

Kanagawa University Economic Society
Discussion Paper No.2021－02

クレジットカード利用情報を用いた
サービス消費のリアルタイム予測

2022 年 3 月

浦沢 聡士

クレジットカード利用情報を用いたサービス消費のリアルタイム予測

浦沢 聡士[†]

要旨

本稿では、経済分析の分野において利活用が進むオルタナティブデータの速報性を活かして経済動向の早期把握を行うため、予測モデルを構築した上で、公的統計の疑似的なリアルタイム予測を行う。その際、コロナ禍以降、これまでの傾向と異なり大きく変動することとなった我が国におけるサービス分野の消費動向に着目し、その予測を行うためクレジットカードの利用情報を用いる。

本稿での検証より得られた結果は、クレジットカードの利用情報が、コロナ禍以降のサービス消費の傾向をいち早く捉えることに有用であることとともに、公的統計の予測を行う上で一定程度の予測力を有することを示している。こうした特徴を持つクレジットカードの利用情報は GDP 統計で示される一国経済全体のサービス消費の動きとも整合的であり、そうであるならば、四半期に一度の頻度でしか補足されない GDP 統計の公表を待つことなく、GDP 統計とも整合的なサービス消費の動きをよりタイムリーに補足する可能性が示唆される。

JEL 分類番号 : E37

キーワード : オルタナティブデータ、公的統計、線形回帰モデル、Lasso 回帰モデル、Ridge 回帰モデル、動学因子モデル

1. はじめに

昨今、経済分析の分野においても、位置情報や検索情報、POS 売上データやクレジットカードの利用情報など、オルタナティブデータ (AD) と呼ばれる従来用いられてきた公的統計と特性を異にするデータの利活用が進んでいる。AD は、公的統計と異なり、特定の調査目的のもと設計、調査、集計されたデータではなく、企業等が日々活動を行う中で派生的に生成されたものと言える。そのため、公的統計と比べ、高頻度で速報性に優れている、また、より詳細な情報を有するといった特性を一般に持つ一方、データの代表性、継続性などについて利用に際し一定の留意が求められている。

こうした特性を持つ AD を経済分析の中で活用する目的については様々あるが、

[†] 神奈川大学経済学部. E-mail: surasawa@kanagawa-u.ac.jp

本稿で示された内容や見解は筆者個人によるものであり、所属する機関のものではない。あり得べき誤りは筆者個人の責に帰するものである。本研究は JSPS 科研費 JP21K20162 の助成を受けたものである。

本稿では、特に、AD の持つ速報性に着目し、経済動向の早期把握を目的とした活用を検討する。AD は、よりタイムリーでありながら信頼性の高い情報を提供することができるのであろうか。

具体的には、AD の速報性を活かすとともに、そうしたデータを用いて経済動向の早期把握を行うため、因子モデルや機械学習の手法を用いて予測モデルを構築し、公的統計の疑似的なリアルタイム予測を行う。その際、コロナ禍以降、これまでの傾向と異なり大きく変動することとなった我が国におけるサービス分野の消費動向に着目し、その予測を行うためクレジットカードの利用情報（「JCB 消費 NOW」）を用いる¹。サービス分野の消費については、本稿で予測の対象とする「サービス産業動向調査」（総務省）を含め公的統計による補足の遅れが指摘されている一方で、業界データやクレジットカードの利用情報、インターネットの閲覧数等公的統計に限らずよりタイムリーなデータの利活用が進んでいる分野でもあり、こうしたデータを有効に活用することが期待される分野と言える。

AD を用いた既存統計の予測に関する取組は、諸外国はもとより我が国においても研究が蓄積されてきている。例えば、景気判断における検索データの利用可能性を筆者の知り得る限り我が国で初めて議論する白木他（2013）では、サービス消費のうち旅行に着目し、検索データ（Google Trend）を用いて旅行取扱額の予測を行っている。Moriwaki（2019）では、ハローワークへの来訪者数を位置情報により計測することで失業率の予測を行っている。他にも、大手電力会社が公表する電力需要の高頻度データをもとに鉱工業生産指数を予測する水門他（2019）や位置情報と施設情報等を組み合わせることで小売、娯楽などのサービス業の売上動向や製造業の生産活動を予測する王他（2021）、トラックカーナビデータ（通行台数データ）をもとに鉱工業生産指数を予測する川村他（2021）などがある。

