

Econoinformatics

経済情報学研究

NO.117 2022

論文

A. Lewbel の操作変数法

目次

- 1 章 はじめに
- 2 章 操作変数法
- 3 章 Arthur Lewbel の操作変数法
- 4 章 応用の可能性
- 5 章 まとめ

令和4年11月

The Association of Econoinformatics, Himeji Dokkyo University

姫路獨協大学

経済情報学会

A. Lewbel の操作変数法

加藤 善昌

要旨

本稿は、計量経済学者の Arthur Lewbel によって提示された操作変数法の応用性について考察した論文である。計量経済学で重視されている「因果関係」の分析ために、被説明変数と説明変数間の内生性を除去する操作変数法は、現代経済学で広く用いられている。そこで、本稿では Arthur Lewbel によって近年提示された新しい操作変数法について言及し、その手法が今後の計量経済学において持つ可能性と課題を述べる。

キーワード；因果関係, 操作変数法, 内生性, Arthur Lewbel

JELコード；B41, C36, I11

1章 はじめに

現代経済学は、おもに二つの側面から成り立っている。まず、一つ目は理論経済学である。これは、現実の社会や経済主体、また、さまざまな事象を対象として、その行動や特性の一般的な法則性を見出す分野である。このような分析そのものは、非常に古くから存在しているが、現在においては分析として数理モデルとその展開が広く用いられる傾向にある。そして、現在では多種多様な分析対象を設定したうえで、数学的に多様な手法も用いながら分析が行われている。

もう一つは、計量経済学である。これは、理論経済学によって提示された経済学の仮説を、現実を対象とするデータによって検証するものである。このような分析も、理論経済学と同様に古い時期から展開されてきた。そして、近年では、理論分析とともに行われるカリブレーションや、機械学習によるシミュレーションといったものも経済学において盛んに行われているが、これらも計量経済学のひとつとしてみなされる。

このように、経済学の分析は理論と計量（あるいは「実証分析」と呼ばれる分野全般）の双方において日々進展している¹。ただし、計量経済学において中心的かつ基礎的な分析は、現実の

1 一般的に、「計量経済学」というと実証分析のための手法についての分野というようにとらえられることが多い。しかし、「計量経済学」においても、計量経済学について理論分析を行う人々も存在し、一方で、計量経済学を用いて既存の経済学の仮説を検証する人々も存在する。したがって、これらの双方が存在することを考慮し、本稿では「実証分析」も含んだうえで「計量経済学」という用語を用いる。

経済データを用いて、主体間の「因果関係」を推測しようとする回帰分析である。これは、統計学において開発された技法であり、現代の計量経済学において必要不可欠な分析手法である。

回帰分析の最も基礎的な手法としては、予測値と実現値の差である残差を最小にしようとする最小二乗法があげられる。しかし、最小二乗法の信用性については、いくつかの仮定があげられる。そして、その仮定が満たされない場合において使用される手法の一つとして「操作変数法」と呼ばれるものがある。しかし、この手法も仮定をいくつか必要とする。そして、その仮定は分析の多くの場合において満たされることが少ない。

本稿は、操作変数法について今一度整理したうえで、近年注目されている操作変数法の応用性について考察したものである。「因果関係」の分析において、操作変数法は有用な可能性を持っている。しかし、その実効性はかなり限定される傾向にある。そして、そのような課題への対処策として、計量経済学者の Arthur Lewbel によって開発された特殊な操作変数法があげられる。本稿で紹介するこの操作変数法は、操作変数法についての分析だけでなく、計量経済学における分析手法の拡大という面においても貢献する可能性を持っていると考えられる。

本稿の構成は以下のようになっている。2節では、操作変数法について今一度整理し、その可能性と課題について述べる。そして、3節では Lewbel の操作変数法について説明する。さらに4節では、Lewbel の操作変数法を用いた分析例を紹介する。最後に5節では、まとめを述べる。

2章 操作変数法

まず、回帰式を以下のように設定する。

$$Y_i = \alpha + \beta X_i + u_i \cdots (1)$$

i はインデックスを意味している。そして、左辺の Y_i は被説明変数に該当し、右辺の X_i は説明変数に該当する。 α と β はパラメータ、あるいは推定量と呼ばれるものであり、 α は定数項、 β は係数である。これらは正規方程式によって導出される。また、 u_i は誤差項と呼ばれ、データでは捉えられないが被説明変数に対して影響を与える要因に該当する。

計量経済学では、パラメータの導出によって説明変数と被説明変数の間の関係を分析することが、おもな目的とされる。そして、最小二乗法は推定において最も基礎的な技法と位置付けられる。このとき、誤差項について以下の五つの条件を満たすならば、最小二乗法によって導出される推定量は最良線形不偏推定量として位置づけられる。

