単語の出現頻度に着目し、ランダムフォレストを用いて分析を行っている。すなわち、両者はいずれも Beige Book のテキスト情報を景気判断に活用する研究であるが、テキスト情報の定量化の方法や推計手法に違いがある。

このように、本稿は既出の DBJ Research を基礎として分析手法や考察を整理・拡充したものである。また、その後に公表された近接研究と比較しても、Beige Book のテキスト情報を用いて景気後退確率を推計するという問題意識は共通する一方、機械学習により景気後退確率の推計を試みている点に特徴がある。

参考文献

- Armesto, M. T., Hernández-Murillo, R., Owyang, M. T., & Piger, J. (2009). Measuring the information content of the beige book: A mixed data sampling approach. *Journal of Money, Credit and Banking*, 41(1), 35-55.
- Banbura, M., Giannone, D., & Reichlin, L. (2010). Nowcasting.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Buckman, S. R., Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2020). News sentiment in the time of COVID-19. *FRBSF Economic Letter*, 8(1), 5-10.
- Chauvet, M., & Piger, J. (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics*, 26(1), 42-49.
- Dunn, A., Hood, K., & Driessen, A. (2020). *Measuring the effects of the COVID-19 pandemic on consumer spending using card transaction data*. US Department of Commerce, Bureau of Economic Analysis.
- Dunn, W., Meade, E. E., Sinha, N. R., & Kabir, R. (2024). Using generative AI models to understand FOMC monetary policy discussions. FEDS Notes December 06, 2024.
- Gascon, C. S., & Werner, D. (2022). Does the Beige Book Reflect US Employment and Inflation Trends?. *Economic Synopses*, (13).
- Gascon, C. S., & Martorana, J. (2024). The beige book and the business cycle: Using beige book anecdotes to construct recession probabilities. Federal Reserve Bank of St. Louis, Research Division.
- Keida, M., & Takeda, Y. (2017). *A Semantic Analysis of Monetary Shamanism: A case of the BOJ's Governor Haruhiko Kuroda*. RIETI.
- Keida, M., & Takeda, Y. (2019). *The Art of Central Bank Communication: A Topic Analysis on Words used by the Bank of Japan's Governors*. RIETI.
- Kliesen, K. L., & Werner, D. (2022). Using Beige Book Text Analysis to Measure Supply Chain Disruptions. *Economic Synopses*, (18).
- Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature.
- Martinez, L. R. (2022). How much should we trust the dictator's GDP growth estimates?. *Journal of Political Economy*, 130(10), 2731-2769.
- Nielsen, F. Å. (2011). Afinn. *Richard Petersens Plads, Building*, 321.
- Renz, I., & Franke, J. (2003). Text Mining.
- Sadique, S., In, F., Veeraraghavan, M., & Wachtel, P. (2013). Soft information and economic activity: Evidence from the Beige Book. *Journal of Macroeconomics*, 37, 81-92.
- Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of forecasting*, 30(6), 565-578.

- 石田基広. (2020). 実践 R によるテキストマイニング: センチメント分析・単語分散表現・機械学習・Python ラッパー.
- 小木しのぶ. (2015). テキストマイニングの技術と動向. *計算機統計学*, 28(1), 31-40.
- 金明哲, & 村上征勝. (2007). ランダムフォレスト法による文章の書き手の同定.
- 白木敬之, 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. *BOJ Reports & Research Papers*
- 水門善之, 和泉潔, 坂地泰紀, 島田尚, & 松島裕康. (2019). 高頻度電力需要データを用いた製造業活動のナウキャストモデリングモデルの構築. *人工知能学会第二種研究会資料*, 2019(SAI-035), 01.
- 高橋悠輔, 大高一樹, & 加藤直也. (2025). 生成 AI の経済シミュレーションにおける応用可能性. 日本銀行リサーチラボ
- 内閣府政策統括官（経済財政分析担当）. (2018). 日本経済 2017－2018－成長力強化に向けた課題と展望－
- 内閣府政策統括官（経済財政分析担当）. (2023). 特別定額給付金の家計消費に与えた影響－リアルタイムに記録される家計簿アプリデータを活用した分析－
- 難波英嗣. (2020). テキスト間の類似度の測定. *情報の科学と技術*, 70(7), 373-375.
- 波部 齊. (2016). ランダムフォレストの基礎と最近の動向. *映像情報メディア学会誌*, 70(9), 788-791.
- 橋本英樹. (2021). 国内企業の決算説明会の感情分析. 第 55 回 2021 年度夏季 JAFEE 大会 (2021 年 8 月 22 日) 発表論文
- 樋口耕一. (2004). テキスト型データの計量的分析—2 つのアプローチの峻別と統合—. *理論と方法*, 19(1), 101-115.
- 保田明夫. (2006). 形態素解析と分かち書き処理. *テキスト・マイニング研究会*

