Quant II

Lab 10: RDD II

Giacomo Lemoli

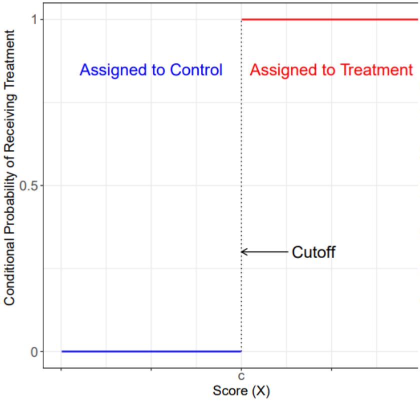
April 13, 2023

Today's plan

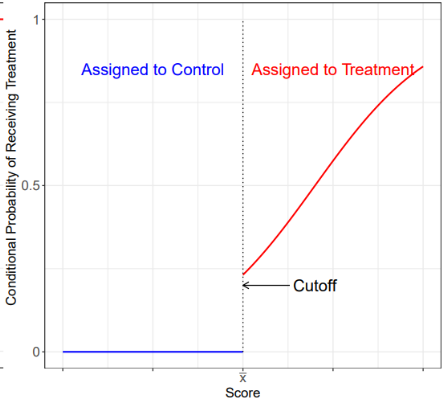
- Fuzzy RDD
- Multiple cutoffs and GRDD
- RDD and heterogeneity: Diff-in-disc
- Local randomization approach
- RDD for the effect of politician characteristics (PCRDD)

Fuzzy RDD

Endogenous treatment take-up or non-compliance, but discontinuous change in treatment probability.



(a) Sharp RD



(b) Fuzzy RD (One-Sided)

Fuzzy RDD

- IV model where the treatment D is endogenous and the cutoff rule gives the instrument. Under the usual IV assumptions:

$$\frac{\lim_{x\to c^+} E[Y_i|X=x] - \lim_{x\to c^-} E[Y_i|X=x]}{\lim_{x\to c^+} E[D_i|X=x] - \lim_{x\to c^-} E[D_i|X=x]} = \tau_{FRD}$$

Fuzzy RDD

- IV model where the treatment D is endogenous and the cutoff rule gives the instrument. Under the usual IV assumptions:

$$\frac{\lim_{x\to c^+} E[Y_i|X=x] - \lim_{x\to c^-} E[Y_i|X=x]}{\lim_{x\to c^+} E[D_i|X=x] - \lim_{x\to c^-} E[D_i|X=x]} = \tau_{FRD}$$

- Under heterogeneous TE and motonicity, τ_{FRD} has a LATE interpretation.

Fuzzy RDD

- IV model where the treatment D is endogenous and the cutoff rule gives the instrument. Under the usual IV assumptions:

$$\frac{\lim_{x \to c^+} E[Y_i|X = x] - \lim_{x \to c^-} E[Y_i|X = x]}{\lim_{x \to c^+} E[D_i|X = x] - \lim_{x \to c^-} E[D_i|X = x]} = \tau_{FRD}$$

- Under heterogeneous TE and motonicity, τ_{FRD} has a LATE interpretation.
- Estimation: once IV regression, today argument `fuzzy` in `rdrobust`

Multi-score RDD

- Treatment assignment depends on more than one RV

Multi-score RDD

- Treatment assignment depends on more than one RV
  - School admission based on multiple test scores

Multi-score RDD

- Treatment assignment depends on more than one RV
  - School admission based on multiple test scores
  - Policy eligibility based on more than one criteria

Multi-score RDD

- Treatment assignment depends on more than one RV
  - School admission based on multiple test scores
  - Policy eligibility based on more than one criteria
  - Treatment based on being on a side of a boundary

Multi-score RDD

Differences relative to standard RDD:

- Continuity along both dimensions

Multi-score RDD

Differences relative to standard RDD:

- Continuity along both dimensions
- Effect at every boundary point: a *treatment effect curve*

Multi-score RDD

Differences relative to standard RDD:

- Continuity along both dimensions
- Effect at every boundary point: a *treatment effect curve*
- One could estimate the effect at different points of the boundary to characterize heterogeneity

