

論 文

オルタナティブデータの可能性と課題*

渡辺 努、大森 悠貴、横山 翔**

＜要旨＞

オルタナティブデータという言葉が新聞や雑誌で頻繁に目にするようになってきた。経済の様子を知るためのデータといえば、これまではGDPなどの政府統計や企業の財務諸表だった。株式市場で売り買いする投資家の意思決定の元になったのはこうしたデータだ。これらは「伝統的」データとよばれている。伝統的データの代替物（＝オルタナティブ）という意味でオルタナティブデータとよばれるものが登場してきた。例えば、スーパーのレジで蓄積されるPOSデータや、クレジットカードの購買履歴データ、スマホの位置情報データなどだ。オルタナティブデータはパンデミック前から存在していたが、一部の金融機関や投資家が使うにとどまり、認知度はさほど高くなかった。しかし、パンデミックを機に広く使われるようになり、パンデミック後も使用が拡大を続けている。本章では、最初に、オルタナティブデータの現状と先行きを展望するとともに、その使用の拡大とともに起きる可能性のあるいくつかの課題について考察する。続いて、パンデミックを題材にとり、オルタナティブデータの使用事例として、（１）スマホの位置情報データを用いた消費者行動の変化に関する分析、（２）クレジットカードの履歴データを用いた消費者・店舗間のネットワーク構造の変化に関する分析を紹介する。

JEL Classification Codes : C80、C81、C82

Keywords : オルタナティブデータ、非伝統的データ、スマホの位置情報データ、クレジットカードの購買履歴データ

* 本章の作成に際しては、家富洋、大西立顕、笹川裕人、久野遼平、水野貴之の各氏との議論が有益であった。また、株式会社ジェーシービーには、本章第4節の分析で使用したクレジットカードの購買履歴データを提供していただいた。記して感謝したい。

** 渡辺 努：東京大学大学院経済学研究科教授、株式会社ナウキャスト創業者・技術顧問。大森 悠貴：東京大学大学院経済学研究科金融教育研究センター特任研究員、株式会社ナウキャスト・データサイエンティスト。横山 翔：東京大学大学院経済学研究科金融教育研究センター特任研究員、株式会社ナウキャスト・アドバイザー、株式会社風音屋・代表取締役。

Opportunities and Challenges in the Use of Alternative Data

By Tsutomu WATANABE, Yuki OMORI and Sho YOKOYAMA

Abstract

The term “alternative data” has become a term frequently seen in newspapers and magazines. Until now, the data used to understand the current state of the economy consisted of government statistics such as GDP data and corporate data such as corporate financial statements. This data was, for example, the source of information investors used to buy and sell securities in financial market. These types of data are called “traditional” data. Now, alternatives to such traditional data called “alternative data” have emerged. Examples include scanner data from supermarket checkout counters, credit card transaction data, and location data from smartphones. While alternative data already existed prior to the pandemic, it was only used by a few financial institutions and investors, and awareness of it was not very high. However, it came into widespread use with the onset of the pandemic, and its use has continued to grow after the pandemic. This chapter first looks at the current state of and future prospects for alternative data and discusses some of the challenges that may arise as its use expands. Then, two examples of the use of alternative data are presented. The first is an analysis of changes in consumer behavior during the pandemic using smartphone location data. The second is an analysis of changes in the network structure between consumers and stores using credit card transaction data.

JEL Classification Codes: C80, C81, C82

Keywords: Alternative data, non-traditional data, smartphone location data, credit card transaction data

1. はじめに

オルタナティブデータという言葉が新聞や雑誌で頻繁に目にするようになってきた。経済の様子を知るためのデータといえば、これまでは GDP などの政府統計や企業の財務諸表だった。株式市場で売り買いする投資家の意思決定の元になったのはこうしたデータだ。

これらは「伝統的」データとよばれている。伝統的データの代替物という意味でオルタナティブデータとよばれるものが登場してきた。例えば、スーパーのレジで蓄積される POS データや、クレジットカードの購買履歴データ、スマホの位置情報データなどだ。

オルタナティブデータはパンデミック前から存在していたが、一部の金融機関や投資家が使うにとどまり、認知度はさほど高くなかった。注目されるきっかけとなったのはパンデミックだ。

最初の緊急事態宣言が発令された直後の 2020 年 4 月 23 日、政府は「月例経済報告」を発表し、景気が「極めて厳しい状況にある」との判断を示した。ここで注目したいのはこの判断がどのようにしてなされたかだ。通常であれば月例報告の判断は政府統計にもとづいてなされる。しかしこのときはそうはならなかった。理由は単純で、この時点で最新の統計は感染拡大前のものだったからだ。

それに代わるものとして政府が頼ったのがオルタナティブデータだ。4 月の月例報告では、クレジットカードデータから算出された支出の実数を用いた図を掲載し、支出の急減が起きていることの証左とした。このときの支出急減は前代未聞のスケールだったが、月例報告にクレジットカード支出の図が掲載されたのもまた前代未聞のことだった。

このようにして、パンデミックを契機に注目を集めるようになったオルタナティブデータだが、この先、オルタナティブデータの使用はどういう分野に、どれだけの速さで拡大していくのだろうか。またそこに死角はないのだろうか。本章ではそうした点について考察することとしたい。

本章の構成は以下のとおりである。第 2 節では、オルタナティブデータの現状と先行きを展望するとともに、その使用の拡大とともに起きる可能性のあるいくつかの課題について考察する。続く第 3 節と第 4 節では、パンデミックを題材に、オルタナティブデータの使用実例を紹介する。第 3 節は、スマホの位置情報データを用いた人々の消費者行動の変化に関する分析である。第 4 節は、クレジットカードの履歴データを用いた消費者-店舗間のネットワーク構造の変化に関する分析である。

2. オルタナティブデータの現状と展望

本節では、オルタナティブデータの現状を整理した上で、その先行きを展望する。さらには、今後オルタナティブデータの使用がさらに拡大する中でどのような社会課題が生じるかを検討していく。なお、本節は『入門オルタナティブデータ』第 6 章に加筆修正した

ものである。

2.1 伝統的データと非伝統的データ

経済がいまどのような状況にあるかを知るにはどうすればよいだろうか。多くの人は政府の発表する統計をみることを考えるだろう。例えば、個人消費であれば家計調査、企業の生産活動であれば鉱工業生産指数というように、用途に合わせて定番の統計が存在する。それらを統合する指標として GDP もある。これらは総称して伝統的データとよばれている。伝統的データの重要な特徴は、原データの収集から集計まですべてのプロセスを政府が取り仕切っているということだ。そのため伝統的データは政府統計とか公的統計とよばれることもある。

これに対してオルタナティブデータは、民間が作るデータだ。伝統的データとの対比という意味で、非伝統的データとよばれることもある。オルタナティブデータの原データは多くの場合、民間企業が業務を進める中で生成されるデータ、つまり業務データを借用しており、それを別の民間企業が加工・集計するという形式をとることが多い。例えば、クレジットカード会社は顧客のカード利用履歴データを業務の一環として生成しているが、それを加工・集計すれば個人消費の指標となる。

2つのデータには作成主体の違い以外にいくつかの重要な差異がある。ひとつは頻度の違いだ。伝統的データは1か月に一度作成・公表されるものが多い。GDP は四半期に一度とさらに頻度が低い。これに対してオルタナティブデータは毎日・毎週といった高頻度で作成される。また、オルタナティブデータは使用する原データのサイズが大きいのも特徴だ。伝統的データの原データの大部分はマニュアルで収集されるのに対し、オルタナティブデータはデジタル化された業務データの流用なのでデータのサイズが桁違いに大きい。この利点を活かしてオルタナティブデータは高粒度で作成される。例えば、クレカデータで消費を見る場合、日本全体の全世代のマクロの消費だけでなく、ある都市の20台の女性の消費活動というように細分化されたセルの様子をみることができる。

2.2 パンデミック下でオルタナティブデータに注目が集まったのはなぜか

冒頭で述べたように、新型コロナの感染が拡大する中でオルタナティブデータへの関心が急速に高まってきた。これには3つの理由がある。

第1の理由は、政府統計がタイムリーに作成・公表されず、政策判断に使えないという状況が起きたからだ。日本政府が行う最も重要な景気判断である「月例経済報告」の2020年4月の文書をみると、新型コロナ感染で景気が急速に悪化し「極めて厳しい状況にある」との判断がなされている。前月の月例報告はここまで踏み込んだ表現になっていなかったため政府がコロナ経済危機をはじめて公式に認めたものと言える。政府はその後、この景気判断をベースに、給付金などの財政的な支援策を決定することになる。

ここで注目したいのはこの景気判断がどのようになされたかだ。通常であれば月例報告

の判断は政府統計にもとづいてなされる。今回の場合、コロナの影響は個人消費、特にサービス支出の減少に顕著に現れたのは明らかなので、昨年4月の景気判断は家計調査の結果をもとになされるべきだった。しかしそうはならなかった。理由は単純で、月例報告が公表された4月23日の時点で最新の家計調査は2020年2月分であり、コロナの影響はほとんど反映されていなかったからだ。

家計調査に代わるものとして政府が利用したのがオルタナティブデータだ。4月の月例報告では、クレカデータから算出された支出（外食の支出やスーパーでの支出など）の実数を用いた図を掲載し、前代未聞の急減が起きていることの証左とした。月例報告にクレカ支出の図が掲載されたことはそれまでなかった。政府統計が間に合わなかったのでやむを得ずオルタナティブデータに頼ったということだろう。しかしこれを契機にオルタナティブデータへの関心が一気に高まることとなった。

オルタナティブデータに頼ったのは政府だけではない。日銀も同様だ。金融政策の材料となる景気判断にオルタナティブデータを頻繁に用いるようになった。日銀の判断を示す文書である「展望レポート」にはクレカのデータや、人の移動に関するGoogleのデータを用いた図が掲載された。マーケットは日銀がこうしたオルタナティブデータをベースに金融政策の判断を行っていることを認識し、日銀が参照するデータに強い関心を示すようになった。

伝統的データがパンデミック下で政策判断の役に立たないという状況は日本に固有ではなく、米国をはじめとする先進各国の政府や中央銀行、さらにはIMFのような国際機関でも、オルタナティブデータの利用が急速に進んだ。例えば、米国の金融政策を決める委員会であるFOMCの議事録からは、クレカ支出やATM利用件数、スマホの位置情報などの非伝統的データと伝統的データを照らし合わせながら政策判断を行っている様子が見える。

オルタナティブデータに関心が集まるようになった第2の理由は、政府統計が単に間に合わないだけでなく、コロナ禍で作成自体が難しいという事情があった。先述のとおり政府統計はマニュアルで作成されるものが多い。データ提供者から対面で情報を聴取するのも少なくない。日本政府の海戦術の限界は今回のコロナ危機で随所に現れたが、統計作成の面でもそれが露呈されたということだ。感染者の急増で医療サービスの提供ができなくなるのが医療崩壊だが、政府の行う統計サービスについても、「統計崩壊」に近い状況に陥った。