経済経営研究目録

(1980年7月より2026年7月まで)

	Vol.	No.	発行年月
◇経済一般理論・実証◇			
経済主体の節度とモラル	45	(1)	2024 . 4
—堀内行蔵氏のビジョンを追って—			
グローバル経済の潮流とアジア～米中摩擦、デジタル化、そして日本の対応	41	(3)	2020 . 8
—2019年度一橋大学・日本政策投資銀行共同シンポジウム抄録—			
技術革新と金融活動～日本経済へのインパクト	40	(2)	2019 . 7
—2018年東大・設研共同共催シンポジウム抄録—			
貨幣・雇用・リベラリズム 大瀧雅之氏の研究を振り返って	40	(1)	2019 . 4
—東大社研・DBJ設研シンポジウム抄録—			
世界の潮流の変化とアジア経済	39	(1)	2018 . 8
—2017年度一橋大学・日本政策投資銀行共同シンポジウム抄録—			
「アジアの世紀」における日本 経済再生とビジネスチャンス	36	(5)	2016 . 3
—2015年度一橋大学・日本政策投資銀行共同シンポジウム抄録—			
日本の景気循環の推計	26	(1)	2005 . 5
—Markov Switching Dynamic Factor Modelを用いた検討—			
経済の情報化とITの経済効果	22	(1)	2001 . 11
日米経済と国際競争	20	(4)	2000 . 3
現金収支分析の新技法	16	(3)	1995 . 11
日米独製造業の国際競争力比較	12	(1)	1991 . 6
—実質実効為替レートを利用した要因分析—			
レーガノミックスの乗数分析	10	(1)	1989 . 5
為替レートのミスアラインメントと日米製造業の国際競争力	9	(1)	1988 . 7
貯蓄のライフ・サイクル仮説とその検証	2	(3)	1982 . 1
今後のエネルギー価格と成長径路の選択	1	(1)	1980 . 7
—期待されるエネルギーから資本への代替—			

◇設備投資◇

Multiple q による投資関数の推計	31	(2)	2010 . 7
—過剰設備の解消過程における資本財別投資行動の考察—			
1990年代不況下の設備投資と銀行貸出	26	(7)	2006 . 3
R&Dのスピルオーバー効果分析	26	(2)	2005 . 6
—日本のハイテク産業における実証—			
1990年代の設備投資低迷の背景について	25	(4)	2004 . 12
—財務データを用いたパネル分析—			
設備投資と不確実性	25	(2)	2004 . 9
—不可逆性・市場競争・資金制約下の投資行動—			
大都市私鉄の運賃改定とその過程の研究	16	(6)	1996 . 1
—1985～1995年—			
大都市私鉄の運賃改定とその過程の研究	16	(2)	1995 . 11
—1966～1984年—			
大都市私鉄の運賃改定とその過程の研究	15	(1)	1994 . 12
—1945～1965年—			
大都市私鉄の投資と公的助成	14	(1)	1993 . 4
—地方鉄道補助法とその評価—			
鉄道運賃・収支と設備投資	13	(2)	1992 . 7
大都市圏私鉄の設備投資について	12	(3)	1991 . 8
設備投資と資金調達	11	(4)	1991 . 2
—連立方程式モデルによる推計—			
土地評価とトービンの q / Multiple q の計測	10	(3)	1989 . 10
我が国の設備機器リース	9	(5)	1989 . 3
—その特性と成長要因—			
設備の償却率について	9	(3)	1988 . 9
—わが国建設機械の計測例—			
設備投資の決定要因	6	(5)	1986 . 3
—各理論の実証比較とVARモデルの適用—			
設備投資研究 '85	6	(4)	1985 . 9
—主要国の設備投資とわが国におけるR&D投資の構造的特色—			

