

Econoinformatics

経済情報学研究

NO.124 2024

論文

計量経済学についてのノート：セレクションバイアスと社会的企業

目次

- はじめに
- 計量経済学
- サンプルセレクション
- 分析の限界：社会的企業の場合
- 結びに代えて

令和6年2月

The Association of Econoinformatics, Himeji Dokkyo University

姫路獨協大学

経済情報学会

計量経済学についてのノート： セレクションバイアスと社会的企業

加藤 善昌

要旨

本稿は、セレクションバイアスとその対応策について、社会的企業の分析を例としながら考察したものである。計量経済学では、分析で使用するデータが何らかの要因によって偏ってしまうセレクションバイアスが存在する。ただし、それへの対応策として Tobit model と Heckit model が存在しており、非常に有効な手法として用いられている。しかし、これらは分析対象そのものが偏っている場合は、その有効性を十分に発揮することができない。これは、「社会的企業」について、分析対象そのものを定義するときやステークホルダーを対象とするときに深刻な問題となる。

キーワード；セレクションバイアス , Tobit model, Heckit model, 社会的企業

JEL Classifications;

1. はじめに

経済学は社会科学の一種であり、科学の一種である。「科学」をどのように定義するかは多様であるが、仮説の検証行為とその全般と位置付けてもよいだろう。したがって、経済学にも仮説とその検証が存在する。そして、仮説は「理論経済学」と呼ばれる分野で考察され、検証は「実証経済学」と呼ばれる。「経済学」をどのように捉えるかは多様であるが、このような過程は一般的であるとみなしてよいだろう。

そして、理論経済学はさまざまな仮定のもとで演繹的に考察を行う。一方、実証経済学は現実のデータを対象に分析を行う。すなわち、実証経済学は理論経済学と現実のデータを対応させるものである必要がある。そして、それは「計量経済学」という分野によって実現される。

しかし、理論経済学と現実のデータは必ず対応するとは限らない。現実のデータは、さまざまな課題を抱えている。その課題を克服するために、われわれはさまざまな統計的手法を用いる。

そして、現実のデータが抱えている課題のひとつに、分析で使用できるデータが何らかの要因によって偏ってしまう現象がある。その結果、計量経済学による分析結果は望ましい結果を出せなくなってしまう。そして、それは計量経済学では「セレクションバイアス」と呼ばれてい

る。ただし、そのような課題にも対応策がある。それらは、現代の計量経済学では Tobit model と Heckit model と呼ばれている。

本稿は、Tobit model と Heckit model の概要、そして、それらの限界を紹介するものである。ただし、その際に、「社会的企業」と呼ばれる組織の分析を部分的に引用する。これまで筆者が述べてきたように、「社会的企業」の分析では分析者の思想等が他の研究対象よりも強く反映される傾向がある。そのため、セレクションバイアスの問題がより深刻になる可能性がある。本稿では、その点を改めて指摘する。

本稿の構成は以下のようになっている。まず、次の 2 節では計量経済学の基本的な枠組みについて整理する。続く 3 節では、そのうえでセレクションバイアスの概要とその対処法である Tobit model と Heckit model について述べる。そして、4 節では Tobit model と Heckit model の限界について、「社会的企業」を対象とした分析を例としながら述べる。5 節では、まとめを述べる。

2. 計量経済学

経済を統計学的手法によって分析する場合、おもに二つの観点からの分析があげられる¹。一つ目は、経済統計学である。これは、現実の事象としての経済を対象にして統計学的手法によって分析するものである。例えば、消費者を性別や学歴にもとづいて標本に編成し、検定によって消費額や貯蓄額の平均値の差異が、統計的に有意性を持つかどうかを調べることは、その典型例である。

もう一つは、計量経済学である。これは、経済学の理論を統計学的手法によって検証することである。すなわち、計量経済学では原則として、経済理論が先に存在し、その理論が正しいかどうかを何らかのデータによって検証することである²。これは、経済統計学が時として分析結果から理論を推測することがあるのと一線を画す³。

そして、計量経済学で最も使用されるツールが回帰分析である。これは、基本的には以下のモデルが使用される⁴。

$$y_i = \alpha + \beta x_i + u_i$$

1 以下は完全に背反あるいは分離されるものではない。例えば、家森他 (2019) は経済理論にもとづいてアンケート調査を行っており、後述する計量経済学としての面も有している。ただし、統計データの特徴を重視している点から、かれらの分析は経済統計学としての面も強く有しているといえるだろう。もちろん、経済統計学と計量経済学について優劣はない。

