

第51回ESRI経済政策フォーラム「マイクロデータを活用した政策研究について」  
平成27年2月27日開催 議事録

1. 開会あいさつ（西川経済社会総合研究所長）

御紹介いただきました、内閣府経済社会総合研究所所長の西川でございます。

本日は、第51回「ESRI－経済政策フォーラム『マイクロデータを活用した政策研究について』」に御参加いただき、ありがとうございます。

今回のESRI経済政策フォーラムでは、マイクロデータを活用した政策研究について取り上げてまいります。

最初に本日のフォーラムの狙いを御説明いたしたいと思います。お手元にお配りの資料5をごらんいただきたいと思います。事務局からの資料を資料5として用意しております。

資料5の2枚目のスライド、1ページ目の下のスライドでございますが、これがこのセミナーの趣旨を、本日の御発表者4名の方に御説明した際の資料でございます。

3つの分野がございます。左から「Ⅰ マイクロデータの作成・収集・提供」、真ん中に「Ⅱ マイクロデータを利用した分析の進展」、さらに「Ⅲ 政策決定過程へのインプット」という3つの分野があるわけでございます。

それぞれの分野では進捗がなされておまして、例えば真ん中の分析の分野では、マイクロ計量分析手法の進歩を背景として、経済学等を中心とした学会で、実証分析の論文が蓄積されております。具体的な例は、後ほど大竹先生や野口先生から御紹介いただけたらと思っております。

また、左のマイクロデータの作成などの分野では、科学のさまざまな機関によるパネルデータとマイクロデータの提供が拡充され、また、統計法の改正により、学術研究等を目的とした政府統計の二次利用、外部提供の拡大が進んでおります。さらには民間におけるビッグデータの利用や、マイナンバー制度の実施などを見据えた、行政記録情報を含めたデータの利活用の検討等、新しい取り組みが始まっています。

右の端の政策決定過程へのインプットの分野では、近年、医療経済の分野をはじめとして、実証分析により、具体的な政策プログラムの評価を行おうという取り組みがございます。

しかしながら、こうした各分野の進捗が、互いに連携がとれているのかということについては、疑問がございます。それぞれの分野の中では、枠内それぞれに記載したような進展があるのですが、各分野の取り組みにおいて、他の分野にどう役立っているのかという認識が十分なのか、3つの分野が全体と

して証拠に基づく政策、Evidence Based Policy Makingという観点から、きちんとしたサイクルになっているのかということが、本日の問題意識でございます。

資料5の中には、各分野の内外の具体的な取組みを幾つか紹介しておりますが、時間が限られておりますので、そのうち、先端的取組みを簡単に触れさせていただきます。

スライドの10枚目では、北欧諸国において、行政記録情報、特に税情報と社会保障関係の記録がリンクされ、利用が盛んになっているということ、さらにスライドの14枚目ですが、それによって、世代間移転を含めた長期的な格差の研究が可能になっていることが紹介されています。

また、資料5の最後のページをごらんいただきたいと思います。上の段、23番目のスライドをごらんください。アメリカ合衆国においては、証拠に基づく政策の議論が進み、昨年11月には、上下両院の予算委員長から、証拠に基づく政策のための法案が議会に提出されております。成立すれば、専門の委員会が設けられ、政策強化に用いるデータの包括的な研究・提供、行政記録情報の一括提供のためのセンターの設置、プライバシー保護などについての検討を行うこととされています。

最後のスライドの下、24番目のスライドにお示ししておりますが、本年1月のアメリカ経済学会では、行政記録情報を用いる専門職団体の協議会、COPAFSが中心となって、行政官、研究者、統計作成者が一堂に集まり、具体的な行政記録のアクセス改善のための議論を進めております。

我が国の実証分析の状況に戻って考えますと、各国と比べて、まだまだ分析発展の余地があるのではないか、あるいは研究成果が政策立案に十分に結び付いているのか、そもそも政策評価に使えるような、効果だけを取り出した情報が充分に入手できるのかなどの指摘がございます。政策に注目した実証研究を推進していく当研究所としては、こうした課題に取り組んでいくことが急務であると考えております。

本日のフォーラムでは、我が国の実証研究分野において、先端的な御活躍をなされている研究者、分析手法、統計収集、労働経済、医療経済などの各分野に造詣の深い専門家をお招きし、御議論をお願いいたします。先端的な研究に携わられている講師の方々の御議論に触れ、今後の研究計画、政策立案評価の進め方、統計の提供・利用に関する議論が深まり、具体的な進捗につながれば幸いです。

ありがとうございます。

## 2. 基調講演 : 東京大学 市村英彦教授「マイクロデータを活用した内外の研究の進展と日本における今後の課題等について」

御紹介いただきました、市村でございます。どうぞよろしくお願いたします。

まず、経済に関する実証分析の現実がどうなっているのか、どういうふうに機能しているのかについて、データを用いて考える知識の体系として計量経済学での進展が政策にも利用できるほど、かなり精緻になってきているということを御説明したいと思います。ここでは、分析に、どういうデータが必要となるのかを、幾つかの例を通して見ていくことで、皆さんの共通の理解を得た後で、いろんな議論が有効にできるのではないかと考える次第です。

タイトルにあるマイクロデータとはどういうものかと言いますと、個人・家計・企業など、意思決定主体と想定される対象に関するデータのことで。

こうしたマイクロデータを用いたマイクロ実証分析は、1950年代より欧米を中心に大きく進展してまいりました。従来の経済分析は、金融政策、財政政策、産業政策、あるいは労働政策などをマクロ的に考えることだというイメージが非常に強いと思いますが、今の世界的な傾向としては、非常に具体的なことについて、マイクロデータを用いて実証分析することが増えてきています。

あとで少人数学級の例で御説明いたしますような効果だとか、あるいは公共財調達などを行うオークションで、どういう仕組みを用いることが収入を最大化するのかといった個々の具体的なテーマについて、マイクロデータを用いた分析が進んでいます。あるいは、具体的な環境政策の効果も分析の対象にのぼってきております。これらの実証分析を通して、現実がどうなっているのかだけでなく、意思決定主体がさまざまな状況で、どういう誘因にどの程度反応するのか、あるいは、ある政策はどれくらい効果を持つのかに関しての知見が蓄えられてきています。のちほど事例を吟味していきたいと思います。

ここで注意しなければいけないことは、我々が測定したいと思うものが、必ずしも測定できるわけではないということです。したがって、実証分析を進めて結果を解釈するときには、それがどういう仮定のもとで行われているのかということを見ていくことが必要です。手法をよく理解していないと、数字が4つ並んでいるときに、その4つを全て同じようなものとして評価してしまうということになりかねません。仮定の強さがどういうふうに関係してくるのかということも、幾つかの例を見ていく中で、感じていただきたいと思います。

測定したいものが、必ずしも測定できるわけではないということの代表的な例として、プログラム評価の問題があります。

あるプログラムに参加したかどうかを示す2値の確率変数を $D$ とし、参加した場合には1、参加しなかった場合には0とあらわすとします。 $Y_1$ は参加した場合

に得られた結果をあらわします。

例えばここで、あるトレーニングプログラムを想定し、参加した場合は $D=1$ で、参加しなかった場合は $D=0$ 。Y1はトレーニングに参加した場合の所得、Y0は参加しなかった場合の所得と考えますと、プログラムに参加することによる効果というのは、 $Y1-Y0$ であらわすことができます。

しかし、同じ人に関しては、プログラムに参加した場合にはY1が観察できて、参加しなかった場合にはY0が観察できるわけですが、どちらか一方は観察できないわけなので、 $Y1-Y0$ というものは、誰にとっても観察できないものになるわけです。ですから、今、プログラムの効果を $Y1-Y0$ で定義したわけですが、 $Y1-Y0$ は測定できないものになります。これが「プログラム評価の問題」と言われているもので、こうした効果をどういうふうに測定するのかということについて、いろいろ工夫がされてきているわけです。

以上のように、個人にとってのプログラムの効果は、原理的に測定できないものですから、平均的なプログラムの効果を測定したらどうかという話になります。平均的なものを測定しようとする、Y1の平均からY0の平均をとればいいということになります。そうしたら、Y1のサンプル平均とY0のサンプル平均を比べた方がいいのではないかということになるわけですが、よく考えてみると、これはそうではないということがわかります。どうしてかと言うと、Y1の平均というのは、結局、プログラムに参加した人たちの平均であって、Y0の平均というのは、プログラムに参加しなかった人たちの平均になります。だから、実際に推定できているものは、プログラムに参加した人たちの平均と、プログラムに参加しなかった人たちの平均になるわけです。プログラムに参加した人たちの平均と、プログラムに参加しなかった人たちの違いを見ているということは、同じ人についての差ではなく、違うグループを比較していることになるわけです。これはプログラムの効果なのか、グループが違うから結果が違うのかが区別がつかないということになります。例えばプログラムに参加する人たちはもともと熱心な人たちだと考えると、これはプログラムの効果を過大評価することになってしまいます。

こういう問題は、薬の効果を測定する場合や肥料の効果を測定する場合でも、同様に起こる問題で、古典的な問題です。農学や生物学をはじめ、医学などでも同様の問題に直面するわけですが、ここではランダムマイゼーションと呼ばれる実験（ランダムイズ実験）が、非常に有効な手段として使われていました。

経済学においても、1960年代から、例えばジェネラスな保険を持った人たちは普通よりも医療支出が増えてしまうかどうかという、モラルハザードの問題などを測定するために、ランド（RAND）研究所が社会実験を行ったりしています。

ランダムマイゼーションでどうして問題が解決されるかということを考えてみます。まず、ここでは人々をランダムにプログラムに割り当てるわけです。そうすると、プログラムに割り当てられた場合のアウトカムとプログラムに割り当てられなかった人たちのアウトカムが見えるという点では前と全く同じわけですが、個人が選んだ結果ではなくランダムに割り当てられたか、割り当てられなかったかだけの違いですので、2つのグループ、プログラムに割り当てられた人たちと、プログラムに割り当てられなかった人たちに違いはないということになります。

この前提が正しいのであれば、プログラムに割り当てられた人たちの平均とプログラムに割り当てられなかった人たちの平均の違いを見ることによって、プログラムの平均的な効果を見ることができることになります。したがって、みんなが自発的にプログラムを選んでいるという現実に行っている観察データではなくて、ランダムに割り当てられた実験データを用いることで、全く同じ手法を使ってプログラムの平均的な効果を測定できることがわかります。

しかし、ランダムマイゼーションで推定可能となるのは平均的な効果であって、プログラム効果の分布などは推定できないことに注意する必要があります。しかし、平均は推定できます。

ただ、ここで重要なことは、実験が非常にうまく機能している場合の話でありまして、肥料を余計に与えられたか、あまり与えられなかったかによってプラントの側は行動を変えたりしないわけですが、人は違うレベルのトレーニングを与えられると、行動を変える可能性があるわけです。だから、この場合には、実験を行っても、うまく機能していない可能性があります。

今、ここで「\*（アスタリスク）」がついている変数というのは、実験下での変数、「\*」がついていないのが、観察下、実際に社会で起こっている変数だとすると、実験下の変数と現実に起こっている、対応しているような変数が、平均の意味では一致している必要があるということになります。

具体的にどういうことが仮定されているのかと言いますと、実験参加者は、割り当てに従うか、従わない事情は、結果とは無関係な事情にあるということが前提になっています。

もっと具体的に考えると、熱心な者がプログラムに割り当てられない場合、他の手段を講じる可能性があるけれども、ここではそういう行動はとらないことが前提になっています。逆に、参加したくない者はプログラムに割り当てられても参加しないかもしれないけれども、そういうこともないということを考えているわけです。また、実験期間は限定されていますけれども、結果に影響しないということも仮定されています。

社会実験を行ってプログラム評価をするというのは、非常に難しい前提が置

かれる必要があるということが、だんだん認識されてきました。

そのほかにも、経済的・政治的にも費用がかかります。例えば少人数クラスの効果について、アメリカでは社会実験を行っているわけですが、これを実際に日本でも行おうとすると、社会的・政治的に相当な費用がかかる。時間がかかるだけではなく、政治的にも非常に難しいプロセスを経る必要があります。したがって、タイムリーにデータを利用できないという欠点があります。

実現可能な場合、有効に実験ができるのであれば、貴重な情報をもたらすわけですが、難しい。

したがって、最近の経済学では、観察データの中から、実験に非常に近い状況を見つけ出してきて、その状況を用いて実証分析を進めることにより、プログラム評価を進めようという趨勢が強いです。いわゆるNatural Experiment、自然実験アプローチと呼ばれているものです。

以下、お配りしているものには5つ書いてありますが、2つだけ御説明させていただきます。

最初は、育児休暇延長がどういうふうに効果を持つのか、出産率に関して、どういう効果があるのかということを中心に話します。

これは、Lalive and Zweimullerという人が、Quarterly Journal of Economicsの2009年に書いている論文で、育児休暇をより充実させることで、どれだけ出生率が上がるか、彼らは女性の労働供給が変わるかどうかも実証的に分析しています。

