

jh181001-NWJ

財務ビッグデータの可視化と統計モデリング

地道 正行（関西学院大学 商学部）

概要 本研究では、企業財務ビッグデータを用いて、企業活動のグローバル化がもたらす負の側面（企業の租税回避、労働者と株主間の付加価値配分、企業の富の偏在等）の実態に関する証拠を提示し、課題を解決するための方策を検討する。この財務ビッグデータは、Bureau van Dijk (BvD)の157カ国・9万社超の上場企業の33年間・86系列の財務データ（Osiris、300万行、1.4GB）および非上場企業を含む2,000万社超の世界最大規模の財務データ（Orbis、2億行、120GB超）である。これをGPGPU環境でApache Spark, Hadoop, HiveとRを連動させて、探索的データ解析（Exploratory Data Analysis）に基づき、時空間の観点からビジュアライゼーション技法を用いたダイナミックでインタラクティブなデータ可視化（Data Visualization）を行う。得られた知見に基づき、企業行動を高精度に予測する統計モデリングと実証分析を行うことでその有効性を検証する。これらの結果を社会に広く還元し、企業行動を持続可能な発展に向けての変革を目指す。

1. 共同研究に関する情報

(1) 共同研究を実施した拠点名

東京大学情報基盤センター

(2) 共同研究分野

- 超大規模数値計算系応用分野
- 超大規模データ処理系応用分野
- 超大容量ネットワーク技術分野
- 超大規模情報システム関連研究分野

(3) 参加研究者の役割分担

地道 正行（関西学院大学 商学部）:

- データ操作・データ整形
- 探索的データ解析にもとづく統計モデリング

宮本 大輔（東京大学大学院 情報理工学系研究科）:

- データ解析環境の構築

阪 智香（関西学院大学 商学部）

- 財務データの経済・会計学的考察

永田 修一（関西学院大学 商学部）

- パネルデータ・時系列データ解析の理論構築

2. 研究の目的と意義

本研究の目的は、次の3つである。

- (1) 経済社会のサステナビリティ（持続可能性）を確保するためには、グローバルレベル

で企業活動を解明し、それが生み出す様々な社会的課題を解決することが欠かせない。そこで、世界の企業財務ビッグデータを用いて、企業活動の実態とその経済的帰結の証拠を提示し、新しい知見と課題を明らかにする。

(2) 本研究で利用する財務ビッグデータは、Bureau van Dijk (BvD)の157カ国・9万社超の上場企業の33年間・86系列の財務データ（Osiris、300万行、1.4GB）および非上場企業を含む2,000万社超の世界最大規模の財務データ（Orbis、2億行、120GB超）である。これらのデータの全てを扱う分析は、世界で初めてであり、GPGPU環境でApache Spark, Hadoop, Hive, Rを連動させて、探索的データ解析（Exploratory Data Analysis）を実施する。さらに、時空間の観点からダイナミックでインタラクティブなデータ可視化（Data Visualization）を行い、データ自身の情報を引き出し、企業活動の実態の証拠と課題をデータに語らせる。

(3) その上で、課題解決の方策を探り、高精度に予測する統計モデリングと実証分析を行うことでその有効性を検証する。これらの結果を社会に広く還元し、企業行動を持続可能な発展に向けて変革するために、シチズンサイエンスとして確立する。

本研究の意義は、次の 3 つである。

(1) 入手可能な財務データとしては、過去最大規模の財務ビッグデータを扱う。120GB 超の複数のテキストファイルは、通常の計算機環境のメモリ容量を超える。既存研究が扱う財務データの量を遙かに凌駕する過去最大規模のものである。

(2) 財務データの構造に最も適した処理と分析を探求する。財務データは、経時観測データ(またはパネルデータ)と呼ばれる構造をもち、複数の個体に対する属性を決算年月日において観測したものである。このような構造から、時間・空間の両面からデータを精査・分析する必要がある。そこで、時間的な変化を見るために、時系列プロットやモーションチャートなどによって可視化を行う。また、クロスセクションデータの対散布図などを用いてデータの時空間的な分布状態を可視化する。さらに、探索的データ解析の視点から、高精度に予測する統計モデリングと検証を実施する。