設備投資研究 '84	5 (1)	1984 . 7
—変貌する研究開発投資と設備投資—		
設備投資研究 '82	4 (2)	1983 . 7
—調整過程における新たな企業行動—		
投資促進施策の諸類型とその効果分析	4 (1)	1983 . 7
設備投資研究 '81	3 (4)	1982 . 7
—研究開発投資の経済的効果—		
税制と設備投資	3 (3)	1982 . 7
—調整費用、合理的期待形成を含む投資関数による推定—		
時系列モデルの更新投資への適用	3 (2)	1982 . 7
設備投資研究 '80	2 (2)	1981 . 7
—投資行動分析の新しい視角—		

◇金融・財政◇

テキスト分析を用いた米景気後退確率の検討	47 (3)	2026 . 7
—Beige Bookの文書データから—		
グローバル経済における資産バブルと経済成長	36 (6)	2016 . 3
—リーマン・ショック前後の世界経済に対する理論的考察—		
リスクマネーの供給と日本経済の持続的成長	36 (1)	2015 . 4
—2014年東大・設研共同主催シンポジウム抄録—		
設備投資研究所設立50周年記念シンポジウム議事録	35 (3)	2015 . 3
第1部 金融・経済篇		
最適負債制御問題	35 (2)	2014 . 10
危機対応業務と環境格付融資の意義	34 (6)	2014 . 3
—DBJ業務が企業価値に与える効果—		
企業の調達流動性に影響を与える要因について	34 (3)	2013 . 9
国際金融の新たな展開と日本企業のダイナミクス	34 (1)	2013 . 5
—2012年度東大・設研共同主催シンポジウム抄録—		
CDSスプレッド指標の決定要因	33 (2)	2013 . 3
—需給構造を考慮した同時方程式推定からの含意—		

銀行ローンシェア構造の決定要因	33	(1)	2013 . 3
—企業—銀行マッチレベルデータからの含意—			
危機後の金融システムはどこに向かうのか	31	(4)	2011 . 3
—2010年東大・設研共同主催シンポジウム抄録—			
企業—銀行間関係の動態的安定性について	31	(3)	2010 . 9
—ハザード関数推計からの含意—			
金融システム・公共政策の課題と展望	30	(1)	2009 . 4
—2008年東大・設研共同主催シンポジウム抄録—			
自由な労働移動のもとの通貨統合の費用	29	(1)	2009 . 3
いわゆる「ゾンビ企業」はいかにして健全化したのか	28	(1)	2008 . 3
貸し手間の協調の失敗と公的政策	27	(1)	2006 . 5
日本企業のガバナンス構造	24	(1)	2004 . 1
—所有構造、メインバンク、市場競争—			
非対称情報下の投資と資金調達	23	(3)	2003 . 2
—負債満期の選択—			
—投資非効率と企業の規模—			
メインバンク関係は企業経営の効率化に貢献したか	21	(1)	2000 . 8
—製造業に関する実証研究—			
ドル・ペッグ下における金融危機と通貨危機	20	(3)	1999 . 8
アメリカ連邦政府の行政改革	20	(1)	1999 . 6
—GPRAを中心にして—			
なぜ日本は深刻な金融危機を迎えたのか	19	(1)	1998 . 9
—ガバナンス構造の展望—			
国際機関投資家の新潮流	16	(4)	1995 . 9
アメリカの金融制度改革における銀行隔離論	13	(1)	1992 . 6
メインバンクの実証分析	12	(4)	1992 . 3
Asset Bubbleのミクロ的基礎	11	(3)	1990 . 12
資産価格変動とマクロ経済構造	11	(2)	1990 . 7
貯蓄・投資と金利機能	11	(1)	1990 . 6
金融構造の変化について	10	(2)	1989 . 8

