



Title	証拠に基づく政策立案（EBPM）と「豊かさ指標」
Author(s)	町野, 和夫
Citation	年報 公共政策学, 14, 59-75
Issue Date	2020-03-31
Doc URL	http://hdl.handle.net/2115/78239
Type	bulletin (article)
File Information	14-07_Machino.pdf



[Instructions for use](#)

【論文】

証拠に基づく政策立案(EBPM)と「豊かさ指標」

町野 和夫*

1. はじめに

日本政府はバブル崩壊後から構造改革の一環として、行政改革を、さらにその中で様々な統計改革を進めてきたが、2016年の経済財政諮問会議ではGDP統計の改革などと並んで、「証拠に基づく政策立案」(EBPM: Evidence Based Policy Making)の必要性が指摘された。EBPMとは、政策立案を従来のように単に過去の事例に基づいて行うのではなく、政策とその効果を結びつけるロジックを明確にし、政策効果を示す定量的なエビデンスを見極めて合理的な政策立案を行うというものである。米国や英国では、1990年代後半にそれぞれ新自由主義的政権に対抗して政権を獲得したオバマ大統領やブレア首相のリベラル派政権が、新自由主義的政策に対抗して従来の支出が膨らみがちの政策から効率的な政策への転換を図る中で推進され始めた。日本では米英に遅れること約20年だが、先進国で突出した累積赤字を抱える財政状況下で政策への支出を国民に納得してもらうためには、EBPMの導入は必要条件であったと言えよう。2017年に設置された統計改革推進会議の最終とりまとめ(同5月)ではEBPMの推進体制構築のために、「統計をはじめとする各種データ」の整備・改善が必要であることが明記されている¹⁾。さらに同年7月にはEBPMを推進する体制として、官民データ活用推進戦略会議官民データ活用推進基本計画実行委員会の下に、EBPM推進委員会が設けられた²⁾。

EBPMが中央政府での政策の実施・評価を含めた政策のPDCA(Plan, Do, Check, Action)サイクル実現の根幹部分と位置付けられていく一方、地方ではそれより少し遡る2014年末には、「まち・ひと・しごと創生総合戦略」が閣議決定されると同時に

* 北海道大学公共政策大学院教授
E-mail: machino@econ.hokudai.ac.jp

- 1) 統計改革推進会議のweb site (<https://www.kantei.go.jp/jp/singi/toukeikaikaku/>) 参照。
- 2) EBPM推進委員会は、「世界最先端IT国家創造宣言・官民データ活用推進基本計画」及び「官民データ活用推進基本計画実行委員会の開催について」第7項の規定に基づき、統計等データを用いた事実・課題の把握、政策効果の予測・測定・評価による政策の改善、その基盤である統計等データの整備・改善を進めることにより、国民により信頼される行政の実現に資するため、関係行政機関相互の緊密な連携の下、政府全体として証拠に基づく政策立案(EBPM)を推進する体制として、官民データ活用推進戦略会議官民データ活用推進基本計画実行委員会の下に設けられた。

地方版総合戦略の策定が各都道府県、市町村に求められた。これら総合戦略でも具体的な数値目標を設定し、PDCAサイクルの中で効果検証と改善の実施が必要とされている。この数値目標の指標がKPI（Key Performance Indicator）という業績評価の主要指標である。

札幌市を例にとると、札幌市の政策体系は、「まちづくり戦略ビジョン」（2013～22年度）という10年の長期総合計画が定められていて、全体的な方向性を定めるビジョン編とそれを達成するための戦略編に分かれている³⁾。ビジョン編では、目指すべき都市像を「北海道の未来を創造し、世界が憧れるまち」と「互いに手を携え、心豊かにつながる共生のまち」としており、まちづくりの分野を、地域、経済、子供・若者、安全・安心、環境、文化、都市空間、に分け、各分野で3つから4つ、計24の基本目標を掲げている。それを達成するための戦略編では、長期的に大きなパラダイムの転換が必要になるテーマとして、暮らしコミュニティ、産業活力、低炭素エネルギー転換、の3つが選択され、3分野それぞれで、2つから3つ計8つの戦略（創造戦略）を掲げ、それらを支える都市計画として「持続可能な札幌型の集約連携都市への再構築を進める」という基本目標のもと、5つの都市空間創造戦略も定めている。

また、これらの10年計画の期間中の2015～19年度には、5年間の中期計画である札幌市戦略ビジョン・アクションプラン2015が作られ、戦略ビジョンで掲げた8つの創造戦略とそれを支える都市空間創造戦略に基いて9つの政策目標を設定し、期間中に実施する全ての政策的事業697をその中に位置付けている⁴⁾。ただし「雇用を生み出す力強い街」、「女性が輝き子供たちが健やかに育つ街」、「いつまでも安心して暮らせる街」、「魅力と活力にあふれた暮らしやすい街」を4つの重点課題については、それに対応する5つのプロジェクトの関連計画事業を示している。このような政策体系の中で札幌市では、8つの創造戦略（即ちそれに対応する8つの政策目標）の26の施策項目について計53個の指標について数値目標（KPI）を設定している。

本稿ではEBPMを現実の公共政策に適用するための課題を、日本の公共政策、特に地方自治体の政策立案において明らかにするとともに、「幸福度」などに代表される主観的指標を併用することで、その課題を一部克服できるのではないかという仮説を提示する。次節では先行研究を基にEBPMの公共政策の立案や評価についての役割と限界を整理する。3節では幸福度等の主観的指標が公共政策の形成のために開発されてきた歴史とその課題について整理し、4節では筆者が札幌市と共同で行った「効果的な指標設定の考え方に関する基礎調査」を基に、地方自治にとって（最終的）目標であるべき住民の幸福度の向上のための政策立案・評価に複数の主観的指標を合成した総合指標を利用する試みについて紹介し、5節で今後の研究課題をまとめる。

