

探索的財務ビッグデータ解析と再現可能研究： 非上場企業のデータ可視化と考察

地 道 正 行
阪 智 香

要 旨

本稿では、Bureau van Dijk 社のデータベース Orbis から抽出されたデータをもとに、世界の全非上場企業の10年間の会計データを用いた探索的データ解析を実施し、売上合計等や、収益性分析（ROE）、安全性分析（自己資本比率）、生産性分析（一人当たり売上高）、付加価値分配、租税回避の蓋然性の実態を示した。上場企業と比較した場合の特徴として、非上場企業ではROEの0（原点）付近の集中度が高いことや、自己資本比率は0%と100%付近の企業が多いこと、また労働分配率が高いことなどを明らかにするとともに、日本の非上場企業の特徴として、収益性と安全性がやや低いこと、労働分配率が緩やかな減少傾向にあること、租税回避の蓋然性が低いことが確認できた。なお、本研究において行われる処理は自動的に実行され、また本稿も動的文書生成によって作成されており、この意味で再現可能研究の立場から行われている。

キーワード：財務ビッグデータ（Financial Big Data）、データラングリング（Data Wrangling）、探索的データ解析（Exploratory Data Analysis）、動的文書生成（Dynamic Documents）、再現可能研究（Reproducible Research）

I はじめに

筆者の研究グループは、Bureau van Dijk¹⁾（ビューロー・ヴァン・ダイク）

1) Bureau van Dijk Web Page: <https://www.bvdinfo.com/en-gb/>

社のデータベース Orbis から抽出された世界の「非上場企業」(unlisted firms and delisted firms) に対する2018年度の連結主体で抽出した財務(諸表) データと単体主体で抽出した財務(諸表) データをラングリングしたものを結合することについて検討するとともに (cf. 地道 (2020-a, b), 地道, 阪 (2022), 地道ら (2022-a)), 地道ら (2022-c) では, 年度を2009年から2018年度の10年間に拡大し, データラングリングを実行したものを可視化することを試みている. さらに, 地道, 阪 (2023) では, mdx 環境²⁾ とローカル環境を協調することによって上記のデータを前処理, ラングリング, 可視化することについて議論している. 本稿では, この結果にもとづき, 特にデータの詳細な可視化を与え, その結果を考察する. なお, この工程は, 規模の大きな財務データに対して探索的データ解析 (Tukey (1977)) を実行することを意味していることに留意しよう.

本稿の構成は以下のようなものである. 利用するデータセットに関する情報を与えたもとで (II 節), データ可視化に関する結果からの考察を行う (III 節). 最後に, 今後の課題などを与える (IV 節).

付録 A には, 本研究で利用しているコンピュータ環境に関する情報を与えており, 付録 B に本稿で利用したデータ可視化のための R スクリプトを与えている. なお, 本研究で行われる前処理, ラングリング, 可視化の全工程は, データ解析環境 R³⁾ と make コマンド (cf. Mecklenburg (2015)) による自動実行によって行われている (詳細については, 地道, 阪 (2023) を参照されたい). また, 本稿も Sweave (cf. Leisch (2002)) による動的文書生成 (dynamic documents) によって作成し, 再現可能性⁴⁾ を確保している.

2) mdx は, 研究環境を用途に合わせてオンデマンドで短時間に構築・拡張・融合できる, データ収集・集積・解析のためのプラットフォームであり, データ活用社会創成プラットフォームとして位置づけられる (<https://mdx.jp>). なお, 詳細は Suzumura *et al.* (2022) を参照されたい.

3) <https://www.r-project.org>

4) R を利用した動的文書生成については, 例えば, Xie (2015) を, また, 再現可能研究に関しては, Peng (2011), Gandrud (2020) を参照されたい.

II データセット

本稿で扱うデータセットについて簡単に説明する。なお、データベースからのデータ抽出や、その処理の詳細は、地道 (2020-a), 地道, 阪 (2022, 2023), 地道ら (2020-a, b, 2022-a, c) を参照されたい。

本研究で利用するデータ属性 (企業情報・会計情報) は、以下のようなものである (地道, 阪 (2022) も参照)。

firmid: 企業名 + BvD ID (例: "ITOHAM YONEKYU HOLDINGS INC. JP
6011001110293"; 伊藤ハム米久ホールディングス + BvD ID)
id: BvD ID (例: "JP6011001110293"; 伊藤ハム米久ホールディングス
の BvD ID)
year: 会計年度 (2009, ..., 2018)
country: 国情報 (例: "Japan"; 日本)
listed: 上場情報 ("Delisted": 上場廃止, "Listed": 上場, "Un-
listed": 非上場)
cons: 連結コード ("C1", "C2", "U1", "U2")
exchange: 主取引所 (例: "New York Stock Exchange (NYSE)"; ニュー
ヨーク証券取引所)
date: 決算年月日
month: 決算月数
practice: 会計基準 ("IFRS": 国際会計基準, "Local GAAP": 国内基準)
infoprov: インフォメーションプロバイダ
assets_fix: 固定資産 (単位: 1,000 US ドル)
assets_cur: 流動資産 (単位: 1,000 US ドル)
assets_total: 総資産 (単位: 1,000 US ドル)
shareholders: 株主資本 (単位: 1,000 US ドル)
debt_long: 固定負債 (単位: 1,000 US ドル)

liabilities_cur: 流動負債 (単位: 1,000 US ドル)
employees: 従業員数 (単位: 人)
operating_revenue: 営業収益 (単位: 1,000 US ドル)
ebit: 営業利益 (単位: 1,000 US ドル)
pl_before_tax: 税引前利益 (単位: 1,000 US ドル)
tax: 税金 (単位: 1,000 US ドル)
net_income: 純利益 (単位: 1,000 US ドル)
costs_employees: 人件費 (単位: 1,000 US ドル)
interest_paid: 支払利息 (単位: 1,000 US ドル)

ここで、国情報は、企業の本社が存在する国を表しており、連結コード cons の種類としては以下のようなものがある：

- C1: 連結財務諸表のみを保有している企業
- C2: 連結財務諸表を保有しており、何らかの理由で単体財務諸表も保有している企業
- U1: 単体財務諸表のみを保有している企業
- U2: 単体財務諸表を保有しており、何らかの理由で連結財務諸表も保有している企業

III データ可視化

本節では、前節までで示した、Bureau van Dijk (BvD) 社のデータベース Orbis から抽出し、前処理・データラングリングを終えた、(最大) 世界164カ国の全非上場企業 (一般事業会社) 2,558万社超⁵⁾ の2009-2018年度 (10年間) の財務諸表データ (CSV形式で約 20GB) を用いて、世界の非上場企業の財務分析、付加価値分配等の探索的データ解析を行う。なお、可視化に

5) 本稿で扱っている企業数は、正確には、25,584,931社である。

用いた各国の非上場企業数を付録 C に与える。

1. 非上場企業の売上合計等

世界の非上場企業の実態をみる上で、まず、各国の非上場企業の経済規模を確認する。各国全非上場企業の2018年度の売上合計について、ドットプロットで可視化したものが図1である。縦軸には国名（アルファベット順、降順）、横軸には各国の全非上場企業の売上合計を示している。図1からは、非上場企業の売上合計が最も大きい国は中国であること、次いで、イギリス、日本、フランス、イタリアの順であることがわかる。

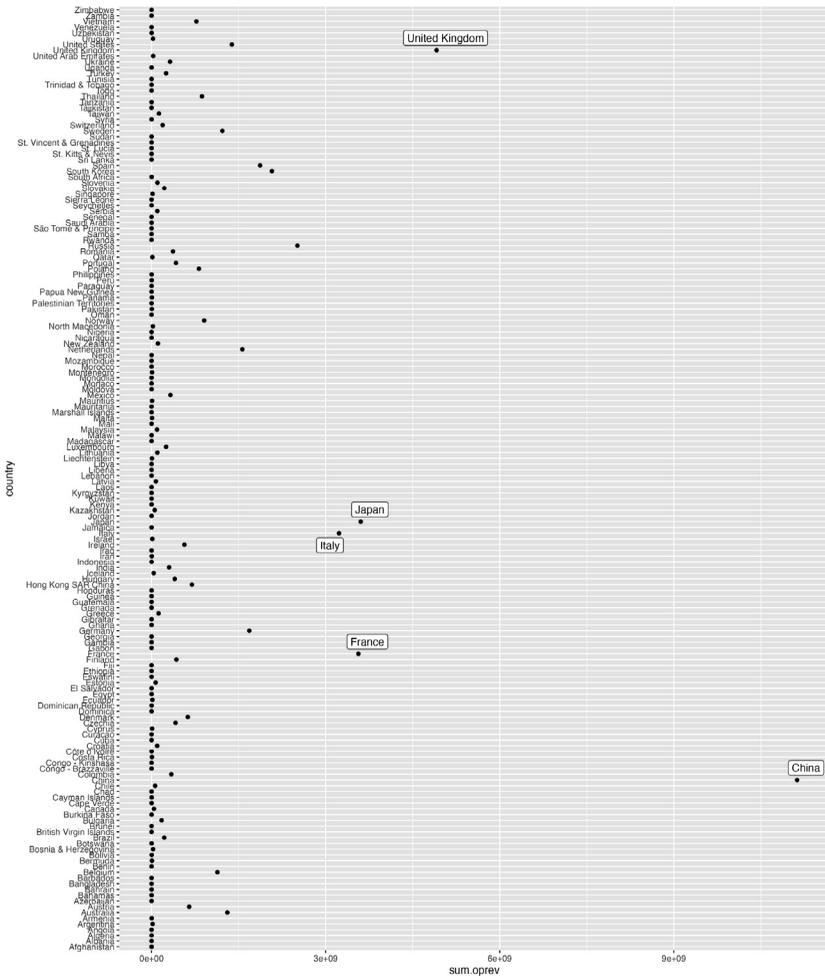


図1：世界164カ国の全非上場企業の売上合計（2018年度）のドットプロット

次に、GDP 上位10カ国における非上場企業の売上分布について確認する。企業の売上額の分布は極端に歪んでいることが確認されているため (cf. Jimichi *et al.* (2018)), 対数表示によって、10年間（2009-2018年度）の国毎の全非上場企業の売上分布をリッジ（ライン）プロット (ridge(line) plot)

で可視化したものが図2である。各プロットの縦軸は会計年度毎の推定された密度、横軸には売上金額（対数）を示している。この結果から、日本企業（JP）や韓国企業（KR）は10年前から売上額の集中度が高いこと、ドイツ企業（DE）やアメリカ企業（US）では売上の集中度が高くなってきたこと、また、アメリカ企業の売上分布は左に歪んでいることがわかった。

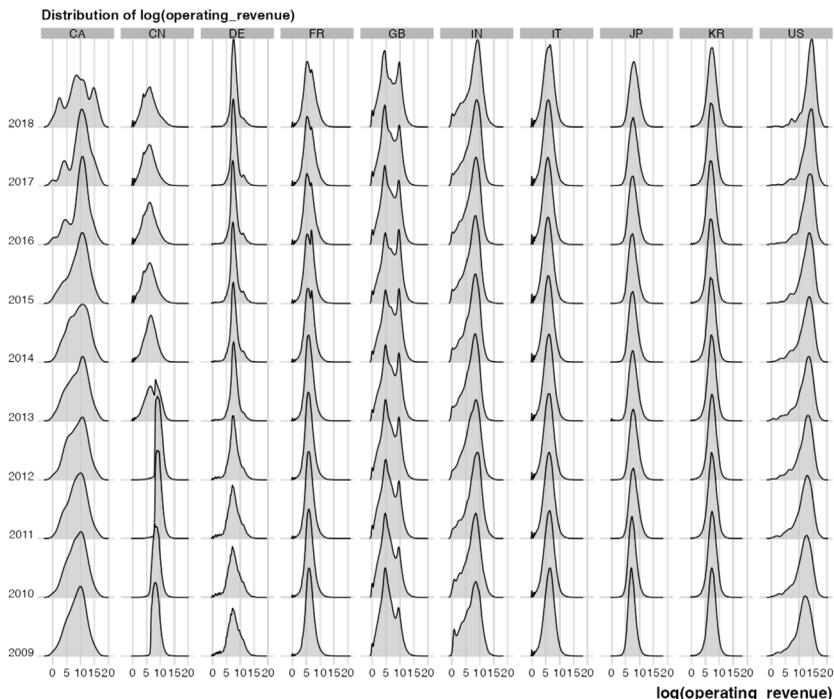


図2：GDP 上位10カ国の全非上場企業の対数売上高のリッジプロット（2009-2018年度）

世界の非上場企業のフロー（売上高）とストック（総資産）の経済規模の分布を確認するために、2018年度の非上場企業に関する対数総資産 $\log(\text{assets_total})$ と対数売上高 $\log(\text{operating_revenue})$ の散布図を与える（図3参照⁶⁾。

