

経済セミナー2019年6・7月号

特集 統計の役割を考える

[インタビュー]

統計問題の本質

2018年末より大きく報道され、国会でも議論が続いてきた「統計不正」問題。さまざまな方向で議論がなされているが、この問題の原因と、今後必要とされる改革について、統計委員会の委員長を務める西村氏に語っていただいた。

西村 清彦 Nishimura, Kiyohiko G.

政策研究大学院大学特別教授／総務省統計委員会委員長

プロフィール

1953年生まれ。東京大学経済学部卒業、イェール大学経済学部博士課程修了(Ph. D.)。東京大学経済学部助教授・教授、内閣府経済社会総合研究所総括政策研究官、日本銀行政策委員会審議委員、同副総裁、東京大学大学院経済学研究科長・経済学部長、政策研究大学院大学教授等を歴任し、2018年より現職。2014年より総務省統計委員会委員長を務める。2015年、紫綬褒章受章。著書: Economics of Pessimism and Optimism: Theory of Knightian Uncertainty and Its Applications, (coauthored) Springer, 2017(2018年度、第61回日経・経済図書賞受賞)。Information Technology Innovation and the Japanese Economy, (coauthored) Stanford University Press, 2010(日本語版は2005年度、第21回テレコム社会科学賞本賞受賞)。

1 統計というインフラが破壊された

—— 2018年12月、厚生労働省が行っている毎月勤労統計調査¹⁾において、全数調査すべき従業員500人以上の事業所に対して、実は東京都では約3分の1の規模での抽出調査がなされていたことが発覚しました。その後、この問題は「統計不正」として、メディアや国会でもさまざまな側面から大きく取り上げられ、今日に至っています。この問題をどのように捉えていますか？

西村 端的に言って、日本の政府統計は危機的状況にあると捉えています。ところが、そのような状況にあることをほとんどの人が認識していません。

私は、「統計」という社会インフラが壊れた状態にあると深刻に受け止めています。比喻を用いれば、東京を出た東海道新幹線が、橋が落ちたために静岡くらいで止まってしまい、名古屋にも、大阪にも行けないような状態になっているということです。

本来ならば、国民の怒りが爆発してもおかしくない状態なのですが、統計というわかりにくいインフラであるために、現状は一見何事もなかったかのような状態になっています。危機的状況にありながら、多くの人々が危機感を感じていない。しかし実際は、経済をよくするために政策を打とうとしても、そもそも経済が現在どのような状態に置かれているのかすら正確にはわからないという、大変な事態に陥っているのです。

なぜこういうことが起こったのかを考えると、まずは統計というインフラ作りを担っていた人員がどんどん削られ、さらに現場がインフラをしっかりとメンテナンスし、より安全で安心なものにしようというインセンティブが働かないような人事システムになっていたことが挙げられると思います。統計データを作る人たちに適切なキャリアパスが整備されておらず、彼らは閉鎖的な環境で外の状況変化に関心も持たず内向きに仕事をしていればそれで済むという中で、統計データが作られてきたのが現状です。

「毎月勤労統計調査」(毎勤)については、それに加えて、言わば「橋にボルトを取り付けた際に場所を間違えた上に、取り付けたボルトが腐食していた。そして、ボルトを付け間違えたことすら、一部の人を除いて知らなかった」という状況です。

2 フィリップス・カーブが作れない

西村 このように今回の統計問題は、重要な社会インフラが毀損したのと同じことなのだと言われたいです。私は経済学者で、経済統計のユーザーでもあるわけですが、広く政府や外部の統計のユーザーたちも、それが使えなくなってしまったのです。当座は、橋が落ちた影響を船などを使ってしのいで、日常生活を送れる程度にはなっているけれども、さらに遠くへ行こうとすると通行できなくなっている（毎勤の場合は、過去にさかのぼろうとするとデータがない）ため、重要な分析などができない状態なのです。たとえば、名目賃金の変化率と失業率の関係を示す賃金フィリップス・カーブを作ろうとしても、適切な形で作ることができなくなっています（図1の「消えたフィリップス・カーブ [1991~2012年]」参照）。これは何としても解決しなければなりません。

