MICROECONOMETRICS
CLASS 8

Wiktor Budziński

TREATMENT EFFECTS

We are often interested in causal effects between two variables

- E.g., effect of higher education on wages
- Regular regressions usually inform us only about correlations between two variables

There are some methods designed to estimate so called treatment effects, which should assure estimation of more causal relationships

TREATMENT EFFECTS

Formally we are interested in how treatment, D_i , will affect outcome Y_i

- This is just dependent/independent variable relationship

If $D_i = 1$, then we will observe Y_i^1, otherwise we will observe Y_i^0

$$Y_i = D_i Y_i^1 + \left(1 - D_i\right) Y_i^0$$

We are interested in $Y_i^1 - Y_i^0$

- Obvious issue: we only observe one state for a given individual. We miss the so-called *counterfactual*

TREATMENT EFFECTS

We usually limit analysis to average treatment effect: $\mathbf{E}\left(Y^1 - Y^0\right)$

If we can conduct a randomized experiment then $Y^d \perp D$, then we can simply calculate:

$$\mathbf{E}\left(Y^1 - Y^0\right) = \mathbf{E}\left(Y \mid D = 1\right) - \mathbf{E}\left(Y \mid D = 0\right)$$

- Often done in medicine, usually not possible in economics

TREATMENT EFFECTS

There are usually two auto-selection issues:

- Auto-selection on observable variables: $Y^d \perp D \mid X$
  - Can be solved with some matching techniques as well as regression models
- Auto-selection on unobservable variables $Y^d \perp D \mid \varepsilon$
  - This is what we would call an endogenous treatment
  - Can be solved with models dealing with endogeneity, regression discontinuity or difference-in-difference model

ENDOGENOUS TREATMENT

If the treatment is endogenous (auto-selection on unobservables) we could use regression techniques for endogenous variables

- For example, 2SLS

Treatment is usually a binary variable, and therefore it is often important to account for it in the model

- 2SLS assumes linear regression model for the first stage
- Estimated with MLE

Similar procedure as in the selection model: we have two equations with correlated error terms:

- One of them is a probit model for a treatment variable

$$y_1 = \mathbf{X}_1\boldsymbol{\beta} + D\gamma + \varepsilon$$

$$D^* = \mathbf{X}_2\boldsymbol{\alpha} + \omega$$

ENDOGENOUS TREATMENT

This model can be relaxed to so called switching regression model

- What if treatment affects how other covariate influence the outcome

$$D_i^* = \mathbf{X}_2\boldsymbol{\alpha} + \omega_i,\ D_i = 1 \text{ if } D_i^* > 0$$

$$Y_i = D_i\left(\boldsymbol{\beta}_1\mathbf{X}_i + \varepsilon_{i1}\right) + \left(1 - D_i\right)\left(\boldsymbol{\beta}_0\mathbf{X}_i + \varepsilon_{i0}\right)$$

$$\begin{bmatrix} \omega_i \\ \varepsilon_{i0} \\ \varepsilon_{i1} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_0\theta_0 & \rho_1\theta_1 \\ \rho_0\theta_0 & \theta_0^2 & 0 \\ \rho_1\theta_1 & 0 & \theta_1^2 \end{bmatrix}\right)$$

EXERCISE 1: TREATMENT EFFECTS

1. Read *me.medexp3.rds* into R. We want to investigate effect of insurance on medical expenditures
2. Estimate model with endogenous treatment and compare results with OLS and 2SLS
3. Estimate switching regression model

ENDOGENOUS TREATMENT

Analogous models can be formulated for non-continuous variables

- Binary, Ordered, Count, Censored, etc.

These would be usually estimated with MLE

- Similarly as in the sample selection, such models are usually difficult to find in most software

For example, for binary model we have

$$\begin{cases} y^* = \mathbf{X}_1\boldsymbol{\beta} + D\gamma + \varepsilon \\ D^* = \mathbf{X}_2\boldsymbol{\alpha} + \omega \end{cases}$$

Where we observe only $\mathrm{y} = \mathbf{1}_{\{\mathrm{y}*>0\}}$, $\mathrm{D} = \mathbf{1}_{\{\mathrm{D}*>0\}}$, and $(\varepsilon, \omega) \sim BN(0, \rho)$

EXERCISE 2: TREATMENT EFFECTS

1. Investigate how the fact of having a private insurance affects the probability of something using medical care
   - Account for the fact that private insurance may be endogenous

REGRESSION DISCONTINUITY DESIGN

Regression discontinuity is a specific design of dataset that can be utilized to estimate local average treatment effect