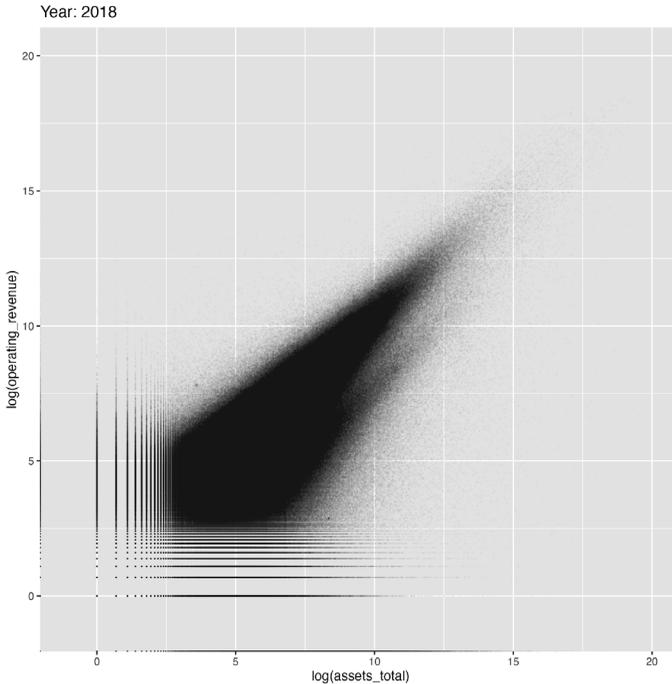


図3：2018年度の非上場企業に関する対数総資産と対数売上高の散布図

図3をよく見ると、峰は少なくとも2つ存在することがわかり、何らかの混合分布に従うことが見て取れる。この要因については、詳細な分析が必要であるため、別稿にゆずることとする。

2. 非上場企業の収益性・安全性・生産性

企業の財務状況を確認するには、企業が利益を稼ぐ能力を示す収益性分析、企業の支払能力や財務的な安全性を示す安全性分析、投入量に対する生産量の割合を示す生産性分析が代表的である。そこで、世界の非上場企業の収益性分析、安全性分析、生産性分析を実施する。

6) Jimichi *et al.* (2018) では、世界の上場企業に関する売上高を従業員数と総資産によって統計モデリングを行うために、散布図を描くことによって知見を得ている。

まず、企業が利益を稼ぐ能力を示す指標である収益性を確認するにあたり、世界118カ国の全非上場企業の2018年度の(1)式で定義される自己資本利益率 (Return on Equity: ROE) を確認する：

$$\text{自己資本利益率} := \frac{\text{当期純利益}}{\text{純資産}} \iff \text{ROE} := \frac{\text{net_income}}{\text{shareholders}} \quad (1)$$

ここでは、「純資産」の代替指標として「株主資本」(shareholders) を利用している。

自己資本利益率のヒストグラムを描くことによって分布状況を調べる。図4の左図は、「 $-10 < \text{ROE} < 10$ 」($\pm 1000\%$)の範囲に限定したものであり、右図は、原点付近の分布状況を確認するために、表示範囲を「 $-0.5 < \text{ROE} < 0.5$ 」($\pm 50\%$)に限定したものである。いずれも各プロットの縦軸は企業数、横軸はROE(表示単位 $\times 100$ で百分率(%)表示となる)をとっている。なお、世界118カ国の全非上場企業のROE(2018年度)の中央値は0.083(=8.3%)であり、全上場企業の場合の中央値5.7%より高い。

図4の左図におけるラグプロット⁷⁾から、世界の非上場企業全体のROEの広がりか極めて大きいことがわかり、ROEの尖度が0付近で極めて高いことがわかる。また、右図からは、分布がROE=0で非対称となっていることが確認できる。ROEの分子である当期純利益は、収益から費用を差し引いたものであるため、正にも負にもなり得るが、右図の0付近の企業が集中している部分に着目すると、0より左(わずかに当期純損失)の企業数は、0より右(わずかに当期純利益)の企業数より明らかに少なく、作為的な会計処理の可能性も考えられる。世界の全上場企業のROEの分布においても同様の傾向は確認できたが、非上場企業の場合は0付近の企業数が極端に多くなっている(cf. 阪(2021-b))。

7) ラグ(敷物)プロット(rug plot)とは、ヒストグラム等において、データ位置に垂直線(1)をプロットしたものである。このプロットによってヒストグラムにデータの分布の密度の情報を付加することができる。

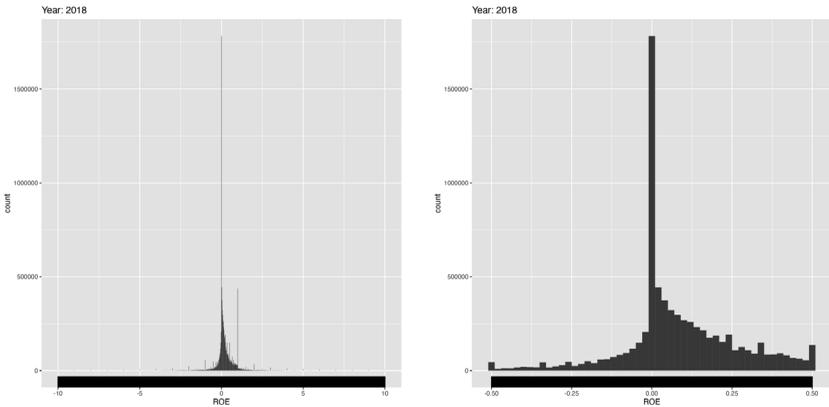


図4：世界118カ国の全非上場企業の自己資本利益率 ROE のヒストグラム（2018年度）：左図：範囲を $-10 < \text{ROE} < 10$ に指定，右図：範囲を $-0.5 < \text{ROE} < 0.5$ に指定

また、国毎の ROE 水準を比較するために、2018年度の日本を含む GDP 上位10カ国の全非上場企業の国別 ROE を “ $-0.5 < \text{ROE} < 0.5$ ”（±50%）の範囲でボックスプロットを描いたものが図5である。ボックスプロットにおいて、ボックスの中の線が中央値、ボックスの左端が第1四分位点 Q_1 、右端が第3四分位点 Q_3 、ボックスのひげの端は $[Q_1 - 1.5Q, Q_3 + 1.5Q]$ ⁸⁾ の範囲で最も遠くのデータ点までの距離を表し、その外側の点は外れ値を示している。図5の縦軸は国名（ISO 国名コード）、横軸には ROE（小数点表示）を示している。

ROE の中央値の上位5位は、フランス企業（FR）13.9%、イギリス企業（GB）13.3%、ドイツ企業（DE）10.2%、アメリカ企業（US）8.5%、イタリア企業（IT）7.8%である。ROE の中央値が低いのはカナダ企業（CA）の -7.1%である。カナダ企業の ROE は四分位範囲が広く（集中度が低い）、一方で、アメリカ企業（US）や中国企業（CN）は ROE の四分位範囲が狭い（集中度が高い）ことがわかる。日本（JP）の非上場企業の ROE の中央

8) ボックスプロットにおいて、 $Q := Q_3 - Q_1$ は四分位範囲と呼ばれる。

値は7.1%であり，非上場企業全体の ROE の中央値8.3%と比較して，収益性はやや低いことがわかる。

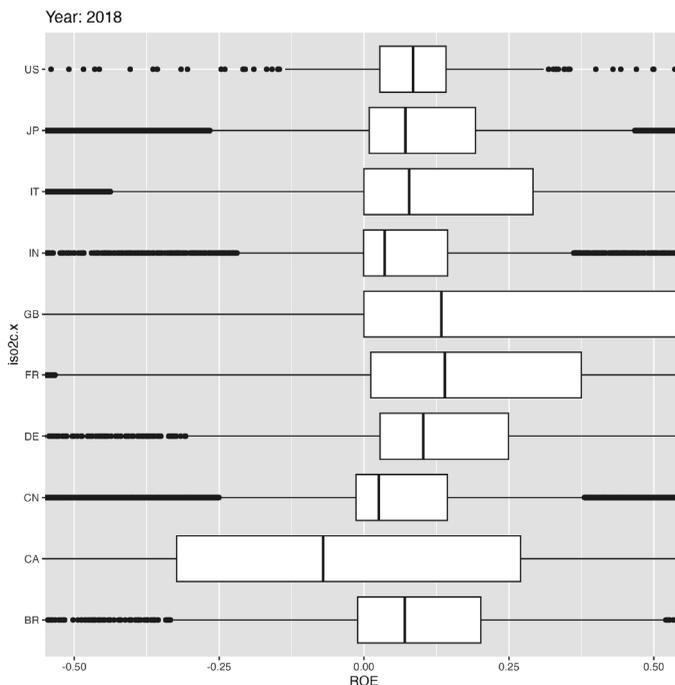


図5：GDP 上位10カ国の非上場企業の ROE の国別ボックスプロット（2018年度）：表示範囲を“ $-0.5 < \text{ROE} < 0.5$ ”（ $\pm 50\%$ ）に指定

次に，企業の支払能力や財務的な安全性を確認するにあたり，まず，自己資本比率（Capital Ratio: CR）を，

$$\text{自己資本比率} := \frac{\text{純資産（株主資本）}}{\text{総資産}} \iff \text{CR} := \frac{\text{shareholders}}{\text{assets_total}} \quad (2)$$

で定義し，世界120カ国の全非上場企業の2018年度に関する自己資本比率 CR のヒストグラムを描くことによって，分布状況を可視化した（図6）．図6の左図は，表示範囲を“ $-10 \leq \text{CR} \leq 20$ ”（ $[-100\%, 200\%]$ ）に指定したヒストグラムである．右図は，表示範囲を“ $0 \leq \text{CR} \leq 1$ ”（ $[0\%, 100\%]$ ）と指定

したものである。いずれのヒストグラムも、縦軸は企業数、横軸には自己資本比率（表示単位×100で百分率となる）を示している。なお、2018年度の世界120カ国の全非上場企業の自己資本比率の中央値は0.429（=42.9%）である。

図6からは、自己資本比率100%（負債が0）の企業が最も多いこと、自己資本比率が0%に近い企業も次に多く存在することがわかる。上場企業の場合は自己資本比率50%付近の企業が最も多く、全上場企業の中央値58.8%と比較すると（cf. 阪（2021-b））、安全性の分布については上場企業と非上場企業に明確な違いがあることがわかる。

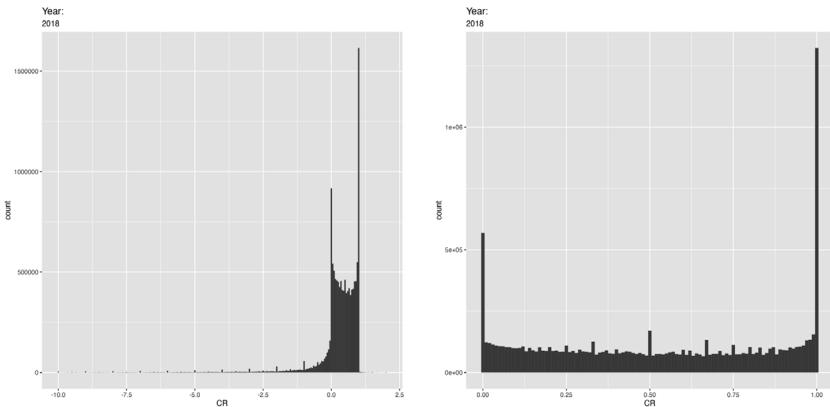


図6：世界120カ国の全非上場企業の自己資本比率（CR）のヒストグラム（2018年度）：左図：範囲を $-10 < CR < 2$ に指定，右図：範囲を $0 \leq CR \leq 1$ に指定

日本を含むGDP上位10カ国の全非上場企業の国別の自己資本比率CRの分布（2018年度）を、ボックスプロットで可視化したものが図7である。縦軸はISO国名コード、横軸には自己資本比率を示している。

図7からは、自己資本比率の中央値は高い順に、ブラジル企業（BR）60.5%、カナダ企業（CA）57.8%、インド企業（IN）48.7%となっている。一方、日本企業（JP）の自己資本比率の中央値29.2%は、イタリア企業（IT）24%に次いで低いことがわかる。カナダ企業は四分位範囲が広く（集中度が低

い), アメリカ企業 (US) は四分位範囲が狭い (集中度が高い), 日本 (JP) の非上場企業の自己資本比率の中央値は29.2%であり, 全上場企業の自己資本比率の中央値42.9%と比較して, 安全性がやや低いことがわかる (cf. 阪 (2021-b)).

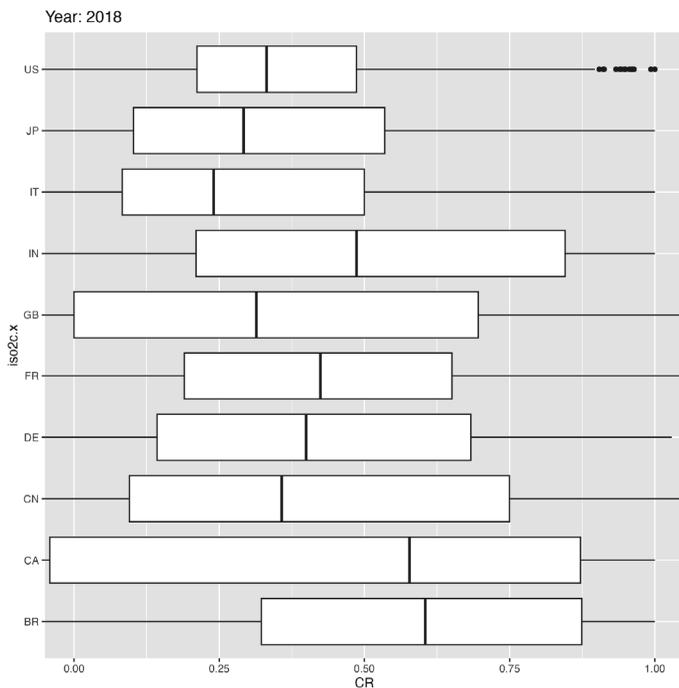


図7：GDP 上位10カ国の非上場企業の自己資本比率 CR の国別ボックスプロット (2018年度)

さらに, 企業の投入量に対する生産量の割合を示す指標である生産性として, 以下の指標を取り上げる:

$$1 \text{ 人あたり売上高} := \frac{\text{売上高}}{\text{従業員数}} \iff \text{OE} := \frac{\text{operating_revenu}}{\text{employees}} \quad (3)$$

国毎の比較をするために, 日本を含む GDP 上位10カ国の全非上場企業の1人あたり売上高の分布をボックスプロットで可視化したものが図8である.

