Econometrics of Policy Evaluation: Regression Discontinuity Design (RD)

Cristian Huse

Main Objective of Program Evaluation

- **Aim:** Estimate the effect of an intervention P on a response variable Y
- **Classical examples:**
  - What is the effect of an increase in the minimum wage on employment?
  - What is the effect of a school meals program on learning achievement?
  - What is the effect of a job training program on employment and on wages?

Indexes and Thresholds

- RDD is a method which compares individuals above vs. below a given threshold, provided some assumptions are satisfied
  - Anti-poverty programs $\Rightarrow$ targeted to households below a given poverty index
  - Pension programs $\Rightarrow$ targeted to population above a certain age
  - Scholarships $\Rightarrow$ targeted to students with high scores on a standardized test

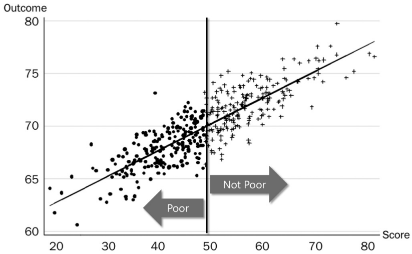
RDD

- When to use this method?
  - The beneficiaries/non-beneficiaries can be ordered along a quantifiable dimension
  - This dimension can be used to compute a well-defined index or parameter
  - The index/parameter has a cut-off point for eligibility
  - The index value is what drives the assignment of a potential beneficiary to the treatment (or to non-treatment.)
- Intuitive explanation of the method:
  - The potential beneficiaries (units) just above the cut-off point are very similar to the potential beneficiaries just below the cut-off
  - We compare outcomes for units just above and below the cutoff

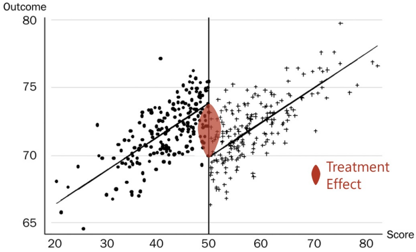
Example

- **Example.** Effect of cash transfer on consumption
- **Goal**
  - Target transfer to poorest households
- **Method**
  - Construct poverty index from 1 to 100 with pre-intervention characteristics
  - Households with a score ≤ 50 are poor
  - Households with a score >50 are non-poor
- **Implementation**
  - Cash transfer to poor households
- **Evaluation**
  - Measure outcomes (i.e. consumption, school attendance rates) before and after transfer, comparing households just above and below the cut-off point

RDD – Illustration



Baseline



Post-intervention

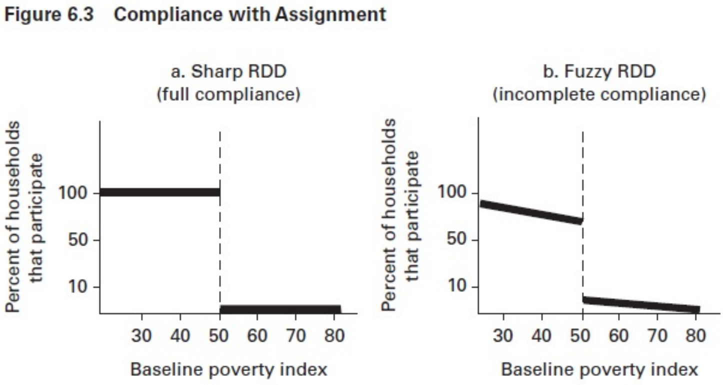
- Focus on right-hand plot:
  - TE holds around the threshold $\Rightarrow$ Local character
  - This "jump" is only a causal effect if we make one crucial assumption: in the absence of the treatment, the relationship between the poverty score (X) and consumption (Y) would have been a **smooth, continuous function across the cutoff**. The only reason for a sharp jump at the cutoff is the treatment itself.

- **Context:** Angrist & Lavy (1999) wanted to estimate the causal effect of class size on student test scores in Israel.
- **The Rule (Maimonides' Rule):** A historical rabbinic rule, used by the Israeli public school system, dictates that a single class cannot exceed 40 students.
- **The Discontinuity:** This rule creates a sharp discontinuity in class size.
  - A school with 40 5th graders has one class of 40.
  - A school with 41 5th graders must split, creating two classes (e.g., 20 and 21 students).
- **The RDD Setup:**
  - **Running Variable:** School enrollment.
  - **Cutoff:** 40 students (and 80, 120, etc.).
  - **Outcome:** Student test scores.
- **Finding:** At the cutoff, class size drops sharply, and test scores jump up. This provided credible evidence that smaller classes cause better test scores.

