

## 速報性のある包括的な消費関連指標の在り方に関する研究会（第2回）議事概要

1 日時 平成28年10月14日（金） 15:30～17:00

2 場所 総務省7階省議室

3 出席者

## 【構成員】

飯塚 信夫	神奈川県経済学部教授
小塩 隆士	一橋大学経済研究所教授
久我 尚子	株式会社ニッセイ基礎研究所主任研究員
座長 国友 直人	明治大学政治経済学部特任教授
熊谷 亮丸	株式会社大和総研調査本部副本部長・チーフエコノミスト (代理 齋藤 勉 株式会社大和総研調査本部エコノミスト)
新家 義貴	株式会社第一生命経済研究所主席エコノミスト
末澤 豪謙	SMBC日興証券株式会社金融財政アナリスト
菅 幹雄	法政大学経済学部教授
関口 博之	日本放送協会解説主幹
永井 暁子	日本女子大学人間社会学部准教授
平嶋 彰英	総務省参与
座長代理 美添 泰人	青山学院大学経営学部プロジェクト教授
渡辺 努	東京大学大学院経済学研究科教授

## 【オブザーバー】

茨木 秀行	内閣府政策統括官（経済財政分析担当）付参事官（総括担当）
長谷川 秀司	内閣府経済社会総合研究所国民経済計算部長
吉牟田 剛	総務省政策統括官（統計基準担当）付統計企画管理官
山澤 成康	総務省統計委員会担当室長
宇波 弘貴	財務省大臣官房総合政策課長 (代理 森田 稔 財務省大臣官房総合政策課経済財政政策調整官)
上田 洋二	経済産業省大臣官房調査統計グループ参事官（総合調整担当） (代理 高辻 育史 経済産業省大臣官房調査統計グループ調査分析支援室長)
肥後 雅博	日本銀行調査統計局参事役（統計担当） (代理 藤原 裕行 日本銀行調査統計局経済統計課統計総務グループ 企画役)

## 【発表者】

堀井 克倫 CCCマーケティング株式会社企画本部事業企画部データサイエンス・ラボ所長

## 【総務省】

高市総務大臣、島田総務大臣政務官、笹島総務審議官、會田統計局長、長屋官房総括審議官、佐伯官房審議官

## 【事務局（総務省統計局統計調査部）】

千野統計調査部長、阿向消費統計課長

#### 4 議事

- (1) 高市総務大臣挨拶
- (2) 東大日次物価指数プロジェクト  
－渡辺構成員
- (3) カルチュア・コンビニエンス・クラブの取組みについて  
－CCC マーケティング株式会社  
企画本部事業企画部データサイエンス・ラボ所長 堀井 克倫 様
- (4) 意見交換

#### 5 議事概要

- (1) 高市総務大臣挨拶
  - ・御多用の中での御参集に感謝。
  - ・第1回の研究会では、皆様から貴重な御意見を頂いた。新たな消費関連指標の開発に向けて、有意義な一步を踏み出せたと感じている。
  - ・前回の研究会から一ヶ月しか経過していないが、この短期間でも政府統計への関心が急速に高まっている。政府や党の様々な場において、経済統計の検討が始まった。それに呼応する形で、この消費関連指標の開発についても、更なる注目が集まっている。
  - ・本日は、東京大学大学院経済学研究科の渡辺努先生、CCC マーケティング株式会社の堀井様からビッグデータを用いた統計作成の最先端を見せていただけるとのこと、大変楽しみにしている。
  - ・国際的にも高く評価され得るような指標の開発に当たっては、政府統計の枠にこだわらずに、研究やビジネスの世界最先端がどこにあるのかを知ることが非常に大切だと考えている。渡辺先生、堀井様、構成員の先生方におかれては、本日もよろしくお願ひ申し上げます。
- (2) 東大日次物価指数プロジェクト

渡辺構成員から資料1に基づき説明が行われた。主な説明は以下のとおり。

  - ・私が研究を進めているビッグデータを用いた物価指数の作成について話す。物価と消費は異なる部分もあるが、どちらも消費者の行動を反映している。物価の研究に用いた POS データは、消費を議論する際にも重要なビッグデータの候補。その意味で、物価の話の主としつつも、消費を考える際の有益な示唆を与えることができると考えている。
  - ・最近では、経済学でもビッグデータを使う研究が非常に多い。私の場合は、各企業から価格に関するデータを網羅的に集めて研究を行ってきた。特に注力しているのが POS データ。日本経済新聞社が集めた POS データを基に物価指数を作成できることが分かり、2013年の春に、東京大学のホームページ上で「東大日次物価指数」として配信を開始した。現在はナウキャストという会社から「日経 CPI ナウ」として配信。