- It does not rely on any heavy assumptions, and should be quite robust

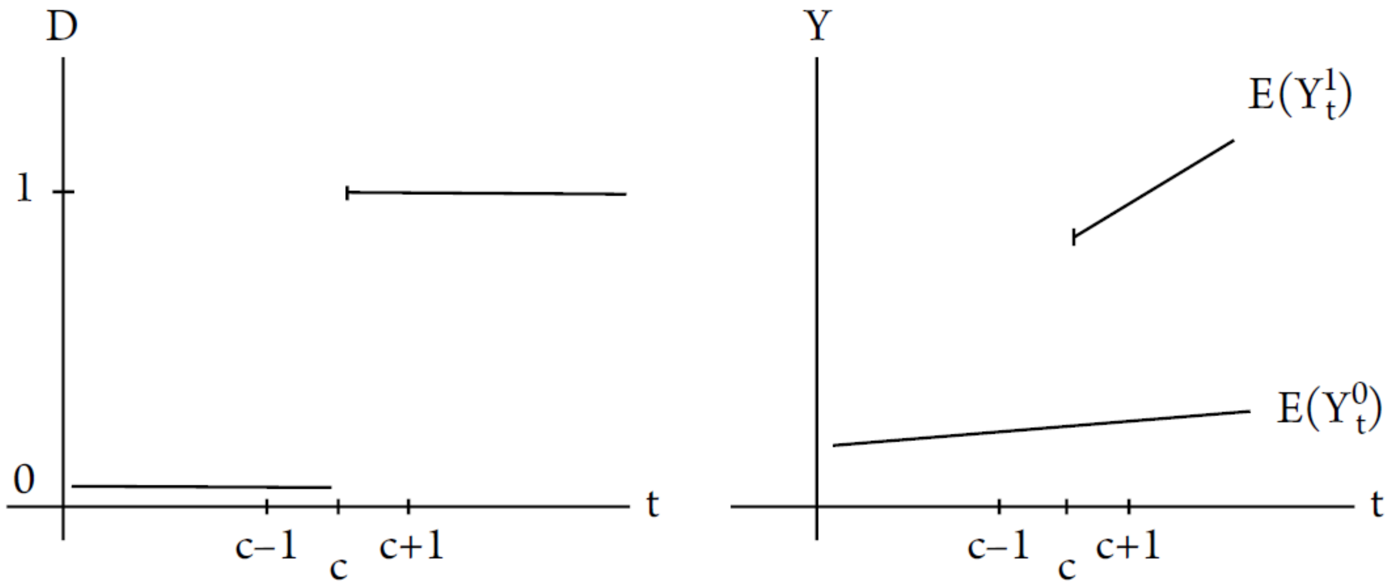
It requires that there exists a running variable S_i, such that $D_i = 1$ if $S_i > c$

- Treatment depends deterministically on the running variable
- For example, some cutoff points as a part of some policy
- Often called „*Sharp RDD*”

REGRESSION DISCONTINUITY DESIGN

Intuitively, individuals close to the cut-off points should be quite similar in both: observable and unobservable covariates

- Cut-off point is arbitrary



REGRESSION DISCONTINUITY DESIGN

We estimate: $E(Y_i \mid S_i) = \beta D_i + m(S_i)$

- where $m(S_i)$ is some polynomial
- We use only observations around the cutoff point c

Disadvantages:

- How close should we be to the cut-off point (*bias-variance* trade-off)
- We only estimate local average treatment effect
  - This should not be generalized for the whole population

EXERCISE 3: RDD

1. Read *me.senate*.rds
2. Investigate how winning in a senate elections changes the chances of winning in a subsequent elections
3. Analyze example 2 in *Sim_examples8.r*

FUZZY RDD

Sometimes assignment rule isn't as clear as in *Sharp RDD*

For example, it only changes the probability of treatment $P(D_i = 1) = F(S_i)$

- It is called Fuzzy regression discontinuity
- Probability is discontinuous at c

In such a case treatment effect can be calculated as: $\tau = \dfrac{m^+(c) - m^-(c)}{\lim\limits_{S_i \to c^+} F(S_i) - \lim\limits_{S_i \to c^-} F(S_i)}$

Where $m(S)$ is some nonlinear function, which may change the slope at c

This is actually equivalent to estimating 2SLS / endogenous treatment model

- And using $\mathbf{1}_{\{S_i > c\}}$ as an instrument for D
- Fuzzy RDD is often used for continuous "*treatments*"

$$y_i = \beta_0 + \beta_1 D_i + \beta_2 (S_i - c) + \beta_3 (S_i - c)\mathbf{1}_{\{S_i > c\}} + \varepsilon_i$$

EXERCISE 3: FUZZY RDD

1. Analyze example 3 in *Sim_examples8.r*
2. Read *Class_size.xls* and use fuzzy regression discontinuity to find the effect of class size on students' performance