(3) 入手可能な全世界の企業財務データの可視化や探索的データ解析は、我々の研究を除きこれまで実施されてきたことがない。一部企業の財務実績のみを取り上げるならば、他の企業行動がミスリードされる可能性もある。我々は、全世界・地域単位・国単位等の様々な視点から、全世界の企業の財務データを探索的に解析することで、企業行動の実態とそれがもたらす社会的影響の分析も行う。これらの証拠を、各種の可視化手法を用いて、わかりやすく社会に提示する。

3. 当拠点公募型共同研究として実施した意義

この研究で扱われるデータセットには、120GB 以上のファイルが複数存在し、その規模から通常のコンピュータ環境では取り扱いが難しい。さらに、これらのデータセットは、抽出してそのまま解析することは不可能

であり、データ形式の変更(ヘッダー部分とデータ部分の分離)、文字・行末コードの変換、欠損値の置換などを行い、データ解析・分析できるファイル形式に変換する、「前処理」(preprocessing)が必要である。この工程は、データを解析することも含めた全体の 50%~90%を占めるともいわれ、近年ビッグデータを扱う分野で最も重要視されている。具体的な処理の方法としては、Unix 系 OS のコマンドやインタプリタ(grep, sed, dos2unix 等)を利用し、ある程度整形した上で、さらにデータ解析環境 R を利用して再整形・ファイル出力を行う必要がある。

このように、データ解析環境で利用できるデータセットを用意すること自体に相応の処理能力を有する計算機環境が必要である。また、データを分析するために拠点へ転送する際や、拠点内でネットワークを介した Hadoop 等の分散環境下で利用するにあたって高速なネットワーク資源が必要となる。さらに、この研究で使用する分析法の一つに、コンピュータ・シミュレーションをベースとした評価法があり、この評価法による研究を進める上で、GPGPU を利用した高速計算を実行可能なコンピュータ環境が必要である。

このような研究を行うにあたり、東京大学情報基盤センターの JHPCN 環境 (FENNEL) のもと、高速な計算機環境・ネットワーク環境という物理的な資源を利用することで、世界規模の企業行動について全く新しい証拠と知見を得ることが期待される。

4. 前年度までに得られた研究成果の概要

2017 年度に採択された JHPCN 環境 (FENNEL) を利用した研究は、Bureau van Dijk 社の Osiris から抽出されたデータ(以下(DS-Osiris-2017)と略す)に基づいている。このデータは、その規模(テキストベースで2GB程度)から通常のコンピュータ環境では扱うことが難しく、Apache Spark と R を連携さ

せることにより分析を行った。実際の処理過程などの検証結果は地道らで発表し、論文にまとめた(2017 年度最終報告書・学術論文[7]参照)。また、財務データを利用し、会計学的・統計学的観点から次の 3 つの研究を行った。

(1) 企業の富の偏在と国際・国内格差
 企業の富の偏在と国際・国内格差の証拠を示すため、(a) 国家間で企業の富が過度に集中し、(b) ストック(資産)の集中はフロー(利益)の集中より大きいこと(図 1 参照)、(c) 企業間格差も拡大傾向にあり、世界全体ではトップ 1% (10%) 企業の売上占有率は 48% (86%) であること、(d) Piketty (2014) が示した格差のメカニズム(資本利益率 $r >$ 成長性 g) が過去 30 年の企業データでも見られること等を明らかにした(2017 年度最終報告書・学術論文[5]参照)。



