

経済セミナー2019年6・7月号

特集 統計の役割を考える

物価統計における民間データの活用

モノやサービスの値段である物価に関する統計（物価統計）は、
経済全体の動きを知るためのセンサーとして重要だ。
技術変化の進んだ現在、統計のデジタル化により、
データ作成の主体は公から民へと移りつつある。
物価統計に求められる正確さとスピードをどのように担保できるか、考えたい。

渡辺 努 Watanabe Tsutomu

東京大学大学院経済学研究科教授

著者紹介

1959年生まれ。東京大学経済学部経済学科卒業。ハーバード大学 Ph. D.（経済学専攻）
取得。日本銀行、一橋大学を経て、2011年より現職。株式会社ナウキャスト創業者・
技術顧問。著書：『新しい物価理論—物価水準の財政理論と金融政策の役割』（一橋大
学経済研究叢書 52、共著、2018年）ほか。

1 経済センサーとしての統計

マクロ経済学で政府統計が登場する有名な例はルーカスの「島の寓話」モデルであ
る。このモデルで登場するのは物価統計である。レタスの生産者はレタスの販売から
得られる収益で生計を立てている。レタスの価格が上昇したとしよう。そのときにレ
タスの生産者は増産すべきかどうか考える。自分の生産するレタスの価格だけが上昇

して、自分が日常的に消費するモノやサービスの価格が変わらないのであれば、労働投入を増やし増産すべきである。自分が生産する商品であるレタスの価格が、自分が消費する商品の価格との対比で上昇しているからだ。

しかし上昇しているのはレタスの価格だけではないかもしれない。中央銀行がおカネの量を増やし、その結果、レタスを含むすべてのモノやサービスの価格が同率で上がっているのかもしれない。そうだとすれば他の商品との対比でレタスの価格は変わっていないのだからレタスを増産すべきでない。

このように、レタスを増産すべきかどうかの意思決定には自分の生産するレタスの値段だけでなく、すべてのモノとサービスの価格を知ることが不可欠だ。経済全体で価格がどうなっているのかを教えてくれるのが物価統計だ。つまり物価統計は経済全体の動きをレタスの生産者に教えてくれる「センサー」だ。

しかし物価統計はリアルタイムで公表されるわけではない。そのため物価統計が公表されるまでの間、生産者は何が起きているかわからない。こうした中で、例えば、経済全体で物価が上がっているにもかかわらず自分の生産物の価格だけが上がっていると錯覚し過剰に生産するということが起き得る。多くの生産者がこうした錯覚をもつと経済全体で無用な景気変動が生じてしまう。

また、物価統計というセンサーからの情報にノイズがたくさん含まれていると、生産者は物価統計を見ても経済全体の動向を正確に把握できず、生産が過剰になったり過小になったりする。

こうしたミスをもっと抑え経済の効率性・安定性を高めるには、高い精度で作成された物価統計がタイムリーに公表される必要がある。高精度と迅速性は物価統計だけでなく、すべての統計に求められる要件である。

2 誰が統計を作成すべきか

では、経済センサーとしての統計を誰が供給するのか。以下本稿では、この点について掘り下げて考えてみたい。

ルーカスのモデルを学生に解説する際には、物価統計を作成するのは政府だと説明することが多い。実際、社会や経済の状況を観測・記録し、それを国民に知らせるのは政府の重要な役割である。例えば、モノやサービスの値段である物価に関する統計の作成が日本で始まったのは戦前であり、現在の消費者物価統計に至るまで 80 年の

歴史がある。物価は各企業が毎年、賃金を決める際の重要な参考指標であり、物価上昇で賃金が目減りするのを防ぐなどの役割を果たしてきた。また、年金や生活保護などの支給額を決める際にも物価統計は活用されてきた。

社会が物価指標を必要としたので政府がそれに応えて提供してきたというのが、事実の整理としては適当であろう。しかし言うまでもなく、社会が何かを必要とすることと、それを政府が提供することとは直接つながらない。統計データに限らず、社会が必要とするが政府は提供していないもので、民間企業などが提供しているというものの方が、むしろ世の中には多くある。