3) 札幌市の該当ホームページ (<https://www.city.sapporo.jp/kikaku/vision/index.html>) 参照。

4) 本稿執筆中の2019年12月下旬に「札幌市まちづくり戦略ビジョン・アクションプラン2019（速報版）」（2019～22年度）が公表されたが本稿の議論は同プラン2015に基づいている。

2. EPBMの現状と課題

2.1 狭義のEBPM

EBPMは今や中央政府だけでなく地方自治体の政策立案に際して必須の条件となっている。しかしEBPMは、元々医療において、ある措置を取った患者群とその措置を取らなかった患者群を比較して、医療措置の効果を客観的な証拠(Evidence)を基に評価する手法を公共政策の立案・評価に取り入れた考え方であり、特定の公共政策(医療、教育、治安など)を医療措置に、政策の効果を医療効果になぞらえて立案・評価する政策立案手法であり、政策の実施・評価を含めた政策のPDCAサイクル実現の重要なツールである。

しかし、ある措置を取った患者群とその措置を取らなかった患者群を比較するというランダム化比較試験(Randomized Controlled Trial: RCT)という手法は、特定の症状を持つ患者への医療措置という狭義のEPBMには適しているが、多様な特性(年齢、性別、人種、学歴など)を持つ市民・住民を対象とする公共政策では、政策が平均的には良い結果を及ぼしたとしても、個人や個別企業によっては十分な結果を得られない場合や、一部の政策対象にマイナスの影響がある場合があるなど、その有効性は十分に吟味されているとは言えない。

Khosrowi(2019)は、EBPMと同様の意味でEvidence-based Policy(EBP)という言葉を使い、EBPが目指すのは、RCTによって政策効果を質の高い客観的証拠で評価されるようにすることだと説明している。EBPの背景にあるのは価値自由あるいは価値中立性(という価値観)であり、重要視されるのは、不偏性(客観性)、正確さ、政策効果のある程度推定できること、であるとしている。価値中立性が可能なのは、(一定の価値観に基づいているはずの)目標設定とEBPによる有効性の追求とは分離できるという前提があるからである。しかしEBPも分配に影響を与えるため価値中立性を実現できていないのではないかというのがKhosrowiの評価である。彼はマイクロ・ファイナンスを例に取って、マイクロ・ファイナンスへのアクセスや結果は、起業前の教育、制度利用前の予算制約、ビジネスプランの妥当性など、多くの個人的事情に影響されるとしている。このようにそれぞれの政策が政策対象の個人や企業に与える影響はその個人や企業の能力や性質によって異なる。

同様にRCTは平均処置効果(Average Treatment Effects(ATEs))という平均値を使って政策の効果を推計するものであり、ある政策があるグループ・地域に有効だったとしても、他の集団にも平均的に同じ効果をもたらすとは限らないし、仮に平均的には同じだったとしてもメンバーの属性の分布は異なるので、その集団全体に有効とは言えない。さらに、政策が平均的には良い結果を及ぼしたとしても、個人や個別企業によっては十分な結果を得られない場合や、一部の政策対象にマイナスの影響がある場合、それをどう評価するのかという問題もある。

こうした問題に対処するためには、サブグループ分析を使うことが考えられる。し

しかしサブグループ分析は、観察できる事前の特徴（年齢、性別、人種、学歴など）によって恣意的に集団を分けることになるので、不偏性（客観性）、正確さ、政策効果のある程度推定できること、といったRCTの良さを脅かすことになる。したがって普通はEBPの解説の中でサブグループ分析は推奨されない。

しかし、EBPの良さをある程度保ちつつ、分配面での結果についての情報も得られる方法があるのではないかとKhosrowiは考える。適切な処置を行ったりサンプルサイズを十分大きくとったりするサブグループ分析がその例である。出来るだけ平等に、というような道徳的価値が目的であれば、EBPの証拠に基づいてその目的を効果的に追求することができる。このように、初めから政策目的が明確な場合、価値観は外生的に決まるので価値自由を実現できるかもしれないが、目的が特に定まっていなくて色々な政策事例を集めるようなケースでは、道徳的価値は直接、評価方法やモデルの選択に影響する。

価値中立という判断基準を満たしているように見える場合でも、例えば「効果的」という評価基準で考えるとEBPでは平均的な値でみて効果的なのであって、最悪の状況の人にとって効果的、ということの意味しない。価値中立を重要視するとしても、少なくともこのような事実を認識しておくべきであるというのがKhosrowiの主張である。

このように問題点はあるものの、個別の具体的な政策では対象集団の属性をある程度把握でき、上述のようなバイアス・問題点も定量的に評価できるという点で、従来のエピソード・ベースの政策立案に比べたEBPMのメリットは明らかである。

2.2 広義のEBPM（米国と英国の先行例）

実際にEBPMで先行する米国と英国での状況を見ると、狭義のエビデンスとしてRCT等の評価だけを認めるのではなく、エビデンスの認識はそれよりも広く、プログラム評価の手法（定性的評価を含む）や、業績測定の結果物もエビデンスとして認識・利用している。集計量にまとめ上げたデータやそのトレンドの分析などもエビデンスとして扱われており、両国のエビデンス供給体制はこの20数年間で進化している。日本の地方自治体がEBPMの遂行にあたって当面抱える最大の問題がエビデンスの不足であることを考えると、EBPM推進体制が整った両国のEBPMの発展の歴史と現状を概観することは日本の進むべき方向を考える上で参考になる。

2.2.1 米国の経験

津田・岡崎（2018）⁵⁾によると、米国では既に1960年代におけるジョンソン政権の「貧困との戦争」以降、連邦政府は貧困対策など数々の社会事業に取り組み、教育省、

