



RIETI Discussion Paper Series 25-J-015

潜在顧客情報の価値：営業履歴データを用いた実証分析

宮川 大介
早稲田大学

柳岡 優希
株式会社東京商工リサーチ

矢澤 広崇
三井住友ファイナンス&リース株式会社

雪本 真治
三井住友ファイナンス&リース株式会社



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

潜在顧客情報の価値：営業履歴データを用いた実証分析*

宮川大介（早稲田大学）

柳岡優希（株式会社東京商工リサーチ）

矢澤広崇（三井住友ファイナンス&リース株式会社）

雪本真治（三井住友ファイナンス&リース株式会社）

要 旨

本研究では、法人向けリース契約を対象とした営業担当者（フィールドセールスパerson）の営業活動において、成約見込みの高い潜在顧客をテレマーケティング担当者（インサイドセールスパerson）が探索した情報の利用が成約確率の向上に寄与するか否かを、その因果関係に注目した上で実証的に検討したものである。国内リース業最大手の SMFL が保有するリース営業活動実績データ、潜在顧客リスト、営業活動の成否実績データを用いて、Kleinberg et al. (QJE 2018)で提案された contraction technique などに基づく分析を行った結果、フィールドセールスパersonが独力で接触する傾向が弱いクラスの潜在的顧客について、インサイドセールスパersonによる追加的な情報提供が成約確率を改善させることを確認した。一方で、前者が独力で接触する傾向の強いクラスの潜在的顧客については、後者による情報提供を通じた成約確率の改善は認められなかった。以上の結果は、異なる方法に基づく情報生産の価値が条件に応じて変化することを意味しており、適切なデータの活用が企業パフォーマンスの改善に繋がることを示唆するものである。

キーワード：情報生産、テレマーケティング、機械学習、選択的ラベル問題

JEL classification: D22, D83, C53

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

*本稿は、独立行政法人経済産業研究所（RIETI）におけるプロジェクト「企業金融・企業行動ダイナミクス研究会」の成果の一部である。本稿の原案は、経済産業研究所（RIETI）のディスカッション・ペーパー検討会で発表を行ったものである。検討会参加者からの有益なコメントに感謝したい。また、本稿の作成にあたっては、深尾京司 RIETI 理事長、富浦英一 RIETI 所長・CRO、荒田禎之研究員から多くの有益なコメントを頂いた。ここに記して、感謝の意を表したい。

1. イントロダクション

生産や販売からなる企業活動の運営に当たっては、様々な情報が利用される。例えば、財やサービスの供給に際して重要となる生産要素について、一定期間に亘ってどのような価格と数量の組み合わせでどのような仕入先から入手可能であるかを把握することは、最適な生産規模を選択する上で欠かせない。同様に、どのような価格と数量の組み合わせの下でどのような販売先から財やサービスに関する引き合いがあるかを見通すことも期待される。実際、こうした情報が準備されていれば、営業担当者の労働資源をより効率的に投入することで企業パフォーマンスを改善することもできるだろう。本研究では、こうした問題意識を踏まえて、販売先の探索における情報の価値に焦点を当てる。特に、当該情報の利用が新規の販売先探索に貢献するか否かを実証的に検討する。

情報を企業活動のインプットとして捉えた研究の代表例としては、情報の蓄積を無形資産の一種とみなした上で計測を試みた一連の研究が挙げられる(Corrado et al. 2005; 2009)。特に、Corrado and Hao (2014)では広告宣伝や社内外のリソースを用いた販売促進活動から構成されるマーケティングを投資の一種と捉えて、米国におけるマーケティング資本を計測しているほか、Sveikauskas et al. (2024)では、米国における集計レベルのマーケティング資本計測を行うことで研究開発資本やソフトウェア資本といった、伝統的に重要な無形資産として認識されていた類型と同水準の資本蓄積が進んでいることを報告している。

こうしたマクロ的観点からの計測を主とする実証研究に対応するミクロ的視点からの実証研究として、例えば、広告宣伝活動がもたらす企業パフォーマンスへの因果効果を推定した実証研究が典型例と言えるが、広告宣伝活動は、上記の無形資産計測が対象としている情報資産のごく一部に過ぎない。例えば、顧客基盤(customer base)(例: Rob and Fishman 2005; Gourio and Rudanko 2014)は企業活動を営む上での重要なインプットの一つであり、特定の情報資本の利用の効果をミクロ的視点から実証的に検討することの価値は依然として高い。

本研究は、こうした問題意識を踏まえて、特に、顧客基盤の獲得や拡大に向けた企業内での活動がどのような効果を生んでいるかを実証的に検討するものであるが、この分析に当たっては、情報の価値を計測する目的で生じる典型的な二つの困難を乗り越える必要がある。第一に、分析に適したデータを得ることの難しさである。本研究が対象とする、顧客獲得に向けた情報収集とその利用がもたらす効果を例にとると、どのようにして収集された情報が、どのタイミングで営業担当者に提供され、その帰結として営業担当者の行動やその後の成果へどのような影響を与えたのかを時系列に沿って収集したデータが一定数確保されなければ実証研究を行うことは難しい。こうしたデータが学術研究目的で提供されることは極めて稀であるなか、本研究では日本国内におけるリース業の最大手である三井住友ファイナンス&リース株式会社(以下、「SMFL」という)との共同研究体制のもとで、かかる詳細なデータへのアクセスがかなったことで、分析が可能となっている。

第二に、情報の提供から成果の計測までの間に、いわゆる選択的ラベル問題を引き起こす

いくつかの要因が存在している。ここで選択的ラベル問題とは、特定の条件を満たすレコードにおいてのみ結果が観測されるという状況を指す。本研究の文脈における例としては、営業担当者がコンタクトしていない先において営業活動の成否が観察されないという事象が挙げられる。この結果、営業担当者に対して成約確率の高い潜在的な顧客の情報提供があったとしても、営業担当者がこの情報が差し示す行動を必ずしも選択しない場合、結果として、情報提供という介入に対応する結果を観測できないこととなる。こうした問題に対して、本研究では、機械学習手法を含む近年の実証手法の発展を踏まえた工夫によって、潜在的顧客に関する情報の利用が販売先探索に貢献するか否かを実証的に検討する。

