

政策分析インタビュー

GDP 統計の改善に向けて

—教育、医療の質の測定—

慶応義塾大学教授

野村 浩二

政策研究大学院大学教授

杉原 茂

現在、「統計改革の基本方針」（平成28年12月経済財政諮問会議決定）や「統計改革推進会議最終取りまとめ」（平成29年5月統計改革推進会議決定）などに基づき、GDP統計を軸にした経済統計の改善に、政府一体となって取り組んでいます。その一環として、経済社会総合研究所では、GDP統計の精度向上に向けた基礎統計の改善のため、医療・介護、教育の質の変化を反映した価格の把握手法に関する研究を行っています。今回は、経済統計をご専門とされる野村教授、また、医療の質を踏まえた医療の効率性や技術波及をご研究されている杉原教授に、医療や教育の分野における質を計測することの意義やGDP統計への反映に向けた課題等についてお話を伺いました。

●医療や教育の分野における質を計測することの意義

——少子高齢化の進展の下、労働生産性の向上が重要となる中、教育の質、教育の効果に対する注目が集まっています。また、高齢化の進展、技術進歩により日本経済における医療産業の重要性が増していることなどを踏まえ、医療サービスがどれだけの価値を生み出すのかという点についても注目されています。今、こうした非市場サービス等の質の計測が求められる意義について、お考えをお聞かせください。

（野村氏）背景には、日本経済における労働生産性の改善に向けた教育への期待があると思います。時間当たりの労働生産性を測定しますと、米国に比べて日本の水準は30～40%ほど低いままにあります。そうしたギャップの根源は何でしょうか。最大の要因は、労働時間当たりの資本サービスの投入量が日本は相対的に少なく、依然として労働集約的であることです。

一方で、日米間で類似的な労働サービスの対価（賃金率）を比較しましても、日本の水準は30%ほど低いのです。より労働使用的事実と整合します。

その意味では、アベノミクスの下、名目賃金を上げていこうという方向は自然です。しかしそうした動きを持続可能なものとするためには、やはり生産性の改善を伴うものでなければなりません。それは、スキルの獲得や改善といった狭義の質向上に限らず、低賃金タスクから資本への代替、高賃金のタスクや産業への労働シフト、また組織において収益の上がらないサービスをやめるといったような点も含め、広い意味での労働の質や働き方のスマートさを高めていくことが必要です。生産性水準が劣位にあることは、日本の成長ポテンシャルとも捉えられます。生産性のキャッチアップに向け、教育は重要な役割を担うと考えられます。

——医療の質については、国際的にどういった関心の下、どのような議論が行われているのでしょうか。

（杉原氏）医療の質については、それぞれの国の特有の問題や政策課題に応じて関心の持ち方や重点の置き方が異なり、それに適したアプローチを採っています。例えば、米国の場合は、20世紀後半に医療技術が飛躍的に発展し、死亡率の低下やQOL（生活の質）の改善に繋がりましたが、その一方で、医療サービスの価格が高騰し医療費が大幅に増加しました。そうした中、医療技術の進歩によってもたらされた医療の質の改善が費用と見合うものなのかといった点が重要な論点となり、医療の質をどう計測するかということが盛んに研究されるようになりました。研究の結果、表面上、医療サービスの価格は大きく上昇していましたが、質の向上を考慮すれば、むしろ価格は低下しているという結論が導き出され、政策的にもインパクトがありました。一方で、個々の病院の医療の質をプロファイリングすることにより、医療技術や治療方法が必ずしも適切に患者に適用されておらず、医療の質が低い病院も多いという研究も多くあります。こうした中で、米国のSNA統計（NIPA）では、質が一定と仮定された分類の中でコストの安い治療方法へシフトしたデフレーターが低下するという形で医療のサテライト勘定が作られるようになりました。

イギリスでは、公的医療制度の下、医療費は抑制されてきましたが、その代わり待機時間が長いことを含

めて医療の質が低いのではないかという問題意識がありました。コストと質を同時に達成する方策として、医療費を抑制しつつ質を高めるといった方向性で議論が進む中、政府サービスの生産性の質を調整した形で計測してSNAとはやや独立した指標群として公表されています。EUについては、加盟国の統計部局の能力に限られる中、データが得やすいということもありDRG分類を利用した細分化による直接推計がSNA本体に組み込まれています。

日本については、研究や政策課題としては、米国と同じように、医療費の増大の妥当性や医療技術の適切な適用、包括支払い制度の下における医療の質の確保などの問題意識があり、素朴な面では、いわゆる「3分間診療」や「医療崩壊」論にみられるように、医療費の抑制の下での医療の質の確保というイギリスの観点があるように思われます。

