

Title	ビッグデータの4Vが企業成果に及ぼす影響
Sub Title	The 4Vs of big data and firm performance
Author	高田, 英亮(Takata, Hidesuke)
Publisher	慶應義塾大学出版会
Publication year	2023
Jtitle	三田商学研究 (Mita business review). Vol.66, No.3 (2023. 8) ,p.195- 206
JaLC DOI	
Abstract	既存研究は、米国のデータを用いて、ビッグデータの3V・4Vが企業の成果・能力に与える影響を分析してきた。それに対して本研究は、日本のデータを用いて、ビッグデータの4Vが企業成果に及ぼす影響を検討する。既存研究と本研究の結果を踏まえると、次の3点が指摘できる。第1に、ビッグデータの4Vのうち、Velocity (データの迅速な分析・活用)とVeracity (正確なデータの分析・活用)が成果の向上に資することである。第2に、Volume (大量のデータの分析・活用)のみでは成果が上がらないことである。この2点は、既存研究と本研究で一貫している。第3に、Variety (多様なデータの分析・活用)と成果の関係については結果が混合的であり、その間に正の関係が成り立つ状況・条件の探索など、今後さらなる検討が必要なことである。
Notes	高橋郁夫教授退任記念号 論文
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234698-20230800-0195

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

ビッグデータの4Vが企業成果に及ぼす影響

高田 英 亮

<要 約>

既存研究は、米国のデータを用いて、ビッグデータの3V・4Vが企業の成果・能力に与える影響を分析してきた。それに対して本研究は、日本のデータを用いて、ビッグデータの4Vが企業成果に及ぼす影響を検討する。既存研究と本研究の結果を踏まえると、次の3点が指摘できる。第1に、ビッグデータの4Vのうち、Velocity（データの迅速な分析・活用）とVeracity（正確なデータの分析・活用）が成果の向上に資することである。第2に、Volume（大量のデータの分析・活用）のみでは成果が上がらないことである。この2点は、既存研究と本研究で一貫している。第3に、Variety（多様なデータの分析・活用）と成果の関係については結果が混合的であり、その間に正の関係が成り立つ状況・条件の探索など、今後さらなる検討が必要なことである。

<キーワード>

ビッグデータ, 4V, Volume, Variety, Velocity, Veracity, 企業成果

1. はじめに

2000年代後半以降、ビッグデータの時代が到来している。パソコンやスマートフォン、モバイル・アプリ、SNS (Social Networking Service)、IoT (Internet of Things) 機器などから、波が次々と押し寄せるようにデータが生み出されている。現在、多くの企業がこれらのデータをうまく活用し、顧客にとって価値のある新しい製品やサービスを作り出そうとしている。また、個々の顧客の関心・行動をめぐる予測の精度を高め、その結果をより良いマーケティング意思決定につなげようとしている。こうした試みにおいて、成功を取めている企業が多く存在する。代表例は、アマゾン・ドット・コムやイーベイ、ターゲット、ネットフリックス、セブン-イレブン・ジャパンであろう。その一方で、失敗に終わっている事例も多く、枚挙にいとまがない。これらの失敗事例では、ビッグデータへの投資に見合った成果を引き出すことができていない。

このことに関連して、調査企業のガートナーは、ビッグデータ・プロジェクトの60～85%近く

が失敗に終わっていると推定している (Asay, 2017; Heudecker & Hare, 2016)。また、ガートナー・ジャパン (2019) は、日本企業におけるビッグデータ活用に関する調査結果を発表している。それによれば、データ利活用の現状について、「全社的に利活用している」と答えた企業が20%、「一部の事業・組織で利活用している」と答えた企業が36%であり、過半数の企業が何らかの形でデータを利活用している。ただし、利活用可能なデータからビジネス成果を得ているかについては、「十分に得ている」という企業がわずか3%、「ある程度得ている」という企業が34%であり、日本においてビッグデータ活用を何らかのビジネス成果につなげている企業は全体の約3分の1に過ぎないといえる。

このように、現実の世界においてビッグデータ活用の成功・失敗が見られる中で、研究の世界においてはその重要な規定要因として(1)ビッグデータ分析能力や(2)ビッグデータの3V・4Vなどに着目した実証分析が行われている。(1)に関する代表的な研究としては、Gupta and George (2016), Wamba et al. (2017), Mikalef et al. (2019, 2020), Ciampi et al. (2021) などが、(2)に関する代表的な研究としては、Johnson et al. (2017), Ghasemaghaei and Calic (2019, 2020), Cappa et al. (2021) などが挙げられる。