具体的には、法人向けリース契約を対象として営業担当者（フィールドセールスパーソン）が営業活動を行っているという環境を対象として、こうした担当者へ、テレマーケティング担当者（インサイドセールスパーソン）が探索した成約見込みの高い潜在顧客情報を提供しているという状況において、かかる情報の提供が成約確率の向上に寄与するか否かを、その因果関係に注目した上で実証的に検討する。既述の問題点を克服するために、案件の探索、営業、成約及び失注に至る案件の詳細情報を利用し、かつ、フィールドセールスパーソンとインサイドセールスパーソンの行動を模倣するモデルを構築した上で、フィールドセールスパーソンの意思決定をインサイドセールスパーソンの意思決定で置き換えた場合に、営業成果に関するどのような変化が生じるかを検討した。

得られた結果から、フィールドセールスパーソンが独力で接触する傾向の弱いクラスの潜在的顧客について、インサイドセールスパーソンによる追加的な情報提供が成約確率を改善させることを確認した。一方で、前者が独力で接触する傾向の強いクラスの潜在的顧客については、後者による当該情報提供による成約確率の改善は認められなかった。以上の結果は、異なる方法に基づく情報生産（潜在的顧客のリスト構築）の価値が条件に応じて変化することを示唆するものである。

以上の実証結果は、より成約確率の高い潜在的顧客を探索するためのマーケティング活動という単一の目標に照らして、フィールドセールスパーソンが慣れ親しんでいるクラスの企業については、インサイドセールスパーソンの情報生産活動が価値を持たないことを意味している。一方で、立地などの事情から必ずしもフィールドセールスパーソンが十分な営業案件のリード情報を持っていないクラスの企業については、インサイドセールスパーソンがテレマーケティング活動を通じて収集した見込み情報が大きな価値を持っている。これらの結果は、財やサービスの販売を目的とした企業活動の実施に際して、異なるテクノロジーに基づいた多様な活動が代替的に用いられることでより高い成果を挙げうることを意味している。

これ以降の構成は以下の通りである。第二節では、データの概要を示す。第三節は分析手法を、第四節は結果とその含意を議論した上で、第五節で本研究の結論をまとめる。

2. データと背景

2.1. 背景

SMFL は、リース取引規模が日本国内最大級の総合リース会社であり、企業向けリースのほか、ファイナンス（延払、金銭消費貸借、立替払）や生命保険、資産管理クラウドサービス等の幅広いサービスを提供している。

SMFL が行う営業活動には、2つのアプローチがある。第一のアプローチは、SMFL の営業部店に所属する営業担当者（フィールドセールスパーソン）が、潜在的な顧客に対して直接往訪するなどの物理的なコンタクトをとり、商談を進める方法である。これに対して、第二のアプローチは、テレマーケティングを活用した営業アプローチである。具体的には、まず、テレマーケティングを専門に行うマーケティング担当者（インサイドセールスパーソン）が潜在的な顧客に架電し、自社サービスに対するニーズをヒアリングする。次に、インサイドセールスパーソンは架電によって取得した投資計画などのニーズ情報をフィールドセールスパーソンに提供する。トスアップと呼ばれるこうした情報提供を受けたフィールドセールスパーソンは、得られた情報を基に見込顧客にコンタクトを取るか否かの判断を行いつつ、具体的な商談を進めることとなる。

両アプローチは、ともに潜在的顧客にフィールドセールスパーソンがコンタクトを取った上で商談を進めるという点で共通している。実際、案件が成約に至るまでにフィールドセールスパーソンのコンタクトが必ず存在するというビジネスモデルが前提とされており、インサイドセールスパーソン単独で成約することはない¹。つまり、インサイドセールスパーソンの役割は、見込顧客の初期のニーズを発掘・精査（リードジェネレーション）し、フィールドセールスパーソンにそれらトスアップ情報を提供することを通じて、成約見込みの高い顧客を選別するというスクリーニングの役割を担っていると整理できる。勿論、インサイドセールスパーソンが架電などを通じて複数回接触することで、潜在的な顧客との関係を構築する役割を果たしている可能性もあるほか、広義のスクリーニング活動として、これまで実績のないエリアや業種の企業に対して営業開拓を試みるという面もあるが、本研究が対象とする SMFL におけるインサイドセールスパーソンの業務実態を踏まえて、以下の分析ではインサイドセールスパーソンが案件のスクリーニング業務を行っていると考ええる。また、第二のアプローチでは、インサイドセールスパーソンからのトスアップ情報に基づいて、フィールドセールスパーソンは当該潜在的顧客にコンタクトするか否かを判断することとなる。一般に、テレマーケティングは物理的な移動が不要であり、フィールドセールスパーソンが現地に往訪して行う物理的なコンタクトに比べて時間的・金銭的成本が低いため、営業プロセスの効率化を図りながら、見込顧客にアプローチすることが可能となると考えられている。なお、本研究では、フィールドセールスパーソンとインサイドセール

¹ フィールドセールスパーソンのコンタクトがない状況で、顧客からの引き合いのみで案件が成約に至るという可能性もあるが、本研究ではこのケースを除外している。

スパークソンの両方が、コンタクトした潜在的顧客に関する高い成約確率を重視して行動しているとの想定を置く。潜在的顧客にコンタクトすることに伴うコストを踏まえれば、こうした目的関数の設定には相応の合理性があると考えられるが、例えば、インサイドセールスパークソンが、より多くの案件情報をその成約確率を度外視してフィールドセールスパークソンへ提供する可能性もある。これは、提供する案件数がインサイドセールスパークソンの評価対象となっている場合において生じうる事象であるが、現実には、提供した情報の質も勘案してインサイドセールスパークソンの業務評価が行われることが一般的であることから、上記の仮定の下で分析を進めることとする。