さらに、毎勤で過去にサンプル入れ替え時に大きな数値の振れが生じたのを直すためにローテーション・サンプリングに移行したのですが、今年に入って一部のサンプルの入れ替えが行われたところ、振れが小さくなるどころか大きな断層ができてしまいました²⁾。サンプルを入れ替えた2019（平成31）年1月の毎勤によると、賃金の伸び率がマイナスになっています。これまでは、振れは大きいもののフィリップス・カーブはそれなりの形を描いていました。しかし今年1月になって、賃金上昇率がトレンドから外れて大きく下がるような動きをしてしまいました（図1の「踏み外したフィリップス・カーブ [2019年]」参照）。

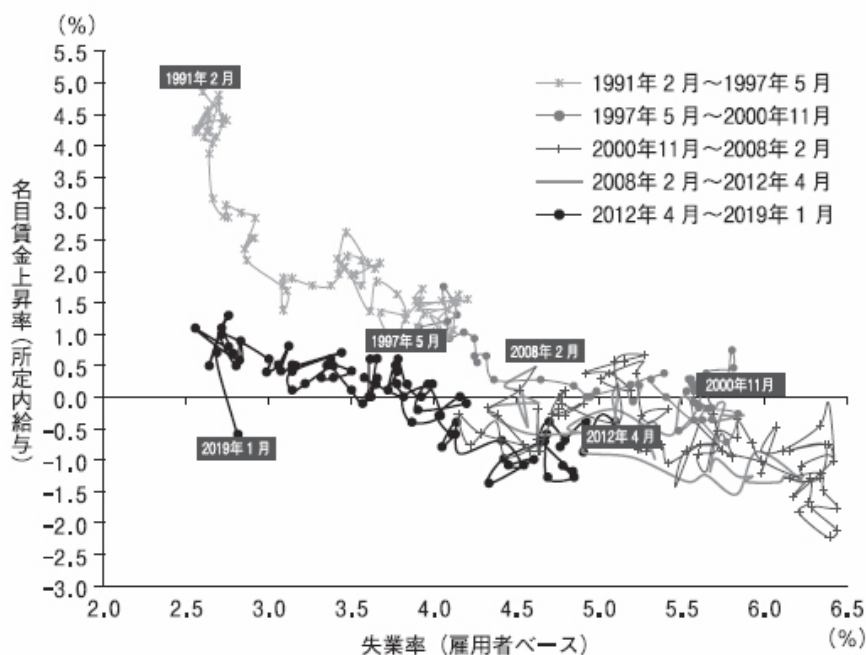
従業員数500人以上の大きな事業所の間には賃金についてかなりのばらつきがあり、これらの事業所の調査を全数調査から抽出調査にしてしまうと、サンプルを入れ替えたときにそれによる大きな変動が生じることがあります。全数調査をしていけばその問題は起こりにくくなります。賃金がわずかしか上がらない近年の日本の状況下では、毎勤は賃金の上下動を測るために必要な精度を確保できていないと言わざるを得ません。

したがって景気指標としては、共通事業所のデータを見て判断する方が無難になります。ただし、共通事業所はサンプルサイズが小さいことに加え、標本の脱落バイアスの影響などを受けることを考慮しなければなりません。

毎勤は、足元の景気を見るための重要な統計です。最近では報道が少なくなり、毎勤

の問題は収束してきたかのように見えますが、問題が解決されたわけではなく、橋は落ち、陸上では移動できなくなったけれど、船は動いているので日常の生活はどうかできるという状態にあるにすぎません。経済状況を分析しなければならない人たちにとっては、真っ暗闇の中を方向も距離も速度もわからないまま、勘を頼りに進むしかないような状態が続いていると言ってよいでしょう。

