Lecture 5b: Evaluating Treatment Effects Using the Regression Discontinuity Method

Reference: Hahn, Todd & Van der Klaauw EMA 2001.

As in the previous two lectures, we are interested in evaluating the effect of some binary treatment variable, here x_i, on outcome y_i.

$x_i = 1$ if *i* is treated and 0 if *i* is untreated. y_{1i} will represent the outcome for agent *i* if treated, and y_{0i} the outcome for *i* if untreated.

As before, our difficulty in identifying the true treatment effect is the missing counterfactual observation: y_{0i} if *i* is treated, y_{1i} if *i* is untreated.

<u>Applications of RD</u>

$\rightarrow$Van der Klaauw (2002) uses an RD approach to estimate the effect of financial aid offers (note: not a binary treatment, aid offers vary) on students' decisions to accept admission to a given college (note: a binary outcome measure). He exploits multiple discontinuities in an administrative formula that determines aid based on SAT score, GPA, & other components.

$\rightarrow$Angrist and Lavy (1999) estimate the effect of class size on student test scores, with identification coming from a rule requiring that one classroom be added in a school whenever average class size exceeds a predetermined threshold. Here class size is a discontinuous (and note: nonmonotonic) function of enrollment in the student's school.

$\rightarrow$Black (1999) uses an RD approach to estimates parents' willingness to pay for school quality by comparing housing prices near school district boundaries. Clearly in this case school quality is discontinuous in geographic space at district borders, and housing prices are the outcome measure of interest.

We will consider two regression discontinuity designs, the sharp design and the fuzzy design.

What they have in common is that (the probability of) treatment depends on some underlying & observable variable z, which will be assumed continuous. Note that *z need not be independent of* y_i. The instrument for treatment here will be the discontinuity in the dependence of x_i on z, not z itself.

Sharp Design

With a sharp regression discontinuity design, x_i depends deterministically on z:

$$x_i = f(z_i)$$

The point z_0 where $f(z)$ is discontinuous is assumed known.

Fuzzy Design

With a fuzzy design, x_i remains random given z_i BUT

$$f(z) \equiv E[x_i \mid z_i = z] = \Pr[x_i = 1 \mid z_i = z]$$

is discontinuous at known value z_0.

The sharp and fuzzy designs differ in that in the sharp design the treatment assignment is deterministic given z, while in the fuzzy design the treatment assignment may depend on additional factors unobserved by the econometrician.

Their common feature is that treatment probability $\Pr[x_i = 1 | z_i = z]$ is a function of underlying continuous & observable variable z, and that function is discontinuous at z_0, i.e.

Assumption (RD): (i) Limits $x^+ = \lim_{z \to z_0^+} E[x_i | z_i = z]$ and $x^- = \lim_{z \to z_0^-} E[x_i | z_i = z]$ exist. (ii) $x^+ \neq x^-$.

Examples: An identifying assumption in Van der Klaauw (2002) would be that the financial aid award given to a student with an SAT score approaching (for example) 1300 from above differs from that given to a student with an SAT score approaching 1300 from below.

In Angrist and Lavy (1999), an identifying assumption would be that the class size for a student in a school with a number of pupils approaching (for example) 800 from above differs from that of a student in a school with a # of pupils approaching 800 from below.

We will look at identification under the fuzzy design, keeping the sharp design in mind as a special case.

Constant Treatment Effects

Rewrite outcome for individual i as

$$y_i = a_i + x_i b_i, \text{ with } a_i \equiv y_{0i} \ \& \ b_i \equiv y_{1i} - y_{0i}.$$

Suppose treatment effect b is constant across individuals. This is the constant treatment effect case.

(Recall that in our previous discussion in the case of variable treatment effects we generally took as our goal the identification of the average treatment effect, or the average treatment effect among the treated, because the identification requirements for measures of the distribution of treatment effects were too great.

Additionally, we saw that when the treatment effect was constant across individuals the average treatment effect among the treated was equal to the average treatment effect.)

One assumption in addition to the discontinuity of *f(z)* at z_0 will be required to establish nonparametric identification of the treatment effect in this case:

Assumption (A1): $E[a_i \mid z_i = z]$ is continuous in z at z_0.