以下では、こうしたテレマーケティング活動が具体的にどのようなフローで行われているのかを概観する。最初のステップとして、テレマーケティング活動を実施する部署が、年当たり数十万件程度の見込顧客リストを作成する。これらの潜在的顧客は、SMFL がこれまでの取引関係等の接触があった顧客のほか、外部情報を通じて取得した情報に基づいている。次に、第二ステップとして、こうした潜在的顧客の立地や業種などを踏まえて担当営業部店（フィールドセールスパークソン）に割り振られ、同営業部店がテレマーケティング活動の実施可否を判断する。このステップは、すでに営業活動を開始しているなどの事情から、テレマーケティング活動の経済的意義が低い先を除くために設けられている。最後に、第三ステップとして、インサイドセールスパークソンは、担当営業部店がテレマーケティング活動の実施が可能であると判断した潜在的顧客リストのなかから、成約見込みが高いと想定する顧客を選択し、架電を行う。この架電時に、設備投資情報など重要な情報を入手した場合、当該顧客の担当営業部店にそれら情報を提供する（トスアップ）。

既述の通り、このトスアップを受けた担当営業部店（フィールドセールスパークソン）は、リストに含まれている潜在的顧客に対して実際にコンタクトを試みるか否かを判断する。この意味で、トスアップがフィールドセールスパークソンの行動を完全に規定しているわけではない点に注意が必要である。この結果として、仮に成約確率が相応に高い潜在的顧客がトスアップされたとしても、フィールドセールスパークソンの判断によってこの情報が有効に活用されない可能性がある。また、こうしたフィールドセールスパークソンの選択によって、トスアップ先の中で最終的なアウトカム（契約の成否）が観測されないケースが生じることになる点も重要である。本研究では、営業活動を促進する目的で行われた情報提供への遵守（compliance）が完全ではないという点と、情報提供の結果が観測されないという点（選択的ラベル問題）を踏まえた実証分析が必要となる。

2.2. データ

本研究で使用するデータは、SMFL の社内データおよび株式会社東京商工リサーチ（以下、「TSR」という）が提供する企業属性情報データである。SMFL の社内データは、テレマーケティング活動履歴、営業活動履歴、成約履歴の3種類である。テレマーケティング活動履

歴および営業活動履歴データは活動単位、成約履歴は契約単位の粒度（レコード）で記録されている。分析を行うにあたって、潜在的顧客単位にデータを集約することで、テレマーケティング活動および営業活動を経て成約に至るプロセスを示すパイプラインデータセットを構築する。つまり、顧客ごとにテレマーケティング活動（コールやトスアップ）や営業活動（コンタクト）を行ったのか否か、そして成約に至ったのか否かを示すデータセットを構築する。そのうえで、各取引先企業に TSR の企業属性情報を紐づける。

以下では、個々のデータに関して説明し、データセットの構築方法を概説する。

2.2.1. SMFL テレマーケティング活動履歴データ

テレマーケティングの対象となる潜在的顧客リストは、約 30 万社の取引先情報（取引先名、連絡先等）によって構成されており、担当営業部店が付与したテレマーケティング実施可否の情報が含まれている。テレマーケティング活動履歴データは、これら潜在的顧客への架電情報として、架電先、架電日、架電者、架電内容、架電結果（トスアップの有無等）等が架電単位のレコードとして、CRM 情報システムに登録されたものである。本研究では、これらテレマーケティング活動履歴情報のうち、2022 年 5 月に作成された潜在的顧客リストのなかで、地方を商圈とする担当営業部店がアウトバウンドセールスでのテレマーケティングを行うことを承諾した潜在的顧客を分析対象とする（5.2 万社）。2022 年 5 月に作成された潜在的顧客リストは、2022 年 6 月から 2023 年 11 月まで使用されたことから、当該期間の架電情報を利用する。

2.2.2. SMFL 営業履歴・成約履歴データ

営業履歴データは、営業担当者の往訪記録といった営業日報データであり、取引先、実施日、担当者、コンタクト方法（往訪、電話・メール等）、コンタクト内容等が記録されている。ただし、すべての往訪が記録されているとは限らない。重要な情報を取得した場合のみ入力されている可能性もあり、コンタクトの有無には何らかのセレクション・バイアスが存在している可能性がある。

また、成約履歴データは、SMFL のリース・延払・金銭消費貸借・立替払のすべての契約が記録され、契約先、契約日、成約担当者、契約種類、成約金額等の情報を保持している。ここで、成約履歴データはリース・延払・金銭消費貸借・立替払のすべての契約情報を蓄積する一方で、テレマーケティングや営業活動では、他のサービス（生命保険や資産管理クラウドサービス等）も扱っている。結果として、本研究で用いるデータでは一部サービスの成約を捕捉することができておらず、インサイドセールスパーソンおよびフィールドセールスパーソンの選択（例：訪問、架電）がこれらサービスでの成約を念頭に置いたものである場合には、後述する各担当者の行動を模倣するモデルを構築する際に、いわゆる **omitted payoff bias** が生じる結果として、観察される各担当者の行動を正確にモデル化できない可能

性がある点に注意が必要である。

本研究では、テレマーケティング開始時期に合わせて、2022 年 6 月から 2024 年 2 月までに実施された営業履歴データ、成約履歴データを利用する。

2.2.3. TSR 企業属性情報データ

TSR が提供する企業属性データには、資本金や従業員数などの企業属性情報が収録されている。また、TSR 独自の手法で調査・作成した企業の信用力等を表す指標（TSR スコア）も含まれている。本研究では、これらの企業属性情報から、資本金、従業員数、設立年、業種、本社所在地、取引先銀行、当期売上高、当期税引後利益、TSR スコアを各モデルの特徴量として使用する。数値変数は各モデル構築データセットの学習データにおける五分位でダミー変数化するほか、欠損値が発生した場合は欠損値カテゴリとして別カテゴリ値とする。なお、利用するデータ時点は、テレマーケティング活動の開始時期を考慮し、2022 年 5 月時点で入手可能な情報を使用する。

2.2.4. 変数定義とデータの分類

分析を行うにあたり、テレマーケティング活動履歴データ、営業履歴データ、成約履歴データを顧客単位に集約することで、テレマーケティング活動および営業活動を経て成約に至るプロセスを示すパイプラインデータセットを構築する。SMFL の社内データは、案件単位で管理されておらず、テレマーケティング活動ないし営業活動がどの契約に至ったかを直接把握することができない。そこで、コールステータスに応じてコンタクトまたは成約の観測期間を設定することで、テレマーケティング活動の時系列の因果を捕捉する。