本研究は(2)の研究群に属するものであり、ビッグデータの4Vが企業成果に及ぼす影響について検討を行う。分析の結果、既存研究と同様に、ビッグデータの4Vのうち、VelocityとVeracityが企業成果と正の関係にあることが示される。その一方で、VolumeとVarietyに関してはそうした正の関係が見られない。本研究では最後に、4Vと成果の関係について押さえておくべき点と今後検討すべき課題を示す。

2. 既存研究のレビュー

まず、本研究の焦点の説明変数であるビッグデータの4Vとは何かを確認する。4Vはビッグデータの特徴づける重要な要素であり、Volume, Variety, Velocity, Veracityからなる(Cappa et al., 2021; Ghasemaghaei & Calic, 2019)。これらに基づく、ビッグデータとは、「桁違いの量(Volume)、豊富な種類(Variety)を有し、時々刻々と高い頻度で生成される(Velocity)データである。ただし、貴重な情報とともにノイズやエラーを多く含むため、クレンジング(洗浄)作業などを通じてそれらの負の要素を取り除いてはじめて正確性の高い(Veracity)ものになる」といえる(原, 2017; 高田, 2020, 2021)。なお、ビッグデータの3Vというとき、それは最後のVeracity以外のVのことを指す(McAfee & Brynjolfsson, 2012)。

これらのVの影響を経験的に検討する研究では、それぞれのVがデータの分析・活用のコンテキストで把握され、次のように定義されている。Volumeは「大量のデータの分析・活用」、Varietyは「多様なデータの分析・活用」、Velocityは「データの迅速な分析・活用」、Veracityは「正確なデータの分析・活用」である。そして、先にふれたが、これらのVと企業の成果・能力との関係について実証分析を行った代表的な研究として、Johnson et al. (2017), Ghasemaghaei and Calic (2019, 2020), Cappa et al. (2021) という4つがある。これらの研究の重要な点を整理

すれば、表1の通りである。この表1の分析結果における図は、それぞれのVの単独の効果を¹⁾整理したものである。

その単独の効果をVごとに見ていく。Volume（大量のデータの分析・活用）は、4つの研究すべてで検討されている。しかし、分析結果は否定的である。企業の成果・能力を表す被説明変数・媒介変数との間に有意な正の関係を見出しているものはない。サーベイ・データを用いた3つの研究では成果・能力との間に非有意な関係、アーカイバル・データを用いた1つの研究では成果（トービンのQ）との間に有意な負の関係が示されている。Variety（多様なデータの分析・活用）も、4つの研究すべてで検討されている。ただし、分析結果は混合的である。サーベイ・データを用いた3つの研究では成果・能力との間に有意な正の関係が確認されている。他方で、アーカイバル・データを用いた1つの研究では成果（トービンのQ）との間に非有意な関係が示されている。

Velocity（データの迅速な分析・活用）は、サーベイ・データを用いた3つの研究で検討されている。分析結果は肯定的であり、それらの研究すべてで企業の成果・能力との間に有意な正の関係が確認されている。Veracity（正確なデータの分析・活用）は、2つの研究で検討されている。こちらも分析結果は肯定的であり、双方の研究で成果・能力との間に有意な正の関係が示されている。

以上の分析結果を簡潔にまとめると、「Volumeは否定的（×）、Varietyは混合的（△）、VelocityとVeracityは肯定的（○）」である。これが既存研究をレビューした結果、押さえておきたい点の1つである。もう1つは、前述の既存研究すべてが「米国のデータを用いている」という点である。米国は日本と比べてデータ基盤の整備、データ人材の育成、データの異分野間流通が進んでおり、製品やサービスの開発・提供におけるビッグデータの活用で先行している。また、AI（Artificial Intelligence）など先進的な分析技術の活用でも先行している（IPA, 2021; 総務省, 2020）。以上の分析結果は、こうしたデータ活用の先進国におけるものであることを認識しておく必要がある。

3. 仮説

前節でレビューした既存研究は、米国のデータを用いて、ビッグデータの3V・4Vが企業の成果・能力に与える影響を分析していた。それに対して本研究は、日本のデータを用いて、ビッグデータの4Vが企業成果に及ぼす影響を検討する。