利用したのは、オーストリアにおける年金局のデータで、育児休暇制度の変更を利用しています。ここでD=1というのは、ある年の政策でD=0というのは違う年の政策になります。2つの政策を比較して、その効果を分析しようというわけです。

制度変更は1990年7月に行われました。

まず、出産休暇というのは、出産前8週間と出産後8週間あります。1990年7月より前、6月あるいはそれ以前の制度では、出産後8週間が終わると、子供が1歳まで育児休暇がもらえて、支給額については、出産後8週間までは産休前3カ月の平均給与がもらえるのですが、育児休暇に入ると一律340ユーロ、これは中央値の約4割だそうですけれども、それだけの額がもらえます。ただし、前回の出産から15カ月半以内の出産の場合には自動的に取得資格がもらえるけれども、1年3カ月半以上開いてしまって、出産前1年間に20週以上働いていないと、「働いていない」という条件が加わるわけです。

1990年7月以降に生まれた子供を持つ母親の場合には、育児休暇の長さそのものが2歳まで延びて、しかも、前回の出産から27.5カ月以内の出産であれば、自動的に育児休暇がもらえるような制度に変わった。この研究では、ジェネラシナ制度に変わったために出生率がどれだけ変わったのかということ进行分析し

ているわけです。

彼らは、1990年6月に子供を産んだ母親と、1990年7月に子供を産んだ母親に着目しています。最初のグループは昔の制度に直面している母親たちで、7月に子供を産んだ母親は、新しいよりジェネラスな育児休暇制度に直面している母親たちです。

2つのグループに共通なのは、次に生まれる子供については、同じ政策に直面しているということです。今の子供の育児休暇の長さが長くなるが、将来の子供については、同じジェネラスな育児休暇に直面している。だから、今の子供に対する効果が、どれだけ出生率に影響するのかということを見ていることになります。

スライド21の左側が6月に生まれた子供の母親、右側が7月以降に生まれた子供の母親の育児休暇の長さを見ていますが、右側は10カ月ぐらい長くなっていることがわかります。

次の図は、この母親たちの中で、もう一人以上子供を産んだ母親の割合がどれぐらい違うかということを示しています。これを比べると、6月に子供を産んだ母親たちよりも、7月に子供を産んだ母親たちのほうが、5%ぐらい、もう一人子供を持つ割合が増えているということが見てとれるわけです。

1つの家計が生涯の間につくる子供の数は変わっていない可能性も吟味する必要がありますが、スライド25の以下の図から、そういうことでもなさそうだということがわかります。

図(A)は、ある子供が生まれてから $t$ 月経った母親が、次の月に子供を産む確率を示しています。旧制度では点線までの間、つまり13.5カ月以内に子供をつくれれば、自動的に育児休暇がもらえたところで、点線のところまでの出産率が高いわけです。新しい政策では、実線で示した26.5カ月まで育児休暇が自動的にもらえるようになったので、新しい政策に直面した、7月に子供を産んだ母親たちのほうが出産率が高くなる。それ以降では、効果があってはおかしいわけですが、点線も実線も子供を産む割合が変わっていないことがわかります。理論的に考えられるとおりの変化が見てとれるということを示しているのが、左の図です。

右側の図は、横軸が時間で、子供を産んでから経っている月数です。120というのは、10年後ですけれども、そこまでの間を見ても、次の子供を産んだ母親の割合というのはコンバージしていないわけですから、5%の効果というのは、長期的にもあったということが、これで見てとれるわけです。

以上の効果は、現在生まれた子供の育児休暇が長くなった効果を見ているわけですが、将来の子供の育児休暇が長くなった効果というのは、どうやって見ているかといいますと、これは1987年6月に子供を産んだ母親と、1990

年6月に子供を産んだ母親とを比較して吟味しています。

両者とも、現在の子供の育児休暇は、前と同じ制度に直面しているわけですから、あまりジェネラスではないわけです。しかし、将来の子供に関しては、1987年6月に子供を産んだ母親は、昔のままだったわけですがけれども、1990年6月に子供を産んだ母親は、新しい制度のもとだということがわかっているわけですから、将来の子供の育児休暇が長くなったということは、知っているということになります。だから、3年間だけに限った話ですがけれども、その間の出産の行動の違いは、将来の子供の育児休暇が長くなった効果だと解釈できます。

そこで、彼らが比較をしてみると、将来の育児休暇を1年延ばすことの出生率への効果は、約7%だということが示されています。したがって、総合的には、この政策をとることによって、出生率は12%変わる、かなり大きな効果があるということを示しているわけです。

ここで見ましたように、一番重要な点は、6月に子供を産んだ母親と7月に子供を産んだ母親の違いがほぼないということです。だから、ニュース、新しい政策が打たれるということ、いつ母親たちが知ることとなったのかということ、吟味することによって、2つのグループの間には差があるとは考えにくいということを議論しています。2つのグループの違いがほとんどないということ、いかにきちんと説明するかということが、こういうデータ分析の非常に重要な点となってくるわけですがけれども、そのところで、専門的な手法などが議論されることとなります。これが1つ目の例です。

2つ目の例として、失業保険の失業期間への影響です。

先ほどと同じデータを使っているんですけども、オーストリアの特定地域で、50歳以上の人たちを対象として、失業給付期間が30週間から209週間に延長されたことを利用した分析をしています。失業期間にどう影響したのかということ、

青い地域に住んでいる人たちは、新しい政策下の人たちで、それ以外の人たちは、新しい政策の影響は受けなかった人たちです。50歳以上かどうかということと、ブルーの地域に住んでいたかどうかということで、政策に影響を受けたグループか、受けなかったグループかということが変わってくるようになります。

Laliveさんはどういうふうに見ているかというと、同一地域、青い地域の中に住んでいる人たちの間で、50歳を超えていなければ新しい政策の影響は受けないわけですから、そういう人たちと50歳をぎりぎりを超えている人たちが1つの比較できるグループで、その比較を見ると、15週間失業期間が延びていることがわかります。よりよい町で職にありつけたのではないかとということもあり



ますが、失業期間が長くなって、それだけ政府支出も増えたわけですから、より根本的なことを考えるのであれば、賃金などがどれだけ上がったのかということを見ていく必要があるわけですが、ここでは失業期間がどれだけ延びたかということだけに着目して分析を見ていきます。

50歳以上について地域ボーダーでの比較を見ることもできます。青い地域と黒い地域との間の細い地域を取り出して、青い地域にぎりぎり入っているのか、青い地域からぎりぎり出ているのかの違いを見て、失業保険の失業期間への影響を分析すると、14週間増えていることがわかります。

スライド31は図で見ているわけですが、左図において50歳よりも上では26週を少し超えた辺りが平均で、50歳を少し下回っているところでは13週間で、きれいにそこでジャンプがあるわけです。それが15週間程度の違いだということから、年齢で見た場合には、失業期間が15週間ぐらい延びるということが見てとれます。

地域で見ると、スライド32における右側がボーダーの中の人たちで、左側がボーダーの外の人たちになりますが、14週間ぐらいの違いがある。いずれにしても、同様の結果が得られているわけです。

一番心配なのは、わざと50歳になることを待って失業したり、引っ越して失業するようなことが起こっているのではないかということです。そういうことが懸念されるわけですが、次の図を見ると、そういうことではなさそうということがわかります。スライド34は何の図かと言いますと、左側は50歳より下の人たちで、該当地域に住んでいる人たちと、該当地域に住んでいない人たちとで、新たに失業してくる人の割合の比を見ています。左側は、50歳以下ですから、該当しないわけですから、その比は1に近いはずですが、実際に1に近いです。

右側、該当地域の50歳以上の人たちは、わざと失業をすれば、失業給付が受けられるようになっているわけです。期間が長くなっているわけですから、わざと失業するインセンティブがある人たちです。該当地域でない人たちは、そういうインセンティブがない人たちですから、もし行動が違うのであれば、この比は1よりも高くなるはずですが、この図では1に近く、該当地域だからといってわざと失業していることはなさそうということがわかるわけです。

今度は地域ボーダーのほうで、年齢の比によって同様の計算をしてみても、男性の場合には、やはり1近辺にあり、行動を変えてわざと失業をしているようにはみえません。年齢・地域のどちらで見ても、同じような結果が出ていることもあり、失業保険期間を延ばしたことの効果は、失業期間を14週間から15週間ぐらい延ばしたということが、かなりの確度で言えるということになります。

女性の結果では、どちらで測定をするかによって、随分大きな違いがあります。年齢で見ると、105週間失業期間が延びたという結果になりますし、地域ボーターの比較で見ると、52週間延びたということになります。

男性と同じ図で見ると、女性の場合、左側は1に近いわけですが、右側は1よりもすごく乖離していることがわかります。この結果から、女性の場合には、わざと失業するインセンティブが強い人たちは、2.5倍以上失業に入ってくる確率が高く、そういう行動をとっているのではないかが懸念されるわけです。こういうことを吟味することによって、年齢を使った分析は、あまり信頼できないということがわかります。ただ、地域で見ると両方とも1に近く、こちらを使ったほうがよさそうだとということがわかります。結果、女性の場合には、52週間延びたとする方を信頼すべきだということがわかると思います。

こういう分析で、失業保険の影響をかなり厳密に分析していくことが可能になります。この分析を支えているのは、年金局の個票のデータになります。

時間の関係もありますので、ほかの例は省きますが、こういう形で、現在のところは、マイクロデータを使った分析が進められているわけです。

計量経済学で重要な研究対象の1つは、以上のような分析で得られた結果が、どれくらい確かなのか、信頼区間がどれくらいの大きさなのかということです。ここでは、今のところ、お話していませんけれども、そういう指標をどういうふうにつくっていくのかということが、盛んに議論されているところです。この20年ぐらいの *American Economic Review*、*Quarterly Journal of Economics*、*Journal of Political Economy* などの雑誌を読むと、こういう分析が数多く見られます。

以上、観察データを用いて、マイクロ実証分析の実例をプログラム評価の枠組みの中で見てきました。データにあまり制約を置くことなく、虚心坦懐にデータに語らせるというアプローチが、現在のマイクロ実証分析の主流となっています。

しかし、このような研究からわかることは、ある時点、ある場所で、どれぐらいの平均的プログラム効果があったかということです。それがわかることを過小評価すべきではないわけですが、同じプログラムを、将来、他の場所で実行したときに、同じ効果を生むのかということについては、何もわかりません。

例えば失業者に対する訓練プログラムの評価を行い、十分な効果があることがわかったとします。しかし、実際に訓練プログラムを労働市場の一部として組み込んでいくと、労働者の行動様式そのものが変わる可能性があります。これがいわゆるルーカス批判と言われている問題です。大学へ進学することのさまざまな効果の実証され、多くの奨学金が整備されたとします。そうすると、大卒の労働者が増加し、高卒との相対賃金は変わってしまうかもしれません。

そうすると、それまで測定されてきた大学へ進むということの経済的リターンが変わっていくこととなります。これが一般均衡論的分析ということで、こういうプログラム評価的な単純分析では、排除されている可能性があります。

どういう行動を、家計・企業・政府などはとっているのか、また、どのようなメカニズムを通して、対象のプログラムが効果を持つのかという点を解明する必要があります。こういう点を解明しようとするのが、**Structural Approach**、構造アプローチというものです。経済分析の重要性は、まさしくこういう分析を行っているところにあります。

もともとは、McFaddenによる静学的な選択のモデルから出発し、1980年代には、Rust、Miller、Wolpinたちによって動学化され、現在はPakesやBerryたちによって、産業政策などを見ていくために、ゲームモデルの推定に利用されています。Heckman、Taber、Wolpin、Leeなどによって、一般均衡のミクロモデルの推定へと拡張されています。

それぞれの対象問題に応じて、部分均衡モデルや一般均衡モデルを基礎に置く確率モデルを作成していく必要があります。こういう構造的なモデルをつくることによって、平均的な効果が、どういう状況のときに大きい効果を持つかということがわかります。

経済モデルのパラメータの大きさがどう依存しているのかを明らかにしていくことによって、効果の安定性、あるいは効果の大小がどういう状況で起こり得るのかということが明らかになっていくわけです。

こういう構造アプローチに関する研究結果は、**Econometrica**、**Review of Economic Studies**といった雑誌で紹介されています。

これでお伝えしたかったことは、今、行われているミクロ実証分析は、一体どういうものなのかということです。ざっと大まかなイメージを持っていたと思いますけれども、ミクロ実証分析を支えているものは、一体どういうものなのかということを考えていきたいと思います。

PhDのトレーニングを受けた想像力のある研修者が第一に必要です。

同じぐらい重要な点は、質の高いミクロデータ、詳細なデータ、母集団に対して代表性があるデータです。

今、お話したように、6月に子供を産んだ母親なのか、7月に子供を産んだ母親なのかといった違いがわかるぐらいのデータでなければ、今のような分析は不可能だったわけです。だから、かなり精度の高いデータが必要だということが、おわかりになると思います。