5) 本節の米国のEBPMに関する紹介は基本的にこのレポートに基づいている。

労働省、保健福祉省等の省庁において、貧困施策の成果を厳密に評価するため、臨床医学の分野を中心にRCTが取り入れられていた。1980年代から1990年代半ばになると、社会保障分野において各州が抱える様々の課題をトップダウンで画一的に対応するには限界があったため、中央政府は一定の裁量を州知事に与え、政策効果を州間で競わせると同時に、裁量や予算を与える条件として効果検証も義務付けた。ここでの効果検証のデファクト・スタンダードがRCTであったため、RCTが一気に普及した。しかし、1996年に社会保障改革法が成立すると、この法律は州政府に大きな裁量や予算を与えたので、州政府はRCTでの評価をする必然性がなくなりRCTの件数も一気に落ち込んだ。この時期、教育と開発経済の分野ではRCTによる頑健な効果検証が盛んに導入されたが⁶⁾、連邦政府施策のほとんどはRCTを原則とする頑健な効果検証を受けていなかった。

オバマ政権になると、訪問型育児支援、健康、ホームレス対策、失業対策、犯罪予防等の幅広い社会施策分野にRCTによる効果検証を拡大していった。例えば、Tiered Grant (階層付き補助金)は政策効果を裏付けるエビデンスの頑健さに応じて補助金を配分する手法であり、幅広い政策分野に適用されていった。それでも頑健性の高い評価が適用されていたのは連邦政府全体ではごく一部の事業に止まっていた。最大のボトルネックは、エビデンス創出に必要な良質なデータが利用できないことにあった。そこで省庁や党派を超えたレベルでの検討が始まり2016年3月に“Evidence-Based Policymaking Commission Act of 2016”が成立し、同9月にCEP (Commission on Evidence-Based Policymaking)が設置され、CEPは2017年9月に最終報告書“The Promise of Evidence-Based Policymaking”を大統領・議会に提出した。最終報告書は、大きく四つの論点、①データへの安全なアクセスの構築方法、②個人情報保護とデータ利用の透明性を改善する方法、③データインフラの近代化、そして④エビデンス創出に向けた制度・機構の能力の強化、で構成されている。

さらにCEPは行政記録情報の利用体制も議論している。日本でも公的統計の二次利用が促進されているが、まだ公的統計の範囲内で利活用を促進することが主眼であり、行政記録情報まで踏み込んだ検討が中心になる段階にはない。日本では地方自治体等が行政記録情報を多く保有しており、住民基本台帳や出生・婚姻・死亡届などは市区町村が保有しているし、許認可情報は権限のある都道府県が持つこともある。CEPは一ヶ所に行政記録情報を集積すると技術的なミスや外部からの攻撃に対して脆弱な体制になることを危惧しており、行政記録情報の管理・保管はそれぞれの組織が行うべきとしているため日本の体制づくりにも参考になる。新たに設立されたNational Secure Data Service (NSDS)は、あくまで承認された研究プロジェクトごと

6) 2019年のノーベル経済学賞はこの分野の成果であり、途上国での教育政策等の効果をRCTの手法で実証したAbhijit V. Banerjee, Esther Duflo, Michael Kremerの3教授が受賞した。

にデータの統合を行い、データ提供に伴う危険性の一元的な管理に徹している。データの出口がNSDSに集約されることで、最先端のプライバシー保護技術を取り入れる役割もNSDSが集中的に担えるので効率的・効果的である。

なおNSDSは米国の統計部局の中で最大で、統計的手法などの技術面でも蓄積が大きいセンサス局に置かれた。日本では最大規模を誇る総務省の統計部局でさえ、米国と比べればかなり規模は小さく、人的資源の確保は大きな課題である。前述した統計改革推進会議の最終とりまとめ（2017年5月）を受け、日本のEBPM推進の要として設立されたEBPM推進委員会は、各府省に置かれたEBPM統括責任者等で構成され、人材確保・育成、省庁横断的な連携を進めようとしている。しかし、1960年代からEBPMを積み重ねてきた米国においては、連邦政府や州・地方政府のみならず、シンクタンクや基金、慈善団体、大学等研究機関における評価人材の蓄積が圧倒的に豊富であり、こうした団体間の人材交流や連携も進んでおり彼我の差は大きい。また人事制度の見直しも必要である。欧米では長期に亘って統計分野で専門性を磨く人材が少なくない。日本でも一定のEBPMを推進するための専門性を身に付けるのに適した人事制度の検討が必要である。

2.2.2 英国の経験

英国では、内山他（2018）⁷⁾によるとEBPMの本格的な導入はブレア政権（1997-2007）で始まった⁸⁾。ブレア政権は、サッチャリズム（新自由主義的路線）でも従来の労働党政権の福祉国家路線でもない「第3の道」を唱道し、政府の役割を尊重しつつその効率化を図った。この方針の下で「Modernising Government」が策定され、EBPMが打ち出された。そうした成果志向に対する意識を反映したものが、包括的歳出見直し(Comprehensive Spending Review: CSR)と公共サービス合意(Public Service Agreements: PSA)に基づく業績管理型政策である。この枠組みの中で、府省別に今後3年間に達成すべき目的、政策目標、業績目標、金銭価値目標が定められ、財務省と各府省との間で合意が取られた。

1998年のPSA導入時は、35分野にわたって600ものターゲットが設けられたが、必ずしも計画的に行われたものではなかったため財政と連動したものではなかった。ブレアの後を引き継いで2007年に誕生したブラウン政権は、複数の省庁にまたがる30の分野横断優先目標を定め、ターゲットを設定したのは153指標のうち1/3に留まるなど

7) 本節の英国のEBPMに関する紹介は基本的にこのレポートに基づいている。

8) 英国においては1972年、医師 Archibald Cochran が著書 *Effectiveness and Efficiency: Random Reflections on Health Services* を出版した。これは医療における RCT の利用を主張するものであり、大きな反響を呼んだ。現在は、医療に関するシステマティック・レビューである Cochrane Reviews が普及しており、“Cochrane Database of Systematic Reviews” として公表されている。