縦軸は国名（ISO 国名コード），縦軸には1人当たり売上高（単位：千 US ドル）を示している．図8からは，1人当たり売上高の中央値は高い順に，カナダ企業（CA），アメリカ企業（US），日本企業（JP）となっており，一方，低い順ではインド企業（IN），ブラジル企業（BR）であることがわかる．アメリカ企業は四分位範囲が広く（集中度が低い），ドイツ企業（DE）やイタリア企業（IT）は四分位範囲は狭い（集中度が高い）．日本（JP）の非上場企業の1人当たり売上高の中央値は287,800 US ドルであり，全非上場企業の1人当たり売上高56,300 US ドルと比較しても，生産性は低くはないことがわかる．なお，全上場企業の1人当たり売上高は1,427,000 US ドルである．

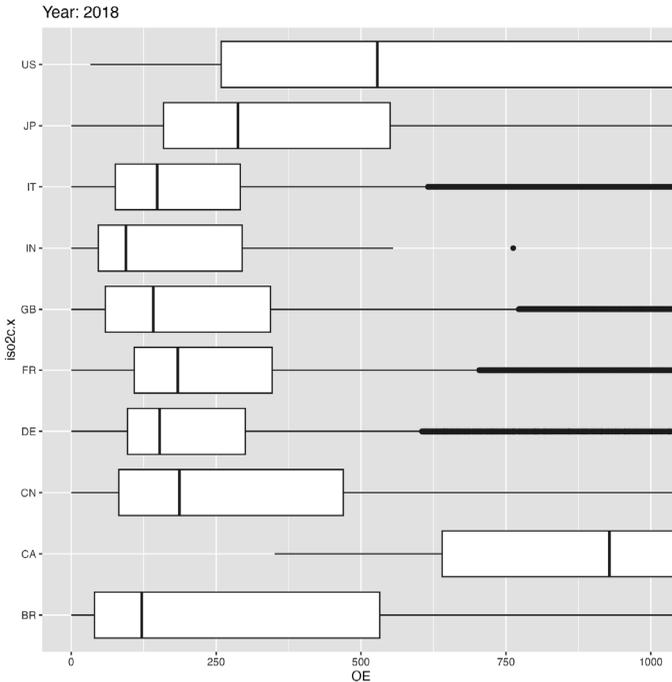


図8：GDP 上位10カ国の非上場企業の労働生産性（1人当たり売上高）の国別ボックスプロット（2018年度）

3. 非上場企業の付加価値分配

ここからは、ビジネスのメインストリームにおいてステークホルダー資本主義が注目されていることから、ステークホルダーの観点からの分析として付加価値分配の実態をみていくこととする。ステークホルダー資本主義については、アメリカの経済団体ビジネス・ラウンドテーブルが、2019年8月に発表した声明において、株主第一主義を脱し、顧客、従業員、サプライヤー、地域社会、株主などのすべてのステークホルダーの利益を尊重する資本主義の実現を目指すとしたことで注目されてきた。日本でも、「新しい資本主義」において、成長と分配の好循環などが掲げられている（内閣官房, 2022）。分配にはさまざまな組織が関与するが、実質的にその多くの部分を担っているのは企業である。企業が生み出した付加価値は、当該活動の関係者であるステークホルダーに対して、債権者に支払利息、従業員に人件費、政府に税金、投資家に配当原資などのもととなる利益として分配される。なお、付加価値とは、企業の経営活動を通じて新たに産出された価値である。付加価値は産出面と分配面の2つの側面から算出することができ、それぞれ控除法および加算法と呼ばれる。ここでは、付加価値を各ステークホルダーへの分配構成要素の合計額として求める加算法に基づき、付加価値の構成要素として固定資産の原価の配分である減価償却費を含めない純付加価値概念を用いる。対象とする主要なステークホルダーとしては、従業員、債権者、国や地方自治体、株主を考慮する（cf. Freeman (2004)）。そこで、付加価値の構成要素を、労働の対価として従業員に分配される人件費、債権者に支払われる借入金や社債の利息（金融費用）、国や地方自治体（政府）への分配額としての租税公課・法人税等、配当や社内留保として最終的に株主に帰属する税引後当期純利益とする（cf. 阪他 (2020)）。

企業が生み出した付加価値が、4つのステークホルダーにどのように分配されてきたかを見るために、過去10年間（2009-2018年度）における105カ国の全非上場企業全体の付加価値分配（割合ベース）の実態をスタック・エリア・プロット（Stacked Area Plot: SAP）で可視化したものが図9である

(範囲：[0, 1] (0%, 100%)). 横軸は、時点 (2009-2018年度) を示している. 縦軸は、付加価値総額 (4つのステークホルダーへの分配額の合計額) を100%とし、上から、支払利息として債権者 (最上段)、人件費として従業員 (上から2段目)、税金として政府 (上から3段目)、利益を通して株主 (最下段) への分配割合を示している. また、図の上の Range の数値は、左が対象期間 (2009-2018年度) における最も企業数が少ない年度における企業数 (多くは2009年度)、右が最も企業数が多い年度 (多くは2018年度) における企業数を示している.

図9から、各ステークホルダーへの分配割合は過去10年間でほぼ変化していないこと、また、従業員への分配が高く維持されていることがわかる. 世界の全上場企業の付加価値分配の実態では、世界全体として従業員への付加価値分配割合が低下し、投資家への分配が増えていることが確認された (阪他, 2020参照) ことと比較すると、その違いは明らかであり、非上場企業では株主からの圧力が少なく、従業員への分配が維持されているという帰結もたらされたと考えられる.

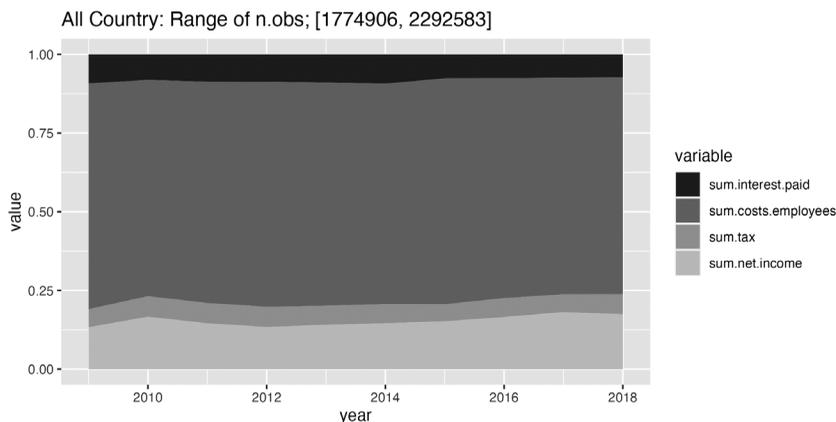


図9：世界105カ国の非上場企業の付加価値分配のスタック・エリア・プロット (2009-2018年度)

次に、主要各国における全非上場企業の過去10年間 (2009-2018年度) の

付加価値分配状況を図10から図16に示す。範囲、横軸、縦軸、Rangeの表記は図9と同様である。図10から図16は、ドイツ企業、フランス企業、イギリス企業、アメリカ企業、日本企業、インド企業、イラン企業の付加価値分配状況のスタック・エリア・プロットを示している。なお、非上場企業では人件費が入手できない国・年などが多いため、10年間の付加価値分配状況が確認できる国が限定される。また、国ごとの図を描画するにあたり、年・国ごとに各付加価値構成要素合計が負になる場合があったため（例えば、当該年・国の非上場企業の税引後当期純利益・損失の合計値が負となる場合）には、0（ゼロ）としてプロットを行った。その場合は他の要素への分配割合が相対的に大きく表示されていることに注意されたい。

図10から図12のプロットより、特に大陸ヨーロッパの企業では従業員への高い分配割合が維持されている一方で、アメリカ（図13）では、過去10年間で、企業は従業員への分配割合を減少させ、利益を増加させてきたことが確認できる。日本企業は、アメリカ企業ほどではないが、従業員への分配割合が緩やかに減少し、株主への分配割合が増加傾向にあることがわかる（図14）。インドやイランでは、債権者への分配割合が比較的大きく、従業員への分配割合が比較的低く、最近の5年間では利益が増加していることがわかる（図15、16）。これらの結果を上場企業の付加価値分配の実態（阪他、2020参照）と比較してみると、ヨーロッパ（特に大陸ヨーロッパ）企業の従業員分配割合率が高いこと、アメリカ企業では従業員への分配割合が減少し株主への分配割合が増加していることは、上場企業でも同様の結果がみられている。一方、イランの上場企業では従業員への分配割合がきわめて低かったことに比べると、非上場企業ではより従業員への分配がなされているといえる。

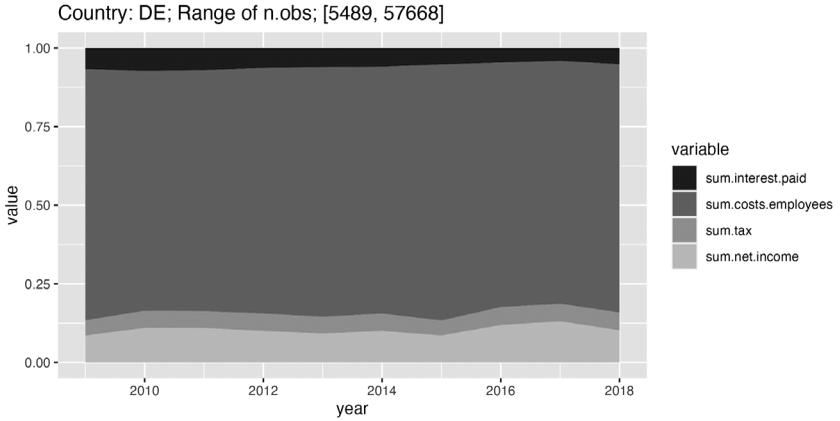


図10：ドイツの非上場企業の付加価値分配のスタック・エリア・プロット（2009-2018年度）

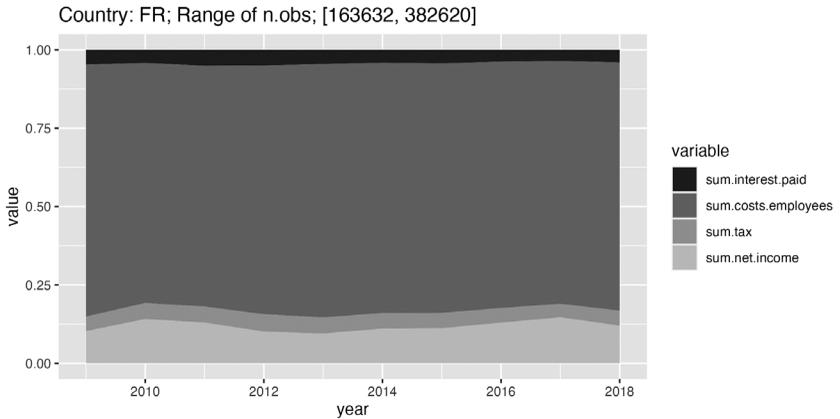


図11：フランスの非上場企業の付加価値分配のスタック・エリア・プロット（2009-2018年度）

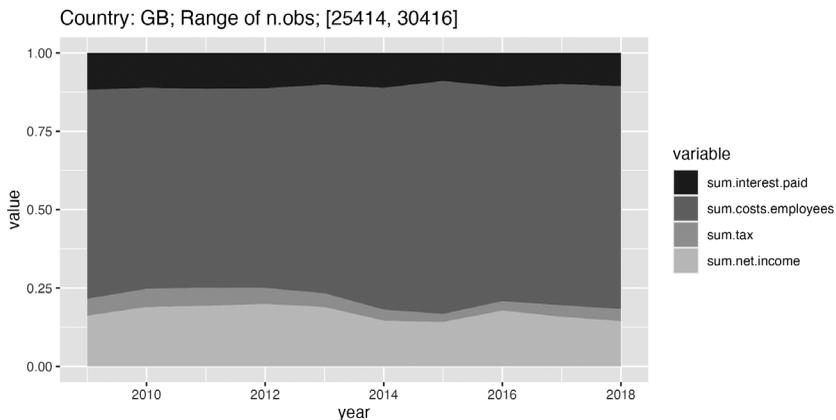


図12：イギリスの非上場企業の付加価値分配のスタック・エリア・プロット (2009-2018年度)

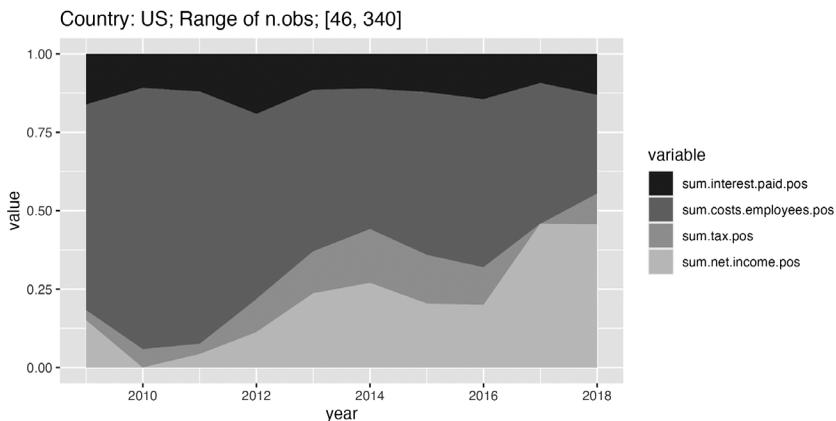


図13：アメリカの非上場企業の付加価値分配のスタック・エリア・プロット (2009-2018年度)

