Applied Econometrics
Prof. Leo Feler
Quiz 4: Regression Discontinuity

Name: Key

1. What is regression discontinuity design? When can we use it? And why do we use it?

Regression discontinuity design is a technique that takes advantage of an arbitrary rule determining assignment to treatment to estimate the effect of that treatment. We use it because we want to approximate random experiments as closely as possible even though we are not able to conduct actual random experiments.

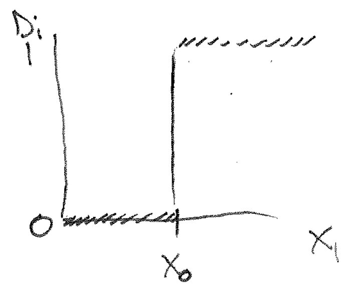
We can use it when treatment is a deterministic and discontinuous function of a covariate X_i:

$$D_i = \begin{cases} 1 & \text{if } X_i \geq X_0 \\ 0 & \text{if } X_i < X_0 \end{cases}$$

X_0 = cutoff value

"deterministic": whether you receive treatment completely depends on your value of variable X_i

"discontinuous": no matter how close you are to X_0, you only get treated once you actually reach X_0



This is a sharp regression discontinuity.

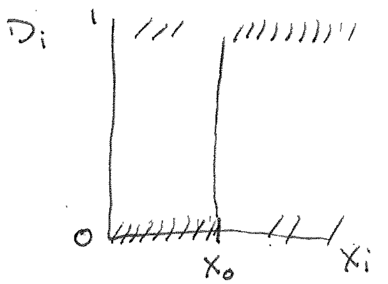
2. How do you implement a regression discontinuity approach to estimating a treatment effect? What are some issues you need to be wary of? And how do you get around these issues to estimate a treatment effect (discuss specifically how you might use instrumental variables to estimate a treatment effect if the treatment discontinuity is "fuzzy")?

One method is to include a dummy variable for treatment: $Y_i = \alpha + \beta X_i + \rho D_i + z_i$

We could also allow for polynomials in the covariates: $Y_i = \alpha + \beta_1 X_i + \beta_2 X_i^2 + \rho D_i + z_i$

we can allow for changing trends with treatment: $Y_i = \alpha + \beta_1 X_i + \rho D_i + \beta_2 X_i D_i + z_i$

An issue to be aware of is when treatment is not completely deterministic.



Some individuals who should not receive treatment do, and some who should do not.

This is a fuzzy discontinuity.

We get around this issue by looking at the probability of receiving treatment, (predicted assignment to treatment) PD_i which is still a discrete, deterministic rule.

Use this probability of treatment as an instrumental variable for treatment.

See Mostly Harmless pp. 259-262 for the proof.

estimate $= \frac{\tilde{\theta}}{\theta}$

Structural

$Y = \gamma + \pi D_i + X\beta$

First

$D_i = \alpha + \theta PD_i + X\beta$

Reduced

$Y = \tilde{\alpha} + \tilde{\theta} PD_i + X\tilde{\beta}$

Another option is to focus on a narrow band of observations around the cutoff.

$X_0 \pm \Delta$, so $E[Y_i | X_0 - \Delta < X_i < X_0] \approx E[Y_{0i} | X_i = X_0]$

$E[Y_i | X_0 \leq X_i < X_0 + \Delta] \approx E[Y_{1i} | X_i = X_0]$

Basically, estimate $Y_i = \alpha + \beta X_i + \rho D_i + z_i$ only using observations in $[X_0 - \Delta, X_0 + \Delta]$

Issues with this approach are that Δ is an arbitrary amount and we lose observations ($\uparrow$ standard errors)