もう1つ重要な点は、学会と政策当局との関係があると思います。

この3点について、議論していきたいと思います。

まず、研究者の現状ですけれども、一定水準を超えたミクロ実証ができる研

究者の数は、国内に30名程度だと思います。米国トップの50の大学それぞれで、10人程度はマイクロ実証を行っていますので、恐らく国内30人と同等か、それ以上の水準の研究者の数は、米国の場合には700人を超えていると思います。

また、欧米・アジアに、国内トップレベル数人と同等か、それ以上の日本人マイクロ実証研究者たちが、少なくとも8名はいます。これらの方々を呼び戻すことが急務だと思います。直接的な研究のインパクトも大きいですが、それにも増して大きいと思いますのは教育を通じての間接効果です。毎年、こういう方々に1人ずつトレーニングしていただいて、大学院で優秀な研究者を育てていただくと、10年の間には80人になるわけです。40歳で帰ってこられて、20年間教育していただくとして、160人になるわけです。そういう効果というのは相乗効果を持ちますので、すごく大きな効果になっていくと思います。

彼らが戻ってこられない大きな原因は給与差です。卑近な話で申しわけないんですけども、その方々を対象に特別待遇のポジションをつくって、米国の大学並みの評価システムで米国並みの給与体系を整えれば、戻ってこられるのではないかと思います。中国・シンガポールのトップ大学では、こういう制度を採用済みです。

幾つかの他分野も、同様に給与差があると聞いていますが、経済では、トップ10に限らず、トップ50でも、同じように大きな給与差があるのが特徴です。数学・物理では、トップ数校だけの給与が高く、それ以外はむしろ日本のほうが高いぐらいです。

世界的な大学をつくるには、コストがかかります。

次に示していますのが、米国経済学部9カ月の給与です。ランク1~15、ランク16~30、ランク31~48、ランク49+のAssistant Professor、講師とあります。ランク49+でも、為替レートが100円だとしても、1,000万円を超えており、このデータは9カ月分の給与ですから、年間の場合はこれに3カ月分不足必要があるわけです。とてもではないけれども、かなり大きな違いがあって、戻ってきてもらえないという状況があります。それが人材に関して、非常に重要なポイントだと思います。

次にデータについて議論したいと思います。

マイクロデータには、同時点で、多数の個人・家計・企業に関するデータを集めたクロスセクションデータ、それを数年間集めたリピーテッド・クロスセクションデータ、多数の個人・家計・企業を追跡調査したパネルデータ、企業とそこで働く労働者の双方を追跡調査したマッチド・パネルデータなどがあります。

政府の集める家計データは、通常、リピーテッド・クロスセクションデータです。月次で集めている家計データの場合には、数カ月間のパネルデータとで

きる場合もありますが、典型的には、リピーテッド・クロスセクションデータの構造をとっています。

企業に関する政府データは、名寄せができれば、パネルデータにできるわけですが、年によっては、名寄せをするための情報がないと聞いているものもあります。センサスデータは、原理的にはパネルデータ化できるはずなんですが、そういう試みは、今のところありません。

政府データの利用が可能となりつつあるのは、すばらしいことで、非常にありがたい話なのですが、幾つかまだ残っている問題点があります。

1つは、院生が教員とは独立に、個人の立場でデータを利用できないという点。論文を投稿した後リバイズを要求されることが多いですが、迅速に対応できないという点。また、先の具体例からも想像できるとおり、分析には詳細な情報が重要になることが多いのですが、個人の特定を防ぐ目的で、しばしば重要な変数が詳細にわかりません。年齢も1歳刻みでわからない場合があります。エリジビリティを議論するために詳細な所得などが必要な場合もあるわけですが、それもカテゴリーでしかわからなかったりする場合があります。

オーダー・メイド集計が盛んに言われているわけですが、オリジナルな研究をするには全くなじまない制度で、生データを利用できることが必須です。また、まだまだ用意する書類が煩雑で、かつデータを受け取るまでに時間がかかることもあります。あと重要なのは、回収率が公にされていない場合があって、データの質がよくわからないという問題があります。

あと、今のところ、税務データ・社会保障関連データの利用にめどが立っていないという状況があります。先ほど御説明した2つの研究例は、両方ともオーストリアの年金局のデータを使ったものです。ああいうものがあれば生涯所得などもわかるわけで、現在の所得の分布ではなくて、生涯所得の分布も議論できるようになってくるわけです。そこが非常に重要な点だと思います。

政府データの進捗に比べて、パネルデータ構築は、欧米に比べて大きくおわれています。パネルデータで分析すべき事柄について、個票を用いて分析する場合には、定常性の仮定が必要になってきます。

例えば今、50歳の人々の10年後の平均所得は、今、60歳の人々の平均所得だということを仮定して分析していくわけですが、こういう例で見ると、男性の場合には、下にいくほど、若い人たちですが、年齢に従って所得がどういうふうが増えていくのかということは、世代によって随分違ってきます。だから、定常性を仮定して見ていくと、大きな間違いを犯す可能性があります。だから、個人を追跡していくことが非常に重要になってくるわけですが、そういうデータで、信頼度の高いものが、非常に限られていることがあります。

若年層を含むデータの場合、限られた質問に絞った阪大の場合を除いて、回収率が実質20%で、これでは国際誌には採用されないレベルのデータになっています。これをどういうふうに上げていくのかということは、きちっと議論して、解決していかなければいけない問題なんですけれども、そこがあまり議論されずに、ずるずるとデータ集めが続いていってしまっているという現状があります。

高齢者の場合には、JSTARというデータがあるんですけれども、60%弱で、欧州各国並みのデータとなっていますが、まだまだアメリカの80%に比較すると、低いわけです。60%というのは、問題が多いデータだと思いますが、どうやって回収率を上げていくのかという議論が、あまりされていない現状があります。

米国では、ランド (RAND) などの研究組織が調査を行うほか、大学が調査を行い得る組織を運営しています。

ミシガン大学のSurvey Research Center(SRC)だとか、シカゴ大学のNational Opinion Research Center(NORC)、オハイオ州立大学のCenter for Human Resource Research(CHRR)という組織があって、会社の組織みたいになっています。例えばNORCなどは、年商100億程度の中規模の企業並みの運営をしています。

調査を丸投げするのではなくて、自前の調査員を抱える場所を持つていくことが、回収率を上げていく、質の高いデータをとっていくためには、どうしても必要なことではないかと思います。私の経験ですが、足立区などで調査をした際、8割の回収率をとってくる調査員もおられるんですが、0という人も相当いるわけです。実際に2時間半ぐらいの面接調査をしているんですけれども、そういうときでも、8割とってくる人が実際におられるわけですから、回収率の高い調査を行うためには、そういう人たちだけを集める作業、調査員選別から行い得る組織が必要だと思います。同時に一定時間内の質問の仕方など、どういうふうに質問するのがよいかということ吟味する研究を行うべきだと思います。SRC、NORC、CHRRのようなセンターが、データ集めのためには必要だと思います。

ビッグデータについても少しスライドに書いてありますが、時間が押していますので、飛ばすことにいたします。

最後に、学会のあり方と政策当局との関係です。エビデンスに基づく政策の必要性は、唱えられて久しいです。膨大な評価は、毎年、実際になされているわけです。その中身を見ていきますと、プログラム参加者に簡単なアンケート調査を実施して、満足度を聞いたものが膨大に毎年つくられています。

そういう現実がどうして生まれるのか、もっときちっとした評価が生まれないのかということを考えていく必要があります。

少なくとも3つの理由が考えられると思います。まず、エビデンスを出すことと、各省庁との利害、予算の獲得が直結していないことが、形式的な書類づくりとなる原因ではないかと思います。

2つ目は、評価を実施できる人材の数が少ないという点も大きいと思います。

3つ目は、学会との関係がきちんと構築されていないということもあると思います。

1つ目に関しては、予算作成過程で実証結果が用いられるようになって、それが新聞などを通してオープンに議論されるようになってくれば、かなり変わってくるのではないかと思います。

2つ目に関しては、必ずしも政策当局者自身が腕まくりをして実証分析をする必要はないと思います。Request For Proposalなどで評価を発注する制度を拡充していく必要があるのではないかと思います。現状、日本で学者が政策担当者とコンタクトを持つ際には、少人数のグループでの集まりとなる傾向が多いように感じます。

もう少しフォーマルな場合には、そこにコンサルティング会社が入ってきますけれども、事務的な仕事はコンサルティング会社がこなし、学者は自由に意見は言いますが、実際の細かな仕事は、ほとんど政策担当者側とコンサルティング会社側で進められる場合が多いです。

政策当局が抱えている問題のうち、中長期的課題に属することは、問題提起を学会にして、学会全体を巻き込むような形で、研究を促すくらいのことを行ってもよいのではないかと思います。政府や日銀の研究所が実務部門の中長期的な課題を吸い上げることで、学会と協働していくような形ができないものかと感じます。もう1つは、政府・日銀内でPhD取得者が何人かおられるわけですが、そういう人たちの役割がどういうものなのかということが、非常に不明確だと感じます。

国内コンサルティング会社などでは、そもそもPhD取得者を雇用していないのではないかと思います。政府のコンサルティング契約への参加要件は、年商や経験年数はあっても、雇用者の質を具体的に担保していないことがあります。コンサルティング契約への参加要件は、参入障壁になっているけれども、質の担保では、あまり役に立っていないのではないかと思います。政府の実務者は、政治家と研究者と学会とを結ぶパイプ役としての役割を担うべきだと思います。だから、自分で腕まくりをして計算していこうという形よりも、むしろそういうところに特化されるのがよいのではないかと思います。

PhDは運転免許証のようなもので、恐らくパイプ役となるためにも、本来は必須要件だと思います。こういう状況は、今のところまだできていないわけですから、そういう状況が定常状態となるような組織変革が望まれると思います。

今、申しましたように、マイクロ実証分析を支える3要件として、研究者、データ、政策当局と学会との関係がありますが、それぞれについて、現状と改善策を議論いたしました。

もちろんマイクロ実証分析には、おのずから限界があります。自分自身を含めて、人々がどういうふうに行動しているのかということに関して、深い理解がまだまだ足りない部分が多い中で実証分析を進めていかなければいけない。まして違う政策をしたときに、人々の行動がどう変わるのかということを理解していかなければいけないわけですが、よくわかっていないからといってデータを見ないで政策を立案していいのかというと、そうではないと思います。ベストを尽くして、どういう政策が一番効果がありそうなのかということ、限られた予算の中で考えていく必要があると思います。

したがって、こういう機会を設けていただいて、より広範な議論と具体的な施策がなされることを期待します。

ありがとうございました。

(休 憩)

### 3. パネルディスカッション

【モデレータ（西川 正郎 内閣府経済社会総合研究所長）】はじめに、パネリストの先生方を紹介します。まず、基調講演をいただきました市村英彦先生、そのお隣が一橋大学の北村行伸先生、大阪大学の太田文雄先生、最後に早稲田大学の野口晴子先生でございます。

最初にパネリストの皆様から、本日のテーマ「マイクロデータを活用した政策研究の推進について」に関して、論点の提起、あるいは具体的な研究の紹介等、それぞれ10分程度で御発表いただきたいと思います。

それでは、北村先生からお願いいたします。

【北村行伸先生】 ただいま御紹介にあずかりました、一橋大学経済研究所の北村でございます。

先ほど市村先生の報告を聞いていて、全ての論点について、もっともだと思いました。非常にコンパクトにマイクロの実証の実態をまとめていただきました。彼とはかれこれ四半世紀、いろんところで、こういうことを一緒にやっているのが感慨深いものもありますし、20年、25年を考えると、データもいろいろ使えるようになったし、実証する人も多少は増えてきたという気持ちもあります。一方で、25年経っても、こうやってマイクロデータの実態が、海外に追いついていないということを発表しないといけないというのは、つらいものがあり



ます。とはいっても、いろいろな形でマイクロデータを使う環境が整っているということは、言えると思います。

2つ大きな流れがあると思います。1つは、統計法が2009年から施行されて、公的統計の二次利用の促進がその中でうたわれていまして、利用者サイドに立った統計という話が出てきています。匿名データの利用、を促進するとか、あるいは公的統計をオンサイトで利用するとか、それをさらに高度化してリモートアクセスの試行とか、いろいろな形のトライが行われていまして、学者・研究者がマイクロデータを使う環境もかなり整ってきている。それが1つ大きな流れだと思っています。

もう1つは、西川さんが最初に御説明されたように、エビデンスベースの政策形成を重要視しましょうという流れが出てきています。最初は、恐らく医学、薬学の話が出てきていたと思うんですけども、今、それが社会科学、特に経済学の分野でしきりに言われるようになっておりまして、実証研究を通して政策形成をしましょうというコンセンサスが、ある程度できつつある。そういうことを政界・学界・民間で連携してやっていくという流れもできていると思います。そういう意味では、ビッグデータの利用環境というのは、大分整ってきていると思います。