図1 消えたフィリップス・カーブ (1991~2012年)、踏み外したフィリップス・カーブ (2019年)



3 ガバナンスを欠いた厚生労働省

西村 統計委員会が統計改革の一環として、毎勤について調べ始めたことで、今回の問題が発覚したのですが、「毎勤があやしい」ということは以前からたびたび指摘されてきたことでした。それまでは、橋桁が落ちたのに、こっそり何食わぬ顔をして応急の橋桁を掛けたものですから、皆、何か変だなとは思いつつながら橋の上を走り続けていたわけです。

重要なインフラが壊れていることが隠され続けてきたことが最も重要な点だと思います。しかし、厚生労働省に隠ぺいの意思が明確にあったかどうかすら、疑わしい状況です。むしろ、統計への関心の欠如から、何が重要で何が重要でないかもわから

ずに、担当者のミスや報告なしの変更を、そのまま放置してきたと言えるでしょう。

この問題に関しては、厚生労働省に組織としてのガバナンスが働いていたとは思えません。橋を作るとはどういうことなのか、ボルトをちゃんと締めるとはどういうことなのか、ボルトの締まり具合を点検するとはどういうことなのか、という基本的なところから見直さなければならない深刻な状況です。

毎勤の場合、府省庁間の分散型統計システムの欠点が典型的に現われてしまったとも言えます。誰も知らないところで、橋を作り、勝手に手を入れ、必要な修理をしてこなかったことが、今回の問題の内実です。組織や制度改革を含めた徹底的な統計改革を進める必要があるのです。

4 統計改革の方向性

—— では具体的に、どのような方向で改革を進めるべきなのでしょう？

西村 まず、きちんとした橋（統計）を作る人材の確保、その人材が適切なインセンティブをもって働ける組織作り、そして問題が起きたときに自分たちで直せるかどうか、直せないのであれば、直せる人を呼び込んで対応できるシステム作りです。省庁ごとにバラバラに統計を作っているいまの仕組みを改めて、最低限のコーディネーションを司る中枢的な組織作りが求められます。

ゼロから新しい仕組みを構築できるのであれば、やりやすいのかもしれませんが、すでにある仕組みを動かしながら改革を進めていく必要がありますので、さまざまな困難があります。厚生労働省だけを見ても、労働以外に福祉・社会保障関連の各種統計も扱っており、それぞれが問題を抱えていますから、改革は一筋縄ではいきません。統計委員会では、これまでも国民経済計算（GDP 統計）を中心にさまざまな改革を行ってきました。大きな改革の1つとしては、統計調査のカバレッジを広げることです。

たとえば企業活動の統計調査の場合、企業・事業所の母集団の名簿を用いて行います。経済センサスで捕捉している企業・事業所の数は約180万社です。ところが、国税庁に法人税を納めている企業・事業所は約260万社あります。さらには、法人番号制度に基づいて国税庁が郵便を送り、郵便が届いた企業・事業所は約310万社です。180万社と260万社の差80万社は何だろうか、260万社と310万社の差50万社は何だろうか。こういう問題が出てきます。いずれにせよ、経済センサスが捕捉する180万社では現在の企業数を正確に捉えることができていない可能性が高く、国税庁から

リストをもらえるのであれば、統計調査のカバレッジを広げることができます。

これまで捕捉漏れしていた企業の経済活動が把握できれば、当然 GDP は増えることになります。先ほどは橋が落ちているという話でしたが、今度は「日本という国の本当の大きさがわかっていなかった」という話になるのです。そうした捕捉されていない企業については、実際に人々が働いて付加価値を生み出している場合もあれば、実態のないペーパーカンパニーである場合もあるでしょう。いずれにせよ、これまでわかっていなかった部分が、改革を進めることで次第に見えてくるようになってきました。