では統計の場合、なぜ政府による提供が始まったのかと考えてみると、その作成費用が非常に大きく、その一方で統計を利用することにより便益を受ける人から代金を回収するのが難しいという事情があったからだろう。例えば物価統計について言えば、さまざまな商品の価格を調べるために調査員を店舗に派遣し、値札を見たり店舗から聞き取りをしたりする必要がある。しかもこれを全国規模で展開しなければならない。調査員が1カ月に集める値札情報は25万件にも上る。この膨大な作業を民間企業がビジネスベースで行うことは不可能だ。他の統計についても事情は同じだ。統計作成を取り巻くこうした環境は長らく変わらなかった。しかし最近になって大きく変化しつつある。物価統計を例にとれば、商品の値段を知るのに調査員を店舗に派遣する必要はもはやない。各店舗が作成する売上台帳はかつては紙ベースであったが今はデジタル化されており、それを入手しそこから値段の情報を取得することができる。調査員方式に比べその方がはるかに安上がりで、ミスも少ない。

例えばTポイントカードを発行・運営するカルチュア・コンビニエンス・クラブ株式会社(CCC)には5000万人のカード保有者の購買履歴がほぼリアルタイムで入ってくる。これは、利用者自らが店頭でTポイントカードを提示することにより購買情報の送信を起動していると見ることができる。その意味で利用者の一人一人が貴重な情報を発信するモニターである。これがあれば利用者や店舗から情報を聞き取り調査する第三者(調査員)はもはや不要だ。

技術環境の変化は物価に関する情報に限らない。例えば、求人サイトの賃金や求人数などのデータを使えば賃金統計などに必要な情報を収集できる。求人サイトには必要なスキルが記述されているので、どのスキルに対する需要が高いのか、そのスキルに対する価格(賃金)はどう変化しているのか、海外の水準と比べてどうなのかといったことをほぼリアルタイムで読み取ることができる。厚生労働省等の作成する現行の労働統計では不可能だ。また、不動産の仲介ビジネスに蓄積されている物件ごとの

データを使えば不動産の売買・賃貸価格を高精度で計測することが可能であり、金融緩和が不動産価格に及ぼす影響をリアルタイムで観察したり、不動産バブルの芽を早期検知したりすることが可能になる。

3 統計デジタル化の具体例

これらの例に共通するのは、民間企業が自らのビジネスを遂行するために収集したデジタル情報は、今までの政府統計よりもはるかに詳細な場合があり、それを政府の統計作成に活用できる可能性があるということだ。「統計のデジタル化」と呼んでもよい。

では、統計のデジタル化は本当に実現可能なのか。筆者は消費者物価指数の推計手法に関する研究の一環として、各国で採られている手法について統計部署の担当者と意見交換する機会が多い。筆者の印象では、欧州の国々はデジタル化に熱心に取り組んでおり、実用化に成功している例も少なくない。

例えばスイスでは連邦統計局が大手スーパーの協力を得て、商品価格のデータを提供してもらい、それをを用いて消費者物価指数を作り始めている。スイスも従来は調査員が店頭で足を運んで商品の価格を調べていたが、その手間が大幅に削減された。店舗側にも、日中の忙しいときに調査員の相手をしなくてすむというメリットがある。スイス統計局ではこうして浮いた人手をデジタル化の難しい分野に回したり、オンライン商取引などこれまで手つかずの分野の統計開発に充てたりして、統計の充実を進めている。同様の取り組みはスウェーデン、オランダなどでも進められている。

スイスなどの取り組みで興味深いのは、データ取得の方法の変更が統計の連続性に与える影響を極力なくそうとしていることだ。例えば、スイス統計局が現在入手しているデジタルデータに含まれている商品の数は以前の何十倍にもなる。逆に言えば、調査員による調査のときには世の中にある商品の数パーセントしか観察することができなかった。せつかくすべての商品に関する情報を入手したのだから、この際、全商品の情報を用いる物価統計に刷新すればよいと思うのだが、実際には、調査員の時代と同じ商品のデータしか使わずその他は捨てることにより、過去の統計との連続性を維持しようとしている。そこまでして統計の連続性にこだわるべきか否かは議論の分かれるところかもしれないが、少なくとも、デジタル化を進めると連続性が犠牲になるという、しばしば聞かれる意見は当たらない。