もう1つ、エビデンスベース・ポリシーの重要性に関しては、そのために統計データが整備されて、利用できることが大事であります。統計のデータには、政府がかなりの資金を使って、毎年あるいは毎月統計調査を行っている公的統計があります。それを十分に活用するという事は、費用の意味でも求められていることですが、これも市村先生がおっしゃったように不十分なところもありますし、精緻な統計になっていない場合は、大学や民間の研究所で新たに統計調査を始めて、本当に学者がほしいような情報を集めるという試みも必要になります。

今、言ったような話を具体的に考えますと、例えば消費者物価指数という、CPIと言われるものは、総務省が集めています。それは月に1回公表されるものです。日本銀行が金融政策の実施において、物価の上昇、たとえば2%を目標にしているようなときには、このCPIを指標にして使うわけですが、実際に政策当局がオペレーションをする場合には、日々金融市場でいろんなものを取引するといった情報が必要になりますし、政策の担当者も今日はどうかとか、先週と比べてどうか、もっとフリークエンシーの高いデータなどが必要です。政府のCPIではなかなか見られないんですけども、昨今、東大の日次物価指数というものを、元同僚の渡辺さんが作っています。一橋大学でも、東大のデータよりは頻度は落ちて週に1回ぐらいなんですけれども、カバーは大きく4,000店舗ぐらいのPOSデータを使って週次のデータベースを作っているということ

があります。そういうデータを使いながら、ある程度1カ月で動くような、物価の動きを見るということをしているわけであります。

政府の統計がいいのか、民間の統計がいいのかという話ではなくて、そういうものを互いに使い分けるといいますか、補完的に使うという話になってきているという感じがいたします。

実際の政策関連の統計の利用の仕方として、公的統計の話について少し申し上げておきたいのは、公的統計というのは長い歴史があるものですから、そんなに簡単に調査方法や、調査項目を変えることはできないんですが、一方で実際にライフスタイル、家族のあり方、高齢化が進むということにより調査が難しくなっているというのは、たしかであります。例えば家計調査に当たった家計は、6カ月間家計簿をつけ続けなければいけない。しかも、買って来たもののグラム数とか、価格とか、そういうものをいちいち記録をつけていくわけですけども、そういう作業が非常に煩雑です。もしそういうことが電子マネーやクレジットカードの情報等に吸い取られて全部記録してくれるのであれば、随分簡単になると思うのですが、手で日記をつけなければいけないという話をずっと続けていくことは、今後可能かどうか。そういう話も考えなければいけないと思います。

それから、時代の要請で、こういう統計をとったほうが良いということが出てきた場合、予算が削減されていく中で、新しい統計を導入することは大変です。スクラップ・アンド・ビルドで古い統計を潰して新しいものにしたらどうかとか、少なくとも一部は統合して浮いた予算を新しい統計に使ってほしいということがあるわけですけども、それは自治体の事情や統計調査の継続性、いろんな面で、難しいところがあるわけですが、こういうことも統計委員会を中心に考え直さなければいけないということはあると思います。

市村さんははしょられたのですが、ビッグデータというのは、今、確かにいっぱい出てきておまして、民間が集めている個人の行動、車の運転履歴などがあります。トヨタとか、ホンダなどの情報処理センターに自動的に情報が集まるということが、実際に起こっております。そうすると、例えば今までは自動車保険や損害保険は、年齢とか、免許歴とか、事故歴とか、そういうことをインプットにして、平均的な保険料金を決めたりしていたわけですけども、もっと個別のドライバーの運転のスタイルとか、どれぐらい危ない運転をしているとか、どれぐらい安全な運転をしているとか、そういうことで、もうちょっときめ細かい保険料の設定ができるという話が出ていました。

あとは、電力の消費でも、スマートメーターというのが、今、アメリカで使われているのですが、そういうもので、実際に個々の家庭の電力消費量がどうなっているかということが、リアルタイムで発電所にいけばわかるんで

す。発電所が電力供給をさらにきめ細かくコントロールできるということが行われていて、ビッグデータというのは、確実に我々の生活を変えていくわけです。

経済施策の中でも、そういうものを使って、いろいろ考えていく余地はかなり多いので、そういうことを取り込んでいくことも1つの課題だと思います。

もちろんこれについては、秘匿といいますか、個人情報の問題とか、いろいろなことがあるんですけども、これを使わないで、従来どおりの政策をやっていくことが効率的かという、そういうことではないだろうということなので、ビッグデータについても、マイクロ統計の枠の中で考えていけばどうかということです。こういうことが、データについての幾つかの論点になるかと思います。

次の論点は、マイクロ計量分析を政策へ反映させるには、どうしたらいいかという話です。これも市村さんが丁寧にお話になったんですけども、実際、学会の中で、マイクロデータを使った実証研究というのは、確実に増えていまして、問題意識とか、アプローチ、それを使うソフトウェアとか、プログラムそのものもかなり共有されてきております。ですから、データの分析の手法なども、いろんな人が使えるような状況になってきている。

同時に、実証したい経済理論に基づくマイクロデータと、調査されているデータが必ずしもマッチしていないということもありまして、例えばパネルデータで見られるものと本当の理論、見たい行動の変化みたいなものとの1対1対応ができていないような場合がありますので、そういうときに、どういうふうに工夫してやっていけばいいのかということが、大きな問題としてあるわけです。

自分たちでデータはつくられないので、今までは与えられたデータを工夫して、理論に合わせたような形で実証していたわけですけども、今のお話にあったような、自分たちで制度設計・調査設計までして、理論的に見たい現象を細かく調査して、それを見ることができれば、より精緻な政策分析ができる。自然実験とか、社会実験と呼ばれていましたけれども、そういうものを使って研究していくという方向性も1つ出てきている。

必ずしも全ての実験をできるような環境にはありませんので、例えば税金についてのリアクションみたいな、特定の年だけ違う税率を適用するとか、免税制度を導入するということは、社会的にしにくいところがありますから、限られていることではあるんですけども、そういう形での政策への反映は考えられています。医療とか、そういうところでは、より使われるようなアプローチかと思います。

実証結果が政策に反映されるためには、政府の問題意識が十分に把握されていて、求められている政策分析を行う必要があるわけですけども、往々にして、学会サイドで持っている問題意識と、政府サイドで持っている問題意識が

一致していない。学会では理論的により精緻な推定方法とか、より安定的なパラメータを求めるとか、政策当局が必要としているものと違うような関心を払うことがあります。

政策当局も関心事が必ずしも安定しているわけではなくて、そのときに応じて、こういう論文はないかとか、こういう現象について研究した結果がないかという形で、学会にアプローチされることが多いと思うんですけれども、そういう問題意識みたいなものをどこかで共有しておく必要がある。どちらかが歩み寄るといよりは、お互いに切磋琢磨して、意識を高めておく。

政策当局については、何かが起こったら慌てて調べるのではなくて、自分たちも常に学習をして、問題意識を事前に予測して実証しておくとか、学会に注文しておくとか、ある程度予見性があるようなことをやっておられると、本当に必要なときにタイムリーに情報が出てきますけれども、何か事が起こって、こういう結果がないかと探すということでは、ちょっと遅い場合もあると思われます。

もう1つ思うのは、我々の大学で政府のPhDを持っていない人を受け入れるとか、PhDのトレーニングを大学で施すということが起こっていますし、大学の先生が、役所なり日本銀行に行って政策提言をしたり、研究員を一定期間やるということが起こっているわけですが、政策を決めるという意味では、政治家の人たちと学会、あるいは政治家と官学の間での意思疎通がどうなっているのかということが、重要なんだろうと思います。学者にとっては、直接政治家と話をして、政策議論をする機会はあまりないので、それは官でつないでもらうというか、その情報がある程度伝えるような役割を担っていただければと思います。

最後に、将来どうするかということですが、結局はマイクロデータの収集とそれを使っての実証研究、それを政策に反映させるということを地道にやっていくしかないと思われます。それを今後も長く続けていくしかないですし、できる限り教育をして、次世代を担うような若手を育てていくことが我々の仕事だと思っています。

ただ、日本は、デフレや高齢化、膨大な公的債務、超金融緩和の結果と、課題をいっぱい抱えている国であります。それをどういうふうに解決して、政策立案に結び付けていって、経済的に問題を起ささないように社会を収めていくかということは、ものすごい大きなチャレンジでありまして、それについて、若い研究者、若い官僚、若い政治家が取り組むということは責務だと思いますし、彼らがこういうことをできれば、本当に世界中の人が耳を傾けてくれるような話になるだろうと思います。

基本的に日本の研究がおくれている中で、課題については一番進んでいるわ

けですし、解かなければいけない問題がたくさんあるということがわかっているわけなので、それにチャレンジしないということはないですし、そういうことをやっていただきたいと思っております。

以上です。

【モデレータ】ありがとうございました。

若い方だけではなく、北村先生には、まだまだやっていただきたいと思えます。次に大竹先生、お願いします。

【大竹文雄先生】 私からは、労働分野を中心に、どんなふうに政府統計のミクロデータが使われていたかという話をする予定です。

政府統計は、最初の市村先生のお話にもありましたけれども、2つのタイプのデータがございます。1つは、政府による標本調査です。労働力調査や賃金構造基本統計調査、ここに挙げた様々な標本調査があります。もう1つは、市村先生から詳しくオーストリアの例で年金のデータのパネルデータを使ったような分析の紹介がありましたけれども、私のスライドでは、それは業務データと言っています。最初の西川さんのお話では、行政記録情報と言ったものが当たりますけれども、業務上生じてくるようなデータです。労働分野では、職業安定所業務統計とか、雇用保険業務統計といったものがそうですし、税金だと、税務統計、社会保険関係だと、厚生年金保険とか、国民年金事業年報で出している、年金の記録、データです。こういった2つのタイプのものがあります。

3ページにいけますけれども、標本調査の個票利用につきましては、昔に比べると、随分使われるようになりました。私は1980年代ぐらいから使いましたけれども、そのときは、一大学教員が使える状態ではなくて、当時は、経済企画庁の経済研究所のプロジェクトとして、特別集計という形でデータを申請して、申請してもらえるようになるまで半年以上かかってというものでした。

それが、統計法が変わって、オープンになってきた。ここに書きましたように、現在、研究者にとって、申請の方法が非常にオープンになって、利用が進んできています。特にこういう情報が入るようになりました。当該学術研究に公的研究費補助金、文科省の科研費とか、厚生労働科研が交付・補助されている場合は、利用できるという規定のものがあります。そうでないものもありますけれども、かなり増えてきました。それで利用が可能になってきて、労働関係でもそういったシステムを使った研究が急激に増えています。

2014年のJournal of the Japanese and International Economicsという雑誌を見ると、こういった政府統計の個票を利用した研究として、4件挙がっています。菅原さんと中村さんの介護保険が労働供給に与える影響を分析した研究は国民生活基礎調査によるものです。別所さんと林さんの配偶者控除の労働供給は、就業構造基本統計調査、中島さんと田中さんの出産奨励と居住選択と出生

という研究は、家計調査を使ったものです。中村さんの親の所得と子供の健康の関係は、国民生活基礎調査によります。

このほかにも、大学が集めているマイクロデータを使ったものや、それ以外のマイクロデータを使ったものがありますけれども、行政関係のデータを扱ったものだけでも、これだけの研究例があります。

ただ、先ほど市村先生がおっしゃったとおり、利用が可能になったとはいえ、その手続きはかなり面倒くさいのは事実で、昔、私が利用していたときには、相当嫌がらせを受けた記憶がありまして、今はそうではないのですけれども、研究する前にあらかじめアウトプット表を全部つくらなければいけない。それをつくったらつくつたで、こんなことをしても意味がないと統計の担当さんに怒られながらやるということで、私自身は二度とやりたくないという、つらい記憶があります。その記憶があるから、最近はあまり使っていません。

今、非常に簡単になったとはいえ、手続きはまだかなり面倒ください。どういうプロポーザルを出さなければいけないのかというのは、ウェブにありますけれども、それはアイデアレベルのものを、これでテストしてみようということができるようなものではありません。

もう1つ、私の言葉で言うと業務データ、行政記録データは、日本ではほとんど研究が進んでいません。

私が知っている例外は、労働分野で、小原さん、佐々木さん、町北さんの3人が一連の研究をやっています。後で詳しくお話しますが、彼らは職業安定業務データと雇用保険データをマッチさせて、同じ人を追跡できるような形の情報を使って、先ほど市村さんがおっしゃったようなことに近い、失業期間と次の仕事の安定性の関係を分析した研究をやっています。ただ、それ以外、こういう研究がなされた、こういう業務データを扱った研究がなされたというのは、知りません。

公的年金に関する業務データが利用可能になれば、同じ人をもっと長く追っていけますから、飛躍的にこの分野の研究は進むのではないかと思います。あと、税に関する業務データも重要だろうと思っています。

例えば最近、話題になった、トップ1%の都民の資産所得の集中も、同じ人がずっとトップ1%にいるのか、そうではないのかというのは、非常に重要なテーマです。非常に少ない人かもしれないのですけれども、そういう人たちを追っていくようなデータがないと、分析できないですし、逆に貧困の問題につきましても、同じ人がずっと貧困のところにいるのかどうかということは、わからないといけません。ところが、パネルデータといっても、貧困の人たちやトップ層の人たちを新たにサンプルに入れるというのは、かなり難しい。そこは業務データの強いところで、そういった情報を研究者が利用できるようにな

