



エビデンスに基づく政策形成と経済学

川口 大司

(東京大学教授)

2016 年の後半からエビデンスに基づく政策形成 (Evidence Based Policy Making, EBPM) の重要性が永田町や霞が関といった政策形成の場で話題になるようになった。EBPM とは限りある政策資源をどの政策に割り振るかという意思決定の過程の中で、エビデンスに基づいて効果の上がっている政策を特定し、その結果に基づいて政策資源を配分しようという考え方である。医学の世界で、医者感覚ではなくて臨床実験の結果に基づいて治療法の選択を行おうというエビデンスに基づく医療という考え方にその起源をもつといわれている。

EBPM の盛り上がりに応ずる形で東京大学大学院経済学研究科にも政策評価研究教育センター (CREPE クレープと発音) が 2017 年 10 月に立ち上がった。センター長に計量経済学者の市村英彦氏、副センター長に筆者が就任した。労働問題の初学者を対象としたこの論文では、CREPE での 1 年半の経験に基づいて研究者の立場から見て EBPM とは何かを紹介したい。そのうえで、その可能性と乗り越えるべき課題を整理してみたい。

I 従来の政策決定との違い

EBPM という話をきいて、それはいったいこ

れまでの政策決定過程と何が違うのかという疑問を持つ方も多いと思う。公的機関の資源配分には、政策実施部局からの予算の要求、財政部局からの予算の査定、予算の執行、監査という流れがあり、特に予算要求の過程においては推進しようとしている政策が重要であることを、何らかの証拠に基づいて論証するという過程が入り込む。実際、世の中には予算請求のためのパワーポイントで作成したポンチ絵なるものがあふれている。たくさんの枠があって、その枠が因果関係を示す矢印でつながっていて、何となくうまくいきそうだなとか、なんだか矢印が強引で筋が悪いなどと思わせてくれるあの図である。そして、ポンチ絵には日本の将来の労働力人口の予測などが入り込んでいる。これはエビデンスではないのか？ あの図に基づき、予算請求が行われ、査定が行われるという図式は EBPM ではないのか？

EBPM は今までのポンチ絵ワールドと何が違うのか。私見では、二つの大きなポイントがある。一つは政策介入が政策目標の達成にどのようにつながるのかという論理構造を持つ点である (ロジックモデルなどといわれる)。たとえば、ある大学部局が大学本部に留学生対応のためのアドバイザーを雇用する予算を請求するとしよう。従来のポンチ絵ワールドでは、その部局の留学生数の推移を示す折れ線グラフが示され、それに伴って発

生している問題が列記され、その対応のために留学生アドバイザーが必要だと論じられる。枠囲みの中の D.K. さんが顔写真入りで「留学生アドバイザー中村さんのおかげで、充実した日本での生活を送ることができました」なんて書いてあったりする。少なくとも D.K. さんは留学生アドバイザーの提供してくれたサービスに満足しているらしいことを示すエビデンスである。このエビデンスの問題点は、この D.K. さんの意見は留学生の代表的な意見なのか、そもそもこの D.K. さんはどのように選ばれたのか、充実した生活とは何か、中村アドバイザーではなくて鈴木アドバイザーの場合はどうだったのか、などなどであり、今後、留学生アドバイザーを増やすと留学生の学習がより良いものになるのか今一つ確信を持ってない。当然のことながら本部職員が真顔で切り込んできて予算がつかない。これが EBPM の世界になると政策介入とその結果を明確にしようというロジックモデルを作ることになる。例えば、留学生アドバイザーの数が増えると、留学生が直面する生活上の困難を解消できる可能性が高まり、勉強に集中できるようになり、留学生の成績が上がるといったロジックである。このような論理構造が明確になると、留学生アドバイザーの数と留学生の成績という 2 つの指標に注目してデータを集めてみようということになる。データがカバーしている範囲は全体であるし、数値化している点で政策介入も結果も客観的かつ明確に定義されている。