もちろんコロナが収束し平穏な日常が戻ればオルタナティブデータに頼る必要性は薄れるのかもしれない。しかし、伝統的であれ非伝統的であれ、統計とはそもそも危機に対応するための道具であり、平時のものではない。平時は統計などなくても誰も困らないし、実際、統計を熱心に見る人もいない。しかしいったん危機が起きれば危機の実像を知るのに統計は不可欠だ。危機時に伝統的データがその役割を果たせないのだからオルタナティブデータで補うしかない。昨今は、地球温暖化に伴う異常気象もあって災害に見舞われる

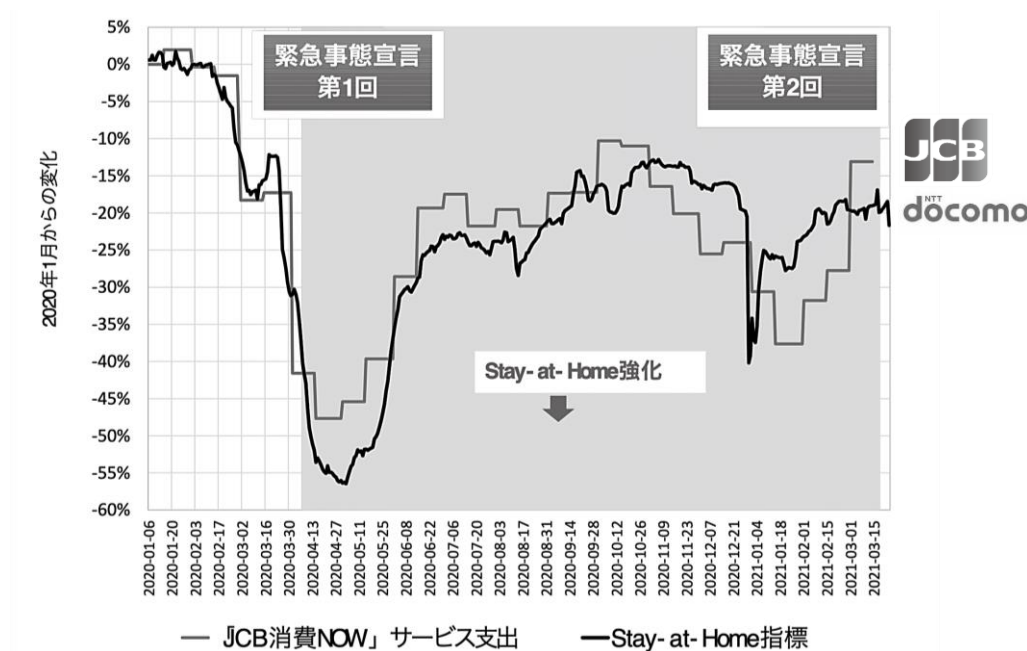
頻度が年々増している。オルタナティブデータは、それらも含めた次の危機に備えるために社会がもつべきインフラである。

第3の理由は、コロナ禍で人の移動にどのような変化が生じているかをモニターする必要性が高まったからだ。感染拡大を抑えるには人の移動を抑えなければならない。しかし人の移動を抑えると経済が停滞する。健康被害と経済被害のバランスを見る上で人の移動をモニターすることは極めて重要だ。しかし政府統計には消費のデータはあっても、それと比較できるタイムスケールでの人の移動のデータは存在しない。

そうした中で筆者（渡辺努）は水野貴之氏（国立情報学研究所）、大西立顕氏（立教大学人工知能科学研究科）と共同でスマホの位置情報データを用いてステイホーム指標を作成し広く公開した。同様の指標は内外の多くの研究者によって作成されている。

図1は、このようにして算出されたステイホーム指標（頻度は日次）の値をサービス支出（頻度は半月）と比較している。サービス支出はJCB社とナウキャスト社がクレカの利用履歴から算出した「JCB消費NOW」を用いている。昨年4月の最初の緊急事態宣言の際には外出はコロナ前との対比で55%減少した。一方サービス支出もコロナ前との対比で50%近く減少した。両者は減少のタイミングとマグニチュードの両方で極めてよく似た動きをしている。また、5月末に緊急事態宣言が解除されると外出も支出もほぼ同じタイミングで同程度の回復を示した。2つの非伝統的データはドコモとJCBという全く異なる業態の異なる企業の業務データを流用したものであるが、どちらも精度の高いセンサーとして機能していることがわかる。

図1：東京都のステイホーム指標とサービス支出



2.3 オルタナティブデータへの経済学者の関与

さて、このようにオルタナティブデータへの関心はパンデミック下で急速に高まったが、オルタナティブデータ自体はパンデミック前から存在し、金融関係者などの間で既に注目を集めていた。

特筆すべきは、オルタナティブデータのこれまでの進展には経済学者が深く関与してきたということだ。筆者の知る限り、オルタナティブデータの最初の事例はミクロ経済学のベストセラー教科書で有名な Hal Varian (カリフォルニア大から 2010 年に Google にチーフエコノミストとして転籍) によるものである。ヴァリアンはネット上にあふれる価格情報を Google の検索技術を使って収集し、Google Price Index (略称 GPI) を開発し 2010 年に発表した。物価指数は数ある経済指標の中でも計測の難しさという点で別格であり、各国統計局の担当者を悩ませてきた。ヴァリアンの手法を使えば各国の物価を高精度・高頻度で計測でき、しかも国際比較も容易だ。筆者は、GPI の構想を初めて聞いたとき、各国の CPI にとってかわる存在になるに違いないと思った。しかし実際には各国の中央銀行が難色を示し GPI 構想が陽の目を見ることはなかった。

これ以外にもオルタナティブデータを用いたプロジェクトは数多くある。筆者（渡辺努）も、GPI に刺激を受け、POS データを用いて物価指数を作成・公開するというプロジェクト（東大日次物価指数プロジェクト）を渡辺広太氏（キャノングローバル戦略研究所と東京大学）と 2012 年に始め、2015 年には株式会社ナウキャストを創業し事業化を始めたという経験がある。

経済学者は実証的な研究の過程でビッグデータを扱うので経済学者がオルタナティブデータの作成に関与することはさほど驚くことではない。しかし経済学者主導、逆に言えば、実務家主導ではないかたちでオルタナティブデータの作成が行われてきたことにはそれなりの必然性がある。

伝統的であれ非伝統的であれ、統計は多くの人に使ってもらってこそ存在意義がある。政府統計の場合は、政府のお墨付きがあるので、多くの人が安心して使う。これに対してオルタナティブデータは、民間企業がいかにも良い指標を作ったとしても、作成方法を巡ってどうしても疑念が残ってしまう。統計作成のレシピを公開するなど工夫の余地はあるがそれでも限界がある。ここが経済学者の貢献できるところなのかもしれない。

例えば、物価指数であれば、効用理論に基づく指数理論の枠組みがあり、そこに修正を加え改善させることは研究テーマとして重要で、今なお多くの論文が書かれている。誰かがこれまでにない作成方法を提案しオルタナティブデータとして公開したときに、それを学術の観点から他の研究者が評価できる体制が整っている。学術的な知見に基づくお墨付きを与えるという経済学者の役割は重要だ。

やや脱線するが、筆者（渡辺努）がナウキャスト社を創業したときに海外の友人から Irving Fisher を目指すのかと冷やかされたことがある。フィッシャーはケインズと同時代を生きた偉大な理論家だ。物価理論でもフィッシャー指数など多大な功績がある。筆者が彼

を目指すことなど到底あり得ない。どういうことかと訝しがっていると、フィッシャーは Index Number Institute という会社を 1923 年に立ち上げ、自らが開発したフィッシャー指数を顧客に販売することを始めたと教えられた。その会社は成功し大きな利益を上げたと伝えられている（ただし株投資の失敗で最終的に利益は消滅）。経済学者が研究成果を活かして新たな経済指標を作り事業展開するのは、実ははるか昔から繰り返されてきたことなのかもしれない。

2.4 オルタナティブデータの今後：統計民営化に向けて

次に、以上の議論を踏まえ、伝統的データと非伝統的データ（オルタナティブデータ）のこれからを考えてみたい。

昭和あるいはそれ以前の時代であれば、統計のもとになるデータはどこにも存在せず、政府が調査員を現場に派遣しミクロの情報を足で収集する以外に手はなかった。集めたデータの加工のノウハウも政府が握っていた。あらゆる意味で政府が民間を凌駕していた。しかし ICT（情報通信技術）の進展に伴って、政府と民間の立場が逆転している。統計のもとになるデータを握っているのは誰かと言えば民間企業だ。そのデータの加工作業も民間企業の方がはるかに長けている。

こうした状況の変化を踏まえれば、鉄道サービスの提供主体が政府から民間企業へと変わっていったのと同様に、統計サービスの提供主体も政府から民間企業に切り替える——つまり伝統的データから非伝統的データ（オルタナティブデータ）へと移行するのは時代の流れだろう。筆者はこれを「統計の民営化」とよんでいる。コロナ禍は民営化を加速させたが、このトレンドは今後も続くと思っておくべきだろう。

ただし、伝統的データから非伝統的データへというのが進むべき方向だとしても、そこに一足飛びに行くというのは非現実的である。民営化という方向感を持ちつつ段階的に統計作成の仕組みを変えていくべきだ。

最初のステップはデータ収集業務の民間企業へのアウトソーシングである。民間企業が自らのビジネスを遂行するために収集した情報を政府の統計作成にも活用することにより効率化をはかることができる。

データ収集業務のアウトソーシングを進めるに当たっては次の 2 つの論点をクリアにする必要がある。第 1 は、政府が情報を民間企業からいかにして取得するかである。具体的には対価をどうするかである。無償か有償か、有償にするならいくらが適切なのか。例えば、物価統計については、欧州や米国で民間流通企業のデータを使うということが始まっているが、これらの国も試行錯誤という段階であり、有償か無償かについても各国で区々となっている。

データ収集業務のアウトソーシングに関する第 2 の論点はサンプルバイアスである。政府統計は、物価や消費などに関する情報を全国民から偏りなく集めるようデザインされている。これに対して民間のビッグデータは民間企業が自らのビジネスを展開した結果とし

て自然発生的に生まれたものであり、事前にデザインされたものではないという意味でオーガニック・データともよばれている。その結果、例えばクレジットカードの購買履歴データには、そのカードを使う人の購買情報は入っていても使わない人の情報は含まれず、全国民から遍く情報を集めるという設計にはなっていない。

民間企業のビジネスの結果として集まったデータを「流用」する以上、サンプルバイアスは避けて通れない課題であり、それに正面から向き合う覚悟が必要である。具体的には、バイアスの度合いの評価とそれを最小化するための研究を政府や大学等の研究機関が中心となって進める必要がある。

2.5 複数の指標で政府の監視が可能に

統計民営化に向けての次のステップでは、統計作成そのものを民間で行うことになる。その際の重要な論点は、民間が統計サービスを提供すると、複数の指標が乱立し、利用者が混乱する可能性があるという点である。