本稿での結果は、クレジットカードの利用情報が、コロナ禍以降のサービス消費の傾向をいち早く捉えることに有用であることを示唆しているが、本稿においても、上述の先行研究と同様、経済動向のリアルタイム把握における AD の有用性を検証することを通じ、AD の利活用を促していくことを企図している。

本稿の構成は、以下のとおりである。次節で、我が国における最近のサービス消費の動向について概観した後、第 3 節で、本稿で用いるクレジットカード利用情報の特性を確認する。第 4 節では、クレジットカード利用情報を用いたサービス消費の疑似的なリアルタイム予測の手法、及び予測結果を報告するとともに、予測パフォーマンスについて評価を行い、第 5 節で、結論を述べる。

2. 我が国におけるサービス分野の消費動向

我が国におけるサービス分野の消費動向を捉える代表的な指標として、GDP 統

¹ 「JCB 消費 NOW」は、株式会社ジェーシービー、及び株式会社ナウキャストにより、統計閲覧会員向けに提供されている。

計のサービスがある（形態別国内家計最終消費支出の一形態として耐久財、半耐久財、非耐久財とともに公表されている）。GDP 統計によれば、GDP に占めるサービス消費の割合は 2020 年には 3 割程度となっており経済活動の一定割合を占めていることがわかる。

こうしたサービス消費の詳細を見るために、GDP 統計の推計にも用いられている「サービス産業動向調査」（サビ動）を見る。サビ動は、月次統計であるが、GDP 統計の速報推計、年次推計のいずれにも基礎統計として用いられており、我が国のサービス消費の動向を測る上で最も重要な指標であることから、本稿でもリアルタイム予測の対象としている。

図 1 では、まず、サービス消費全体の動向を見るため、サビ動におけるサービス産業全体の売上高（前年比）を GDP 統計のサービス（名目（除く帰属家賃）、前年比）とともに示している。両者は当然に整合的な動きとなるが、いずれを見ても、2020 年第 2 四半期（2020Q2）以降、コロナ禍において緊急事態宣言の発令、解除が繰り返される中、過去の傾向と異なり大きく変動していることが確認できる。

表 1 では、2020 年におけるサービス産業全体の平均的な売上高の伸び（-10.2%）とともに分野別の売上高の動向を示しているが、コロナ禍において経済、社会活動が抑制される中、娯楽業、飲食店、宿泊業、旅客輸送業といった分野において、特に落ち込みが顕著であったことがわかる。

感染症拡大の影響を強く受けるこうしたサービス消費の動きをリアルタイムで把握することは、感染症対策と経済政策の両立を目指す上でも、今後、益々重要となる。しかし、サビ動には速報値であったとしても 2 か月程度のレポートラグが伴われるため、サービス消費の分野において、今起こっている変化をタイムリーに把握することができない。こうした中で、AD には公的統計に伴われる補足の遅れを補う役割が期待されている。

図 1 では、GDP 統計、サビ動に加え、本稿で用いるクレジットカードの利用情報（「JCB 消費 NOW」）の動き（四半期値については、各四半期における月次データの平均値として計算）を示しているが、図 1 より、「JCB 消費 NOW」はコロナ禍での変動を含め GDP 統計、サビ動の動きを十分に捉えていることが確認できる。こうしたことは、「JCB 消費 NOW」がサビ動のリアルタイム予測に高い予測力を有する可能性とともに、GDP 統計において示される一国経済全体のサービス消費とも整合的であることを示唆している。そうであるならば、「JCB 消費 NOW」は、四半期に一度の頻度でしか補足されない GDP 統計の公表を待つことなく、2 週間に一度の頻度で、GDP 統計とも整合的なサービス消費の動きを補足することを可能とし、GDP 統計を中心に据えて行われる経済動向の評価、ひいては経済政策運営の迅速化に資することも期待できる。