ることが、この分野の研究を促進する上では大事だと思っています。

もう1つ、サンプルデータの中で、政府がパネルデータをつくっているものもあります。厚生労働省の21世紀出生児縦断調査とか、21世紀成年者縦断調査・中高年者縦断調査というものがあるんですけども、これは利用者が非常に少ないです。野口さんが使われていますけれども、社人研、国立社会保障・人口問題研究所以外では、フルのデータセットを出しているところは少ない。少し例はありますけれども、せっかく政府がつくっているパネルデータがあるにもかかわらず、利用がまだ難しい。

平成24年度に厚生労働科研プロジェクトで、これについての研究課題を公募するということがありましたけれども、普通ではまだ使えない状態です。せっかくある試算なのに、非常にもったいないと思っています。

そういうことで、データの利用は十分にできるようになってきたんですけども、まだまだ障害が大きい。潜在的に日本の行政データには、かなりの価値があるにもかかわらず、その利用がまだ進んでいないという形になっているかと思います。

【モデレータ】 ありがとうございます。それでは、野口先生、お願いします。

【野口晴子先生】 お手元の資料前半部分は、市村先生がお話になった内容の実例といいますか、事例のような形のお話になっております。時間も限られておりますので、主として、後半部分を中心にお話をいたします。

私は医療について何か話せというお題をいただいたので、医療は何が問題かといいますか、医療でアウトカム、プログラム評価をする場合、あるいはコスト・ベネフィット・アナリシスなどをする場合、何が一番難しいかという、やはり医療サービス市場の不透明性なんです。これは質の継続が非常に難しいということになります。

医療評価の指標は、かなりいろいろな指標が発展的に形成されているんですが、いまだに基本はDonabedianの1969年の3つの指標、構造指標、経済指標、効果指標が医療の質をはかるための基本として定着しています。

まず、構造指標（Structure）の部分です。例えば病院や介護施設の土地や建物の面積、病室・病床数といった物的資本、あるいは患者1人当たりに対して、医師が何人いるか、看護師が何人いるかといった人的資本です。こういった構造指標でまず測るということです。この部分については、日本もデータを70年代ぐらいから、ずっと集めております。

2番目は、経過指標です。これは、最近、病院ランキングみたいなことをして、その病院では、1年間にどのぐらい手術をしているかとか、どのぐらいの治療をしているかとか、そういったことでランキングをしているような本もござ

いますけれども、いわゆる治療の内容です。ここでは便宜的にhigh-tech treatmentsとして、開発費用、固定費用と限界費用が巨額にのぼるような治療と、主に薬剤によるlow-tech treatmentsに一応分けております。こういったものが、どのぐらいその医療施設で実施されるのかということを目印にしています。

日本は、医療サービスがマーケットメカニズムによって決まっておられません。公定価格ですので、中医協という現場で、いわゆる医者と患者と保険者が集まって、議論を喧々諤々でやって、公定価格を決めるというプロセスがあるんですけども、そこで最も問題になっているのが、効果指標です。例えば診療報酬に対する価格の上限がどのぐらいアウトカムに影響を与えるか、あるいはあるプログラムがどのぐらいアウトカムに影響を与えるか、そのアウトカムを一体何ではかるのかということ、非常に大きな議論になっているわけです。これが非常に難しいわけです。

医療の質に関する事で、専門家間でのコンセンサスと理解が一致していなかったということがあります。ただ、これは、昨今、医療経済学をやる人たち、いわゆる社会科学系の人たちと、お医者さん、医療系、看護師さん、薬剤師さん、そういう人たちが集う場が徐々に増えてきていて、そういった意味では、コンセンサス、合意形成が、徐々にではありますけれども、できつつあると思います。例えば市村先生がやっていらっしゃる、JSTARというプロジェクトがありますけれども、このプロジェクトにおいては、医師あるいは自然科学系の方も参画されておりますので、そういった部分で、いろいろ場ができてきていると思います。

2番目はデータです。これは一番最後でレセプトのお話をしますが、レセプトに関しても、2000年にレセプトの検討会が立ち上げられたわけですが、そのころから比べると徐々にではありますけれども、整備が進んでおりますし研究者による研究も始まっているという状況です。

3番目は、市村先生が中心にお話された、統計学的手法の確立に関する問題です。資料の12ページ目を見ていただくと、実験データと非実験データの違いがあります。医療に関しては、先ほど北村先生のお話にもありましたけれども、ある治療が有効かどうかという治験を実験でやるのが、ゴールドスタンダードになっておりますので、実験のデータがほかの社会科学の現象に比べると手に入るということで、実験のデータと非実験のデータ、例えばこれはアメリカのレセプトデータを使った事例ですが、比較可能なのです。どのぐらい非実験データに問題があるかということが、目に見えて、推定することができます。

これを見ていただくと、これはあるハイテク治療が、どのぐらいアウトカムに効果があるかということをはかるために、1つは、左側の2行になりますけ



れども、いわゆるRCT、実験をやったケースです。ランダムイズされていますので、患者さんの属性、年齢、女性、こういった基本的な属性から、いろいろリスク要素があるんですけども、実験ですから、当然ですが、ランダムイズされて、2群でほとんど差がない。これは先ほど市村先生がおっしゃっていた、実験によるランダムイズになります。

隣の2つを見ていただくと、これがいわゆるアメリカのレセプトデータになります。ですから、非実験データです。これを見ていただくと、明らかに積極的な治療をやったほうが左、やらなかったほうが右なんですけれども、明らかに積極的な治療をやったほうが、やらなかった人よりも若くて、女性が少ない。女性の場合は、血管が非常に細いので、リスクが伴うわけです。しかも、ほかのリスク指標を見ていただいても、積極的治療をやった人たちのほうが、圧倒的に重症度が低いし、疾患もないということで、一言で言ってしまうと、症状が軽い人たちです。普通の観察データであると、明らかに積極的な治療が死亡率を引き下げる。積極的な治療が、患者のアウトカムの向上に役立つという結果が出てくるわけです。

ところが、ランダムイゼーションによる結果を見ると、死亡率に対して、統計学的に有意な影響が全くないか、あるいは死亡率を引き上げてしまうという結果が出ています。これは先ほどの市村先生のお話にあった、観察するべきものがはかられない、あるいは非実験データを使って、非常に単純なやり方をしてしまうと、恐ろしい。この場では、患者さんの命にかかわることですけども、これが社会におけるプログラム評価においても、多大な人数の人たちに影響を与えることですので、非常に恐ろしい話だという、1ついい事例だと思います。

また話を飛ばしますが、それを統計的に工夫してあげる。例えば操作変数法というやり方、あるいはNatural Experimentを使うことによって、ランダムイズを疑似的にすることができるということで、結果もランダムイゼーションのRCTとほぼ変わらないような結果を出すことも可能ですという事例です。

先ほどの実験の例で、これは実験の結果をお見せするというよりは、これがどういうふうに関米政策に反映したのかということを中心にお話したいのですが、これはMITのFinkelsteinという非常に有名な経済学者がいますけれども、オバマケアのときに、国民皆保険制度を入れることが、アメリカの国民にとって有益であるのか。しかも、コストベネフィットに見合っているのかどうかということ、オレゴン州の実験によって確かめたといった実験であります。

この実験の結果をまとめると、低所得層の精神面・心理面での健康の改善には役に立つ。しかも、貧困層の家計に占める医療費の割合を軽減する、あるいはその人たちの経済的な負担を軽減するという意味では、非常に役に立つとい

う結果が出ています。ただし、その場合、例えば糖尿病の比率であるとか、病気の比率であるとか、そういったものには影響がない。あるいはアクセスはすごく上がりますから、保険を与えることによって、医療サービスに対するアクセシビリティが上がりますから、医療費は非常に伸びてしまう。そういう結果が、この実験で得られています。

結果はともかくとしても、我々がここで知るべきは、これがどう使われたかなんです。これがNew England Journal of Medicineという非常に権威のあるアメリカの雑誌に載ったのが2013年3月になります。この結果が、同年5月のオバマ大統領の大統領レポートで掲載されている。実験ですから、ピュアな環境の中でやられたもので信頼性は高い。そういった実証結果、エビデンスを受け入れて、いわゆる反対派と賛成派が議論をしている。

例えば反対派にとってみれば、身体面での健康増進の効果が認められなかったことをもって反対をした。一方で、賛成派は、家計負担の軽減が認められ、低所得者の医療需要が増大し、医療サービスのアクセスの向上を評価する。つまり今は医療費が多少かさむかもしれないけれども、低所得者層の子供たちが、将来的に人的資源を充足していく可能性があり、それはアメリカにとっていいことである。同じ客観的・科学的なエビデンスを持って、政策の議論がなされている典型的な事例だと思います。これができるアメリカはすばらしいと思います。

時間がなくなりましたので、現状をちょっとお話しますと、日本もレセプトが徐々に使われるようになります。研究者に使われるための体制は、あと10年ぐらいかかるかもしれませんが、整いつつあるという、これが法的な状況です。

ただ、地域によっては、紙媒体で集められているところも多く、地域差があるということです。歯科にいたっては、まだ紙媒体で収集されているという状況です。

提供されるデータの種類は、特別抽出から集計表情報ということで、軽さ、重さがあります。特別抽出というのは、非常にストイックです。これは先ほど大竹先生もおっしゃっていましたが、まだまだ限られています。特にお医者さん、大学病院しかアクセスできないような状態です。集計表情報、あるいは基本データセットは、経済学者でもアクセスが可能な情報です。大竹先生も先ほど述べられていましたが、特別抽出のデータは、探索はできないような状態で収集しなければいけない。治療の効果を自分がある程度わかっているような状態でしか、申請ができない、非常に厳しいデータです。ですので、これを探索的にしていく必要があるということです。

それに対して、これはアメリカのレセプトデータの使用状況ですけれども、

アメリカではここに見られるように、Federal/Local Governmentの人も利用しているし、民間企業も利用しています。レセプトの民間企業への提供は、今、厚生労働省でも議論が始まっています。

これが意外に驚くべきことですが、利用者の割合のうちUniversityとCollegeは全体の21%しかないんです。それぐらい、アメリカはレセプトデータをみんなが活用して、医療政策に資するようなエビデンスを出そうと努力しているということだと思います。日本は、今、社会保障が抱える問題が大きくて、例えば高齢者はあのサービスもほしい、このサービスもほしい、自己負担で払うのは絶対に嫌と騒いでいる。一方で、若者たちはいっぱい高齢者を背負っていたり、金銭的な介護負担とか、老老介護とか、これらが非常に問題になっている。医療者は、慢性的な人的資源不足で、低賃金、過酷な労働環境にあり、縦割りです。いわゆる無謬性という議論が出ていましたけれども、失敗が許されない、修正は許されない世界で闘っています。研究者は研究者で、市村先生から怒られるかもしれませんが、病的にスタンダードジェネレーションにこだわるということで、この変数がない、あの変数がない、スタンダードジェネレーションが0.1上がった、下がったが大事ということで、全員でばらばら感があるということで、こういった人たちをまとめていく、一人一人がやっている重要さを納得して、みんなが何らかのコンセンサスをつくり上げていくことが、1つ大きな今後の課題だと思いました。

以上です。

【モデレータ】 野口先生、ありがとうございます。

ランダムマイゼーションの具体的な例とか、オバマ経営の実情といった、具体的なことをわかりやすく示していただいて、ありがとうございます。

最後のわかりやすいイラストは、私が最初に説明した、データの整備、研究、政策の3つの分野が縦割りで、ばらばらにやっているといったものと共通するものがあって、強く共感いたしました。

それでは、パネルディスカッションに移りたいと思います。

最初、私が整理した資料で恐縮ですけれども、3つの分野の整理に基づきながら、補足的に皆さんの御意見あるいは御質問をしていただきたいと思います。

資料5の2枚目のスライドで、一番左に書いてあります「I ミクロデータの作成・収集・提供」に関して、皆さんの報告の中で私の言葉では行政記録情報、行政データの有用性について、いろいろと議論がありました。今年1月、アメリカの経済学会で議論では、1つはデータ収集のコストが小さい。もう1つは、市村先生も強調されたランダムマイゼーションについて、全数調査なのでパネルデータなどと比べると、はるかに偏りが少なく、中立的な結果が得やす

いということがありました。日本で行政記録情報を使って何ができるのか、あるいは具体的にこういうことがわかるようになってきたという例はあるのでしょうか。大竹先生は説明を略されてしまったので、追加的に御説明していただきたいと思います。

それから、統計委員会の議論などを傍聴していると思うんですが、アメリカの場合、行政記録情報の利用について、個々人の研究者ではなく集団でアクションをとっておられるように思うんですけれども、今後、集団でアクションをとるような可能性があるのかということについて、どなたからでもよろしいですけれども、少しお話いただければと思います。

