VAM: Regression Discontinuity Design

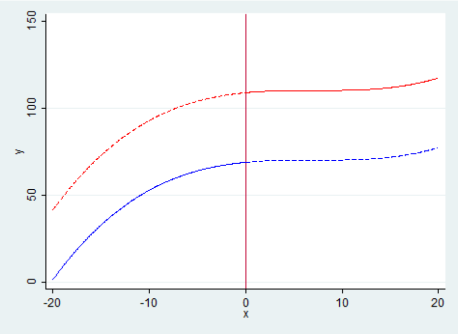
Salomo Hirvonen

University of Bristol

RDD Theory

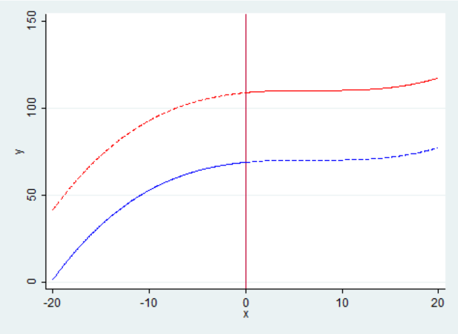
- Regression Discontinuity Design is a quasi-random method with a growing popularity in past ten years.
- The fundamental problem of causal inference: only $Y_i(1)$ or $Y_i(0)$ is observed.
- In RD-design treatment status T_i is determined by forcing variable X_i E.g. (e.g. test score, vote margin etc.) and a cut-off c.
- Thus (for sharp RDD) we observe: $Y_i = (1 - T_i) \cdot Y_i(0) + T_i \cdot Y_i(1)$

$$Y_i = \begin{cases} Y_i(0) & \text{if } X_i < c \\ Y_i(1) & \text{if } X_i \geq c \end{cases}$$



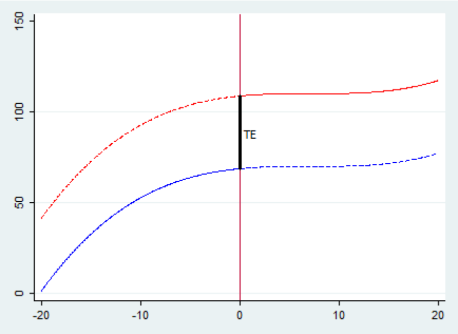
RDD Theory

- ATE given forcing variable:
  $\mathbb{E}[Y_i(1)|X_i = x] - \mathbb{E}[Y_i(0)|X_i = x]$
- If functions continuous at $x = c$:
  $\mathbb{E}[Y_i(1) - Y_i(0)|X_i = c] = \lim_{x \downarrow c} \mathbb{E}[Y_i|X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i|X_i = x] = \tau_{rdd}$



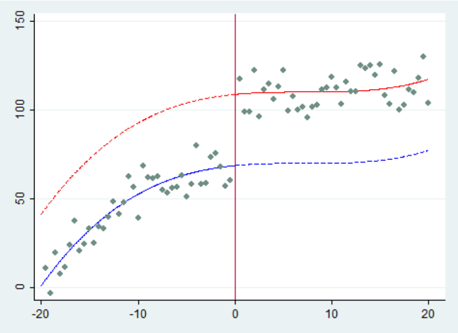
RDD Theory

- ATE given forcing variable: $\mathbb{E}[Y_i(1)|X_i = x] - \mathbb{E}[Y_i(0)|X_i = x]$
- If functions continuous at $x = c$: $\mathbb{E}[Y_i(1) - Y_i(0)|X_i = c] = \lim_{x \downarrow c} \mathbb{E}[Y_i|X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i|X_i = x] = \tau_{rdd}$
- Can be interpreted as local average treatment effect for treated.



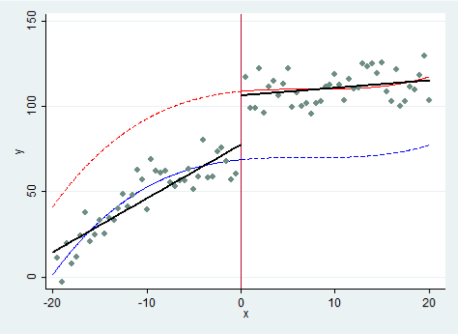
Bandwidth Choice

- In practice we have to choose how to measure the "jump" at the cut-off.
- In terms of bandwidth choice there is a trade-off between bias and variance.
- A large bandwidth implies less variance but more bias.



