

IMES DISCUSSION PAPER SERIES

ファイナンス・ワークショップ
「データ・サイエンスの企業分析への活用」の模様

Discussion Paper No. 2022-J-1

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

日本銀行金融研究所

〒103-8660 東京都中央区日本橋本石町 2-1-1

日本銀行金融研究所が刊行している論文等はホームページからダウンロードできます。

<https://www.imes.boj.or.jp>

無断での転載・複製はご遠慮下さい。

備考：日本銀行金融研究所ディスカッション・ペーパー・シリーズは、金融研究所スタッフおよび外部研究者による研究成果をとりまとめたもので、学界、研究機関等、関連する方々から幅広くコメントを頂戴することを意図している。ただし、ディスカッション・ペーパーの内容や意見は、執筆者個人に属し、日本銀行あるいは金融研究所の公式見解を示すものではない。

ファイナンス・ワークショップ 「データ・サイエンスの企業分析への活用」の模様*

1. 概要

日本銀行金融研究所は、2021年11月5日に「データ・サイエンスの企業分析への活用」と題するファイナンス・ワークショップをオンライン開催し、学界、民間金融機関等から約150名が参加した¹。近年発展著しいデータ・サイエンスの手法を用いてわが国の企業分析を行った2本の研究論文が報告され、指定討論者、参加者、および報告者との間で活発な議論が行われた。また、「データ・サイエンスの発展とファイナンス研究の方向性」をテーマとする対談セッションでは、データ・サイエンスに基づく研究の最近の発展状況を概観し、同分野の手法を用いた新たな研究の方向性や課題について展望した。本稿では、本ワークショップの模様を順に紹介する。

2. 開会挨拶

金融研究所長の**副島豊**（日本銀行）は、金融研究所におけるファイナンス研究の歩みや過去のワークショップを振り返ったうえで、今回のテーマ設定について以下のように説明した。

金融研究所では、フィナンシャル・エンジニアリングの研究部署を立ち上げた1994年以降、折々の課題を捉えて多数の研究論文を発表してきた。具体的には、金融自由化に伴うイールドカーブ・モデル、金融派生商品の開発とそれに伴う市場・信用リスクの計測・管理手法、金融危機後のカウンターパーティ・リスクや市場間の相関、テール・リスクへの注目の高まりを受けた信用評価調整（CVA）、コンピュータ、極値理論など様々なファイナンス分野の研究で貢献を行ってきた。このほか、金融市場構造の変化を踏まえたマーケット・マイクロストラクチャー分析や、高頻度データという市場データ面での技術発展を受けたボラティリティ計測手法の精緻化、マクロプルーデンスやシステムミック・リスクの分析など金融システムを分析する手法の発展にも幅広く貢献してきた。

また、ファイナンス・ワークショップを開催し、金融危機後の金融工学の展開、国

* 本稿に示された意見はすべて発言者たち個人に属し、その所属する組織の公式見解を示すものではない。

¹ プログラムは別添1を、参加者リストは別添2を参照。所属はワークショップ開催時点のもの。

債市場の流動性とシステミック・リスク、ビッグデータ・人工知能（Artificial Intelligence: AI）を活用したリスク計測・分析等、その時々に関心事項を採り上げ、研究者や実務家と活発な議論を続けてきた。

今回のワークショップのテーマとした「データ・サイエンス」は、前回採り上げたビッグデータやAIなどを包含するより広いコンセプトであり、分析対象、分析手法、そして分析に用いるデータはますます多様化していることを反映したテーマ設定となっている。データ・サイエンスをファイナンス研究にどう活用していくか活発な議論が展開されることを期待する。