2 辻村 (2008) を参照。

3 もちろん、計量経済学でもそれらは必要とされる。しかし、それらは予備的分析に位置づけられることが大半である。

4 以下では単回帰で述べるが、説明変数が複数存在する重回帰でも基本的には同様である。また、パネル分析でも個体特有の効果について考慮する必要はあるものの、基本的には同様である。

i は個体を識別するインデックスである。そして、左辺の y_i は被説明変数であり、従属変数、あるいは、内生変数と呼ばれる。何らかの要因によって、決定づけられる変数である。そして、 x_i は説明変数であり、独立変数、あるいは外生変数と呼ばれる。これは、被説明変数である y_i を決定づける要因のうち、データとして観測できる部分である。そして、 u_i は誤差項と呼ばれる要因であり、被説明変数である y_i を決定づける要因のうち、データとして観測できない部分である。計量経済学では、経済理論にもとづいて被説明変数と説明変数を設定し、最小二乗法、すなわち、データの平均値と観測値の差が最小になるように正規方程式を設定し、 α と β を推定することが行われる。また、 α と β については、一致性と不偏性、効率性の有無が重視され、これらが満たされるときは最良線形不偏推定量とされる⁵。

3. サンプルセレクション

このような枠組みが、計量経済学の基本的な分析ツールである。ただし、計量経済学はいくつかの点で課題を抱えている。その一つが、サンプルセレクションと呼ばれるものである。これは、サンプルの採集において何らかの特徴によってサンプルが偏り、その結果、推定値 α と β は最良線形不偏推定量としての性質を満たさなくなることである。

この場合、通常の最小二乗法は頑健な手法ではなくなる。そこで、以下のモデルによる分析が用いられる⁶。

$$y_i = \alpha + \beta x_i + \gamma \lambda_i + u_i$$

右辺の第三項である λ_i は、逆ミルズ比と呼ばれるものである。これは、サンプルの採集において偏りが発生しているかどうかを示す指標である。この逆ミルズ比 λ_i は以下のように定義される。

$$\lambda_i = \frac{f(x_i)}{F(x_i)}$$

右辺の分母である $F(x_i)$ は分布関数であり、被説明変数 y_i の観察される確率を表している。そして、分子である $f(x_i)$ は正規分布の確率密度関数である。すなわち、逆ミルズ比は被説明変数 y_i の観察されにくさを表している。

サンプルセレクションの問題は、古くから指摘されてきた。しかし、それを計量経済学において本格的に分析されるようになったのは、James Tobin や James J. Heckman、そして、雨宮健による功績が大きい。Tobin は 1958 年の論文において、何らかの要因によって観測値が 0 から実

5 詳細は岩田 (1983) を参照。

6 詳細は山本 (2015) を参照。

現値に極端に偏るモデルをあげ、その対処法を述べた⁷。このモデルは、現在では Tobit model として現在の経済学では一般的に用いられている⁸。そして、雨宮もデータが観測されないことについて理論分析を行った。さらに、雨宮の分析を下地にしながら二段階の推定方法を Heckman は 1976 年と 1978 年の論文で考察した。かれの考案した推定方法も、Heckit model として一般的に用いられている⁹。

このように、サンプルセクションは計量経済学において非常に有用なツールとして用いられている。

4. 分析の限界：社会的企業の場合

このように、Tobit model や Heckit model はセクションバイアスが生じている状態において有効である。ただし、限界がある。以下では、社会的企業の分析例とともにいくつか紹介する。

まず、「社会的企業」としての分析対象そのものの偏りである。例えば、「社会的企業」という組織の場合、利潤と社会問題の解決の両立を目指す組織全般を指す。そのため、その対象には法人形態としての営利企業や非営利組織が含まれることとなる。この場合、例えば、労働時間等についてデータを収集している場合、そのデータについてのセクションバイアスに対しては Heckit model は有効であるだろう。また、データについて何らかの潜在的な閾値が存在し、その閾値にしたがって観測変数が観測されるかどうか決定される場合は、Tobit model が有効であるだろう。しかし、利潤や社会問題の解決についてある程度の潜在的な閾値が存在し、その閾値にしたがって回答する対象のセクションバイアスが生じているとする。その場合、対象自体が偏っているために Heckit model は有効ではない。したがって、また別の分析方法が必要となる。「社会的企業」としての対象が明確でないことが大半である分析では、重要な課題である。

