

行政データの研究利用

大阪大学大学院国際公共政策研究科准教授
小原 美紀

行政データを使った分析が急増している。経済学分野で主要誌に挙げられる *Econometrica*, *American Economic Review*, *Journal of Political Economy*, *Quarterly Journal of Economics* において、行政データ (administrative data) をキーワードに持つ論文の掲載数は、1995年から2005年にかけて6本であったのに対して、2006年から2015年5月時点にかけて17本に上っている (キーワードに入っていないものはもっと多い)。この傾向は筆者が専門にする労働経済学分野ではさらに顕著で、*Journal of Labor Economics* と *Labour Economics* では1995年から2005年にかけて8本であったが、2006年から2015年には42本にまで増えている。

行政データの研究利用に注目しているのは学会だけではない。多くの先進諸国で政府統計局が主導となって行政データの研究への活用が積極的に呼びかけられている。行政データの利用がいち早く進んだ北欧諸国、とくにデンマークやスウェーデン、フィンランドでは、行政データの活用促進のためにホームページを充実させているし、イギリスやオランダ、ドイツなど他のヨーロッパ諸国でも行政データを活用した研究が数多く紹介されている。ヨーロッパほどではないものの、カナダやアメリカでも研究への活用は増加の一途にある。

行政データは行政のために登録された個人情報収集したデータである。当然のことながら必ずしも悉皆調査ではなく標本調査である場合も多い。また、複数の行政にまたがり情報を集約することを目指したものではないし、そもそも分析目的のために集められたデータでもない。このようなデータをどのように研究に利用できるのだろうか。第一に、行政データを使えば様々な経済現象に関して長期動向を計測できる。行政データは決められたテーマに関して長期間にわたって

記録されていることが多い。とくに先進国の統計局による調査は、調査項目の一貫性が保たれているので分析が行われやすい。Schmillen and Möller (2012) はドイツの雇用者登録データを用いて、個人の生涯労働期間に対する失業期間の長さをとった生涯失業率を計算し、初職就職時点の景気の悪さが生涯失業率を高めることを示している。

第二に、複数の行政データを接合させれば、経済主体の様々な属性や環境の差を取除きながら知りたい政策の効果を計測することができる。この場合には接合するための識別番号が必要になる。たとえば Corak, Limndquist and Mazumder (2014) は、アメリカの社会保障番号に基づき接合された所得補助プログラムと社会保険に関する行政データを使って、父親と息子の賃金分布の位置と階層移動を分析している。Lechner and Wunsch (2013) は、ドイツにおいて雇用、失業保険、職業訓練、職業紹介に関する4つの行政記録を接合させたデータを用いて、職業訓練の就業促進効果が小さいことを示している。

行政データの利用が近年になって急増している理由の一つは、因果効果や政策の処置効果を抽出するために、「政策プログラムへの無作為割付」を使うことの重要性が高まっているためである。内生性を取除きながら真の因果効果を抽出し政策評価をすることは近年のマイクロ計量経済学の中心テーマの一つである。古くから *Natural Experiment* と呼ばれる外生変数 (たとえば生まれ月や性別、天候) を操作変数に利用した計量分析が行われてきたが、近年では、個人が何らかの政策プログラムに無作為に割り振られたという事実を利用して、外生的な環境差がもたらした結果に注視することで因果効果を炙り出す方法がとられるようになってきている。社会実験と呼ばれるこの方法を用いて真の処置効果を捉えられることができれば誘導形モデルの分析であっても価値の高いものになるだろう。Chetty (2009) が言うように、処置効果が正しく把握された誘導形モデルの分析は、影響の径路がわかりやすいという意味で、様々な仮定を置きながら行動モデル全体を表現しようとする構造形モデルの分析にはない価値を持つ。ここで、社会実験の状況を分析するのに行政データの利用が役立つ。評価しようとする政策割付に直結した行政データが手に入る場合はもちろん、そうでなくても、政策ターゲットとなったグルー

ブを識別する情報があれば、処置や割付を受けた効果を捉えられるからである。

もちろん、注目する政策に対して社会実験の状態が常に存在するわけではない。その場合には非実験データを利用することになるが、ここでも行政データの利用が役に立つ。実際のところ、政策の影響を受けたグループは限定的であることも多い。そしてそれら影響を受けたと考えられる層ほど調査には回答しない可能性がある。行政データは調査を行う組織が確立されており調査対象者が多いうえ、行政が調査をすることで通常は回答率の低い層も回答してくれる可能性がある。影響を受けた層を標本に含みながら全体の傾向を捉えられ得る。たとえば、筆者らが厚生労働省から失業者と雇用者に関する行政データを接合したデータの提供を受けて行った失業者の職探しに関する研究では、失業者は時間をかけるほど定着率の高い仕事に就けるようになるが、失業給付が切れる直前に行動を大きく変えた人にはこの様子は当てはまらないことが示された。この研究は行政データの提供がなければ行えなかった。失業前後の経路を追跡することは多大なコストがかかり、一研究者が時間とお金を費やして行うには限界がある。

利用にあたっては研究者が注意しなければならないことも多い。第一にプライバシー保護の遵守である。このために行われる事前や事後の審査の煩雑さは回避できない。第二に分析対象の妥当性で、調査対象や標本属性が注目する分析テーマと合致しているかを確認することが必要となる。たとえば、格差の分析をするのに特定のグループばかりが捕捉されている調査回答を使えば全体の姿は見えないだろう。第三に分析結果の説明の十分性で、得られた結果がどのような調査対象をもとに得られていて、どのような範囲までが言えるのか（言えないのか）を明確に伝える必要がある。繰り返しになるが行政データは研究分析のために作られたものではない。行政データそのものの問題ではなく、それをうまく使わなかったために問題が生じることもある。調査対象を把握したうえで分析テーマに合った利用を行い、得られた分析結果が調査対象に依存している可能性や留保すべき点を明らかにすることが求められる。

ヨーロッパで行政データを利用した研究が先駆的に行われた理由の一つに、ヨーロッパにはアメリカより

も中央集権的な政策を多く採用している国が多く、行政データの蓄積がスムーズであったことが挙げられる。行政データを使えるヨーロッパ諸国での研究が活発に行われるようになり、遅れをとったと感じるアメリカの研究者らがその必要性を訴えるという流れもある (Card et al. (2010))。行政データを研究に利用できるという理由で、ある国のある政策に関して多くの研究者が参加して精度の高い研究が進められれば、その国にとってより良い政策インプリケーションが導かれる。これは行政データが存在する日本についてもあてはまる。加えて、日本には多くの興味深い経済・社会環境があり、他国には無い先進的な政策も多い。日本の知見は日本だけでなく海外でも活かされる。行政データの今後の利用は大いに期待されている。

参考文献

- Card, David, Raj Chetty, Martin Feldstein, and Emmanuel Saez (2010) "Expanding Access to Administrative Data for Research in the United States," *NSF SBE 2020 White Paper*.
- Chetty, Raj (2009) "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451-488.
- Corak, Miles, Matthew J. Lindquist and Bhashkar Mazumder (2014) "A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States," *Labour Economics*, 30, pp.185-200.
- Lechner, Michael and Conny Wunsch (2013) "Sensitivity of matching-based program evaluations to the availability of control variables," *Labour Economics*, 21, pp.111-21.
- Schmillen, Achim and Joachim Möller (2012) "Distribution and determinants of lifetime unemployment," *Labour Economics*, 19, pp.33-47.

小原 美紀 (こはら みき)