- ・ 指数の公開を始めた理由は、文部科学省から研究費を得ていたこともあり、何らかの形で社会に還元するのが望ましいと考えたから。
- ・ 足元の経済状況をリアルタイムで把握する試みである経済ナウキャストという概念が直近2、3年で現れた。私たちが研究している物価指数もナウキャスト指標の一つであり、本研究会で議論される消費指標もその一つだと理解。
- ・ 株価や為替は秒あるいはそれよりも短い単位の高頻度で捉えられているのに比べて、消費者物価指数は総務省統計局から一ヶ月に一回公表されるのみで、物価の捕捉頻度は低かった。物価についても高頻度で情報が得られれば、各企業は効率的な経済活動を営めるのではないかと考えた。
- ・ POS データの集計は現在、ほぼ完全に自動化されており、日次での集計と翌々の公表が可能。ただし、POS データはスーパーマーケットから集めているデータのため、カバーしている範囲は食料品や日用雑貨などに偏っている。
- ・ POS データの動きを一部紹介する。  
資料1のP.7の上のグラフは、ある店舗のある商品の価格変動を表したもの。黒線が実際の価格の変動、赤線が価格の基準線。週に一回程度、特売が実施されており、黒線が頻繁に下に振れている。一方、P.7の下のグラフは対応する販売数量を表している。特売日には青線が上に振れており、特売日の売れ行きは、平時に比べて桁違いに大きいことが読み取れる。このように、マイクロで見たときには非常に特異な動きをしている場合がある。価格も数量も、変数は連続的に動いておらず、単純に足すのは非常に危険。消費だから足せばよい、というほど単純ではない。
- ・ 我々が研究の初期に真剣に考えたのは、作成方法を確立した上で、それを開示すること。作成方法が信頼できないと誰も見ず、秘匿されていれば、信じて良いか分からない。そこで私たちは、学術的な背景を明確にした上で、独自の工夫を加えて計算のアルゴリズムを作成し、公開した。同じデータがあれば結果を再現できるようにすることで品質の保証ができ、多くの人が指数を使う動機付けになると考えた。
- ・ また、過去のトラッキングをきちんと見せる必要があると考えた。できるだけ過去の動きを既存の統計と比べて、どのような状況でどのような違いが生じるのかを示し、ユーザーに十分な注意事項となるようなメッセージを発することを心がけた。私たちの指数は、総務省の消費者物価指数と比較し、過去にさかのぼってよく似た動きをしている一方で、概して指数が低く出る傾向がある。これは、総務省では特売価格を除外するのに対し、私たちの指数ではこれを算入しているため。このように既存の統計と「似ている」ことを強調すると同時に、手法やデータの違いを反映して「違う」点も主張し、切り分けることによって、安心して使ってもらえるように心がけた。
- ・ オランダやスイスといった人口規模の小さな国では、公的統計の作成部署もPOSデータから物価指数を作成している。人口規模の小さな国だからこそ出来ている部分もあるが、ビッグデータというのは、必ずしも民間だけのものではなく、政府でも使うところが出てきている。
- ・ 我々の日次物価指数について、いくつかの使用事例を紹介する。