簡素化を図り、個々の府省に対しては新たに「府省戦略目的」を導入することとした。つまり、政権が各省の取組を管理し、成果重視の行政管理を行う仕組みとなった。

CSRとPSAの枠組みのメリットは、省庁間の協力体制によって全体の整合性が確保されていたこと、各省が目的を共有して政策を行えたこと、公共部門のデータが提供されていたこと、中央政府とサービス提供者の間に緊張感をもたらしたことなどである。しかしデメリットとして、本質的な改革ではなく目標の達成が重視されてしまったこと、アウトカム目標の多くが政府のできる範囲を超え、政策とアウトカムとの因果関係がはっきりしなかったこと、柔軟性のない仕組みであったため改革意欲が薄れてしまったこと、事務負担が増加したことなどが指摘されている。

こうした課題を踏まえて、2010年に誕生した保守党のキャメロン政権は、透明性の向上、権限移譲、ビジネスライクという3つの原則を掲げ、PSAを廃止して、政権の優先目標に沿って各府省における事業計画を導入した。その後、2015年には単独府省計画が導入され、各省の政策及び施策について、4年程度の期間の成果目標が設定され、それに対する評価が実施されている。

内山他(2018)によると英国政府におけるエビデンス供給を支える要因としては、第一に、エコノミスト、社会調査職、オペレーショナル・リサーチ職、統計職などの分析専門職が政府内において確固たる地位を占めていることがある。第二に、省庁の枠を超えた専門職ネットワークの存在が挙げられる。こうした横割的集団の存在により、省庁の枠を超えた共通理解が形成される。各専門職は、広範な専門職コミュニティに所属することからの恩恵を享受しつつ、各省庁へも十分に統合されている。また省庁間の異動も頻繁に行われ、研修もこうした専門職集団が実施することで分析スキルの質を担保していると考えられる。第三に、各省ではChief AnalystないしはChief Scientific Adviserと呼ばれる分析部門統括職が各省の分析部門を統括しておりエコノミストや社会調査等の各専門職の枠を超えて総合的に判断することが期待されている。なお、首席エコノミスト(Chief Economist)は経済学者等政府外の人材が登用されるという省庁が多い。第四に、財務省の役割が重要で、その一つは、Green Book、Magenta Book等を通じ、政策分析のガイドラインを各省に浸透させていることである。予算要求書には政策の事前分析が必要であるが、各省のエコノミストによる分析を財務省のエコノミストがチェックする体制がとられることにより、分析の質を担保している。第五に、行政府外の機関の役割も重要である。議会における常任委員会は基本的に各省ごとに設置され、超党派で政府を監視する存在である。また、National Audit Office(NAO)は、公的支出を精査する機関として議会に付属するが、こうした外部機関の存在がエビデンスの質を監視し、その質を担保している。中でもWhat Works Centre(WWC)は英国政府によって2013年に開始された取組で、歳出プログラムや現場での実践が、利用可能で最良のエビデンスを踏まえたものなることを目的とした機関である。現在、WWCは7つメンバー機関と3つのアフィリエイトで構成されてお

り、多様な活動領域やアプローチでエビデンスの創出、紹介、実行可能なガイダンスづくり、意思決定の支援という機能を担っている。

以上、EBPMの先進国である米英の経験を概観すると、当初はあまりに多数の目標を設定していたり、数値目標が独り歩きしたりと日本との共通点も少なくないが、人的資源の量的格差のような重要な相違点も見られる。

人的資源の格差ほど目立たないが、さらに重要な違いは試行錯誤に対する覚悟の違いである。CEPの会長でもあったBaron (2013)によれば、2002～12年の間にIES (Institute of Education Sciences)が行ったRCT 90件のうち、90%は殆ど効果が確認されなかった。また、1992年から2012年までに労働省が行った13件のRCTに関しても約75%には殆ど効果が確認されていない。さらに医学分野で簡易な試験で効果が認められた研究成果の50-80%はRCTで効果が認められなかった。しかし、Baronの報告ではグーグルとマイクロソフトが行った13,000の新製品開発や新戦略の試行でも80-90%は失敗であることも指摘されており、頑健なテストと並んで多くの試行錯誤の必要性も強調している。試行錯誤の重要性は誰もが認識していることであるが、日本の公的機関でこのレベルの成功確率の低さを受入れられるかどうかは疑問である。

3. 非経済指標や主観的指標の歴史とその課題

本稿の「問い」は、EBPMは本来の強みを生かして個別政策の目標設定や評価に特化し、総合的な政策の方向設定と評価指標は幸福度のような主観的指標（群）を活用するという新たな政策評価指標の体系を構築できないかというものである。前節でも明らかにしたようにEBP(M)の背景にあるのは価値中立性であり、不偏性（客観性）が重視される。しかし政策目標はアローの不可能性定理を持ち出すまでもなく、全ての国民あるいは住民の間で合意できるものではなく、その社会の価値観を反映せざるを得ない。従って問題は社会の価値観をどのように（正しく⁹⁾）反映させるか、ということになる。

国全体が貧しい時代であれば、国全体の所得でもあるGDPを目標とすることに大きな異論はなかった。しかし、ある程度の豊かさを実現した1960年代以降は、GDPが国民の幸福や福祉水準を示す唯一の指標でないことが広く国民の間で意識されGDPを含む新たな指標（群）によって政策目標の設定や政策評価を行おうとする試みも続けられてきた。こうした動きが始まったのは、公害や都市への過度の人口集中など、急激な経済成長の負の側面が社会問題化してきた先進国においてであるが、発展途上国における開発の在り方という視点からも、経済成長はあくまでも手段であり、最終目標は個々人の福祉水準の向上であるという考え方が1980年代後半から広まり浸

9) 何が正しいかという別の重要なテーマは本稿では議論しない。

透してきた。この考え方は1990年以降毎年発表されている国連開発計画（United Nations Development Program）の人間開発指数（HDI: Human Development Index）にも反映されている。