1) Johnson et al. (2017), Ghasemaghaei and Calic (2019), Cappa et al. (2021) は、それぞれのVの単独の効果のみならず、何らかの交互作用効果も検討し、次のような結果を見出している。Johnson et al. (2017) は、3Vと顧客の変化の交互作用を分析し、顧客の変化がVolumeと成果の関係を負に調整する一方、Velocityと成果の関係を正に調整することを示している。Ghasemaghaei and Calic (2019) は、Volumeとそれ以外の3Vの交互作用を検討し、Volumeがその他3Vの正の効果を強めることを見出している。Cappa et al. (2021) は、VolumeとVarietyの交互作用を分析し、Varietyのレベルが非常に高いとき、Volumeと成果の負の関係が正に変わりうることを示している。

表1 ビッグデータの3V・4Vが企業の成果・能力に及ぼす影響に関する研究のまとめ

	Johnson et al. (2017)	Ghasemaghaei & Calic (2019)	Ghasemaghaei & Calic (2020)	Cappa et al. (2021)	本研究 (2023)
データ	サーベイ	サーベイ	サーベイ	アーカイバル	サーベイ
サンプル	米国	米国	米国	米国	日本
産業(%) / 企業	消費財 (29.9%) 金融サービス/コンサルティング (20.3%) テクノロジー/通信 (19.5%) 医療/医薬 (9.6%) 運輸/物流 (6.1%) 生産財・化学 (2.7%) その他 (11.9%)	製造 (生産財) (13.7%) 製造 (消費財) (18.3%) サービス (53.7%) 金融 (11.2%) 公益事業 (3.1%)	製造 (33.4%) サービス (50.2%) 公益事業 (13.5%) 金融 (2.9%)	Google Play Store で公開されているモバイル・アプリが1つしかないB2C企業	製造 (生産財) (37.3%) 製造 (消費財) (30.2%) サービス (32.5%)
サイズ	261	280	239	114	209
分析手法	潜在変数得点を用いたSEM	PLS-SEM	PLS-SEM	OLS	SUR
被説明変数 Y	新製品売上高	イノベーション能力	企業成果	トービンの Q	企業成果 (市場・財務成果)
媒介変数 M		データ・ドリブン・インサイト	イノベーション成果		
分析結果					
統制変数	事業規模 事業年齢 製品タミナー	企業規模 産業 収益	企業規模 産業	売上高 負債/資産 企業年齢 ペータ値 産業の ROA 産業の R&D 集約度 事業分野	企業規模 企業年齢 B2C タミナー

(注) サンプルの産業 (%) は、各論文に記載された順番で示している。分析手法に関しては、SEM: Structural Equation Modeling (構造方程式モデリング) ; PLS-SEM: Partial Least Squares-Structural Equation Modeling (部分最小二乗法に基づく構造方程式モデリング) ; OLS: Ordinary Least Squares (最小二乗法) ; SUR: Seemingly Unrelated Regression (見かけ上無関係な回帰)。分析結果に関して、+ の実線: 正で有意、- の実線: 負で有意、破線: 非有意。

既存研究との違いは、まさしく分析対象国が日本という点である。前述のように、日本は米国と比べてデータ基盤の整備とデータ人材の育成の面で遅れている。ただし、これらは全体的に見た場合の傾向であり、個々の企業のレベルで見れば、中にはクラウド対応が進み、効果的・効率的なビッグデータ活用基盤を持つ企業や、社内IT（Information Technology）教育プログラムが充実し、優秀なデータ人材を多く抱える企業が存在する。これらの企業はその物的・人的資源に基づいて、データ分析・活用の迅速さ・正確さを高め、ひいてはマーケティング・ケイパビリティの向上を図ることができる（Erevelles et al., 2016; Takata, 2016）。よって、日本を分析対象とする本研究でも、米国を分析対象とする既存研究と同様に、VelocityとVeracityについては企業成果との間に肯定的な結果が得られると予想する。

もう少し具体的に見ていく。Velocity（データの迅速な分析・活用）に関して、例えば、テレビ番組で人気タレントが紹介したある製品が、その後すぐにX（旧ツイッター）で話題になったとする。小売企業は、この情報をクラウドを活用したリアルタイム・データ基盤を通じて即座に把握し、その製品をいち早く多く店頭に並べることができれば、販売機会を逃さずに済む（高田, 2022）。

