

第10講 選択バイアス問題

ヘックマンの2段階推定法

Heckman(1974)はGronau(1973)のモデルに労働時間を加えて、賃金と労働時間が同時決定されるモデルを考察し、プロビット法と最小二乗法によって推定できる2段階推定法(ヘキット)を提示した。すなわち、

$$y_i = x_i' \beta + \sigma \lambda(x_i' \alpha) + \varepsilon_i \quad y_i > 0$$

ここで $\alpha = \beta/\sigma$ 、 $E(\varepsilon_i) = 0$ 、 λ は逆ミルズ比である。

このとき誤差項の分散は次のように表せる。

$$Var(\varepsilon_i) = \sigma^2 - \sigma^2 x_i' \alpha \lambda(x_i' \alpha) - \sigma^2 \lambda(x_i' \alpha)^2$$

ヘックマンは次の2段階の推定法を提案した。

第1ステップ：労働供給に関する二項選択モデルについてプロビット推定し、パラメータ $\hat{\alpha}$ を得る。

第2ステップ：パラメータ $\hat{\alpha}$ を(1)式に代入し、 $y_i > 0$ のサンプルについて最小二乗推定し、パラメータ $\hat{\beta}$ と σ を得る。これは一致推定量である。

この場合、Whiteのrobust推定を行って不均一分散の問題に対処しておくべきである。ヘックマンの2段階推定法¹は広く使われているが、縄田(1997)が指摘しているように、第一段階のセレクションで用いる説明変数 x_{1i} と第2段階の推定で用いる説明変数 x_{2i} がかなり重複しているような場合には、多重共線性の問題がおこりヘックマンの方法は最尤法に比べて、誤差の分散が大きくなり望ましくない。逆に x_{1i} と x_{2i} が別の変数である場合には、最尤法と比べても誤差分散は小さくなく、推定方法の簡便さを考えると、ヘックマンの方法を用いることが正当化される。²

2つの変数 y_1 と y_2 が2変量正規分布に従う

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right)$$

¹STATAではheckmanというコマンドを用いることでヘックマンの2段階推定法を行うことができる。またheckprobというコマンドでセレクション用のプロビット最尤法推定ができる。

²以下は、星野(2009, pp.146-155)から引用している。

ρ は y_1 と y_2 の相関係数。

y_2 がある閾値 c を超えた場合にのみ y_1 が観測されるとする。

$$p(y_1 | y_2 > c) = \int_c^\infty \frac{p(y_1, y_2)}{(y_2 > c)} dy_2$$

$$E(y_1 | y_2 > c) = \mu_1 + \rho\sigma_1 g(\alpha) > E(y_1)$$

$$V(y_1 | y_2 > c) = \sigma_1^2(1 - \rho^2 h(\alpha)) < V(y_1)$$

$\phi(\cdot)$ 標準正規分布の確率密度関数

$\Phi(\cdot)$ 累積分布関数

$\alpha = (c - \mu_2)/\sigma_2$, $g(\alpha) = \phi(\alpha)/(1 - \Phi(\alpha))$, $h(\alpha) = g(\alpha)(g(\alpha) - \alpha)$.

対象者の選択や観測されるかどうかはランダムではなく「 y_1 がある閾値 c を超えた場合にのみ y_1 が観測されるとする」といった特定の被験者の変数の値のみ観測されるメカニズムを 選択メカニズム と呼び、またこれを無視して単純な推定を行うことで生じるバイアスを 選択バイアス と呼ぶ。

ヘックマンのプロビット選択モデル

結果変数の線形回帰モデル

$$y_{i1} = x_{i1}^t \beta_1 + \varepsilon_{i1} \quad (1)$$

結果変数はある特定の調査対象者でのみ観測される。観測可能性は x_2 に依存する。

$$y_{i2} = x_{i2}^t \beta_2 + \varepsilon_{i2} \quad (2)$$

y_{i2} を調査対象者 i の潜在的な状態； $y_{i2} > 0$ なら y_{i1} が観測される。この式は選択方程式と呼ぶ。

ε_1 と ε_2 については次のような仮定をおく。

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1 \\ \rho\sigma_1 & 1 \end{pmatrix} \right) \quad (3)$$