条件 1：誤差項 u_i の平均値が 0

条件 2：誤差項 u_i の分散が一定 (σ^2)

条件 3：誤差項同士で相関しない

条件 4：誤差項 u_i が説明変数 X_i と相関しない

条件 5：誤差項 u_i が条件 1, 2 を満たす正規分布に従う ($u_i \sim N(0, \sigma^2)$)

最良線形不偏推定量とは、推定値が真の値と等しい不偏性、さらに、推定値のなかで最も分散が小さい効率性、また、サンプルサイズが大きくなると推定値が真の値に収束するという一貫性の三つの性質を満たす推定量のことを示す。そして、以上の 5 つの条件を満たしたときに最小二乗法の推定量が最良線形不偏推定量になることを、計量経済学では一般的に「ガウス・マルコフの定理」という。

このような条件が成立しているならば、最小二乗法による推定は問題がない。しかし、説明変数と被説明変数の間に双方向の因果関係が生じている場合、この二者の間には「内生性」と呼ばれる問題が生じている。これは、係数である β に対して偏り（バイアス）が生じ、推定値はサンプルサイズが大きくなっても真の値に収束しなくなってしまう。その結果、最小二乗法によって導出される推定値は一貫性を満たさなくなってしまう、推定値は最良線形不偏推定量ではなくなる。

このような問題に対応するため、計量経済学ではさまざまな手法が導入されている。そして、その手法の一つが操作変数法である。内生性の解決方法として、説明変数である X_i を一度被説明変数として回帰分析してから、内生変数として二段階で推定する方法があげられる。そして、そのような方法の代表例が操作変数法であり、以下のように表すことができる。

$$\hat{X}_i = \gamma + \pi Z_i + \varepsilon_i \cdots (2)$$

すなわち、 Z_i が操作変数である。そして、操作変数は以下の条件を満たしている必要がある。

条件 1：第一段階の被説明変数である X_i と相関している

条件 2：第二段階の推定式における誤差項 u_i と相関しない

ここで、条件 2 は以下の条件とほぼ同義として扱われることもある。

条件 2'：第二段階の推定式である Y_i に直接相関しない

これらが操作変数法の定義として位置づけられる。そして、操作変数法を用いた経済学の分析で最も有名なのが Angrist (1991) である。Joshua D. Angrist は、ベトナム戦争への従軍や就学年数が労働者の賃金に対してどの程度影響を与えるのかを実証分析しようとした。このとき、操作変数として注目されたのが、従軍についてのくじ引きである。くじ引きは、一般的にはランダムのもので、すなわち、外生的なものとして位置づけられる。Angrist はくじ引きの結果を操作変数として位置づけ、ベトナム戦争への従軍がその後の労働者の賃金にどの程度影響を与え

たのかを検証した²。この分析は、操作変数を用いた分析の中でもとくに有名なものである。そして、操作変数法を用いた分析はその後、非常に多く展開された。

さらに、近年の研究の例をあげると、Gentzkow and Shapiro (2008) も操作変数法を用いた著名な分析である。Matthew Gentzkow と Jesse M. Shapiro は、テレビが子どもの学力を下げるかどうかを問題意識として、テレビが子どもの学力に与える影響を推定した。このときに、かれらは 1950 年代のアメリカを対象として、それぞれの地域において放送の免許があるかどうかを操作変数とした。放送免許の有無は各地域の放送局の設立に対して影響するが、学力と直接的に相関するわけではない。そのため、かれらは放送免許の有無を操作変数としてテレビと子供たちの学力の関係を検証した³。この研究は、特に教育経済学において重要なものの一つとして引用されることが多い。

しかし、経済学において操作変数法が過度に多用されることについて批判的な見解が寄せられることもある。特に有名なのが、James J. Heckman による批判である。かれは操作変数法を完全に否定するわけではないが、経済学的な基礎を欠いた使用については批判的な見解をしている⁴。Heckman によると、経済学者が分析しようとするイベントやプログラムの影響は、データとしては観測できない非常に多くの要因から強く影響を受けている。操作変数法を用いる際には、経済学者が分析しようとする効果とデータでは捉えられない要因についてあらかじめ考慮することが不可欠である。そのため、単に因果関係の検証を目的とした方法としての操作変数法は有効でないことを Heckman (1997) において指摘している。そして、この指摘は近年の研究を考慮すると的確なものであるといえるだろう⁵。

このように、操作変数法については有用性と危険性の両方が同時に存在する。分析の目的に沿って、適切な手法を用いることが科学においては不可欠である。分析が与える影響を過度に重視し、このような点について考慮せずに操作変数法を使用することは避けるべきである⁶。

2 分析の結果、従軍は賃金を 15% ほど引き下げることがこの研究では判明している。

3 分析の結果、テレビは子どもたちの学力に対して統計的有意性を持たないことが確認された。この結果は、テレビは子どもの学力を直接的には引き下げないことが意味している。