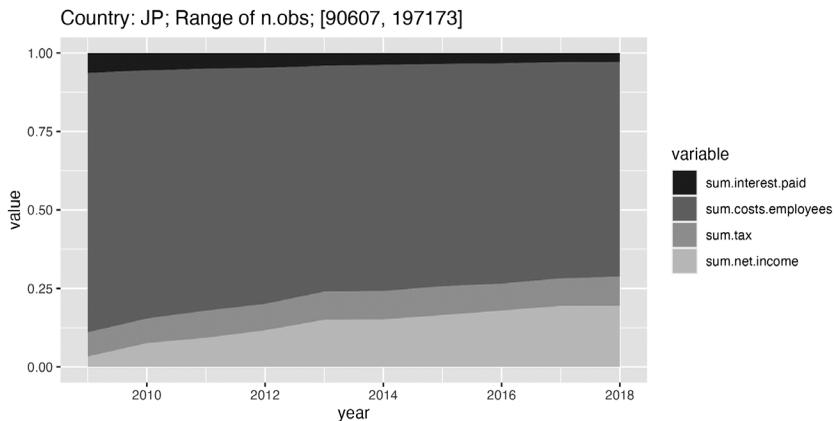


図14：日本の非上場企業の付加価値分配のスタック・エリア・プロット（2009-2018年度）

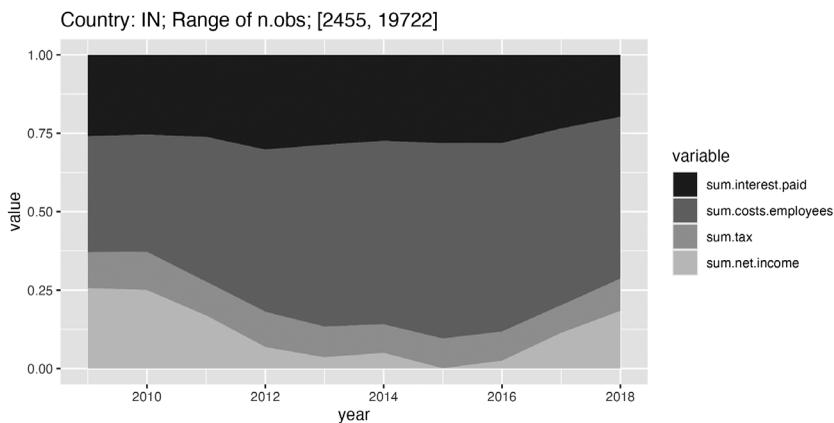


図15：インドの非上場企業の付加価値分配のスタック・エリア・プロット（2009-2018年度）

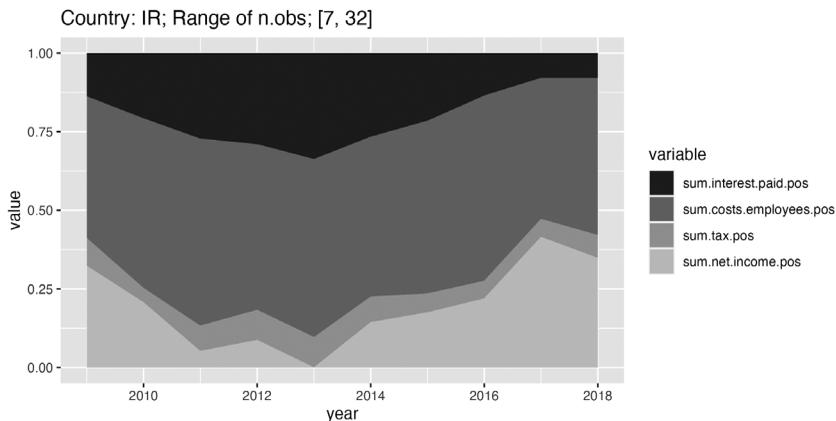


図16：イランの非上場企業の付加価値分配のスタック・エリア・プロット（2009-2018年度）

ステークホルダーへの付加価値分配の中でも，社会的平等に最も関連する労働分配率に焦点を当てる．図10-16のスタック・エリア・プロットでは誌面の都合上，全ての国を掲載することはできなかつたため，各国の非上場企業の以下で定義される労働分配率（Labor Share: LS）の世界地図（World Map）に描いたものが図17である．

$$\begin{aligned} \text{労働分配率} &:= \frac{\text{人件費合計}}{\text{付加価値合計}} \\ &:= \frac{\text{人件費合計}}{\text{支払利息合計} + \text{人件費合計} + \text{税金合計} + \text{純利益合計}} \end{aligned} \quad (4)$$

$$\begin{aligned} \Leftrightarrow \text{LS} &:= \frac{\text{sum.costs_employees}}{\text{VA}} \\ &:= \frac{\text{sum.costs_employees}}{\text{sum.interest_paid} + \text{sum.costs_employees} + \text{sum.tax} + \text{sum.net_income}} \end{aligned} \quad (5)$$

なお，付加価値合計（sum of Value Added: VA）は，年と国毎の各指標の合計を計算しており，労働分配率 LS は [0, 1] (= [0%, 100%]) の範囲に限定している．図17では，2018年度時点における非上場企業の労働分配率が低い

国から、グラデーションで表わされている。なお、一部の国の労働分配率は極端に高くまたは低くなっているが、これは非上場企業では人件費が入手できない国・年などが多いため、労働分配率が確認できる国が限定されるためである（例えば、当該国の人件費合計が0の場合は図では黄色で描画される）。これらの国を除けば、概ね、ヨーロッパ諸国の労働分配率が高く、その他は労働分配率が低いという全体の傾向が確認できる。

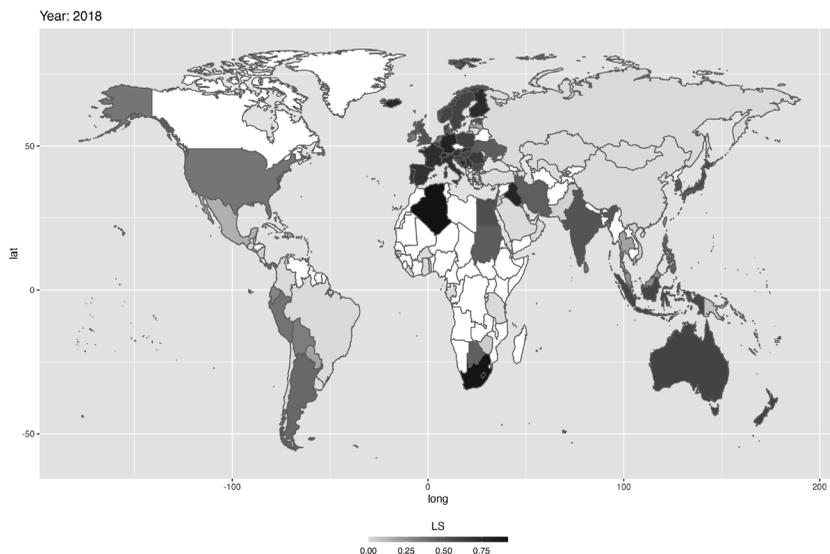


図17：世界120カ国の非上場企業の労働分配率（Labor Share: LS）の世界地図（2018年度）

付加価値を取り巻く企業の状況を明らかにするために、付加価値とそれに関連する財務指標の10年間（2009-2018年度）の相関を確認する。ここでは、データ行列における変量間の相関を色の濃淡として可視化するヒートマップ（heat map）を用いる。2009-2018年度の毎年のヒートマップを作成し、経年変化を見るためにアニメーション化したが、紙面に示すという制約上、2009年度と2018年度の2時点のヒートマップを示したものが図18である。ヒートマップ（図18）では、上段が2009年度時点（77カ国）、下段が2018年

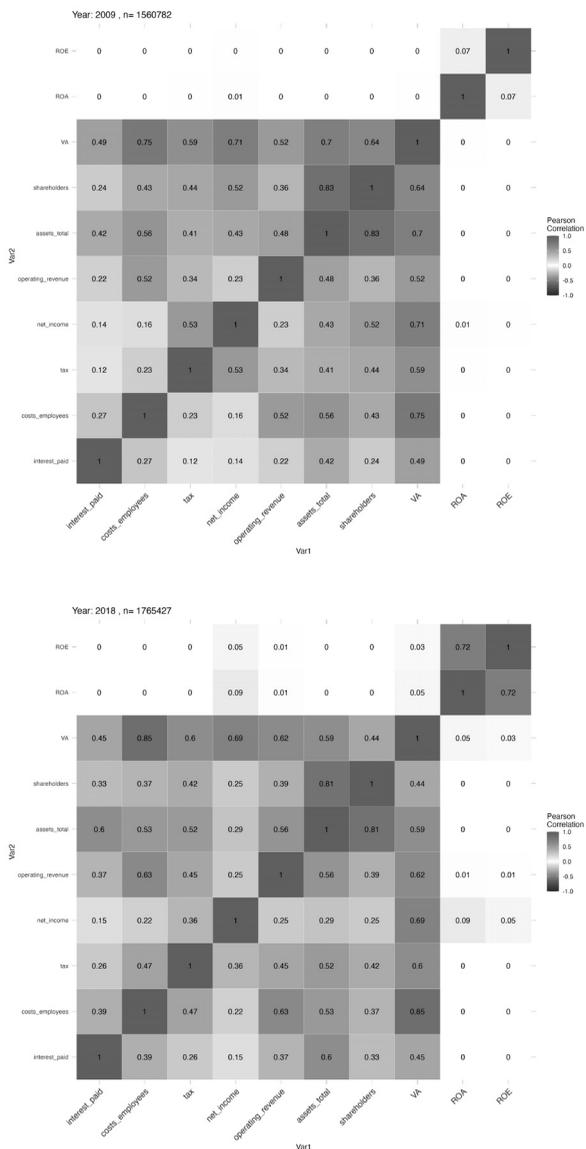


図18：非上場企業の付加価値と関連する財務指標の相関（上段：2009年度（77カ国），下段：2018年度（72カ国））

度時点（72カ国）の（データが入手可能な）非上場企業全体における、（図18の横軸左から）4つの付加価値である支払利息（interest_paid）、人件費（costs_employees）、支払税金（tax）、当期純利益（net_income）、売上（operating_revenue）、総資産（assets_total）、純資産（shareholders）、付加価値合計額（VA）、総資産利益率（Return On Assets: ROA）（(6)式参照）、自己資本利益率（ROE）（(1)式参照）を加えて合計10指標の相関を示している。なお、総資産利益率は

$$\text{総資産利益率} := \frac{\text{当期純利益}}{\text{総資産}} \iff \text{ROA} := \frac{\text{net_income}}{\text{assets_total}} \quad (6)$$

で定義される。

ヒートマップでは、色の濃淡で相関の高低を示しており、相関係数が数値として表示される。ヒートマップのタイトルの数値（n）は、その年において10指標全てのデータが入手可能な企業数を示している。図18からは、例えば、次の指標について、10年間の相関の変化を読み取ることができる。

売上と付加価値合計額の相関	2009年度	0.52	→	2018年度	0.62
売上と人件費の相関	2009年度	0.52	→	2018年度	0.63
売上と支払税金の相関	2009年度	0.34	→	2018年度	0.45
売上と当期純利益の相関	2009年度	0.23	→	2018年度	0.25
付加価値合計額と人件費の相関	2009年度	0.75	→	2018年度	0.85

まず、非上場企業における売上と付加価値合計額の相関が、2009年度に0.52、2018年度は0.62であったことから、2018年の方が売上が増加した場合に付加価値の増加に結び付きやすくなっているといえる。ただし、上場企業の売上と付加価値合計額の相関は、2009年度は0.92、2018年度は0.73であったことから、上場企業ではその相関が低下し、非上場企業ではその相関が増加したにもかかわらず、非上場企業の相関は上場企業より低く、上場企業ほど売上の増加に伴い付加価値が増加しないという実態がわかる。なお、以下で定義される企業毎の付加価値率（Value Added Ratio: VAR）について、全上場企業の中央値が22.2%、非上場企業の中央値が27.3%であることから、

売上から得られる付加価値の割合は、非上場企業の方が高いということがわかる。

$$\text{付加価値率} := \frac{\text{（企業毎の）付加価値合計額}}{\text{売上}} \quad (7)$$

$$\begin{aligned} \Leftrightarrow \text{VAR} &:= \frac{\text{VA}}{\text{operating_revenue}} \\ &= \frac{\text{interest_paid+costs_employees+tax+net_income}}{\text{operating_revenue}} \end{aligned} \quad (8)$$

なお、売上と付加価値合計額の相関と同様のことは、売上と人件費の相関、売上と支払税金の相関、売上と当期純利益の相関についてもみられ、上場企業ではその相関が低下し、非上場企業ではその相関が増加したにもかかわらず、非上場企業の相関は上場企業より低く、上場企業ほど売上の増加に伴い人件費や支払税金や当期純利益が増加しない。ただし、売上と人件費の相関については、非上場企業の2018年度の相関は、上場企業の相関（0.64）とほぼ同じレベルとなっており、また、付加価値合計額と人件費の相関も高くなっていることから、非上場企業では従業員への支払が（他のステークホルダーへの支払よりも）重視されているのではないかと推察できる。