日本でも政府の国民生活審議会の調査部会で経済学者や社会学者を中心に1970年に「社会指標」作成の作業が始まった。欧米の定義の最大公約数として「社会指標」を「国民生活の諸側面或いは社会的諸目標分野の状態を包括的かつ体系的に測定する非貨幣的統計を中心とする統計指標体系」と定義し、また、その目的として「国民の福祉水準の全体的判定」、「社会報告の作成」、「社会計画の策定」の三つを挙げている。

この「社会指標」では、①健康、②教育・学習・文化、③雇用と勤労生活の質、④余暇、⑤所得・消費、⑥物的環境、⑦犯罪と法の執行、⑧家族、⑨コミュニティ生活の質、⑩階層と社会移動、の10分野を目標分野、それを27の根源的な社会的課題に分け、77の副次的課題、さらに188の下位課題に分け、それに対応する形で368の指標を採用したが、こうした膨大な情報を生かし切れず無駄になっているなど様々な問題点が指摘された。また、所得水準が上昇し、人々の関心は多様化、高度化し、より次元の高い精神的、文化的な欲求を重視するようになったので、見直しが必要であるという認識が高まり、1986年に「国民生活指標」として各種指標を再編成した新たなる総合指標が作られた。

この指標にも健康、安全、収入といった生活の基礎的条件と家庭、勤労、学校といった生活の場に関するもの、地域・社会活動、学習・文化活動といった生活活動そのものを表す分野が混在しており、十分整理しきれていないなどの問題点が残った。そこでこれを個人生活の充実という視点から見直し、客観的な指標だけではなく主観的な指標との結合をはかり、さらに、豊かさの評価軸も、基礎的なものから自由や快適さまで多角的な構成を考え、世代間や男女間などによる評価の違いも考慮した指標として、1996年に「新国民生活指標（豊かさ指標）」が開発された。

その後も、小泉政権の構造改革の成果を国民生活の視点から、ゆとりや安心などの非貨幣的な面も含めて評価し、暮らしの改革がどの程度進んでいるかを明らかにすることを目的として2002年に作成された「暮らしの改革指数」や、民主党政権の「新成長戦略」（2010年6月）に盛り込まれた「新しい成長及び幸福度に関する調査研究」を推進するため、有識者からなる「幸福度に関する研究会（2011）」で作られた幸福度指標試案があるが、いずれも前政権の政策を否定して成立した新政権が自らの政策の新規性をアピールするという色彩が強かった。

ただし、幸福度は最近の社会心理学や行動経済学の最新の成果を取り入れた概念であり、幸福の経済学という分野の勃興も、同様の指標開発の動きを活性化させている。ノーベル経済学賞受賞者のStiglitzやSenなどが、2009年にフランスのサルコジ大統領の依頼を受けて、生活の質を測るための新たな指標について行った提言（2009）が代表的であり、OECD（経済協力開発機構）を中心に国際的にも連携を取って開発を進

める体制になりつつある。

GDPは引き続き政策目標としての地位を保ってきたが、所得（≒GDP）という定量的評価が厚生指標として理論的に正当化できる代表的方法は「等価所得（WTP）」を計算し、各個人のWTPの加重平均を社会的厚生測定値とするものである（Fleurbaey, 2009）。もちろんWTPの測定や個人ごとの加重平均のウェイトをどう決めるかなどの問題のため実用化は難しい。GDP（所得）が、人々が幸福を実現するために（理論的ではなく）実際には大きな要因であることは否定できないが、GDPを修正してよりよい政策指標を作り出すためには、既述の問題以外にも、世代を超えた社会的望ましさの測定（資源消費や環境悪化、あるいは富の蓄積の長期的影響の評価が必要）をどうするかなど様々な実践上の課題も多く、現実には使える段階ではない。

そこで、代替的な方法として考えられるのが、後述のような幸福度を測る指標（群）、Senなどが提唱している潜在能力（Capability）アプローチ、国連の人間開発指数（HDI）のような総合指数の3つの方法である。このうち第3の総合指数については、統計の整備が不十分な場合には必要なものではあるが、理論的には個人的な効用や幸福（well-being）に基づくものでもなく、総合指数を算出する際の関連指標の重みづけに理論的根拠があるわけでもない。第2のセンのアプローチによる測定については本稿の範囲を越える議論が必要である。

幸福度については、まず主観的幸福は測定できるのかという問いに答えなければならないが、近年では、アンケート調査（全体的に言ってご自身の生活にどの程度満足していますかという問いにとっても満足している、満足している、あまり満足していない、全然満足していない、の選択肢で答える）、経験抽出法（1日の内のランダムな時点での気分の申告）、1日再現法（気分を思い出して申告）、心理学的計測（ホルモンの濃度、皮膚の伝導性、など）、神経科学的計測（脳の活動）、行動観察（微笑みなど）などの組合せで、かなり測定できるようになっている（Fleurbaey, 2009）。個人間比較可能性についての疑問はまだあるが、さらに進歩することは予想される。

近年の進歩の中で心理学的研究でわかった重要なことの一つは、主観的幸福の多面性であり、特に認知的評価と感情的状態の違いは重要である。このことを示すもっとも有名な例は、生活水準がある程度高くなると、それ以上経済が成長しても主観的幸福は高くないという「イースタリン（Easterlin）・パラドックス」である。従って効用を主観的な感情ととらえる標準的な厚生主義では、一貫性のある幸福度の指標化は難しい。

過去40年の幸福の経済学の研究の歴史をまとめたClark (2018)によると、主観的な幸福度のデータを使った研究は、主に(a)何が幸福度を高めるかという研究と(b)幸福度の高い人は何をするかという研究の2タイプの研究に分類される。公共政策の目標の参考となるのは前者の研究であるので、その概要を紹介する。このタイプの研究では幸福度を被説明変数として扱い、それがどんな変数によって説明されるかを研究