まず初めに、分析対象期間において、個々の潜在的顧客に対して、テレマーケティング活動であるコールまたはトスアップを行ったかを確認する。具体的には、インサイドセールスパーソンが、架電したか（コール）、そしてフィールドセールスパーソンに情報提供をしたか（トスアップ）を、テレマーケティング活動履歴データより捕捉する。この際、各潜在的顧客は、コールステータスとして、①コール無し、②コール有りかつトスアップ無し、③コール有りかつトスアップ有りの 3 群に分けられる。次に、各群に観測期間を設定し、コンタクトおよび成約の有無を判定する。具体的には、①コール無し群では、分析対象期間において担当営業部店が当該潜在的顧客に対してコンタクトおよび成約したかを、営業履歴データおよび成約履歴データより捕捉する。同様に、②コール有りかつトスアップ無し群では担当営業部店がコール後に当該潜在的顧客に対してコンタクトおよび成約したかを、③コール有りかつトスアップ有り群では担当営業部店がトスアップ後に当該潜在的顧客に対してコンタクトおよび成約したかを捕捉する。こうしたデータ処理には、インサイドセールスパーソンの活動によって得られた情報を踏まえて、フィールドセールスパーソンが営業活動を行っているという時系列のフローを整える意図がある。

以上の手続きを通じて、コール、トスアップ、コンタクトの有無に応じて、顧客は6つのグループに分類することができる（表1）。表1は、各グループの成約率を示している。グループ0,2,4はフィールドセールスがコンタクトしていない顧客群から成るグループであるため、成約率は0%となっている。また、コール・トスアップ無しかつコンタクト有りのグループ1とコール・トスアップ有しかつコンタクト有りのグループ5の成約率を比較すると、グループ1の方がグループ5より高いことが分かる。

最後に、各顧客にTSRの企業属性情報を紐づける。表2は、企業属性情報の基本統計量を示している。

3. 分析手法

本研究における実証分析の目的は、インサイドセールスパークソンによるトスアップ情報の提供によって、フィールドセールスパークソンの営業活動成果（成約率）が改善するか否かを明らかにすることにある。この目的に照らして、重要な実証上の課題が二点存在する。第一に、成約の有無は、フィールドセールスパークソンが実際にコンタクトした顧客に対してのみ観測可能である。コンタクトされない顧客については、成約していないという事実は確認できるものの、もしコンタクトしていたら成約していたかは観測ができない（選択的ラベル問題）。また、一般にフィールドセールスパークソンが潜在的顧客のもとに往訪するなどのコンタクトを取るコストは、インサイドセールスパークソンが架電するコストよりも高いため、インサイドセールスパークソンからのトスアップがないにも関わらず、フィールドセールスパークソンが直接コンタクトを取るような顧客は、そもそも成約見込みが高い顧客の可能性がある。

第二に、コンタクトするか否かはフィールドセールスの選択であり、インサイドセールスがトスアップした潜在的顧客に対して、必ずしもコンタクトするとは限らないという点である。つまり、インサイドセールスがトスアップしたすべての潜在的顧客に対して成約の有無を観測することができない。

以上の実証上の問題点を乗り越えるため、本研究では各主体の選択行動を学習した機械学習モデルを構築し、構築した各主体の選択行動を模倣するモデルを対象としてパフォーマンスを評価することで、トスアップ情報によってフィールドセールスの営業活動の成果が改善するかを明らかにする。具体的には、インサイドセールスパークソンのコール・トスアップ行動と、フィールドセールスパークソンのコンタクト行動をデータについて機械学習手法を用いた予測モデルを構築する。そのうえで、個々のモデルの予測結果に対して、後述の「序列比較」および「収縮法比較」を行うことで各主体の行動を評価する。

3.1. モデル構築

本研究では、インサイドセールスパーソンおよびフィールドセールスパーソンの各主体の選択行動をスコア（確率）で表すために、各主体の行動を予測変数とした機械学習モデルを構築する。具体的には、インサイドセールスパーソンのコールおよびトスアップ、フィールドセールスパーソンのコンタクトに関する 2 値選択をモデル化する。

各モデルの構築にあたり、インサイドセールスパーソンおよびフィールドセールスパーソンの選択行動を比較するため、他の主体の行動の影響を受けていないデータセットを利用する。具体的には、コンタクト選択モデルはフィールドセールスパーソンのコンタクト有無を予測変数としている。ここで、利用するデータセットはグループ 0,1 のみであり、インサイドセールスパーソンがコールしているグループ 2-5 はコール内容をフィールドセールスパーソンが把握可能なため、モデルの学習には利用しない。他方、トスアップ選択モデルは、インサイドセールスパーソンのトスアップ有無を予測変数としており、グループ 0,1,4,5 を利用する。また、コール選択モデルは、インサイドセールスパーソンのコール有無を予測変数としており、全グループのデータを利用する。

データは各グループで 70%を学習セット、30%を評価セットとしてランダムに分割する。学習セットを用いてモデルを学習し、評価セットでモデルの予測性能を検証する。前処理として、カテゴリ特徴量における欠損値は欠損カテゴリとし、数値特徴量については学習セットの五分位に基づき **one-hot encoding** を行っている。

機械学習モデルとしては、ランダムフォレストモデルを用いる。ハイパーパラメータの調整は、学習セットを用いたクロスバリデーションのもとでベイズ最適化を用いて行う。

3.2. 評価

3.2.1. 序列比較

選択的ラベル問題により、成約の有無はフィールドセールスパーソンがコンタクトした顧客しか観測することができない。そこで、序列比較では、成約の有無を観測することができた顧客群であるグループ 1,5 それぞれについて、機械学習モデル（トスアップ・コール選択モデル、コンタクト選択モデル）を用いて各主体の行動を評価する。

具体的には、トスアップ選択モデル、コール選択モデル、コンタクト選択モデルそれぞれを用いて各群の顧客に予測確率を付与し、モデル毎にそれぞれ行動をとる確率が高い順にコンタクトを取った場合における成約率を比較する。この時、各モデルの行動確率が高い顧客は成約見込みが高いと各主体が判断したことを表しており、横軸に予測確率の上位パーセンタイル点、横軸に成約率を取ると、曲線は右下がりとなることが想定される。また、曲線が上に位置するほど当該主体の選択行動が有効であることを意味する。