図 1 世界 157 か国、全上場企業の資産合計・利益合計の国別分布のジオチャート

(2) 企業の租税回避行動の研究

企業の付加価値分配の 1 つとして税金支払に着目し、①企業の租税回避行動の証拠を、支払税金ゼロ企業の存在(図 2 参照)と、58 カ国中の 2/3 で

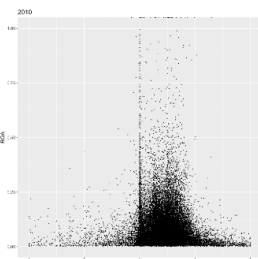


図 2 利益率 (ROA、縦) と実効税率 (横) 散布図 (2015 年)

「実効税率－法定税率」が負であることなどにより示した。②企業の租税回避と各国税率の引下競争により、過去 15 年間に企業の実効税率と各国法定税率が世界規模

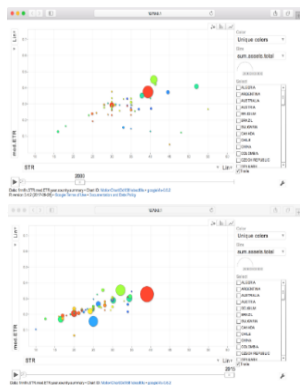


図 3 実効税率(縦軸)、法定税率(横軸)(上 2000 年, 下 2015 年)

で下方に収斂した実態(図 3 参照)を示した。これらの成果は学会・シンポジウム(2017 年度最終報告書・会議発表[4], [5], [7])で報告し、論文にまとめた(学術論文[3], [8]参照)。

(3) 生産関数の推測の研究

経済学の基本問題である生産関数の推定を、世界規模の企業財務データ(DS-Osiris-2017)で検証した。(a) 売上高、従業員数、資産合計の可視化を行った結果、「歪み」があることを確認、(b) データの対数をとることにより非対称正規分布で近似できること(昨年度最終報告書[1]参照)、さらに裾の部分で当てはまりに問題があること、(c) 売上高に対して非対称テール分布を当てはめたところ赤池情報量規準等によって精度の改善が認められ(2017 年度最終報告書[2]参照)、(d) 売上高を従業員数と資産合計で説明するために対数線形モデルを当てはめたところ、K-分割交差検証法によって誤差項に非対称テール分布を仮定した対数線形モデルが最も予測精度がよいことがわかった。これらの結果を会議で報告し、論文にまとめた(2017 年度・会議発表[6]、学術論文[1], [2], [7]参照)。

5. 今年度の研究成果の詳細

今年度の研究成果は、次の 5 つである。

(1) データセット(DS-Osiris-2017)の前処理とデータ解析環境への読み込み(データラングリング)、そしてデータ解析結果が再現可能性であるかを検証した。その結果を得るための工程を Unix 環境(主に make コマンドを利用)で自動実行することにより、データの前処理、データラングリング、データ解析、モデル選択、モデル評価、さらに文書作成までの全工程を東京大学の FENNEL 上で再現可能なものとして生成できるかについて検討し、結果として再現性が確保されていることがわかった。これらの成果を、学会・シンポジウム(国内会議発表[1], [3], [4], [7], [12], [13])で報告し、前処理の再現性確保の考察を論文

にまとめた(学術論文[2]参照)。さらに、学術論文[4]では、データ解析、モデル選択、モデル評価、さらに文書作成までの工程を再現可能なものとするための方法をまとめた。

(2) 2017 年度得られた主要な結果の一つである、コブダグラス型生産関数を用いて売上高を従業員数と資産合計を説明するための統計モデリングに関する結果が、学術雑誌 *Japanese Journal of Statistics and Data Science* に採択された(学術論文[1]参照)。