例えば物価について政府も含めて複数の主体が物価指標を公表したとして、全ての指標が同じ方向を指していれば問題ないが、ある指標は物価上昇、別な指標は物価下落となったとき利用者はどの指標を信用すればよいかわからなくなり、混乱が生じるかもしれない。

議論を整理すると、計測精度が極端に低い統計が流通するという問題とそれ以外の問題は分けて考えた方がよい。計測精度が低くノイズの大きい指標は、利用者によって選別され、やがては淘汰されるであろう。問題は、そうした悪質な指標が淘汰されたとしても、複数の指標が別な方向を指し示す可能性がなお残るということである。計測誤差が全くない指標というのは存在しないからだ。どの指標にも多少なりともノイズが入り込んでいるので、利用者は各指標の変動を、意味のある動き（シグナル）とノイズに分けるという、容易でない作業を迫られることになる。

この手間をどうとらえるかだが、ひとつには、エコノミストなど民間の専門家がシグナルとノイズの仕分けを行うということが考えられる。専門家にとっても手腕を競う場が増えるという意味でビジネスチャンスであろう。

悪名高き大本営発表は、遠隔地の戦況という、入手の難しい情報を政府が独占的に取得・公表していたことに端を発する。今であれば人工衛星からの画像を人工知能（AI）で処理して遠隔地の戦況を伝える民間企業がたくさん出て来るであろう。軍部によるバイアスのかかった発信で正確な情勢判断ができないという事態に陥ることはなかっただろう。

経済統計もこれと同じだ。政府によって作成された統計には、意図せざるノイズと、意図的なバイアスが含まれる可能性があるという認識すべきだ。これらの誤差に経済活動が振り回されるのを避けるためにも、政府だけでなく民間も統計作成に関与し、利用者が複数の指標を比較できるようにすることが望ましい。

筆者（渡辺努）が東大指数の公開を始めた頃、多くのご意見・ご批判をいただいたが、最も多かったのは「複数の指標を作るのは無駄であり、混乱のもとだ」という意見だ。政

府の作る伝統的データがあるにもかかわらず民間がわざわざオルタナティブデータを作るのはリソースの無駄という趣旨だ。また、伝統的データである総務省の物価指数と東大指数が異なる動きをすると、人々はどちらを見てよいかわからなくなり、社会が混乱するというご批判も多かった。

しかしパンデミックの経験はこれらの批判が当たらないことを私たちに教えてくれた。確かに同じ目的の指標が複数あることは平時には無駄かもしれない。しかし今回のような危機の局面では政府統計の入手が遅れたり、作れなくなったりする。バックアップとしてオルタナティブデータを用意しておくことは不可欠だ。

また、複数の指標は社会を混乱させるというご意見は、伝統的データは無謬との前提のもとでなされていたように思う。しかし政府統計が無謬という考え方にはリスクがある。そのことに気づかせてくれたのは2019年初の賃金統計の不正問題だ。不適切な集計方法に誰も気づかぬままに間違った数字が公表される状況が長く続き、景気判断や雇用保険の給付等に甚大な影響を及ぼした。政府統計は決して無謬ではない。

賃金統計の単純なミスに誰も気づかなかったのはなぜなのか。この指標と比較できる別な指標が存在しなかったからだ。この不正事例から学ぶべき最も重要な教訓はこれだ。あ のときオルタナティブデータを用いた賃金指標が複数存在していれば、政府統計との乖離に多くの人が気づき、傷の浅いうちに原因究明できたはずだ。オルタナティブデータには、政府統計の監視という重要な役割がある。

2.6 統計インフラの再構築に向けて

精度の高い統計は、国民が安心して暮らせる社会を支える不可欠のインフラだ。2019年の統計不正に続いて、今回のパンデミックで政府統計の機能不全が起きたことを踏まえると、統計インフラの修復が急務という点について異論はないだろう。

しかし具体的にどこをどのように修復するかとなると意見が分かれている。例えば、政府の統計作成部署の人員をこれまで削りすぎたのが問題だから、人員を元に戻すべきとの意見も少なくない。しかし厳しい財政状況を踏まえれば、統計部署の人員増は選択肢としてあり得ない。

また、統計作成の厳格性を世界に誇っていたかつての日本を取り戻すべきとの声も少なくない。しかし統計作成の最先端はICT（情報通信技術）の急速な進歩に伴い大きく変容しつつある。昭和の、人海戦術型の統計制度に戻すという選択肢はないと考えるべきだ。

統計インフラ再建のカギを握るのは、民間のもつ大量のデジタルデータと優れたノウハウであり、それをベースに急成長を遂げたオルタナティブデータの活用だ。これにより、低コストで、しかもこれまでより格段に精度の高い統計を国民に届ける仕組みを整備できる。統計インフラの再建に向けて官民の密接な連携が不可欠だ。

3. スマホ位置情報で捉える消費者行動の変化

本節と次節では、オルタナティブデータの利用事例として、パンデミック下での消費者行動に関する分析例を紹介する。まず最初に、本節ではパンデミック下での人々の行動変容についてスマホの位置情報データを用いて考察する。具体的には、行動変容はどの程度生じたのか、どのようなメカニズムで生じたのか、感染拡大を防ぐという観点から行動変容はどの程度有効だったのか、行動変容は経済活動にどのような影響を及ぼしたのか、といった点について考察する。なお、本節の分析は、Watanabe and Yabu (2021a, 2021b) と Mizuno et al. (2021) で行われたものである。分析の詳細については原論文を参照されたい。

3.1 感染と行動変容

新型コロナウイルスの感染拡大と縮小は人々の外出行動と密接に関係している。各国のデータを用いたこれまでの研究で明らかになったように、感染が拡大すると政府は人々に外出抑制を促す様々な措置をとる一方、人々は感染を恐れ外出を控える。この2つが相まって感染拡大に歯止めがかかる。多くの国でこうした行動変容が起きた。日本も例外ではなく、人々は感染の3つの波に対して行動を変容させ、それが感染拡大の抑制に一定の効果をあげてきた。

しかし日本が他国と異なるのはその過程で政府の果たす役割が相対的に小さかったという点である。中国や米国、イタリア等の欧州諸国では、感染が拡大すると政府がロックダウンを宣言し、国民に対して外出禁止の「命令」を出した。これに対して日本の緊急事態宣言は、中国や欧米のロックダウンと異なり、法的な拘束力をもたず、緊急事態宣言中に外出しても罰金や逮捕などの懲罰が課されることはなかった。日本政府は国民に対し外出を抑制するよう口頭で「要請」したに過ぎない。日本国民はそうした要請に応じて外出を抑制した。この意味で日本の緊急事態宣言は「自発的なロックダウン」とみることができる。

日本政府の施策の特異性は、オックスフォード大学の公表している Stringency and policy index で確認できる。同大学の Blavatnik School of Government は各国政府の感染対策の強さを指標化し、Stringency and policy index として公表している。日本で緊急事態宣言が発動中の4月末の時点でみて、日本の値は47.22であり、フランス(87.96)、米国(72.69)、英国(75.93)、ドイツ(76.85)、イタリア(93.52)、カナダ(72.69)と比べて、stringencyの低さが際立っている。日本はスウェーデン(46.30)と同水準である。スウェーデン政府は集団免疫の獲得を目指し感染拡大を放置するという方針を初期の頃にとったが、日本政府の施策はそのスウェーデンと同じくらい緩かったということである。

コロナ感染拡大に対して法的拘束力のない施策しか行っていなかった日本で国民の行動変容が起きたという事実は、日本以外の国での行動変容の原因が(法的拘束力のある)ロックダウンではなかった可能性を示唆している。

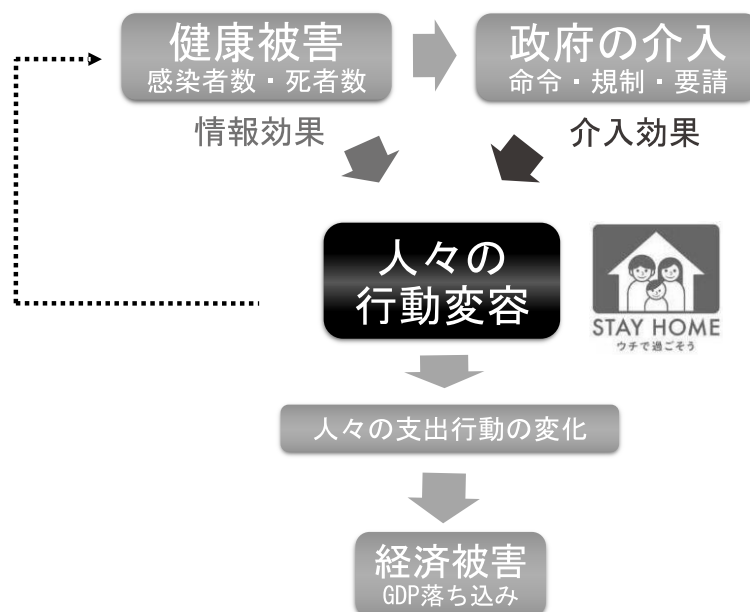
3.2 行動変容の2つのチャネル

行動変容に関するこれまでの研究は次の2つのチャネルに注目してきた。第1のチャネルは、政府の命令や要請が人々の行動を変化させるというものであり、「介入(intervention)効果」とよばれている。武漢のロックダウンは最初の事例であるが、中国や米国、欧州各国では政府が国民の移動の自由を奪う法的な権力をもつ。これに対して日本の緊急事態宣言は先述のとおり法的拘束力のない「要請」である。

第2のチャネルは「情報効果」である。人々は感染に関する情報（新規感染者数や死者数など）を取得し、それを踏まえて自ら進んで行動を変容させるというものである。国民がよりどころとする情報の発信源としては様々なものが考えられる。しかし日本を含む先進各国では、感染者のプライバシー保護の観点から、感染の実態の詳細が国民に開示されることは稀である。そのため政府のもつ情報量が多く（少なくとも多くの国民はそうように信じている）、政府から発信される情報が通常の場合よりも大きな影響力をもつと考えられる。

図2はコロナ危機で経済被害が生じる仕組みを示している。新規感染者数や死者数の増加といった健康被害が生じるとそれに対して政府が介入的な施策を行う。これは人々の外出抑制など行動変容を促す効果がある。これが介入効果である。一方、人々は健康被害を目の当たりにして、感染に対する恐怖心をもち、行動を（自発的に）変化させるということもある。これが情報効果である。この2つの効果によって外出抑制などの行動変容が生じると、飲食店などでの人々の支出が減少する。この結果、GDPが落ち込むなどの経済被害が起きる。

図2：健康被害と経済被害



3.3 介入効果の識別方法

人々の行動変容が介入効果と情報効果で生じているとして、どちらがどの程度の役割を果たしているのだろうか。それを知るには2つの効果を分離する必要がある。どうすれば分離できるだろうか。Watanabe and Yabu (2021a) は、こうした問題意識から、2つの効果を分離する手法を提案し、日本のデータを用いて実際に分離を行った。