する。個人レベルでの研究成果としては、

- ・ 高所得者は幸福度が高いが所得が上がるにつれて限界的な変化は通減する、
- ・ 失業が幸福度の低さと繋がっている、

などは広く認められているが、年齢、性別、民族、生まれた国など政策で変更できない個人の性質に関しても、様々な研究が進められている。重要視されている、健康、家庭生活、教育に関しては、以下のような結果が出ている。

- ・ 主観的健康は調査時点での気分左右されやすく正確にとらえられないので、健康問題の数、既往症の数、病院に行った回数、入院日数、など客観データなどとの相関で関係が検証されている。
- ・ 家庭生活では、結婚している人はそうでない人（独身、離婚者、寡婦(夫)など）より幸福度が高い（因果関係は逆の可能性も高い）。子供の有無も幸福度と関係あるが所得水準によって負から正に変わるという研究もある。
- ・ 教育に関しては様々な結果があり学歴が高ければ所得は増えるが期待も高まる。

集計データに関する研究成果では、最も有名なものは、上述の一人当たりGDPに関する「イースタリン・パラドックス」であるが、その他にも、失業とインフレ、不平等、環境汚染、緑地や公園、気温や湿度、福祉政策や累進課税の度合いといった制度や政策、社会全体の雰囲気、例えば犯罪の多さ、他人への信頼、寛大さ（寄付の多さなど）、政治的腐敗、利他性、社会的支援の充実度なども幸福度に影響する、という研究結果がある。

次節では国や地方の公共政策の方向付け、優先順位の決定を行う際にこうした成果をどのように利用するかを考えていくが、予め幸福度研究の知見からの次のような注意点を押さえておきたい。

①短期的にはプラスだが長期的には効果が消えてしまう可能性、②ある人にはプラスでも他に人にはマイナスになる可能性、には気をつけなければいけない。①については以下のような例がある。

- ・ 現在の所得を自分の過去の所得と比較するので、過去の所得が高いとマイナスの効果があり、
- ・ 失業も続くと慣れてしまい（しかし消えてしまう訳ではない）、
- ・ 結婚、離婚、子供の誕生、配偶者との死別も同様のことが言える。
- ・ しかし、障害、整形、不安定な職などは効果が小さくなりにくい。

②に関しては、以下のような例がある。

- ・ 自分自身の所得はプラスに働くが社会的参照グループの所得はマイナスに影響する、
- ・ 家のサイズも同様、
- ・ 失業も失業率が高い地域ではあまり影響が大きい。（失業による自殺率も）
- ・ 体重や身長も比較グループ(やパートナー)との相対的な値が幸福度に影響する。

応用研究の一つとして、幸福度を被説明変数として重回帰分析し、その説明変数の係数を使って代替率を計算するというものがある。一人ひとりについての推計はできないが、何かの基準でグループ分けすればそのグループの式を（従って係数を）推定でき、二つの説明変数間の代替関係を定量的に評価できる。

以上のように幸福の経済学あるいは経済学以外も含めた一般的な幸福研究にはかなり歴史があり、研究も蓄積されている。さらに最近ではニューロサイエンス、あるいはビッグデータやAIなど自然科学からの研究も急速に進んでいる。しかし、社会科学や人文科学のギリシャ哲学から続く蓄積や、倫理面での配慮も踏まえて最新の自然科学的成果を活かしていくべきであろう。

4. 「豊かさ指標」の利用可能性

4.1 「豊かさ指標」のデータ

第2節ではEBPMが個別の公共政策を効果的に行うための必要条件であることを示し、第3節では、幸福度などの主観的指標の決定要因を分析することで、その国その地域の公共政策全体の方向付けとして使えるのではないかという可能性を示唆した。本節では、札幌市と筆者が共同で行った「効果的な指標設定の考え方に関する基礎調査」を例にとって複数の主観的指標の合成指標を政策立案・評価のために利用する可能性について議論する。

まず、この「豊かさ指標」開発のきっかけとなった西部他（2013）における北海道の地域別「豊かさ」指標作成で筆者が当初想定したのは、人々の生活の様々な側面で重要度が高い質問項目それぞれに対応する客観指標を選択することであった。しかし、実際には市町村レベルで個々の質問に対応付けられる客観的な統計データは少なく、使えたのは「豊かさ指標」を合成するのに使った10個の個別指標の内、「納税義務者当り課税所得」、「人口1万人当り医師数」、「完全失業率」の3個のみであった。他の項目については客観的指標ではなく、「道民ニーズ調査」¹⁰⁾の各質問項目に対する満足度で代用した。従って、ほとんどの項目で満足度という主観的データを使うことになった。

すると、指標を構成する大半の項目が主観的データを使うようになって、わざわざ少数の客観的データを含める必要があるかという疑問が生じてきた。客観的データと主観的データの混在した指標の解釈は難しく、整合性の観点から考えるとすべての個別データを満足度とすれば問題はなくなる。このような経緯で、「豊かさ指標」は主観的データのみで構成することになり、指標の意味は明確で作成も容易（異なる指標の基準化は不必要）になり、直接幸福感を0から10の数値で尋ねるような幸福度と

10) マーケティング分析の一つである消費者満足度分析（Consumer Satisfaction (portfolio) Analysis: CS分析）の応用であり、多くの自治体で住民のニーズを知るために用いられている。

も比較しやすくなった。

以下では、この手法を使って筆者が札幌市と共同で行った上述の調査の中で作成した札幌市の「豊かさ指標」について説明し、その分析結果を踏まえて効果的な政策評価手法のあり方を検討する。

まず「豊かさ指標」の基になるデータは札幌市在住の18歳以上の市民1万人に対する市民の意識調査である¹¹⁾。この調査では市民の生活シーンを、「健康・医療」、「家族や友人」、「経済」、「仕事」、「福祉」、「子育て・教育」、「余暇・生涯学習」、「生活の利便性」、「交通」、「環境」、「安全」、「地域」の12分野に分け、札幌市の長期ビジョン等の政策体系の分類も参考にしながら、市民にとって重要と思われる48項目の質問項目について重要度と満足度を尋ねた¹²⁾。