4. 非上場企業の租税回避の蓋然性

前節で取り上げたステークホルダーへの付加価値分配の中でも、パナマ文書等によって租税回避が注目されたこともあり、付加価値分配のうちの政府への分配（税金支払）に焦点を当てる。パナマ文書等では特定の企業や個人が注目されたが、世界の非上場企業全体の実態について明らかにする。なお、租税回避は節税と脱税の中間に位置しており、その境界がはっきりせずグレーな部分も存在するが、ここでは先行研究で企業の租税回避の指標として用いられてきた以下で定義される実効税率（Effective Tax Rate: ETR）を用いて租税回避の蓋然性を確認する：

$$\text{実効税率} := \frac{\text{支払税金}}{\text{税引前利益}} \iff \text{ETR} := \frac{\text{tax}}{\text{pl_before_tax}} \quad (9)$$

過去10年間（2009-2018年度）の各年度の全非上場企業（損失企業を除く）の実効税率 ETR を横軸とし、総資産利益率 ROA を縦軸として、

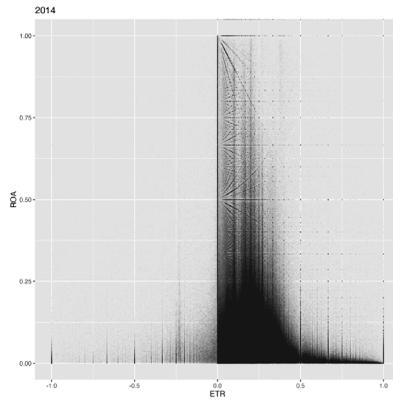
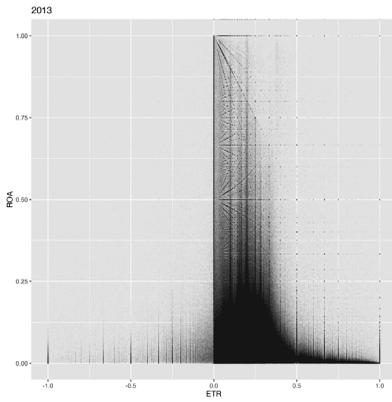
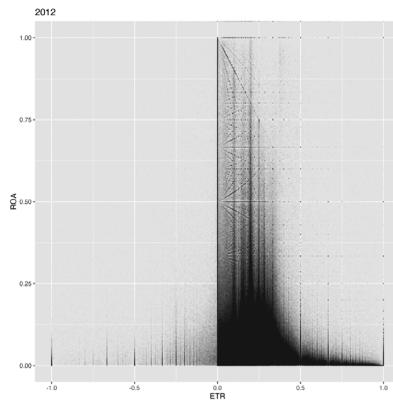
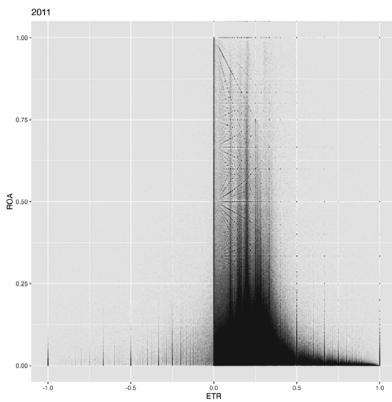
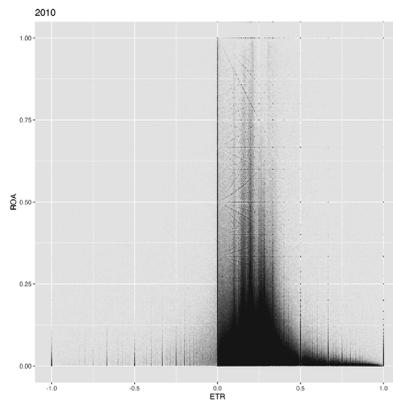
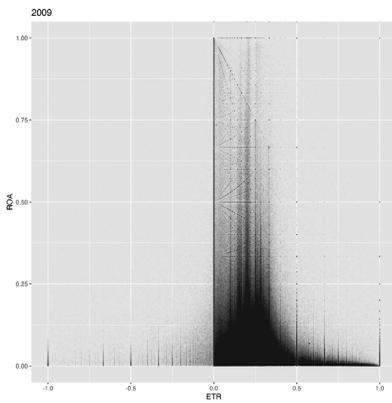
$$(\text{ETR}, \text{ROA}) \in [-1, 1] \times [0, 1] (= [-100\%, 100\%] \times [0\%, 100\%])$$

の範囲で散布図を描いたものが図19である。

国によって法定税率が異なるため実効税率の分布には広がりがあるものの、図19からは、実効税率ゼロ（ETR=0）付近に企業が集中していることがわかる。税額控除や繰延欠損金等の理由で企業の支払税額がゼロとなること自体はありうるが、利益率の相当高い企業（縦軸 ROA の最大値は 1（=100%）である）にも、実効税率ゼロ（ETR=0）の企業が相当数存在することから、作為的な租税回避行動の蓋然性が確認できる。なお、上場企業で確認した場合も同様の現象を確認している（Saka et al. (2019), 地道, 阪 (2020), 阪 (2021-a) 参照）。

図19の2018年度における ETR=0（支払税額がゼロであることを意味する）企業は1,515,991社存在し、その国別の内訳を上位10カ国について表1に与える。

表1ではランク外であるが、ETR=0 の日本企業は4,096社である。2009年度以降毎年の ETR=0 である企業を確認しても、日本企業は少ないことから、日本企業では租税回避行動の蓋然性が低いといえる。なお、上場企業においても、日本企業の租税回避行動の蓋然性は低いことが確認されている（阪 (2021-a) 参照）。



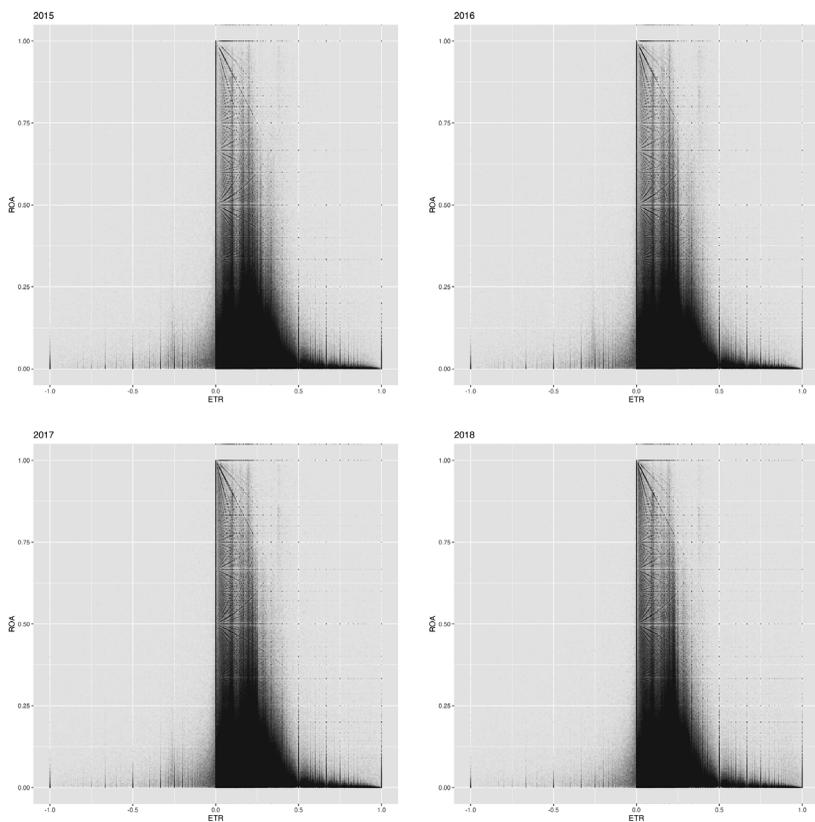


図19：最大147カ国の全非上場企業の利益率（縦軸：ROA）と実効税率（横軸：ETR）（10年間，2009-2018年度）

表 1：2018年度国別実効税率（ETR）がゼロの企業数（obs）：降順

	country	obs
1	Russia	415,423
2	Romania	129,416
3	Bulgaria	106,269
4	Vietnam	106,107
5	Ukraine	95,735
6	France	78,093
7	Hungary	72,940
8	Norway	49,923
9	China	46,006
10	Italy	43,610

IV おわりに

本稿では、データベース Orbis から抽出された非上場企業の財務データを可視化し、それらの結果を得る工程を再現性を確保することによって行った。探索的データ解析の実行によって、売上合計等や、収益性分析（ROE）、安全性分析（自己資本比率）、生産性分析（一人当たり売上高）、付加価値分配、租税回避の蓋然性の実態を示した。上場企業と比較した場合の特徴として、非上場企業では ROE の 0（原点）付近の集中度が高いことや、自己資本比率は 0%と100%付近の企業が多いこと、また労働分配率が高いことなどを明らかにするとともに、日本の非上場企業の特徴として、収益性と安全性がやや低いこと、労働分配率が緩やかな減少傾向にあること、租税回避の蓋然性が低いことが確認できた。本稿で扱った処理工程を全て `mdx` 環境で行うことや、非上場企業に対する財務データの可視化（の結果）についての考察、及び得られた知見にもとづく統計モデリングについては今後の課題としたい。

（筆者らは関西学院大学商学部教授）

参考文献

- [1] Freeman, R. E. (2004) *A Stakeholder Theory of Modern Corporations, Ethical Theory and Business, 7th ed.*, Prentice Hall, NJ.
- [2] Gandrud, C. (2020) *Reproducible Research with R and RStudio, Third Edition*, CRC

Press.

- [3] 地道正行 (2018) 『データサイエンスの基礎：Rによる統計学独習』, 裳華房.
- [4] 地道正行 (2020-a) 『探索的財務ビッグデータ解析：前処理の並列化』, 商学論究, 第67巻, 第3号, pp. 1-19, 関西学院大学商学研究会.
- [5] 地道正行 (2020-b) 『探索的財務ビッグデータ解析：PG-Stromによるデータラングリングの並列化』, 商学論究, 第68巻, 第1号, pp. 1-34, 関西学院大学商学研究会.
- [6] Jimichi, M., D. Miyamoto, C. Saka, and S. Nagata (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, No. 2, pp. 347-371, <https://doi.org/10.1007/s42081-018-0019-1>
- [7] 地道正行, 阪智香 (2020) 『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』, 日本経営数学会誌, 第39巻, 第1・2号, pp. 1-23.
- [8] 地道正行, 宮本大輔, 阪智香, 永田修一 (2020-a) 『探索的財務ビッグデータ解析：PG-Stromによるデータラングリングの並列化』, 日本計算機統計学会, 第34回大会, オンライン開催, 2020年5月31日(日), 講演論文集, pp. 41-44.
- [9] 地道正行, 宮本大輔, 阪智香, 永田修一 (2020-b) 『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12回シンポジウム, オンライン開催, 2020年7月9日(木), 発表用ポスター. <https://jhpcn-kyoten.itc.u-tokyo.ac.jp/abstract/jh191002-NWJ>
- [10] 地道正行, 宮本大輔, 阪智香, 永田修一 (2020-c) 『探索的財務ビッグデータ解析：PG-Stromによるデータラングリングの並列化』, 国際数理科学協会, 2020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, 大阪大学, オンライン開催, 2020年8月22日(土), 配付資料.
- [11] 地道正行, 宮本大輔, 阪智香, 永田修一 (2021-a) 『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13回シンポジウム, オンライン開催, 2021年7月8日(金), 発表用ポスター. <https://jhpcn-kyoten.itc.u-tokyo.ac.jp/abstract/jh211001-NWJ>
- [12] 地道正行, 宮本大輔, 阪智香, 永田修一 (2021-b) 『Rによる財務ビッグデータのラングリング再考』, 2021年度「データ解析環境Rの整備と利用」研究集会, オンライン開催, 2021年12月18日(土), 発表用資料 https://drive.google.com/file/d/1s6gA3m0T9F_xy6nce3nEGuD7GKuEtZKt/view
- [13] 地道正行, 阪智香 (2022) 『探索的財務ビッグデータ解析と再現可能研究：非上場企業のデータラングリング』, 商学論究, 第69巻, 第3・4号, pp. 83-120, 関西学院大学商学研究会.
- [14] 地道正行, 阪智香, 宮本大輔, 永田修一 (2022-a) 『探索的財務ビッグデータ解析と再現可能研究：非上場企業のデータラングリング』日本計算機統計学会, 第36回大会, 愛媛県民文化会館本館 (ハイブリッド開催), 2022年5月22日(日), 講演論文集, pp.

87-90.