こうして、顧客のニーズや市場のトレンドをめぐるデータをリアルタイムに収集・統合・分析する一連のインテリジェント・システムを開発し、最新のエビデンスに基づく適切なマーケティング戦略（例えば、トレンドに敏感な顧客が求めているホットな製品の提供）を迅速に実行できる企業は、より良い成果を得ることができるであろう（Ghasemaghahi & Calic, 2020）。リアルタイム・データに基づく迅速なマーケティング行動は現在、すべての企業ができるわけではない。そのため、これができる企業は、できない企業よりも、変化する顧客・市場に素早く対応でき、競争上有利になるであろう。以上より、次の仮説が設定される。

仮説1：Velocity（データの迅速な分析・活用）は、企業成果と正の関係にある。

Veracity（正確なデータの分析・活用）に関して、これはビッグデータのコンテキストで重要な要件である。分析ツールが大幅に進歩したにもかかわらず、正確性の低いデータから高い成果を得ることは依然として困難である（Ghasemaghahi & Calic, 2019）。既述したように、ビッグデータはSNSやIoT機器などから生み出され、貴重な情報とともにノイズやエラーを多く含む。そのため、社内IT教育を受けた優秀なデータ人材（データ・エンジニア）がクレンジング作業などを行うことで、データを詳細な分析に使用できる状態にする必要がある。加えて、詳細な分析に使用できるデータが得られたとしても、それを適切に分析することは容易ではない。自社のビジネスにおける顧客の購買行動や特性の理解、因果関係の推定、結果の解釈、自社独自のマーケティング戦略の提案などを適切に行うべく、自社と分析技術の双方のドメイン知識を持った優秀なデータ人材（データ・サイエンティスト）が必要である（高田, 2020）。

こうした優秀なデータ人材（データ・エンジニアとデータ・サイエンティスト）に基づく正確なデータの分析・活用は現在、すべての企業ができるわけではない。そのため、これができる企業は、できない企業よりも、厳密な分析を通じて的確な意思決定を行い、競争優位を確保し、高い成果を得ることができるであろう。以上より、次の仮説が設定される。

仮説2：Veracity（正確なデータの分析・活用）は、企業成果と正の関係にある。

以上のように本研究は、既存研究において肯定的な結果（○）が確認された Velocity と Veracity について仮説を設定する。その一方で、既存研究において否定的な結果（×）が示された Volume と混合的な結果（△）が示された Variety については以下の理由により仮説を設定しない。

Volume（大量のデータの分析・活用）に関して、これはビッグデータの“ビッグ”と直接関係する部分であり、重要な活動であると思われるが、しかし前述の通り、既存研究はこの活動のみでは成果が高まらないことを示している。それは文字通り大量のデータを使用することからデータ分析の精度向上に寄与するであろうが、それ単独では最終的に高い成果につながるほど決定的な役割は果たさないものと考えられる（Ghasemaghahi & Calic, 2019, 2020）。顧客や市場に関する大量のデータは現在、一定の価値を有するものの、一般的に豊富に存在する資源であり、多くの企業にとって利用可能であるため、それのみで競争優位を築くことは難しいであろう。

Variety（多様なデータの分析・活用）に関して、今日の企業は、数値で示される伝統的な構造化データから、映像、画像、非言語、テキスト、音声といったより現代的な非構造化データまで、様々な種類のデータ（Balducci & Marinova, 2018）を分析し、異なるデータ間の新しい関係やパターンを発見し、その結果として顧客にとって価値の高い創造的な製品やサービスを作り出すことができる。しかし、このプロセスは容易ではない。むしろかなり難しく、その達成には“ある状況”が整っている必要があると考える。その状況とは、データの異分野間流通と AI の本格的な活用であろう（他にも、さらなる条件が必要かもしれない）。分野の異なるデータを AI を高度に活用してこれまで誰も検討したことのないような形で複雑に組み合わせることで、企業は高い成果につながる独自の顧客インサイトを引き出すことができると考える。ただし、こうした状況は日本ではまだ整っていないと思われる（IPA, 2021; 総務省, 2020）。よって、日本を分析対象とする本研究では、Variety と成果の間に直接的な正の関係は見られないと予想する。

4. 実証分析

本節では、前節で提示した仮説の経験的妥当性を検討すべく実証分析を行う。本研究は、既存研究（Ghasemaghahi & Calic, 2019, 2020; Johnson et al., 2017）と同様に、様々な産業の企業に対するサーベイ（質問票調査²⁾）を通じてデータを収集する。サンプリング・フレームは、日経 NEEDS に収録されている日本の上場企業である。

