

政府統計の信頼回復に 民間活力の導入急げ

東京大学大学院
経済学研究科教授

渡辺 努

■信頼が大きく失墜した 政府統計

厚生労働省が作成している毎月勤労統計の調査が2004年以降、不適切に行われていたことが発覚し、政府統計への信頼が大きく揺らいでいる。不正を一言で表せば、東京にある常用雇用者500人以上の規模の事業所1500の全部を調べるのが大変なので手を抜いて500しか調べなかったということだ。単純な手抜き以上の悪意があったか否かは筆者にはわからないが、手抜きの意図があったことは間違

いない。

毎月勤労統計の調査は厚生労働省が用意する1枚の調査票を、企業に埋めてもらうことで行われる。東京の500人以上の規模の企業であれば、ほぼ間違いないく、各人の給与額などのデータはデジタル化されている。企業の担当者は自社のデータベースから必要な数字を取り出し、それを基にその企業の給与総額など厚生労働省が求める数字を計算する。それを調査票に転記し、厚生労働省に郵送する。調査票を受け取った厚生労働省は数値をデータ化した上で、コンピュータ上で集計作業を行う。こうい

った手作業が毎月行われている。

企業の担当者が最も嫌な作業は、企業内で既にデジタル化されている情報をわざわざ調査票に手書きで転記する部分だ。ここに重い負担を感じ、多くの企業が回答を渋る。厚生労働省も嫌がる企業に無理やり回答させるわけにもいかず、結果として調査企業数を減らさざるを得なくなる。不正の背景にはこうした事情がある。

企業側でデータをデジタル化しているにもかかわらず、それを直接省庁に渡せない非効率な状態は、実はさまざまな統計で起きている。例えば、筆者

が頻繁に使用する消費者物価統計の場合、店舗側ではどの商品がいくらで何個売れたかをデータとして保有している。少し先進的な店舗であれば、その商品の購入者がどういう属性の人で過去にどのような購買履歴があるかもデータとして保有し、マーケティングに活用している。

しかし、消費者物価統計を作成する総務省統計局がデジタルデータを利用することはほとんどない。総務省の手法はICT（情報通信技術）普及前と同じであり、調査員が店舗を訪問し、店頭の価格を目で見て、それを手元の調査票に記載する。

筆者の手元には昭和25年当時の調査票があるが、現在の調査票と基本的には同じだ。例えば、胃腸薬の項目をみると、当時の「太田胃散缶入り81g」が今は「第一三共胃腸薬細粒」に代わっているが、違いはそれだけで、今も昔も調査員が店頭で胃腸薬の値段を調べて

来るというルールになっている。

不正の原因に、統計作成に携わる人員の削減を指摘する声が少なくない。財政再建のために人員削減が行われ、そのために統計作成の現場が疲弊しているのは事実である。しかし、少ない人数でやれないのかと言えば決してそうではない。

既にデジタル化された企業のデータを省庁がそのまま受け取ることができれば、大幅な省力化が可能だ。それを阻んでいるのは、昔ながらの調査手法を変えようとしないうる省庁の姿勢だ。

■統計作成のデジタル化が 急務

では、統計のデジタル化は本当に実現可能なのか。筆者は消費者物価統計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、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済むメリットがある。スイス統計局は、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。

米国では、労働統計局が今年3月より、アパレル商品などを対象に流通業者からデジタルデータを入手し、それを用いて消費者物価統計を作成するこ

とを始めた。

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されていることだ。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。

せっかくなかたきさん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だが、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って、しばしば聞かれる意見は当たらない。

デジタル化の具体例をもう一つ紹介しよう。筆者の研究室では、日本経済え、いまなお膨大な数の調査員を抱え、統計作成にかなりの人件費をかけている。また、調査員が日中の忙しいときに店舗を訪問することで報告者の側にも費用負担が発生している。データ購入費用と比べるべきは、これら全ての費用の合計である。消費者物価について言えば、筆者の概算では、デジタル化でおつりがくるはずだ。

■民営化は時代の流れ

今回の問題で明らかに変わったのは、省庁は民間に比べて統計作成に強みがあるとは言えないということだ。統計の基になるデータを握っているのは誰かと言えば民間企業で、データの加工作業も民間の方がはるかに長けている。

もちろん昔からこうだったわけではない。昭和やそれ以前の時代であれば、統計の基の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

新聞社の協力を得て、13年春から「東大日次物価指数」を配信してきた(16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、東大側は完全な無人で、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。

当時われわれが扱っていたのはスーかった。また、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の進展に伴って、政府と民間の立場が逆転している。

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、そこに一足飛びにいくというのは非現実的である。民営化の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。データ収集の

パーで扱う品目に限定され、総務省の消費者物価統計がカバーする品目の約2割だ。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。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。

この話を省庁ですると、最初に出る反応は、民間のデータは高くて手が出ないので実現性が低いというものだ。確かに東大指数で使っている購買履歴データは、商用利用を前提に値段がついているので高い。省庁の理屈で言えば、いまはタダで取得できている情報に高額を払うのは筋が通らないということかもしれない。

しかし省庁は、近年減少したとはいアウトソーシングは一部の統計で既に始まっているが、今後、さまざまな統計についてこうした動きが出てくるだろう。

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、次の論点をクリアする必要がある。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だ。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、先述のとおり、物価統計については、欧州のいくつかの国で民間流通企業のデータを使うことが始まっているが、これらの国も試行錯誤の段階であり、有償か無償かについても各国で別々となっている。

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があり、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、カードを使う人の購買情報は入っていない人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ふ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などの研究機関が中心となって進める必要がある。

■複数指標で政府の監視も可能に

で明らかにしたように、間違った方向を指し示す可能性はある。

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事態に陥る。そのリスクを軽減できるといふ意味で、複数の主体による複数の指標の公表は望ましいことである。

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で、遠隔地の戦況を伝える民間企業がたくさん出てくるであろう。当時、それができれば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事態にはならなかったはずだ。

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

統計民間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点だ。

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るれば問題ないが、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ほ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。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という意見もしばしば聞かれる。

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。

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。

■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のインフラだ。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は異論がないだろう。

しかし、具体的にどこを修復するかとなると意見が分かれている。政府の統計作成部署の人員を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。ただ、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

問題は、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ことである。計測誤差が全くない指標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ふ、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、一つには、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意味では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、今回の労働統計の事例

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選択肢はない。

統計インフラの再建には、民間の持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。

わたなべ・つとむ

東京大学経済学部卒。経済学博士（ハーバード大学）。19年4月より東京大学大学院経済学研究科長・経済学部長。日本銀行大学院。エコノミストを経て、一橋大学、京都大学、ポツダム大学、コロンビア大学等で客員准教授・教授などを歴任。日本学術振興会の大型プロジェクト「対話型中央銀行制度の設計」の研究代表者。統計改革推進会議委員、総務省統計研究所客員教授、キヤノングローバル戦略研究所研究主幹等を兼務。論文・著書多数。