Sharp and Fuzzy Discontinuity

- **Sharp discontinuity**
  - The discontinuity precisely determines treatment
  - Equivalent to **random assignment in a neighborhood**
  - e.g., social security payment depend directly and immediately on a person's age
- **Fuzzy discontinuity**
  - Discontinuity is highly correlated with treatment
  - e.g., rules determine eligibility but there is a margin of administrative error
  - Use the assignment as an IV for program participation

Sharp and Fuzzy Discontinuity cont'd

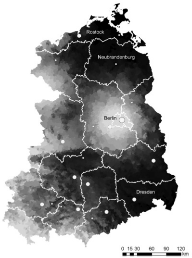


- **Incomplete compliance:**
  - Not every poor household is treated (participates in program)
  - Not every rich household is non-treated

Sharp and Fuzzy Discontinuity: Example

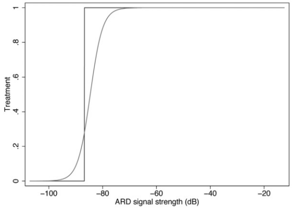
- Bursztyn and Cantoni (2016): geographic features determine access to Western TV and impact composition of consumption
  - Exposure to Western broadcasts affects composition (but not level) of consumption; choices biased towards categories of goods with higher pre-reunification advertisement. The effects vanish by 1998

FIGURE 1.—SIGNAL STRENGTH (ARD) IN EAST GERMANY, 1989



Dark: weak signal. Dots represent major cities. District borders are (*Bezirke*) superimposed as white lines.

FIGURE 4.—TREATMENT DEFINITIONS



The figure compares the binary treatment definition, based on the threshold of –86.7 dB, and the continuous treatment definition, based on the fit of actual viewership data as in figure 2, upper panel.

- Authors calculate strength of TV signal for the entire former GDR
- The treatment area comprises all regions with <u>a positive probability of reception of Western TV broadcasts</u> (lots of interesting details in paper)

Identification for Sharp Discontinuity

- Estimate regression

$$Y_i = \beta + \delta P_i + f(X_i) + \epsilon_i$$

where

$$P_i = \begin{cases} 1, & \text{if household } i \text{ participates in program} \\ 0, & \text{if household } i \text{ doesn't participate in program} \end{cases}$$

X is the **running variable**; $f(X_i)$ is a continuous function around the **cut-off point** X_0 (to account for nonlinearities in the relationship between r.v. and outcome); ϵ_i is a random error term

- We assume that assignment is strictly enforced, without any exceptions of either side of the cut-off, e.g. assume X is a test score, eligibility is below 50

$$P_i = \begin{cases} 1, & \text{score of } i < 50 \\ 0, & \text{score of } i \geq 50 \end{cases}$$

Identification for Sharp Discontinuity

Identification: The Continuity Assumption in Practice

- Recall $Y_i = \beta + \delta P_i + f(X_i) + \epsilon_i$. For a unit just at the cut-off X_0:

$$Y_{i0} = \beta + \delta \times 0 + f(X_0) + \epsilon_{i0}$$

- For a unit just τ_i below the cut-off: (location: $X_0 - \tau$)

$$Y_{i1} = \beta + \delta \times 1 + f(X_0 - \tau_i) + \epsilon_{i1}$$

- The difference between treated and non-treated is

$$Y_{i1} - Y_{i0} = \delta + f(X_0 - \tau_i) - f(X_0) + (\epsilon_{i1} - \epsilon_{i0})$$

- Since the function $f(.)$ is continuous, as we get closer and closer to the cut-off X_0, $f(X_0 - \tau_i)$ will tend to $f(X_0)$, and their difference will tend to 0
- Therefore, δ estimates the **local average treatment effect** (LATE) at the threshold. This relies on the **continuity assumption** that $f(X_0 - \tau_i)$ and $f(X_0)$ become equal as $\tau_i \to 0$.