4 Heckman は現代において非常に多様な分析を行っているが、かれ自身の問題意識は常に一貫している。人的資本が個人の賃金に与える影響と、それを計量経済学によって可能な限りに精緻に分析することである。そして、かれはいつの時代においてもこれらを追求している。このような研究姿勢は、経済学者のみならずすべての科学者が参考にするべきものであるだろう。

5 筆者も Kato (2019) において操作変数法を用いて分析を行ったことがある。その分析では、個人の動機と賃金の因果関係の検証を目的としたものである。そして、操作変数としては職位と資格を用いた。しかし、これらは当然のことではあるが、賃金を通じることなく個人の動機に影響を及ぼすものである。そのため、この分析は操作変数法としての経済学的な基礎を欠いた分析であるといえる。

6 操作変数法について重要な指摘をしているほかの例として Bound, et al. (1995) があげられる。この論文では、操作変数として経済学の研究において用いられているものが説明変数と弱い相関しかもっていないことを問題点として指摘している。このように、操作変数として適切なものを探すのは非常に困難な作業である。

3章 Arthur Lewbel の操作変数法

このように、適切な操作変数を探すのは非常に困難である。ところで、操作変数法について重要な指摘を行ったのが Arthur Lewbel である。かれは 2012 年の *Journal of Business and Economic Statistics* に投稿した論文において、誤差項と相関しない外生的な操作変数を用いることができなくても、操作変数法を使用できることを示した⁷。その操作変数を(2)式の Z_i を使用して表すと以下ようになる。

$$Z_i = (X_i - \bar{X}) * \varepsilon_i \cdots (3)$$

すなわち、第二段階の推定式である(1)式の説明変数である X_i とその平均値である \bar{X} の差に第一段階の推定式の残差である ε_i の積を操作変数として使用するのである。そして、この操作変数は以下の条件を守ったら使用可能であることも述べられている。

条件 3：操作変数 Z_i は ε_i の二乗項と相関している

条件 4：操作変数 Z_i は ε_i と u_i の積と無相関である

つまり、操作変数 Z_i は第一段階の推定式の誤差項である ε_i と相関しているが、第二段階の推定式である(2)式の誤差項の u_i とは無相関でなければならないというものである。この手法は、操作変数を用いるための条件である条件 1 と条件 2 をかなりの程度緩和したものである。したがって、使用するための検定としては、第二段階の説明変数である X_i が第二段階の誤差項である u_i と相関するかどうかを Breusch-Pagan 検定によって確認すればよい⁸。さらに、かなり強い条件である条件 1 と条件 2 を緩和したうえで β を一致推定量として推定することが可能であることを示したため、論文が公刊された当時は非常に注目を集めた。そして、現在では Stata においても使用可能である⁹。

4章 応用の可能性

この研究を応用したのが Courtemanche, et al. (2021) である。かれらは、BMI を被説明変数としたうえで人々の運動量がどの程度肥満を軽減させるのかを分析対象とした。このときに問題となるのが、観察されない個人の特徴が BMI とかれらの運動量に対して同時に影響を与えて

7ただし、これはどのような状況においても使用できる操作変数というわけではない。後述のように、この操作変数も守らなければならない条件がある。したがって、あくまで外生的な操作変数を見つけることができない場合において、この手法は有効な方法として使用できるというものである。なお、萌芽的な発想は 2000 年の論文 (Lewbel, 2000) においても示されている。

8 Breusch-Pagan 検定は、変量効果モデルとプーリング回帰モデルのどちらが適切であるかを検定するものである。なお、同様の検定として、変量効果モデルと固定効果モデルのどちらかが適切であるかを検定する Housman 検定があげられる。

9 使用コマンドは `ivreg2h` である。なお、Lightstone 社から無料でダウンロードして使用することが可能である。

いる可能 Lewbel の操作変数法を用いて分析を行っている。まず、第一段階の構造形を以下のように設定する

$$BMI = X\beta_1 + M^*\gamma + \alpha_1\mu + v_1$$

基本的には、(1)式を応用したものである。被説明変数として BMI を設定し、さらに、個人のコントロール変数を X とし、 M^* を平均的な運動時間とした。ただし、誤差項については個人について観察されない固定効果を μ とし、その他の誤差項を v_1 としている。そして、 M^* は以下のように書き換えることができる。

$$M^* = X\beta_2 + \alpha_2\mu + v_2$$

v_2 は、 v_1 とは異なるが、平均的な運動時間に対して影響を与える誤差項である。そして、パネルデータを用いた分析においては、これらの関係は以下のように書き換えることができる。