This assumption is valid where we have reason to believe that persons close to threshold z_0 are similar, and thus would experience similar outcomes absent treatment.

Theorem 1: Suppose $b_i = b \ \forall i$, and (RD) and (A1) hold. Then

$$b = \frac{y^+ - y^-}{x^+ - x^-},$$

where $y^+ \equiv \lim_{z \to z_0^+} E[y_i \mid z_i = z]$ and $y^- \equiv \lim_{z \to z_0^+} E[y_i \mid z_i = z]$.

Proof: Let $e > 0$ be an arbitrarily small number. The mean difference in outcomes for those above and below the discontinuity point is

$$
\begin{aligned}
&E[y_i \mid z_i = z_0 + e] - E[y_i \mid z_i = z_0 - e] \\
&\quad = b \cdot \left\{ E[x_i \mid z_i = z_0 + e] - E[x_i \mid z_i = z_0 - e] \right\} \\
&\quad + \left\{ E[a_i \mid z_i = z_0 + e] - E[a_i \mid z_i = z_0 - e] \right\}.
\end{aligned}
$$

Under (A1),

$$
\begin{aligned}
&\lim_{z \to z_0^+} E[y_i \mid z_i = z] - \lim_{z \to z_0^-} E[y_i \mid z_i = z] \\
&\quad = b \cdot \{\lim_{z \to z_0^+} E[x_i \mid z_i = z] - \lim_{z \to z_0^-} E[x_i \mid z_i = z]\}.
\end{aligned}
$$

The conclusion follows from here. The denominator is nonzero by (RD).

For the sharp design, $x^+ = 1$ and $x^- = 0$. Here the treatment effect is identified simply by

$$b = y^+ - y^-.$$

<u>Variable Treatment Effects</u>

This is the case in which b_i need not be common across individuals.

See Hahn, Todd and Van der Klaauw for a discussion of the additional assumption required for nonparametric identification of the treatment effect at z_0 in this case (actually, 2 possible assumptions are presented, each of which suffices to identify the average treatment effect at z_0).

An important result for the variable treatment effects case is that only treatment effects at z_0 can be identified nonparametrically using the regression discontinuity method.

Estimation

In both the sharp & fuzzy designs, the treatment effect is identified by

$$\frac{y^+ - y^-}{x^+ - x^-}.$$

Therefore with consistent (& nonparametric) estimators of the four one-sided limits $\hat{y}^+, \hat{y}^-, \hat{x}^+, \hat{x}^-$, the treatment effect can be estimated consistently (& nonparametrically) by

$$\frac{\hat{y}^+ - \hat{y}^-}{\hat{x}^+ - \hat{x}^-}.$$

Consider one possibility for the nonparametric estimator for the limits, the one-sided uniform kernel. Here the estimates of the limits are equivalent to

$$\hat{y}^{+} = \frac{\sum_{i\in\Lambda} y_i w_i}{\sum_{i\in\Lambda} w_i}, \quad \hat{y}^{-} = \frac{\sum_{i\in\Lambda} y_i (1-w_i)}{\sum_{i\in\Lambda} (1-w_i)},$$
$$\hat{x}^{+} = \frac{\sum_{i\in\Lambda} x_i w_i}{\sum_{i\in\Lambda} w_i}, \quad \hat{x}^{-} = \frac{\sum_{i\in\Lambda} x_i (1-w_i)}{\sum_{i\in\Lambda} (1-w_i)},$$

where Λ is the subsample such that $z_0 - h < z_i < z_0 + h$, $h > 0$ is the bandwidth, & $w_i \equiv 1(z_0 < z_i < z_0 + h)$.

An interesting note is that this estimator is numerically equivalent to an IV estimator for the regression of y_i on x_i in subsample Λ & using $w_i \equiv 1(z_0 < z_i < z_0 + h)$ as an instrument.

Hahn, Todd & Van der Klaauw note that the one-sided uniform kernel estimator for the limits causes the treatment effect estimate based on these limits to have undesirable properties, and propose an alternative approach estimating the limits by local linear regression.

[to Van der Klaauw (2002) RD application estimates/graphs]