GRDD

- Effect of geographic boundaries: a long interest of political science

GRDD

- Effect of geographic boundaries: a long interest of political science
- Political/administrative boundaries

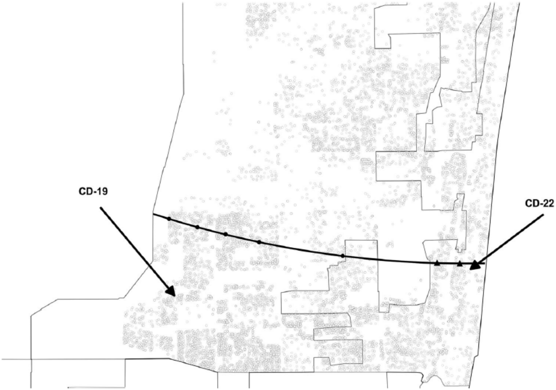
GRDD

- Effect of geographic boundaries: a long interest of political science
- Political/administrative boundaries
- Media market boundaries

GRDD

Velez and Newman (2018)

FIGURE 4 **Congressional District Boundaries and Voter Locations**



Note: The bold line represents the coverage boundary. The FCC reception boundary contains several neighborhoods and two congressional districts.

GRDD

Dell (2010)

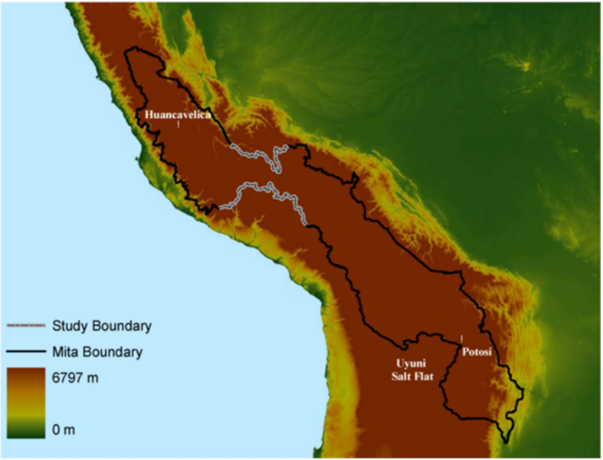
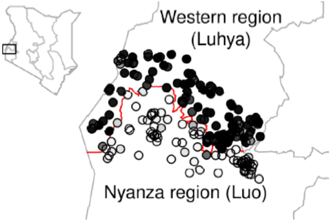


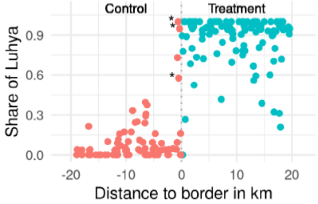
FIGURE 1.—The *mita* boundary is in black and the study boundary in light gray. Districts falling inside the contiguous area formed by the *mita* boundary contributed to the *mita*. Elevation is shown in the background.

GRDD

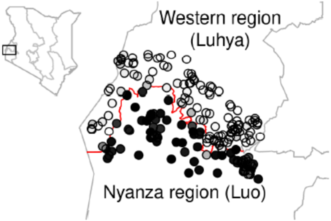
Müller-Crepon (2021)



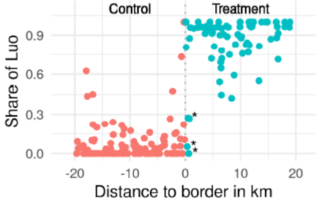
(a) Luhya share



(b) Luhya share



(c) Luo share



(d) Luo share

GRDD

Rozenas, Schutte, and Zhukov (2017)

