Econometrics

Regression Discontinuity Designs (RDD)

Paul P. Momtaz

The Anderson School
UCLA

Outline

Outline

Introduction

Assignment Probability and Local Continuity Assumption

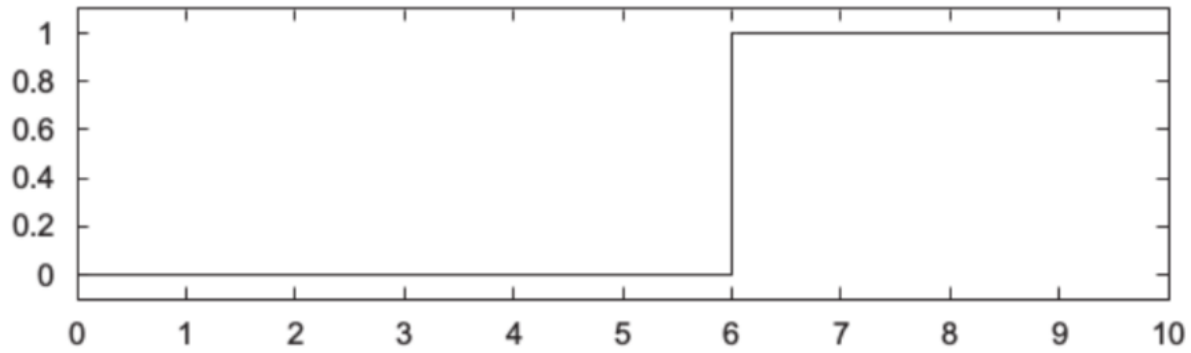


Fig. 1. Assignment probabilities (SRD).

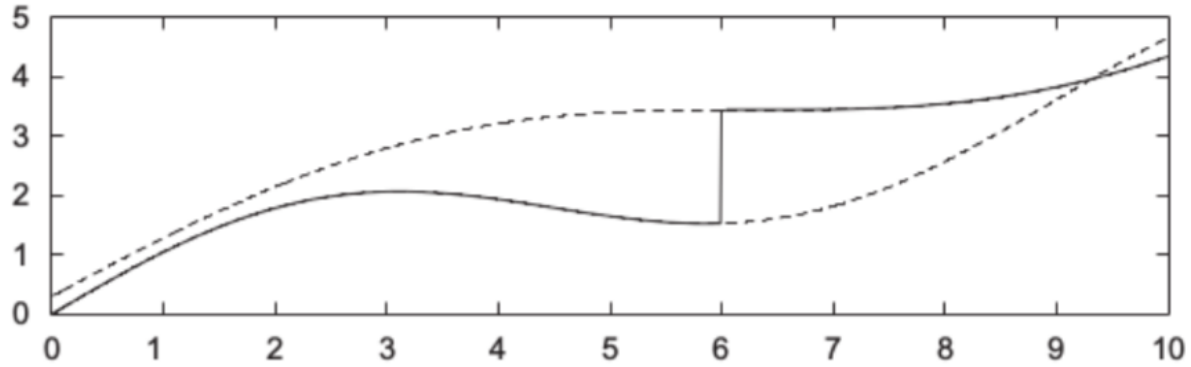
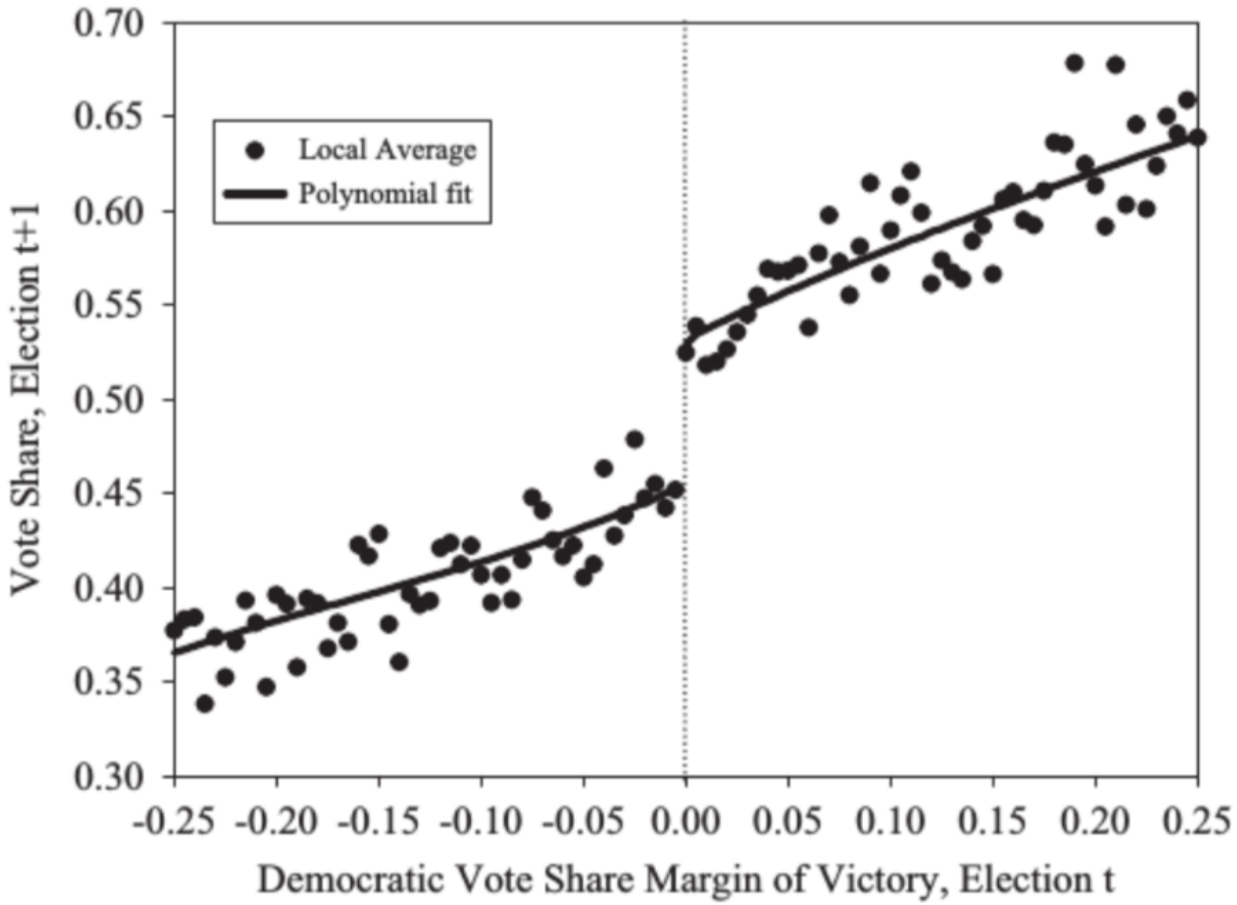