Modern RDD Implementation: Local Linear Regression

- While we can model $f(X_i)$ using a high-order polynomial for the entire dataset (as in the Auffhammer example below), this is sensitive to the polynomial's order.
- The modern, preferred method (Imbens & Lemieux, 2008; Lee & Lemieux, 2010) is **local linear regression**:
  1. **Choose a narrow bandwidth (window)** around the cutoff (e.g., only use data for scores between 45 and 55).
  2. Run a simple linear regression within that window, allowing the slope to differ on each side. The regression is:

     $$Y_i = \beta_0 + \delta D_i + \beta_1(X_i - c) + \beta_2(X_i - c) \times D_i + \epsilon_i$$

     (where c = cutoff and D_i =treatment dummy)
  3. **The coefficient** δ is our estimate of the LATE at the cutoff.
- This method is less sensitive to non-linearities far from the cutoff.

Lab: RDD Implementation

- Will estimate $\mathit{Health} - \mathit{expenditures}_i = \beta + \delta \mathit{eligible}_i + f(X_i) + \epsilon_i$
  - Households with a score below a certain cut-off (here, 58) are chosen to participate in the program. Households with a score above that cut-off do not participate
  - We assume that program eligibility rule is strictly enforced
  - In addition to the value of the index measured at baseline and prior to the program roll-out, we also measure the outcome of interest (health expenditures) after the end of the program.
  - We start by normalizing the poverty index threshold to 0 and create dummy variables for households with a poverty-targeting index to the left or right of the threshold
    - i.e., we allow the relationship between the outcome and running variables (the poverty index) to have different slopes on either side of the threshold
  - We then run a regression of health expenditures on a dummy variable capturing exposure to the program, as well as the two dummies for whether households have a poverty index to the left or to the right of the threshold

RDD Implementation

- This example implements a version of RDD by fitting two separate lines, one on each side of the cutoff. This is known as a **global linear fit**. The **local linear** approach is a more flexible alternative.
- We obtain $\hat{\delta} = -11.19$ as the estimate of the treatment effect

```
Stata Example 12. Regression Discontinuity Design Estimates

* REGRESSION DISCONTINUITY DESIGN
* In this context, you compare health expenditures at follow-up between households
just above
* and just below the poverty index threshold, in the treatment localities.

*Select the relevant data
use "evaluation.dta", clear
keep if treatment_locality==1

*Normalize the poverty index
gen poverty_index_left=poverty_index-58 if poverty_index<=58
(8570 missing values generated)

replace poverty_index_left=0 if poverty_index>58
(8570 real changes made)

gen poverty_index_right=poverty_index-58 if poverty_index>58
(11257 missing values generated)

replace poverty_index_right=0 if poverty_index<=58
(11257 real changes made)

reg health_expenditures poverty_index_left poverty_index_right eligible if round ==1

      Source |       SS       df       MS              Number of obs =    4960
-------------+------------------------------           F(  3,  4956) =  843.52
       Model |  257911.257     3  85970.4191           Prob > F      =  0.0000
    Residual |  505111.412  4956  101.919171           R-squared     =  0.3380
-------------+------------------------------           Adj R-squared =  0.3376
       Total |   763022.67  4959  153.866237           Root MSE      =  10.096

------------------------------------------------------------------------------------
health_expenditures |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------------+---------------------------------------------------------------
 poverty_index_left |   .1755687   .0302796     5.80   0.000     .1162073      .23493
poverty_index_right |   .2202764   .0315631     6.98   0.000     .1583987    .2821541
           eligible |  -11.19171   .4661828   -24.01   0.000    -12.10564   -10.27779
              _cons |   20.55449   .3376327    60.88   0.000     19.89258     21.2164
------------------------------------------------------------------------------------
```

Identification for Fuzzy Discontinuity

- Estimate regression

$$Y_i = \beta + \delta P_i + f(X_i) + \epsilon_i$$

where

$$P_i = \begin{cases} 1, & \text{if household } i \text{ participates in program} \\ 0, & \text{if household } i \text{ doesn't participate in program} \end{cases}$$

- But treatment depends on whether X_i (say, test score, as above) is $>$ or ≤ 50
  - In addition to endogenous factors

Identification for Fuzzy Discontinuity

- Under fuzzy RD, the eligibility threshold does not fully determine participation in the program
- Assume that the running variable is X_i, and that a threshold X_0 contributes to determine eligibility
- Under fuzzy discontinuity, the threshold only partially determines program participation
  - e.g. units above the threshold are more likely to participate in the program. However, there are units on both sides of the threshold that participate in the program
- The case of fuzzy discontinuity can be analyzed in an instrumental variable framework