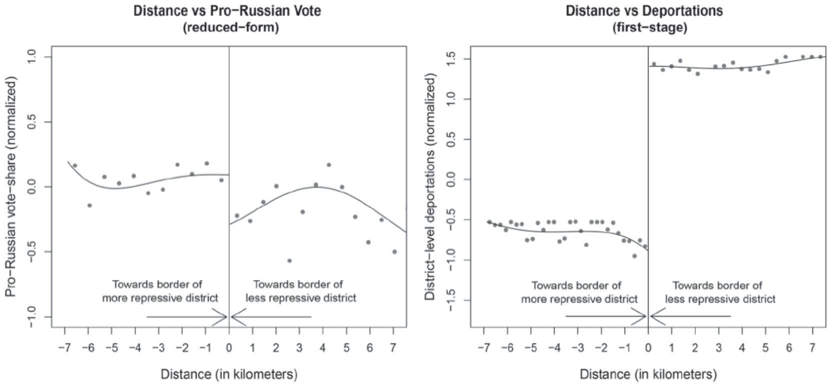


Figure 3. Reduced-form relationships between the instrument (distance from the contiguous district with more repression), deportations, and pro-Russian vote.

GRDD

Discussion in Keele and Titiunik (2015)

Challenges:

- Compound treatments: many things change at the same boundary

GRDD

Discussion in Keele and Titiunik (2015)

Challenges:

- Compound treatments: many things change at the same boundary
- Geographic sorting

GRDD

Discussion in Keele and Titiunik (2015)

Challenges:

- Compound treatments: many things change at the same boundary
- Geographic sorting
- Sparsity around the boundary

GRDD

Discussion in Keele and Titiunik (2015)

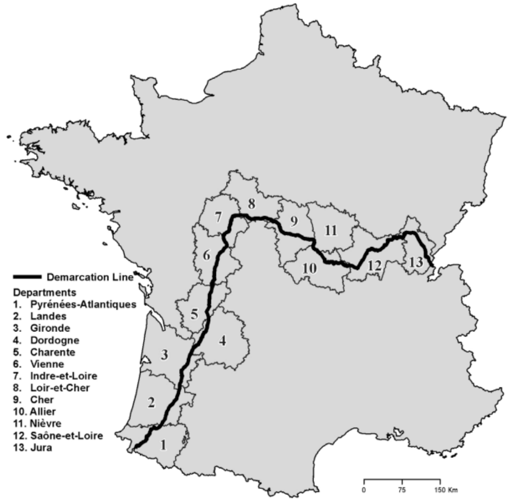
Challenges:

- Compound treatments: many things change at the same boundary
- Geographic sorting
- Sparsity around the boundary
- Non-exogeneity of the boundary

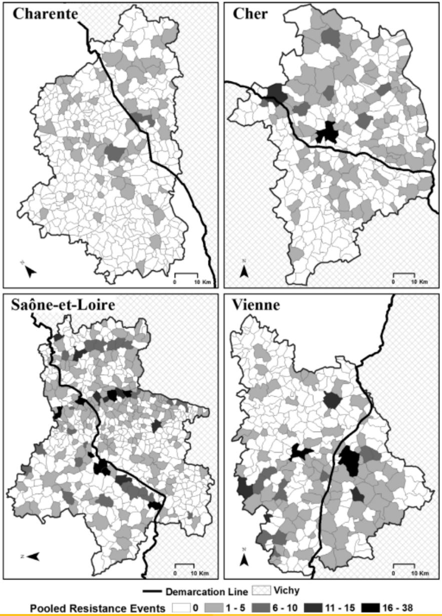
Non-exogeneity of the boundary

Ferwerda and Miller (2014)

FIGURE 1. Map of the Demarcation Line across Intersected French Departments



Non-exogeneity of the boundary



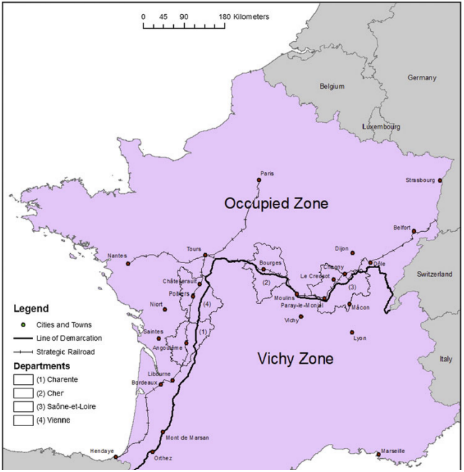
Non-exogeneity of the boundary

Kocher and Monteiro (2016)