ポイント 1：EBPM では政策介入と政策結果を明確に定義し数量化する

留学生アドバイザーの数と、留学生の成績のデータを集めてみたでしょう。ここで留学生アドバイザーの数と留学生の平均的な成績はともに上昇傾向にあったでしょう。このような状態のことを留学生アドバイザーの数と留学生の平均的な成績の間には正の相関関係があるという。しかし、この正の相関関係は留学生アドバイザーを増やすと留学生の成績が上昇するという因果関係があることを意味しない。たとえば、その大学の留学生

からの人気が高まり、必要に応じて留学生アドバイザーを増やしていった一方で、そのこととは無関係に留学生向け入試の競争が激化し質の良い留学生が入学するようになっただけなのかもしれない。このような相関関係が観察されているだけだとすると、留学生アドバイザーを増やしたとしても留学生の成績は上がらない。経済学者が折に触れて相関関係と因果関係は違うと強調するものだから、耳にタコができたという向きもあるだろう。しかし、あえて言いたい。ここで目的は留学生アドバイザーを増やすための予算をとってくることにあるのだから、大切なのは因果関係だと。

実をいうと観察されたデータから因果関係を読み取るのはとても難しい。ある年から留学生アドバイザーが増員され、その年から留学生の成績が上がっていたとしても、先に述べたように、その年から留学生の入試競争が激化し、留学生の質が上がっているのかもしれない。このような状況に対応するためには様々な方法があるが、たとえば、同じ外国大学を卒業した入学年度が異なる学生数名のデータをとってきて、留学生アドバイザーが多かった時期に入学した学生と少なかった時期に入学した学生の 1 年生修了時の成績を比較するといったことが行われる。観察されたデータから因果関係を読み解くにあたってのキーワードは比較である。政策介入を受けた状態の結果と、政策介入を受けなかった状態の結果を比較することで政策が結果に与える因果関係を知るという考え方である。ここで違う人の結果を比較してもしようがないので、理屈の上では D.K. さんという留学生が留学生アドバイザーからあまりアドバイスが受けられない状態といつでもアドバイスを受けられる状態を比較することが必要になる。しかし、D.K. さんの人生は一回きりであり、彼はどちらかの状態しか経験していないことになる。仮に D.K. さんが留学生アドバイザーのアドバイスをいつでも受けられる状態にあるとすると、留学生アドバイザーのアドバイスがほとんど受けられなかった状態の結果は観察することができない。これがデータから因果関係を読み解くにあたっての根本的な問題だ。この問題を反実仮想という仮に D.K. さんがアドバイザーからのアドバイスが

あまり受けられなかったときに実現したであろう状態を推測することによって乗り越えていくのが因果推論のテクニックだ。これについては中室・津川（2017）や伊藤（2017）といった入門書がわかりやすく解説している。

ポイント2：EBPMでは政策介入が政策結果に与える因果関係を推定する

ポイント3：因果関係を推定するにあたっては、政策介入があったときとなかったときの結果の比較をする。

3つのポイントに示したように、予算請求の資料につけるエビデンスをより系統だったものとするのがEBPMの従来の政策決定との違いである。政策を実行することで結果が伴うということをやより説得力のあるエビデンスを使って示すということであり、説得材料の質を上げることがポイントになるわけなので、今までの政策決定の流れと違いはないともいえる。違いは因果のストーリーをサポートする、データの質、因果推論の質を上げることであり、地道な取り組みだ。しかし、地道に説得材料の質を上げていけば、その小さな滝はだんだん集まり大きな滝となり、政策資源の配分を大きく変え、気づいたら社会の在り方を大きく変えていたということにもなりうる。大切なのは一つ一つの資料のエビデンスの質を上げていくという地に足の着いた地道な取り組みである。

なお、この観察されたデータから因果関係を読み解くという作業は、経済学者が1980年代半ばから精力的に取り組んできたテーマである。今日でも次々に新しい概念や推定手法が開発され、実際のデータに適用され研究が発展しつつある。しかしながら、実際の政策評価の現場で使われる技術は本質的な部分では「枯れた」技術であり、筆者が大学院で勉強していた20年前と本質的な意味では変わらない。学術的には「枯れた」技術を、重要なテーマ、よりよいデータに応用し、新しい知見を作り出すことが重視されていると感じる。

