

## 第6講 操作変数法による政策評価

### 1 はじめに

人間の活動は相互依存的である。ある意思決定は他の意思決定次第であるということは、日常的に経験する事実である。このことはミクロ計量経済学にも当てはまる。すなわち、経済主体毎に集められたマイクロデータの多くは全く偶発的に発生しているものではなく、多かれ少なかれ内生的に決まっているものである。

例えば、消費行動を分析する場合には、所得や資産は所与として扱っているが、実際には、労働時間や職業選択、投資計画などを通して所得は内生的に決まっているものである。また貯蓄は所得と消費の残差として受け身的に決まれていると考えるより、ある程度、長期的な経済見通しに基づいて主体的に決めているので、これも内生的に決まっていることになる。ミクロ計量経済学で推定するモデルの右辺の説明変数が内生である場合に、それを考慮せずに外生変数として扱うと推定パラメータにバイアスがかかる、より厳密に言えば一致推定量が得られないということになる。本章ではこの問題にどのように対処すればいいのかを説明したい。

計量モデルで関心のある説明変数の効果を調べたい時に、それが内生変数であれば、そのバイアスを調整する必要がある。ではその変数が内生かどうかはどのように調べればいいのか。後ほど説明するように、統計的には内生性検定の方法が提示されており、それを適用すればいいのだが、計量モデルが依拠している理論モデルからも内生であるのかどうかはある程度見通しをつけることができる。既に述べたように、ミクロ計量経済学の扱う変数の多くは内生であると考えて推定方法を工夫すべきなのである。

次に、内生変数を説明する操作変数を選ぶことが必要になるが、これがまた難問である。後に見るように、適切な操作変数とは本来、推計したいモデルの誤差項とは無相関であり、しかもそのモデルで用いられている内生変数と強い相関があるような変数でなければならない。計量モデルで扱おうとしている経済システムの外の変数で、かつシステム内の内生変数と関係があるものとは、いわば無い物ねだりのようなものである。実際、マイクロデータ自体が特定の経済目的で集められており、そのシステムとは関係のない変数に

ついて調査項目が入っていることはまれである。それだけに適切な操作変数を探すのに苦労するのである。また無意味な操作変数を用いることは、追加的なバイスをもたらすので注意しなければならない。適切な操作変数の選択に関しては近年その手法が多く開発されており、ここでもできるだけ最新のものまで紹介するつもりである。

## 2 操作変数法の考え方

連立方程式を解いてパラメータを推計するという作業は計量経済学ではよく行われている。また経済理論上、経済変数が同時に決定されるということもよくある。このような場合には変数の内生性を考慮した推計が必要になる。操作変数法はそのような問題に対処するための推計方法である<sup>1</sup>。

操作変数法の最も簡単な説明は次のようなものである。

$$y = \beta_0 + \beta_1 x + u$$

ここで  $x$  が外生変数ではなく、 $u$  と相関しているとすれば、次のような関係が見出される。

$$\text{Cov}(x, u) \neq 0$$

この場合、パラメータ  $\beta_0$  と  $\beta_1$  の推計値は  $x$  の内生性のために一致推定にはならない。そこで、 $x$  とは相関しているが、 $u$  とは無相関な変数  $z$  を導入する。

$$\text{Cov}(z, u) = 0$$

$$\text{Cov}(z, x) \neq 0$$

この変数  $z$  に相当するものが内生変数  $x$  に対する操作変数と呼ばれるものである。操作変数は次の条件を満たしていることが想定されている。(1) 操作変数  $z$  は誤差項  $u$  と無相関である。(2)  $z$  は内生変数  $x$  と相関している。(3)  $z$  と  $x$  は強く相関している。

ここで  $\text{Cov}(z, u) = 0$  を直接テストする方法はないが、 $\text{Cov}(z, x) \neq 0$  をテストするには、次の式を推計してパラメータ  $\pi_1 = 0$  を検定すればよい。すなわち、

$$x = \pi_0 + \pi_1 z + v$$

<sup>1</sup>操作変数法に関する基本文献は Bowden and Turkington (1984) である。近年、後に説明するように弱相関問題などを巡って新しい研究が蓄積されている。より詳しくは北村(2009、第5章)を参照されたい。