【大竹文雄先生】 1つは、日本の研究で、小原さん、佐々木さんたちがやった研究で、ハローワークの職業安定業務統計と雇用保険業務統計が用いられています。先ほどの市村先生の話だと、失業期間が長いと、よりよいマッチングになるはずだということなんですけれども、本当にそうかという研究です。できる範囲のコントロールで分析すると、長く失業した人は再就職後の勤続期間が短く、必ずしもマッチングがよくなっていないということが、彼らの研究結果です。

次に出していますのは、アメリカ、カナダ、スウェーデンのデータなどを使っているもので、長期の所得変動についての分析です。子供の所得と親の所得の関係、階層間移動の関係を分析しています。同じ人をずっと継続して観察できるような社会保障庁、あるいは税のデータを使った分析がなされています。

先ほどのプレゼンでも申し上げましたけれども、今、所得が高い人はずっと高い、あるいは今、所得が低い人がずっと低いかどうかということが、大きな問題ですけれども、普通のデータではうまくとれません。そこは年金や所得を毎年記録できるようなデータを使っていくことが、かなり大事だと思っております。

【モデレータ】 ありがとうございます。

例えば世代間の格差という言葉を使ったときに、今、40代の人と、今、20代の人が40代になったときの世代という言葉は、区別して使われなければいけない。そういう時系列的なもの、特に最近の経済学では、ライフサイクル的な観点を入れた状況の認識や判断みたいなものは、非常に重要です。そういう意味でも、行政記録情報を時系列的に使うことは、非常に重要かと思っています。

北村先生、統計委員会に御出席されていますが、特にございますか。

【北村行伸先生】 今、大竹さんが説明された、小原さんと佐々木さんの業務統計については、データ処理を私のところでやったんですけれども、業務統計は統計分析をするのに、データの入れ方を変換するのに、とてもコストがかかります。統計として使うことをあまり想定していないデータなんです。業務統

計なり、行政記録を統計データとして使うためには、ある程度、二次利用とか、三次利用も含めたフォーマットを、統計として変換しやすいような形にしておいていただかないと、なかなか進まないだろう。しかも、それが各都道府県やその他の自治体、役所で違うということになると、読み解くだけで、ものすごいコストがかかることになりますので、その辺の議論がすごく大事だと思っています。

あと、この間、統計委員会で、地方自治体の公務員の方の給料というのはちゃんと記録があるので、わざわざ基幹統計で調べる必要があるのかというのがちょっと議論になりました。そういうデータについては、恐らく簡単にトランスファーできるようになっていて、既に利用できる。すぐに利用できる形の行政記録と、統計に変換するのに大変なものがあって、利用促進をこの面で本当に進めるのであれば、統計として使う可能性も含めた入力の仕事みたいなもの、基本的なフォーマットを考える必要があるのではないかと思います。

【モデレータ】 ありがとうございます。

本格的な管理、利用促進のための仕分け、性格のふり分け、様式といったことの議論が必要だという御紹介だと思います。

野口先生、どうでしょうか。医療の分野で、行政記録情報を使って、具体的に研究をしておられる立場から、お願いします。

【野口晴子先生】 まさしく、今、北村先生がおっしゃったとおりで、例えばレセプトデータが整備されつつありますけれども、それには、膨大な時間と人的資源を投じる必要がある。ただ、その点に関しては、アメリカのメディケアなども最初は次官がかかっていましたけれども、最近は、行政の側、例えばCMS、Centers for Medicare & Medicaid ServicesにPhDの人たちが10人ぐらいいて、彼らがデータを使ってログを書き、データのクリーンナップを行っている。行政の中に研究者を取り込んで、そこでクリーンナップをするというシステムができ上がっているのです。90年代からです。

例えばそういったシステムを日本で、厚生労働省の中に、計量統計なりを学んだPhDの学生が10人ぐらい入れるポストをつくっていただければ、そういうシステムができるのではないかと考えております。

【モデレータ】 ありがとうございます。

市村先生、何かございますか。

【市村英彦先生】 業務データを使ってできる分析としては、大竹さんも話しておられましたけれども、生涯所得が計算できるようになれば、生涯を通じての所得にどういう違いがあるのかということが、世代ごとにわかってくるわけです。そういうことがわかれば、相当いいと思います。しかも、非常に高い所得の人たちも、非常に貧しい人たちも、全面的にわかってくるわけですから、

すごくインパクトの大きな研究になってくるのではないかと思います。

それから、パネルデータなどを集める際に、所得の情報というのは、回答者は覚えておられない場合も多いですし、集めにくいデータなので、こういう質問に関しては、アメリカのデータなどでは個人が許可してくれた場合に、年金局のデータと結び付けて集めることができるようになっているわけです。そのように、個人がいいと言ってくれた場合には、政府のデータと突合してもいいという形にしていただけると、データ集めの際のコストが下がりますので、ありがたいと思います。

【モデレータ】 北村先生、どうぞ。

【北村行伸先生】 今の市村さんの話に関連しているのですが、北欧の統計システムはいろんなものにマッチングできるという話で、有名になっています。どうやってできるかという、マッチングしたデータがあるわけではなくて個々の統計が別々にあり、マッチングしようと思えば、例えば所得のデータと年金のデータをくっ付けることができる。業務統計は簡単にマッチングできるような仕組みがそもそもつくられていて、また、それを想定した上で統計がデザインされているところがあります。

北欧でなぜそういうことが進んだかということ、人手不足ということがあると思います。予算がないとか、統計の行政にコストがかけれないといった場合でも、行政記録を使えば、それにかわるような統計が得られますという仕組みを想定してデザインされているところがあるので、そういうところは、学ぶべきだと思います。

【市村英彦先生】 もう一点、私が日ごろ思っていることですが、今、死んだときに、自分の体を献体できます。それと同じように、政府側でもってもらえるデータを、私が許可した場合、私に関するものは、全部ZIPか何かに入れて提供できるという形をつくっていただければ、10年もすれば、相当なデータが集まるのではないかと思います、データ集めの苦勞が相当減るのではないかと思います。

【モデレータ】 今、まさに市村先生のお話を聞いていて思いましたが、献体をする方というのは、医学教育の現場に対する強い信頼があって、はじめて献体をされるので、そのために大学の医学部などはすごく注意を払っているわけです。行政記録情報を献体のように出そう、個人は社会保障の情報を出そうとか、そういうことを考えたときに、分析に使ってもらえるとありがたいと、分析に対する強い信頼があってからだと思います。

次に「Ⅱ ミクロデータを利用した分析の進展」に移したいと思います。分析の中で、皆さんが御発表されたときに、2つぐらい論点があったと思います。1つは、「政策決定」ともかかわってきますが、分析をしたことによって、ど

ういう経済的な費用対効果のメリットがあるのか。行政記録情報を整備するのに、時間とお金がかかるということでございますし、それによって、具体的に労働政策、あるいは医療、年金といった社会保障の改革にどういうメリットがあるのかについての説得力ある説明が、分析をされる方、あるいはそれを使われる人たちに必要になってくると思います。それが1点です。

もう1つは、今の献体の話も適当かと思えますけれども、自分のデータ、いろいろな情報を使われることに関する倫理的なプライバシー、あるいはランダムマイズされた実験をしていいのかといったことについて、お気づき、御関心の点があれば、お願いしたいと思えます。

大竹先生、どうぞ。

【大竹文雄先生】 分析をすることのコストベネフィットがどのくらいあるかということですが、それは間違った政策をしたときに、どれだけコストが大きいかということだと思います。政策が間違っていることを明らかにして、実施しないようにできるわけです。それを明らかにすることが、まさに研究の役割だと思っています。

【モデレータ】 間違っていることを直すだけではなくて、間違っていることを直したときのメリットの大きさが十分大きくないと、手間がかかると考えられるのではないのでしょうか。

【大竹文雄先生】 政策によって、これだけ大きな損失が生じているということがわかることがベネフィットなわけです。それから先、その政策が直るかどうかは知りませんが、それだけ間違っているということを明らかにすることは、大きなベネフィットだと思います。

【市村英彦先生】 いろんな政策のあり方が考えられる中で、どうしましょうかといういろいろ探っているわけですが、何もデータを見ないで物事が決められていくことをできるだけ避けて、今、何かわかる可能性があるわけですから、ベストを尽くして、わかる範囲のことは、一体何なのかということを考えていくためにデータを使いましょうということだと思います。

【北村行伸先生】 政策に関して、決定的に正しいパラメータであるとか、正しい結果を求めるのは非常に難しいと思えます。例えば税については、労働供給の所得弾力性とか、消費の価格弾力性をはかれば、最適な消費税率が決められますという議論があったとしても、特定のデータで、特定の期間について推定したら、こういう値が出ましたが、ちょっと変えると違いますという結果になる。先ほど市村さんがおっしゃられたように、パラメータが状況に応じて変わってしまうと、安定的なことは言えないわけなので、それについては、どういう解決策があるかという、これも市村さんがおっしゃいましたが、一般均衡的なのか、構造推計みたいなものをするというのは、1つの考え

方でしょうし、あとは、いろんな時期、いろんなデータについて、いっぱい結果を出して、それで平均的なところを見るということも大事かと思います。

ただ、たとえ構造推定みたいなことが、一般均衡的なモデルでできたとしても、本当の意味でのみんなの考え方が一気にひっくり返ってしまうようなショックが起こったときに、何ができるかというところ、そこはなかなか難しいところなので、そういうときに、常識というか、ある程度、政策判断ができるために何をすればいいのかというのは、普段の蓄積をするしかありません。本当に構造がひっくり返ってしまうようなことが起こると、手の施しようがないところはあって、そこは限界もあると思います。

【モデレータ】費用対効果について、今、何も見ないで、間違っているかもしれないと考え、費用、まさに機会費用をベンチマークに考えるべきだという御趣旨だったと思います。また、統計的な因果推論に一定の幅や条件、コンディショナリティーがつくことについて、広い理解をちゃんと求めていくことが大事だと思いました。

最初に申し上げた、分析を行っていくときの倫理上の問題、特に医療の分野などは、具体的に生のデータを扱う場合、いろんなことが発生しているのではないかと思います。他の分野でデータの利用が広がるときに、例えばこんなことに気をつければいいのではないかとすることがありましたら、野口先生、お願いします。

【野口晴子先生】確かに医療の問題は、プライバシーに抵触することが多々あると思います。マイナンバー制度では、得られる情報には現在のところ医療情報を載せないということに、なって、集約されて出てくる情報の中に、医療がどれだけ入るのかはわかりません。日本国民の感覚としては、医療というのは、一番のプライバシーだと思います。

国民的な感情、あるいは感覚を変えるというか、社会保障政策、医療保険財政、あるいは介護財政というのは、このまま放っておくと、明らかに破綻を来してしまうし、今の医療制度であるとか、介護制度を今後も持続可能なものにするためには、どうしてもコストベネフィットを図っていかなければいけない。そのためには、プライバシーに気をつけなければいけないような個人情報をもとに、いわゆる価格弾力性といったものはかっていかなければいけないということになってくると、そこをどういうふうに国民に対して、あるいは医療政策の担当者に対して説得していくかというのは、非常に重要な問題です。

私自身も手がかりというのは、まだよくわかりませんが、一方で、持続可能なものにするためのコストだ、自分の医療情報とか、自分の健康情報を、重々プライバシーには注意した上で、効果に対して提供するということが、やがては自分たちの制度を守るための1つの手段であるということ、国民一人



一人が納得すれば、私は可能性があると思います。

スウェーデン、あるいは北欧諸国、デンマークなどもそうで、個々の国民が個人情報の提供を認めているというのは、一人一人が、そういった情報を提供することによって、今の制度の持続可能性が高まると考えてのことだと思いませんので、そういった意識の転換を図っていくことが、非常に重要だと思います。

【モデレータ】 ありがとうございます。

今、野口先生がおっしゃったようなことを少しまとめさせていただきますと、個人の情報を使って分析をして政策に生かして、それが情報を提供した国民のために、社会保障なら社会保障の改革に有用であったと得心されるようなサイクルを描かないといけないと思います。資料5にある絵は、極めて専門家が考えているような言葉で説明をしていますので、こういうサイクルを国民に理解してもらおうという観点から整理されていないと、反省いたしました。

では、最後の「Ⅲ 政策決定過程へのインプット」の議論に移りたいと思います。

1つは、人づくりの話、特にこのサイクルを担う人の話が結構ありましたけれども、本日は、パネリストの方には直接関係者はいらっしゃらないのですが、例えば公共政策大学院の卒業生の方で、国家公務員になられて霞が関で働いている方もずいぶん多くなりました。そういう教育、人づくりが直接私どもの職場で生かし切れているかという、そうではなさそうに思います。他方、政策は、ここ10年、20年ぐらいの間に、PDCAの観点から、何をしようとしているのか、いつまでにしようかということ、かなり考えるようになっていて、いつ何が課題かみたいなことは、割と見えるようになっています。

例えば社会保障などは、何年のここらで法案を変えるとか、そういうアジェンダ設定は、一昔前に比べると格段に明るくなっているのではないかと感じております。そういう中で、事前に実証研究を貯めてインプットすることが、必ずしも着実に進んでいるわけではないという御指摘だったと思います。北村先生から、分析あるいは政策の部門というのは、お互いに切磋琢磨していく、あらかじめ用意していくことが大事だという御指摘をいただきました。