ε_2 の分散は 1 とする。

このモデルをヘックマンのプロビット選択モデル (Heckit) と呼ぶ。

尤度関数

$$L = \prod_{i=y_{i2} \leq 0} [1 - \Phi(x_{i2}^t \beta_2)] \times \prod_{i=y_{i2} > 0} \left[\Phi \left(\frac{1}{\sqrt{1 - \rho^2}} \left\{ x_{i2}^t \beta_2 + \frac{\rho}{\sigma_1} (y_{i1} - x_{i1}^t \beta_1) \right\} \right) \times \frac{1}{\sigma_1} \phi \left(\frac{y_{i1} - x_{i1}^t \beta_1}{\sigma_1} \right) \right] \quad (4)$$

選択メカニズムを無視した尤度関数は

$$L = \prod_{i=y_{i2} \leq 0} [1 - \Phi(x_{i2}^t \beta_2)] \times \prod_{i=y_{i2} > 0} \Phi[x_{i2}^t \beta_2] \times \frac{1}{\sigma_1} \phi\left(\frac{y_{i1} - x_{i1}^t \beta_1}{\sigma_1}\right) \quad (5)$$

ヘックマンは(2)をプロビット推定で $\tilde{\beta}_2$ を計算し、それを用いて(1)を推定する方法、ヘックマンの二段階推定法(Heckman's Two Stage Estimation Method)を提案した。

$y_2 > 0$ という条件の下での y_1 の期待値

$$\begin{aligned} E(y_{i1} | y_{i2} > 0) &= E(x_{i1}^t \beta_1 | y_{i2} > 0) + E(\varepsilon_{i1} | y_{i2} > 0) \\ &= x_{i1}^t \beta_1 + E(\varepsilon_{i1} | \varepsilon_{i2} > -x_{i2}^t \beta_2) \\ &= x_{i1}^t \beta_1 + (\rho \sigma_1) \frac{\phi(x_{i2}^t \beta_2)}{\Phi(x_{i2}^t \beta_2)} \end{aligned} \quad (6)$$

選択バイアス修正のために付加された項をコントロール関数(control function)と呼ぶ。

第1段階(1)「 y_1 が観測されるかされないか」をダミー変数化し、このダミー変数を x_2 説明するプロビットモデルから β_2 の推定値 $\tilde{\beta}_2$ を得る。

第1段階(2)推定値 $\tilde{\beta}_2$ を用いて擬似的な説明変数 $\tilde{\lambda}_i = \phi(x_{i2}^t \tilde{\beta}_2) / \Phi(x_{i2}^t \tilde{\beta}_2)$ を計算する。

第2段階 y_{i1} を x_{i1} と $\tilde{\lambda}_i$ に回帰し、 β_1 と $\rho \sigma_1$ を推定する。

ヘックマンの二段階推定は一致性はあるが以下の問題点が指摘されている。

- (1) 漸近有効性がない推定量
- (2) $V(\varepsilon_{i1} | y_{i2} > 0)$ は不均一分散であり、最小二乗法の推定値の分散は修正する必要がある
- (3) β_2 のうち定数項に対応する部分以外の値がある程度大きな値になる必要がある
- (4) 共変量に関して制約がある

プロビット選択モデルの因果効果推定への拡張

$y_{i2} > 0$: 介入を与えられた場合の結果変数 y_{i1} が観測されるだけでなく、介入を与えられない場合の結果変数 y_{i0} は観測されない。

$y_{i2} \leq 0$: y_{i0} は観測され、 y_{i1} は観測されないとする。結果変数についての回帰モデルに

$$y_{i1} = x_{i1}^t \beta_1 + \varepsilon_{i1}, \quad y_{i0} = x_{i1}^t \beta_0 + \varepsilon_{i0}$$

$$\begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N \left(0, \begin{pmatrix} \sigma_0^2 & \sigma_{01} & \rho\sigma_0 \\ \sigma_{01} & \sigma_1^2 & \rho\sigma_1 \\ \rho\sigma_0 & \rho\sigma_1 & 1 \end{pmatrix} \right)$$

ここでの関心は

$$E(y_1 - y_0) = \{E(x_1)\}^t (\beta_1 - \beta_0)$$

処置群での因果効果

$$\begin{aligned} TET &= E(y_1 - y_0 \mid z = 1) \\ &= \{E(x_1 \mid z = 1)\}^t (\beta_1 - \beta_0) + (\rho\sigma_1 - \rho\sigma_0) E \left(\frac{\phi(x_{i2}^t \beta_2)}{\Phi(x_{i2}^t \beta_2)} \mid z = 1 \right) \end{aligned}$$