●質の計測に関する最近の研究動向

——欧州の各国統計局では、教育に関するGDPの計測に際し、学校種別に細分化したデータを用いるほか、英国のONS（英国統計局）の生産性分析においては、テストスコアを活用して明示的に質を調整するなどの研究が進んでいます。現在、どのような観点からの研究が主流で、どこまで進んでいるのでしょうか。

（野村氏）質の把握については、様々なアプローチがありますが、そのベースは生徒数や時間など測定可能な単位によって、教育上のアウトプットを計測するという手法です。これまでは、インプットの集計値によってアウトプットを把握してきましたが、実質アウトプットを直接計測するといった考え方です。その際、アウトプットの測定単位をどう細分化していくかといった課題に直面します。質の計測のためにもっとも重要なことは、例えば教育分野であれば、教育サービスを層化し、細分化していくことです。それにより、アウトプットがより同質的な単位による測定へと接近することが期待されます。

その上で、次なる問題として、細分化した、それぞれに異なる質の教育サービスをどう集計するのかといった課題が生じます。医療と教育の計測の難しさは、そのサービスの便益が長期にわたり持続すること、その耐久性にあります。とくに教育では顕著です。もし教育サービスが消費的であるならば質の観察

もずっと容易かもしれませんが、人的資本として蓄積され将来にわたって労働や人生の質を高めると考えられます。その近似として、生涯所得の期待値の増加を考慮した集計も欧米での研究蓄積があります。

ONSなどでは、テストのスコア等を活用し、直接観察されるアウトプットの品質調整を行うことを試みています。その測定値を紹介しますと、1996年から2013年までの間、こうした明示的な質的調整の前では、英国でのアウトプットの伸びは年率0.4%であり、その間のインプットの伸びが1.9%であったので、両者の差である生産性は1.5%悪化したと推計されておりました。しかし、テストスコアを用いて調整したアウトプットをみると、2.1%に増加し、生産性についても0.2%のプラスに転じています。こうした結果の正当性には議論があります。教育の成果とは、問題の発見や解決能力、協働するチームの構築など、広い意味で捉えられるべきとすれば、テストのスコアはその一面であることは明らかです。また学校以外での教育機会や家庭教育も大きく影響しますが、そうした外的環境の統御は容易ではありません。教育全体の成果や質を測定・評価するためには、まだまだ議論・研究の蓄積が必要です。測定可能な単位による細分化と集計に関する課題を中心として、まずは構成変化の把握によって、教育サービス全体の質の変化へと接近していくことが重要であると思います。

もう1つのアプローチとしては、インプットの測定をあらたに見つめなおすことです。教育の改善に向けた取組として、例えば、教員1人あたり児童生徒数の縮小、IT機器の導入や実験設備の改善、学生同士が議論を行うような場所をキャンパス内に設けるなど、そういったインプット側の取組による質の改善もあります。研究部門では、企業内研究開発の生産量を（質の違いを補正した）特許数などで測る試みもありましたが、SNAでは特許に結実せずとも、さまざまな成果や失敗も知識の増大であるとしてより広義に研究のアウトプットを捉えています。教育部門でも類似的に捉えれば、生徒数やスコア補正などよりも、教育機関のさまざまな取組を評価できるような集計インプットのほうが、SNAにおける教育サービスのアウトプットとして適切かもしれません。そのアウトプットがどのように教育の成果（アウトカム）につながるのかは、サービス生産の次のプロセスなのです。アウトプットとインプット両面から、測定可能な分野での計

測を進め、質を反映した指標と生産性指標とを同時に検討していくということが、教育の質を総合的に評価するために必要だと思います。



——医療については、eurostatやOECD等の国際機関、英国や米国等の欧米各国政府機関をはじめ、米国の医療経済学者であるDavid Cutler氏等の研究者の研究蓄積がありますが、どのような観点からの研究が主流で、これまでの研究はどのように整理できるでしょうか。

(杉原氏) これまで研究者や統計担当者が行ってきたことは、大きく分けて2つの軸で整理できます。1つの軸は、質を調整する際に依拠するアプローチは何かで、分類の細分化によるのか、あるいは統計的手法により調整を行うのかといったことがあります。もう1つの軸は、質を調整する対象として何を用いるのかということで、直接的にアウトプットを推計する過程で質を調整するのか、または質の調整を行ったデフレーターを用いるのかといったことです。この2つの軸を組み合わせるとマトリックスになりますが、例えば、分類の細分化とアウトプットの直接推計という組み合わせはeurostat、細分化の考えをデフレーターに適用したのが米国のサテライト勘定、統計的な手法による質調整とアウトプットの直接推計を組み合わせたのがイギリ