Non-exogeneity of the boundary

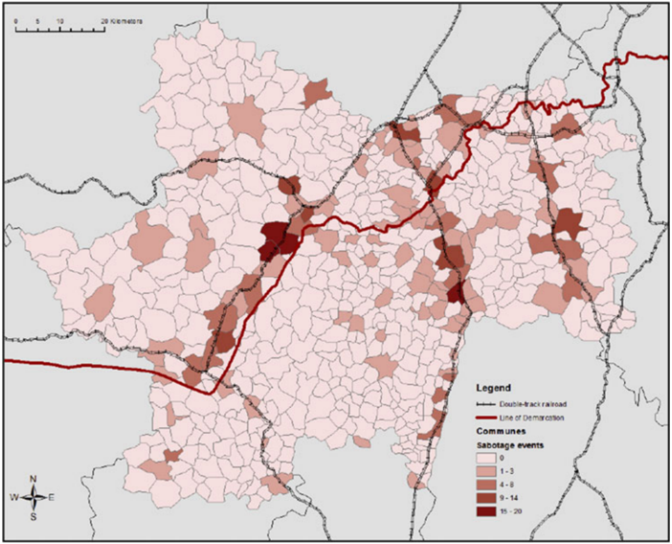
Kocher and Monteiro (2016)

Map 1
The Line of Demarcation (LoD) and Nantes-Tours-Belfort and Paris-Tours-Bordeaux railroads



Non-exogeneity of the boundary

Map 4
Double-track railways and sabotage in Saône-et-Loire, 1940–1944



GRDD

Example

Difference-in-discontinuities

- Add a dimension of heterogeneity (e.g. time) to RDD

Difference-in-discontinuities

- Add a dimension of heterogeneity (e.g. time) to RDD
- Addresses problems of compound treatment if one of the treatments is introduced *after* the other

Difference-in-discontinuities

- Add a dimension of heterogeneity (e.g. time) to RDD
- Addresses problems of compound treatment if one of the treatments is introduced *after* the other
- Estimate moderation effects

Difference-in-discontinuities

- Add a dimension of heterogeneity (e.g. time) to RDD
- Addresses problems of compound treatment if one of the treatments is introduced *after* the other
- Estimate moderation effects
  - Assumptions not very formalized

Difference-in-discontinuities

- Add a dimension of heterogeneity (e.g. time) to RDD
- Addresses problems of compound treatment if one of the treatments is introduced *after* the other
- Estimate moderation effects
    - Assumptions not very formalized
- Usually standard regression models within a bandwidth, more parametric assumptions

Difference-in-discontinuities

Grembi, Nannicini, and Troiano (2016)

The diff-in-disc estimator can be implemented by estimating the boundary points of four regression functions of Y_{it} on P_{it}: two on both sides of P_c, both before and after t_0. We apply a local linear regression, following Gelman and Imbens (2014).[23] The method consists in fitting linear regression functions to the observations distributed within a distance h on either side of P_c, both before and after t_0. Formally, we restrict the sample to cities in the interval $P_{it} \in [P_c - h, P_c + h]$ and estimate the model

$$\begin{aligned} (1) \qquad Y_{it} = {} & \delta_0 + \delta_1 P_{it}^* + S_i(\gamma_0 + \gamma_1 P_{it}^*) + T_t[\alpha_0 \\ & + \alpha_1 P_{it}^* + S_i(\beta_0 + \beta_1 P_{it}^*)] + \xi_{it}, \end{aligned}$$

where S_i is a dummy for cities below 5,000 capturing treatment status, T_t an indicator for the posttreatment period, and $P_{it}^* = P_{it} - P_c$ the normalized population size. Standard errors are clustered at the city level. The coefficient β_0 is the diff-in-disc estimator and identifies the treatment effect of relaxing fiscal rules, as the treatment is $R_{it} = S_i \cdot T_t$. We present the robustness of our results to multiple bandwidths h, optimally computed first following the algorithm developed by Calonico, Cattaneo, and Titiunik (2014a, b), and then implementing the cross-validation method proposed by Ludwig and Miller (2007).[24]