3. 研究報告セッション

(1) 「CO₂ 排出量と企業パフォーマンス：Double Machine Learning を用いた日本の実証研究」

報告者の**有賀涼**（日本銀行）は、最初に、CO₂ 排出量と企業パフォーマンスの関係については数多くの先行研究が存在するものの、コンセンサスが得られていないことを紹介した²。その要因として、先行研究で用いられている線形回帰モデルには、目的変数（被説明変数）とコントロール変数の間に非線形な関係がある可能性を捨象することによるバイアスや、潜在的に含めるべき重要な変数を採用しないことに起因するバイアス（欠落変数バイアス）が生じている可能性を指摘した。そこで、従来の線形回帰モデルによる分析に加えて、これらのバイアスに対処するため、Chernozhukov *et al.* [2018]が提案する Double Machine Learning (DML) を採用したセミパラメトリック・モデルによる分析を行った。具体的には、目的変数として、長期・短期の企業パフォーマンス、および資本コストの観点から様々な指標を採用し、企業の CO₂ 排出量とこれら目的変数の関係について複数のモデルを用いて評価および頑健性チェックを行った³。

分析の結果、CO₂ 排出量の少ない企業ほど (1) 長期的な企業パフォーマンスが良好となること、(2) 株主資本コストが低くなることが確認された。この結果について、CO₂ 排出量の少ない企業ほど将来の気候変動に関連する経営リスク（例えば、将来の規制導入による業績悪化）が低いと評価され、当該企業の市場価値が高くなる、あるいは株主の要求プレミアムが低くなる（資本コストが低くなる）という解釈を与えた。

このほか、CO₂ 排出量の少ない企業ほど短期的な企業パフォーマンスが高いことや、負債コストが低いことも示唆された。もっとも、長期の企業パフォーマンス分析と異なり、コントロール変数を拡張した DML による分析では、線形回帰モデルによる分

² 詳細は有賀・五島・千葉 [2021] を参照。

³ 本研究では、長期の企業パフォーマンスとして市場価値を表す指標、短期の企業パフォーマンスとして直近の収益性指標、資本コストとして株主資本コストと負債コストを採用した。

析に比べて推計結果の有意性が低下することが確認された。この結果から、CO₂排出量と短期的な企業パフォーマンスや負債コストとの関係については、線形回帰モデルによる分析では欠落変数バイアスが生じる可能性があるとして指摘した。

指定討論者の**渡辺安虎**（東京大学）は、内生性の問題に対処しながら CO₂ 排出量の影響を正確に推計することが課題だが、DML を上手く活用した本研究では頑健性の高い結果が得られていると評価した。今後の研究の発展に向けて、①企業固定効果を導入した分析、②コントロール変数に財務変数以外の定性情報（例えば各社の IR 資料における ESG に関する記述）の組入れ、③業種や規模といった企業毎の異質性を考慮した追加分析、④資本コスト低下のメカニズム、⑤結果の時系列変化の背景の考察を提案した。これに対して**有賀**は、①と③～⑤については、サンプル・サイズの制約などから本研究では対応が難しいが、将来的にデータの蓄積が進んだ後に研究を掘り下げていくうえで重要な視点であるとして同意した。また、②についてはテキスト解析などが必要となるため、別途の研究課題としたいと回答した。

また、同じく指定討論者の**湯山智教**（金融庁）は、企業パフォーマンスと CO₂ 排出量の関係に注目した研究が最近多くみられる中、本研究では従来の線形回帰モデルと DML による分析を並行して実施し、また複数の目的変数を用いて評価を行うことで頑健性の高い結果を導出した点を高く評価した。そのうえで、推計期間中の構造変化の可能性や、依然として内生性の問題が残存している可能性について質問した。また、本研究の結果から、CO₂ 排出量が高い企業は高リスク・低リターンとなっているように見えるがファイナンス理論的には高リスク・高リターンとなるのが通常であることを考えると謎でありどう解釈すべきかという質問と、CO₂ 排出量は（ファーマ・フレンチの 3 ファクターのように）株価リターンを説明する継続的なファクターとなっている、すなわち効率的市場仮説が成立しない傍証になっていると解釈することの妥当性に関する質問があった。これらに対して**有賀**は、いずれも有益な示唆であり、追加データ収集や新たな分析フレームワークを考えるうえで大変参考になると応答した。