(3) 今年度に抽出した Bureau van Dijk 社の Orbis データセット(以下(DS-Orbis-2018)と略す)を連結(Consolidated)主体のもの(以下(DS-Orbis-C-2018)と略す)と非連結(Unconsolidated)主体のもの(以下(DS-Orbis-U-2018)と略す)に区分し、前処理した。さらに Apache Spark と R を用いてデータラングリングを行うことによってデータ解析できる形式に変換し、これらの工程を再現可能とするために、make コマンドを利用して自動実行することを検証した。さらに、従来から行われているリレーショナルデータベース(MySQL)と R を用いたラングリングも行い、Apache Spark と R を利用した場合との利点・欠点の比較を行った(図 4, 5 参照)。これらの結果は、国内会議発表[5], [11]で報告した。

(4) 企業活動の実態解明を行うために、(a)

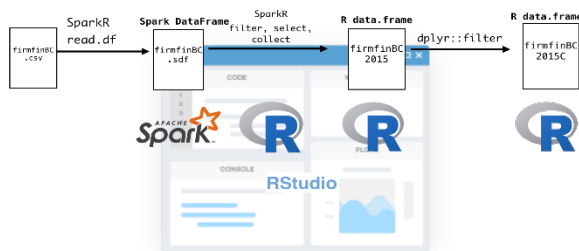


図 4 Apache Spark と R を用いたラングリング

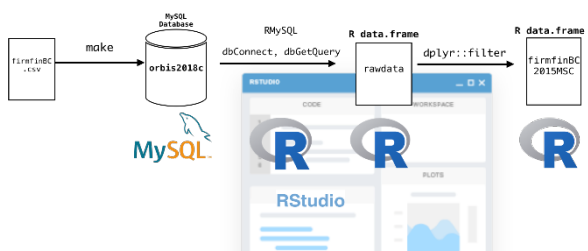


図 5 MySQL と R を用いたラングリング

企業の富の偏在と国際・国内格差、(b)付加価値の分配、(c)企業の租税回避についての考察結果を、国内会議発表[2], [4], [9], [10], [14]で報告し、海外学術雑誌 *Meditari Accountancy Research* に採択された(学術論文[7]参照)。さらに、データベースとデータ解析の実行環境は異なるが、国内会議発表[6], [8], [15]は、東京証券取引所第一部上場企業(以下東証一部上場企業と略)の財務データの探索的データ解析を、再現可能性とともに考察した研究報告である。この結果は、データセット(DS-

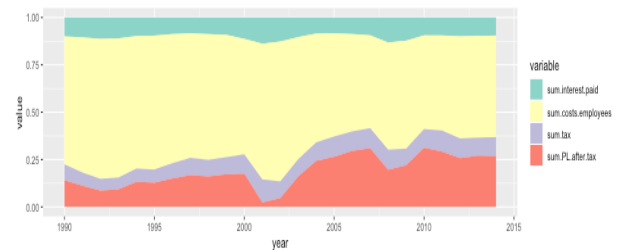


図 6 付加価値分配 (157 カ国、25 年間)

Osiris-C-2018)を利用し、日本の東証一部上場企業と同様の結果が得られることも既に結果として得られており、今後発表を予定している。

(5) 企業が生み出した付加価値が、ステークホルダーにどのように配分されているかを確認するために、世界の全上場企業の付加価値分配の実態を分析した(図 6 参照)。支払利息として債権者(緑)、従業員人件費(黄)、政府への税金(紫)、利益を通して株主(赤)への分配を可視化したところ、過去 25 年間で企業は従業員への分配を減少し利益を増

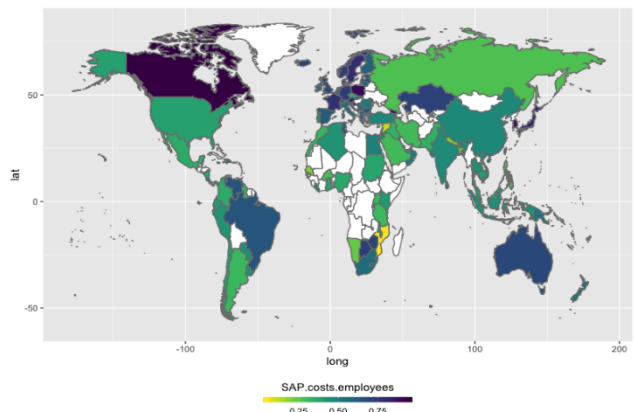


図 7 国別労働分配率 (143 カ国、2015 年)