3.2.2. 収縮法比較

前項の序列比較は、アウトカムである成約が観測可能なグループに限定したものであり、成約の有無が観測されていないグループを考慮した上で評価を行うことが出来ない点に注意が必要である。例えば、インサイドセールスパーソンがトスアップした顧客に関して、コンタクトするか否かの判断はフィールドセールスパーソンが行っているため、トスアップされた顧客に対してコンタクトしないのはフィールドセールスパーソンが当該顧客の成約見込みが低いと判断したためかもしれない。そうした顧客の成約率が実際に低く、適切な意思決定が行われている可能性がある反面、フィールドセールスパーソンの判断を上書きし、そうした顧客にコンタクトをすることで、結果として成約率が改善する可能性もある。このため、成約の有無が観測されていないグループを何らかの形で考慮に入れた上で情報提供の意義を検討することが重要となる。

そこで、本研究では営業部店毎のコンタクト率の差異に着目し、Kleinberg et al. (2018)、Lakkaraju et al. (2017)によって提案された収縮法 (Contraction Technique) を適用する。収縮法は、大きく 6 つのステップに分けられる。まず、第 1 のステップとして「セル化」を行う。これは、グループ 4,5 のトスアップされた顧客群において、トスアップのタイミング (トスアップ日および曜日)、当該取引先へのアクセスの容易さ (企業が所在する都道府県に SMFL の事業所があるか) に基づいて顧客を分類し、データを分割するプロセスである。こうした分割の結果、12 個のセルが構成され、分割された各セル内における顧客は均質的となることが期待される。第 2 のステップでは、各セル内において、顧客に紐づく担当営業部店毎にコンタクト率を算出し、このコンタクト率を基準に営業部店を四分位 (Q1~Q4) に分類する。第 3 のステップでは、各セルにおいて四分位に分類された営業部店の顧客を、その四分位 (Q1~Q4) ごとに集計し、各グループでのコンタクト率および成約率を算出する。ここで、成約率は「非コンタクト先を含む全顧客数のうち、成約した顧客の割合」として定義する。この定義を採用したことによって、各四分位グループでの顧客の特徴が均質であれば、コンタクト率が上昇すると成約率が機械的に上昇することとなる。第 4 のステップでは、コンタクト率が高かった営業部店の顧客群である Q4 グループのコンタクト先に対してトスアップ選択モデル、コール選択モデル、コンタクト選択モデルの各予測確率を付与する。第 5 のステップでは、この Q4 グループのコンタクト率を、Q1~Q3 グループのコンタクト率と一致するようにコンタクト数を調整する。この際、各モデルの予測確率が低い顧客から順に仮想的な非コンタクト先として扱い、これら仮想的な非コンタクト先は実際の成約の有無にかかわらず、成約率の計算において非成約先として取り扱う。最後のステップでは、Q4 グループ内の調整後のコンタクト率に応じた成約率を算出し、Q1~Q3 の成約率と比較する。この方法により、コンタクト率を統一した条件下で、Q1~Q3 グループそれぞれで実績とモデルの成約率を公平に比較することが可能となる。結果として、モデルの選択がより有効である場合、成約率が高くなるため、曲線が上に位置するほど当該主体の選択行動が有効であることを意味する。なお、この分析手法を用いることの追加的なメリットとして、実

際に行われた営業活動との比較を行うことができるという点もある。

4. 分析結果

本節では、第2節で説明したデータを用いて、第3節の分析を行った分析結果を述べる。また、各主体の行動を予測する機械学習モデルの構築結果を示す。次に、構築した機械学習モデルを用いて、各主体の行動を序列比較および収縮法比較で評価した結果を示す。

4.1. モデル評価

表3は、各主体の行動を予測変数として構築した機械学習モデルの予測精度を示している。予測精度の指標であるAUC (Area Under Curve)²は、トスアップ選択モデル、コール選択モデル、コンタクト選択モデルのいずれにおいても、十分に高い水準（高い予測精度）を示している。

実際、各モデルの予測確率を評価セットの顧客群に対して付与し、予測確率を百分位ごとに分割し、予測変数の発生率を描画した図1を見ると、明確な右上がりの傾向があり、各主体の行動を相当程度模倣できていることが分かる。

4.2. 序列比較

成約の有無を観測したグループである、フィールドセールスパーソンがコンタクトしたグループ1,5の顧客群それぞれについて、機械学習モデル（トスアップ選択モデル、コール選択モデル、コンタクト選択モデル）を用いて各主体の行動を評価する。

図2の上段は、インサイドセールスパーソンがコールをせず、フィールドセールスパーソンがコンタクトを取ったグループ1の顧客群に関する序列比較の結果を示している。各モデルの予測確率が高い顧客群同士で比較すると、コンタクト選択モデルの方がコール選択モデルおよびトスアップ選択モデルよりも成約率が高い。つまり、グループ1については、フィールドセールスパーソンの選択に従って予測確率が高い顧客からコンタクトを取った方がインサイドセールスパーソンの選択に従うよりも成約率が高いことを示している。

一方で、図2の下段は、インサイドセールスパーソンがトスアップし、フィールドセールスパーソンがコンタクトを取ったグループ5の顧客群に関する序列比較の結果を示している。グループ5の顧客群においては、コール選択モデルおよびトスアップ選択モデルに従って予測確率が高い顧客からコンタクトを取った方がコンタクト選択モデルに従うよりも成約率が高いという結果になっている。

² AUCはランダムな予測では0.5になり、完全な予測では1.0となる精度評価指標である。

これらの結果は、フィールドセールスパーソンとインサイドセールスパーソンそれぞれが、異なる顧客群をターゲティングするための行動に特化しており、これら各主体の特化した行動をターゲティング顧客群以外に適用すると成約率が低下するなど営業効率が悪化する可能性があることを示唆している。

既述の通り、序列比較による結果は、アウトカムである成約の有無が観測された顧客に限定されている。特に、インサイドセールスパーソンがトスアップした顧客全体であるフィールドセールスパーソンがコンタクトしないと判断した顧客を含めて評価することの意義が高い。