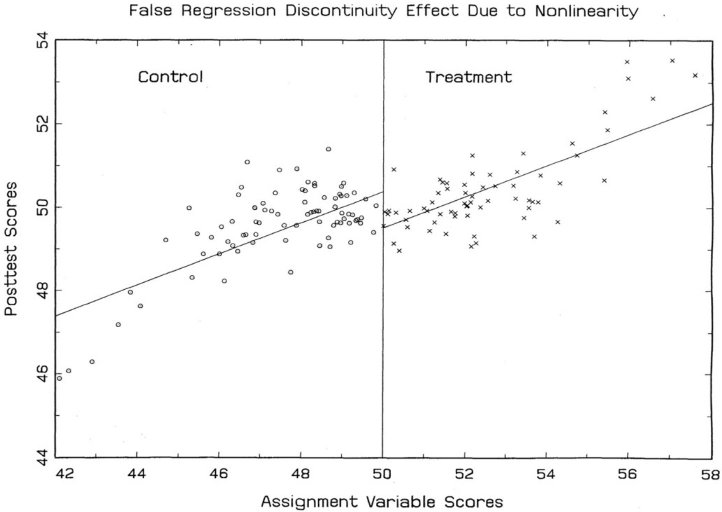
Identification for Fuzzy Discontinuity

- Recall sharp RD, $y_i = \beta_0 + \beta_1 P_i + \delta(score_i) + \epsilon_i$
- IV estimation is performed in two steps:
  1. $P_i = \gamma_0 + \gamma_1 \mathbb{1}(score_i > 50) + \eta_i$
  2. $y_i = \beta_0 + \beta_1 \hat{P}_i + \delta(score_i) + \epsilon_i$
  - Step 1: construct a dummy $1(X_i > X_0)$. This dummy strongly influences, but does not determine participation. It is used as an instrument for program participation
  - In words: *Program participation is influenced – but not entirely determined – by whether score is above 50. Thus estimate this relation.*
  - Step 2: use the predicted participation ($\hat{P}$) from Stage 1 to estimate the program impact at the threshold
- As before, this provides an estimate of local average treatment effects

Potential Disadvantages of RD

- RDD provides a **local** estimate of average treatment effects
  - We estimate the effect of the program around the cut-off point
  - This is not always generalizable, i.e. doesn't necessarily hold for points far away from the cutoff point
- Power
  - The effect is estimated at the discontinuity, so we generally have fewer observations than in a randomized experiment with the same sample size
- Specification can be sensitive to functional form
  - Make sure the relationship between the assignment variable and the outcome variable is correctly modeled, including:
    1. Nonlinear Relationships; and
    2. Interactions

Potential Disadvantages of RD



Advantages of RD

- RD yields an unbiased estimate of treatment effect **at the discontinuity**
- Can take advantage of a known rule for assigning the benefit
  - This is common in the design of social interventions.
  - No need to "exclude" a group of eligible households/ individuals from treatment.
  - In particular, it does not rely on **selection on observables** (like Matching, in a few weeks). The crucial **continuity assumption** is required, but it is often more plausible and partially testable.

Checklist

- RD requires that the eligibility index be continuous around the cut-off and that units be similar in the vicinity of the cut-off score.
  - Is the index continuous around the cut-off at the time of the baseline?
- Is there any evidence of non-compliance with the rule that determines eligibility for treatment?
  - Test whether all eligible and no-ineligible units have received the treatment. If you find non-compliance, you will need to combine RD with an IV approach to correct for this “fuzzy discontinuity”.
- Is there any evidence that index scores may have been manipulated in order to influence who qualified for the program?
  - Test whether the distribution of the index score is smooth at the cut-off point – this is the **McCrary (2008) density test**. If you find evidence of “bunching” of the index either above or below the cut-off, this might indicate manipulation.
- Is the cut-off unique to the program being evaluated, or is the cutoff used by other programs as well?