- [15] 地道正行, 宮本大輔, 阪 智香, 永田修一 (2022-b) 『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13回シンポジウム, オンライン開催, 2022年7月7日 (木), 発表用ポスター. <https://jhpcn-kyoten.itc.u-tokyo.ac.jp/abstract/jh221001>
- [16] 地道正行, 阪 智香, 宮本大輔, 永田修一 (2022-c) 『探索的財務ビッグデータ解析と再現可能研究：非上場企業のデータ可視化』, 日本計算機統計学会, 第34回シンポジウム, ハイブリッド開催, 2022年11月27日 (日), 講演論文集, pp. 160-163.
- [17] 地道正行, 阪 智香 (2023) 『探索的財務ビッグデータ解析と再現可能研究：mdx環境とローカル環境の協調による非上場企業データのラングリングと可視化の自動化』, 商学論究, 第70巻, 第3号, pp. 123-173, 関西学院大学商学研究会.
- [18] Leisch, F. (2002) *Sweave: Dynamic generation of statistical reports using literate data analysis*, In Wolfgang Härdle and Bernd Rönz, editors, *Compstat 2002-Proceedings in Computational Statistics*, pp. 575-580. Physica Verlag, Heidelberg. ISBN3-7908-1517-9.
- [19] Mecklenburg, R. (2005) *Managing Projects with GNU Make, Third Edition*, O'Reilly Media, Inc.
- [20] 内閣官房・新しい資本主義実現会議 (2022) 『新しい資本主義のグランドデザイン及び実行計画』, 令和4年6月7日.
- [21] Peng, R. D. (2011) Reproducible research in computational science, *Science*, Vol. 334, pp. 1226-1227.
- [22] Saka, C., T. Oshika, and M. Jimichi (2019) Visualization of tax avoidance and tax rate convergence: Exploratory analysis of world-scale accounting data, *Meditari Accountancy Research*, Vol. 27, No. 5, pp. 695-724.
- [23] 阪智香・國部克彦・地道正行 (2020) 『会計と不平等：付加価値分配率の探索的データ解析』, 国民経済雑誌, 第221巻, 第4号, pp. 1-20.
- [24] 阪智香 (2021-a) 『企業の格差と分配：探索的財務データ解析』, ディスクロージャー & IR, 第17巻, pp. 154-163.
- [25] 阪智香 (2021-b) 『アカデミック・フォーサイト：世界の上場企業の財務データ解析』, 会計・監査ジャーナル, 第792号, pp. 76-84.
- [26] 阪智香 (2022) 『DXの進展と会計研究・教育』, 会計, 第201巻, 第6号, pp. 67-79.
- [27] Suzumura, T., A. Sugiki, H. Takizawa, A. Imakura, H. Nakamura, K. Taura, T. Kudoh, T. Hanawa, Y. Sekiya, H. Kobayashi, Y. Kuga, R. Nakamura, R. Jiang, J. Kawase, M. Hanai, H. Miyazaki, T. Ishizaki, D. Shimotoku, D. Miyamoto, K. Aida, A. Takefusa, T. Kurimoto, K. Sasayama, N. Kitagawa, I. Fujiwara, Y. Tanimura, T. Aoki, T. Endo, S. Ohshima, K. Fukazawa, S. Date, and T. Uchibayashi (2022) mdx: A cloud platform for supporting data science and cross-disciplinary research collaborations, *The 8th IEEE International Conference on Cloud and Big Data Computing (CBDCom 2022)*, doi10.48550/ARXIV.2203.14188, <https://arxiv.org/abs/2203.14188>

- [28] Tange, O. (2022) *GNU Parallel 20221022 ('Nord Stream')*, Zenodo, DOI:10.5281/zenodo.7239559, URL: <https://doi.org/10.5281/zenodo.7239559>
- [29] Tukey, J. W. (1977) *Exploratory Data Analysis*, Addison-Wesley Publishing Co.
- [30] Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis, Second Edition*, Springer. (石田基広, 石田和枝共訳 (2011) 『グラフィックスのための R プログラミング: ggplot2 入門』, シュプリンガー・ジャパン株式会社.)
- [31] Wickham, H. and G. Grolemund (2016) *R for Data Science*, O'Reilly. (黒川利明訳, (2017) 『R ではじめるデータサイエンス』, オライリー・ジャパン.)
- [32] Xie, Y. (2015) *Dynamic Documents with R and knitr, Second Edition*, CRC Press.

謝辞

本研究の一部は以下の助成を得ている。

 科学研究費基盤研究 C: 「共有価値創造 (CSV) のための社会環境会計の構築」(2019年~2022年), 課題番号: 19K02006

 科学研究費基盤研究 C: 「企業価値報告とサステナビリティ報告の統合的フレームワークの構築」(2023年~2025年), 課題番号: 23K01689

 2017年度~2023年度学際大規模情報基盤共同利用・共同研究点 (JHPCN) 課題: 「財務ビッグデータの可視化と統計モデリング」, 課題番号: jh171002-NWJ, jh181001-NWJ, jh191002-NWJ, jh201003-NWJ, jh211001-NWJ, jh221001, jh231001

 2023年度統計思考院公募型人材育成事業ワークショップ課題: 「探索的ビッグデータ解析と再現可能研究」, 課題番号: 2023-思考院-7004

 2023年度統計数理研究所公募型共同利用一般研究 2 課題: 「財務ビッグデータの統計モデリングと可視化に関する研究」, 課題番号: 2023-ISMCRP-2017

 関西学院大学図書館図書費 B, 研究設備費 (III), 個人研究費

Moody's 社の増田歩氏にはデータ抽出に関してご足労を賜った。ここに感謝の意を表する。

付録 A コンピュータ環境

mdx 環境

本研究を行うために、東京大学柏キャンパスに敷設された mdx 環境を利用した。各ノードの簡単な仕様を以下に与える⁹⁾：

CPU ノード：

OS: Ubuntu 20.04

CPU: Intel® Xeon® Platinum 8368 Processor (38core, 2.4 GHz) × 2
個

CPU Cores: 152 (=38×2×2)

Main Memory: 229.78 GB

GPU ノード:

OS: Ubuntu 20.04

CPU Cores: 18

Main Memory: 57.6 GB

GPU Unit and Memory: NVIDIA® Tesla® A100, 40 GB

ストレージ: lustre (高速内部ストレージ (DDN ES400NVX) : 8TB, 大容量内部ストレージ (DDN ES7990X) : 8TB)

ローカル環境

Machine: MacBook Pro 2021

OS: macOS Ventura (13.3.1)

CPU: Apple M1 Max (arm64), 10 Cores

Main Memory: 64 GB

Machine: Mac Studio 2022

9) 今回利用しているノードは、ログインノード (1), CPU ノード (3), GPU ノード (1) の合計 5 である (ただし, () 内はノード数である).

OS: macOS Ventura (13.3.1)
CPU: Apple M1 Ultra (arm64), 20 Cores
Main Memory: 128 GB

ソフトウェア環境

- R (R. Ihaka, R. Gentleman, R Core Team, <https://www.r-project.org/>)
- R Packages
 - **{arrow}** (N. Richardson, <https://arrow.apache.org/docs/r/>)
 - **{countrycode}** (V. Arel-Bundock, <https://vincentarelbundock.github.io/countrycode/>)
 - **{RPostgreSQL}** (J. Conway *et al.*, <https://github.com/to-moakin/RPostgreSQL>)
 - **{tidyverse}** (H. Wickham *et al.*, <https://www.tidyverse.org>)
- RStudio (RStudio, <https://www.rstudio.com/>)
- Sweave (F. Leisch, <https://leisch.userweb.mwn.de/Sweave/>)
- PostgreSQL (The PostgreSQL Global Development Group, <https://www.postgresql.org>)
- PG-Strom (K. Kaigai, <https://heterodb.github.io/pg-strom/ja/>)

R 関数 `sessionInfo` を実行することによって、本稿を執筆すること
に利用した R に関する環境情報を以下に与える：

sessionInfo による情報

- R version 4.2.2 (2022-10-31), aarch64-apple-darwin20
- Locale: ja_JP.UTF-8/ja_JP.UTF-8/ja_JP.UTF-8/C/ja_JP.UTF-8/ja_JP.UTF-8
- Running under: macOS Ventura 13.3.1
- Matrix products: default
- BLAS: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRblas.0.dylib
- LAPACK: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRlapack.dylib
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils
- Other packages: arrow 11.0.0.3, countrycode 1.4.0, DBI 1.1.3, dplyr 1.1.2, forcats 1.0.0, ggplot 2.3.4.2, lubridate 1.9.2, purrr 1.0.1, readr 2.1.4, RPostgreSQL 0.7-5, stringr 1.5.0, tibble 3.2.1, tidyr 1.3.0, tidyverse 2.0.0, xtable 1.8-4
- Loaded via a namespace (and not attached): assertthat 0.2.1, bit 4.0.5, bit 64 4.0.5, cli 3.6.1, colorspace 2.1-0, compiler 4.2.2, fansi 1.0.4, generics 0.1.3, glue 1.6.2, grid 4.2.2, gtable 0.3.3, hms 1.1.3, lifecycle 1.0.3, magrittr 2.0.3, munsell 0.5.0, pillar 1.9.0, pkgconfig 2.0.3, R6 2.5.1, rlang 1.1.0, scales 1.2.1, stringi 1.7.12, tidyselect 1.2.0, timechange 0.2.0, tools 4.2.2, tzdb 0.3.0, utf8 1.2.3, vctrs 0.6.1, withr 2.5.0

付録 B データ可視化のための R スクリプト

本稿で利用したデータ可視化のための R スクリプトをスクリプト 1 に与える。

スクリプト 1 : makepng.R

```

1 #-----
2 # パッケージの読み込み
3 #-----
4 library(tidyverse)
5 library(arrow)
6 library(countrycode)
7 library(ggrepel)
8 library(maps)
9 #-----
10 # データの読み込み
11 #-----
12 x <- read_parquet("../DW-Local/Orbis2019c-ul-sub.parquet")
13 x <- x %>%
14   mutate(country = countrycode(iso2c.x, origin = 'iso2c', destination = '
      country.name'))
15 x2018 <- x %>% filter(year == 2018)
16 #-----
17 # 点プロット (dot plot):
18 # 国毎の売上合計
19 #-----
20 x2018.sum.oprev <- x2018 %>%
21   group_by(country) %>%

```

```

22 | summarise(obs = n(),
23 |           sum.oprev = sum(operating_revenue, na.rm = TRUE)) %>%
24 | arrange(country)
25 | #
26 | p.2018cu.sum.oprev <- x2018.sum.oprev %>%
27 | ggplot(aes(country, sum.oprev, label = country)) +
28 | geom_point() +
29 | coord_flip() +
30 | geom_label_repel(data = subset(x2018.sum.oprev, sum.oprev >= 3*10^9))
31 | #
32 | ggsave(
33 |   filename = "./figures/sum.oprev2018cu.png",
34 |   plot = p.2018cu.sum.oprev,
35 |   device = "png", width = 30, height = 35, units = "cm")
36 | #-----
37 | # リッジプロット (ridge plot):
38 | # GDP 上位 10  国の売上合計 (operating_revenue)
39 | #-----
40 | ggplot.ridge.sales <- function(df = x, cntry = "JP")
41 | {
42 |   require(tidyverse)
43 |   require(ggridges)
44 |   x <- df %>%
45 |     select(year, operating_revenue, iso2c.x) %>%
46 |     filter(year >= 2009, year <= 2018, operating_revenue > 0, iso2c.x %in%
47 |           cntry)
48 |   p <- x %>%
49 |     ggplot(aes(x = log(operating_revenue), y = factor(year)))
50 |     p <- p + geom_density_ridges(alpha = 0.6, fill = "lightblue", scale =
51 |           1.5) +
52 |           scale_x_continuous() +
53 |           scale_y_discrete(expand = c(0.01, 0)) +
54 |           labs(x = "log(operating_revenue)", y = NULL,
55 |             title = paste("Distribution_of_log(operating_revenue)") +
56 |             theme_ridges() +
57 |             theme(title = element_text(size = 16, face = "bold")) +
58 |             facet_grid(~ iso2c.x)
59 |           print(p)
60 | }
61 | # GDP Top 10
62 | p.GDP10.sales <- ggplot.ridge.sales(cntry = c("US", "CN", "JP", "DE", "GB",
63 |       "IN", "FR", "IT", "CA", "KR"))
64 | ggsave(
65 |   file = "./figures/RP.GDP10.sales.png",
66 |   plot = p.GDP10.sales,
67 |   dpi = 100, width = 30, height = 25, units = "cm")
68 | #-----
69 | # ヒストグラム (histogram):
70 | # 自己資本利益率 (Return Equity: ROE)
71 | # 自己資本利益率 := 当期純利益/純資産 x 100 %
72 | #-----
73 | ROE2018 <- x2018 %>%

```

```

71 mutate(country = countrycode(iso2c.x, origin = 'iso2c', destination = '
    country.name')) %>%
72 filter(shareholders !=0) %>%
73 mutate(ROE = net_income/shareholders) %>%
74 select(firmid, country, iso2c.x, net_income, shareholders, ROE)
75 #
76 p.histROE2018 <- ROE2018 %>%
77 filter(ROE >= -0.5, ROE <= 0.5) %>%
78 ggplot(aes(ROE)) +
79 geom_histogram(binwidth = 0.02) +
80 geom_rug() +
81 ggtitle("Year:_2018")
82 ggsave(file = "./figures/hist-ROE50-2018cu.png",
83         plot = p.histROE2018,
84         width = 20, height = 20, units = "cm")
85 #
86 p.histROE2018.pm10 <- ROE2018 %>%
87 filter(ROE >= -10, ROE <= 10) %>%
88 ggplot(aes(ROE)) +
89 geom_histogram(binwidth = 0.02) +
90 geom_rug() +
91 ggtitle("Year:_2018")
92 #
93 ggsave(file = "./figures/hist-ROE1000-2018cu.png",
94         plot = p.histROE2018.pm10,
95         width = 20, height = 20, units = "cm")
96 #-----
97 # ボックスプロット (box plot):
98 # GDP 上位 10 か国における 2018 年度の ROE
99 #-----
100 plot.boxplot.GDP10 <- function(df = ROE2018, rg = c(-0.5, 0.5))
101 {
102   require(ggplot2)
103   require(dplyr)
104   GDP10 <- c("US", "CN", "JP", "DE", "IN", "GB", "FR", "BR", "IT", "CA")
105   p <- df %>% filter(iso2c.x %in% GDP10) %>%
106     ggplot(aes(ROE, iso2c.x)) + #, color = as.factor(iso2c.x)) +
107     geom_boxplot() +
108     coord_cartesian(xlim = rg) +
109     ggtitle("Year:_2018")
110   print(p)
111 }
112 #
113 plot.boxplot.GDP10()
114 ggsave(file = "./figures/BP-GDP10-ROE50-2018cu.png",
115         plot = plot.boxplot.GDP10(),
116         width = 20, height = 20, units = "cm")
117 #-----
118 # ヒストグラム (histogram):
119 # 自己資本比率 (Capital Ratio: CR)
120 # 自己資本比率 := 純資産/総資産 x 100 %