フロアから**渡辺努**（東京大学）が、CO₂ 排出量と目的変数の間に非線形の関係が存在する場合の解釈について質問した。**水門善之**（野村証券）は、各企業の属するセクターによって CO₂ 排出量が企業パフォーマンスに及ぼす影響の度合いが異なりうる点や、ポストコロナのデータを用いた場合にはこの間の産業構造の変化によって分析結果が異なる可能性について指摘した。さらに、**宇野淳**（早稲田大学）は、結果を解釈するうえで、検証したパフォーマンス変数を企業の CO₂ 排出量に対する各ステークホルダー（投資家や銀行、経営者等）の反応として考察することが有用であるとコメントした。また、分析対象をグローバルに拡張することで、各法域の規制の違いや世論の関心度の違いが結果に与える影響を考察することも期待されると述べた。

(2) 「中小企業の生産性ダイナミクス：退出形態のデータを用いた実証分析」

報告者の**宮川大介**（一橋大学）は、企業合併が中小企業の実生産性に及ぼす影響について、大規模データを用いた因果推論の研究成果を報告した。欧米の先行研究では、企業の退出はマクロレベルの実生産性を高めるとの実証結果（正の退出効果）が得られている。これは、産業内の競争により実生産性の低い企業が退出していくと解釈されている。一方、わが国を対象とした先行研究では、実生産性の高い企業が何らかの理由で退出することでマクロレベルの実生産性が低下するという実証結果（負の退出効果）が報告されており、長らくパズルとされてきた。

宮川は、上述のパズルの背景を解明するため、合併、休業業、倒産といった個別企業の退出形態に関する情報を含む約 100 万社の中小企業データを用いて、企業退出が中小企業全体の実生産性に及ぼす影響を退出形態別に分析した。分析の結果、中小企業の退出に伴う中小企業全体の実生産性低下の主因は、高実生産性中小企業の休業業や倒産によるものではなく、高実生産性中小企業が他の企業に合併されて退出すること（被合併）によるものであったと報告した。

ここで追加的な関心事となるのは、高実生産性企業を合併した側の企業の実生産性の変化である。合併した側の企業の実生産性は改善することが期待されるが、そのプラス効果と被合併に伴って生じるマイナス効果（負の退出効果）の差し引きでマクロの実生産性が決まる。この点について、**宮川**は以下のような手法を用いて分析した。まず、存続企業の合併後の成長率と、当該企業が仮に合併しなかった場合の成長率推計値との差分を「合併効果」とみなし、これを推定した。具体的には、標準的な因果推論手法である逆確率重み付け推定法（Inverse Probability Weighting: IPW）を用いて、処置群（ここでは合併を行った存続企業）が仮に「合併しなかった場合の成長率」を推計し、存続企業の合併後の成長率との差分として「処置群における平均処置効果」（Average Treatment Effect on the Treated: ATT）を算出し、これを合併効果とみなした。そのうえで、合併効果と被合併による実生産性低下効果を比較し、合併効果のほうが上回ることを示した。この結果から、高実生産性企業の被合併を通じた退出に伴う実生産性の低下を打ち消す形で、合併を行った存続企業の実生産性が高まっており、合併企業を考慮すれば、マクロの実生産性は実際には上昇していると述べた。

指定討論者の**深尾京司**（一橋大学）は、わが国の実生産性動学分析で企業買収情報を活用した先行研究はほとんどなく、この分野における重要な貢献になりうると評価した。そのうえで、分析上の大きな問題として、各企業の実生産性を集計する際のウェイトがアウトプットになっており、実生産性の高い企業に過大なウェイトを与え得る点や、業種別のアウトプットの計測方法が売上になっており、卸小売業のウェイトが過大になる点を指摘した。そのうえで、こうした偏りを抑制する方法として、実生産性のウェイトは労働投入で、卸小売業のアウトプットは商業マージンでそれぞれ計測すべきと提案した。また、高実生産性企業が合併により退出する理由を検証することも重要であ

るとコメントした。さらに、中小企業間の合併は、社内調整が比較的容易な株式会社が中心となりやすいが、株式会社の生産性は他の企業形態と比べて高い傾向があるため中小企業に絞った分析では合併の効果を過大評価してしまう危険性があると述べ、大企業を含めた全企業を対象とすることが望ましいと指摘した。

