第7講　差分の差による政策評価

パネル調査において、介入後の「処置群と対照群の差」と介入前の「処置群と対照群の差」の差を指す。[1]

- 「処置群での介入後と介入前の差」と「対象群での介入後と介入前の差」の差に等しい。
  - 介入プログラムの効果+誤差。パネルデータ分析における**固定効果推定**と同じ。

y_a: 政策が実施される前の結果変数

y_{1b}: 「もし政策の対象となる集団に所属した場合における」政策実施後の結果変数

y_{1b}: 「もし政策の対象となる集団に所属しない場合における」政策実施後の結果

t 時点でのグループ所属インディケーター z_t

$y_t = zy_{1b} + (1-z)y_{0b}$

δ を b 時点での測定値であれば $\delta = 1$、a 時点の測定値であれば、$\delta = 0$ とするインディケータとすると、

$y = \delta y_b + (1-\delta)y_a = \delta\left\{zy_{1b} + (1-z)y_{0b}\right\} + (1-\delta)y_a$

$DID = \{E(y \mid z=1, \quad \delta=1) - E(y \mid z=1, \quad \delta=0)\} - \{E(y \mid z=0, \quad \delta=1) - E(y \mid z=0, \quad \delta=0)\}$

$= E(y_{1b} - y_a \mid z=1) - E(y_{0b} - y_a \mid z=0)$

処置群での因果効果（TET）

$TET = E(y_{1b} - y_{0b} \mid z=1)$

$E(y_{0b} \mid z=1)$ は観察できない。

「もし政策の対象とならなかったときの経時変化が 2 つのグループ間で等しい」

$\Rightarrow E(y_{0b} - y_a \mid z=1) = E(y_{0b} - y_a \mid z=0)$

ならば

$DID = E(y_{1b} - y_a \mid z=1) - E(y_{0b} - y_a \mid z=1) + \{E(y_{0b} - y_a \mid z=1) - E(y_{0b} - y_a \mid z=0)\}$

---

[1] 以下は、星野（2009, pp.103-111) からの引用である。

$= E(y_{1b} - y_a \mid z = 1) - E(y_{0b} - y_a \mid z = 1)$

$= E(y_{1b} - y_a \mid z = 1)$

$= TET$

何も介入しなければ結果変数が時間的な変化を起こさない場合（成長なし）

$y_{0b} = y_a$ ならば、この条件は満たされる。

対照群の因果効果 TEU が DID と等しくなるためには「もし政策を実行したときの経時変化が 2 つのグループ間で等しい」ことが条件になる。

$$DID = \frac{1}{N_1}\sum_{i:z_i=1}^{N_i}(y_{bi} - y_{ai}) - \frac{1}{N_2}\sum_{i:z_i=0}^{N_2}(y_{bi} - y_{ai})$$

セミパラメトリックな"差分の差"推定

Abadie(2005)

共変量 x を考える

$TET = E(y_{1b} - y_{0b} \mid z = 1) = Ex(E(y_{1b} - y_{0b} \mid z = 1, x)$

$DID = Ex(E(y_{1b} - y_{0a} \mid z = 1, x) - E(y_{0b} - y_a \mid z = 0, x)$

$= Ex(E(y_{1b} - y_a \mid z = 1, x) - E(y_{0b} - y_a \mid z = 1, x) + \{(E(y_{0b} - y_a \mid z = 1, x) - E(y_{0b} - y_a \mid z = 0, x)\}$

$TET = DID$ である条件は

$E(y_{0b} - y_a \mid z = 1, x) = E(y_{0b} - y_a \mid z = 0, x)$

さまざまな共変量を共通にしたとき、「政策の対照にならなかったときの経時変化が 2 つのグループ間で等しい」という条件になる。

Abadie(2005) は、回帰関数を推定する代わりに傾向スコア $e = E(z - 1 \mid x)$ を用いて TET を表現する。

$E(y_{rb} - y_{0b} \mid z = 1) = \int E(y_{1b} - y_{0b} \mid z = 1, x)p(x \mid z = 1)dx = Ex\left[E(\rho(y_b - y_a) \mid x)\frac{p(z=1|x)}{p(z=1)}\right]$

$= E\left(\frac{y_b - y_a}{p(z=1)}\frac{z-e}{1-e}\right)$

ただし $\rho = \frac{z-e}{e(1-e)}$ とすると、$\rho = \frac{1}{e}$ $(z = 1$ のとき$)$、$\rho = \frac{-1}{1-e}$ $(z = 0$ のとき$)$。

$E(\rho(y_b - y_a) \mid x) = E(\rho(y_b - y_a) \mid z = 1, x) \times e + E(\rho(y_b - y_a) \mid z = 1, x) - E(y_b - y_a) \mid z = 0, x)$

TET を傾向スコアで重み付けした推定量

$$\frac{1}{N}\sum_{i=1}^{N}\frac{y_{bi} - y_{ai}}{p(z=1)}\frac{z_i - e_i}{1 - e_i} \quad \text{or} \quad \frac{\sum_{i=1}^{N}(y_{bi} - y_{ai})\frac{z_i - e_i}{1 - e_i}}{\sum_{i=1}^{N} e_i}$$