3. クレジットカード利用情報の特性

本稿では、我が国におけるサービス消費のリアルタイム予測を行うため、クレジットカードの利用情報（「JCB 消費 NOW」）を用いる。以下では、まず、「JCB 消

費 NOW」の特性について概観する。

クレジットカードショッピング信用供与額（日本国内のクレジットカード会社が発行したクレジットカードを利用して消費者がクレジットカードショッピングを行った額）の 2020 年の値は 74.5 兆円程度（日本クレジット協会）であり、その額は増加傾向にある。GDP 統計の民間最終消費支出に占める割合を見ても、2013 年 14%程度から、2020 年には 26%程度にまで増加し、クレジットカードを用いた消費の役割が高まっていることが窺える。本稿で用いる「JCB 消費 NOW」は、JCB グループのカード発行会社が発行するカードを利用する会員のクレジット利用情報を用いたものであるが、JCB が行う事業における取扱額が 2020 年に 33.8 兆円程度（株式会社ジェーシービー）であることに鑑みると、我が国における市場規模の一定程度を占めていると考えられる。

「JCB 消費 NOW」は、上記業務データをもとに、データ分析への利用を図るため、株式会社ナウキャストにより、加工、集計されたものであり、具体的には、以下による多段階での加工・調整が行われている。

- ・ 約 1000 万サンプルのクレジットカード利用情報を匿名加工（調査対象消費金額は年 4 兆円以上。利用金額とともに、財、サービス別に細分化された業種（60 業種以上）、居住地（都道府県別）、年代（14 区分別）、性別といった情報もあわせて記録）
- ・ 集計に際し、カードホルダーのウエイトの偏りを調整するため、細分化された業種、及び利用者の属性別に合計支出金額の変化率を人口推計データのウエイトを用いて調整
- ・ また、業種ごとの支出の偏りを補正するため、「商業動態統計」（経済産業省）、「消費活動指数」（日本銀行）から得られる業種間ウエイトを利用し補正

こうして集計されたデータは、半月ごとに集計され、概ね 2 週間のレポーティングラグで公表されている（2021 年 11 月にリニューアルされたデータは 2016 年 4 月分より利用可能）。

「JCB 消費 NOW」は、カード利用者と加盟店の間でクレジット利用に基づく取引が成立した時点での情報を記録することからも、消費者を対象とした需要側データであるとともに、事業者を対象とした供給側データとも考えられるが、多量のサンプルを用いることにより企業や事業所を調査対象（約 37000 サンプル）とするサビ動より得られる情報を補完する可能性が考えられる。また、本稿での分析目的を超えた視点となるが、需要側の情報（誰が、何を買ったのか）と、供給側の情報（どの企業・事業所が、何を売ったのか）の同時把握が可能となることで、それらを組み合わせることにより、“どこに住む誰が、いつ、どこで、誰から、何を、いくら買ったのか”、といった公的統計では捉えきれない詳細な消費動向の把握も可能となる。

次節では、こうした「JCB 消費 NOW」が持つ情報を用いて公的統計であるサビ動を予測するが、それに先立ち、ここでは、両者の動きを、サービス消費合計とともに、コロナ禍において特に変動の大きい娯楽業、飲食店、宿泊業、旅客輸送業について確認する。図 2 では、サビ動に概念上対応する「JCB 消費 NOW」（マクロ

セクター)の項目として、娯楽業については「娯楽(映画、スポーツ施設、遊園地)」、飲食店については「外食」、宿泊業については「宿泊(ホテル、旅館)」、「旅行」、そして旅客輸送業を含む運輸業等に対応するものとして「交通(鉄道旅客、バス、タクシー、航空旅客)」を用いて両者の比較を行っている。

サービス消費合計については、図1で見たように、感染症拡大の影響を受ける以前については、概ね一致した動きとなっていたが、コロナ禍において両者に乖離する動きが観察されるようになった。こうした乖離は足もとにかけて引き続き見られるが、前年比がマイナス、またはプラスに転じる時期が基本的に一致するなど両者が整合的に動く傾向に変わりはない。

その一方、図2において、こうした動きを分野別に見ると、2020Q2を底として前年比が加速し、2021Q3に再び前年比が減速するといった傾向は同様であるが、サービス消費合計の場合とは異なり、両者に傾向的な乖離、また、コロナ禍において特に大きな乖離が発生して分野もある。例えば、宿泊について、サビ動では2020Q2に-70%以上の落ち込みとなったが、「JCB消費NOW」では-20%弱程度にとどまっている。こうした乖離を生む要因の1つとして、サビ動にはインバウンドによる影響を含む一方で、「JCB消費NOW」にはそうした海外需要の影響を受けないといったことが考えられる(サービス消費合計についても、インバウンドによる影響の有無により、サビ動では、「JCB消費NOW」に比べ、振れが大きくなっている可能性が考えられる)。