具体的には、最も重要な行動変容であるステイホームに注目し、スマホの位置情報データを用いて、人々のステイホームの度合いを示す指標（外出者数×外出時間がコロナ前の2020年1月との対比で何%低下したか）を、県別に日次で計測し、パネルデータを構築した。その上で、政府の介入施策（具体的には、緊急事態宣言と学校閉鎖）の開始と終了のタイミングが各県で異なることを利用して2つの効果の識別を行った。

例えば、埼玉県に緊急事態宣言が出たのは2020年4月7日であったが、その時点では、その隣の群馬県には発出されなかった。群馬県に緊急事態宣言が発出されたのは4月16日である。したがって、7日から15日までの間は群馬県で介入効果はない。しかし群馬県の人々は埼玉で緊急事態が発令されたことは知っているので情報効果は存在し、その分だけ外出が減少する。これに対して埼玉の人々には介入効果と情報効果の両方の要因から外出抑制が働く。したがって、埼玉と群馬の人々が感染について同じ情報を持ち両県の情報効果に差がないという前提の下では、両県のステイホーム指標の差をみることにより介入効果を抽出することができる。

緊急事態宣言は発動のタイミングが県の間でずれただけでなく、解除も3回に分けて行われたので、そうしたタイミングのずれを利用して介入効果を識別できる。同様に学校閉鎖についても、学校閉鎖の解除のタイミングが県の間で大きなばらつきがあるので、これを利用して介入効果を識別できる。

3.4 スマホの位置情報データ

新型コロナウイルスは人と人との接触によって感染する。そのためコロナ禍の分析では人の移動に関する情報が多用された。具体的には、スマホの位置情報データを用いて、①人の移動が感染拡大にどのように影響したのかという、移動→感染拡大の方向の因果と、②感染拡大に伴って人々がステイホームするなど、移動をどの程度抑えたかという、感染拡大→移動の方向の因果の両方に関する分析が内外で行われてきた。各国の研究事例については Watanabe and Yabu (2021a, 2021b) で紹介しているので参照されたい。

Mizuno et al. (2021) は、スマホの位置情報データを用いて人々の外出抑制を測定する手法を提案している。Mizuno et al. (2021) では、位置情報データとして、株式会社ドコモ・インサイトマーケティングが提供する「モバイル空間統計」を用いる (<https://mobaku.jp/>)。モバイル空間統計は、約7800万台のドコモの携帯電話がどこにあるかを10分ごとに記録したデータである。具体的には、各地にある携帯電話の基地局は、その地域内にどの携帯電話があるかを把握している。そのデータをもとに、日本国内を500m×500mのメッシュに

区切った上で、あるメッシュにおいて、ある時刻（10分刻み）に、そのメッシュ内に何台の携帯電話があるか、それらの携帯電話の持ち主の年齢や性別、居住市町村を記録したデータが作成・公開されている。ただし、メッシュ内の台数が非常に少ないメッシュについては、個人情報保護の観点から、除外されている。

このデータを利用して、次の2つのステップでステイホーム指標を作成する。第1のステップは住宅地の検出である。あるメッシュについて、深夜0時から5時までの時間帯に平均的に何人いるかをカウントし、それをそのメッシュの夜間人口とする。同様に、朝9時から17時までの時間帯に何人いるかをカウントし、それをそのメッシュの昼間人口とする。住宅地は、昼間人口が夜間人口よりも多ければビジネスや商業の地域と考えることができ、その逆に昼間人口が夜間人口よりも少なければ住宅地である。Mizuno et al. (2021) では、コロナ感染が拡大する前の2020年1月の時点で、昼間人口／夜間人口 <0.8 を満たすメッシュを住宅地と定義している。なお、閾値を0.9または0.7としても計算結果は定性的に同じということが確認されている。

第2のステップは、外出率の計算である。第1のステップで住宅地と判定されたメッシュについて、ある日のある時刻（10分刻み）における夜間人口と昼間人口をカウントし、夜間人口から昼間人口を差し引くことにより外出者数を計算する。それを時刻とメッシュについて集計することにより、県別・日次の外出者数を計算する。

最後に、コロナ感染拡大前の2020年1月（2020年1月6日から1月31日まで）の外出者数を平常時の外出者数とし、県別日次の外出者数が平常時の外出者数から何パーセント乖離しているかを算出する。外出率の乖離率に-1を掛けたものがステイホーム指標である。

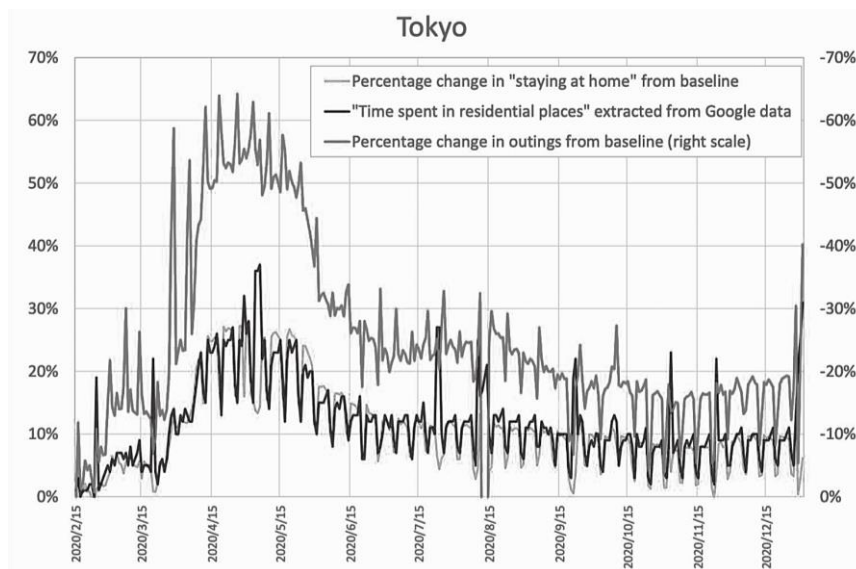
Mizuno et al. (2021) の手法について2点コメントしておきたい。第1に、使用しているデータは基地局のデータであり、GPSデータではないという点である。駅や繁華街などにおける人の混み具合を計測したいのであればGPSデータは適している。なぜなら、そうした繁華街では人々はGoogle Mapなどのアプリを稼働させ、そこでGPS接続がなされるからだ。しかしGPSデータは人々の在宅中はあまり役に立たない。在宅中はGPS接続があまり頻繁ではないからだ（例えば、自宅でGoogle Mapを使う頻度は低い）。それに対してスマホは在宅中も基地局とは頻繁に交信しており、基地局データは在宅中も着実に蓄積されている。本章の目的はステイホームの度合いを測ることであるから、GPSデータではなく、基地局データが分析に適している。

第2に、Mizuno et al. (2021) のデータをGoogleが提供している位置情報データである「コミュニティ・モビリティ・レポート (<https://www.google.com/covid19/mobility/>) と比較する。Googleは様々な場所のカテゴリについて何台の端末があるかという情報を県別に提供しており、場所のカテゴリのひとつに「住宅地（“time spent in residential places”）」がある。これは概念的にはMizuno et al. (2021) のステイホーム指標に近い。

図3は、Googleデータにより作成されたステイホームの度合いと、ドコモのデータで作成されたステイホームの度合いを比較したものである。対象地域は東京である。ここでは、

Google データの定義に合わせるため、ドコモのデータを用いて推計した「日中家にいる人口」がコロナ前と比べてどう変化したかを新たに計算した。両者は非常によく似た動きをしていることが確認できる。この結果は、ドコモと Google というデータの出所の違いはそれほど大きくないことを示唆している。ただし、この2つの指標は、Mizuno et al. (2021) のステイホーム指標とは概念的に異なる点には注意が必要である。Mizuno et al. (2021) の指標は、「日中外出している人口」がコロナ前と比べてどう変化したかの変化率（に-1を掛けたもの）である。図3にはこれも含め3つの指標を示してあるが、Mizuno et al. (2021) のステイホーム指標は Google データとほぼ同じ動きをしているものの完全には一致していないことがわかる。

図3：Google データとの比較（東京都）



3.5 介入効果の日米比較

Watanabe and Yabu (2021a) は、Mizuno et al. (2021) のステイホーム指標を用いて、日本政府の介入施策（緊急事態宣言と学校閉鎖）の効果を計測した。その結果、2020年4月～5月の緊急事態宣言は、人々の外出をコロナ前に比べ8.5%抑制させる効果があったとの結果を得た。一方、Goolsbee and Syverson (2021) は、米国の位置情報データを用いて、米国のロックダウンに伴う介入効果は7.6%であったと報告している。法的拘束力の有無という点で日米の措置は大きく異なったにもかかわらず介入効果の推計値が概ね同じオーダーであったことは非常に興味深い。ただし、Watanabe and Yabu (2021a) は、日本の推計結果を、法的拘束力のない措置であるにもかかわらずそれなりに大きな効果があったと読んだ。これに対して Goolsbee and Syverson (2021) は厳しい法的拘束力をもつ措置をとったにもかかわらず効果は限定的だったと評価した。数値は同じでもその評価は対照的であった。

政府の強制措置の限界を示す結果は他の研究でも指摘されている。例えば、Rojas et al. (2020) は、学校閉鎖のタイミングが州の間で異なることに注目し、ある時点で学校閉鎖が行われている州とそうでない州で失業保険申請件数に差があるか否かを調べた結果、統計的に有意な差はないことを示した。米国民の行動変容は（法的拘束力のある）ロックダウンに起因するものでない可能性は Chetty et al. (2020) でも指摘されている。これらの研究は、米国の行動変容は、政府によって行われた法的拘束力のある施策によってもたらされたものではなく、感染に対する米国民の自発的な対応の結果と指摘している。

米国以外のデータを用いた研究としては、Sheridan et al. (2020) が、感染拡大のために政府が外出や経済活動に法的な介入を行ったデンマークと、それらの介入のなかったスウェーデンを比較して、銀行の取引データに観察される経済活動の落ち込みは大きく異ならず、したがって政府による介入は経済活動停滞の主因ではないと主張している。

3.6 情報効果の推計

スマホの位置情報データを用いた研究では、日米をはじめとする各国において、人々の外出がパンデミック前に比べて半減またはそれ以上減ったと報告されている。しかしそのうち介入効果で説明できるのは7～8%程度に過ぎない。このことは人々の外出を抑制した要因が他にあったことを示唆している。

図4は東京都のステイホーム指標の要因分解を行った結果を示している。東京都のステイホーム指標は緊急事態宣言中の5月1日に55%とピークをつけた。コロナ前に比べ外出が半減したことを意味している。このうち緊急事態宣言と学校閉鎖という2つの介入政策の寄与を推計すると11%ポイントである。学校閉鎖を加味することで政府による介入施策の寄与は多少高まっているが、それでもなお介入で説明できる部分は限定的である。では外出抑制の主因は何だったのか。図からわかるように、ステイホーム指標の変動の多くを説明するのは情報効果である。5月1日の値のうち情報効果で説明できる部分は43%ポイントである。

