Regression Discontinuity II: Extensions

Lecture 8 - Introduction to Causal Inference

Kevin Li

Parametric Estimators

While the most common RDD estimators use the machine learning local linear regression (see last class), we can also fit a normal linear regression or polynomial on both sides of the cutoff.

You should generally not do this unless you have very small sample sizes, in which local linear regression performs worse.

▶ This is because local linear regression weights units closer to the cutoff more, which allows for more accurate estimation of the jump at the cutoff.

▶ Normal linear models and polynomial models are also sensitive to the choices we make when modelling them - rather than local linear regression using machine learning to optimise.

Linear Parametric Model

We can fit linear lines on both sides of the cutoff as follows:

$$Y_i = \alpha + \tau D_i + \beta_1 \tilde{X}_i + \beta_2 \tilde{X}_i D_i + \varepsilon_i$$

▶ The estimate of τ is our causal estimate of the LATE.

▶ β_1 is the slope of the line fit for values $\tilde{X}_i < 0$ (below the cutoff, the side with no treatment).

▶ $\beta_1 + \beta_2$ is the slope of the line fit for values $\tilde{X}_i > 0$ (above the cutoff, the side with treatment).

The linear model also allows us to add control variables, or fixed effects if we are dealing with panel data.

Quadratic Parametric Model

Perhaps the linear lines fit on both sides don't capture the true relationship between $\tilde{X}$ and Y well, which could result in bad estimates for the "jump" $\hat{\tau}$.

Instead, we can fit quadratics on both sides:

$$Y_i = \alpha + \tau D_i + \beta_1 \tilde{X}_i + \beta_2 \tilde{X}_i^2 + \beta_3 \tilde{X}_i D_i + \beta_4 \tilde{X}_i^2 D_i + \varepsilon$$

▶ The estimate of τ is our causal estimate of the LATE.

▶ $\beta_1 \tilde{X}_i + \beta_2 \tilde{X}_i^2$ is the polynomial fit for values $\tilde{X} < 0$ (below the cutoff, the side with no treatment).

▶ $(\beta_1 + \beta_3)\tilde{X}_i + (\beta_2 + \beta_4)\tilde{X}_i^2$ is the polynomial fit for values $\tilde{X} > 0$ (above the cutoff, the side with treatment).

Regression Kink

Regression Discontinuity focuses on if there is a "jump"/"discontinuity" in Y at some cutoff.

What if there isn't a "jump", but a change in the relationship (slope) between the running variable and outcome?
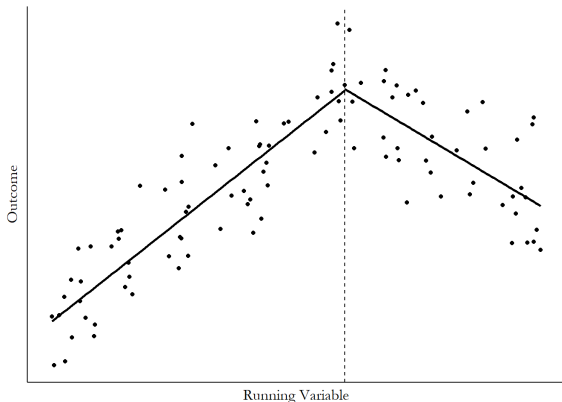
Example

Imagine you are interested in the relationship between income (X) and how many hours an individual works (Y).

Let us that individuals under $X = £40,000$ have a 20% income tax rate, while thus above have a 40% income tax rate.

▶ We probably don't expect that right below or above £40,000, people will suddenly work significantly different hours (no jump). Thus, a discontinuity design isn't very useful.

▶ However, we might expect that since individuals above the cutoff face higher taxes, they might have less incentive to work more - so the relationship between income X and hours worked Y will change.

Regression Kink Setup

Our setup for regression kink is very similar to standard RDD:

▶ We have a treatment D and a outcome Y.

▶ Treatment D is assigned based on some **running** variable X. Some cutoff $X = c$ determines if a unit is treated or not treated.

Difference: Instead of looking for a jump in Y, we will look for a slope shift in the relationship between X and Y.

Thus, our goal is not to find the LATE, but instead, to find the **local average response**.

▶ Local average response is how much the relationship between X and Y changes at the cutoff $X = c$.

Derivatives and Discontinuity

Our goal in regression kink is to find a change in the slope. Let us first define the relationship between Y and X as a function:

$$f(X) = Y$$

▶ We will not specify any function form for now - this is just for illustration purposes.

What is the slope between X and Y? Well, it is the first derivative of the function $f(X) = Y$.

This implies that our goal in regression kink is to **find a discontinuity in the first derivative** $f'(X)$ at the cutoff point $X = c$.

Assumption: Continuity of Derivative

In regression discontinuity, one of the assumptions was that if there was no treatment at $X = c$, then there would be no "jump"/discontinuity in Y (so $Y_i(0)$ is continuous).

Our goal in regression kink is to find a discontinuity in the first derivative $f'(X)$ at the cutoff point $X = c$.

▶ Thus, our identification assumption in regression kink is that if there was no treatment, $f'(X)$ is continuous at the cutoff.

▶ Or in terms of potential outcomes, if $f(X) = Y_i(0)$, then $f'(X)$ must be continuous at the cutoff $c = 0$.

If this is not true (i.e. there is a "jump"/"discontinuity" in the slope/derivative at $X = c$ even without treatment), then we do not know if the "jump"/"discontinuity" we observe is a result of the treatment D.

Estimation

Estimation is done in a similar way as regression discontinuity:

▶ Fit a best-fit line on both sides of the cutoff.

▶ Then, find the derivative of the best-fit line.

▶ Finally, check for any discontinuity in the derivatives.

The recommended procedure is to use the local linear regression method, just like in regression discontinuity.

▶ The **rdrobust** package allows for a simple option to turn a regression discontinuity into a regression kink.

We can also estimate this with a parametric model (linear, quadratic), but again, this is not recommended unless your sample size is very small.