- 資料1のP.13に示した図は、近年の日経CPIナウの動き。2013年4月に異次元金融緩和が始まって以降、消費税率の引上げや追加緩和の実施等のタイミングで物価の動きが変化していることが明確に読み取れる。このように、日次でデータを見ることによって、政策が物価に与える影響がかなり見やすくなる。消費についても、恐らく同じことが言える。
- 資料1のP.14に示した図は、東日本大震災時の物価と売上げの動き。売上げは買いだめなどの影響で震災直後に急激に上昇しているが、物価の動きは鈍く、売上げに比べて緩やかな上昇となっている。震災による需要で物価が上昇しているものの、実態としては、通常価格の値上げというよりも、特売の頻度や値引率を小さくする動きが見られた。
- バーコードの中には製造企業の情報も含まれているため、企業別の物価指数を作ることができる。資料1のP.18に示した表は、左側が2012年9月からの3年間に商品価格が上がった企業のグループ、右側が下がった企業のグループ。この表から、飲料業界では近年、価格が下落傾向にあり、過当競争と言ってもよい状況が生じていることが読み取れる。こういった分析も、金融政策を考える上でひとつの材料になればという思いから行っている。
- POSデータからは、どの店舗のデータか特定できるため、地域別の指数を作ることができる。日本国内であっても、地域間における嗜好の違いのようなこともPOSデータから読み取ることができる。
- 最後に、物価指数を作成する中で得た教訓を、消費を意識して整理したい。
- 1点目、ビッグデータの一番の魅力は非常に細かい粒度で分析が可能だということ。物価の例で言えば、商品単位や店舗単位などの指数を作ることができる。細かいレベルで見ると、各々の違いが見えてくる。環境や家族構成など、色々な条件に応じて消費のパターンは異なり、単に平均値を見ればよいわけではない。異質のものは異質なまま処理をしながら、何らかの情報をくみ取るという発想が大事。それはビッグデータだからこそできるものと考えます。
- 2点目、ビッグデータは無闇に集計してはいけないということ。集計をする前にデータの特性を十分に吟味する必要がある。先に挙げた特売の例のように、特売の頻度や割引率は、集計するとその情報が消えてしまう。ミクロのレベルで十分に検討した上で集計する。この順番を間違えてはいけない。
- 3点目、ビッグデータにはサンプルバイアスがあること。ビッグデータは、サンプリングをきちんとしている訳ではなく、業務用に作られたデータに過ぎない。バイアスを除く努力は必要だが、完全に除くことはほぼ不可能。残ったバイアスについて、政府統計の利用者にどのように理解してもらおうかが、コミュニケーションの問題として重要。
- 4点目、ビッグデータの議論をする際によく出てくるのが、精度と迅速性。月に一度の公表では、関心を持って統計を見続けるということは難しいが、東大日次物価指数のように毎日更新されれば、ユーザーはそれを毎日確認するようになる。これに併せて、利用のされ方も変わってくる。利用者から統計に関する御意見を頂く頻

度も変わる。

- ・ 5点目、広く使ってもらうには、作成方法を十分に開示すること、過去データがあるということがまず重要。過去データを使ってトラッキングが十分になされていると提示することが、新しい指標を世に問うときには大事なプロセス。

### (3) カルチュア・コンビニエンス・クラブの取組みについて

CCC マーケティング株式会社企画本部事業企画部データサイエンス・ラボ所長 堀井様から資料2に基づき説明が行われた。主な説明は以下のとおり。

- ・ 私は CCC マーケティング株式会社で、データにまつわる研究を行っている。当社はこれまで、Tカードを通じて ID 付きの POS データを皆様からお預かりし、ビジネスを進めてきた。この事業開始の背景を御紹介したい。
- ・ CCC (カルチュア・コンビニエンス・クラブ) は、大きく TSUTAYA、カルチュア・エンタテインメント、CCC マーケティング、CCC デザインカンパニーから構成される。CCC 全体に共通して言えるのは、生活提案を実現するような企画を様々に実施してきたということ。その中で、Tカードのデータがあるということを強調したい。
- ・ 当社では、Tポイント物価指数 (TPI:T-POINT Price Index) を公表している。この指数のベースとなるTカードデータの会員構成をまず解説する。
- ・ Tカードは、分散していたポイントカードやクレジットカードを一つにまとめ、TSUTAYAの会員証にそれらの機能を付与するという生活提案の一つとして始まった。
- ・ 現時点で、1年間に御利用いただいた会員数(名寄せ後)は5,981万人に上る規模。これは日本の総人口の47%に相当。
- ・ Tカードの提携先は現時点で158社、56万4,123店舗に及び、スーパーマーケット、外食チェーン、ドラッグストア、百貨店等が提携企業として入っている。家計調査の10大費目のほぼ全てをいずれかの企業が何らかの形でカバーしている。
- ・ 都道府県別にTカードの会員比率を見ると、全国に会員が分布しているが、特に、一都三県や鹿児島、沖縄が高い。
- ・ Tカードの月間利用件数は3億4,452万件。これは伝票レベルでの数値であり、TPIの作成に当たり、実際に集計している品名レベルのレコード数は更に多い。
- ・ Tカードが関与した売上額は月間で5,277億円、年間で約6兆円の規模の消費データをお客様からお預かりしている。
- ・ 日本人の4割を超える規模にTカードの利用者が拡大したのであれば、これを社会に還元していくのが責務ではないかと考えた。社会価値の還元として始めたのが、T-POINT Price Index (TPI)。集計、計算の方法については渡辺努先生、渡辺広太先生にアドバイスを頂きながら、T経済圏の中での社会経済指数を公表している。
- ・ TPIの大きな特徴は、次の3点。
- ・ まず1点目は性別・年齢別の物価指数が見えること。もともとTカード(TSUTAYAのレンタル会員証)という形で、顧客の属性情報をしっかりと把握してきた、その