また、統計のデジタル化はスイスのような小回りの利く小国では可能だが大国では難しいとされてきた。しかしその通説も最近、覆りつつある。米国の労働統計局（Bureau of Labor Statistics : BLS）は、アパレル商品の価格データを民間企業から入手し、それをもとに消費者物価指数を作成することを、2019年3月の公表値から始めた。今後、ガソリンなど他の品目についても民間データへの切り替えが計画されており、民間データのカバー率は消費者物価の30%程度に達する見込みである。

デジタル化の具体例をもう1つ紹介しよう。筆者の研究室では、日本経済新聞社の協力を得て、2013年春から「東大日次物価指数」（東大指数）を配信してきた（2016年より株式会社ナウキャストに移管）。これには物価指数の計測手法に関する研究成果を社会還元するという目的とともに、統計のデジタル化の実証実験という側面もあった。

どの店でどの商品がいくらで何個売れたかという情報を、日経と契約した約300の店舗がその日の夜に日経のサーバーに送り、それをさらに東大のサーバーに転送する。東大では、われわれが開発したアルゴリズムを用いて物価指数を計算し、それを翌々日にはホームページにアップする。この間、東大側は完全に無人で、誰かがデータに触るのは年に数回起きるシステムの不具合のときだけだ。物価指数の更新頻度も、総務省の消費者物価指数が月1回なのに対して、われわれの指数は毎日更新される。

スタートした当時われわれが扱っていたのはスーパーで販売される品目に限定されており、総務省の消費者物価統計がカバーする品目の約2割だった。それらの品目の物価指数を作成するために総務省とその作業を実際に担当する各県でどれだけのマンパワーが割かれているのかは不明だが、その部分だけでも無人化できればマンパワーを大幅に削減できるはずだ。スイスのように、デジタル化が容易でない分野にその人材を回すことにより、省庁全体として、統計作成の生産性を向上させることができる。

4 統計の民営化は時代の流れ

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一つまり統計の民営化が時代の流れだろう。

ただし、進むべき方向が民営化だとしても、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、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、先述の通り、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まちまちとなっている。

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」とも呼ばれている。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。

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、すべ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。そんな混乱が生じるくらいなら民間が手出ししない方がよい。東大指数の配信を開始した当初、そうした意見をしばしば頂戴した。

議論を少し整理すると、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って、容易でない作業を迫られることになる。

この手間をどう捉えるかだが、1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。

複数の指標があると余計な手間がかかるように見えるが、実際にはその手間は経済の状況を正確に知る上で十分コストに見合うものだ。政府だけが統計サービスの供給者であれば煩雑な作業は不要だ。しかし政府の出す統計も無謬ではなく、今回の毎月勤労統計の事例で明らかになったように、間違った方向を指し示す可能性がある。そうなったときに政府の統計しかない、その統計が間違っているか否かさえ評価できないという事態に陥る。

太平洋戦争時の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。今であれば人口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。そのため、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はずだ。

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。

実例を1つ挙げよう。図1では、1990年代前半における東大物価指数の動きを総務省統計局の作成する消費者物価指数（CPI）と比較している（東大指数は太線、CPIは細線）。ここでは、総務省が公表しているCPIの品目別指数を用いて、東大指数で採用されている品目（食料品と日用雑貨）だけから構成されるCPIを作成している。