4.3. 収縮法比較

前節の序列比較と異なり、アウトカムである成約の有無が観測されていない顧客を含めた、インサイドセールスがトスアップした顧客全体で各主体の行動を評価するために、収縮法による比較を行う。

図3は、インサイドセールスパーソンがトスアップした顧客群であるグループ4,5での収縮法の比較を示している。具体的には、セル化を通じてコンタクト率が最も高い営業部店グループの顧客群について、各機械学習モデルの予測確率に従って、四分位グループ（Q1～Q3）のコンタクト率と一致するようにコンタクト数を調整し、成約率を算出している。ここで我々がおいている重要な想定として、各四分位グループの顧客群の特性が平均的には類似しているため成約率を公平に比較することができるといものがある。この点に関しては、実際に、Q1～Q4についてコンタクトした顧客先に占める成約先数の割合を確認したところ、それぞれ7.4%、11.2%、10.1%、9.7%と概ね同水準にあると言える。

図3における青線の各点は、各四分位グループのコンタクト率に応じた成約率である（実績）。また、赤線の各点は、コンタクト予測モデルによる予測確率に基づいてコンタクトする顧客を選択した場合の成約率である（コンタクト選択モデル）。コンタクト選択モデルの結果は、仮にフィールドセールスパーソンがインサイドセールスパーソンからのトスアップ情報がない状況においてグループ0,1の顧客群と同じようにコンタクトを取る顧客を選択した場合の結果を表していると解釈することができる。このコンタクト選択モデルの結果が実績よりも下に位置していることから、トスアップ情報およびトスアップ情報を受けたフィールドセールスパーソンの行動変容によって、成約率が改善することを示唆している。

また、黄線および緑線の各点は、インサイドセールスパーソンのコール選択モデルおよびトスアップ選択モデルの予測確率に基づいてコンタクトを取る顧客を選択した場合の成約率であり、これらが実績やコンタクト選択モデルよりも上に位置している。つまり、インサイドセールスパーソンがトスアップする顧客群に関しては、フィールドセールスパーソンの判断ではなく、インサイドセールスパーソンの判断を重視してコンタクトを取る顧客を

選択することで成約率をより改善できることを示唆している。

最後に、拡張として、インサイドセールスおよびフィールドセールスが成約獲得を目的として行動した場合を考える。Kleinberg et al. (2018)では、人間と機械学習モデルの判断を比較しており、本研究においても成約を予測変数とした機械学習モデルを構築し、比較を行う。この目的を成すために、グループ 1,3 のフィールドセールスがコンタクトしたアウトカムが観測可能な顧客群のデータを用いて学習を行った成約予測モデルを利用する³。図 3 における黄緑線の各点は、成約予測モデルでの予測確率に基づいてコンタクトを取る顧客を選択した場合の成約率を表しており、最も上位に位置している。インサイドセールスの役割はトスアップを行うことであり、インサイドセールスがトスアップした後にフィールドセールスがコンタクトし、さらに成約に至るかどうかには直接関与することはできない。そのため、成約獲得ではなくトスアップを行うこと自体が目的となり得るため、成約予測モデルの予測確率に基づいてコンタクトする顧客を選定の方が望ましい可能性がある。

5. 結語

本研究では、法人向けリース契約を対象としたフィールドセールスパークソンがより高い営業成果を上げるために、インサイドセールスパークソンが生産した情報がどのような意味を持つかを実証的に検討した。実際のリース営業活動データと潜在顧客リストなどの通常は研究に用いることが難しいデータを用いた分析から、フィールドセールスパークソンが必ずしも十分なリード情報を持っていないと考えられる企業クラスについて、インサイドセールスパークソンによる追加的な情報提供が成約確率を改善させることを確認した。こうした結果は、企業全体としてより高い成果を上げるという目的の実現にとって複数の情報生産技術が貢献することを示唆するものである。

将来的な研究課題としては、第一に、観察データから推論した上記の結果について、ランダム化比較実験を通じて追試することが考えられる。この時、どのようなクラスの企業について、どのようなバックグラウンドを持ったインサイドセールスパークソンによる情報生産が効果的か、といったより詳細な因果効果の識別が期待される。第二に、本研究では十分に扱うことのできなかった、リース案件の成約以外のアウトカムを分析に含めることで、フィールドセールスパークソンとインサイドセールスパークソンの行動様式をより正確にモデル化するとともに、マーケティング活動に係る資源の配分を最適化することが期待される。いずれの課題も学術研究と企業実務の密接なコラボレーションによってのみ実現されるものであり、引き続きかかる方針に基づく研究が蓄積されることが期待される。

³ 成約予測モデルの評価セットにおける AUC は 73.3%である。

参考文献

- Corrado, C., & Hao, J. (2014). Brands as productive assets: Concepts, measurement, and global trends,” WIPO Economic Research Working Paper No. 13.
- Corrado, C., Hulten, C., & Sichel, D. (2005). Measuring capital and technology: An expanded framework. In C. Corrado, J. Haltiwanger, & D. Sichel (Eds.), *Measuring Capital in the new Economy*, studies in income and wealth (Vol. 65, pp. 11–46). University of Chicago Press.
- Corrado, C., Hulten, C., & Sichel, D. (2009). Intangible capital and U.S. economic growth. *Review of Income and Wealth*, 55, 661–85.
- Gourio, F., and Rudanko, L. (2014). Customer Capital. *Review of Economic Studies* 288(3), 1102-1136.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *Quarterly Journal of Economics*, 133(1). doi:10.1093/qje/qjx032
- Lakkaraju, H., Kleinberg, J., Leskovec, J., Ludwig, J., & Mullainathan, S. (2017). The selective labels problem: Evaluating algorithmic predictions in the presence of unobservables. *Proceedings of the ACM SIGKDD International Conference on Knowledge. Discovery and Data Mining Part F129685*.
- Rob, R. and Fishman, A. (2005). Is bigger better? Customer base expansion through word-of-mouth reputation. *Journal of Political Economy* 113(5), 1146-1162.
- Sveikauskas, L., Soloveichik R., Garner, C., Meyer, P. B., Bessen, J. & Russell, M. (2024). Marketing, other intangibles, and output growth in 61 United States industries. *Journal of Income and Wealth* 70(4), 1190-1215.