基盤が性別・年齢別の TPI につながったと考えている。国の経済政策が、どの世代、どの性別に届いたか、ということが可視化されれば、我々の置かれている生活環境も少し変わるのではないかと考えた。

- ・ 2 点目は、買値での物価であるということ。
- ・ 3 点目は、幅広いカバレッジを持っているということ。現在はスーパーマーケット、ドラッグストア、コンビニエンスストア、外食チェーンでの取引に限定されているが、提携企業は今もなお拡大しており、全体の指数を作成するのが当面の目標。
- ・ 最後に、これはお客様にも明言しているが、個人情報とデータベースマーケティングで扱う情報は厳密に分けて管理している。個人のプライバシーに配慮しながら、社会価値の還元につなげていこうと取り組んでいる。

#### (4) 意見交換

##### ○議事 (2) について (渡辺構成員に対する質疑)

- ・ 東大物価指数を作成する上でのウエイトはどのようにしているか。

→前年比を計算しているので、対象となる日と前年同日の両方に存在した商品を探して、その価格変動を計算する。トゥルンクヴィスト式で用いるウエイトは、対象となる日の商品シェアと前年同日のシェアを足して 2 で割ったもので、フィッシャーのような指数。通常、消費者物価指数などでは現時点のウエイトは使えず、基準時点のウエイトを使うことになるが、現時点のウエイトを使えることが POS データの強み。

- ・ 資料内のグラフでは、短期的には大きな変動があるので、移動平均を取っているように見えるが、その辺りを御説明いただきたい。

→大きく変動している線が生データ。生データには特売の影響やそれぞれの店舗又は商品に固有な事情など、様々なものが反映されている。当初は季節性を綺麗に除去することを真剣に考えていたが、店舗は特売日などを巧みに変えたりするので、季節性を綺麗に除去することは、非常に困難であると理解した。今のところ最大限嘘を付かずにできるものとして、一週間の移動平均をかけている。

- ・ 資料 1 の P. 26 で特に大事だと思うのは、ビッグデータは無闇に集計してはいけないということ。データが豊富にあるということは、様々な集計が可能で、色々な結果が出る。それをどう読むのか、どう集計するかは相当な訓練を積んでいないと難しい。この研究はこの点を考えていることが素晴らしい。

→集計方法については、半年くらいかけて数千くらいパターンを試行錯誤した。例えばウエイトを変えるとどのような傾向が出るのか、一つ一つ確認した。その中でどれが他者に説明する際に一番説明しやすいかを考え、最も正しいと思えるものを採用した。試行錯誤は必要。

- ・ 資料 1 の P. 26 の「ビッグデータにもサンプルバイアスはつきもの」とあるが、これは適切な認識。「ビッグデータは全数だからサンプリングの議論は要らない」などと言う人もいるが、それに対して、ビッグデータはサンプルそのものであるという、