調査は2022年3月に実施した。1,100社に対して質問票を郵送し、218社から返送があった。そのうち有効回答は209票であり、有効回答率は19%であった。質問票への回答は、マーケティング戦略におけるデータ分析・活用に精通している担当者に依頼した。回答企業の従業員数の平均

2) このサーベイは、近年、ミズーリ科学技術大学の府川信幸先生と取り組んでいる研究プロジェクトに関連して、それとは別に筆者が行ったものである。日頃よりご指導いただいている府川先生に深く感謝の意を表したい。

値は4,139人、中央値は932人であった。

調査によって得られたサンプルに関して、3つのチェックを行った。第1に、無回答バイアスを検討するために、早期回答者（上位25%のグループ）と後期回答者（下位25%のグループ）を比較した。本研究で考慮するすべての変数について t 検定を行った結果、5%水準で有意な差は見られなかった。第2に、キー・インフォーマント・コンピテンシーをチェックするために、回答者の適切性を2つの質問項目を用いて測定した。10段階のリッカート尺度において平均値（標準偏差）は8.02（1.55）と7.62（1.65）であり、それぞれ高い値であった。第3に、コモン・メソッド・バイアスを検討するために、Harmanの単一因子検定を実施した。潜在的な構成概念のすべての質問項目を対象として探索的因子分析を行った結果、寄与率の大部分を占める単一因子はなかった。以上の結果は、本研究のサンプルに重大な問題がないことを示唆している。

本研究で用いた測定尺度は表2の通りである。焦点の説明変数であるビッグデータの4Vに関する質問項目は、Johnson et al. (2017) と Ghasemaghaei and Calic (2019, 2020) の尺度に基づいている。被説明変数である企業成果は、Mu et al. (2018) に依拠して、市場成果と財務成果を通じて把握された。市場成果は市場シェアの伸びなど「市場ベースの成果目標の達成」、財務成果は営業利益など「財務ベースの成果目標の達成」を捉えるものである。質問項目は、Vorhies and Morgan (2005) の尺度に基づいている。そして、本研究と同様に、成果を被説明変数とする既存研究 (Guo et al., 2018; Mu et al., 2018; Takata, 2016) を踏まえて、企業規模、企業年齢、B2C デミーという3つが統制変数として考慮された。

複数の項目からなる測定尺度に関して、信頼性・妥当性の確認を行った。まず、Cronbachの α 係数とCR (Composite Reliability) を算出した。いずれも基準値の0.70以上であり、尺度の信頼性が確認された。また、それらの補完的指標であるAVE (Average Variance Extracted) を算出した。すべてが基準値の0.50以上であった。続いて、確認的因子分析を行った。その結果、適合度指標は受容可能な値であった ($\chi^2/df=1.87$, RMSEA=0.07, CFI=0.96, TLI=0.96, SRMR=0.05)。すべての項目は、それぞれの潜在的な構成概念との間に1%水準で有意な因子負荷量を示した。以上より、尺度の収束妥当性が確認された。さらに、Fornell and Larcker (1981) にしたがって、ある概念のAVEの平方根とその概念と他の概念間の相関係数を比較した。すべての概念において前者が後者を上回り、尺度の弁別妥当性が確認された。表3はこれらの測定尺度の信頼性・妥当性指標とともに、変数の記述統計量と変数間の相関行列を示す。

本研究は、見かけ上無関係な回帰 (SUR: Seemingly Unrelated Regression) モデルを用いて、仮説をテストする。SURを用いる理由は次の通りである。前述のように、本研究は、被説明変数である企業成果を市場成果と財務成果を通じて把握する。分析では、2つの回帰モデルを推定する。1つは市場成果を被説明変数とする回帰モデル、もう1つは財務成果を被説明変数とする回帰モデルであり、双方で説明変数は同じである。よって、この2つの回帰モデルは関連しており、誤差項に相関がある可能性が高いが、SURを用いることで、双方のモデルを1つの回帰システムとして推定し、誤差項の相関を考慮し、推定精度の改善を図ることができる (Homburg et al., 2015)。Breusch-Pagan 検定の結果は、その2つの回帰モデル間の誤差項が無相関ではないこと