ε_1 と ε_0 との共分散 σ_{01} は推定できないので最尤法推定はできない。因果効果や TET には σ_{01} が入っていないのでヘックマンの二段階推定法によって得た β_1 などの推定値を代入すれば推定できる。

「因果効果推定のためのプロビット選択モデル」は、ルービンの因果モデルをパラメトリックに表現したものであると考える。

参考文献

- [1] 北村行伸 (2005) 『パネルデータ分析』、岩波書店
- [2] 縄田和満 (1997) 「Probit, Logit, Tobit」、養谷千鳳彦・廣松毅監修 『応用計量経済学 II』、多賀出版、第 4 章、pp.237-298.
- [3] 牧厚志 (2001) 『応用計量経済学入門』、日本評論社
- [4] Amemiya, T.(1985) *Advanced Econometrics*, Harvard University Press.
- [5] Cameron, A.C. and Trivedi, P.K.(2005) *Microeconometrics: Methods and Applications*, Cambridge University Press.
- [6] Goldberger, A.S.(1964) *Econometric Theory*, Wiley.
- [7] Gronau, R.(1973) “The Effects of Children on the Housewife’s Value of Time”, *Journal of Political Economy*, 81(2-2), S.168-S.199.
- [8] Gronau, R.(1974) “Wage Comparisons-A Selectivity Bias”, *Journal of Political Economy*, 82(6), pp.1119-1144.

- [9] Heckman, J.J. (1974) “Shadow Prices, Market Wages, and Labor Supply”, *Econometrica*, 42(4), pp.679-694.
- [10] Heckman, J.J.(1979) “Sample Selection as a Specification Error”, *Econometrica*, 47(1). pp.153-161.
- [11] Maddala, G.S.(1983) *Limited-Dependent and Qualitative Variables in Economics*, Cambridge University Press.
- [12] Mroz, T.A.(1987) “The Sensitivity of an Empirical Model of married Women’s Hours of Work to Economic and Statistical Assumptions”, *Econometrica*, 55(4), pp.765-799.
- [13] Tobin, J.(1958) “Estimation of Relationships for Limited Dependent Variables”, *Econometrica*, 26. pp.24-36.
- [14] Winkelmann, Rainer and Boes, Stefan.(2006) *Analysis of Microdata*, Springer.
- [15] Wooldridge, Jeffrey. M.(2003a) *Introductory Econometrics*, Thomson.
- [16] Wooldridge, Jeffrey. M.(2003b) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press
- [17] 星野嵩宏 (2009) 『調査観察データの統計科学』、岩波書店
- [18] 北村行伸 (2009) 『ミクロ計量経済学入門』、日本評論社
- [19] Anderson, T.W. (1984) *Introduction to Multivariate Statistical Analysis*, Wiley.
- [20] 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.
- [21] Andrew, Donald W.K. and Stock, James H.(2005) “Inference with Weak Instruments”, NBER Technical Working Paper 313.
- [22] Angrist, J.D.and Krueger, A.B.(1991) “Does Compulsory School Attendance Affect Schooling and Earnings?”, *Quarterly Journal of Economics*, 106, pp.979-1014.
- [23] 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.

- [24] 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.
- [25] 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.
- [26] 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.
- [27] Baum, Christopher. (2006) *An Introduction to Modern Econometrics Using Stata*, Stata Press.
- [28] Blackburn, McKinley and Neumark, David.(1992) “Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials”, *Quarterly Journal of Economics*, 107(4), pp.1421-1436.
- [29] 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.
- [30] Bowden, R.J. and Turkington, D.A.(1984) *Instrumental Variables*, Cambridge University Press.
- [31] Breusch, Trevor., Qian, Hailong., Schmidt, Peter., and Wyhowski, Donald.(1999) “Redundancy of Moment Conditions”, *Journal of Econometrics*, 91, pp.89-111.
- [32] Cameron, A.C.and Trivedi, P.K.(1998) *Regression Analysis of Count Data*, Cambridge University Press.
- [33] Cameron, A.C. and Trivedi, P.K.(2005) *Microeconometrics: Methods and Applications*, Cambridge University Press.
- [34] Chao, John.C. and Swanson, Norman R.(2005) “Consistent Estimation with a Large Number of Weak Instruments”, *Econometrica*, 73(5), PP.1673-1692.