これまで日本の GDP の四半期速報 (QE) は、その時点で利用可能だったデータを組み合わせる形で作成されており、各時点で異なった作り方をしてきました。そのために、大きな変動や振れが生じることがたびたび指摘されてきました。しかしそろそろこのやり方から卒業する必要があります。これを「シームレス化」と言っているのですが、各時点において同じ手法で算出したデータを蓄積していくことで、変動を小さくしていこうとしています。

この考え方の根本にあるのが、「経済構造実態調査」です 3)。5年に1度のペースで経済センサスの調査が行われ、それが基準年の数字のベースになります。製造業・非製造業をあわせた経済構造実態調査は年1度行われ、QE ではそれと整合的な生産動態統計を用いて算出します。この三者の推計方法や基礎データのベースを合うようにすれば、統計の振れを小さくできます。そのときに重要となるのが、先ほど申し上げた母集団名簿です。企業は毎年10万社くらい新しくできては倒産していきます。それを反映して、事業所母集団名簿を年次ごとになるべく正確にスクラップ・アンド・ビルトしていかなければなりません。

5 誰も何もチェックせず、15年

西村 実は、こうした統計改革を進める中で発覚したのが毎勤の問題だったのです。毎勤はずっと長い間、統計委員会およびその前身の統計審議会の審議の対象になっていませんでした。2009年に新しい統計法が施行される以前は、同じ手法で統計を取り続けている限り、審議の対象とする必要はなかったからです。毎勤は1990年に調査対象を30人以上の事業所から5人以上の事業所に拡大して以降、まったく見直しを

してきませんでした。

ところが、その間に東京都の従業員数 500 人以上の事業所を全数調査から抽出調査に申請もせずに変えていたのです。変えるのであれば、定められた手続きで申請し、統計委員会で審議する必要がありました。厚生労働省は統計法が定める手続きに反していたわけです。なぜ変えたのか、本当のところはよくわかっていません。

しかし実は、同様の話はほかにもあるのです。その要因は、政府統計のコンピュータ・システムが大掛かりである上に非常に古いことにあります。その昔(二十数年前)に多額の費用をかけて一気通貫の「素晴らしいシステム」を作ったのはよかったのですが、長い年月を経る中で、途中で何が起きているか誰にもわからない状態になってしまいました。これが、現在の日本政府の情報システムのアキレス腱になっています。さて、毎勤の話に戻すと、二十数年もの間、状況が変化する中で何の対応もしなければ、統計の精度が下がるのは当たり前です。しかし、それらのメンテナンスを担うべき人々が統計の精度に関心を持っていなければ、彼らは何もしない方が楽なので、そのままになってしまいます。統計委員会では、4 年ぐらい前から、まったく諮問されていない統計を 1 つひとつ取り上げて検証するようになりました。毎勤もその作業の中で、ローテーション・サンプリングを導入したり、共通事業所系列を作ったりといくつかの改善がなされました。その後もフォローアップを続け、その中で 15 年前に行われた不正が明るみに出たということです。

私が理解できなかったのは、内部でのチェック機能がまったく働いていなかったことです。どんな仕事だって、システムを変更したら、責任者がそれを理解し、ちゃんと動くかどうか、チェックするでしょう。ところが、誰も何もチェックしていなかった。抽出調査に移行するのであれば、復元処理をするのが基本中の基本であるにもかかわらず、それすら行われていない。証言によると、復元のプログラムを作るように指示した人がいるらしいのですが、結局は復元されたかどうかをチェックしてきませんでした。これでは、指示したことになりません。

6 不足するデータの専門家

—— 外部の研究者がアクセスできるように個票データを公開していれば、今回のようなミスに気づく可能性は高まったのではないかという指摘もありますが。

西村 それはそうですが、個票は個人情報の固まりです。情報を公開するには、セキュリティが十分に担保されていなければなりません。ある個票データを別の個票データに組み合わせることで、個人を特定できてしまう可能性が高まります。仮に、ミネソタ大学などが運営する IPUMS International 4) のような機関に情報を提供する場合も、日本の統計法上、生のデータを提供することはできません。