この認識は妥当である。

- ビッグデータといっても全数ではないのでサンプルバイアスの問題は当然ある。その上で、バイアスを除去するために最大限の工夫をすべき、ということを経験したが、それは CCC のように企業でやることと、政府でやることの間には大きな違いがあると思う。我々は、Tポイントカードが使われている世界、経済圏という意味で「Tポイントワールド」、「T経済圏」という言葉を使う。この経済圏の物価を表していると言えば、バイアスのことは考えなくて良い。少なくとも CCC としては、これで十分。一方で政府が行う場合は当然そのようにはいかない。あらゆる策を尽くしてサンプルバイアスを除去することが必要。この研究会の中には統計学の専門家もいるので、お知恵を拝借しながらバイアスのないものを作っていくことはできると思う。ただ、私の知る限りでは、そこまでの完成度でバイアスのないものを作った例というのは内外ともない。非常にチャレンジングな取組だと思う。
- 資料1のP.5「足元の予測には失敗がない、ビッグデータを用いて緻密に計算すればするほど精度が上がる」との記述があるが、これは間違いだ。ビッグデータは、母集団の一部に対応するという意味であくまでもサンプルであり、緻密に計算してもその意味には限界がある。例えば前年の同日に取引のあったものは、一部の品目、一部の店舗に限られる。これは統計学の立場ではサンプルだから、推計の際には母集団との対応を明確に意識することが重要。
  - トゥルンクヴィスト指数は経済学の視点では最良指数であるが、最良指数の定義と解釈は、決して易しいものではないし、経済学者にも誤解がたくさんある。今回の研究で言えば、日次でシェアをウエイトにするのは、統計的な視点からすると安定性の問題があり、かなり危険である。ロバストネスの問題になるが、外れ値の発生メカニズムまで考えて議論をする必要があり、これはビッグデータでは特に難しい問題だと思う。
  - 資料1のP.25、資料2のP.16に、家計調査のカバレッジは300億円、Tカードは5兆円だとある。家計調査のカバレッジは小さい、調査対象が9,000世帯だから精度が低いというのは、よくある勘違い。無限母集団でサンプルサイズが9,000というのは、恐らく学生では扱えないくらい大きなサンプルであるが、母集団が5,000万世帯と聞いてしまうと、途端に割り算をしたくなって、たった300億円、GDPの家計最終消費支出のたった0.01%というコメントが出てくるが、これにより精度が低いと結論付けるのは、基本的な統計学の認識が不足していることによる誤り。高名なエコノミストや経済学者の中にも時々同じような議論をする人がいるので、表現に注意していただきたい。
  - ビッグデータをマクロの統計に用いるのは非常に難しいと感じた。無闇に集計してはいけないという中で、それをマクロの消費を示す指標にどう活用していけばいいのか、今後検討を進めていく際にかかなり困難な点になるだろうと思う。ビッグデータの持っている速報性、属性別データが得られるという特性はミクロの分析に役立つと思うが、ビッグデータの特性を新指標にどう生かしていくのか、非常に困難な課題だと思う。

→無闇に集計しないという意味は、集計してしまうと、特売のようなマイクロのよい情報を消してしまうことになるということ。新しい指標では、家計調査以外の観点から作ろうとしているので、単に家計調査と同じような集計値だけでは寂しい。もう少し肌ざわりがあるような形の新指標にするためには、やはりマイクロのレベルでの工夫が必要だと思う。

・消費の指標を作っていく上では、同時に物価の指標を作ることが必要であると思う。景気を見ていく上では実質消費を見るのが基本で、名目消費金額を物価でデフレートしなければならない。その際に使う物価が、消費者物価指数ベースのものなのか、それとも、Tポイント物価指数のようなある種の消費単価に近いものなのかで、その意味するところはかなり変わってくると思う。

→実質化についても、ビッグデータ、とりわけ POS データの強みはあると思う。公表はしていないが、我々は価格の指数と同時に、それと対となる数量の指数を、POS データから直接作ることができる。

・マクロの消費動向を把握する上で、ビッグデータはどこまで使えるか。

→困難さをやや強調しすぎてしまったかもしれないが、机上の議論よりも、とにかく実物を作ってみることが大事。批判や助言を頂きながら、徐々に改善していくというのが良いアプローチだと思う。

#### ○議事（3）について（CCC マーケティング株式会社 堀井様に対する質疑）

・資料2のP.22で対象品目数について「品目数580品目のうち191品目を活用」とあるが、限定している理由と、品目の範囲について教えていただきたい。得られるものは全て使っているという理解で良いか。

→580品目は総務省の消費者物価指数で対象としている品目数。端数を丸めて580品目と記載している。191品目となっているのは、現状TPIではTポイント提携先のうち、マーケット、コンビニエンスストア、ドラッグストア、外食チェーンに限ってデータを使用しているため。したがって食料、外食、家事用品・身の回り用品が主な品目の範囲。今後拡張の余地はある。

・TPIは価値還元企業の理念に基づいているという発言があったが、仮に「価値還元の一つとして総務省にデータを提供してほしい」と言われたらどこまで協力可能か。

→当社は生活提案を大事にしているので、それにつながるのであれば可能と考える。物価に関する取組は進めているが、まだ生活提案からは遠いと感じている。これをもう少し近づける努力があるならば、そこに参画することは可能と思っている。

#### ○島田総務大臣政務官感想

・渡辺先生及び堀井様の発表、本当に感謝。

・私はよく「ビッグデータの活用」という言葉を使うが、この活用が今日の議論の中で大変難しいものであると改めて感じた。

・委員の皆様方の格段の御理解、精力的な御議論を今後ともお願い申し上げます。