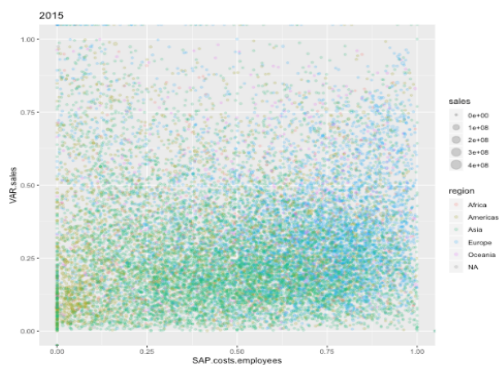


図 8 全企業労働分配率 (143 カ国、地域別 (青: ヨーロッパ、緑: アジア、黄緑: アメリカ、赤: アフリカ)、2015 年)

加させていることがわかった。特にアメリカ諸国や新興国ではそれが顕著であること、一方でヨーロッパ諸国では従業員への高い分配割合を維持していることが明らかとなった。国レベル・企業レベルで同様の結果が得られた (図 7、8 参照)。

6. 今年度の進捗状況と今後の展望

2018 年度研究進捗を時系列的に列挙する。

2018 年 4 月:

- ・4/28(土) 2018 年度キック・オフ・ミーティング開催(今年度の研究計画の確認)
- ・探索的財務ビッグデータ解析を、データセット(DS-Osiris-2017)に対して前処理、データラングリング、データ可視化、統計モデリング、モデル選択、モデル評価、再現可能研究の全工程の観点から実行することを検討

2018 年 5 月:

- ・今年度の上場企業データセット(DS-Osiris-C-2018(連結決算)、DS-Osiris-U-2018)の検証
- ・5/26(土) 第 2 回ミーティング開催(テーマ: 『R 関数の GPGPU 環境下でのチューニングについて+GPU の活用事例』、『配当金支払金額の探索的データ解析』)
- ・5/31(木) 41st Annual Congress of the European Accounting Association で発表

2018 年 6 月:

- ・6/2(土) 日本経営数学会で発表
- ・6/22(金) 23rd International Euro-Asia Research

Conference で発表

- ・6/22(土) 日本組織会計学会第 2 回全国大会、統一論題で発表

- ・データセット(DS-Orbis-C-2018)検証

2018 年 7 月:

- ・7/12(木) JHPCN で口頭・ポスター発表
- ・7/21(土) 応用経済時系列研究会で発表
- ・データセット(DS-Orbis-U-2018)の検証
- ・論文『探索的財務ビッグデータ解析—前処理, データラングリング, 再現可能性』投稿

2018 年 8 月:

- ・論文『探索的財務ビッグデータ解析—データ可視化による企業活動の実態解明と統計モデリング』投稿
- ・8/4(土) 第 3 回ミーティング開催(テーマ: 『探索的財務ビッグデータ解析—前処理, データラングリング, 再現可能性—』、『配当金支払金額の探索的データ解析』)

- ・データセット(DS-Orbis-U-2018)の検証

- ・8/25(土) 国際数理科学協会 2018 年度年会で発表

- ・Japanese Journal of Statistics and Data Science 誌への投稿論文が採択

2018 年 9 月:

- ・非対称正規誤差をもつ線形回帰モデルに対する竹内情報量規準の理論面の計算
- ・9/8(土) 統計関連連合大会 (中央大学後楽園キャンパス)で発表

- ・データセット(DS-Orbis-C-2018)の FENNEL 環境におけるクラスター化 (Hadoop, Hive)

- ・論文『探索的財務ビッグデータ解析—データ可視化、統計モデリング、モデル選択、モデル評価、動的文書生成、再現可能研究』投稿 (9/30)