スの生産性指標群という形でしょう。統計的な質調整でデフレーターを計測したのがCutlerたちの研究で、これは、生計費指数の考え方に基づいて、医療の質の上昇により消費者の効用水準が高まることは、同じ水準の効用を得るために支払う金額が少なくなるということから質を調整したデフレーターを算出します。

こうしたアプローチや対象には、それぞれトレードオフがあります。例えば、細分化による調整は、疾病や治療行為によって分類を細かく設定することによって、分類内の質を均一化し、分類間のシフトを質の変化ととらえて質の調整を行うものですが、統計的なモデルに依存せず、比較的頑健な調整方法だと考えられます。ただし、分類を細かくするのにも限度がありますので、質の調整自体には、不十分な面が残ってしまう可能性もあります。また、分類内では医療の質は時間とともに変化しないと仮定しますので、技術進歩等が勘案されにくいほか、質の改善運動のようにコストをかけて医療の質を改善しようとする計測される生産性が低下してしまうという問題があります。他方、統計的手法により、明示的に質の調整を行う場合、データの利用可能性にもよりますが、かなり詳細な質調整が可能である一方、統計的なモデルに依存するため、頑健性が失われる可能性があります。質の調整対象として何を用いるのかといった点については、質を調整したデフレーターを計測し、名目産出額をデフレートして実質化するというアプローチの方が、市場型サービスの計測の原則に則った形であり統計としてGDP統計との親和性が高いと思いますが、市場価格が存在しない中で医療の質を金銭換算する必要があり、データや手法上の課題を克服していく必要があります。

質の調整を行う手法については、唯一絶対の手法が存在せず、様々な手法を試し、感度分析を重ねることで検討を進めていくことがよいと考えています。統計改革の重点として生産性の計測をしっかり行うということがありますが、そうした目的に合致した計測手法となっているかをきちんと評価する必要があります。なお、アウトプットの直接推計は市場が存在しない非市場型サービスにおける計測手法として提案されているものですが、価格や供給量が規制されている下では消費者の評価を反映したデフレーターの質の調整が一筋縄でいかないことを考えると、日本の医療のような市場型サービスでも代替的な経路として、直接推計か

ら質調整を試みて検証する意義はあると思います。

●我が国のGDP統計への反映に向けた課題

——研究動向を踏まえ、我が国のGDP統計への反映に向け、どのような検討課題が考えられますか。

(野村氏) 我が国のGDP統計という視点から見ても、まずは適切な細分化と集計が求められます。高等教育では学科の細分化も検討課題ですし、集計においては先ほどの生涯所得からのアプローチに加えて、コストウェイトでの接近法も相互の精度検証のためにも重要です。そのためにも投入表(U表)における教育業の投入ベクトルを、細分化された教育サービスへとマトリックス展開するような時系列データベースの構築が必要です。

今回、統計改革の中で進める、教育のアウトプットの質的な変化を把握するという試みは新しい課題ですが、インプットとしての労働における教育の質的变化を把握することは、生産性統計では長い測定の歴史があります。SNA勧告には労働投入の測定もありますが、2008 SNAでは質を調整した労働投入指標の計測が推奨されています。インプット面では、労働サービスというフロー量に対する価格(賃金率)が観察できますので、直接的に教育の質の代理変数となります。アウトプットに着目した今回の研究では、その積分値にフォーカスしますので、両者には密接な関係性があります。設備投資と資本ストックのように、価格面におけるフローとストックの関係性を整合的に計測することも重要な課題です。

教育分野における生産は、現状のJSNAではその多くが非市場型の産出であるとされていますが、授業料の支払いを通じて家計は一定の負担をしているわけですので、非市場サービスであるとして価格の観察を断念するのも時期尚早かもしれません。米国経済分析局でも最近、奨学金などを差し引いた私立大学での授業料による価格指数も推計され、高等教育のCPIとの乖離も指摘されています。奨学金や補助金などの制度要因を調整した、基本価格としての価格指数へと修正しながら、実質量をインプリシットに定義していく可能性も残されているかもしれません。JSNAでは測定可能性を軸としながら、価格と数量、アウトプット指標とインプット指標を総合的に分析していく必要があります。

ます。教育の把握の相違により、生産性や産業の効率性、労働生産性の変化や劣位性の源泉などに関する認識も変わりますが、内部整合性のある体系的な測定と相互チェックが、JSNAの更なる改善に向け重要になってきていると思います。



——医療分野については如何でしょうか。

(杉原氏) 決定的な手法はないとは言うものの、医療分野については、質の概念がQOLということで比較的明確にされており、質を計測するための手法についてもかなり研究が進んでいます。加えて、日本の場合、国民皆保険制度の下、保険者が統一的な枠組みの下にあることから、データも収集しやすいと言えます。そうした蓄積やデータを活用すれば質の調整にも期待が持てます。