表2 測定尺度

構成概念	質問項目
焦点の説明変数：ビッグデータの4V	
Volume (Johnson et al., 2017; Ghasemaghaei & Calic, 2019, 2020)	1. 貴社では、大量のデータを分析している。 2. 貴社が調査するデータの量は、膨大である。 3. 貴社は、大量のデータを使用している。 4. 貴社では、膨大な量のデータを精査している。
Variety (Johnson et al., 2017; Ghasemaghaei & Calic, 2019, 2020)	1. 貴社では、洞察を深めるために、いくつかの異なるデータ・ソースを使っている。 2. 貴社では、多くの種類のデータを分析している。 3. 貴社は、データの活用を可能にするデータベースを多数所有している。 4. 貴社では、多数のソースを使ってデータを検証している。
Velocity (Johnson et al., 2017; Ghasemaghaei & Calic, 2019, 2020)	1. 貴社は、データを受け取ると直ちに分析する。 2. 貴社は、新たなデータの入手から分析まで短時間でこなしている。 3. 貴社は、データを迅速に探索する。 4. 貴社は、データの分析を迅速に行う。
Veracity (Ghasemaghaei & Calic, 2019)	1. 貴社では、正確で信頼できるデータを扱っている。 2. 貴社は、質の高いデータを分析する。 3. 貴社は、信頼性の高い、整合の取れたデータを処理する。
被説明変数：企業成果	
市場成果 (Vorhies & Morgan, 2005)	1. 競合他社と比較した市場シェアの伸び 2. 売上高拡大 3. 新規顧客の獲得 4. 既存顧客に対する売上の増加
財務成果 (Vorhies & Morgan, 2005)	1. 収益性（営業利益） 2. 投資利益率（ROI） 3. 売上高利益率（ROS） 4. 財務目標の達成
統制変数	
企業規模	従業員数の自然対数
企業年齢	設立年からの経過年数
B2C ダミー	貴社の代表的な製品・サービスが消費者に販売するものであれば1、法人や官公庁などの組織に販売するものであれば0とするダミー変数

注) 焦点の説明変数の各質問項目は「1：全くそう思わない～7：非常にそう思う」という7点尺度、被説明変数の各質問項目は「1：主な競合他社よりもずいぶん悪い～7：主な競合他社よりずっと良い」という7点尺度で評価された。

を強く示し ($\chi^2(\text{d.f.}=1)=97.53, p<0.01$)、SURの採用を正当化した。

表4は分析結果を示す。仮説1は、Velocity（データの迅速な分析・活用）と企業成果の正の関係を予想する。この予想と一致して、結果は、Velocityが市場成果および財務成果の向上に有意に貢献することを示した（それぞれ $\beta=0.15, p<0.05$; $\beta=0.35, p<0.01$ ）。ゆえに、仮説1は支持された。また、仮説2は、Veracity（正確なデータの分析・活用）と企業成果の正の関係を予想する。結果は予想通り、Veracityが市場成果および財務成果を有意に高めることを示した（それぞれ $\beta=0.18, p<0.05$; $\beta=0.23, p<0.05$ ）。ゆえに、仮説2も支持された。

以上のように、本研究で仮説を設定したVelocityとVeracityについて、それぞれ正の効果が確認された。その一方で、本研究で仮説を設定しなかったVolume（大量のデータの分析・活用）と

表3 記述統計量, 相関行列, および信頼性・妥当性指標

	1	2	3	4	5	6	7	8	9
1 Volume	0.88								
2 Variety	0.76	0.82							
3 Velocity	0.55	0.61	0.91						
4 Veracity	0.56	0.63	0.60	0.93					
5 市場成果	0.22	0.24	0.32	0.31	0.76				
6 財務成果	0.18	0.20	0.37	0.31	0.72	0.93			
7 企業規模	0.36	0.25	0.09	0.10	0.08	0.04	—		
8 企業年齢	0.00	-0.04	-0.19	-0.06	-0.13	-0.11	0.28	—	
9 B2C ダミー	0.23	0.22	0.24	0.30	0.08	0.06	0.10	-0.16	—
平均値	4.19	4.12	3.64	4.54	4.59	4.32	6.87	66.62	0.48
標準偏差	1.40	1.22	1.35	1.23	1.05	1.42	1.56	29.66	0.50
α	0.93	0.89	0.95	0.95	0.84	0.96	—	—	—
CR	0.93	0.89	0.95	0.95	0.85	0.96	—	—	—
AVE	0.78	0.67	0.83	0.86	0.58	0.87	—	—	—