1つの方法としてはノイズを入れた匿名情報として提供することです。それでも、ノイズをかなり大きくしておかないと特定化されてしまう。しかしノイズを大きくすると、今度は実態から乖離して、何のためのデータなのかという話になる。あるいは個票から再度ランダムにリサンプリングするという方法もありますが、これもデータが実態から乖離してしまう可能性があります。それらのデータにどれほどの意味があるのか。そういうデータを使う場合は、その程度のものだということをふまえた上で分析・解釈しなければなりません。

こうしたセキュリティ等への関心が比較的薄い国は積極的にデータを提供していますが、今度はそういう国の回答者がはたしてまじめに回答しているかどうか、という問題にぶつかります。たとえばアメリカはデータについて開放的ですが、背景にある事情はさまざまなので、「アメリカがやっているから日本も見習うべきだ」というわけにはいきません。

—— 最後に、経済学者等の専門家は統計にどのように関わっていくべきだとお考えでしょうか？

西村 私も経済学者ですが、経済学者は統計データのユーザーであって、統計の専門家、データの専門家ではありません。いまわれわれが必要としているのは、データの専門家です。どのように調査したらよいかを熟知し、その調査がその他のデータと比べてどのような特徴を持っているかを理解しているデータ・サイエンティストが不足しているのです。

しかしだからといって、データ・サイエンティストという一群の人たちを特別に養成すれば解決する問題だとは思いません。それよりも、経済学者や社会学者などのユーザーとしての専門家、いま述べたデータの専門家、政策立案者、さらにはビジネスサイドの人たちとの対話を通じて、ユーザー、メーカー、データ・サイエンティスト三者間のシナジーを形成することがまず求められていると思います。データ・サイエンティストが不足していることは事実ですが、むしろ、ユーザーやメーカーの側がデ

ータ・サイエンティストになることを目指すべきです。

経済学者も、統計データのユーザーであり続け、ただデータがほしいと言っていればよい時代ではなくなったということです。データがないから分析ができないという言い訳は、これからは通用しないでしょう。また実験データなど、経済学者も自分でデータを作る時代になっています。自分でデータを作る能力がないと、データを使うことだって本当はできないのです。そしてデータを使わないと作る意欲もわかない。したがって今後は、関係者みんながデータを作る社会になっていくと思います。そして、そうやって作られたデータを国の情報資産の中に入れて大きなデータベースを作っていかなければなりません。

とはいえ、その中でも政府統計はクオリティの高さやカバレッジの広さといった観点から格別の位置を占めており、今後も中核的な役割を果たすものであることに変わりはありません。ですから、確固たる仕組みを構築しておく必要があるのです。今回明らかとなった統計問題は、日本の政府統計に対する国内外の信頼性を根幹から揺るがす大きな問題であり、改革の必要性を改めてわれわれに突きつけたものだったと言えるでしょう。

[収録日：2019年4月15日収録]

注

- 1) 毎月勤労統計調査について、詳細は本誌 pp. 41-45 (労働統計と経済学) 参照。
- 2) 厚生労働省「毎月勤労統計調査(全国調査・地方調査): 結果の概要」(<https://www.mhlw.go.jp/toukei/list/30-1a.html>) 参照。
- 3) 経済構造実態調査は、5年ごとに実施する「経済センサス-活動調査」の中間年の実態を把握することを目的に毎年実施する調査。2019年から開始される調査で、サービス産業動向調査(総務省)、商業統計調査(経済産業省)、特定サービス産業実態調査(経済産業省)を統合・再編し、製造業やサービス産業の実態を明らかにするとともに、国民経済計算(GDP統計)の精度向上に貢献することを目的としている。
- 4) IPUMS International について、詳しくは本誌対談 p. 11 を参照。