ここで  $\pi_1 = Cov(z, x)/Var(z)$  であるので、 $Cov(z, x) \neq 0$  が成立するためには、 $\pi_1 \neq 0$  であることが必要十分条件となるのである。しかし  $\pi_1 \neq 0$  であることは必ずしも操作変数が強相関を持つことは意味していない。この問題については後ほど論じる。

ところで、操作変数を用いた推定式のパラメータは次のように表せる。

$$\hat{\beta}_{IV} = \frac{Cov(z, y)}{Cov(z, x)} = \frac{\sum_{i=1}^n (z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^n (z_i - \bar{z})(x_i - \bar{x})} = \frac{r_{zy} \sqrt{\mathbf{y}'\mathbf{y}}}{r_{zx} \sqrt{\mathbf{x}'\mathbf{x}}}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_{IV}\bar{x}$$

ここで  $z = x$  であれば、推計値は最小二乗法と一致する。上 2 式が満たされるとすると、操作変数法による推計パラメータは一致推定となる。すなわち  $p \lim(\hat{\beta}_{IV}) = \beta_1$  となる。もし、上 2 式が満たされない場合は推計パラメータは一致推計とはならない。とりわけ、 $x$  と  $u$  が相関していれば、推計パラメータはバイアスを持つ。特に標本数が少ない場合にはかなり大きなバイアスを持つことが知られている。

2 段階最小二乗法 (two stage least squared method: 2SLS) は操作変数が複数あって、内生変数 ( $k$ ) と操作変数 ( $l$ ) の数が必ずしも一致しない ( $l \geq k$ ) 場合の推定方法であり、次のように考える。

まず、第 1 段階では内生変数を操作変数によって回帰する。

$$x = \pi_0 + \pi_1 z_{i1} + \pi_2 z_{i2} + \omega_i$$

推定パラメータを用いて内生変数を予測する。行列式で書くと次のように表せる。

$$\hat{\mathbf{X}} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}$$

ここで簡便化のために推定パラメータの行列式を  $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$  と定義すると、第 2 段階での推定パラメータは次のようになる。

$$\begin{aligned} \hat{\beta}_{2SLS} &= (\hat{\mathbf{X}}'\mathbf{X})^{-1}\hat{\mathbf{X}}'\mathbf{y} \\ &= \{\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\}^{-1}\{\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}\} \\ &= (\mathbf{X}'\mathbf{P}_Z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_Z\mathbf{y} \end{aligned}$$

$\hat{\beta}_{2SLS}$  が識別されるためには、 $\mathbf{Z}'\mathbf{Z}$  が  $l \times l$  非特異行列であり、 $\mathbf{Z}'\mathbf{X}$  は行列階数 (ランク)  $k$  の行列であることが必要条件である。ここで  $l = k$  の場合、2SLS は操作変数法 (IV) と一致する。ここでは便宜的に 2 段階で推定を行っ

ているように表現しているが、実際には第1段階と第2段階を同時に推定していることに注意すべきである。すなわち、誤差項は次のように定義されるべきである。

$$\hat{u}_i = y_i - \mathbf{X}\hat{\beta}_{2SLS}$$

ここで  $\mathbf{X}$  の代わりに推定された  $\hat{\mathbf{X}}$  を用いると一致推定量が得られない。上記の推定パラメータ  $\hat{\beta}_{2SLS}$  の導出式を再確認していただきたい。