$$BMI = X\beta_1 + M\gamma + \varepsilon_1$$

ただし、誤差項 ε_i はより詳細に記すと以下ようになる。

$$\varepsilon_1 = \alpha_1\mu + v_1 - \gamma e_d$$

パネルデータにおいて生じる、短期の誤差を e_d として表記している。そして、 e_d は運動時間 M に対して正に作用するので、以下のように表記される。

$$M = X\beta_2 + \varepsilon_2$$

そして、上記のように、誤差項 ε_2 はより詳細に記すと以下ようになる。

$$\varepsilon_2 = \alpha_2\mu + v_2 + e_d$$

そして、Lewbel の操作変数法上の仮定は、この場合では以下のように記すことができる。

$$Cov(Z, \varepsilon_2^2) \neq 0$$

すなわち、操作変数である Z は、運動量 M についての誤差項である ε_2 とは相関しているということである。そして、もう一つの条件は以下ようになる。

$$Cov(Z, \varepsilon_1\varepsilon_2) = 0$$

つまり、操作変数である Z は、運動量 BMI についての誤差項である ε_1 とは無相関であるということである。これらの条件を満たす操作変数として、平均気温と降水量があげられる。これらは BMI とその誤差項に対しては直接的には影響しないが、運動量とその誤差項に対しては直

接的に影響する要因としてあげられる。

この方法によってかれらは、BMI に対して運動量がどのように影響を与えるのかを分析した。そして分析の結果、特に男性について、運動が BMI に対して与える影響は予想よりも小さいことが判明した。この結果からかれらは、運動によって消費されたカロリーを補償するための栄養摂取が、特に男性について大きいことを指摘している。この研究は、BMI と運動の間の因果関係を検証したのものとして重要な研究であるといえるだろう。

5章 まとめ

計量経済学は現代経済学において、必要不可欠な分野である。ただし、計量経済学が何を分析目的とすべきか、そもそも、何をもって「計量経済学」と定義するべきかについては、古くから議論が展開されている。「経済学」という分野についても多様な見解が提示されている状況を考慮すると、「計量経済学」についても多様な見解が提示されても不思議ではない。

ただし、経済理論を検証するために統計学を応用する分野として計量経済学を位置づけることは、かなり高い一般性を持つといえるだろう。そして、そのように計量経済学を位置づけるならば、経済モデルにしたがってデータを用いて回帰分析を行うことは、最も基礎的な分析とみなすべきである。よって、「因果分析」の明確な分析が可能であるかどうかは議論が分かれるとしても、回帰分析が計量経済学において必要な分析であることには異論はないだろう¹⁰。

そして、操作変数法は現在も経済学においては重要視されている技法のひとつである。しかし、適切な操作変数を見つけるのは非常に困難な作業である。経済学的な理論と乖離した状況での操作変数の使用は慎むべきであるが、現実的にはそれも非常に難しい。

本稿で紹介した Lewbel による操作変数法は、確かに既存の操作変数法に比べたら条件はかなり緩和されているものではある。しかし、いずれにしても、経済学的な理論にもとづいているべきである点は最も普遍的な原則であるだろう。すなわち、計量経済学の最先端の技法によって実証分析を行うといっても、その根幹として理論分析を常に経済学者は据え置いておくべきである点は変わらない。

参考文献

- Angrist, J. D. (1990) "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence Social Security Administrative Records", *American Economic Review*, Vol. 80, 313-336.
- Bound, J., Jaeger, D. A. and Baker, R. M. (1995) "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak", *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 443-450.

10この点について蓑谷(2007)では、日本と慶應義塾大学における計量経済学の発展とともに問題提起を行っている。

- Courtemanche, C., Pinkston, J. C. and Stewart, J. (2021) "Time Spent Exercising and Obesity: An Application of Lewbel' s Instrumental Variables Method", *Economics and Human Biology*, Vol. 41, 100940.
- Gentzkow, M. and Shapiro, J. M. (2008) "Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study", *Quarterly Journal of Economics*, Vol. 123, No. 1, pp. 279-323.
- Heckman, J. J. (1997) "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations", *Journal of Human Resources*, Vo. 32, No. 3, pp. 441-462.
- Kato, Y. (2019) "Crowding Out of Intrinsic Motivation by Wage: The Case of Long-term Care Workers in Japan", *Econoinformatics*, No. 107, pp. 1-24.
- Lewbel, A. (2000) "Semiparametric Qualitative Response Model Estimation with Unknown Heteroskedasticity or Instrumental Variables", *Journal of Econometrics*, Vol. 97, Vol. 1, pp. 145-177.
- Lewbel, A. (2012) "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models", *Journal of Business and Economics Statistics*, Vol. 30, No. 2, pp. 67-80.
- 箕谷千凰彦 (2007) 「慶應計量経済学派の胎動、確立および発展」, 『三田学会雑誌』, 第 100 卷, 第 1 号, pp. 79 - 140.