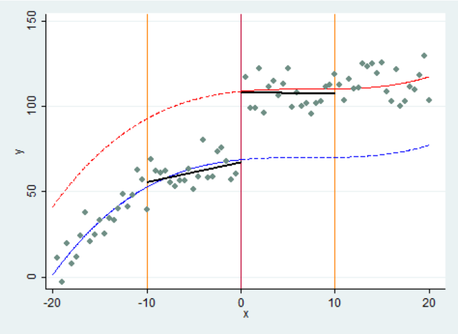
Bandwidth Choice

- In practice we have to choose how to measure the "jump" at the cut-off.
- In terms of bandwidth choice there is a trade-off between bias and variance.
- A large bandwidth implies less variance but more bias.



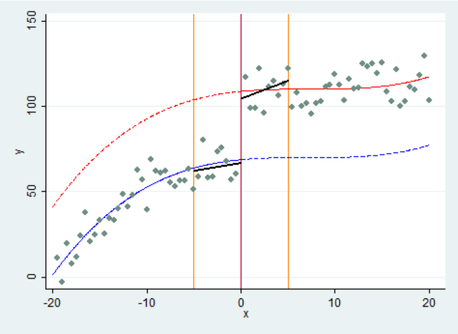
Bandwidth Choice

- In practice we have to choose how to measure the "jump" at the cut-off.
- In terms of bandwidth choice there is a trade-off between bias and variance.
- A large bandwidth implies less variance but more bias.



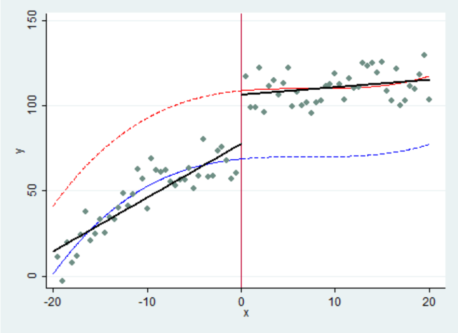
Bandwidth Choice

- In practice we have to choose how to measure the "jump" at the cut-off.
- In terms of bandwidth choice there is a trade-off between bias and variance.
- A large bandwidth implies less variance but more bias.
- For a small bandwidth it is other way around.
- Current convention is to choose a (theoretically appealing) MSE-optimal (data driven) bandwidth.



RDD More Choices

- We have to choose also the order of polynomial for the regression functions on both side of the cut-off. In practice one should limit to the low order polynomials (Gelman and Imbens (2017)) i.e. first and second order.
- Triangular weighting within the bandwidth is a common choice in practice.
- However remaining bias, which depends on the curvature of the underlying function, imposes a problem for inference.



Bias-corrected Robust Inference

- Calonico et al. (2014) provides a robust bias correction method for valid confidence intervals.
- It involves estimating the bias by using a polynomial of order $p + 1$ and having a new variance formula for the Studentization, which takes bias correction in the account.
- In practice pilotbandwidth can be set equal to the main bandwidth.
- Hyytinen et al. (2018) shows that CCT-method, unlike conventional not bias corrected, recovers lottery estimate of incumbency advantage in Finland.
- CCT can be implemented by rdrobust package, which is available for both R and STATA.

Plotting Data

- rdrobust comes with a companion command: rdplot.
- Visualising data important (first) step for rd-analysis.
- Evenly-spaced bins: keeps length of the bin same (for respective sides of the cut-off).
- Quantile-spaced bins: keeps number of observations about the same across bins (across the side of the cut-off), varying in length.
- Data driven methods for number of bins: Mimic Variance and IMSE-optimal.

Validation/Falsification

- Why we should be concerned on following things?
- Density of the running variable near the cut-off
- Covariate smoothness
- Placebo cut-offs