```

```

121 #-----
122 CR2018 <- x2018 %>%
123   filter(month == 12, assets_total > 0) %>%
124   mutate(CR = shareholders/assets_total) %>%
125   select(firmid, iso2c.x, shareholders, assets_total, CR)
126 #
127 plot.CR.hist <- function(df = CR2018, yr = 2018, bw = 0.05, rg = c(-10, 2))
128 {
129   p <- df %>%
130     filter(CR >= rg[1], CR <= rg[2]) %>%
131     ggplot(aes(CR)) +
132     geom_histogram(binwidth = bw) +
133     coord_cartesian(xlim = rg) +
134     ggtitle("Year:", yr)
135   print(p)
136 }
137 #
138 ggsave(file = "./figures/hist-CR2018uc.png",
139         plot = plot.CR.hist(),
140         width = 20, height = 20, units = "cm")
141 ggsave(file = "./figures/hist-CR2018uc-0-1.png",
142         plot = plot.CR.hist(bw = 0.01, rg = c(0, 1)),
143         width = 20, height = 20, units = "cm")
144 #-----
145 # ボックスプロット (box plot):
146 # GDP 上位 10 国における
147 # 2018 年度の自己資本比率 (CR)
148 #-----
149 plot.boxplot.CR.GDP10 <- function(df = CR2018, rg = c(0, 1))
150 {
151   require(ggplot2)
152   require(dplyr)
153   GDP10 <- c("US", "CN", "JP", "DE", "IN", "GB", "FR", "BR", "IT", "CA")
154   p <- df %>% filter(iso2c.x %in% GDP10) %>%
155     ggplot(aes(CR, iso2c.x)) + geom_boxplot() +
156     coord_cartesian(xlim = rg) +
157     ggtitle("Year:_2018")
158   print(p)
159 }
160 ggsave(file = "./figures/BP-GDP10-CR-2018cu.png",
161         plot = plot.boxplot.CR.GDP10(),
162         width = 20, height = 20, units = "cm")
163 #-----
164 # ボックスプロット:
165 # 生産性指標の定義:
166 # 1 人あたり売上高 (OE) := 売上高/従業員数
167 #-----
168 OE2018 <- x2018 %>%
169   filter(month == 12, operating_revenue > 0, employees > 0) %>%
170   mutate(
171     country = countrycode(iso2c.x, origin = 'iso2c', destination = 'country
     .name'),

```

```

172   OE = operating_revenue/employees
173 ) %>%
174 select(firmid, country, iso2c.x, OE)
175 #
176 plot.boxplot.OE.GDP10 <- function(df = OE2018, rg = c(0, 10^5))
177 {
178   require(ggplot2)
179   require(dplyr)
180   GDP10 <- c("US", "CN", "JP", "DE", "IN", "GB", "FR", "BR", "IT", "CA")
181   p <- df %>% filter(iso2c.x %in% GDP10) %>%
182     ggplot(aes(OE, iso2c.x)) + geom_boxplot() +
183     coord_cartesian(xlim = rg) +
184     ggtitle("Year:_2018")
185   print(p)
186 }
187 ggsave(file = "./figures/BP-OE-2018uc.png",
188         plot = plot.boxplot.OE.GDP10(rg = c(0, 10^3)),
189         width = 20, height = 20, units = "cm")
190 #-----
191 # スタック・エリア・プロット (SAP)
192 # 支払利息 + 人件費 + 支払税金 + 当期純利益が
193 # 付加価値合計 (Value Added: VA)
194 # に占める割合
195 #-----
196 # 世界全体
197 #-----
198 `%nin%` <- Negate(`%in%`)
199 #-----
200 y.all <- x %>%
201   dplyr::filter(firmid %nin%
202     c("EURL_IMPOLUX_DZ09B0543183",
203       "JOHNSON_&_JOHNSON_POLAND_SP._Z.O.O._PL006934330",
204       "SOUF_BOIS_SAKER_ET_CIE_DZ05B0542751")) %>%
205   dplyr::filter(month == 12) %>%
206   dplyr::filter(
207     !is.na(interest_paid),
208     !is.na(costs_employees),
209     !is.na(tax),
210     !is.na(net_income)) %>%
211   dplyr::group_by(year) %>%
212   dplyr::summarize(
213     obs = n(),
214     sum.interest.paid = sum(interest_paid),
215     sum.costs.employees = sum(costs_employees),
216     sum.tax = sum(tax),
217     sum.net.income = sum(net_income),
218     sum.interest.paid.pos = ifelse(sum.interest.paid >= 0, sum.interest.
219       paid, 0),
220     sum.costs.employees.pos = ifelse(sum.costs.employees >= 0, sum.costs.
221       employees, 0),
222     sum.tax.pos = ifelse(sum.tax >= 0, sum.tax, 0),
223     sum.net.income.pos = ifelse(sum.net.income >= 0, sum.net.income, 0)

```

```

222 | )
223 | #
224 | plot.SAP.all <- function(df = y.all)
225 | {
226 |   require(dplyr)
227 |   require(ggplot2)
228 |   require(reshape)
229 |   n.obs <- df %>% pull(obs)
230 |   p <- df %>%
231 |     data.frame() %>%
232 |     select(year, sum.interest.paid, sum.costs.employees, sum.tax, sum.net.
233 |             income) %>%
234 |     melt(id.vars = c("year")) %>%
235 |     ggplot(aes(x = year, y = value, fill = variable)) + geom_area(position
236 |             = "fill") +
237 |     ggtitle(paste0("All_Country:_Range_of_n.obs;_[" , min(n.obs) , " , "_", max(n.
238 |             obs) , "]" ))
239 |   print(p)
240 | }
241 | #
242 | plot.SAP.pos.all <- function(df = y.all)
243 | {
244 |   require(dplyr)
245 |   require(ggplot2)
246 |   require(reshape)
247 |   n.obs <- df %>% pull(obs)
248 |   p <- df %>%
249 |     data.frame() %>%
250 |     select(year, sum.interest.paid.pos, sum.costs.employees.pos, sum.tax.
251 |             pos, sum.net.income.pos) %>%
252 |     melt(id.vars = c("year")) %>%
253 |     ggplot(aes(x = year, y = value, fill = variable)) + geom_area(position
254 |             = "fill") +
255 |     ggtitle(paste0("All_Country:_Range_of_n.obs;_[" , min(n.obs) , " , "_", max(n.
256 |             obs) , "]" ))
257 |   print(p)
258 | }
259 | #
260 | plot.SAP.standard.all <- function(df = y.all)
261 | {
262 |   require(dplyr)
263 |   require(ggplot2)
264 |   require(reshape)
265 |   n.obs <- df %>% pull(obs)
266 |   p <- df %>%
267 |     data.frame() %>%
268 |     select(year, sum.interest.paid, sum.costs.employees, sum.tax, sum.net.
269 |             income) %>%
270 |     melt(id.vars = c("year")) %>%
271 |     ggplot(aes(x = year, y = value, fill = variable)) +
272 |     geom_area() +
273 |     ggtitle(paste0("All_Country:_Range_of_n.obs;_[" , min(n.obs) , " , "_", max(n.
274 |             obs) , "]" ))

```

```

267 | print(p)
268 | }
269 | #
270 | plot.SAP.pos.standard.all <- function(df = y.all)
271 | {
272 |   require(dplyr)
273 |   require(ggplot2)
274 |   require(reshape)
275 |   n.obs <- df %>% pull(obs)
276 |   p <- df %>%
277 |     data.frame() %>%
278 |     select(year, sum.interest.paid.pos, sum.costs.employees.pos, sum.tax.
279 |            pos, sum.net.income.pos) %>%
280 |     melt(id.vars = c("year")) %>%
281 |     ggplot(aes(x = year, y = value, fill = variable)) +
282 |     geom_area() +
283 |     ggtitle(paste0("All_Country:_Range_of_n.obs;_[" , min(n.obs) , "_", max(n.
284 |                   obs) , "]" ))
285 |   print(p)
286 | }
287 | #
288 | ggsave(file = "./figures/SAP-all.png",
289 |         plot = plot.SAP.all(),
290 |         width = 20, height = 10, units = "cm")
291 | ggsave(file = "./figures/SAP-standard-all.png",
292 |         plot = plot.SAP.standard.all(),
293 |         width = 20, height = 10, units = "cm")
294 | #-----
295 | # スタック・エリア・プロット
296 | # (Stacked Area Plot: SAP):
297 | #-----
298 | # 国別プロット関数の定義
299 | #-----
300 | y <- x %>%
301 |   filter(firmid %nin%
302 |          c("EURL_IMPOLUX_DZ09B0543183",
303 |            "JOHNSON_&_JOHNSON_POLAND_SP._Z_O.O._PL006934330",
304 |            "SOUF_BOIS_SAKER_ET_CIE_DZ05B0542751")) %>%
305 |   dplyr::filter(month == 12) %>%
306 |   dplyr::group_by(year, iso2c.x) %>%
307 |   dplyr::filter(
308 |     !is.na(interest_paid),
309 |     !is.na(costs_employees),
310 |     !is.na(tax),
311 |     !is.na(net_income)) %>%
312 |   dplyr::group_by(iso2c.x, year) %>%
313 |   dplyr::summarize(
314 |     obs = n(),
315 |     sum.interest.paid = sum(interest_paid),
316 |     sum.costs.employees = sum(costs_employees),
317 |     sum.tax = sum(tax),
318 |     sum.net.income = sum(net_income),

```

```

317   sum.interest.paid.pos = ifelse(sum.interest.paid >= 0, sum.interest.
      paid,0),
318   sum.costs.employees.pos = ifelse(sum.costs.employees >= 0, sum.costs.
      employees, 0),
319   sum.tax.pos = ifelse(sum.tax >=0, sum.tax, 0),
320   sum.net.income.pos = ifelse(sum.net.income >= 0, sum.net.income, 0)
321 ) %>%
322 arrange(iso2c.x, year)
323 #
324 plot.SAP <- function(country = "JP", df = y)
325 {
326   require(dplyr)
327   require(ggplot2)
328   require(reshape)
329   df <- df %>%
330     filter(iso2c.x == country)
331   n.obs <- df %>% pull(obs)
332   p <- df %>%
333     data.frame() %>%
334     select(year, sum.interest.paid, sum.costs.employees, sum.tax, sum.net.
      income) %>%
335     melt(id.vars = c("year")) %>%
336     ggplot(aes(x = year, y = value, fill = variable)) +
337     geom_area(position = "fill") +
338     ggtitle(paste0("Country:_", country, ";_Range_of_n.obs;_",min(n.obs),"
      ,_",max(n.obs),"]"))
339   print(p)
340 }
341 #
342 plot.SAP.standard <- function(country = "JP", df = y)
343 {
344   require(dplyr)
345   require(ggplot2)
346   require(reshape)
347   df <- df %>%
348     filter(iso2c.x == country)
349   n.obs <- df %>% pull(obs)
350   p <- df %>%
351     data.frame() %>%
352     select(year, sum.interest.paid, sum.costs.employees, sum.tax, sum.net.
      income) %>%
353     melt(id.vars = c("year")) %>%
354     ggplot(aes(x = year, y = value, fill = variable)) +
355     geom_area() +
356     ggtitle(paste0("Country:_", country, ";_Range_of_n.obs;_",min(n.obs),"
      ,_",max(n.obs),"]"))
357   print(p)
358 }
359 #
360 plot.SAP.pos <- function(country = "JP", df = y)
361 {
362   require(dplyr)

```

```

363 require(ggplot2)
364 require(reshape)
365 df <- df %>%
366   filter(iso2c.x == country)
367 n.obs <- df %>% pull(obs)
368 p <- df %>%
369   data.frame() %>%
370   select(year, sum.interest.paid.pos, sum.costs.employees.pos, sum.tax.
371     pos, sum.net.income.pos) %>%
372   melt(id.vars = c("year")) %>%
373   ggplot(aes(x = year, y = value, fill = variable)) +
374   geom_area(position = "fill") +
375   ggtitle(paste0("Country:_", country, " ;_Range_of_n.obs;_[" ,min(n.obs) ,"
376     , "_" ,max(n.obs) ," ]"))
377 print(p)
378 }
379 plot.SAP.pos.standard <- function(country = "JP", df = y)
380 {
381   require(dplyr)
382   require(ggplot2)
383   require(reshape)
384   df <- df %>%
385     filter(iso2c.x == country)
386   n.obs <- df %>% pull(obs)
387   p <- df %>%
388     data.frame() %>%
389     select(year, sum.interest.paid.pos, sum.costs.employees.pos, sum.tax.
390       pos, sum.net.income.pos) %>%
391     melt(id.vars = c("year")) %>%
392     ggplot(aes(x = year, y = value, fill = variable)) +
393     geom_area() +
394     ggtitle(paste0("Country:_", country, " ;_Range_of_n.obs;_[" ,min(n.obs) ,"
395       , "_" ,max(n.obs) ," ]"))
396   print(p)
397 }
398 #-----
399 # 国別 SAP
400 #-----
401 countries <- c("US", "JP", "DE", "FR", "GB", "IN", "RU", "ID", "IR")
402 #
403 for(i in countries)
404 {
405   ggsave(file = paste0("./figures/SAP-", i, ".png"),
406     plot = plot.SAP(country = i),
407     width = 20, height = 10, units = "cm")
408 }
409 #
410 for(i in countries)
411 {
412   ggsave(file = paste0("./figures/SAP-", i, "-standard.png"),
413     plot = plot.SAP.standard(country = i),
414     width = 20, height = 10, units = "cm")
415 }