Fig. 2. Potential and observed outcome regression functions.

Introduction

Lee 2008



Introduction

Treatment is deterministic and discontinuous function of remaining variable

$$D_i = \begin{cases} 1 & \text{if } x_i \geq x_0 \\ 0 & \text{if } x_i < x_0 \end{cases}$$

$$y_i = f(x_i) + \rho D_i + \eta_i$$

Interaction Terms:

$$\mathbb{E}[y_{0i}|x_i] = f_0(x_i) = \alpha + \beta_{01}\tilde{x}_i + \cdots + \beta_{0\rho}\tilde{x}_i^{\rho}$$
$$\mathbb{E}[y_{1i}|x_i] = f_1(x_i) = \alpha + \beta_{11}\tilde{x}_i + \cdots + \beta_{1\rho}\tilde{x}_i^{\rho}$$

Where, $\tilde{x}_i = x_i - x_0$ (centring)

Introduction

$$\mathbb{E}[y_i|x_i] = \mathbb{E}[y_{0i}|x_i] + \mathbb{E}[y_{i1} - y_{i0}|x_i]D_i$$

Substituting in yields

$$y_i = \alpha + \beta_{01}\tilde{x}_i + \cdots + \beta_{0\rho}\tilde{x}_i^{\rho} + \rho D_i + \beta_1^* D_i\tilde{x}_i + \cdots + \beta_{\rho}^* D_i\tilde{x}_i^{\rho} + \eta_i$$

$$\text{where } \beta_1^* = \beta_{11} - \beta_{01}$$

$\Rightarrow$ No restriction on conditional mean functions

Outline

Discontinuity Sample

Test idea:

- Is discontinuity an unaccounted-for-nonlinearity?

$$\lim_{\Delta \to 0} \mathbb{E}[y_i | x_0 < x_i < x_0 + \Delta] - \mathbb{E}[y_i | x_0 - \Delta < x_i < x_0]$$
$$= \mathbb{E}[y_{1i} - y_{0i} | x_i = x_0]$$

Doest not depend on

- correct specification of $\mathbb{E}[y_{0i}|x_i]$ model
- constant effects assumption, $y_{1i} - y_{0i} = \rho_0^*$

But requires

- good estimate of mean of y_i
- enough data

$\Rightarrow$ Bins too narrow $\rightarrow$ imprecise; too wide $\rightarrow$ bias

Outline

McCrary (2008) Density Test

Idea: Test whether aggregate distribution of running variable discontinuous

Procedure:

1. Partition running variable into equally sized bins and compute frequency.
2. Frequency count as dependent variable in local linear regression

This test can fail if upward jumps set off by downward jumps !