ただし、いずれにせよ、経済で起こっている変化をリアルタイムで把握することを目的とする場合、「JCB消費NOW」は、程度の差(前年比の幅に差)はあれ、サビ動と同様、これまでの傾向と異なり大きく変動することとなったサービス消費の動きを捉えており、その信頼性とサビ動のリアルタイム予測に用いることの妥当性を示していると言える。

4. 疑似的リアルタイム予測

以下では、「JCB消費NOW」を用いてサビ動を予測対象とする疑似的なリアルタイム予測を行い、その精度を評価していくが、それに先立ち、まず、「JCB消費NOW」が有する予測力を検証するため、予測対象とするサビ動のサービス産業全体の売上高(前年比)を被説明変数とした線形回帰モデルの推計を行う。

(「JCB消費NOW」の予測力)

線形回帰モデルの説明変数としては、被説明変数の1期ラグに加え、「JCB消費NOW」における、外食、宿泊、旅行、交通、娯楽に加え、サービス消費に占めるシェアが大きい通信、医療といった変数(いずれも前年比)を用いる(モデル(1))。また、こうした個別分野における利用額を集計したサービス消費合計(前年比)を説明変数とするモデルの推計も行っている(モデル(2))。

表2は、推計期間を2017年5月～2021年6月までとする月次データを用いた推計結果を示している(参考として、コロナ禍の影響を除く2019年12月までを推計期間とした結果も報告)。個別分野の変数を用いたモデル(1)の推計結果

を見ると、推計期間にコロナ禍における動きを含むことで、外食、宿泊、娯楽などコロナ禍において特に大きな変動が見られた分野を中心に一定程度の予測力を有していることがわかる。なお、モデル (1) と (2) の推計結果より得られた RMSE を比較することで、サービス消費合計といった集計された情報を用いるより、個別分野の動きを表す細分化された情報を用いることで予測の精度が高まることが確認できる。

(予測モデル)

以下では、上述の結果も踏まえ、「JCB 消費 NOW」を用いたサビ動の疑似的リアルタイム予測を行い、その精度を評価していく。予測に用いるモデルについては、上述の最小二乗法による 2 つの線形回帰モデルに加え、正則化項付最小二乗法による Lasso 回帰、Ridge 回帰モデル、更には、動学 1 因子モデル (DFM) を用いる。なお、モデルの推計に際し、線形回帰モデルを除きデータの標準化を行っている。また、正則化項の強さを決めるパラメータについては、交差検証により予測誤差が最小となる値を利用している。

一般に、適切な予測モデルの選択は、予測対象とする変数のデータ生成過程の影響を受ける。本研究では、データが sparse (疎) な過程に従う場合、dense (密) な過程に従う場合の双方を念頭に、前者に対応するモデルとして Lasso 回帰モデル、後者に対応するモデルとして Ridge 回帰モデル、DFM を用いている。加えて、参考として、自己回帰モデルを用いた予測に加え、更には、モデルを用いずに「JCB 消費 NOW」のサービス消費合計の値をサビ動のサービス産業全体の売上高の予測値として用いた場合の結果も報告する。後者の見方は、「JCB 消費 NOW」とサビ動の動きが一致すると仮定し、先んじて公表される「JCB 消費 NOW」の動きからサビ動の動きを簡便的に予測するといったより実務に近いアプローチと言えるが、こうした実務的な利用と比べ、モデルを用いることにより予測精度がどの程度高まるかを併せて確認していく。

線形回帰モデル

線形回帰では、以下の残差 2 乗和を最小化することにより、モデルの推計が行われる。

$$\hat{\beta}^{OLS} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{t=1}^N (y_t - \beta_0 - \sum_{i=1}^P x_{ti}\beta_i)^2 \right\}$$

ここで、 y は被説明変数 (予測対象とする変数)、 x は説明変数を表す。説明変数については、上述のモデル (1)、(2) の中でそれぞれ用いた変数を利用する (OLS (1)、OLS (2))。

