MQE: Economic Inference from Data: Module 6: Regression Discontinuity

Claire Duquennois

6/9/2020

Regression discontinuity research designs

Introduced in other fields as far back as the 1960s.

Gain popularity in economics in the past 20 years or so as economists:

- increasingly focus on causal inference
- large administrative datasets became more widely available

When correctly applied, RD designs are very transparent in how they acheive causal identification which makes them very appealing.

Regression discontinuity research designs

RD designs leverage the researchers knowledge of a rule or policy that determins treatment.

Identification comes from how some rules are applied in a fairly arbitrairy way.

This arbitrairy application generates randomness we can exploit to estimate causal effects.

The Set Up

Suppose that we want to estimate the effect of some binary treatment D_i on an outcome Y_i. Using the potential outcomes framework:

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

Suppose that the value of D_i-i.e. whether or not an individual gets treated- is completely (or partially) determined by whether some predictor X_i lies above or below a certain threshold, c.

The Set Up

The predictor X_i need not be randomly assigned. In fact, we assume that it is related to the potential outocmes $Y_i(0)$ and $Y_i(1)$, but that this relationship is smooth, i.e. $Y_i(0)$ and $Y_i(1)$ do not jump discontinuously as X_i changes.

$\Rightarrow$ any discontinuous change in Y_i as X_i crosses c can thus be interpreted as a causal effect of D_i. We call X_i the "running variable".