II 統計改革とのリンク

政府の中ではEBPMの推進と政府統計の整備は車の両輪であるかのように語られることがある。しかし政府統計そのものは特定の政策の有効性を示すエビデンスとなりえないことは明らかなので、二つは分けて考えたほうが良い。ただし、エビデンスを作り出すために重要な役割を果たす政府統計の品質を向上させることは、EBPMを効果的に推進していくためには欠かせないのは確かである。

2018年12月に基幹統計『毎月勤労統計』のサンプリングが仕様書通りに行われておらず、結果としてサンプリングウェイトが正しく振られていなかったという問題が発覚した。多くの識者が指摘しているように、統計法に反して勝手にサンプリングを変えた、サンプリングウェイトを適切に割り振らなかった、2018年年初からこそっと正しいサンプリングウェイトに修正したという3点が問題である。法令違反をしながら業務を進めるリスクを行政官は認識しているはずであり、なぜそのような異常事態が起こってしまったのかを調べ、問題点をあぶりだすことが必要である。同時に問題の範囲を画定し、統計不信が政府統計全体に広がっていかないような取り組みも求められているといえよう。

『毎月勤労統計』の不正の発覚までに時間がかかった理由の一つに当該統計の個票を用いた研究を行っている研究者が少なかったことがあげられるだろう。『毎月勤労統計』は事業所単位での賃金支払い総額と、総労働時間が尋ねられているだけで、労働者個人の給与額や労働時間は記録されていない。そのため、学歴・年齢・勤続年数といった労働者属性と賃金の関係といった労働経済学者が関心を持つ分析を行うことはできない。そのため、筆者自身も含めて多くの研究者は『賃金構造基本統計調査』の個票を用いた研究を行ってきた。『毎月勤労統計』で起こったように東京都の大規模事業所のサンプルサイズがある調査時点から一気に1/3になり、振られているサンプリングウェイトが変更されていないとなると、その不

自然さに気付いた研究者がいたとしても不思議ではない。今回の『毎月勤労統計』における不正は、全数調査の対象となっているはずのグループにサンプル替えに伴う平均値の断層があったことをきっかけにして明らかになったという。これは特殊な例であって、個票を見ることなくして、統計調査が設計書通りに行われているかどうかを確認するのは極めて難しい。また、個票を外部の研究者が使う可能性があるということ自体が、外部に説明できないような行動を抑止することにつながるはずである。その意味で、統計の品質向上のためには、政府統計個票の研究者による利用を促進することが重要である。

個人的な経験で恐縮なのだが、大学院生時代にアメリカのNLSY79というパネルデータを使って博士論文を書いていた時に、回答から間接的に計算される勤続年数がある年だけ異様に動いていることに気づいたことがあった。コードブックやデータを何度見直してもよくわからないので、勇気を振り絞って作成しているオハイオ州立大学に問い合わせたところ、担当者がコードを確認してくれて、バグがあったことを発見し、修正したデータを送ってくれたことがある。返信には、「どうもありがとう。こうやってデータをちょっとずつ直してよいものにしていっているのです」というコメントがついていた。日本でもデータを使う研究者とデータを作るものがこのような関係を築けないものかと思う。データを作ったり加工したりという経験がある人々は膨大な作業の中でどうしても確率的に間違いが発生し、ミスはあるものだと思っている。それを前提にして、間違っていた場合に、間違っていたと素直に認めて修正できるような体制を構築する必要がある。

ポイント4：統計の作成にミスは起こる。それを前提に、多くの人が個票を利用し、間違いが見つかれば修正できる仕組みを設計すべき。

なお、今回の不正の根源的な理由となった回答者負担の問題は不正の原因解明とは別に取り組まれるべき問題だろう。一定規模以上の事業所では労働者情報はほぼ電子的なシステムで管理されて