具体的な検討課題を挙げますと、例えば細分化については、医療の質が一定となる分類を設定することが決定的に重要ですから、分類軸をどうするか、また、どこまで細分化する必要があるかを様々な候補を含めて検討することが必要です。また、分類軸は、質を一定にするだけでなく計測に歪みを生じないことも重要です。例えば疾病に加えて治療方法を分類軸とすると、治療方法には医療供給者が選択するという側面があり、例えば、コストはかかるが効果は薄い治療方法が普及すると医療の質が改善したとカウントされてしまうとい

うこともあります。統計的な質の調整は、何と云っても、適切に質を計測する手法を開発することが課題です。これまでバイズ的な階層モデルなど有力な手法が提案されていますが、利用可能なデータでどこまで詳細なモデルで質の計測を行うべきか、逆にどこまで手を抜いても大丈夫か、あるいは、粗い情報でどこまで代替できるかなどを検証する必要があります。さらに、生計費指数アプローチを含めた全体的な課題として、何を質の指標とするかという問題があります。正統的には治療後のQOL（生活の質）の向上を採るということですが、QOLのデータを統計作成用に収集することは困難です。代表的な指標として死亡率を採ることが多いですが、死亡率が高くない疾病などうまくカバーするためには、それ以外に、治療後の合併症、再手術、日常生活活動（ADL）のレベル、主観的な健康観などを組み合わせて工夫していくことが望ましいでしょう。

実用化に向けては、実際の統計作成過程において全ての作業を厳格に行うことは困難であるため、個票データ等を用いて様々な感度分析を行うことなどを通じて、統計として、必要となる精度を確保することができる作業のレベルを把握することが研究の1つの役割だと思います。質の調整には不確実性が伴いますので、シミュレーション等によって不確実性の程度についても把握しておくことが重要です。また、医療分野のデータには、DPC（診断群分類）データやレセプトデータなど幾つか利用可能なものがありますが、それぞれに情報量やカバレッジが異なることから、両者を上手く組み合わせることにより質の計測を行う、または両者の違いを検証するなどの柔軟な取組が重要になると思います。なお、DPCデータは患者の重症度等について詳細な情報を含むものですが、すべての病院がDPCに参加しているわけではなく、また、年々参加病院が増えていくという2重の意味でサンプルの選択が行われています。医療全体のアウトプットやデフレーターを計測するには脱落によるバイアスを調整するために相当の工夫が必要ですが、統計学的にも大変興味深い研究課題です。なお、DPCデータの持つ詳細な情報を利用して限定した範囲の病院で精緻な推計ができたとしても、それを情報の少ない入院全体の推計に拡張しようとする大きな不確実性が生じるので、各手法を評価する際にも、そうしたカバレッジのギャップに起因する不確実性もシミュレーション等で良く把握して評価する必要があります。

最後に外来や診療所等の入院外の医療サービスの計測について簡単に述べたいと思います。日本には特定健診という制度がありますので、そこから検査値など詳細な情報と広いカバレッジを持つデータが得られます。検査値から将来の疾患の発症や合併症などが予測できれば、検査値の変化を通じて入院外の医療の質を推計することがかなりの程度できることになります。あるいは、カバレッジは限定されますが、国民生活基礎調査のように身体・健康の状態や日常生活動作の状況等把握できれば、それをQOLに換算することにより、医療の質を測ることも考えられます。ただ、いずれの方法も転帰やQOLへの変換をどのように行うか更なる研究が必要です。

——質の測定を行う際、医療と教育の分野では、どういった違いがありますか。

（野村氏）医療と教育では、インプットの質の変化スピードに大きな違いがあると考えています。医療技術の向上や、生産性の改善の多くは、医療用機械や医薬品、言い換えれば、資本財や中間財のプロダクト・イノベーションに起因しているのではないのでしょうか。とくにそれが新製品であるとき、測定の困難性を伴います。質の評価として、医療ではとくに生産に利用されるプロダクトの分析へと拡張が必要なのだろうと思います。

（杉原氏）インプットのようなものを基に質を測定・調整するというアプローチは重要であると思います。先行研究でも、薬剤や手技を適用した場合の死亡率等を臨床研究により把握して、それらの薬剤や手技がどれだけ広範に適用されるようになったかを掛け合わせることで、医療の質の変化を測定する試みがあります。

——本日は貴重なお話をいただきました。ありがとうございました。

（聞き手：経済社会総合研究所上席主任研究官 市川恭子）

（本インタビューは、平成29年11月8日（水）に行いました。なお、インタビューの内容は、以下のページからご覧いただけます。

http://www.esri.go.jp/jp/seisaku_interview/seisaku_interview2012.html