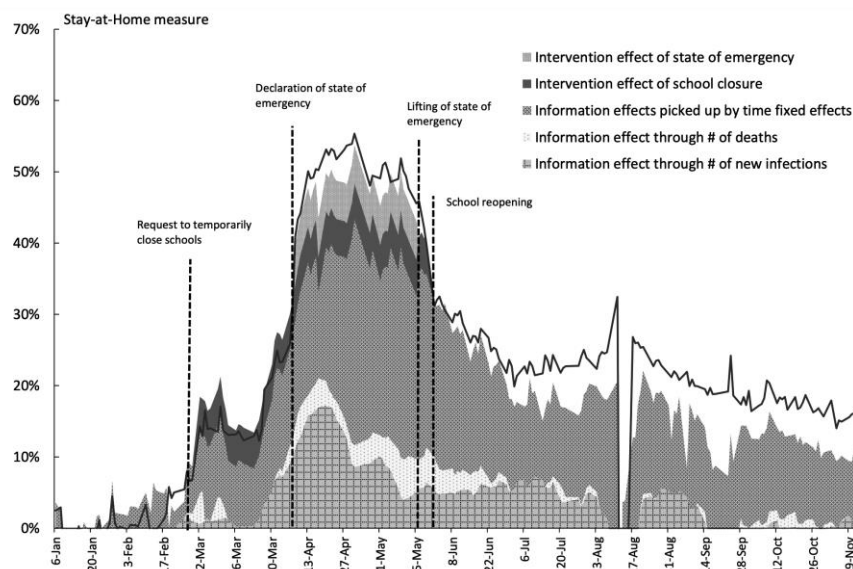
ここでは情報効果を2つの方法で捉えている。第1は、人々が新規感染者数の増加などの「情報」に接し外出を抑制する効果を推計している。具体的には、各県の新規感染者数や死亡者数の増加に反応してその県のステイホーム指標が上昇する度合いを計測している。5月1日のステイホーム指標の値のうち新規感染者数の増加に対する反応の寄与は9%ポイント、また死亡者数の増加に対する反応の寄与は3%ポイントである。この2つを足すと介入効果（緊急事態宣言と学校閉鎖の和）に匹敵する。

情報効果を捉える第2の方法は時間ダミーである。具体的には、全国を7つの地域に分け、各地域ごとの時間ダミーを作成し、それによって捉えられる時間効果を情報効果とみなしている。これは、ある地域に属する各県の住民が取得する感染情報（感染者数や死亡者数以外の感染情報）は同一であり、したがって同一の反応がステイホーム指標に現れるという仮定にもとづくものである。5月1日の値のうち時間効果として捉えられる情報効

果は 31%ポイントである。以上をまとめると、外出抑制のうち介入効果の寄与は約 2 割、情報効果の寄与は 8 割であり、行動変容の支配的なチャネルは情報効果であったことがわかる。

情報効果を通じた行動変容はどのようなメカニズムで生じるのだろうか。カギを握るのは感染に対する人々の恐怖心である。Eichenbaum et al. (2020) や Cochrane (2020) など消費者や労働者の意思決定を SIR モデルに組み込んだモデルでは、人々は、消費や労働の場での他人との接触によって感染することを恐れ、消費活動や生産活動を控えると想定されている。具体的には、Aum et al. (2020) の設定では、感染の恐怖心が不効用を生む。その下で人々は最適化問題を解き、その結果、外出の抑制や在宅勤務が選択される。また、Brotherhood et al. (2020) は、人々の時間の使い方として、職場での労働、在宅労働、外出を伴う余暇、在宅の余暇という 4 種類を想定し、人々は効用最大化の結果としてこの 4 種類への時間配分を決めると考える。この設定の下では、仮に政府の介入がない場合でも、人々は（特に重症化・死亡リスクの高い中高年層は）感染に対する恐怖から在宅労働や在宅余暇を自ら選択する。図 4 で推計されている情報効果は、Brotherhood et al. (2020) などのモデルが表現している「恐怖心にもとづく自発的な外出抑制」を捉えていると解釈できる。

図 4：東京都のステイホーム指標の要因分解



3.7 介入・情報効果の年齢別推計

新型コロナウイルスの最も重要な特徴のひとつは、感染の重症化率や死亡率が年齢によって大きく異なることである。厚生労働省の公表した数値によれば、感染者のうちで重症化する割合は、70 代では 20 代の 156 倍であり、極めて大きな格差がある (MHLW 2020)。日本以外でも重症化率や死亡率が年齢によって大きく異なるとの結果が報告されている。

Acemoglu et al. (2020) は、この異質性に注目して、すべての年齢層に対して一律の外出

抑制を求める（uniform lockdown）のではなく、高齢層に照準を合わせて外出抑制を求める（targeted lockdown）方が、経済活動への影響を抑えつつ感染を効率よく抑制できるという意味でより望ましいということを明らかにした。同様の主張は医学研究者の論文でも展開されている。こうした研究を踏まえ、Watanabe and Yabu（2021b）は、緊急事態宣言や学校閉鎖など日本政府による介入政策が、高齢者の外出抑制に強く効くという意味で効率的なものであったか否かの検証を行っている。

Watanabe and Yabu（2021b）のもうひとつの目的は「恐怖効果」の推計である。恐怖効果とは、人々が感染情報（例えば、新規感染者数の増加など）に接した結果、感染の恐怖心が強まり、それを踏まえて外出抑制を行うことであり、情報効果の一部とみることができる。Brotherhood et al.（2020）などの理論研究が指摘しているように、恐怖効果の重要な含意は年齢依存性である。つまり、高齢層と若年層では、重症化・死亡リスクが異なるので、同じ感染情報に接したとしても、彼らの抱く恐怖心のレベルが異なり、外出抑制の度合いも異なる。

恐怖効果はどうすれば識別できるだろうか。先述のとおり、Watanabe and Yabu（2021a）では、ステイホーム指標の県間の差異を用いて介入効果の識別を行った。これに対して、Watanabe and Yabu（2021b）は、県別・日次のステイホーム指標をさらに性別・年齢別に細分化した上で、ステイホーム指標の年齢グループ間の差異を用いて恐怖効果の識別を行う。例えば、ある日において東京都の新規感染者数が大きく増えたとして、そこに住む20代の人たちと70代の人たちのステイホーム指標がそれにどう反応したかを比較する。これらの人々の居住地域は同一であり（この場合は東京都）、したがって感染情報も同一であるが（この場合は新規感染者数）、重症化・死亡リスクが異なるので、恐怖心のレベルが異なる。20代と70代のステイホーム指標の差を見ることで恐怖効果を識別できる。

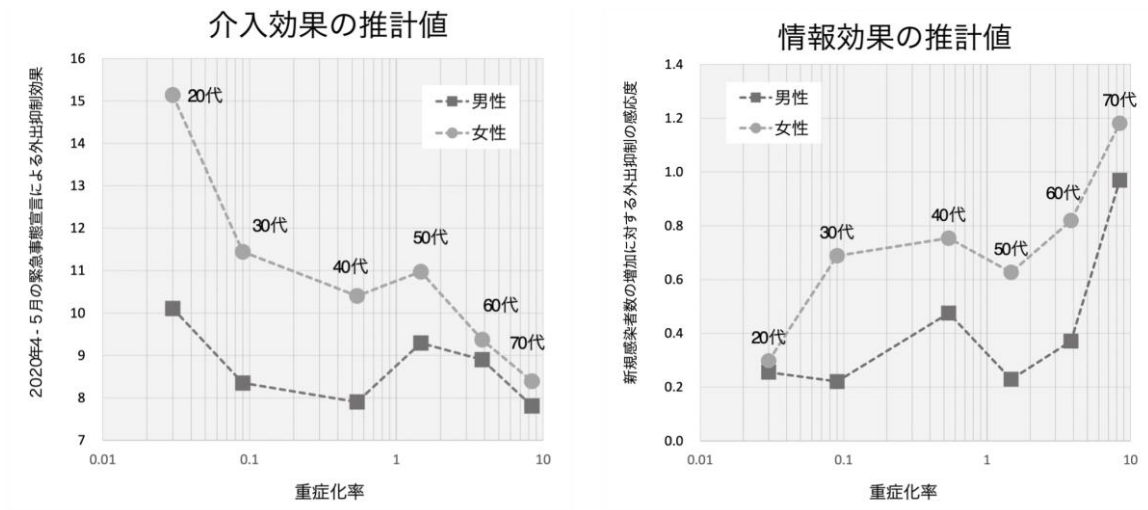
図5は、このようにして推計した介入効果と情報効果の年齢別の値を示している。図の横軸は各年代の重症化率であり、縦軸は各年代の介入効果と情報効果の推計値である。次のことが読み取れる。第1に、2020年4月から5月にかけて発出された緊急事態宣言の介入効果の年齢プロファイルを見ると（図3の左側）、重症化・死亡リスクの高い高年齢層の外出抑制の度合いは若年層との対比で低かったことがわかる。この結果は、緊急事態宣言は、Acemoglu et al.（2020）の提唱する targeted lockdown とは真逆の効果をもったことを示しており、健康被害と経済被害の両方を効率的に抑えるという観点からは望ましくなかった可能性がある。

右下がりの年齢プロファイルの背後では、対面型サービスの供給サイドの変化が重要な役割を果たした可能性がある。すなわち、緊急事態宣言の対象地域では、レストランや劇場、小売店など繁華街にある商業施設の営業時間短縮や営業停止の要請が出されたが、これら繁華街の商業施設の主な利用者は20代から50代の世代である。対象地域に住むこの世代の利用者は、非対象地域に住む同世代との対比では、商業施設の営業時間短縮や営業停止という供給サイドで異なっており、これが外出抑制につながったと考えられる。これ

に対して 60 代超の高年齢層は、元々こうした商業施設の利用が少なく、緊急事態宣言の供給サイドの影響をあまり受けなかったと考えられる。

次に、情報効果の年齢プロファイルを見ると（図 3 の右側）、介入効果と異なり、年齢とともにステイホームの度合いが高まる傾向が週末・休日にみられる。人々は、感染に関する情報を日々更新し、それに自らの重症化・死亡リスクを加味して外出抑制の意思決定を行ったことを示している。Brotherhood et al. (2020) は、理論モデルを用いて、高齢者は最適化の結果、自発的に外出を抑制することを示しているがここでの結果はそれと整合的である。また、Acemoglu et al. (2020) の提唱する targeted lockdown は、緊急事態宣言に代表される政府の介入によっては起きなかったが、人々の自発的な行動の結果、起きたと言える。

図 5：介入効果と情報効果の年齢別推計値



4. クレジットカードデータで捉える消費者 - 店舗のネットワーク構造の変化

4.1 マクロの三面等価とミクロの三面等価

現時点で最も広く活用されているオルタナティブデータはおそらくクレジットカードの購買履歴データであろう。国内だけでなく、米国など海外も同じ事情だ。消費という景気循環の最重要の指標を精度高く捉えることができるのがその理由だ。