### 3 操作変数法の政策評価への導入

ルービンの因果モデルの考え方と2値変数の操作変数による推定法<sup>2</sup>

割り当て  $d(=1,0)$  を説明する操作変数  $z(=1,0)$  が存在し、 $z=1$  であれば得られる割り当て変数は  $d_1$ 、 $z=0$  であれば  $d_0$  とする。結果変数についても  $d=1$  であれば  $y_1$ 、 $d=0$  ならば  $y_0$  が得られるとする。

$$d = zd_1 + (1-z)d_0$$

$$y = dy_1 + (1-d)y_0$$

ここで操作変数の仮定として ( $\Pi$  は独立を表す)

$$(y_0, y_1) \Pi z \mid d, \quad d_z \Pi z$$

すなわち、「割り当て変数を条件付けると潜在的結果変数と操作変数は独立、操作変数は直接結果変数に影響を与えるのではなく割り当て変数を通してのみ影響を与える (除外制約 exclusion restriction)

結果変数に直接影響を与えないような変数を操作変数として選べば、除外制約は成立する。

割り当てと操作変数の短調性 (monotonicity) が成立していることが必要。

Imbens and Angrist (1994) は、局所的平均処置効果 (local average treatment effect: LATE) を  $LATE = E(y_1 - y_0 \mid d_1 = 1, d_0 = 0)$  と定義した。

これは「操作変数の値と割り当て変数の値が同じになる ( $z=0$  で  $d=0$ 、 $z=1$  で  $d=1$ )」という部分集団における因果効果を意味している。

$z=1$  の群と  $z=0$  の群での  $y$  の期待値の差は、(1)~(3) から次のように表せる。

<sup>2</sup>以下は、星野 (2009, pp.96-99) からの引用である。

$$\begin{aligned} E(y | z = 1) - E(y | z = 0) &= E(dy_1 + (1 - d)y_0 | z = 1) - E(dy_1 + (1 + d)y_0 | z = 0) \\ &= E(d_1y_1 + (1 + d_1)y_0 | z = 1) - E(d_0y_1 + (1 - d_0)y_0 | z = 0) \end{aligned}$$

ここで  $d_1 - d_0$  は  $1, 0, -1$  であるが、単調性の仮定から  $p(d_1 - d_0 = -1) = 0$  である。

$$\begin{aligned} E(y | z = 1) - E(y | z = 0) &= E((d_1 - d_0)(y_1 - y_0)) \\ &= \sum_{a=-1,0,1} aE(y_1 - y_0 | d_1 - d_0 = a)p(d_1 - d_0 = a) \\ &= E(y_1 - y_0 | d_1 - d_0 = 1)p(d_1 - d_0 = 1) \end{aligned}$$

単調性より、 $d_1 - d_0 = 1$  は  $d_1 = 1, d_0 = 0$  を表しており、 $p(d_1 - 1) = p(d_1 = 1, d_0 = 1) + p(d_1 = 1, d_0 = 0)$ 、 $p(d_0 = 1) = p(d_0 = 1, d_1 = 1)$  である。さらに  $E(d | z = 1) = p(d_1 = 1)$ 、 $E(d | z = 0) = p(d_0 = 1)$  より

$$LATE = \frac{E(y | z = 1) - E(y | z = 0)}{E(d | z = 1) - E(d | z = 0)}$$

と表現することができる。

ここで、 $\bar{y}_z = 1$  を  $E = 1$  に割り当てられた群での結果変数の平均  $d_{z=1}$  を割り当てられた群で  $d = 1$  となる比率

(4) 式は

$$LATE = \frac{\bar{y}_{z=1} - \bar{y}_{z=0}}{\bar{d}_{z=1} - \bar{d}_{z=0}}$$

と書き換えることができる。これは  $z$  が 2 値データの場合の操作変数推定量である。

操作変数推定では「独立変数と相関が高い操作変数を選ぶ」ことが重要であるが、両者とも 2 値変数の場合に相関が高いということはこの式の分母が 1 に近づくということであり、その場合には  $z$  で群分けした場合の平均差そのものが  $LATE$  になる。

## 参考文献

- [1] 星野高宏 (2009) 『調査観察データの統計科学』、岩波書店
- [2] 北村行伸 (2009) 『ミクロ計量経済学入門』、日本評論社
- [3] Anderson, T.W. (1984) *Introduction to Multivariate Statistical Analysis*, Wiley.