4.2 「豊かさ指標」と幸福感

今回の「豊かさ指標」は各項目の満足度を重要度で加重平均した値を（アンケートでの5段階評価から幸福感の質問と同じ）0-10スケールに変換した値である。直接0から10の段階で幸福感を尋ねた問2への回答を札幌市民全体で加重平均すると6.57となり、世界幸福度報告（Helliwell et al. 2018）での2019年版の日本の値5.886より高かった。「豊かさ指標」も6.56で、直接尋ねた幸福感とほぼ同じ値となっている。また、男女別に見ると女性の方が幸福感も「豊かさ指標」の値も高くなっている。

しかし年代別の「豊かさ指標」を見ると10代が最も高く、70代以上が2番目。最下位と下から2番目は50代と60代であるのに対し、幸福感を見ると年齢階級ごとに大きな違いはなく、しかも30代が一番高く20代が一番低い。

この違いは幸福感の判断基準や判断の際に重視した項目が年齢によって異なるためだと解釈できる。問3で尋ねた幸福感の判断基準として年齢によって大きく変わるのは「将来への期待」で、これを挙げる人は年齢と共に低下する。「豊かさ指標」に比べて20代の幸福感が低いのは将来への不安が大きいためかもしれない。また、問4で尋ねた幸福感の判断で重視した事項の中で20代が他の年代より明らかに多く選んだ項目は、「自由な時間」、「精神的なゆとり」、「趣味・社会貢献などの生きがい」、「友人関係」、「職場の人間関係」である。こうした精神的な豊かさに関わる項目は重要度・満足度を尋ねる48項目では捉えきれなかったのかもしれない。

その他の分析は紙幅の関係で詳しく言及できないが、世帯人数別の「豊かさ指標」と幸福感を見てみると、どの指標でもひとり暮らしの世帯の値が明らかに低い。さら

11) 2019年1月実施。対象者は1行政区あたり千人、男女半数ずつで有効回答数4032件（回収率40.32%）。

12) アンケートの質問票及び回答の集計表、基本的分析については、札幌市ホームページ「まちづくり基礎調査」平成30年度調査（<http://www.city.sapporo.jp/somu/machikiso/seika30.html>）の「効果的な指標設定の考え方に関する基礎調査」を参照。

に一人世帯の年齢階級別の幸福感は10代、20代、70代以上では市全体の数値とあまり変わらないが、30代以降60代までは明らかに低い。これは家庭を持つことと幸福感との相関を推測させるが、因果関係までは明らかでない。

以上のように「豊かさ指標」は単なる幸福度ではなく、その構成要素と想定される各種の要因（個人の属性や質問項目で問われる生活の様々な側面での満足度）という豊富な情報を含むデータである。因みにSano (2020) は、同じ札幌市の調査データを使って、前節で紹介した重回帰分析や主成分分析で、幸福感が「豊かさ指標」を構成する各項目の満足度で統計学的に十分なレベルで説明できることを確かめている。

4.3 「豊かさ指標」の項目別分析

まず札幌市全体で、需要度の高い順に質問項目を見ると、最も重要度が高いのは「自分や家族が健康であること」であり、その他、上位には健康・医療や家族、家計、安全、エネルギー等の項目が並ぶ。さらに各項目の重要度（横軸）と満足度（縦軸）の散布図を4つの象限に分けて分析した。政策の観点から最も注目すべきなのは、言うまでもなく重要度が高いのに満足度が低い項目である。その中でとくに重要度の高い項目は「犯罪や事故などの心配がなく安心して生活できること」、「地震、水害、雪害、火災などへの対策がしっかりしていること」、「費用の心配をせずに、治療が受けられること」であり、とくに満足度が低い項目は「エネルギーや上下水道が安いこと」、「除雪がゆきとどいていること」、「高齢者が経済的に不安なく生活できることである」である。「地震…」の満足度の低い理由は、一昨年度の胆振東部地震やその際のブラックアウトが影響しているかもしれない。「除雪…」についてはアンケート実施時期が冬であったことが関係している可能性もある。その他の項目は高齢化に伴う問題や、東日本大震災以降続くエネルギー価格の問題など、長期的かつ札幌市単独では解決が難しい課題である。

以上のように生活の様々な分野の重要度や満足度という主観的指標の分析は、政策の優先度や年齢別・行政区別など特定グループごとの政策課題を発見するための有益な方法の一つと言えよう。客観的な指標を見つけられない分野も含む全ての政策分野でその分野の重要度や満足度という指標を得られるため、分野別の政策に対する満足度だけでなく政策全体に対する総合評価を得ることもできる。

しかし上述のように、その時々々の出来事（地震、台風による大きな被害、年金や税金の制度変化、など）によって重要度や満足度が大きく変化することは避けられない。また、統計学的にも主観的幸福度で比較分析をするのはかなり例外的な場合に限られることも最近の研究で明らかになっている。(Bond and Lang, 2019)

5. 今後の研究課題（結論に変えて）

EBPMは元々医療におけるRCTのような、厳密には限られた対象で特定の目標設定

や評価の指標として、個々の政策・事業が特定の目的を果たせたかどうかを示すものであった。しかし、多くの場面で広義のエビデンスが作られEBPMが進展するのは客観的な目標設定・評価の情報を提供するという意味で必然的であり望ましいことである。日本でも2節で紹介した中央省庁だけでなく、例えば北海道(2019)のように地方自治体でもEBPMに対する理解は急速に浸透している。しかしこれも2節で議論したように、人的資源の不足や試行錯誤の量などEBPMの広範囲な実現へのハードルは高い。

これに対して主観的指標は3節で紹介したように具体的な目標実現や評価のための道具ではなく、国民や住民が何を求めているかという視点から目標を設定したり評価したりするための道具となり得る。また4節で議論したように、複数の主観的指標を総合した「豊かさ指標」のような合成指標は、個々の項目を適切に選択できれば国民や住民の意識を適切に反映した目標を設定でき(従って評価もでき)EBPMのための個別の目標設定を容易にすることもできる。