Difference-in-discontinuities

Larreguy, Marshall, and Querubín (2016)

TABLE 4. Effect of Split Polling Station by Distance

	Turnout (1)	PRI Vote Share (2)	PAN Vote Share (3)	PRD Vote Share (4)
Split	0.0084***	0.0034***	0.0042**	0.0006
	(0.0014)	(0.0011)	(0.0018)	(0.0011)
Distance	−0.0067	0.0058	−0.0089***	−0.0017
	(0.0049)	(0.0036)	(0.0032)	(0.0020)
Distance squared	−0.0003	−0.0006	0.0003	0.0000
	(0.0007)	(0.0007)	(0.0006)	(0.0003)
Split × Distance	0.0014	0.0110***	−0.0035	−0.0002
	(0.0047)	(0.0036)	(0.0028)	(0.0031)
Split × Distance squared	−0.0009	−0.0030**	0.0010*	−0.0005
	(0.0014)	(0.0013)	(0.0006)	(0.0007)
Observations	27,420	27,420	27,420	27,420

Notes: All specifications include district-year fixed effects, and are estimated with OLS. All results are for a 20 voter bandwidth. Block-bootstrapped standard errors are clustered by state (1,000 resamples). Locality-weighted distance to the polling station was unavailable for 347 electoral precincts. *denotes $p < 0.1$, **denotes $p < 0.05$, ***denotes $p < 0.01$.

Local randomization

- Alternative approach to RDD

Local randomization

- Alternative approach to RDD
- Instead of continuity, assume random assignment in a neighborhood of the cutoff

Local randomization

- Alternative approach to RDD
- Instead of continuity, assume random assignment in a neighborhood of the cutoff

Local randomization

- Alternative approach to RDD
- Instead of continuity, assume random assignment in a neighborhood of the cutoff

Given a window W_0 around the cutoff:

- Distribution of RV in W_0 is known, the same for all units, and does not depend on the potential outcomes: $F_{X_i|X_i \in W_0}(x) = F(x)$

Local randomization

- Alternative approach to RDD
- Instead of continuity, assume random assignment in a neighborhood of the cutoff

Given a window W_0 around the cutoff:

- Distribution of RV in W_0 is known, the same for all units, and does not depend on the potential outcomes: $F_{X_i|X_i \in W_0}(x) = F(x)$
- Inside W_0, the potential outcomes depend on the running variable solely through the treatment indicator $T_i = \mathbb{I}(X_i > c)$, but not directly: $Y_i(X_i, T_i) = Y_i(T_i) \forall i \text{s.t.} X_i \in W_0$

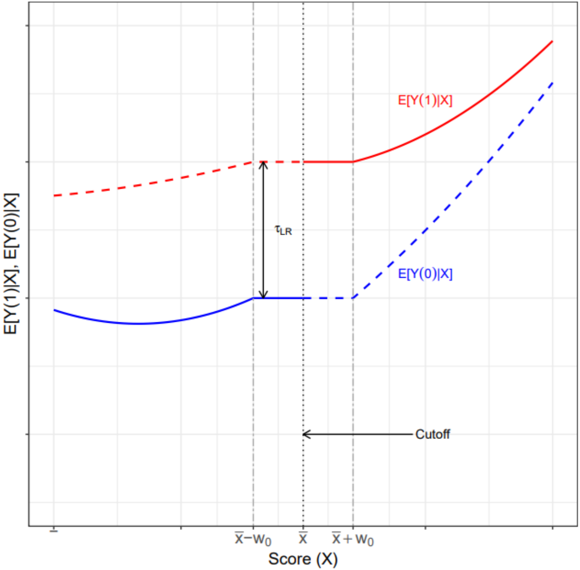
Local randomization

- Alternative approach to RDD
- Instead of continuity, assume random assignment in a neighborhood of the cutoff

Given a window W_0 around the cutoff:

- Distribution of RV in W_0 is known, the same for all units, and does not depend on the potential outcomes: $F_{X_i|X_i \in W_0}(x) = F(x)$
- Inside W_0, the potential outcomes depend on the running variable solely through the treatment indicator $T_i = \mathbb{I}(X_i > c)$, but not directly: $Y_i(X_i, T_i) = Y_i(T_i) \forall i \text{s.t.} X_i \in W_0$
- **More** restrictive than continuity, but finite sample inference methods

Local randomization



Local randomization

- Estimation and inference with the package `rdlocrand`

Local randomization

- Estimation and inference with the package `rdlocrand`
- Data-driven approach to find the window: specify pre-determined covariates, choose larger windows until T and C units are no longer balanced on them

Local randomization

- Estimation and inference with the package `rdlocrand`
- Data-driven approach to find the window: specify pre-determined covariates, choose larger windows until T and C units are no longer balanced on them
- Fisherian methods for inference: randomization inference for the sharp null

Estimation

To choose the window, specify:

- Covariates to use for balance

Estimation

To choose the window, specify:

- Covariates to use for balance
- Test statistic to use for balance (e.g. difference in means)

Estimation

To choose the window, specify:

- Covariates to use for balance
- Test statistic to use for balance (e.g. difference in means)
- Randomization mechanism

Estimation

To choose the window, specify:

- Covariates to use for balance
- Test statistic to use for balance (e.g. difference in means)
- Randomization mechanism
- Minimum no. observations in the smallest window

Estimation

To choose the window, specify:

- Covariates to use for balance
- Test statistic to use for balance (e.g. difference in means)
- Randomization mechanism
- Minimum no. observations in the smallest window
- Significance level

Estimation

To choose the window, specify:

- Covariates to use for balance
- Test statistic to use for balance (e.g. difference in means)
- Randomization mechanism
- Minimum no. observations in the smallest window
- Significance level
- Then, difference-in-means within the selected window + randomization inference

Application to Islamic mayors

```r
library(rdlocrand)
Z <- cbind(data$i89, data$vshr_islam1994, data$partycount, data$lpop1994,
           data$merkezi, data$merkezp, data$subbuyuk, data$buyuk)
colnames(Z) = c("i89 ", "vshr_islam1994", "partycount", "lpop1994",
                "merkezi", "merkezp", "subbuyuk", "buyuk")
```

Application to Islamic mayors

```r
out <- rdrandinf(Y, X, covariates = Z, seed = 50, d = 3.019522)
```

```
Number of obs      =         2629
Order of poly      =            0
Kernel type        =      uniform
Reps               =         1000
Window             =   rdwinselect
H0:           tau  =        0.000
Randomization      = fixed margins

Cutoff c =     0.000   Left of c  Right of c
      Number of obs        2314         315
 Eff. number of obs          21          26
    Mean of outcome      12.933      16.049
    S.d. of outcome       8.752      10.815
             Window      -0.875       0.875

==============================================================================

                                   Finite sample         Large sample
                               ------------------  -------------------------
----
          Statistic          T         P>|T|        P>|T|    Power vs d =
3.020
==============================================================================

     Diff. in means      3.116         0.306        0.275
0.185
==============================================================================
```

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics?

John Marshall Columbia University

Abstract: *Politician characteristic regression discontinuity (PCRD) designs leveraging close elections are widely used to isolate effects of an elected politician characteristic on downstream outcomes. Unlike standard regression discontinuity designs, treatment is defined by a predetermined characteristic that could affect a politician's victory margin. I prove that, by conditioning on politicians who win close elections, PCRD estimators identify the effect of the specific characteristic of interest* and all compensating differentials—*candidate-level characteristics that ensure elections remain close between candidates who differ in the characteristic of interest. Avoiding this asymptotic bias generally requires assuming either that the characteristic of interest does not affect candidate vote shares* or *that no compensating differential affects the outcome. Because theories of voting behavior suggest that neither strong assumption usually holds, I further analyze the implications for interpreting continuity tests and consider if and how covariate adjustment, bounding, and recharacterizing treatment can mitigate the posttreatment bias afflicting PCRD designs.*