Lasso 回帰モデル

Lasso 回帰では、上述のモデル (1) の説明変数を用い、以下の L1 正則化項を含む残差 2 乗和を最小化することにより、モデルの推計が行われる (Lasso)。

$$\hat{\beta}^{Lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{t=1}^N (y_t - \beta_0 - \sum_{i=1}^P x_{ti} \beta_i)^2 + \lambda \sum_{i=1}^P |\beta_i| \right\}$$

なお、 λ は係数の縮小の程度を表す複雑度パラメータを表す。

Ridge 回帰モデル

Ridge 回帰では、上述のモデル (1) の説明変数を用い、以下の L2 正則化項を含む残差 2 乗和を最小化することにより、モデルの推計が行われる (Ridge)。

$$\hat{\beta}^{Ridge} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{t=1}^N (y_t - \beta_0 - \sum_{i=1}^P x_{ti} \beta_i)^2 + \lambda \sum_{i=1}^P \beta_i^2 \right\}$$

動学 1 因子モデル

動学 1 因子モデルでは、以下の観測方程式、状態方程式より得られる共通因子をもとに予測が行われる。

$$y_{it} = \gamma_i C_t + u_{it} \quad (\text{観測方程式})$$

$$\phi(L)C_t = \eta_t \quad (\text{状態方程式})$$

ここでは、 y を被説明変数及び説明変数（上述のモデル (1) の説明変数に加え GDP 統計のサービスを利用）を含む観察されるデータとして表わす。 C は観察されることのない共通因子であり、その推計に際してはモデルを状態空間表現により表わした上で、カルマンフィルターを用いた逐次計算を行う。なお、 γ は因子負荷ベクトル、 u は固有ベクトルを表す。また、 C 、 u はそれぞれ自己回帰モデルに従うものとする。

動学 1 因子モデルについては上述の標準モデル (DFM) に加え、予測に用いる変数を予測力という観点から Lasso によりスクリーニングしたモデル (Lasso-DFM) も用いる (Lasso-DFM では、変数選択の結果、上述のモデル (1) の説明変数のうち外食、旅行、通信が除外)。

(予測の枠組み)

以上の 6 つのモデル（及び参考として報告する 2 つ）を用いてサビ動のサービス産業全体の売上高（前年比）の予測を行う。

予測期間についてはコロナ禍を含む 2020 年 1 月～2021 年 6 月とし、この期間における 1 期先予測の結果をもとにモデルの予測精度を評価する。予測の 1 期目となる 2020 年 1 月の予測については、2017 年 5 月～2019 年 12 月までのデータを用いて推計されるモデルをもとに行い、2020 年 2 月以降については、予測を行うたびにデータ期間（学習期間）を 1 か月分追加していくことで行う。

こうして予測されるサビ動の 1 期先予測が得られるタイミングは、予測対象

月終了後の1か月後となり、サビ動の公表より1か月早い時点となる。例えば、2021年6月の予測については、2021年5月のサビ動が公表される2021年7月末に、既に7月中旬に公表済である「JCB消費NOW」の6月データを用いて実施される。

なお、本来、リアルタイム予測については、日々、公表・更新される最新のリアルタイムデータを予測に取り込んで行うが、過去に遡ってリアルタイムデータを構築することは容易ではなく、本稿でも、2021年11月時点で利用可能なデータにより構築される2017年5月～2021年6月までのデータセットを用いて、擬似的にリアルタイム予測を行っている。

(予測精度)

以下では、コロナ禍を含む2020年1月～2021年6月を対象として実施した1期先予測の精度を、上述のモデルについて、RMSFE、MAFE等の誤差統計をもとに評価する。

表3を見ると、予測誤差が最も小さい予測モデルは、Ridgeモデルであり、次いで、Lasso-DFMとなっている。MAFEを見ると、Ridgeモデルでは2.45%、Lasso-DFMでは2.57%と、2%台半ば、RMSFEで見ても、Ridgeモデルでは2.81%、Lasso-DFMでは3.23%と、3%前後となっている。

経済動向のリアルタイム把握を目的とする場合、前年比で見て2%～3%程度の誤差は、サービス消費が経済全体（GDP）に占める割合（3割）を鑑みても決して小さくはないが、上記期間におけるサービス産業全体の売上高（前年比）の標準偏差が10%程度と、コロナ禍において過去に例のないほど大きく変動したことを踏まえれば、モデルによる予測に限界があることも否めない²。なお、参考までに、GDP統計のサービス（名目、前年比）について、1次速報値（3か月目情報についてはサビ動の実績が得られず補外推計）から2次速報値（3か月目情報についてもサビ動の実績を反映し推計）への改定幅を2020Q1～2021Q2について見ると、改定差の絶対値平均で0.5%程度（最も改定差が大きかった2021Q2でも1.2%）となっており、本稿で検討する予測モデルを実務で活用するためには改善の余地も大きい。