札幌市の「まちづくり戦略ビジョン」では26の施策ごとに一つから三つの成果指標を挙げているが、施策レベルでは、例えば「市民の孤立を防ぐ支えあいの環境づくり」というようにまだ抽象的な目標になっており、この施策に関する指標項目は、「孤立死について心配していない市民の割合」と「福祉推進委員会を組織している単位町内会の割合」である。もちろん「孤立死について心配していない市民の割合」は「市民の孤立を防ぐ支えあいの環境づくり」がうまくいっているかどうかを推測する一つの材料とはなり得るし、「孤立死」は孤立の最悪の結果なのでそれに注目するのは理解できる。しかし、心配しない市民の割合が増えても、「市民の孤立を防ぐ支えあいの環境づくり」以外の理由でそうなったかもしれない。もう少し「孤立死について心配していない」ための具体策まで考えるべきであろう。例えば郵便や宅配などの事業者と連携を取り、一人暮らしの高齢者のチェックするシステムが作られカバーされる世帯数が増えることが重要かもしれない。孤立死は定量的に測ることができるが、「環境づくり」は、様々な努力(行政側の事業や日常業務、地域の能動的・受動的取組)によって総合的に達成できるものである。従って、個々の政策目標や施策を実現するために、いかに効果的な事業を企画できるかが最も重要であり、それができれば、その事業の達成度合いを測る(客観的あるいは主観的)指標がその施策にとって大事な成果指標(KPI)となる。

以上のようにEBPMの運用の問題点と主観的指標の問題点に関するこれまでの研究成果を踏まえ、両者の利点を活かして政策評価体系の改善を図り、さらにEBPMの本来の客観的成果指標としての特長と、多様な分野の評価が可能な主観的指標の特長を組み合わせた総合的な厚生(well-being)指標(群)によって、新たな政策評価体系を構築することが本研究の最終目的である。十分なデータや人的資源が揃っていない状態で形だけのEBPMを進めることは無駄な仕事を増やすだけになる恐れがある。幸

福度などの主観的指標については、一時のブームは終わり、既述のような統計理論的な限界も明らかになっているが、筆者の「豊かさ」指標のように地域別に異なる重要性を反映した、しかも多様な分野をカバーした総合的指標は情報量も多く、有用性は十分にある。新たな工夫によってEBPMとも融合させて総合的な政策評価システムを構築しようとする試みはトライする意義が十分にあるのではないだろうか。

<参考文献>

- Baron J. (2013) “Statement to the House Committee on Ways and Means, Human Resources, Hearing on What Works,” Washington, D.C.: Coalition on Evidence-Based Policy
- Bond and Lang (2019) “The Sad Truth about Happiness Scales” *Journal of Political Economy*, 2019, vol. 127, no. 4, pp 1629-40.
- Fleurbaey, Marc (2009) “Beyond GDP: The Quest for a Measure of Social Welfare” *Journal of Economic Literature*, 47:4, 1029–1075.
- 北海道「エビデンスに基づく政策展開の推進」調査研究チーム（2019）「最終報告書『本編』、『別冊：「エビデンスに基づく政策展開の推進」のための手引』」
- Khosrowi, Donal (2019) Trade-offs Between Epistemic and Moral Values in Evidence-Based Policy *Economics and Philosophy*, 35 49–78.
- 西部他（2013）「地域の多元的価値を測定するための豊かさ指標の開発とその地域政策への応用可能性の研究」北海道開発協会開発調査総合研究所、平成24年度助成研究論文集 (https://www.hkk.or.jp/kenkyusho/file/jyosei_rep24-05.pdf)
- Sano, Yusuke (2020) “The effect of satisfaction and importance in various kinds of areas concerning daily life based on the questionnaire conducted in Sapporo” Master Thesis, Graduate School of Economics and Business, Hokkaido University
- 札幌市、町野（2019）「効果的な指標設定の考え方に関する基礎調査」（第3章 本調査に関する考察） (http://www.city.sapporo.jp/somu/machikiso/documents/h30machikiso_shihyou02.pdf)
- Stiglitz, Joseph E., Sen, Amartya, and Fitoussi, Jean-Paul (2009) Report by the Commission on the Measurement of Economic Performance and Social Progress, (<http://www.stiglitz-sen-fitoussi.fr/en/index.htm>).（要約の邦訳：福島清彦訳（2012）『暮らしの質を測る－経済成長率を超える幸福度指標の提案－』金融財政事情研究会）
- 津田広和、岡崎康平（2018）「米国における Evidence-based Policymaking（EBPM）の動向」RIETI Policy Discussion Paper Series 18-P-016
- 内山融 他（2018）「英国におけるエビデンスに基づく政策形成と日本への示唆－エビデンスの「需要」と「供給」に着目した分析－」RIETI Policy Discussion Papers Series 18-P-018
- 札幌市まちづくり戦略ビジョン (<https://www.city.sapporo.jp/kikaku/vision/index.html>)
- 札幌市まちづくり戦略ビジョン・アクションプラン2015 (<https://www.city.sapporo.jp/chosei/actionplan2015.html>)

Evidence-Based Policy Making (EBPM) and “Affluence Indicators”

MACHINO Kazuo

Abstract

“Evidence-based policy making” (EBPM) is a method to evaluate and to improve the effectiveness of individual policy projects. It is not a method to evaluate the policy purposes.

In order to compensate EBPM the author proposes the appropriate use of the “affluence indicator”, which is a value obtained by weighted average of the resident’s satisfaction levels on almost all aspects of ordinary citizen’s daily life. Attempts to build a comprehensive policy evaluation system by integrating “affluence indicator” with the EBPM may well be worth trying.

Keywords

Evidence-based policy making, Randomized Controlled Trial, happiness, well-being