図表

表1 グループ別成約率

グループ	コール	トスアッ プ	コンタク ト	取引先数	成約先数	成約率
0	無し	無し	無し	33,702	0	0.0%
1	無し	無し	有り	1,894	414	21.9%
2	有り	無し	無し	14,409	0	0.0%
3	有り	無し	有り	1,057	198	18.7%
4	有り	有り	無し	619	0	0.0%
5	有り	有り	有り	447	45	10.1%
総計				52,128	657	1.3%

(注) 表は見込顧客リストの顧客をコール、トスアップ、コンタクトの有無に応じて、6グループに分類し、各グループの取引先数および成約先数、成約率を示している。

表2 企業属性情報の基本統計量

	欠損	平均	標準偏差
資本金	7,474	209,725	5,566,716
従業員数	3,235	83	509
社齡	767	39	19
当期売上高	3,718	3,085,733	28,661,372
前期売上高	5,665	3,144,187	29,584,765
当期利益	16,061	86,369	904,057
前期利益	16,404	80,644	1,026,487
当期売上高成長率	5,885	103	115
前期売上高成長率	7,773	103	135
当期利益 / 当期売上高	16,634	-0.01	3.34
当期利益 / 当期売上高	16,811	-0.08	14.49
取引先銀行フラグ	353	0.07	0.26

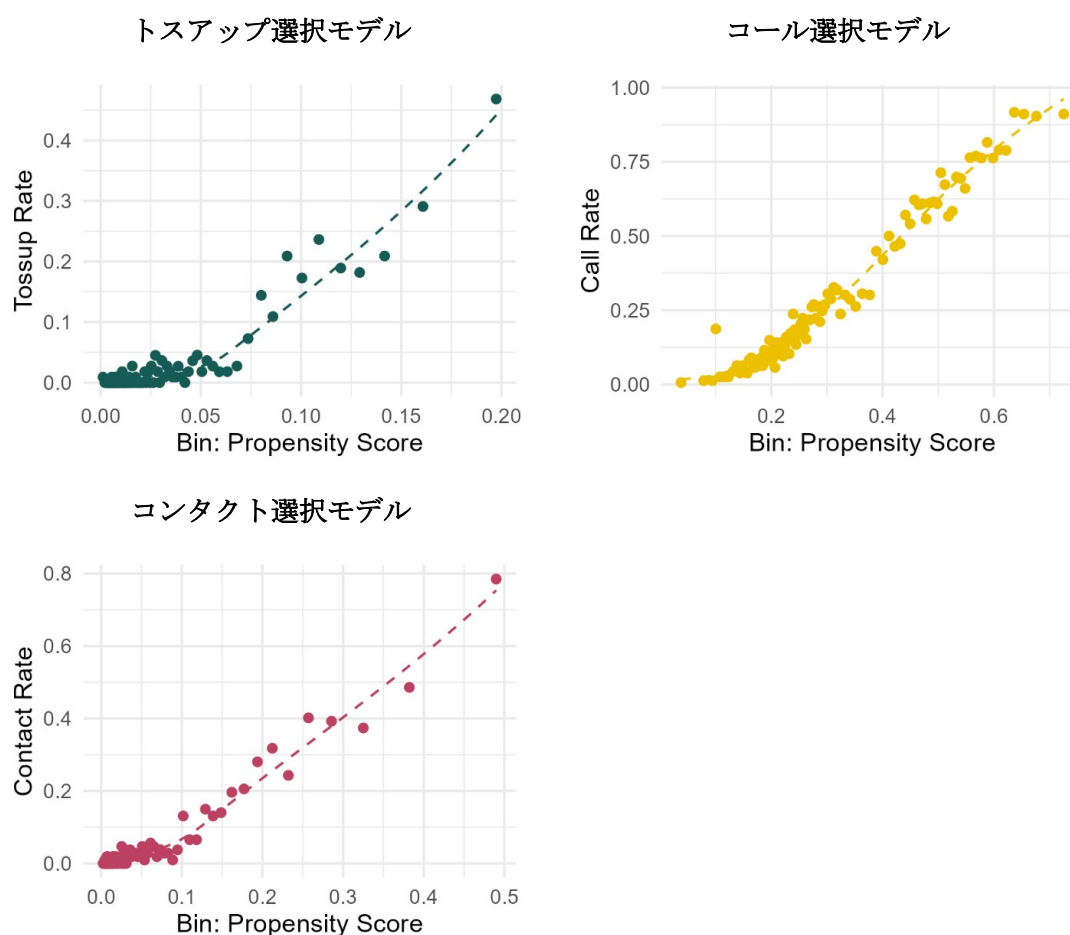
(注) 表は、企業属性情報の基本統計量を示している。

表3 予測精度

モデル	予測変数	データセット	AUC(CV)	AUC(Test)
トスアップ選択モデル	トスアップ	グループ{0,1,4,5}	90.6%	90.9%
コール選択モデル	コール	全グループ	82.8%	83.0%
コンタクト選択モデル	コンタクト	グループ{0,1}	90.4%	90.8%