次の指定討論者の**家森信善**（神戸大学）は、まず論文全体の貢献について、合併を通じた資源配分が機能していることを示す結果であり、中小企業政策や金融機関の支援に関する政策のあり方に対する示唆に富んだ分析であると評価した。そのうえで、海外との比較可能性や、被合併による生産性低下効果が日本で顕著な理由、近年の非正規雇用の増加が生産性の計測に及ぼす影響、景気回復期である2014年から2018年にかけて生産性が低下した背景の評価について質問した。続いて、合併は案件ごとの異質性が高いことを踏まえると、合併が企業が生産性に及ぼす影響度の評価は、平均的な効果ではなく合併事例ごとに計測し、その分布によって評価を行うほうがよいのではないかと質問した。さらに、合併には仲介手数料などの多大な負担がかかることから合併の便益が大きい高生産性企業が合併対象として選ばれやすいと考えられるため、合併の効果を評価する際は、こうした暗黙の選別について注意する必要があると指摘した。これに関連して、合併コストの軽減には合併に対する政府補助金が有効であるが、他方で合併のメリットが小さい企業も合併対象とされやすくなるため、結果として合併のプラス効果が弱まるかもしれないと述べた。最後に、企業が自社の価値を十分に認識しておらず、合併のメリットを把握できていない印象を持っており、合併を通じてわが国の生産性を高めていくには、取引銀行といった社外から事業承継を働きかけることが有効との見解を示した。

宮川は、上記の意見に同意したうえで、生産性のウエイト付け手法、一部業種のアウトプットの計測手法、全企業を対象とした分析、海外における合併との比較については、今後の追加分析で確認するとした。このほか、被合併企業の特性に関する追加分析を行うことで、合併により退出した企業が合併を選択した理由や、合併が実現する条件、（合併ではなく）廃業にいたる条件などを解明していくことが可能になると述べた。

フロアからは、**河原光雄**（東京商工リサーチ）から、企業データを提供する側の立場から今後どのような企業データが研究上有用となるかとの問いがあった。**宮川**は、一例として取引額といった企業間の取引関係の大きさに関する情報を挙げた。**渡辺（努）**は、本報告が先行研究と異なるデータセットを用いた点について、データ利用に関する制約がなければ同一のデータセットに基づいて検証を行う方が本研究テーマに関するコンセンサス形成の観点では望ましい、とのコメントを行った。**宮川**は、そうした検証は今後の課題として受け止めると述べた。

4. 対談セッション「データ・サイエンスの発展とファイナンス研究の方向性」

モデレーターの**副島**は、**大橋和彦**（一橋大学・東京工業大学）と**渡部敏明**（一橋大学）に対し、近年のデータ・サイエンス発展がファイナンス研究にどのような影響をもたらしているか展望コメントを求めた。

大橋は、最近のファイナンス関係の主要誌におけるデータ・サイエンスに関する論文をみると、特に、テキストデータを用いた研究が目立つと指摘した。具体的には、中央銀行によるアナウンスメントの効果や政策的含意に関する内外中央銀行⁴による研究を紹介した。また、機械学習を用いて財務報告書からイノベーションやチームワークなど企業文化に関わる情報を定量化し、これらが企業業績に与える影響を分析した研究（Li *et al.* [2021]）を採り上げ、企業の ESG への取り組み状況の分析などにも応用できる可能性を指摘した。続いて、気候変動に関するニュースを自然言語処理で分析し、株式市場における気候変動リスクの評価とそのヘッジ手段を提案した Engle *et al.* [2020]について、気候変動リスクを捉えるための新しいアプローチとして興味深いと評価した。さらに、新しい分野への活用という観点では、個人や企業の信用評価や、資産価格の予測、資産価格モデルの検証など、応用分野が着実に広がっている点に言及した。

渡部は、近年ポピュラーな手法となっているベイズ統計学の概説を行い、コンピュータの計算速度の飛躍的な向上によってファイナンス研究でのベイズ統計学の利用が拡大しており、機械学習や自然言語処理といった新しい分析手法のアルゴリズムにも利用されていると述べた。また、計量ファイナンス研究における新しいデータの活用については、高頻度データの活用が進んでおり、例えば、資産価格のボラティリティの推定に、近年、日中の高頻度資産価格から計算されるリアライズド・ボラティリティが用いられていることなどを紹介した。