公的部門の金融活動	9 (4)	1988 . 10
—米国での動きとわが国との対比—		
クラウドディング・アウトについての研究	8 (1)	1987 . 11
—国債発行の国内貯蓄および金融仲介への影響—		
アメリカの金融システムの特徴と規制緩和	7 (1)	1986 . 10
アメリカの金融自由化と預金保険制度	6 (3)	1985 . 6
西ドイツの金融自由化と銀行収益および金融制度の安定	6 (2)	1985 . 7
西ドイツの公的金融 —その規模と特徴—		
アメリカの公的金融	6 (1)	1985 . 7
—フェデラル・ファイナンス・バンクと住宅金融—		
金融市場の理論的考察	5 (2)	1984 . 7
債券格付に関する研究	2 (1)	1981 . 7
資本市場に於ける企業の資金調達	1 (2)	1980 . 10
—発行制度と資金コスト—		

◇資源・環境・社会的共通資本◇

銀行を中心とした金融における水害リスクの管理	47 (2)	2026 . 6
新型コロナウイルス感染症が都市に及ぼした影響と 今後の見通し	44 (2)	2023 . 11
産業公害の克服と金融の役割の再検討	41 (2)	2020 . 7
—1970年代開銀公害防止融資の実証分析—		
設備投資研究所設立50周年記念シンポジウム議事録	36 (3)	2015 . 5
第3部 社会的共通資本篇		
CSR経営が企業価値に及ぼす効果	34 (2)	2013 . 6
環境配慮活動の決定要因と企業価値	31 (1)	2010 . 4
—環境格付融資事例による分析—		
温暖化対策の経済評価	30 (3)	2010 . 2
—わが国の中期目標における選択肢—		
二酸化炭素排出と環境グズネツ曲線	27 (3)	2007 . 3
—ダイナミック・パネルデータ推定による検証—		
カーボンファイナンスの評価と今後の可能性	25 (5)	2004 . 12
—モンテカルロ法によるシミュレーション分析—		

地域経済と二酸化炭素排出負荷	24 (4)	2004 . 3
エネルギー問題に関する理論および実証のサーベイ	1 (3)	1981 . 2

◇企業・財務・会計◇

取締役会の性別構成と環境情報開示	42 (1)	2021 . 8
比較制度分析の視点から見た企業統治改革～資本市場，所有構造 および支配権の在り方～ —2019年度早稲田大学高等研究所・日本政策投資銀行設備投資研究所共同主催シンポジウム抄録—	41 (1)	2020 . 7
統合報告書の公表企業像とその非財務情報の特徴 —統合報告書の公表企業へのアンケート調査分析—	39 (2)	2019 . 1
日本企業のコーポレートガバナンス： 産業の新陳代謝，サステナビリティ経営に向けた課題と展望 —2016年東大・設研共同主催シンポジウム抄録—	37 (2)	2017 . 3
有価証券報告書における定性情報の分析と活用 —リスクの多様化にともなう望ましい対話のあり方—	37 (1)	2016 . 5
国際インフラ投資セミナー議事録 —インフラ投資における昨今の潮流・ESG及びベンチマーク—	36 (8)	2016 . 3
リスク情報の統合開示 —統合報告にみる新しい財務報告の視座—	36 (7)	2016 . 3
現代の株式会社が社会的価値を創造するには？ —不完備契約理論からの考察—	36 (4)	2016 . 2
設備投資研究所設立50周年記念シンポジウム議事録 第2部 経営・会計篇	36 (2)	2015 . 4
統合報告の制度と実務	35 (1)	2014 . 7
コーポレート・ガバナンスと多角化行動 —日本の企業データを用いた実証分析—	34 (5)	2014 . 2
資本構成の調整手段について —日本の上場企業データによる実証分析—	34 (4)	2013 . 9
ストック・オプションと企業パフォーマンス —オプション価格評価額に基づく実証分析—	30 (4)	2010 . 3
ドイモイ（刷新）政策導入後のベトナムに於ける資本・金融 自由化政策概観	27 (4)	2007 . 3