問題点

この推定量は、差分の差という単純な形にはなっていない。

$z = 0$ の場合の共変量の情報を有効に活用していない。

クロスセクションデータを利用した差分の差推定
b 時点の調整であれば $\delta = 1$
a 時点の調整であれば $\delta = 0$
処置群の調整であれば $z_a \begin{cases} = 1 \\ = 0 \end{cases}$
対照群の調整であれば $z_b \begin{cases} = 1 \\ = 0 \end{cases}$
$y_b = z_b y_{1b} + (1 - z_b) y_{0b}$
a 時点では介入がないので、$z_a = 0, 1$ どちらでも y_a が得られる。

$y = \delta \left\{ z_{by1b} + (1 - z_b) y_{0b} \right\} + (1 - \delta) y_a$
$DID = E(y \mid \delta = 1, z_b = 1) - E(y \mid \delta = 0, z_a = 1) - \{E(y^- \mid \delta = 1, z_b = 0) - E(y \mid \delta = 0, z_a = 0)\}$
$= E(y_{1b} \mid \delta = 1, z_b = 1) - E(y_a \mid \delta = 0, z_a = 1) - \{E(y_{ab} \mid \delta = 1, z_b = 0) - E(y_a \mid \delta = 0, z_a = 0)\}$
推定量

$$\frac{\sum_{i=1}^{N} \delta_i z_{bi} y_i}{\sum_{i=1}^{N} \delta_i z_{bi}} - \frac{\sum_{i=1}^{N} (1-\delta_i) z_{ai} y_i}{\sum_{i=1}^{N} (1-\delta_i) z_{ai}} - \left\{ \frac{\sum_{i=1}^{N} (1-z_{bi}) y_i}{\sum_{i=1}^{N} \delta_i (1-z_{bi})} - \frac{\sum_{i=1}^{N} (1-\delta_i)(1-z_{ai}) y_i}{\sum_{i=1}^{N} (1-\delta_i)(1-z_{ai})} \right\}$$

以下の３条件が成立すれば、$DID = TET$ となる。
　２時点間で調査対象者は等質である
$E(y_{1b} \mid z_b, \delta = 1) = E(y_{1b} \mid z_b, \delta = 0)$
$E(y_{0b} \mid z_b, \delta = 1) = E(y_{0b} \mid z_b, \delta = 0) = E(y_{0b} \mid z_b)$
$E(y_a \mid z_a, \delta = 1) = E(y_a \mid z_a, \delta = 0)$

　介入しなかった場合の δ 課変数の変化が b 時点における処理群と対照群が等しい。
$E(y_{0b} - y_a \mid z_b = 1) = E(y_{0b} - y_a \mid z_b = 0)$

　a 時点での γ 課変数の平均は２つのデータの処理群で共通であり、対照群でも共通。
$E(y_a \mid z_b) = E(y_a \mid z_a)$

$DID = E_x \left[E(y_{1b} \mid \delta = 1, z_b = 1, x) - E(y_a \mid \delta = 0, z_a = 1, x) - \{E(y_{0b} \mid \delta = 1, z_b = 0, x) - E(y_a \mid \delta = 0, \right.$
が $TET = E(y_{1b} - y_{0b} \mid z_b = 1)$ と一致する。

Triple differences (TD): Difference in Differences in Differences[2]

DD assumes that all people in region 1 at tiem b were treated. TD is relevant if only some qualified people (g=1, e.g. age) in regional at time b are treated.

---

[2] 以下は、Lee(2005, pp.111-115）より引用。

$$g_i = \begin{array}{ll} 1 & if\ treatment\ qualified\ in\ group 1 \\ 0 & if\ in\ group 0 \end{array}$$

$$r_i = \left.\begin{array}{ll} 1 & if\ living\ in\ region 1 \\ 0 & if\ living\ in\ region 0 \end{array}\right\} d_{it} = g_i r_i t_t\ for\ treatment$$

$$t_t = \begin{array}{ll} 1 & if\ t = b \\ 0 & if\ t = a \end{array}$$

y_{jit} is the potential outcome for individual i at time t, $j = 0, 1$.

$$\left.\begin{array}{l} E(y_{1b} - y_{0a} \mid g = 1, r = 1) \\ E(y_{0b} - y_{0a} \mid g = 1, r = 0) \end{array}\right\} first\ DD$$

$$\left.\begin{array}{l} E(y_{0b} - y_{0a} \mid g = 0, r = 1) \\ E(y_{0b} - y_{0a} \mid g = 0, r = 0) \end{array}\right\} \text{second } DD$$

DDD or TD

$= E(y_{1b} - y_{0a} \mid g = 1, r = 1) - E(y_{0b} - y_{0a} \mid g = 1, r = 0) - \{E(y_{0b} - y_{0a} \mid g = 0, r = 1) - E(y_{0b} - y_{0a} \mid g =$