具体的に、Ⅱ「研究・分析」とⅢ「政策への反映」のフェーズをどううまく連携していくかについて、改めて皆さんのお考えを伺いたいと思いますが、最初に市村さん、よろしいでしょうか。

【市村英彦先生】 既に始まっていることですがけれども、日本経済学会に政策統計の方々に来て、こういうディスカッションの場で議論していただいたデータがあります。そういう場で、具体的にこういう政策を長期的には考えたいので、そういうことに関する知見を集めたい。例えばTPPに関して、過去から随分話はあったわけですから、日本経済学会でそういう研究を進めておいて

ほしいということが、例えば5年ぐらい前からあれば、今、相当な知見があるはずです。数人の先生方を巻き込むだけではなくて、学会全体を巻き込んで、詳しいデータ、各省庁で使っているデータなども開放して、全体として研究を進めていくような体制を構築できたら、相当迫力のある結果が出てくるのではないかと思います。

【大竹文雄先生】 私も同じように思います。先ほど市村先生もおっしゃったのですが、役所の中で頑張ってもらうことも大事ですけども、やはり学会を利用してもらって、今こういうアジェンダがあってデータも提供しますとして、資金援助もしてもらったら、もっとありがたいので、各省庁、政府側のアジェンダを科研費の特別プロジェクトにきちっと出して、しばらくしたら、ちゃんと業績がたまるような仕組みをうまく組み込んでいくことが、一番大事ではないかと思います。

【北村行伸先生】 今のお二人の意見に賛成ですけども、いろんな知見を集めて公開していくような、例えば個別実証結果が出たら、それを要約して、こういうパラメータでやると、こういう値が出ていますということが、あまりまとめられていないのです。個別の研究でつくったデータは、それはそれでここに置いて、ほかの人がやるときに、また新たにやらなければいけないということで、蓄積型になっていないというか、学会も政策担当者にとっても、情報をシェアする仕組みができていないのです。そういうことは大学サイドと一緒にやって、内閣府の研究所の中で集めてもらったり、サーベイしてもらったりすることは、重要なことだと思います。

【モデレータ】 1つは、こういうふうにしてつくと、マイクロデータ、データベースがつくれます、あるいはそれを預けてみんなで共有しましょうといったことは、私も少しずつやっていますし、他の研究所でも進めているところではあります。

【市村英彦先生】 逆にお伺いしたかったことがあるのですけれども、例えば特定健診や女性の社会進出などに関して、大々的に政策を打っておられるわけですが、それを打つに際して、どういう実証分析に基づいてなされたのか、あるいはそれが十分になされていないのであれば、その理由は何だと西川さんは考えておられるのかという辺りを、お話を伺わせていただけたらと思います。

【モデレータ】 私というか、政策に近い現場にいる知人等といろいろお話をしてみて、こんな感触があります。まず実証分析の研究を見ようと思うと数が少ない。例えば1個しかない、あるいは2～3個しかない。それを見て整理しても、一意に確定しないと落ち着かないということが1つです。ばらつくという感じがあります。

2つ目は、途中でも御指摘がありましたが、インハウスでやったものでない

と、いけないというわけではないのですけれども、誰がやったかわからないから、自分でやるということです。そこを進めていくと、政策の具体的な分野を見ますと、特定の業界、あるいは業種、職業の方と利害があるようなものがある、そういう分野の偏りがあるかもしれないということについては、懸念もあります。そういうことを言って、結局、数が少なくても、見たほうが、とんでもない間違いをしないよりはいいのではないかということとともに、特定のものに引っ張られて、間違えてはいけないという学説に引っ張られているように思います。

さらにもう1つ、そうしたことの背後には、まさにこういうセミナーをしていただいたことの狙いでもありますけれども、こうした実証分析をきちっとやらないと、どういう機会費用があるのか、そういう問題について、行政のトップも含めて、必ずしも十分な理解がされていないのではないかと考えています。

【市村英彦先生】 数が少ないといっても、例えば女性の社会進出に関しては、相当な研究の蓄積があります。日本だけではなくて、海外でもあるわけです。方向性として、それが間違っていると考えているわけではないですけれども、民間に対して相当な変革を要求しているわけですから、社会に対して相当インパクトのある政策を打つに際しては、相当な情報収集なり、何かに基づいたことをされないといけないと思います。それができる仕組みを、我々でつくっていく必要があるのではないかと思います。

【モデレータ】 先ほどのお答えは、いきなりボトムラインのところでお答えしてしまいましたが、その点では政策の透明性の向上のために様々な情報発信をして、皆様からオープンなコメントをいただく機会を設け、研究会のホームページなどにも、一昔前に比べると格段の進歩で、どういうものを参照したかということが公開されているようになってきていると思います。ただ、こうしたパブリックコメントや御意見募集が、形式に流されていないか、あるいは実質的に最も重要な実証分析みたいなものを見ているかということにつきましては、うまくまとめて見せられるようなものが格段に少ないと思います。例えば欧米の研究で、ハンドブックシリーズなどのサーベイなどを見ますと、特に労働の分野などは、4桁にのぼるようなものをサーベイしても、このぐらいの世代間の格差の継承というのは、0.7ぐらいなのではないかといったことが積極的にまとめられています。そこに日本の文献があるかということ、統計の不足もあって不十分であると思います。

#### 4. フロアとの質疑応答

【モデレータ】 そろそろお時間も近づきましたので、フロアから御質問があり

ましたら、よろしくお願いたします。

御質問の際には、お名前と所属を明らかにされた上で、御質問をしていただけると幸いです。

【フロアから】 株式会社価値総合研究所の工藤と申します。本日は、大変興味深いお話を聞かせていただきまして、先生方にお礼を申し上げたいと思います。

質問というよりも、個人的な意見になってしまうかもしれませんが、今、スライドで表示されております、政策決定過程へのインプットに関連する部分、その前の分析にも関連する部分について、先ほど先生方のお話で、献体のような形で、自分の個人的なデータが社会的に役に立つのであれば、それが1つインセンティブになって、データが提供されるのではないかといったお話があったかと思います。その部分に興味を持ちました。

最近、若者の政治への無関心が言われておりますけれども、その理由の1つに、自分たちが関わらないところで政策が決まっている。つまり自分たちは何もできないという無力感があるのではないかと思います。逆に自分たちの情報を提供して、分析して、活用することで、それが政策に生かされて、それが行く行くは自分や将来の世代にメリットになるのであれば、今の若者にとって、政治に関心を持つという意味での動機づけになるのではないかと思います。

そのためには、今、はやりのキーワードを入れますと、見える化が非常に大事だと思います。自分が経験したものが、どういうふうに分析された、それがどういうふうに政策に生かされるかというプロセス全体の見える化が必要だと思います。見える化は、行政だけでとか、分析を行った研究者の方だけで行うのではなくて、全体的に見える化をしていくような仕組みづくりとおっしゃっていましたが、それが1つマイクロデータを政策に生かす上で、必須なのではないかと思います。

あとは、ビッグデータの話が、先生方のお話で出てきましたけれども、それらのデータを活用して、それを企業戦略に生かすことについては、今、知識を普及している状況ですし、若い世代にも、ビッグデータのエネルギーは伝わっている状況だと思いますので、そういう意味では、今、非常にマイクロデータを活用する機運が高まっている状況だと捉えられますので、いろんな観点からアプローチをして、仕事づくりをしていただけて、単に自分のことを知られたくない、個人情報を守りたいというところで終わってしまうのではなくて、国民全体が自分たちのために、自分のデータを生かしてもらえることを実感できるような仕組みをつくっていただければと思います。

【フロアから】 私は京都大学の市村でございます。

長年、マクロの日本経済モデルのことをやってきまして、今日は、ミクロの

データについて、大変勉強になりましたが、両方とも似たような問題を抱えている感じが強くあります。

マクロのモデルは、各種各様のものがたくさんできましたけれども、主なものを整理するという仕事は、やった人もいませんし、今もないと思います。

日本経済研究センターがNEEDSというデータベースをつくりまして、それを売り出し、それは非常に役に立つデータベースですけれども値段が高くて、我々学者はそれを買うだけの研究費がなくて、結局、苦勞して自分自身でつくって、しかも、それは各大学の小さなサークルの中だけで使われて一般には開放されない状況がずっと続いているわけであります。

経済企画庁、あるいは日本銀行がデータのベースをかなりきちっとつくっておられるように思いますけれども、それもほかの人に公開されているとは言えません。たまたま許可を得て使ったら、むしろ非難されることも、過去にありました。ですから、データの公共性をどういう形でつくるのかということは、学会と官界と業界を合わせての非常に重要な問題ではないか。そういうものを、先ほども意見が述べられておりましたけれども、パブリックヘルスとして、うまく提供できるような仕組みは、かなり優れた人たちが寄り合って、工夫を凝らす必要があるのではないかと思います。これは非常に重要な問題だと思います。それが1つです。

もう一点、本日のような会議が開かれたことは、非常に有意義だと思いますし、もっとたくさんの方が出席すれば、学術的だと思いますが、要するにサーベイ論文、それぞれの重要な分野に関するサーベイをして、かなり整備されたサーベイが学会及び一般界に提供されるということが、もっとなされるべきではないかと思います。

サーベイを書く人は、かなりの能力と時間を必要としますので、それに対して、報われるところはあまりありません。1つの学問分野の進歩について、ある段階になって話したときに、アメリカの学会では、サーベイをするということを、有力な雑誌の編集者がかなりのお金を使いまして、特定の人に頼むということをやったことが、何度もありますが、アメリカではなさいますけれども、日本では、私が知る限り、そういうことはほとんどなされません。たまに日本経済学会などがやりましたけれども、あまり先行的ではなかったように思います。

しかし、サーベイをしませんと、若い人が、今、進行中の保険金は、大体どんな段階まで来ているのかということが、わかりません。大学院生等は、自分のやった分野のことだけしか知らないで、ほかの分野のことを知りません。したがって、サーベイをすることに関して、もっといい仕組みをつくりらないといけないのではないかと。

特に学会は、この点に責任があると思います。私を知る限り、今の経済学会の話では、こういった経済的な分野だけではなくて、ほかの分野についても、そういう仕組みがうまくでき上がっていないように思います。それぞれの学会で努力されて、幾つかのものがあることは知っておりますけれども、非常に不足しているのではないかと。データの公表とか、それらの蓄積をどうするか、学問的成果をサーベイすることの重要性に、一工夫がほしいという感想を持ちました。

【モデレータ】 どうもありがとうございました。

最初に御質問いただきました、工藤さんからの若者社会の関心といった、広いコンテキストの中で、統計あるいは情報の収集みたいなことを捉え、見える化みたいなことを考えなければいけないのではないかというお話でしたが、こういう分野ですと、広い社会的な観点でいろいろと研究をしておられる、大竹先生、コメントがありましたら、お願いします。

それから、フロアからの市村先生のお話は、御質問というより、御指摘かと思いましたが、パネリスト側から北村先生、市村先生、個別にありましたら、お願いいたします。

どうぞ。

【大竹文雄先生】 データ提供した人たちに、どんなふうにもその成果があったかということは、示していかないといけないと思いますけれども、個別にどうかというのは、なかなか難しいと思います。

先ほど言い忘れたのですが、経済学の分野は、自然科学の分野と比べると、研究成果をニュースリリースする、記者発表するという習慣・慣習がないので、専門論文を書いて、それで終わるといふところがかなりありました。

特にマイクロデータでの実証分析というのは、いろんな限定はつきながらも、どういう結果であるかというのは、比較的一般向けにもわかるように説明できる分野ですから、そういう一般向けへの広報のあり方というのは、組織的にも必要だと思います。

例えばESRIもそうだと思いますけれども、ワーキングペーパーなどを出されたら、一般向けにわかるような、例えば英語の論文であっても、日本語でちゃんと解説をつけるとか、そういうことをされていません。それは個々の研究者はやっていませんので、そういう努力をすると、政策担当者の目にもより触れるようになるかもしれませんし、データ提供をしてくださる一般の方にも、それを通じて、また、それがマスコミを通じて、より広く知られることになればいいと思っています。そういうことは、今後、学会でも考えていくべきだと思います。

【モデレータ】 データのことについて、北村先生、どうぞ。

【北村行伸先生】 市村先生からのコメントは、おっしゃるとおりだと思っております。特に学会でつくられているプログラムとか、そういうものを共有する仕組みがありません。

1つ例ですけれども、大竹さんもやっていたらよかったと思いますが、全国消費実態調査を使って、可処分所得を計算しないといけないときに、全国消費実態調査には、税とか、社会保障のデータが入っていないので、それを推定してつくらなければいけない。そういうことを、データを使う人たちはずっと続けてきたわけですけれども、なかなかプログラムが共有されなくて、私も若い生徒と一緒にそういうことをやってきましたが、我々がわかる範囲内で、とりあえず、今の時点では、こういうプログラムを書きましたというものがあつたとしても、5年後にまた新しい調査が出たときには、制度が変わっていて、違うことをやらなければいけないときに、また新しく誰かが始めないといけないという、本当に無駄なことをやっていました。こうしたプログラムについて公開して、新しく研究に入ってくる人は、古い情報を改善して、さらによくしていくという、オープンソースみたいな形で、プログラムを改善していくような仕組みというのは、ぜひ必要だと思います。