- [35] Cragg, John G. and Donald, Stephen G.(1993) “Testing Identifiability and Specification in Instrumental Variable Models”, *Econometric Theory*, 9, pp.222-40.
- [36] Davidson, Russell and MacKinnon, James G.(2004) *Econometric Theory and Methods*, Oxford University Press.
- [37] Durbin, J.(1954) “Errors in variables”, *Review of the International Statistical Institute*, 22, pp.23-32.
- [38] Griliches, Zvi.(1976) “Wages of Very Young Men”, *Journal of Political Economy*, 84(4. Part 2), pp. S69-S85.
- [39] Griliches, Zvi.(1977) “Estimating the Returns to Schooling: Some Econometric Problems”, *Econometrica*, 45(1), pp.1-22.
- [40] 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.
- [41] Hahn, Jinyoung and Hausman, Jerry. (2002a) “A New Specification Test for the Validity of Instrumental variables”, *Econometrica*, 70(1), pp.163-189.
- [42] Hahn, Jinyoung and Hausman, Jerry. (2002b) “Notes on Bias in Estimators for Simultaneous Equation Models”, *Economics Letters*, 75, pp.237-241.
- [43] Hahn, Jinyoung and Hausman, Jerry. (2003) “Weak Instruments: Diagnosis and Cures in Empirical Econometrics”, *American Economic Review*, 93(2), pp.118-125.
- [44] 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.
- [45] 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.
- [46] Hansen, Lars.P (1982) “Large Sample Properties of Generalized Method of Moments Estimators”, *Econometrica*, 50(4), pp.1029-1054.
- [47] Hausman, Jerry. (1978) “Specification tests in econometrics”, *Econometrica*, 46, pp.1251-72.

- [48] Hausman, Jerry., Stock, James H. and Yogo, Motohiro.(2005) “Asymptotic Properties of the Hahn-Hausman Test for Weak-Instruments”, *Economics Letters*, 89, pp.333-42.
- [49] Hayashi, Fumio.(2000) *Econometrics*, Princeton University Press.
- [50] Imbens, Guido W. and Angrist, Joshua D.(1994) “Identification and Estimation of Local Average Treatment Effects”, *Econometrica*, 62(2), pp.467-475.
- [51] Koenker, Roger.(1981) “A Note on Studentizing a test for Heteroscedasticity”, *Journal of Econometrics*, 17., pp.107-112.
- [52] Koenker, Roger. (2005) *Quantile Regression*, Cambridge University Press.
- [53] 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.
- [54] 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.
- [55] Pagan, A.R. and Hall, D. (1983) “Diagnostic Tests as Residual Analysis”, *Econometric Reviews*, 2(2), pp.159-218.
- [56] Ruud, P.A. (2000) *An Introduction to Classical Econometric Theory*, Oxford University Press.
- [57] Sargan, J.D. (1958) “The Estimation of Economic Relationships Using Instrumental Variables”, *Econometrica*, 26(3), pp.393-415.
- [58] Shea, John.(1997) “Instrument Relevance in Multivariate Linear Models: A Simple Measure”, *Review of Economics and Statistics*, 79(2), pp.348-352.
- [59] Staiger, Douglas. and Stock, James.H. (1997) “Instrumental Variables Regression with Weak Instruments”, *Econometrica*, 65(3), pp.557-86.
- [60] Stock, James H. and Wright Jonathan H. (2000) “GMM with Weak Identification”, *Econometrica*, 68(5), pp.1055-96.
- [61] 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.

- [62] 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.
- [63] Winkleman, Rainer and Boes, Stefan. (2005) *Analysis of Microdata*, Springer.
- [64] White, Halbert. (1980) “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica*, 48(4), pp.817-838.
- [65] White, Halbert. (1982) “Instrumental Variables Regression with Independent Observations”, *Econometrica*, 50(2), pp.483-499.
- [66] Wooldridge, Jeffrey. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press
- [67] Wu, D-M. (1973) “Alternative tests of independence between stochastic regressors and disturbances”, *Econometrica*, 41, pp.733-50.