注) 相関行列の対角線上の値は AVE の平方根。

表4 分析結果

	市場成果	財務成果
Volume	-0.01 (0.08)	-0.05 (0.10)
Variety	-0.02 (0.09)	-0.11 (0.12)
Velocity	仮説 1 0.15 (0.07) *	0.35 (0.09) **
Veracity	仮説 2 0.18 (0.08) *	0.23 (0.10) *
企業規模	0.06 (0.05)	0.05 (0.07)
企業年齢	-0.00 (0.00)	-0.00 (0.00)
B2C ダミー	-0.11 (0.14)	-0.17 (0.19)
定数項	3.26 (0.40) **	2.58 (0.53) **
R ²	0.14	0.16
Breusch-Pagan 検定		
H ₀ : 誤差項が無相関	$\chi^2(1) = 97.53^{**}$ ($p = 0.00$)	

注) **: 1%水準で有意, *: 5%水準で有意。
各説明変数の値は回帰係数 (標準誤差)。

Variety (多様なデータの分析・活用) は双方とも, 市場成果および財務成果と有意な関係を示さなかった。3つの統制変数も, 2つの成果と有意な関係を示さなかった。

5. おわりに

以上, 本研究は, 日本のデータを用いて, ビッグデータの4Vが企業成果に及ぼす影響を検討した。サンプルや各変数, 分析結果など, 本研究の重要な点を表1の最後の列に示した。そして, この表1の分析結果の行を見れば, ビッグデータの4Vと成果の関係について, 既存研究と本研究を踏まえて押さえておくべき点と今後検討すべき課題を確認することができる。

押さえておくべき点は、次の2つである。第1に、Velocity（データの迅速な分析・活用）と Veracity（正確なデータの分析・活用）が成果の向上に資することである。第2に、Volume（大量のデータの分析・活用）のみでは成果が上がらないことである。この2点については、既存研究と本研究で一貫した結果が得られている。以上を踏まえると、企業は強力なりアルタイム・データ基盤の構築と有能なデータ人材の育成に投資を行い、データ分析・活用の迅速さ・正確さを高める一方、大量のデータの収集のみに焦点を当てることは避けるべきである。

今後検討すべき課題は、Variety（多様なデータの分析・活用）と成果の関係である。本研究ではその間に直接的な正の関係が見られる状況が日本ではまだ整っていないと考え、仮説を設定せず、分析結果はその考えを支持するものであった。これは現在の日本において分野の異なるデータを複雑に組み合わせて独自の顧客インサイトを引き出すことの難しさを示唆しており、注目すべきものである。また、Varietyと成果・能力の関係について、本研究を含めこれまで5つの研究が分析している。そのうち有意な正の関係を示すものが3つ、非有意な関係を示すものが2つであり、結果は混合的である。この原因は前述のように、米国と日本という仮説をテストした状況（国）の違いにあるかもしれない。関連して、Varietyの効果を顕在化させる重要な条件があり、それが考慮されれば正の効果が（より）明確に確認されるかもしれない。あるいは、Varietyの測定尺度に問題があるかもしれない。例えば、質問項目を「貴社では、新たな洞察を得るために、非構造化データに着目し、それを構造化データと組み合わせて分析している」といったより具体的なものにすれば、結果が異なるかもしれない。これらの詳細な検討が必要である。

以上のように、今後検討すべき課題は残るものの、企業がビッグデータを活用する際の前提の1つとして、既存研究と本研究を通じて明らかになった「4Vの単独の効果の違い」を理解しておくことが重要である。その違いとは、「Volumeは×（正で有意ではない。これのみのレベルアップでは成果が高まらない）、Varietyは？（さらなる検討が必要。これを高い成果につなげることはかなり難しい?）、VelocityとVeracityは○（正で有意。双方ともレベルアップを図れば成果が高まる）」ということである。

付記

高橋郁夫先生にはじめてお目にかかったのは、私の大学2年次の「商業学」においてであった。毎回身近なマーケティング・流通現象をめぐる新しい発見があり、毎週楽しみにしていた授業であり、この授業をきっかけに大学3・4年次には商業学を専門的に学びたいと思うようになった。思い返せばこのときから現在まで20年以上、先生には何らかの授業を通じて常にご指導いただいている。先生に教えていただいたことは、研究・教育についてはもちろん、人生において大切なこと、きめ細やかなコミュニケーション、グローバルな視点や中長期的な目標・目的意識を持つことの大切さなど多岐にわたる。また、自分一人では判断が難しい問題に直面しご相談した際、先生はいつも親身になって進むべき正しい方向性を示してくださった。まさに“父”のような高橋先生は、私にとって唯一無二の特別な存在である。これまでいただいた学恩に心より感謝申し上げますとともに、先生のさらなるご活躍とご健康をお祈りしたい。そして、先生の思いを受け継ぎ、慶應義塾における研究者・教育者としての務めを精一杯果たしていきたい。