しかしクレジットカードの履歴データの魅力は単に消費指標が作れるということだけにとどまらない。クレジットカードの履歴データには、どの消費者がどの店でいくら使ったが記載されている。消費の動向を知るのが主目的であれば、「どの店で」という情報は不要だ。例えば、ナウキャストの「JCB 消費 NOW」はそうやって作られている。

一方、「どの消費者が」を無視して「どの店で」だけに注目するということもあり得る。例えば、クレジットカードは規模の小さい店でも使われているので、規模の小さい、個々の店の売上を集計することにより、中小零細企業の売上の指標を作ることにもできる。中小

企業庁とナウキャストが作成した「中小企業売上指数」はそうやって作られたものだ。

クレジットカードの履歴データの使用をこのように整理してみると、「どの消費者が」と「どの店で」の両方の情報を、どちらも一切捨てることなく、利用し尽くすということとはできないのだろうか、という疑問が湧いてくる。本節の分析が意図しているのは正にそれだ。ある消費者がある店で買い物をするというのは、その消費者とその店がその買い物によって結ばれた、少し違った言い方をすると、「リンク」が生まれたことを意味する。

つまり、各消費者と各店舗は、このようにして生成されたリンクによってつながっており、ネットワークが存在する。本節の目的は、消費者と店の間のネットワークはそもそもどのような性質をもっているのか、その性質はパンデミック下でどのように変質したかを調べることである。

ところで、SNA で言えば、「どの消費者が」というのは支出の場面での情報であり、「どの店で」というのは販売（あるいは生産）の場面での情報である。したがって、クレジットカードの履歴データは、三面等価のうちの、「支出＝生産」を我々に見せてくれているともいえる。ただし、SNA では、支出サイドを捉える際の統計と、生産サイドを捉える際の統計は同一ではない。したがって、ある消費者の支出が支出サイドを捉える統計に含まれているとして、その消費者のその支出が生産サイドを捉える際の統計にも含まれている確率はほぼゼロだろう。

これに対して、「どの消費者が」と「どの店舗で」の両方の情報をもつクレジットカードの履歴データでは、「支出」を構成する購買取引のひとつひとつと「販売（生産）」を構成する取引のひとつひとつが完全に一致しており、この意味で、三面等価を（マクロではなく）ミクロのレベルで観察できるという特徴がある。

本節では、東京都新宿区の飲食店とそこで支出する消費者たちに注目し、取引ネットワークを生成した上で、その特徴を概観する。次に、ネットワークのコミュニティ抽出の手法を用いて、取引ネットワークをコミュニティに分割し、各コミュニティに属する店舗と消費者の特徴を確認する。さらに、コミュニティの時系列特性をみることで、パンデミックがネットワークにどのような影響を及ぼしたかを調べる。

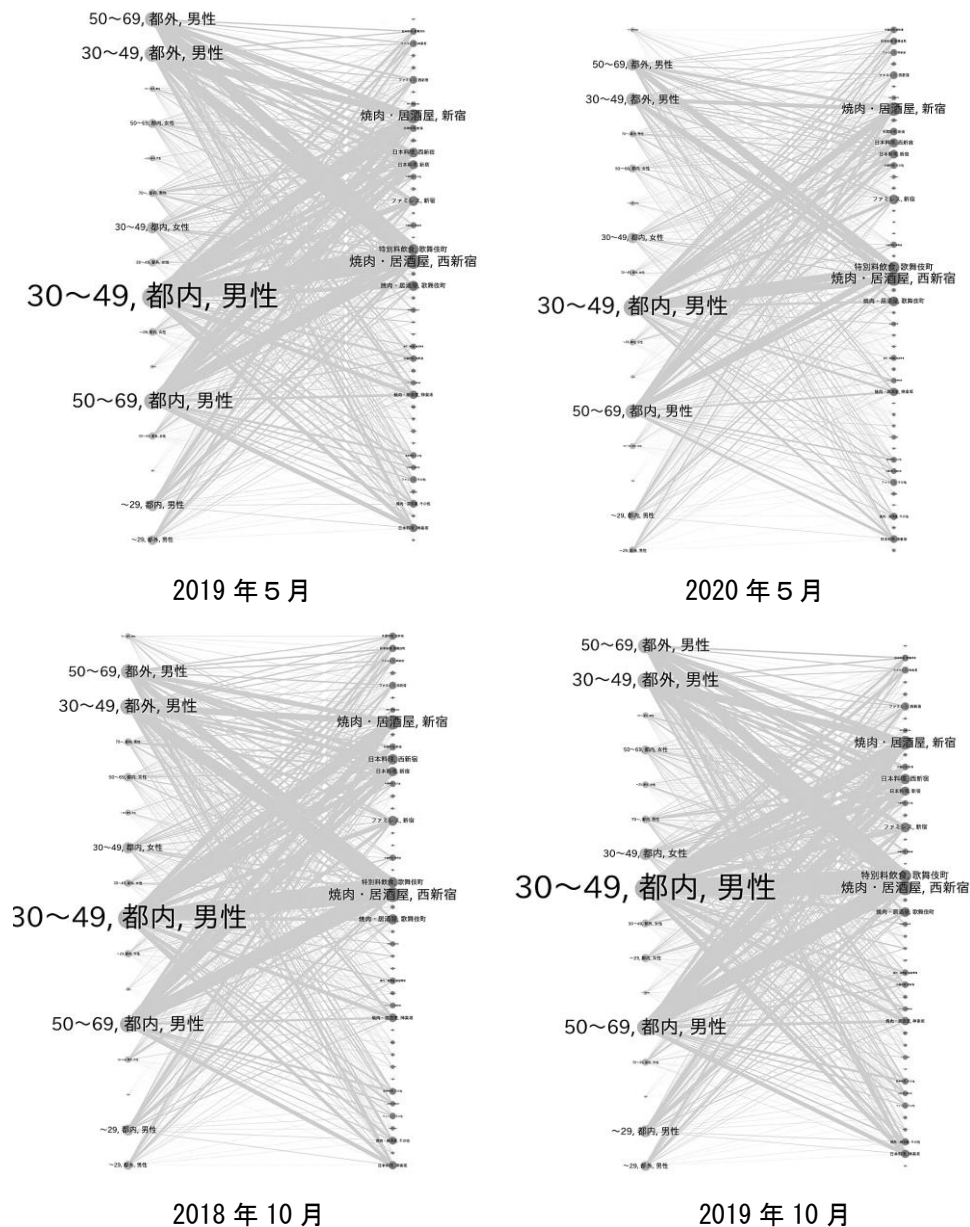
4.2 新宿区飲食店の取引データ

本節で用いるデータはJCB の新宿区の飲食店における取引データである。具体的には、東京都新宿区の飲食店のうち売上総額の上位 1,000 店舗と、その店舗で決済した消費者を月ごとに抽出した。「新宿区の飲食店」に含まれる店舗には、新宿区に本社があり、全国に店舗展開しているような企業も含まれ、その場合は、新宿区内の販売と区外の販売を区別できない。本節の分析では、これらの事例を除外した。また、抽出した店舗データには「四谷」「神楽坂」「歌舞伎町」といったエリア情報が付与されているが、西新宿はビルによってエリア表記に揺れがあるため、「西新宿」という 1 つの地域にまとめた。消費者については、店舗の抽出後、対象店舗への決済金額が正となる消費者を抽出した。

4.3 消費者属性と店舗属性の相関

新宿区の飲食店を取り巻く消費行動について、直感的に特徴を把握できるように、同じ属性の店舗・消費者を集約して2部グラフを描画する。該当する消費者の人数や飲食店の店舗数が多いほどノードをより大きく描き、該当ノード間の決済総額が大きいほどエッジをより太く描いている。

図6：消費者の属性と店舗の属性の対応関係



消費者については、年齢を～29歳、30～49歳、50～69歳、70歳～で、居住地を都内、都外で、性別を男性、女性で分けた。店舗については、飲食店の中でもさらに細かい業種、

東京都新宿区の中でもさらに細かいエリアで分けた。集計元データの業種分類「焼肉・ビアホール」を「焼肉・居酒屋」に変える、店舗数が少ないエリアを「その他」でまとめる、といった加工を行っている。

図6の上段は、2018年10月と2019年10月の2時点について、消費者属性と店舗属性の相関を調べた結果を示している。2019年10月は消費税率引き上げの直後であり、消費税率引き上げの消費への影響を抑制するために経産省によるキャッシュレス・ポイント還元事業が行われていた。2時点を比較することにより、これらの施策の効果を確認できる。図6の下段は、パンデミックが始まり第1回の緊急事態宣言が出ていた2020年5月とその前年の2019年5月を比較している。

図6からわかるように、消費者や店舗の中でも中心的な役割を占める属性は変わっておらず、客層、業種、エリアの入れ替わりは起きていない。しかし、消費税率引き上げ・ポイント還元事業の際には取引が活発になりネットワークが密になっていることがわかる。その逆にパンデミックの際には取引が減ってネットワークが疎になった。

なお、ネットワークのノード数とエッジ数をカウントすると、2019年5月はノード数が36,064、エッジ数が45,084であったのに対して、パンデミック中の2020年5月は、ノード数が26,264、エッジ数が31,388と大きく減少している。2018年10月と2019年10月の比較では、前者がノード数33,061、エッジ数41,175に対して、後者はノード数37,765、エッジ数47,152となっており、キャッシュレス・ポイント還元事業の影響でノード数とエッジ数ともにわずかながら増加している。

4.4 コミュニティ抽出

次に、消費者と飲食店で構成される取引ネットワークの特徴を確認するために、コミュニティ抽出を行う。連結成分を確認すると、9割以上の消費者、店舗は1つの大きなサブグラフに属しており、他のサブグラフにはそれぞれ1～2店舗しか含まれなかったため、最大のサブグラフをコミュニティ抽出の対象とした。

Clauset et al. (2004) の手法を用いて10個のコミュニティに分割した。繋がりがコミュニティ内でどれだけ閉じているかの指標であるモジュラリティは、2017年5月から2020年12月までの期間で0.81～0.85であった。モジュラリティは1に近いほどコミュニティ内で繋がりが完結していることを意味し、0.8以上であればコミュニティを適切に分割できていると言える。なお、今回の分析では、新宿区の飲食店に限定したが、そのことでモジュラリティが過剰に高い値になっている可能性がある点には注意が必要である。地域・業種を広げると、より多様な消費者・店舗の活動がネットワークに反映されて、モジュラリティが下がることが考えられる。