いる。そのため調査票情報の電子的提出を積極的に進めていくことが必要だ。試みに『賃金構造基本統計調査』を調べてみるとe-Govウェブページの情報によれば、2019年2月現在、調査票の電子的提出は認められていない。回答者が電子情報を紙に印刷し、統計当局がその紙の情報を電子化するという悪い冗談のような状況が起こっているのではないだろうか。回答者負担の軽減、統計作成業務の軽減の両方の観点から、電子化を検討すべきだろう。また、厚生労働省は雇用保険や年金といった所得階層に応じて保険料を徴収するシステムを持っている。これらの数字を使って賃金の動きを補足的にとらえるということもできるはずである。民間の求人業者、派遣業者団体などと協力して賃金動向を把握するというのも技術的には可能だろう。複数の統計を組み合わせると一つの系列を作り出すことには、様々な困難が伴うが興味深い課題である。

新しい調査設計を行うためには統計に関する深い専門知識が必要になる。日本には統計の専門家がいないとの指摘もあるが、私の知る限り総務省統計局は統計畑で採用された専門家が多く、局長まで上がっていけるようなキャリアパスが用意されている。また、各省庁には、国内外の経済学、公共政策、ビジネススクール、工学など関連分野で修士号を取得した統計の基礎的な知識を持った職員が一定数いる。彼らの処遇を改善するとともに、高技能の専門家の採用を加速していくことが今後の課題であろう。

ポイント5：世の中の電子化が進んでいることに合わせて調査設計も変えていく必要がある。進めるためには既存人材の活用も含めた高度人材の投入が必要。

Ⅲ 行政と学界の協働

東大CREPEが始動してから中央官庁・地方自治体の方々と仕事をする機会が増えた。以前から役所の研究会に出席するような機会もあったのだが、違いは行政官が大学にやってきて研究会に参加することである。データを提供してもらい、大

学院生にデータ分析をしてもらい、その成果を報告してもらう。それを行政官が聞いてコメントするという構図である。英文査読誌への論文出版が研究業績評価のすべてとなっている経済学の分野において、統計の質の改善のような地味なテーマを研究成果にどのようにつなげるのか、CREPEにかかわる教員の力量が問われている。

一方の行政官の側にも事情がある。短い期間で成果を出すことが求められるし、研究成果は実際の政策決定の中の一要素に過ぎない。統計の調査設計を変えろという一見すると技術的な問題についてすら、調査負担を負う地方自治体との調整の中で、学術的に望ましい統計設計とのバランスを取りながら行政官は仕事をしている。研究者が発表する研究成果について、その発表の時期や形態に敏感にならざるを得ないのは理解できる。

行政と学界が協働していくためには、まずそれぞれ相手が何を目的に行動しているのかを正確に理解する必要がある。初期時点での率直で密なコミュニケーションが必要だ。協働による余剰が発生しないのであれば、その共同研究はしないほうがいい。また、相手の利得構造を理解することは相手の要求を丸呑みすることではない。研究者はどこまでが妥協できる範囲なのかを決める必要があるが、研究倫理に抵触しないか、同業者に自分の判断を説明できるか、といったところが最終的な判断基準になるだろう。密なコミュニケーション

ンはとりながらも緊張関係を保つというつかず離れずの関係を築く必要があり、ケースバイケースで関係構築していくしかない。

ポイント6：行政と学界の共同研究は相手の目的を正確に理解する。そのうえでつかずはなれずの緊張関係を保ちながら仕事を進める。

IV ま と め

始まったばかりのEBPMであるが、地道な活動が続ければ、今後徐々に定着し日本の政策決定の在り方に影響を与えていくであろう。必要以上に悲観することなく、楽観することもなく、深く考えたうえで目の前の課題に取り組んでいくことが、成果を生み出す。そのことが共感を生み、取り組みの輪を広げていくと期待したい。

参考文献

- 伊藤公一朗（2017）『データ分析の力 因果関係に迫る思考法』光文社。
中室牧子・津川友介（2017）『「原因と結果」の経済学——データから真実を見抜く思考法』ダイヤモンド社。

かわぐち・だいじ 東京大学大学院経済学研究科教授。
最近の主な著書に『労働経済学——理論と実証をつなぐ』（有斐閣、2018年）。労働経済学専攻。