副島は、こうした動きについて、①大規模データや非構造データという新しいデータの登場とその分析技術の発展、および、②新しい分析対象への関心の広がりという2つの方向性が組み合わさり、ファイナンス研究におけるデータ・サイエンスの活用は進んでいると整理した。次いで、今後のファイナンス研究の発展の方向性や、留意すべき課題について見解を求めた。

渡部は、AI や機械学習の実務での利用は一層拡大していくと予想し、貸出の信用評価といった人間による既存業務の一部を置き換えていく可能性もあると指摘した。一方、機械学習を利用する際の課題として、分析結果の解釈の難しさや、サンプル・データへの過学習の可能性、相関関係と因果関係の識別の必要性、過去に起きていない事象や経済構造の変化等を踏まえたモデルの見直しの重要性などを指摘した。

大橋は、**渡部**が提示した論点に加えて、ビッグデータ・AI の活用が、ファイナンス

⁴ 具体的には、風戸 [2020]、風戸・黒崎・五島 [2019]、Munday and Brookes [2021]。

以外の領域にも拡がり、ビジネスチャンスを生み出す可能性を指摘し、こうした動きはファイナンス研究の分析対象の拡大にもつながるとの見解を示した。一方で、データの活用にかかるプライバシーの問題や、機械学習による信用評価が人種差別につながりうる可能性など、新しいデータ・サイエンスがもたらす負の側面も意識しつつ、活用を進めることが重要であると指摘した。

5. 閉会挨拶

理事の**貝塚正彰**（日本銀行）は、報告者・登壇者・出席者の参加や議論への貢献に感謝の意を示したうえで、一連の議論を踏まえた、データ・サイエンスの活用にあたっての留意点を述べた。具体的には、ブラックボックス化しやすいデータ分析が示す結果が、どのようなストーリー（文脈や因果性）のもとで生じているのかをしっかりと考えることが重要であり、近年の分析手法の一部にはこれが難しいものがあるのではないかとの見解を述べた。また、データ分析によって得られた知見を政策遂行に活用するなど、価値の創造につなげていくことが重要であると指摘した。さらに、データ分析の基礎となるデータの作成自体コストがかかるものであり、こうしたワークショップにおける研究発表や様々な情報発信を通じて、高品質なデータが継続的に提供されるような環境をサポートしていくことの重要性を指摘した。そのうえで、日本銀行金融研究所としては、今後もデータ・サイエンス研究に邁進し、有意義な政策運営につなげていく所存であると述べ、ワークショップを締めくくった。

報告論文

有賀涼・五島圭一・千葉貴司、「CO₂ 排出量と企業パフォーマンス : Double Machine Learning を用いた日本の実証研究」、金融研究所ディスカッション・ペーパー No. 2021-J-11、日本銀行金融研究所、2021 年

Ito, Yojiro, and Daisuke Miyakawa, “Performance of Exiting SMEs in Japan: An Empirical Analysis Using Exit Mode Data,” IMES Discussion Paper Series, 2022, forthcoming.

参考文献

風戸正行・黒崎哲夫・五島圭一、「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No.2019-J-16、日本銀行金融研究所、2019 年

風戸正行、「中央銀行のコミュニケーションを巡るテキスト分析：政策的含意と今後の課題」、『金融研究』第 39 巻第 3 号、日本銀行金融研究所、2020 年、23～46 頁

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins, “Double/Debiased Machine Learning for Treatment and Structural Parameters,” *Econometrics Journal*, 21(1), 2018, pp. C1–C68.

Engle, Robert F., Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, “Hedging Climate Change News,” *Review of Financial Studies*, 33(3), 2020, pp. 1184–1216.

Li, Kai, Feng Mai, Rui Shen, and Xinyan Yan, “Measuring Corporate Culture Using Machine Learning,” *Review of Financial Studies*, 34(7), 2021, pp. 3265–3315.

Munday, Tim, and James Brookes, “Mark My Words: The Transmission of Central Bank Communication to the General Public via the Print Media,” Staff Working Paper No. 944, Bank of England, 2021.