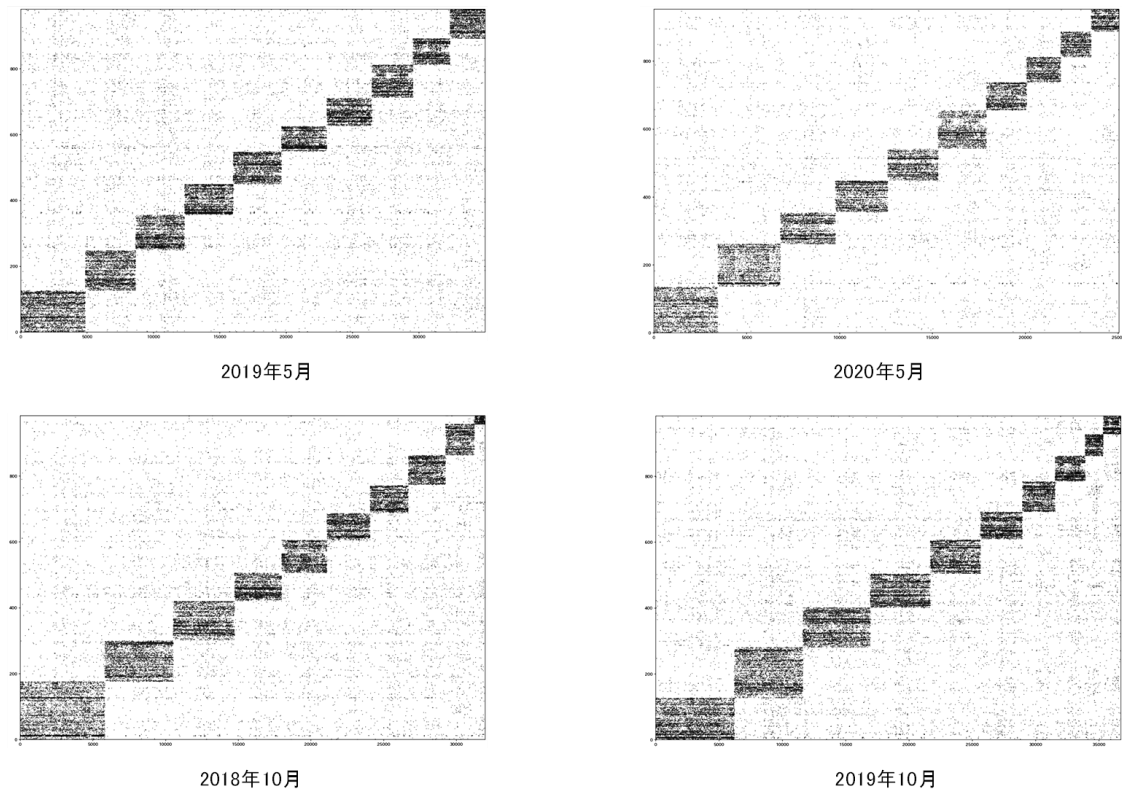
4.5 消費者と店舗の隣接行列

次に、Larremore et al. (2014) に倣い、消費者と店舗の隣接行列をプロットして、抽出し

たコミュニティの可視化を行った。結果は図7に示されている。図7の横軸は個々の消費者であり、その数は約3万人である。各消費者は、①所属するコミュニティの番号（1から10まで）、②居住地（東京都か否か）、③性別、④年齢層、⑤決済金額でソートしてあり、所属するコミュニティを含め、同質性の高い消費者が隣り合うようにしている。図の縦軸は店舗であり、その数は約1,000店舗である。店舗についても、属性の近い店舗が近い位置となるように、①所属コミュニティ、②新宿区のエリア名、③業種名、④決済金額でソートしてある。

図7では、横軸のある消費者が縦軸のある店でその月に購買を行った場合にドットを打っている。図からわかるように、対角成分のところにドットが集中しており、あるコミュニティに属する消費者たちはそのコミュニティに属する店舗で購買を行う傾向が強いことが確認できる。ただし、すべての購買がコミュニティ内でのみ起きているわけではなく、消費者が自分の属するコミュニティとは異なるコミュニティの店舗で購買するということも少ないながら起きている。

図7：消費者と店舗の隣接行列



あるコミュニティに所属する消費者と店舗の特徴を確認するために、店舗については、エリア名、業種名、決済金額、顧客数、平均単価でピボット集計した。決済総額、顧客数、平均単価は絶対値ではなく、ネットワークに存在する他の店舗と比べたときの位置付けを

求め、決済総額が相対的に多ければAランク、少なければDランクといったラベルを与えて四分位で表現した。顧客数が多ければ人気店、単価が高ければ高級店と解釈できる。

消費者についても同様に、居住地、性別、年齢、決済金額、購買履歴のある店舗数、平均単価でピボット集計した。決済金額と平均単価は四分位で表現した。取引店舗数については、新宿の飲食店に限定すると1店舗しか取引実績のない消費者が多数存在し、四分位では集計できなかったため、次数分布を踏まえて閾値を設定し、Aランク～Dランクで表現した。

「2018年10月」の図の最も右端にあるコミュニティを例にとると、そのコミュニティに属する店舗のうち76%が歌舞伎町、53%が特別料飲食店（いわゆるホストクラブやキャバクラ）、65%が高単価な店舗と判定された。そのコミュニティに属する消費者の約83%は男性で、他のコミュニティに比べて相対的に取引店舗数が多く、支払総額も相対的に高い消費者が多い。

同様に、「2020年5月」の図の最も左端にあるコミュニティに注目すると、20代～30代の男性、西新宿の居酒屋が多い。若手の会社員による飲み会のようなケースが考えられる。また、「2020年5月」の左から2番目のコミュニティでは、相対的に支出金額が低い若年男性が多く、高田馬場、神楽坂、西新宿といったエリアの居酒屋を訪れている。さらに左端から数えて3番目のコミュニティは、女性の比率が相対的に高く、新宿駅の周辺にある人気店を訪れていることが分かる。「2019年5月」の図の左から4番目のコミュニティをみると、相対的に支出金額の低い若年層の男性が、高田馬場に位置する低単価の居酒屋を複数訪れている。

4.6 エッジの増減から見たコミュニティ構造の変化

キャッシュレス・ポイント還元事業やパンデミックの影響でコミュニティ構造はどのように変化しただろうか。その点を詳しくみるために、以下の計算を行った。例えば、パンデミック前の2019年5月のデータを用いてコミュニティ抽出を行い、各消費者、各店舗の所属するコミュニティを確定させ、それに応じて、隣接行列における各消費者と各店舗の位置も確定させる。これが図8の「2019年5月」の図である。次に、各消費者と各店舗の位置を固定したままで、その平面に、パンデミック中の2020年5月における取引をプロットする。これが図8の「2020年5月」の図である。

「2019年5月」の図と「2020年5月」の図の差分は、①2019年5月には存在せず2020年5月に新たに生成された取引関係（図8の「新規に生成されたパス」）と、その逆に、②2019年5月には存在したが2020年5月には消えてしまった取引関係（図8の「失われたパス」）である。このようにして、パンデミック前後でネットワークがどのように変化したかを可視化することができる。図8ではキャッシュレス・ポイント還元の前後でも同様の可視化を行っている。

図8：コミュニティ構造の変化

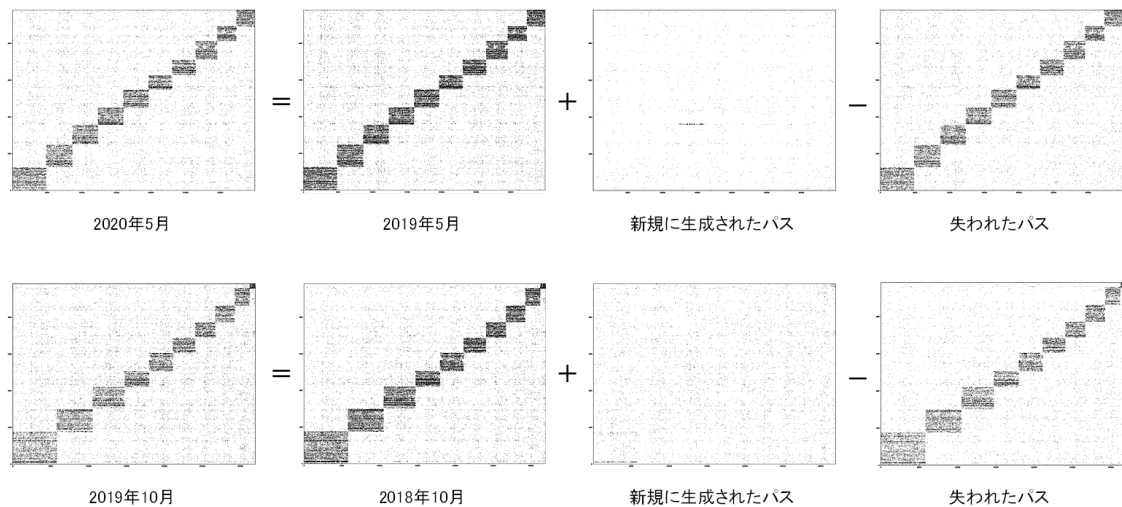


図8の結果をみると、新規に生成されたパスよりも失われたパスのほうが多く、取引ネットワークは1年後には疎になっていることがわかる。これは、1年前の時点においてあるコミュニティに属していた消費者の大部分は一年後もそのコミュニティに属しているものの、別なコミュニティへと鞍替えする消費者も一定数いることを意味している。図には示していないが、このようなコミュニティ構造の変化は、パンデミックやキャッシュレス・ポイント還元プログラムのような、大きなイベントのない時期でも起きている。

図8をもう少し詳しく見ると、キャッシュレス・ポイント還元事業では、図の最も左端のコミュニティにおいて「新規に生成されたパス」の一部が濃くなっている。このコミュニティに属する店舗の特徴を詳しくみると、新宿駅、大久保、高田馬場といったエリアに位置する居酒屋で、顧客単価が低く、顧客数が多いということがわかる（一方、消費者の年齢、性別、居住地に目立った特徴はない）。キャッシュレス・ポイント還元事業は、こうした特徴をもつ店舗での消費支出を促進する効果が大きかった可能性を示唆している。

また、パンデミックの前後では、左から4番目のコミュニティで「新規に生成されたパス」の一部が濃くなっている。特徴としては、高田馬場の飲食店の64%がこのコミュニティに所属している。他にも大久保、西新宿、神楽坂の飲食店も含まれているが、総じて低単価の居酒屋が多い。消費者は支出金額の低い若年層の男性が多い。パンデミック前は高田馬場エリアに出かけていた若年男性たちが、パンデミックを機に、西新宿や神楽坂へと移動した可能性がある。例えば、従来は大学やオフィスの近くで大人数の歓迎会を行っていたが、パンデミックで店舗営業の自粛が始まり、ターミナル駅周辺や生活圏での少人数の食事会にシフトした、といったケースが考えられる。