```

```

411 }
412 #
413 for(i in countries)
414 {
415   ggsave(file = paste0("./figures/SAP-pos-",i,".png"),
416         plot = plot.SAP.pos(country = i),
417         width = 20, height = 10, units = "cm")
418 }
419 #
420 for(i in countries)
421 {
422   ggsave(file = paste0("./figures/SAP-pos-standard-",i,".png"),
423         plot = plot.SAP.pos.standard(country = i),
424         width = 20, height = 10, units = "cm")
425 }
426 #-----
427 # 世界地図:
428 # 労働分配率 (Labor Share: LS)
429 #-----
430 # 国・地域情報の追加
431 cntnreg <- read_csv("country-region.csv")
432 cntnrvDOSiris <- read_csv("countryBvDOSiris.csv")
433 fullcntr <- full_join(cntnrvDOSiris, cntnreg, by = c("ALP2" = "alpha-2"))
434 #
435 world.map <- map_data("world") %>%
436   filter(region != "Antarctica") %>%
437   fortify()
438 #
439 LS.country <- x %>%
440   select(year, iso2c.x, month, interest_paid, costs_employees, tax, net_
         income) %>%
441   dplyr::filter(month == 12) %>%
442   dplyr::group_by(iso2c.x, year) %>%
443   dplyr::summarize(
444     obs = n(),
445     sum.interest_paid = sum(interest_paid, na.rm = TRUE),
446     sum.costs_employees = sum(costs_employees, na.rm = TRUE),
447     sum.tax = sum(tax, na.rm = TRUE),
448     sum.net_income = sum(net_income, na.rm = TRUE),
449     VA = sum.interest_paid + sum.costs_employees + sum.tax + sum.net_income
         , # Value Added
450     LS = sum.costs_employees/VA # Labor Share
451   ) %>%
452   left_join(fullcntr, by = c("iso2c.x" = "ALP2")) %>%
453   left_join(world.map, by = c("region"))
454 #
455 ggplot_map <- function(df = LS.country, yr = 2018)
456 {
457   require("maps")
458   require("viridis")
459   world.map <- map_data("world") %>%
460     filter(region != "Antarctica") %>%

```

```

461   fortify()
462   p <- world.map %>% ggplot(aes(x = long, y = lat))
463   p <- p + geom_map(data = world.map,
464                   map = world.map,
465                   aes(group = group, map_id = region),
466                   fill = "white", colour="#7f7f7f", linewidth = 0.5)
467   df <- df %>%
468     filter(year == yr, LS >= 0, LS <= 1)
469   p.map <- p + geom_map(data = df,
470                      map = world.map,
471                      mapping = aes(fill = LS, map_id = map_data),
472                      colour="#7f7f7f", linewidth = 0.5) +
473     theme(legend.position = "bottom") +
474     labs(title = paste("Year:", yr)) +
475     scale_fill_viridis(
476       option = "viridis",
477       direction = -1,
478       name = "LS",
479       guide = guide_colorbar(
480         direction = "horizontal",
481         barheight = unit(2, units = "mm"),
482         barwidth = unit(50, units = "mm"),
483         draw.ulim = F,
484         title.position = 'top',
485         title.hjust = 0.5,
486         label.hjust = 0.5
487       ))
488   plot(p.map)
489 }
490 for(i in 2009:2018)
491 {
492   ggsave(
493     paste("./figures/LS.world.map-", i, ".png", sep = ""),
494     ggplot_map(yr = i),
495     device = "png", width = 30, height = 20, units = "cm")
496 }
497 #-----
498 # ヒートマップ:
499 # 変量間の相関係数
500 #-----
501 hmap_df <-x %>%
502   select(year,
503          month,
504          iso2c.x,
505          firmid,
506          interest_paid,
507          costs_employees,
508          tax,
509          net_income,
510          operating_revenue,
511          assets_total,

```

```

512     shareholders,
513     pl_before_tax) %>%
514 dplyr::filter(month == 12) %>%
515 dplyr::filter(
516   !is.na(interest_paid),
517   !is.na(costs_employees),
518   !is.na(tax),
519   !is.na(net_income),
520   !is.na(operating_revenue),
521   !is.na(assets_total),
522   !is.na(shareholders),
523   !is.na(pl_before_tax)) %>%
524 dplyr::filter(
525   operating_revenue > 0,
526   assets_total > 0,
527   shareholders > 0) %>%
528 mutate(
529   interest_paid = as.numeric(interest_paid),
530   costs_employees = as.numeric(costs_employees),
531   tax = as.numeric(tax),
532   net_income = as.numeric(net_income),
533   operating_revenue = as.numeric(operating_revenue),
534   assets_total = as.numeric(assets_total),
535   shareholders = as.numeric(shareholders),
536   VA = interest_paid + costs_employees + tax + net_income, # Value Added
537   ROA = net_income/assets_total, # Return On Assets
538   ROE = net_income/shareholders # Return On Equity
539 ) %>%
540 select(year,
541        iso2c.x,
542        firmid,
543        interest_paid,
544        costs_employees,
545        tax,
546        net_income,
547        operating_revenue,
548        assets_total,
549        shareholders,
550        VA,
551        ROA,
552        ROE)
553 #
554 plot.heatmap <- function(df = hmap_df, yr = 2018)
555 {
556   require(tidyverse)
557   require(reshape2)
558   if(yr == 2010)
559   {
560     df <- df %>%
561       filter(year == yr, iso2c.x != "DZ") %>%
562       select(-year, -iso2c.x, -firmid)

```

```

563 } else if(yr == 2017)
564 {
565   df <- df %>%
566     filter(year == yr, net_income > -10^7) %>%
567     select(-year, -iso2c.x, -firmid)
568 } else if(yr == 2018)
569 {
570   df <- df %>%
571     filter(year == yr, net_income > -10^8) %>%
572     select(-year, -iso2c.x, -firmid)
573 }
574 else
575 {
576   df <- df %>%
577     filter(year == yr) %>%
578     select(-year, -iso2c.x, -firmid)
579 }
580 p <- melt(cor(df, use = "p")) %>%
581 ggplot(aes(
582   x = Var1,
583   y = Var2,
584   fill = value)) +
585 geom_tile(color = "white") +
586 scale_fill_gradient2(low = "blue",
587   high = "red",
588   mid = "white",
589   midpoint = 0,
590   limit = c(-1,1),
591   space = "Lab",
592   name="Pearson\nCorrelation") +
593 theme_minimal() +
594 theme(axis.text.x = element_text(angle = 45, vjust = 1, size = 12,
595   hjust = 1)) +
596 coord_fixed() +
597 ggtitle(paste("Year:",yr,",_n=",dim(df)[1])) +
598 geom_text(aes(label = round(value, digits = 2)), color = "black")
599 print(p)
600 }
601 for(i in seq(2009, 2018))
602 {
603   ggsave(
604     paste("./figures/heatmap-", i, ".png", sep = ""),
605     plot.heatmap(yr = i),
606     device = "png", width = 30, height = 30, units = "cm")
607 }
608 #-----
609 # 散布図:
610 # 軸 x: 利益率 (ROA),
611 # 軸 y: 実効税率 (ETR)
612 #-----
613 ROA.ETR <- x %>%

```

```

613 filter(pl_before_tax > 0) %>%
614 mutate(ETR = tax/pl_before_tax,
615         ROA = net_income/assets_total) %>%
616 select(firmid, year, iso2c.x, ETR, ROA)
617 plot.ROA.ETR<- function(df = ROA.ETR, YEAR)
618 {
619   require(ggplot2)
620   require(dplyr)
621   p <- df %>%
622     filter(year == YEAR) %>%
623     ggplot(aes(ETR, ROA)) +
624     geom_point(size = 0.01, alpha = 0.01) +
625     xlim(-1, 1) + ylim(0, 1) +
626     labs(title = YEAR)
627   print(p)
628 }
629 for(i in 2009:2018)
630 {
631   ggsave(paste0("./figures/SP-ROA-ETR-", i, ".png"),
632         plot.ROA.ETR(YEAR = i),
633         width = 20, height = 20, units = "cm")
634 }
635 #-----
636 # 散布図:
637 # x 軸: 対数総資産 log(assets_total)
638 # y 軸: 対数売上高 log(operating_revenue)
639 #-----
640 plot.log.AT.OR <- function(df, YEAR)
641 {
642   require(ggplot2)
643   require(dplyr)
644   p <- df %>%
645     filter(year == YEAR) %>%
646     select(assets_total, operating_revenue) %>%
647     ggplot(aes(log(assets_total), log(operating_revenue))) +
648     geom_point(size = 0.002, alpha = 0.01) +
649     scale_x_continuous(limit = c(-1, 20)) +
650     scale_y_continuous(limit = c(-1, 20)) +
651     ggtitle(paste("Year:", YEAR))
652   print(p)
653 }
654 for(i in 2009:2018)
655 {
656   ggsave(paste0("./figures/SP-log-AT-OR", i, ".png"),
657         plot.log.AT.OR(x, i),
658         width = 20, height = 20, units = "cm")
659 }

```

付録 C 国別企業数

国別の企業数は、本稿で扱っている期間を通じて共通であり、表2によって与えられる。

表2：国別企業数 (obs)

	country	obs
1	Afghanistan	1
2	Albania	353
3	Algeria	70,568
4	Angola	1
5	Argentina	468
6	Armenia	174
7	Australia	27,366
8	Austria	151,439
9	Azerbaijan	10
10	Bahamas	5
11	Bahrain	3
12	Bangladesh	109
13	Barbados	6
14	Belgium	440,437
15	Benin	7
16	Bermuda	123
17	Bolivia	36
18	Bosnia & Herzegovina	47,089
19	Botswana	37
20	Brazil	13,822
21	British Virgin Islands	60
22	Brunei	1
23	Bulgaria	508,696
24	Burkina Faso	16
25	Canada	2,336
26	Cape Verde	5
27	Cayman Islands	182
28	Chad	5
29	Chile	550
30	China	2,513,565
31	Colombia	2,722,402
32	Congo - Brazzaville	4
33	Congo - Kinshasa	10
34	Costa Rica	32
35	Croatia	132,963
36	Cuba	7
37	Curaçao	2
38	Cyprus	5,374
39	Czechia	285,988
40	Côte d'Ivoire	19
41	Denmark	258,774
42	Dominica	1

43	Dominican Republic	10
44	Ecuador	164
45	Egypt	520
46	El Salvador	11
47	Estonia	147,977
48	Eswatini	2
49	Ethiopia	16
50	Fiji	2
51	Finland	202,371
52	France	1,233,618
53	Gabon	12
54	Gambia	10
55	Georgia	3
56	Germany	716,794
57	Ghana	9
58	Gibraltar	2
59	Greece	36,636
60	Grenada	2
61	Guatemala	12
62	Guinea	3
63	Honduras	2
64	Hong Kong SAR China	693
65	Hungary	458,602
66	Iceland	36,746
67	India	93,822
68	Indonesia	337
69	Iran	86
70	Iraq	22
71	Ireland	168,710
72	Israel	1,370
73	Italy	1,082,187
74	Jamaica	8
75	Japan	580,060
76	Jordan	41
77	Kazakhstan	3,210
78	Kenya	5
79	Kuwait	37
80	Kyrgyzstan	13
81	Laos	1
82	Latvia	122,847
83	Lebanon	29
84	Liberia	13
85	Libya	4
86	Liechtenstein	23
87	Lithuania	15,431
88	Luxembourg	55,681
89	Madagascar	11
90	Malawi	12
91	Malaysia	276,560
92	Mali	12
93	Malta	23,123
94	Marshall Islands	9
95	Mauritania	3
96	Mauritius	4,130

97	Mexico	15,253
98	Moldova	63,029
99	Monaco	1
100	Mongolia	74
101	Montenegro	17,805
102	Morocco	255,079
103	Mozambique	4
104	Nepal	1
105	Netherlands	715,559
106	New Zealand	3,546
107	Nicaragua	9
108	Nigeria	37
109	North Macedonia	65,496
110	Norway	301,883
111	Oman	25
112	Pakistan	188
113	Palestinian Territories	5
114	Panama	82
115	Papua New Guinea	8
116	Paraguay	178
117	Peru	1,823
118	Philippines	22,032
119	Poland	288,224
120	Portugal	454,282
121	Qatar	13
122	Romania	862,639
123	Russia	2,274,053
124	Rwanda	12
125	Samoa	3
126	Saudi Arabia	69
127	Senegal	1
128	Serbia	162,735
129	Seychelles	20
130	Sierra Leone	2
131	Singapore	63,684
132	Slovakia	233,949
133	Slovenia	130,028
134	South Africa	2,767
135	South Korea	509,696
136	Spain	1,167,066
137	Sri Lanka	109
138	St. Kitts & Nevis	1
139	St. Lucia	1
140	St. Vincent & Grenadines	1
141	Sudan	4
142	Sweden	487,908
143	Switzerland	582
144	Syria	2
145	São Tomé & Príncipe	1
146	Taiwan	3,381
147	Tajikistan	6
148	Tanzania	37
149	Thailand	723,055
150	Togo	14

151	Trinidad & Tobago	5
152	Tunisia	5
153	Turkey	57,185
154	Uganda	1
155	Ukraine	570,334
156	United Arab Emirates	59
157	United Kingdom	2,986,416
158	United States	7,870
159	Uruguay	4,181
160	Uzbekistan	208
161	Venezuela	20
162	Vietnam	687,071
163	Zambia	24
164	Zimbabwe	30