日本銀行金融研究所 ファイナンス・ワークショップ
(テーマ：データ・サイエンスの企業分析への活用)

▽ 日程と開催方式

(日程) 2021年11月5日(金)

(開催方式) オンライン

▽ プログラム

開会挨拶：副島 豊 (日本銀行 金融研究所長)

研究報告1 「CO₂排出量と企業パフォーマンス：

Double Machine Learning を用いた日本の実証研究」

報告者：有賀 涼 (日本銀行)

—— 五島 圭一 (東京大学エコノミックコンサルティング)、

千葉 貴司 (三井住友銀行) との共同研究

指定討論者：渡辺 安虎 (東京大学)、湯山 智教 (金融庁)

研究報告2 「中小企業の生産性ダイナミクス：退出形態のデータを用いた実証分析」

報告者：宮川 大介 (一橋大学)

—— 伊藤 洋二郎 (日本銀行) との共同研究

指定討論者：深尾 京司 (一橋大学)、家森 信善 (神戸大学)

対談セッション「データ・サイエンスの発展とファイナンス研究の方向性」

登壇者：大橋 和彦 (一橋大学・東京工業大学)、渡部 敏明 (一橋大学)

モデレーター：副島 豊 (日本銀行 金融研究所長)

閉会挨拶：貝塚 正彰 (日本銀行 理事)

参加者の氏名および所属（所属はワークショップ開催時点のもの。敬称略、五十音順）

氏名	所属
有賀 涼	日本銀行
家田 明	マネーフォワードケッサイ
和泉 潔	東京大学
磯貝 明文	三菱UFJ トラスト投資工学研究所
伊藤 洋二郎	日本銀行
岩壺 健太郎	神戸大学
植田 和男	共立女子大学・東京大学
内山 朋規	東京都立大学
内山 雅浩	ゴールドマン・サックス・アセット・マネジメント
宇野 淳	早稲田大学
太田 晴康	みずほリサーチ&テクノロジーズ
大橋 和彦	一橋大学・東京工業大学
小田 信之	杏林大学
小野寺 敬	日本経済研究センター
貝塚 正彰	日本銀行
梶田 脩斗	日本経済研究センター
片岡 剛士	日本銀行
加藤 大輔	みずほリサーチ&テクノロジーズ
金田 規靖	日本銀行
亀山 智行	三井住友銀行
河原 光雄	東京商工リサーチ
北岸 郁雄	マネーフォワード
北村 行伸	立正大学
木全 友則	日本銀行
楠岡 成雄	東京大学・明治大学
五島 圭一	東京大学エコノミックコンサルティング
齊藤 誠	名古屋大学
坂地 泰紀	東京大学
塩路 悦朗	一橋大学
塩野 剛志	クレディ・スイス証券

氏名	所属
新谷 幸平	日本銀行
新谷 元嗣	東京大学
水門 善之	野村證券
須田 真太郎	三菱 UFJ トラスト投資工学研究所
関根 敏隆	一橋大学
副島 豊	日本銀行
高橋 大志	慶應義塾大学
千葉 貴司	三井住友銀行
中川 秀敏	一橋大学
仲田 泰祐	東京大学
中村 豊明	日本銀行
畠 篤史	三菱 UFJ 銀行
原 尚子	日本銀行
肥後 雅博	東京大学
平木 一浩	日本銀行
深尾 京司	一橋大学
福田 慎一	東京大学
藤井 健司	グローバルリスクアンドガバナンス
藤木 裕	中央大学
藤原 大資	三菱 UFJ 銀行
保坂 豪	東京証券取引所
牧本 直樹	筑波大学
水田 孝信	スパークス・アセット・マネジメント
宮川 大介	一橋大学
武藤 一郎	日本銀行
森平 爽一郎	慶應義塾大学
森本 祐司	キャピタスコンサルティング
柳岡 優希	東京商工リサーチ
山内 浩嗣	三菱 UFJ トラスト投資工学研究所
山中 卓	青山学院大学
家森 信善	神戸大学
湯田 清敬	住友生命保険

氏名	所属
湯山 智教	金融庁
吉羽 要直	東京都立大学
若田部 昌澄	日本銀行
和田 健治	日本銀行
渡辺 努	東京大学
渡部 敏明	一橋大学
渡辺 安虎	東京大学

※ このほか、傍聴者が約 90 名（うち日本銀行職員が約 15 名）参加。