その一方で、モデルを用いて予測を行うことの意義は依然として存在する。1つに、モデルを用いることで、モデルを用いずに「JCB消費NOW」の動きから簡便的に予測を行うことに比べ、予測誤差を縮小させることが可能となる。表3によれば、「JCB消費NOW」を用いた簡便的な予測（JCB消費NOW）の誤差は、MAFEで3.41%、RMSFEで4.24%であるので、モデルを用いることで1%程度予測誤差が縮小されている。

また、予測期間を通じた平均的な誤差による評価に加え、サービス消費の動きをリアルタイムで把握するといった観点から評価すると、モデルによる予測は、

²コロナ禍における影響を除いてモデルによる予測の精度を確認するため、予測期間を2019年1月～2019年12月として1期先予測を実施したところ、予測誤差が最も小さくなる場合（Lasso回帰モデル）で、MAFEは0.62%、RMSFEは0.76%となった。なお、同期間におけるサービス産業全体の売上高（前年比）の標準偏差は1.5%程度。

コロナ禍以降、これまでの傾向と異なり大きく変動することとなったサービス消費の傾向をいち早く捉えることができる。

図3では、予測対象であるサービス産業全体の売上高（前年比）について、予測期間中の実績値と各月の1期先予測の結果を、Ridgeモデル（パネルA）と「JCB消費NOW」を用いた簡便的な予測（パネルB）で比較している。初めてとなる緊急事態宣言の発令のもと経済、社会活動が人為的に抑制されたことを背景として見られた、2020年4月以降の消費の急激な落ち込みとその後の回復といったコロナ禍でのリアルタイム予測を行う上で最も関心の高い時期について見ると、モデルによる予測は、2021年4月に-18.2%（実績値：-20.3%）、5月に-20.3%（-23.2%）、そして6月に-16.9%（-13.0%）と、実績値で見られた急激な落ち込みとともに、5月を底に、6月に前年比のマイナス幅が縮小するといった回復傾向を捉えることができる。他方、簡便的な予測では、4月（-19.4%）、5月（-17.3%）、6月（-10.7%）と、4月の落ち込みを適切に捉える一方で、5月の時点で前年比のマイナス幅が縮小するといった実績とは異なる動きを示す。こうした傾向は、同期間を前年の“裏”とする2021年4月～6月にかけても見られ、月々の動きをリアルタイムで把握するといった観点から評価しても、モデルを用いることで、より適切な予測を行うことが可能となる。

5. まとめ

本稿では、オルタナティブデータの速報性を活かして経済動向の早期把握を行うため、我が国におけるサービス分野の消費動向に着目し、そのリアルタイム予測を行うためクレジットカードの利用情報（「JCB消費NOW」）を用いた。

本稿での検証を通じ、クレジットカードの利用情報は、コロナ禍以降のサービス消費の傾向をいち早く捉えることに有用であること、その結果、予測精度には改善の余地もあるが、1か月という公的統計の半分のレポーティングラグで得られるクレジットカードの利用情報を用いた予測は、公的統計の予測を行う上で一定程度の予測力を有することが示された。繰り返しになるが、こうしたクレジットカードの利用情報が事後的にGDP統計で示される一国経済全体のサービス消費の動きと整合的であり続けるのであれば、四半期に一度の頻度でしか補足されないGDP統計の公表を待つことなく、GDP統計とも整合的なサービス消費の動き、特に、その大きな変化に関する兆候をよりタイムリーに補足する可能性が示唆され、その結果、経済政策を含む意思決定に資するより有用な情報を提供することも可能となる。

なお、今回の検証については、2017年5月以降の比較的限られたサンプルをもとに、コロナ禍といった特に変動の大きい時期について行った。引き続き、検証を行っていくことで、時間の経過とともにサンプル数が増加し、また、サービス消費の動きが安定化する中で、モデルによる予測精度も向上していくことが期待される。

引用文献

王悠介, 須合智広, 高橋耕史, 松村浩平, 2021. 位置情報データによる経済活動のナウキャストニング. 日本銀行ワーキングペーパーシリーズ No. 21-J-2.