2018 年 10 月:

- ・10/6(土) 国際会計研究会・大阪会計研究会合同大会で発表

- ・10/25(木)『データサイエンスの基礎 R による統計学独習』, 裳華房 出版

- ・10/27(土) 日本社会関連会計学会第 31 回全

国大会統一論題で発表

2018 年 11 月:

- ・ 11/3(土) 第 5 回ミーティング開催(テーマ: 『探索的財務ビッグデータ解析』、『配当金支払金額の探索的データ解析』)
- ・ 11/11(日) 日本計算機統計学会第 32 回シンポジウムで発表
- ・ データセット(DS-Orbis-C,U-2018)の前処理の並列化検証

2018 年 12 月:

- ・ 12/8(土) 2018 年度統計数理研究所共同研究集会「データ解析環境 R の整備と利用」で発表
- ・ 12/15(土): 科研費シンポジウム「多変量データ解析法における理論と応用」で発表
- ・ 12/15(土) 日本経営分析学会西日本部会で発表

- ・ 12/16(日) 第 6 回ミーティング開催(テーマ: 『GNU parallel による Osiris データの前処理』、『経時観測データとしての Osiri2017 データ』)

2019 年 1 月:

- ・ 1/16(日) 第 7 回ミーティング開催(テーマ: 『Osiris2017 データの欠測の観点からの考察』)

2019 年 2 月:

- ・ 再抽出されたデータセット(DS-Orbis-C,U-2018)の前処理を並列化及び、FENNEL 環境におけるクラスター化を検証
- ・ 2/21(木) JHPCN サブミーティング開催 (テーマ: 『データセット(DS-Orbis-C,U-2018)の可視化による検証』)

2019 年 3 月:

- ・ 3/3(日) 第 8 回ミーティング開催(テーマ: 『Orbis 2018 データ (2019/02 抽出バージョン)の並列化前処理とデータラングリング』)
- ・ *Meditari Accountancy Research* 誌への投稿論文が採択

以上が 2018 年度の進捗であるが、本年度解決されていない問題と次年度での取り組み

を述べることによって、今後の展望を行う。

(1) まず、今年度のテーマとして挙げていた時間的な推移を考慮した観点からのモデリングについては、2018 年 12 月と 2019 年 1 月のミーティングで中間的な結果の報告を行ったが、経時的な観点から利用できるデータ量や、欠損情報などのデータの品質についての詳細な検討が必要であることが分かったため、次年度に向けて引き続き有効なモデリングを検討する予定である。

(2) 次に、データセット(DS-Orbis-C, U-2018)のデータを前処理とラングリングから、実際にデータを可視化することによって検証する過程 (図 9 参照) で、精緻にデータを分析するためには、データ提供母体であるインフ

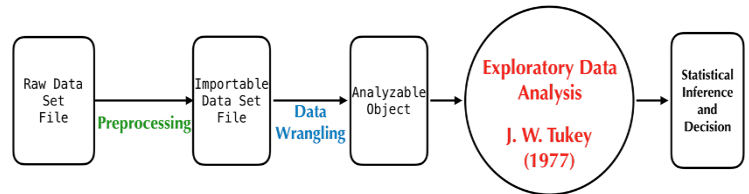


図 9 前処理、データラングリング、探索的データ解

オメーションプロバイダの情報が必要であることがわかったため、データの再抽出を BvD 社に依頼し、2 月に納品された。これらのデータセットに対して、改めて前処理を行う必要があったが、その規模から従来の処理法では、一回の処理に 10 時間程度の時間を要するという時間的な問題があった。この問題に対して、11 月から検討を始めた並列化による処理を、まずは規模の小さいデータセット(DS-Osiris-2017)に対して GNU parallel を利用して試み、さらにその知見を応用してデータセット(DS-Orbis-C, U-2018)に対して並列処理を行った。その結果、(実験段階ではあるが) 10 分の 1 の時間で前処理が完了するという結果を得た。この結果は、3 月におけるミーティングで共有した。これは、ビッグデータを扱う上で重要となるキーワードの一つである速度 (velocity) の観点から注目すべき改善が得られたことを意味する。なお、この結果については、さらに詳細に検討を行い、