- 
- [4] Anderson, T.W. and Rubin, H.(1949) “Estimators of the Parameters of a Single Equation in a Complete Set of Stochastic Equations”, *Annals of Mathematical Statistics*, 21, pp.570-82.
- [5] Andrew, Donald W.K. and Stock, James H.(2005) “Inference with Weak Instruments”, NBER Technical Working Paper 313.
- [6] Angrist, J.D.and Krueger, A.B.(1991) “Does Compulsory School Attendance Affect Schooling and Earnings?”, *Quarterly Journal of Economics*, 106, pp.979-1014.
- [7] Angrist, Joshua, D., Imbens, Guido, W. and Rubin, Donald B.(1996) “Identification of Causal Effects Using Instrumental Variables”, *Journal of the American Statistical Association*, 91(434), pp.444-455.
- [8] Angrist, Joshua D. and Krueger, Alan B.(2001) “Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments”, *Journal of Economic Perspectives*, 15(4), pp.69-85.
- [9] Angrist, Joshua, D., Bettinger, Eric., Bloom, Erik., King, Elizabeth., and Kremer, Michael. (2002) “Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment”, *American Economic Review*, 92(5), pp.1535-1558.
- [10] Basman, R.L. (1960) “On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics”, *Journal of the American Statistical Association*, 55(292), pp.650-59.
- [11] Baum, Christopher. (2006) *An Introduction to Modern Econometrics Using Stata*, Stata Press.
- [12] Blackburn, McKinley and Neumark, David.(1992) “Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials”, *Quarterly Journal of Economics*, 107(4), pp.1421-1436.
- [13] Bound, John., Jaeger, David.A. and Baker, Regina. 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*, 90(430), pp.443-50.
- [14] Bowden, R.J. and Turkington, D.A.(1984) *Instrumental Variables*, Cambridge University Press.

- [15] Breusch, Trevor., Qian, Hailong., Schmidt, Peter., and Wyhowski, Donald.(1999) “Redundancy of Moment Conditions”, *Journal of Econometrics*, 91, pp.89-111.
- [16] Cameron, A.C.and Trivedi, P.K.(1998) *Regression Analysis of Count Data*, Cambridge University Press.
- [17] Cameron, A.C. and Trivedi, P.K.(2005) *Microeconometrics: Methods and Applications*, Cambridge University Press.
- [18] Chao, John.C. and Swanson, Norman R.(2005) “Consistent Estimation with a Large Number of Weak Instruments”, *Econometrica*, 73(5), PP.1673-1692.
- [19] Cragg, John G.and Donald, Stephen G.(1993) “Testing Identifiability and Specification in Instrumental Variable Models”, *Econometric Theory*, 9, pp.222-40.
- [20] Davidson, Russell and MacKinnon, James G.(2004) *Econometric Theory and Methods*, Oxford University Press.
- [21] Durbin, J.(1954) “Errors in variables”, *Review of the International Statistical Institute*, 22, pp.23-32.
- [22] Griliches, Zvi.(1976) “Wages of Very Young Men”, *Journal of Political Economy*, 84(4. Part 2), pp. S69-S85.
- [23] Griliches, Zvi.(1977) “Estimating the Returns to Schooling: Some Econometric Problems”, *Econometrica*, 45(1), pp.1-22.
- [24] Griliches, Zvi., Hall, Bronwyn., and Hausman, Jerry.(1978) “Missing Data and self-Selection in Large Panels”, *Annales de L'INSEE*, XXX-XXXI, pp.137-76.
- [25] Hahn, Jinyoung and Hausman, Jerry. (2002a) “A New Specification Test for the Validity of Instrumental variables”, *Econometrica*, 70(1), pp.163-189.
- [26] Hahn, Jinyoung and Hausman, Jerry. (2002b) “Notes on Bias in Estimators for Simultaneous Equation Models”, *Economics Letters*, 75, pp.237-241.
- [27] Hahn, Jinyoung and Hausman, Jerry. (2003) “Weak Instruments: Diagnosis and Cures in Empirical Econometrics”, *American Economic Review*, 93(2), pp.118-125.