川村健史, 久保達郎, 淀谷恵実, 2021. トラックカーナビデータを用いた鉱工業生産指数の予測—機械学習を用いた試み—. 経済財政分析ディスカッションペーパー No. 21-1, 内閣府.

白木紀行, 松村浩平, 松本梓, 2013. 景気判断における検索データの利用可能性. 日本銀行調査統計局調査論文.

水門善之, 和泉潔, 坂地泰紀, 島田尚, 松島裕康, 2019. 高頻度電力需要データを用いた製造業活動のナウキャストニングモデルの構築. 人工知能学会第二種研究会資料 2019 巻 SAI-035.

Moriwaki, D., 2019. Nowcasting Unemployment Rates with Smartphone GPS Data. in International Workshop on Multiple-Aspect Analysis of Semantic Trajectories, eds. K. Tserpes, C. Renso, and S. Matwin, 21-33. Springer.

図1 サービス消費の動向（前年比、%）

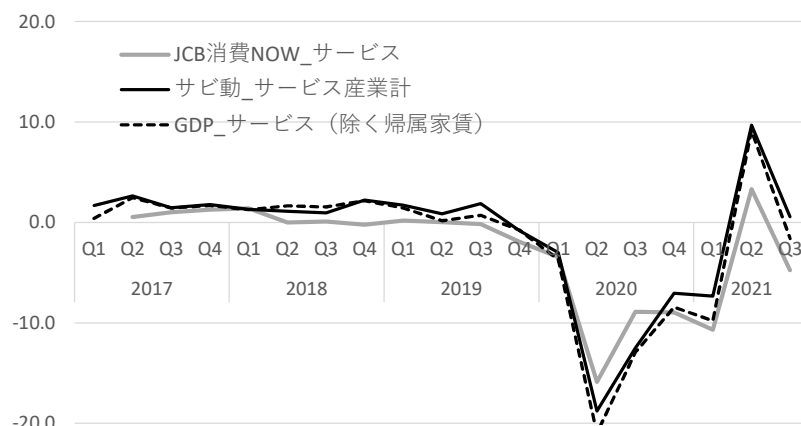
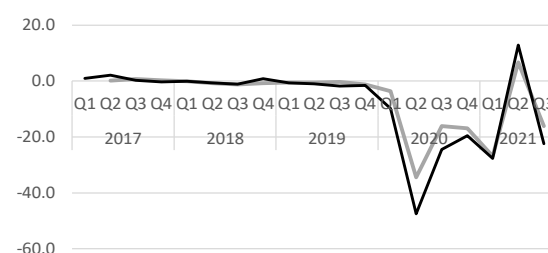