2019 年度に公表する予定である。

(3) 前処理が行われたデータセットのファイルをデータ解析ソフトウェア R に読み込む段階、すなわちデータラングリング (図 9 参照) を行う工程を、R, Spark, Hadoop, Hive 環境と GPGPU 環境を連動させることによって高速化するというテーマについては、Spark の最新の Java 環境への対応の遅れによるものと思われる問題から実現できていない。この問題に対しては、次年度にデータベースサーバ PostgreSQL を GPGPU 環境下で利用可能にするツールである PG-Strom を用いることによって、高速なラングリングを実現することを試みる。なお、この方法の実行可能性については、テスト環境での実験を既に行っており、良好な結果を得ている。

(4) 今年度得られた結果には、規模の大きなデータセット(DS-Orbis-C, U-2018)のデータ処理の再現性とスピードを改善させることに注力したため、このデータセットに対して、時空間の観点からのダイナミックかつインタラクティブなデータ可視化に本格的に取り組む段階に至らなかったが、規模の小さなデータセット(DS-Osiris-2017, 2018)に対しては、データ解析環境 R における googleVis, Shiny パッケージを利用してダイナミックにデータの構造を可視化することについての結果を得ており、次年度には、これらの結果を公表する予定である。

7. 研究成果リスト

(1) 学術論文

- [1] Jimichi, M., Miyamoto, D., Saka, C., and Nagata, S. (2018) “Visualization and statistical modeling of financial big data: double-log modeling with skew-symmetric error distributions”, *Japanese Journal of Statistics and Data Science*, Vol. 1, pp. 347-371.
- [2] 地道 正行「探索的財務ビッグデータ解析

ー前処理、データラングリング、再現可能性ー」、『商学論究』、第 66 巻第 1 号、関西学院大学商学研究会、pp. 1-31、2018 年 9 月。

- [3] 阪 智香「会計ビッグデータの可視化」『企業会計』第 70 巻第 4 号、中央経済社、pp. 4-5、2018 年 4 月。
- [4] 地道 正行「探索的財務ビッグデータ解析ーデータ可視化、統計モデリング、モデル選択、モデル評価、動的文書生成、再現可能研究ー」、『商学論究』、第 66 巻第 2 号、関西学院大学商学研究会、pp. 1-41、2018 年 12 月。
- [5] Saka, C., Oshika, T. and Jimichi, M. (2018) “Financial KPIs for Sustainability: Evidence from Japanese Long-lived Firms”, *International Review of Business*, No. 19, pp. 1-18.
- [6] 阪智香「長寿企業の財務的特徴ー収益性、財務報告の質、付加価値分配の分析ー」『商学論究』、第 66 巻第 4 号、関西学院大学商学研究会、pp. 435-452、2019 年 3 月。
- [7] Saka, C., Oshika, T. and Jimichi, M. (2018) “Visualization of Tax Avoidance and Tax Rate Convergence: Exploratory Analysis of World-scale Accounting Data”, *Meditari Accountancy Research*, forthcoming (accepted).

(2) 国際会議プロシーディングス

(3) 国際会議発表

- [1] Saka, C., Oshika, T. and Jimichi, M. “Does Tax Avoidance Diminish Sustainability?”, 41st Annual Congress of the European Accounting Association, Milan, Italy, 1 June 2018.
- [2] Saka, C., Oshika, T. and Jimichi, M. “Does Tax Avoidance Diminish Sustainability?”, 23rd International Euro-Asia Research Conference, Kobe, Japan, 22 June 2018.