Inspect baseline covariates: Replace dependent variable with each of the observed baseline covariates to check whether they are locally balanced or each side of the threshold

Fuzzy RDD Motivation

Keys Mukherjee Seru Vig 2010

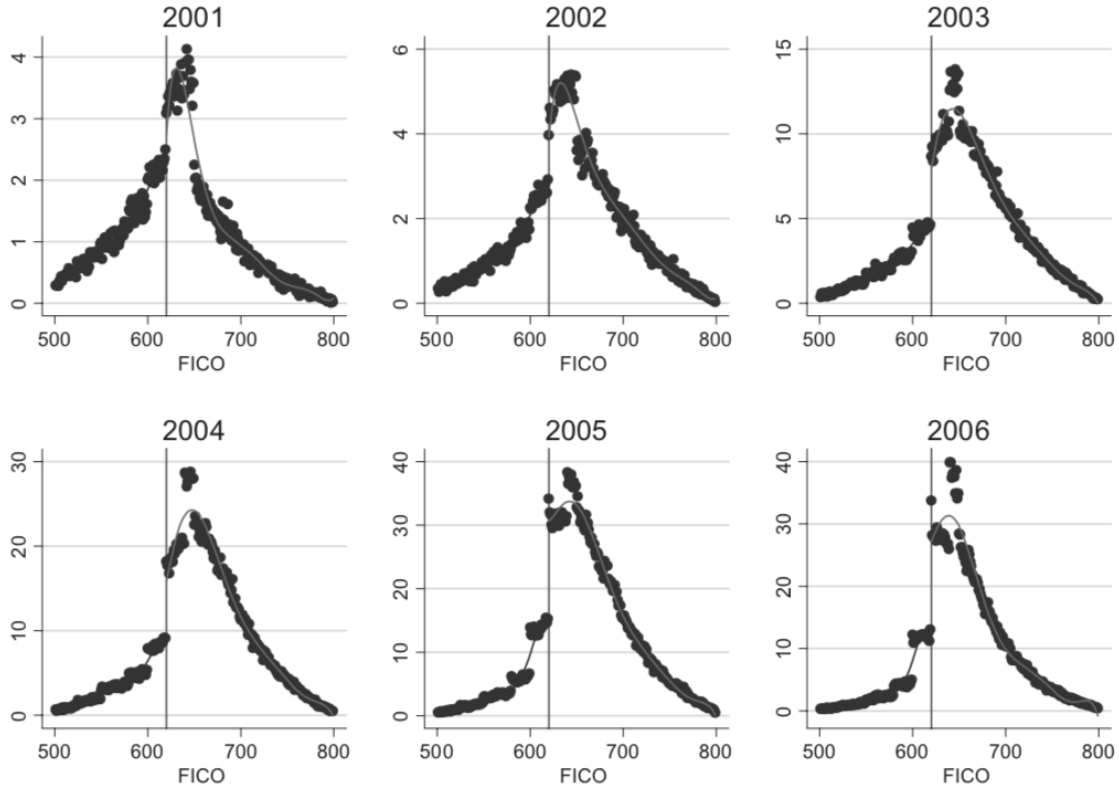


FIGURE II

Number of Loans (Low-Documentation)

The figure presents the data for number of low-documentation loans (in '00s). We plot the average number of loans at each FICO score between 500 and 800. As can be seen from the graphs, there is a large increase in the number of loans around the 620 credit threshold (i.e., more loans at 620^+ as compared to 620^-) from 2001 onward. Data are for loans originated between 2001 and 2006.

Fuzzy RDD = IV

Idea: Exploit discontinuities in probability of treatment conditional on covariate.

Discontinuity = IV

$$P[D_i = 1|x] = \begin{cases} g_0(x_i) & \text{if } x_i \geq x_0 \\ g_1(x_i) & \text{if } x_i < x_0 \end{cases}$$

$$\mathbb{E}[D_i|x_i] = P[D_i = 1|x_i] = g_0(x_i) + [g_1(x_i) - g_0(x_i)]\,T_i$$

Where $T_i = 1(x_i \geq x_0)$

Fuzzy RDD = IV

First Stage:

$$D_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \cdots + \gamma_\rho x_i^\rho + \pi T_i + \xi_{1i}$$

Fuzzy RDD Reduced Form:

$$y_i = \mu + u_1 x_i + u_2 x_i^2 + \cdots + u_\rho x_i^\rho + \rho\pi T_i + \xi_{2i}$$

where $\mu = \alpha + \beta\gamma_0$ and $u_j = \beta_1 + \rho\gamma_j$

Interaction terms: Center polynomial term around x_0
First Stage:

$$D_i = \gamma_{00} + \gamma_{01}\tilde{x}_i + \cdots + \gamma_{0\rho}\tilde{x}_i^\rho + \gamma_0^* T_i + \gamma_1^* \tilde{x}_i T_i + \cdots + \gamma_\rho^* \tilde{x}_i^\rho T_i$$

Second Stage: Same as in sharp RDD

Special Types of RDD

Special types of RDD:

- Geographic
- Categorical Running Variable
- Donut RDD
- Kink RDD
- Dynamic RDD

Comments on RDD

Comments:

- Sharp RDD $\rightarrow$ ATE, fuzzy RDD $\rightarrow$ ATT
- fuzzy RDD: variation in treatment near threshold randomized.
- Bandwidth choice for graphical representation