図2 分野別のサービス消費の動向（前年比、%）

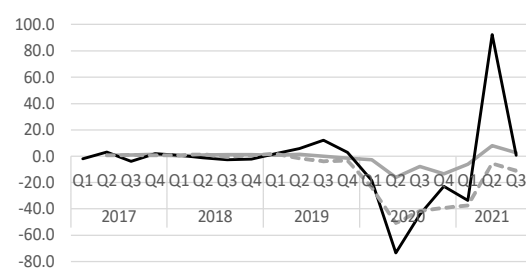
A. 娯楽分野



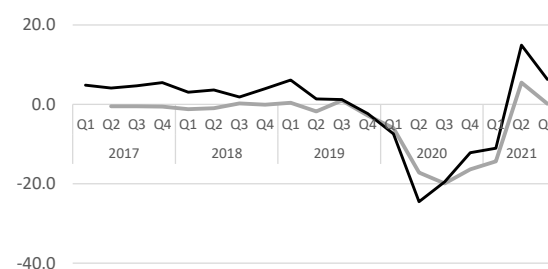
B. 飲食分野



C. 宿泊分野



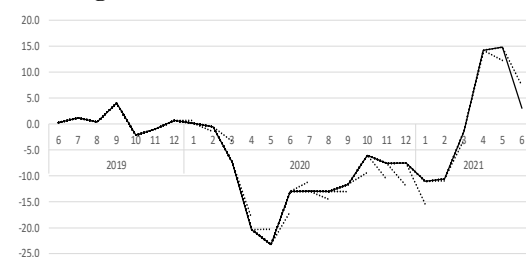
D. 運輸分野



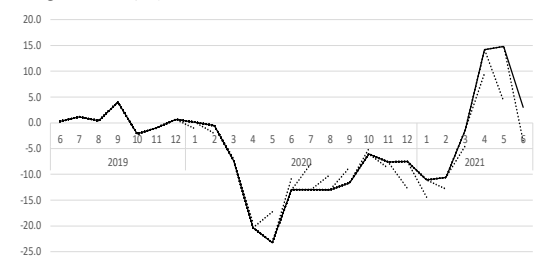
（備考）パネルCでは、「JCB 消費 NOW」の「旅行」の動きを破線で参考表示。

図3 サービス消費の実績値と1期先予測の比較（前年比、%）

A. Ridge



B. JCB 消費 NOW



（備考）実線はサビ動のサービス産業全体の売上高（前年比）、破線はRidge 回帰モデル、「JCB 消費 NOW」による各月の1期先予測値。

表 1 2020 年におけるサービス消費の動向

	2019年	2020年	
	売上高、兆円	売上高、兆円	前年比、%
サービス産業計	33.7	30.2	-10.2
情報通信業	5.0	4.9	-2.3
運輸業、郵便業	5.4	4.5	-15.8
鉄道業	0.7	0.4	-39.4
道路旅客運送業	0.3	0.2	-31.6
航空運輸業等	0.3	0.1	-52.2
不動産業、物品賃貸業	4.4	4.2	-3.1
学術研究、専門技術等	2.9	2.8	-3.2
宿泊、飲食サービス	2.8	2.0	-28.1
宿泊業	0.5	0.3	-40.1
飲食店	2.0	1.5	-27.2
生活関連サービス、娯楽業	4.5	3.3	-26.4
娯楽業	3.1	2.4	-22.0
教育、学習支援業	0.3	0.3	-12.8
医療、福祉	5.0	4.8	-3.6
他に分類されないもの	3.6	3.4	-5.5

(備考) 各年の売上高は当該 12 か月における月間数値の平均値。

表 2 「JCB 消費 NOW」を用いた線形回帰モデルの推計結果

被説明変数: サビ動_サービス産業計 (前年比)	(1)				(2)			
	2017年5月-21年6月		2017年5月-19年12月		2017年5月-21年6月		2017年5月-19年12月	
	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
説明変数								
サビ動_サービス産業計 (1期ラグ)	0.14 **	0.06	0.15	0.12	0.29 ***	0.06	0.03	0.14
JCB_サービス計					0.91 ***	0.07	0.57 ***	0.14
JCB_外食	0.19 ***	0.04	0.00	0.12				
JCB_宿泊 (ホテル、旅館)	0.21 **	0.09	0.21	0.13				
JCB_旅行	0.00	0.04	0.00	0.06				
JCB_交通 (鉄道旅客、バス、タクシー、航空旅客)	0.14	0.09	0.39 ***	0.09				
JCB_娯楽 (映画、スポーツ施設、遊園地)	0.19 ***	0.06	0.07	0.05				
JCB_通信 (ISP業、移動通信業)	-0.29	0.45	0.68 *	0.37				
JCB_医療	0.18 **	0.08	0.18 **	0.07				
定数項	1.66 ***	0.30	1.53 ***	0.24	1.29 ***	0.32	1.19 ***	0.26
Adjusted R-squared	0.96		0.72		0.92		0.33	
RMSE	1.44		0.60		2.04		0.93	
観測数	50		32		50		32	

(備考) *, **, ***は、それぞれ有意水準 10%、5%、1%で有意であることを示す。

表 3 予測モデルに基づく 1 期先予測の誤差

サビ動_サービス産業計 (前年比) の 1 か月先予測								
予測期間: 2020 年 1 月 ~ 2021 年 6 月								
予測モデル								
Evaluation statistics	OLS (1)	OLS (2)	Lasso	Ridge	DFM	Lasso- DFM	参考	
							JCB消費 NOW	AR
RMSFE	3.24	3.65	3.04	2.81	3.90	3.23	4.24	7.84
MAFE	2.90	3.04	2.73	2.45	3.35	2.57	3.41	5.56
SMAPE	45.69	50.42	44.20	40.81	46.63	38.70	59.17	90.33