(4) 国内会議発表

- [1] 地道 正行、宮本 大輔、阪 智香、永田 修二「探索的財務ビッグデータ解析—データ(ファイル)操作、データ可視化、統計モデリング、モデル選択、モデル評価、再現可能研究—」、日本経営数学会第 40 回(通算 60 回) 研究大会、関西学院大学上ヶ原キャンパス、2018 年 6 月 2 日。
- [2] 阪 智香「SDGs 時代の会計の役割—会計から CSV (共有価値創造) を考える—」日本組織会計学会第 2 回全国大会、統一論題、明治大学、2018 年 6 月 23 日。
- [3] 地道 正行、宮本 大輔、阪 智香、永田 修二「財務ビッグデータの可視化と統計モデリング」、学際大規模情報基盤共同利用・共同研究拠点(JHPCN) 第 10 回 シンポジウム、THE GRAND HALL (品川)2018 年 7 月 12 日。
- [4] 阪 智香、地道 正行「財務データの可視化による企業活動の実態解明と SDGs への課題」、第 35 回応用経済時系列研究会、慶應義塾大学三田キャンパス、2018 年 7 月 21 日。
- [5] 地道 正行、宮本大輔、阪 智香、永田 修二「探索的財務ビッグデータ解析—前処理、データラングリング、再現可能性—」、国際数理科学協 2018 年度年会「統計的推測と統計ファイナンス」分科会研究集会、関西学院大学梅田キャンパス、2018 年 8 月 25 日。
- [6] 柳 麻衣、阪 智香、地道 正行「配当金支払金額の探索的データ解析」、国際数理科学協会 2018 年度年会「統計的推測と統計ファイナンス」分科会研究集会、関西学院大学梅田キャンパス、2018 年 8 月 25 日。
- [7] 地道 正行、阪 智香、宮本大輔、永田 修二「探索的財務ビッグデータ解析」、2018 年度年会統計関連学会連合大会、中央大学後楽園キャンパス、2018 年 9 月 12 日。
- [8] 柳 麻衣、阪 智香、地道 正行「配当金支払金額の探索的データ解析」、2018 年度統計関連学会連合大会、中央大学後楽園キャンパス、2018 年 9 月 12 日。
- [9] 阪 智香「Visualization of Tax Avoidance and Tax Rate Convergence: Exploratory Analysis of Accounting Big Data」国際会計研究会・大阪会計研究会合同大会、グランフロント大阪、2019 年 10 月 6 日。
- [10] 阪 智香「S と G を会計データから考える—探索的会計ビッグデータ解析—」、日本社会関連会計学会第 31 回全国大会、統一論題、関西学院大学西宮上ヶ原キャンパス、2018 年 10 月 27 日。
- [11] 地道 正行、宮本大輔、阪 智香、永田 修二「探索的財務ビッグデータ解析—前処理、データラングリング、再現可能性—」、日本計算機統計学会第 32 回シンポジウム、滋賀大学データサイエンス学部、2018 年 11 月 11 日。
- [12] 地道 正行「探索的財務ビッグデータ解析」、統計数理研究所共同研究集会 2018 年度「データ解析環境 R の整備と利用」、2018 年 12 月 8 日。
- [13] 地道 正行、宮本大輔、阪 智香、永田 修二「探索的財務ビッグデータ解析」、科研費シンポジウム「多変量データ解析法における理論と応用」、広島大学西条キャンパス、2018 年 12 月 15 日。
- [14] 阪 智香「探索的財務ビッグデータ解析」、日本経営分析学会第 1 回西日本部会、関西大学吹田キャンパス、2018 年 12 月 15 日。
- [15] 柳 麻衣、阪 智香、地道 正行「配当金の探索的データ解析」(ポスター発表)、2018 年度日本統計学会春季集会、日本大学経済学部、2019 年 3 月 10 日。
- (5) その他 (著書)
- [1] 地道 正行『データサイエンスの基礎: R による統計学独習』、裳華房、2018 年 10 月 31 日発行。