日本のM&A	26	(6)	2006 . 3
—イベント・スタディによる実証研究—			
ベトナム私法整備の経緯と日本支援の役割	26	(5)	2006 . 3
—社会的共通資本としての法学の視点から—			
DIPファイナンスの実証研究	26	(4)	2006 . 3
税効果会計と利益操作	25	(6)	2005 . 3
—倒産企業による実証分析—			
コーポレート・ガバナンスの世界的動向	25	(3)	2004 . 9
—欧米、中国・韓国における法制度を中心とする最近の展開ならびに 「会社法制の現代化に関する要綱試案」の動向—			
コーポレート・ガバナンス改革の現状と課題	24	(5)	2004 . 3
—経営機構改革の具体例の検討、内部統制システム等に関する考察を中心として—			
利益の質による企業評価	24	(3)	2004 . 3
—利質分析の理論と基本的枠組み—			
企業の再生と挫折	24	(2)	2004 . 3
—UALにおけるターンアラウンド戦略の評価—			
商法改正後の新しいコーポレート・ガバナンスと企業経営	23	(6)	2003 . 3
—社外取締役、監査役会など米国型機構、従来型機構の検討を中心として—			
日本の製造業	23	(5)	2003 . 3
—長期データに基づく収益力の再検証—			
利益操作の研究	23	(4)	2003 . 2
—不当な財務報告に関する考察—			
バブル崩壊後の企業財務の推移と課題	18	(3)	1998 . 3
連結決算20年のデータで見る日本企業の資本収益性低下	18	(2)	1998 . 3
日米医療NPO（非営利組織）の経済分析	17	(2)	1997 . 3
企業のリストラクチャリングについて	16	(1)	1995 . 5
日本主要企業の資本構成	12	(2)	1991 . 7
企業における情報行動の分析	7	(2)	1987 . 3
—職場における情報行動に関する調査報告—			
ビジネス・リスクと資本構成	3	(1)	1982 . 4

◇産業構造・労働◇

ヘルスケア分野の開発における社会実装の要諦	47	(1)	2026 . 6
持続可能な国内物流の確保に向けて	46	(2)	2025 . 6
ー「トラックにおける2024年問題」の形成過程と 持続可能な国内物流の確保に向けた視点ー			
ライフサイエンス・エコシステムの形成と進化	46	(1)	2025 . 5
ー社会実装プロセスと人材開発ー			
医療と介護のサステナビリティ	45	(2)	2024 . 6
ーポスト2025における提供体制の再構築ー			
ライフサイエンスにおけるエコシステム形成の構図と創意工夫	44	(1)	2023 . 4
ーベルギーでのシステム形成と日本の課題対応			
輸出産業の生産性上昇と均衡失業率	30	(2)	2009 . 6
防衛的技術進歩	26	(3)	2005 . 7
ーグローバル経済下の内生的技術進歩ー			
技術進歩と人的資本	25	(1)	2004 . 5
ースキル偏向的技術進歩の実証分析ー			
我が国の半導体産業とイノベーション	23	(7)	2003 . 3
ーイノベーション経営研究会報告書ー			
我が国製造業の打開策を探る	23	(2)	2002 . 11
ープロダクション・ニューパラダイム研究会報告書ー			
貿易と雇用	23	(1)	2002 . 11
ーグローバル化の産業と地域への影響ー			
グローバル化と労働市場	21	(2)	2000 . 11
ー日本の製造業のケースー			
偏向的技術進歩と日本製造業の雇用・賃金	20	(2)	1999 . 6
ーコンピュータ投資にみる技術進歩の影響ー			
戦間期日本における農工間賃金格差	19	(3)	1998 . 12
日本の労働市場と失業	9	(2)	1988 . 8
ーミスマッチと女子労働供給の実証分析ー			
産業調整問題に関する理論および実証	3	(5)	1982 . 8

◇地域政策◇

ハイテク型産業クラスターの形成メカニズム	27	(2)	2006 . 10
—フィンランド・オウルICTクラスターにおける歴史の実証—			
地域・目的別社会資本ストックの経済効果	19	(2)	1998 . 11
—公共投資の最適配分に関する実証的分析—			
地域間所得移転と経済成長	18	(1)	1998 . 3
アジアにおける地域の国際ネットワーク化試論	17	(1)	1997 . 3
—ネットワークの理論的考察とその応用としてのアジア重層ネットワーク構想—			
新しい町づくりの試みサステイナブル・コミュニティ	16	(5)	1995 . 10
—真のベター・クオリティ・オブ・ライフを求めて—			
首都圏を中心としたハイテクゾーンの現状と将来	6	(6)	1986 . 3