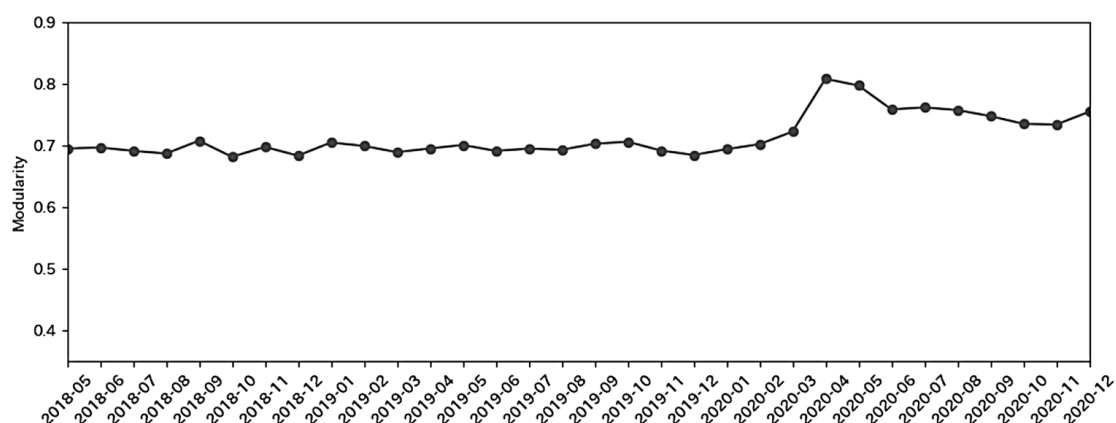
4.7 モジュラリティの時間的变化

次に、繋がりがコミュニティ内でどれだけ閉じているかの指標であるモジュラリティを

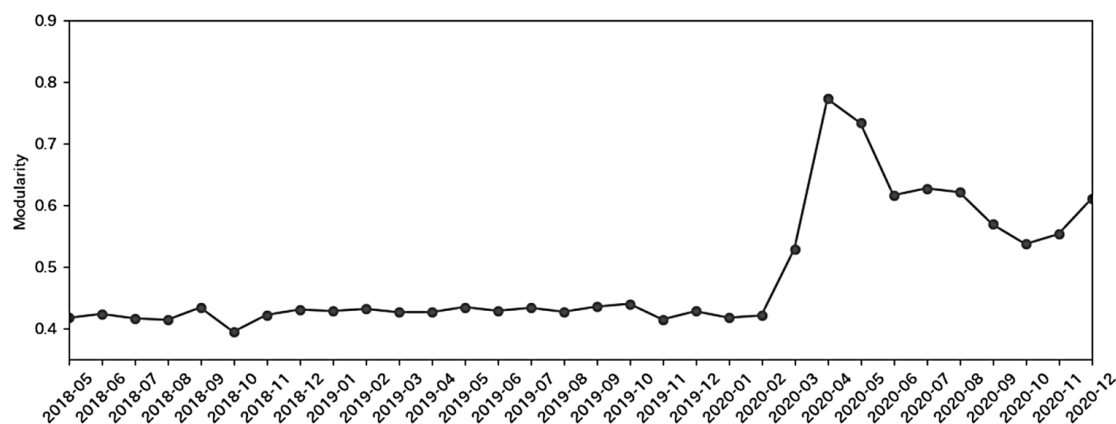
計算し、コミュニティが時間とともにどのように変化したかをみてみよう。ここでは、これまでの分析と同じく、ある時点の1年前の月のデータを用いてコミュニティ抽出を行う。その上で、抽出されたコミュニティを所与として、1年後の取引データを用いてモジュラリティを計算する。こうすることで、1年間の間でコミュニティの構造に変化があったか否かを確認することができる。例えば、1年経ってもコミュニティ構造があまり変化していなければモジュラリティの値は1に近いはずだ。一方、1年間でコミュニティ構造が大きく変化した場合にはモジュラリティの値は1を大きく下回るはずだ。

この方法によるモジュラリティの計算に際しては、1年前に存在せず、新たに生まれたノード（消費者と店舗）をどう扱うか、その反対に、1年前には存在していたが消えてしまったノードをどう扱うかを決める必要がある。まず、新規のノードについては、①新規のノードをネットワークから外し計算の対象外とする、②新たなコミュニティに属する（既存のコミュニティの数は10個なので、新たに11番目のコミュニティが存在し、新規のノードはそこに属すると考える）という2つの方法で対処した。一方、消滅したノードについては計算の対象外とすることとした。

図9：モジュラリティの時間的变化



1年前に存在しなかったノードを計算の対象外とした場合



1年前に存在しなかったノードを11番目のコミュニティに判定した場合

図9はこのようにして計算したモジュラリティの値を示したものだ。上段は新規ノードを無視した場合の結果、下段は新規ノードを新たなコミュニティに割り付けた場合の結果である。上段の図では、2018年5月から2019年末の間、モジュラリティの値は0.7前後と比較的高い値で推移しており、コミュニティ構造に大きな変化がなかったことを示している。しかしモジュラリティの値は2020年4月には0.8まで上昇しており、その後もしばらくの間、それまでの時期に比べると高めの水準を維持した。2020年4月はパンデミックに伴う最初の緊急事態宣言が出された時期であり、これがきっかけとなってモジュラリティの値が高まった可能性がある。緊急事態宣言の発令とともにモジュラリティの値が高まるという傾向は下段の図でも同様であり、新規ノードの処理方法に依存しないことを示している。

図6で確認したように、緊急事態宣言の発令とともにノード数とエッジ数は顕著に減った。これは消費者が自宅にこもり、新宿区の繁華街に出かけなくなったことを意味しており、スマホの位置情報データを用いた前節の分析で確認されたとおりである。しかしノード数とエッジ数が減れば必ずモジュラリティの値が上昇するかと言えば決してそんなことはない。ノード数とエッジ数の減少がランダムに起きたとすれば、モジュラリティの値は2019年までの値と大差なかったはずだ。しかし実際にはモジュラリティの値は顕著に上昇したのだから、それぞれのコミュニティ内における消費者と店舗繋がりがこの時期、より一層緊密になったことを意味している。

前節の分析結果と合わせて考えると、緊急事態宣言の発令によって多くの消費者はステイホームを実施し、新宿区の繁華街に出かける人は減った。しかし減ったとは言っても皆無になったわけではなく、一部の消費者は緊急事態宣言を無視して、引き続き新宿区の繁華街に出かけていった。その際、それらの消費者が訪れたのは、典型的には、馴染みの店（1年前にも足繫く通っていた店）であり、馴染みでない店（1年前に行っていなかった店）を新たに訪れるということではなかった。これらの消費者がなぜ馴染みの店を訪れたのかはここでの分析からはわからないが、馴染みの店だからこそ、緊急事態宣言を無視してでも訪れたいと考えた可能性は否定できないだろう。また、当時、営業を停止する飲食店が少なくなかったことを考えると、馴染みの店だからこそ営業中という情報を入手できた（馴染みでない店は営業しているのかどうかわからなかった）ということもあるかもしれない。

前節でも議論したように、ウイルスの感染拡大を防ぐ最善の手段は人と人の接触をなくすこと、つまり隔離だ。ステイホームは家族以外との接触をなくす隔離政策だ。この理屈から言えば、日本国民全員が家に閉じ籠るのが理想だ。しかし国民の一定層、特に重症化率・致死率の低い若年層にとっては、外出を抑えることの費用がその便益に見合わないの、外出したいと考える。

では、こうした若年層の外出はある程度仕方ない、認めざるを得ないとして、その前提の下で感染拡大をできるだけ抑える策、いわば次善の策は何だろうか。それは、こうした消費者たちを特定の少数の店舗に集中させる、その意味で彼らを一般社会から隔離するこ

とだろう。Acemoglu et al. (2020) は、重症化率・致死率の高い高齢層に照準を絞って自宅に籠ることを要求するというかたちでの隔離を提唱したが (targeted lockdown)、重症化率・致死率の低い若年層を特定の店舗に集中させるというのも、類似の隔離効果をもつと考えられる。本章の分析で確認された、緊急事態宣言直後のモジュラリティの上昇は、そうした隔離が実際に行われた可能性を示唆している。

参考文献

- Acemoglu, Daron, Victor Chernozhukov, Ivan Werning, and Michael D. Whinston (2020), “Optimal targeted lockdowns in a multi-group SIR model.” No. 27102. National Bureau of Economic Research.
- Aum, Sangmin, Sang Yoon Tim Lee, and Yongseok Shin (2020), “Inequality of fear and self-quarantine: Is there a trade-off between GDP and public health?” No. 27100. National Bureau of Economic Research.
- Brotherhood, Luiz, Philipp Kircher, Cezar Santos, and Michele Tertilt (2020), “An economic model of the Covid-19 epidemic: The importance of testing and age-specific policies.” CESifo Working Paper No. 8316, June 2020. Available at <https://ssrn.com/abstract=3618840>
- Chetty, Raj, John N. Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team (2020), “How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data.” No. 27431. National Bureau of Economic Research.
- Clauset, A., M.E.J. Newman, C Moore (2004), “Finding community structure in very large networks,” *Physical Review E* 70.
- Cochrane, John (2020) “A SIR model with behavior.” Available at <https://johnhcochrane.blogspot.com/2020/05/an-sir-model-with-behavior.html>
- Eichenbaum, Martin S., Sergio Rebelo, and Mathias Trabandt (2020), “The macroeconomics of epidemics.” No. 26882. National Bureau of Economic Research.
- Goolsbee, Austan, and Chad Syverson (2021), “Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020,” *Journal of Public Economics*, Volume 193, January 2021, 104311.
- Larremore, D. B., A. Clauset, A. Z. Jacobs (2014), “Efficiently inferring community structure in bipartite networks,” *Physical Review E* 90.
- Ministry of Health, Labour and Welfare (2020), “10 things to know about the COVID-19 as of right now.” December 2020. Available at: <https://www.mhlw.go.jp/content/000727050.pdf>
- Mizuno, Takayuki, Takaaki Ohnishi, Tsutomu Watanabe (2021), “Visualizing social and behavior change due to the outbreak of COVID-19 using mobile phone location data,” *New Generation*

Computing, 39, 453–468.

Rojas, Felipe Lozano, Xuan Jiang, Laura Montenovo, Kosali I. Simon, Bruce A. Weinberg, and Coady Wing (2020), “Is the cure worse than the problem itself? Immediate labor market effects of COVID-19 case rates and school closures in the U.S.” No. 27127. National Bureau of Economic Research.

Sheridan, Adam, Asger Lau Andersen, Emil Toft Hansen, and Niels Johannesen (2020), “Social distancing laws cause only small losses of economic activity during the COVID-19 pandemic in Scandinavia.” *Proceedings of the National Academy of Sciences*.

Watanabe, Tsutomu, and Tomoyoshi Yabu (2021a), “Japan's Voluntary Lockdown.” *PLoS ONE* 16(6): e0252468. <https://doi.org/10.1371/journal.pone.0252468>. Published: 10 June 2021.

Watanabe, Tsutomu, and Tomoyoshi Yabu (2021b), “Japan's Voluntary Lockdown: Further Evidence Based on Age-Specific Mobile Location Data,” *Japanese Economic Review*, 72(3), 333-370.