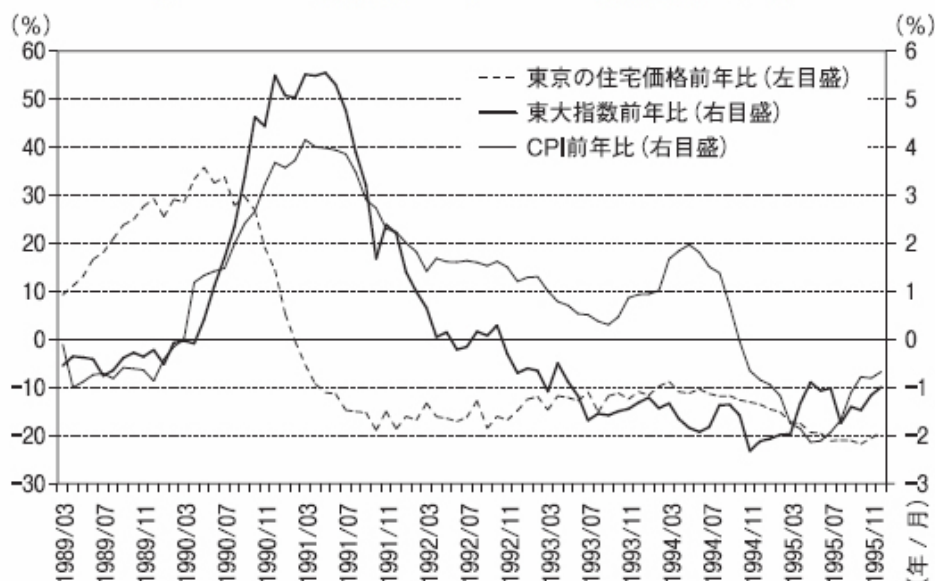
図1の破線は住宅価格の前年比であり、1990年夏をピークに低下していることがわかる。そこから1年ほど遅れてCPIと東大指数の前年比の低下が始まっている。ここで注目したいのはこれらの指数の前年比がマイナスになるタイミングである。CPI前年比がマイナスになるのは1995年初だ。つまりCPIでは1995年からデフレが始まった。一方、東大指数の前年比がマイナスに突入するのは1992年春だ。つまり、CPIと東大指数ではデフレ突入のタイミングが約3年違っている。

この違いが何に起因するのかを知るには2つの指数の原データを子細に比較する必要があり、非常に難しい。しかし、不動産バブルの崩壊から5年後にようやくデフレが始まったというCPI統計の結果は経済理論では説明しにくい。消費の多くが大型店にシフトするという流通構造の変化をCPI統計が捉え損ねた可能性がある。

今から振り返ると、この時期のCPIの動きは日本経済のその後を決める大事な意味をもっていた。CPIでみると、1992～1993年の前年比は1%から2%の間であり、多くの中央銀行が理想とする物価上昇率の水準に近かった。実際、1993年2月の講演で日銀の三重野康総裁（当時）は「インフレとの戦いという点ではわが国は非常にうまくいっている」と語っている。

バブル崩壊後の日銀の政策については金融緩和が遅かったと指摘されることが多いが、その背景には物価に関するこうした認識があったと考えられる。日銀が本格的な金融緩和を開始するのは1990年代末であり、ここで後手に回ったことがデフレを長期化させた可能性が高い。20年以上経ってから「もし」と言ってみても仕方がないことだが、もし日銀がデフレ突入を早期に検知できていれば、その後の金融政策も、そして日本経済の軌跡も変わっていたかもしれない。

図1 バブル崩壊時におけるCPIと東大指数の比較



注) 東大指数は、同指数の作成アルゴリズム (特許第6395307号「物価指数推定装置、物価指数推定プログラム及び方法」) に関する株式会社ナウキャストとのライセンス契約にもとづき、同社に移管され、現在は「日経CPINow」として配信されている。

6 政府統計の信頼回復に向けて

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なインフラだ。しかし、今回の不正でそれが損なわれてしまった。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今回の不正の原因であるとして、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の急速な進歩に伴い大きく変容しつつある。昭和の統計制度に戻すという選択肢はないと考えるべきだ。

統計インフラの再建には、民間のもつ大量のデジタルデータと優れたノウハウを有効に活用することが不可欠だ。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。政府統計の信頼回復に向けて、官民の密接な連携を期待したい。