- [28] Hall, Alastair R., Rudebusch, Glenn D. and Wilcox, David W.(1996) “JUdging Instrument Relevance in Instrumental Variables Estimation”, *International Economic Review*, 37(2), pp.283-298.
- [29] Hall, Alastair R. and Peixe, Fernanda P.M.(2000) “A Consistent Method for the Selection of Relevant Instruments”, A paper presented at Econometric Society World Congress 2000.
- [30] Hansen, Lars.P (1982) “Large Sample Properties of Generalized Method of Moments Estimators”, *Econometrica*, 50(4), pp.1029-1054.
- [31] Hausman, Jerry. (1978) “Specification tests in econometrics”, *Econometrica*, 46, pp.1251-72.
- [32] Hausman, Jerry., Stock, James H. and Yogo, Motohiro.(2005) “Asymptotic Properties of the Hahn-Hausman Test for Weak-Istruments”, *Economics Letters*, 89, pp.333-42.
- [33] Hayashi, Fumio.(2000) *Econometrics*, Princeton University Press.
- [34] Imbens, Guido W. and Angrist, Joshua D.(1994) “Identification and Estimation of Local Average Treatment Effects”, *Econometrica*, 62(2). pp.467-475.
- [35] Koenker, Roger.(1981) “A Note on Studentizing a test for Heteroscedasticity”, *Journal of Econometrics*, 17., pp.107-112.
- [36] Koenker, Roger. (2005) *Quantile Regression*, Cambridge University Press.
- [37] Nelson, Charles R. and Startz, Richard.(1990a) “The Distribution of the Instrumental Variables Estimator and Its t-Ratio When the Instrument is a Poor One”, *Journal of Business*, 63(1, Part.2), pp. S125-S140.
- [38] Nelson, Charles R.and Startz, Richard.(1990) “Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator”, *Econometrica*, 58(4), pp.967-76.
- [39] Pagan, A.R. and Hall, D. (1983) “Diagnostic Tests as Residual Analysis”, *Econometric Reviews*, 2(2), pp.159-218.
- [40] Ruud, P.A. (2000) *An Introduction to Classical Econometric Theory*, Oxford University Press.
- [41] Sargan, J.D. (1958) “The Estimation of Economic Relationships Using Instrumental Variables”, *Econometrica*, 26(3), pp.393-415.

- [42] Shea, John.(1997) “Instrument Relevance in Multivariate Linear Models: A Simple Measure”, *Review of Economics and Statistics*, 79(2), pp.348-352.
- [43] Staiger, Douglas. and Stock, James.H. (1997) “Instrumental Variables Regression with Weak Instruments”, *Econometrica*, 65(3), pp.557-86.
- [44] Stock, James H. and Wright Jonathan H. (2000) “GMM with Weak Identification”, *Econometrica*, 68(5), pp.1055-96.
- [45] Stock, James H., Wright, Jonathan H. and Yogo, Motohiro. (2002) “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments”, *Journal of Business and Economic Statistics*, 20(4), pp.518-29.
- [46] Stock, James H. and Yogo, Motohiro. (2005) “Testing for Weak Instruments in Linear IV Regression”, in Andrews, D.W.K. and Stock, J.H.(eds) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press. pp.80-108.
- [47] Winkleman, Rainer and Boes, Stefan. (2005) *Analysis of Microdata*, Springer.
- [48] White, Halbert. (1980) “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica*, 48(4), pp.817-838.
- [49] White, Halbert. (1982) “Instrumental Variables Regression with Independent Observations”, *Econometrica*, 50(2), pp.483-499.
- [50] Wooldridge, Jeffrey. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press
- [51] Wu, D-M. (1973) “Alternative tests of independence between stochastic regressors and disturbances”, *Econometrica*, 41, pp.733-50.