参 考 文 献

- Balducci, B., & Marinova, D. (2018). Unstructured data in marketing. *Journal of the Academy of Marketing Science*, 46(4), 557-590.
- Cappa, F., Oriani, R., Peruffo, E., & McCarthy, I. (2021). Big data for creating and capturing value in the digitalized environment: Unpacking the effects of volume, variety, and veracity on firm performance. *Journal of Product Innovation Management*, 38(1), 49-67.
- Ciampi, F., Demi, S., Magrini, A., Marzi, G., & Papa, A. (2021). Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research*, 123, 1-13.
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897-904.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Ghasemaghaei, M., & Calic, G. (2019). Does big data enhance firm innovation competency? The mediating role of data-driven insights. *Journal of Business Research*, 104, 69-84.
- Ghasemaghaei, M., & Calic, G. (2020). Assessing the impact of big data on firm innovation performance: Big data is not always better data. *Journal of Business Research*, 108, 147-162.
- Guo, H., Xu, H., Tang, C., Liu-Thompkins, Y., Guo, Z., & Dong, B. (2018). Comparing the impact of different marketing capabilities: Empirical evidence from B2B firms in China. *Journal of Business Research*, 93, 79-89.
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, 53(8), 1049-1064.
- 原隆浩 (2017). 『ビッグデータ解析の現在と未来：Hadoop, NoSQL, 深層学習からオープンデータまで』, 共立出版.
- Homburg, C., Vomberg, A., Enke, M., & Grimm, P. H. (2015). The loss of the marketing department's influence: Is it really happening? And why worry? *Journal of the Academy of Marketing Science*, 43(1), 1-13.
- IPA (独立行政法人情報処理推進機構) (2021). 『DX 白書2021』.
- Johnson, J. S., Friend, S. B., & Lee, H. S. (2017). Big data facilitation, utilization, and monetization: Exploring the 3Vs in a new product development process. *Journal of Product Innovation Management*, 34(5), 640-658.
- McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, 90(10), 60-68. (有賀裕子訳 (2013). 測定できれば, マネジメントできる ビッグデータで経営はどう変わるか. 『Diamond ハーバード・ビジネス・レビュー』, 38(2), 42-53.)
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment. *British Journal of Management*, 30(2), 272-298.
- Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, 57(2), 103169.
- Mu, J., Bao, Y., Sekhon, T., Qi, J., & Love, E. (2018). Outside-in marketing capability and firm performance. *Industrial Marketing Management*, 75, 37-54.
- 総務省 (2020). 『令和2年版 情報通信白書』.
- Takata, H. (2016). Effects of industry forces, market orientation, and marketing capabilities on business performance: An empirical analysis of Japanese manufacturers from 2009 to 2011. *Journal of Business Research*, 69(12), 5611-5619.
- 高田英亮 (2020). マーケティング戦略におけるビッグデータの活用. 『三田商学研究』, 63(4), 137-149.
- 高田英亮 (2021). マーケティング戦略と流通チャネル研究の要諦② ビッグデータを上手に使うには. 『日経ビジネス』, 6月28日号, 74-77.
- 高田英亮 (2022). ビッグデータ活用の4V分析. 『日経広告研究所報』, 325, 80-81.
- Vorhies, D. W., & Morgan, N. A. (2005). Benchmarking marketing capabilities for sustainable competitive advantage. *Journal of Marketing*, 69(1), 80-94.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365.

参考記事

Asay, M. (2017). 85% of big data projects fail, but your developers can help yours succeed. TechRepublic. <https://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed/>

ガートナー・ジャパン (2019). ガートナー、企業におけるデータ活用に関する調査結果を発表。プレスリリース。 <https://www.gartner.co.jp/ja/newsroom/press-releases/pr-20190527>

Heudecker, N., & Hare, J. (2016). Survey analysis: Big data investments begin tapering in 2016. Gartner Research. <https://www.gartner.com/en/documents/3446724>