$E(y_{1b} - y_{0a} \mid g = 1, r = 1) - E(y_{0b} - y_{0a} \mid g = 1, r = 1) + \{E(y_{0b} - y_{0a} \mid g = 1, r = 1) - E(y_{0b} - y_{0a} \mid g = 1,$
$- \{E(y_{0b} - y_{0a} \mid g = 0, r = 1) - E(y_{0b} - y_{0a} \mid g = 0, r = 0)\}$

If we assume the same time effect in $\{\cdot\}, \{\cdot\}$
then
$TD = E(y_{1b} - y_{0b} \mid g = 1, r = 1)$
The same time-effect (DD identification condition) is
$E(y_{0b} - y_{0a} \mid g = 1, r = 1) - E(y_{0b} - y_{0a} \mid g = 1, r = 0) = 0.$

Linear Models for TD
$y_{ijt} = \beta_1 + \beta_g g_i + \beta_r r_i + \beta_t t_t + \beta_{rt} r_i t_t + \beta_{aj} + u_{jit}$
$j = 0, 1$, $i = 1, ..., N$, $t = a, b$ $E(u_{jit)=0}$
$E(u_{jit} \mid g, r) = 0\ \forall g \& r.$
$E(y_{1b} - y_{0a} \mid g = 1, r = 1) = \beta_1 + \beta_g + \beta_r + \beta_t + \beta_{rt} + \beta_d - (\beta_1 + \beta_g + \beta_r) = \beta_t + \beta_{rt} + \beta_d$
(Time, region*time, treatment)

$E(y_{0b} - y_{0a} \mid g = 1, r = 0) = \beta_1 + \beta_g - (\beta_1 + \beta_g) = \beta_t$ (time)
$E(y_{0b} - y_{0a} \mid g = 0, r = 0) = \beta_1 + \beta_g + \beta_t + \beta_{rt} - (\beta_1 + \beta_r) = \beta_r + \beta_{rt}$ (time, region*time)
$E(y_{0b} - y_{0a} \mid g = 0, r = 0) = \beta_1 + \beta_t - \beta_1 = \beta_t$ (time)
First time difference is $\beta_{rt} + \beta_d$
second time difference is β_{tt}

$TD = \beta_d$

General regression model

$y_{jit} = \beta_1 + \beta_\varphi g_i + \beta_r r_i + \beta_t t_t + \beta_{gr} g_i r_i + \beta_{gt} g_i t_t + \beta_{rt} r_i t_t + \beta_{dj} + u_{jit}$
Four differences in TD are

$E(y_{1b} - y_{0a} \mid g = 1, r = 1) = \beta_1 + \beta_g + \beta_r + \beta_t + \beta_{gr} + \beta_{gt} + \beta_{rt} + \beta_d + E(u_{1b} \mid g = 1, r = 1)$;
$E(y_{0b} - y_{0a} \mid g = 1, r = 0) = \beta_1 + \beta_g + \beta_t + \beta_{gt} + E(u_{0b} \mid g = 1, r = 0) - \{\beta_1 + \beta_g + E(u_{0a} \mid g = 1, r = 0)\}$
$= \beta_t + \beta_{gt} + E(u_{0b} - u_{0a} \mid g = 1, r = 0)$

$E(y_{0b} - y_{0a} \mid g = 0, r = 1) = \beta_1 + \beta_r + \beta_t + \beta_{rt} + E(u_{0b} \mid g = 0, r = 1) - \{\beta_1 + \beta_r + E(u_{0a} \mid g = 0, r = 1)\}$
$= \beta_t + \beta_{rt} + E(u_{0b} - u_{0a} \mid g = 0, r = 1)$;

$E(y_{0b} - y_{0a} \mid g = 0, r = 0) = \beta_1 + \beta_t + E(u_{0b} \mid g = 0, r = 0) - \{\beta_1 + E(u_{0a} \mid g = 0, r = 0)\}$
$= \beta_t + E(u_{0b} - u_{0a} \mid g = 0, r = 0)$

DD in the first two differences is

$DD_f = \beta_{rt} + \beta_d + E(u_{1b} - u_{0a} \mid g = 1, r = 1) - E(u_{0b} - u_{0a} \mid g = 1, r = 0)$
DD in the second two differences is

$DD_s = \beta_{rt} + E(u_{0b} - u_{0a} \mid g = 0, r = 1) - E(u_{0b} - u_{0a} \mid g = 0, r = 0)$
Then
$TD = DD_f - DD_s$
$= \beta_d + E(u \mid q - u_{0a} \mid g = 1, r = 1) - E(u_{0b} - u_{0a} \mid g = 1, r = 0) - \{E(u_{0b} - u_{0a} \mid g = 0, r = 1) - E(u_{0b} - u_{0a} \mid g = 0, r = 0)\}$

If these error therms are zero, then $TD = \beta_d$

参考文献

[1] 星野嵩宏 (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 Paramters 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 Weal 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, Guid, 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] Basmann, 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", *Quaterly 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 Varaible 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 Internatinal 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 Hausoman, 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”, *Economerica*, 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 Instrumetns”, *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] Winklemann, 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.