次に、ステークホルダーについてのセクションバイアスである。ただし、これは標本としてのステークホルダーではなく、分析者側の課題である。「社会的企業」について、そのステークホルダーは多岐に渡る。そのなかには、経営者である社会的企業をはじめ、消費者や地域住民、そして、労働者や出資者が含まれる。そのため、分析者が何をしらべるかにしたがって、おもな分析対象となるステークホルダーが決められる。そのため、分析者の目的にしたがってセクションバイアスが生じることとなる。事業が多面的な影響をもたらす「社会的企業」については、不可分の問題である。

7 詳細は Tobin (1958) を参照。

8 Tobit model のその後の展開については Amemiya (1980) を参照。

9 詳細は Heckman (1976, 1978) や Lee (1983) を参照。

5. 結びに代えて

計量経済学は、経済理論を検証するために統計学的手法を援用する分野である。したがって、分析者がどのような視点で分析するかが最も肝要である。経済理論をどのように捉えているかが、計量経済学としての分析に決定的な影響を与える。したがって、経済学の分析においては理論が改めて重要である。

そして、経済理論を検証するための方法として、計量経済学は存在する。しかし、計量経済学は決して万能の分析ツールではない。データには常に何らかの問題がつきまとう。特に、分析対象の偏りであるサンプルセレクションバイアスは分析に対して根幹から影響を与える問題である。

そのようなサンプルセレクションバイアスに対応する Tobit model や Heckit model は、非常に有用な分析ツールである。しかし、分析対象自体が偏っている場合、その有用性は著しく損なわれる。特に、経済理論の設定においてイデオロギーが伴いやすい「社会的企業」は、サンプルセレクションが生じやすいと考えられる。この点に対する留意は、非常に重要である。

さて、研究の発展は、論文の発刊が非常に重要である。近年は、インパクト・ファクターなどの客観的指標による評価が重要視される。もちろん、そのような指標は非常に重要かつ必要である。また、発刊物にも多様な掲載物が重要ではある。しかし、特定の研究機関による発刊物としての特色が強かったとしても、そこから学問が不偏性を持ちながら発展するケースも決して珍しくない¹⁰。したがって、インパクト・ファクターなどに必ずしも反映されることがなかったとしても、研究が後続の研究に影響を与えることがある。そのため、紀要のような刊物も研究のすそ野を広げるといふ点から必要なのである。

参考文献

- Amemiya, T. (1980) "Tobit Models: A Survey", *Journal of Econometrics*, Vol. 24, No. 1 - 2, pp. 3 - 61.
- Heckman, J. J. (1976) "The Common Structure of Statistical Models of Truncation: Sample Selection and Limited Dependent Variables and a Sample Estimator for Such Models", *Annals of Economic and Social Measurement*, Vol. 5, No. 4, pp. 475 - 492.
- Heckman, J. J. (1979) "Sample Selection Bias as a Specification Error", *Econometrica*, Vol. 47, No. 1, pp. 153 - 161.
- Lee, L. (1983) "Generalized Econometric Models with Selectivity", *Econometrica*, Vol. 51, No. 2, pp. 507 - 512.
- Tobin, J. (1958) "Estimation of Relationships for Limited Dependent Variables", *Econometrica*, Vol. 26, No. 1, pp. 24 - 36.

10 蓑谷 (2007) でも述べられているように、三田学会雑誌を起点として日本の計量経済学において重要な分析が非常に多く見られることがその例である。

- 岩田暁一(1983)『経済分析のための統計的手法(第二版)』,東洋経済新報社.
- 辻村江太郎(2008)『計量経済学』,岩波書店.
- 蓑谷千風彦(2007)「慶應計量経済学派の胎動,確立,及び発展」,『三田学会雑誌』,第100巻,第1号,pp.79-140.
- 家森信善,上山仁恵,柳原光弘(2019)「わが国の高等学校における金融・証券教育の現状と課題:高等学校教員に対する調査結果に基づく考察」,『経済経営研究』,第69巻,pp.71-144.
- 山本勲(2015)『実証分析のための計量経済学』,中央経済社.