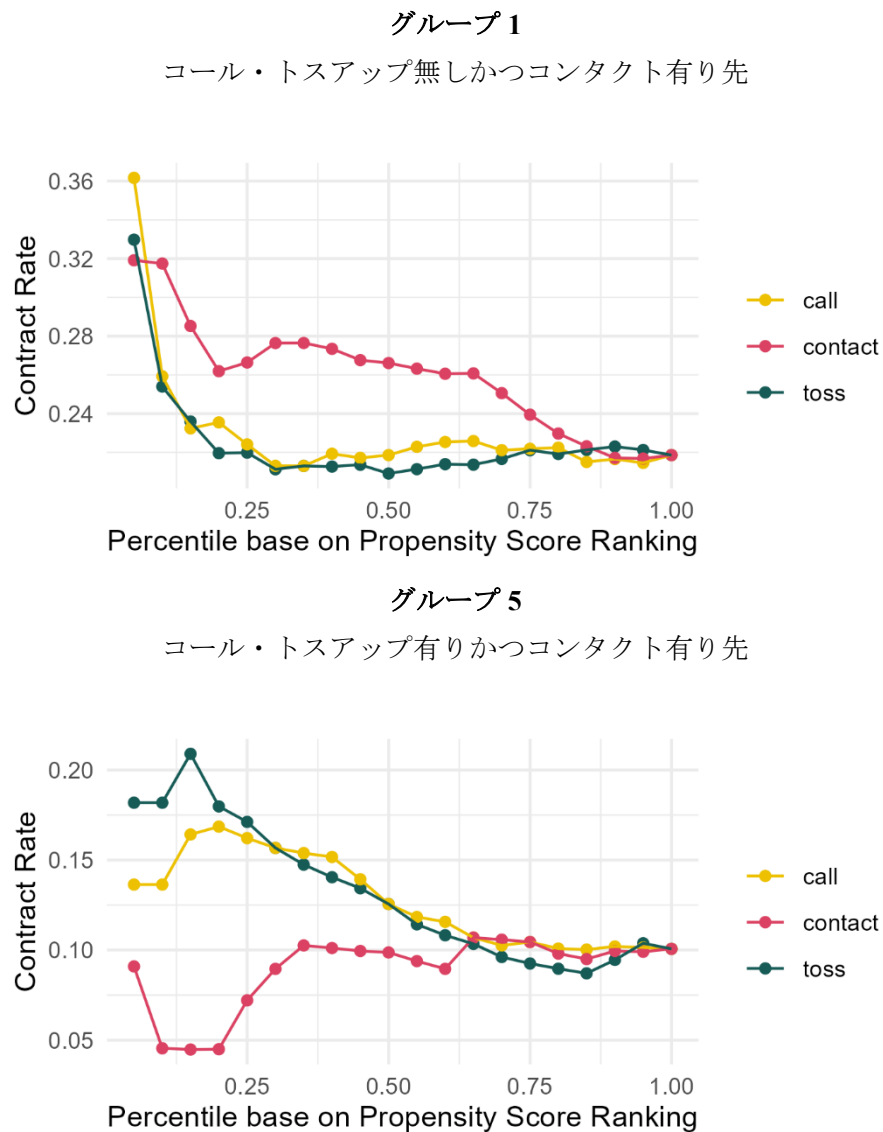
(注) 表は、インサイドセールスおよびフィールドセールスの行動を予測変数として構築した機械学習モデルの予測精度 (AUC, Area Under Curve) を示している。AUC(CV)は学習セットでのクロスバリデーションセットにおける平均 AUC を、AUC(Test)は評価セットでの AUC を示している。

図1 予測値と実績値



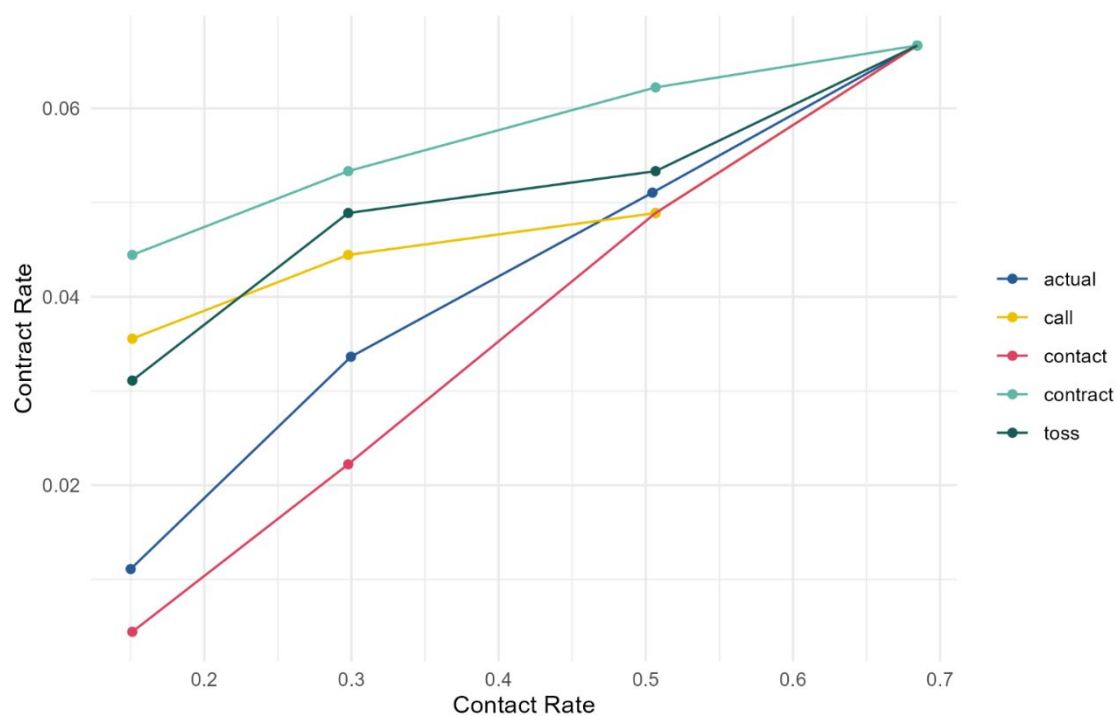
(注) 各モデルによる予測確率を百分位ごとに100ビンに分割し、各ビンの予測変数の発生率を描画した散布図。横軸は予測確率のビンごとの平均値、縦軸は予測変数の発生率を示している。データは、各モデルの構築データセットでの評価セットを利用している。

図2 序列比較



(注) 各予測モデルの予測確率に基づいてデータを降順に並べ、上位の累積パーセンタイルごとの成約率をモデルごとに描画したグラフ。横軸は対象データの上位パーセンタイル点、縦軸はその範囲における成約率を示している。曲線が上に位置するほど、予測が有効であることを意味する。

図3 収縮法比較



(注) 図は収縮法による比較結果を示している。各点は、コンタクト率が最も高い営業部店グループの顧客群について、各機械学習モデルの予測確率に従って、四分位グループ（Q1～Q3）のコンタクト率と一致するようにコンタクト数を調整したうえでの成約率を描画している。横軸はコンタクト率、縦軸は成約率をそれぞれ表す。成約率は当該グループの顧客数に占める（仮想）コンタクト先のうち成約した顧客先数の割合として算出している。