それから、サーベイは私も重要だと思います。学会のこの人に任せれば、バランスのとれたサーベイをやってくれるという人が、日本にもおられると思いますけれども、最近、そういうことをやろうとする人は少なく、個々の研究に没頭している、あるいは何百、何千もの論文を読み込んで、バランスをとった見方を出すというのは、非常に労力が要る割には、報われないことがあると思います。そういうことはぜひ必要でありまして、そういうことこそ、ESRIの経済分析とか、そういうところでやっていただければと思います。

一橋の経済研究の中でも、調査論文という形で、小さな形ではありますが、一応簡単なサーベイをやっていますが、本格的なハンドブックのような大部なサーベイをやっているという仕組みは、日本の中にはないと思いますので、そういうことはやっていく必要があると思います。

【モデレータ】 例えばアメリカですと、サーベイを書くことが非常に名誉であり、また、非常に有用なことだと思います。

【市村英彦先生】 最初のコメントに関して、私もどういふふうに、皆さんの興味というか、データ提供に関する興味を喚起できるかと思っています。

すでに議論した点ではあるのですけれども、政府の予算の議論をマスコミなども通して、各省庁が財務省とどういふ折衝を行っているのかという点を、かなりオープンにして、真っ当に議論を闘わせることになると、国民も相当興味を持って、この話を聞くのではないかと思います。

大昔、担当だった方々に、研究資金の正当性を議論し、財務省に対して、ど

ういう苦勞したかということを知ると、かなり真つ当な闘いを両者でやっているという感じを受けますので、あれを新聞などで報道し出すと、国民も相当興味を持って聞くのではないかと思います。

2番目のコメントについては、私も同様に思います。JSTARに関しては、4回目までデータを公開して、自由に使えるようになってはいますが、それを利用する際のプログラムまで提供しているかということ、そういう制度はありません。特に拒んでいるわけではありませんが、制度としてまだ整備されていない面があります。今後プログラムまで提供するように工夫していきたいと思っています。サーベイに関しては、全く同様に思います。

【モデレータ】 ありがとうございます。今日、パネルディスカッションに上がっていただいた方は、分析をしておられる研究者の方々です。統計あるいはマイクロデータ、分析、政策といったことを、全体像としてどう考えていくべきかということ議論したのは、我々も初めての試みだと思っております。

最後に一言コメントを、野口先生から順にお願いできますでしょうか。

【野口晴子先生】 本日は、貴重な機会を与えていただき、どうもありがとうございました。

フロアからのコメントにもありましたように、マイクロデータは社会的なインフラだと思っています。学問的な意味もそうですけれども、政策的にも、こうしたインフラを我々は構築していく必要がある。そうでない限り、少なくとも医療と介護、社会保障制度に関しては、持続可能性などを科学的に検証することは、不可能であると信じております。

また、フロアからのコメントにもありましたサーベイですけれども、医療に関しては、メタアナリシスとあって、例えばある治療、ある資源に関して、100とか80とかいう論文を横に並べて分析するという研究も、トップジャーナルなどで非常に評価されるわけです。例えば『Science』であるとか『Nature』、『The Lancet』であるとか、そういうところで医療に関しては評価される文化があります。そういったところを多分、社会科学である経済学でも、ある程度日の当たる研究にしていくべきなのかなと感じました。

もう1つは、政策と研究をつなぐ人として、先ほど市村先生がおっしゃったみたいに、すぐくつなぐのが難しくて、そこにも Ph.D が必要だと思うのです。そういった橋渡しをするインタープリターとしての人材の育成というものが、研究者の育成のなかでやっていかなければいけないのではないかと感じました。本日はどうもありがとうございました。

【大竹文雄先生】 今日の議論を聞いていて、データがないからその研究者が要る。研究者が少ないから分析結果も少ない。分析結果が少ないから政策にも使われないという悪循環になっていると思うのです。それをどこかで断ち切ら



ないと事実に基づいた政策につながる、いい循環にいかないと思います。1つは市村さんもおっしゃったように、潜在的には優秀な人材は海外にもたくさんいるので、そういう人材を集めてくる。それから、データをもっと使いやすくするという、そういう何らかのビッグプッシュがないと悪循環を断ち切ることができないと思いました。こういうシンポジウムがそういうきっかけになればというふうに思った次第です。

【北村行伸先生】 今、大竹さんがおっしゃったようにデータの二次利用というのは非常に重要だと思っています。私はそういうところで責任を持って仕事をしているので、進めていきたいと思えますし、マイクロデータの実証を積み重ねていくというのは地道なことで大変なことなのですが、それを続けていくしかないと思っておりますので、それを続ける。

それから、ミクロの話とマクロの話をつなぐような話も大事なわけで、個別の医療とか教育とか家計とか年金とか税とか、そういう話はもしかしたらミクロのレベルで完結するのかもしれませんが、もう少し大きな金融政策とか財政政策とかでマクロな話と、どうやってつないで全体で見るかという話。そういう話はぜひ政府のほうで、ESRIでもやっていただければと思っています。

【市村英彦先生】 基本は人とデータと両方必要なのですが、今、人に関して徐々に教育されつつある段階で、ここで海外から優秀な若手に入っていただけるような環境を私は期待しているのは私立大学の方なのです。米国をみても、国立大学はバークレーやミシガンで、トップはハーバード、イエール、シカゴなわけです。民間のほうが自由度が高くて、いろいろと雇用面などを変革していく力があるはずなのですから、私立の大学にしっかりやっていただいて、若手を呼び戻すような仕組みをつくっていただきたいと私は思います。

それから、データに対して私は少し厳しいことを言ったのは自覚していますが、これはパネルデータを日本で初めてつくっていかれた方々はすごく大変な努力をされて、日本記録が回収率 20%だという段階だという感じだと思います。若い人たちを含めたデータを集めるのは非常に難しいということが今のところわかっているのが現状で、日本記録を世界記録までどこまで近づけていけるかということ、みんなでそういう関係をつくっていかねばいけないというところだと思います。そのためには海外では企業組織のようなものを大学の中につくって、年商 100 億程度の中規模の会社を運営しているような状況なわけですから、それを日本でどうやって実現していくのか。真似事で済むのかという話もありますけれども、まずそれを試してみるのがいいのではないかと私は思います。

【モデレータ】 4名先生方、大変御示唆に富むまとめをしていただきました。私ども経済社会総合研究所、ESRI のいろいろな事業にも重要な御示唆をいただ

いたと思いますし、他の政策部局にも今日の御議論を共有させていただくようにしたいと思います。

それでは、最後に名誉所長である清家先生からお話を申しあげたいと思います。どうぞこちらにお願いします。

## 5. 総括

【清家篤 経済社会総合研究所名誉所長】 今日のパネリストの方、大変興味深いディスカッションありがとうございました。

政策とマイクロデータは関与しているのではないかという話が出ておりましたけれども、私の理解ですと、もともとこのマイクロデータというのは政策的な要請から出てきた部分があると思います。冒頭に北村さんが25年前のことを考えるとおっしゃったのですが、多分、私は北村先生よりも10歳ぐらい年上なので、35年ぐらい前の話も知っているのですけれども、私が大学の学部か、あるいは大学院に入ったぐらいのころ、先ほど市村先生が言われたサーベイ論文、**Income maintenance and labor supply** というサーベイ論文がありました。これはもともと1960年代にアメリカでワーキングプアといいますか、所得補償を受けている人のほうが所得補償を受けられないで働いている人よりも所得が高い。所得補償を受けないで働いている人のほうが所得補償、生活を受けている人よりも貧しいという問題が非常に強い政策的な関心を集めて、そのときに、そうしたらどういうふうに政策をつくったらいいのかというものです。例えば **New Jersey income maintenance experiment** とか、そういうものはまさに実験が行えるようなマイクロデータがつくられて、さまざまな分析が始まったのが1960年代の終わりごろから70年代初めだったのではないかと思います。

最初は非常にプリミティブなものでしたけれども、たしか70年代初めぐらいに論文が出されたりして、だんだんモデルも精緻化してきたということがあったように思います。そういう面では、もともとマイクロデータというのは政策的なニーズから生まれてきているということがあることを指摘したいと思います。

もう1つは、この研究がアメリカですごく進んだのは、データを処理するプログラミングがすごく進んだということです。私も1980年代の初めごろ、私の指導教授が当時の企画庁の経済研究所の客員主任研究官になったので、私はその下働きとして計算のプログラムをつくっていました。当時はパッケージがなかったので **Fortran** とかでプログラムをつくって、特に例えば所得制約線が屈折する場合の均衡点がどんなところに決まるかというものの推定をしたりするプログラムが、それをつくるだけでもほとんど時間が使われてしまうという時代でした。その後、急に **SAS** とかが出てきて、私が80年代の終わりにしばらく

仕事をしたときにはプログラムができていて、要するにヘックマンの2段階推計そのものができてしまうとか、そういうような急速な進歩があって、それがマイクロデータ分析をみんながやるようになった背景にあったと思います。

そういう面では今日皆さん言われたように、1つはこれから日本でやらなければいけないのはマイクロデータをもっと拡充していくということですが、そのときに予算とか資源の制約がありますから、私は1つももっとも積極的にやっていいのではないかと思うのは、市村さんも言われましたけれども、既存データの中で例えば事業所統計などは簡単にと言うとあれですが、パネル化しやすいわけです。事業所番号がついているとか、あるいはセンサスは理論的に言えばパネル化できるわけですから、パネルデータというものを一からつくるコストというのはかなり大変ですが、既存のデータをパネル化したりすることが比較的技術的には余裕でできるので、それは何とか制度的にできるようにやっていけないかなと思います。

同時にそういったデータのフロンティアを拡大すると同時に、我々はそのデータの制約というか限界というものをしっかり、特にこれはこんなことができますよということと同時に、データにはこういう限界がありますということをパブリックにはよく理解してもらう必要があると思います。もちろん市村さんが最初に言われたプログラム評価の場合などもありますし、いろいろなサンプルセレクションの問題があります。データをクリーニングすると矛盾があるデータが全部消えていくわけですが、そうすると、それは真面目に答えた人のデータだけが残ってくるということなのです。そうすると私も以前そういうことをやったことがあって、どうかなと思いましたが、要するに世の中には真面目に答える人とそうでない人両方いるわけで、クリーニングした後のデータからわかるのは、比較的真面目なコンシステントな行動をとる人のデータだということになります。合理的な行動をとる人のデータだけわかるような経済分析には適していると言えばそうかもしれませんが、そういうような面はあって、データの限界を正直に示すべきではないかと思います。

そういう中で私は今、政策的なニーズが非常に高いのは、野口先生が発表されたような医療の分野では、既にレセプトを使って最適な診療報酬を探しているということを考えようとしているわけですが、特に医療の場合には医療サービスを公的に供給するか、あるいは市場に任せるか両極端があるわけですが、恐らくその中間にデータを使って最適な、最も社会全体にとってコストパフォーマンスのよい医療サービスの提供体制をつくっていくことがキーになるかなと思います。

市村さんが最後に質問された点は、とても私は大切だと思っています。しっかりとした分析に基づいて政策が行われる。そのときに何が大切かという点、

要するに非常にイージーなファインディングをするか、ちょっと見たらこうだという観察に基づいて政策がつくられたりするのが一番危険だと思うのです。こういうものは大げさな言い方かもしれませんが、例えば天体の動きを見ると、我々がふだん観察しているのは圧倒的に天動説のほうがフィットしますね。つまり朝起きると東から日がのぼって、夕方になると西に沈んでいく。夜になると星がぐるぐる回る。結構そういう単純な観察に基づいて天動説的な思い込みで政策が主張されたりすることもあると思います。しかし、事実そうではなくて、我々が予想もしなかったような地球は丸くて傾いた地軸の周りをぐるぐる自転していて、太陽の周りを公転している。それは要するに天文学という学問を通じて初めて真理がわかるわけで、そういう何か事実に基づいて政策を進めるといっても、今、見えていることに基づいて政策を進めようということではなくて、サイエンスによって導かれた心理というか、いつでもそれは仮設の段階かもしれませんが、単純なオブザベーションではなく、例えば経済学に基づいてわかったこと。そういうことに依拠した政策が求められるのではないかと思いました。そういう面で私は今日、皆さん方のお話を伺って大変勉強になりましたし、昔、若いときにまだ大学院生だった大竹さんに厳しいコメントをされたなどか、そういったことも思い出してとても感慨深かったです。どうもありがとうございました。

【モデレータ】最初の市村先生の言葉で言えば、構造モデルによる議論をしっかり踏まえて観察をなさいということだったと思います。市村さんをはじめ、4人のパネリストの方々、どうもありがとうございました。

以上