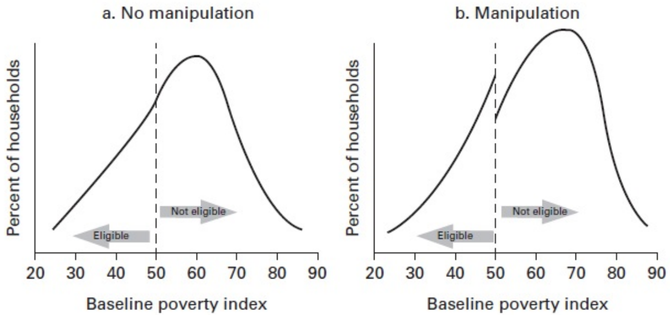
Manipulation of Running Variable

- For an RDD to yield an unbiased LATE estimate at the cutoff, it is important that the eligibility index not be manipulated around the cutoff so that an individual can change treatment or control status.
- Manipulation of the eligibility criteria can take many forms.
  - Enumerators who collect data that are used to compute the eligibility score could change one or two responses of respondents.
  - Respondents may purposefully lie to enumerators if they think that doing so would qualify them for the program.
  - Manipulation of the scores might get worse over time as enumerators, respondents, and politicians all start learning the "rules of the game."
    - e.g. farmer alters land title or misreports the size of their farms to be qualified as a small farmer.

Manipulation of Running Variable – Visual Diagnostic

The McCrary Density Test

Figure 6.4 Manipulation of the Eligibility Index



- **Example:** Refuse job offer, promotion, salary increase to remain below the threshold.

Overview

- Paper: Auffhammer and Kellogg (2011). Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality. American Economic Review
    - Research question: Effect of US gasoline content regulations on ozone pollution
    - Motivation: Ozone has been linked to asthma, increased susceptibility to pneumonia and bronchitis, and damage to crops and natural vegetation; even marginal short-term changes in ozone concentrations can have substantial human mortality impacts
    - Empirical strategy: Use introduction of regulations and an RD methodology
    - Results:
        - Flexible federal gasoline standards, which allow refiners in choosing a compliance mechanism, did not improve air quality (Intuition: refiners minimize costs, no impact on air quality)
        - Inflexible Californian regulations, which require the removal of particularly harmful compounds significantly improved air quality – reduction in ground-level ozone concentrations by 16 percent in the severely polluted Los Angeles-San Diego area

Institutional Background

- Policies: Restrictions on the chemical composition of gasoline primarily intended to reduce VOC emissions from mobile sources
  - Substantial geographic variation in regulation: EPA tightness varies across states/counties, some local governments have implemented their own more stringent regulations
- Key policies of interest
  - RVP: Reid vapor pressure regulation (# refers to the VP limit)
  - RFG: Federal reformulated gasoline
  - CARB RFG: California Air Resources Board reformulated gasoline

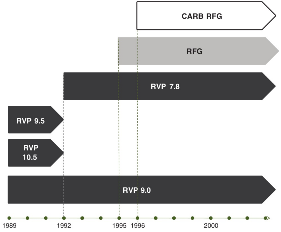


FIGURE 1. REGULATORY TIMELINE

Notes: RVP: Reid vapor pressure regulation (the number refers to the vapor pressure limit). RFG: Federal reformulated gasoline. CARB RFG: California Air Resources Board reformulated gasoline.

Institutional Background

- Cross-sectional variation in regulations (2006)

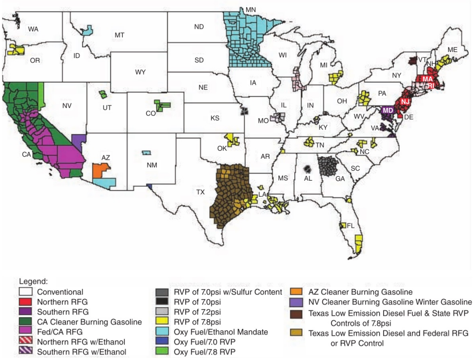


FIGURE 2. MAP OF RVP PHASE II AND RFG REGULATIONS AS OF 2006

Notes: Unshaded "conventional" gasoline areas are subject to the summertime RVP Phase II standard of 9.0 psi. Shaded areas in Minnesota, Colorado, Utah, and Montana have oxygenated gasoline for control of carbon monoxide pollution but do not have RVP or RFG. This study does not evaluate the effect of oxygenates on carbon monoxide pollution.

Source: EPA (Dec. 2006)

Data

- Data on ambient air concentrations of ozone from the EPA's Air Quality Standards database for 1989–2003
    - Hourly readings from the EPA's network of air quality monitors
    - Construct the daily maximum concentration and the daily eight-hour maximum
- Weather data measurements from the National Climatic Data Center's Cooperative Station Data (NOAA 2008), which provide daily minimum and maximum temperatures, rain, and snowfall. To assign to each air monitoring station the corresponding weather data...
    1. Identify the ten closest weather stations to each pollution monitor, provided that each is less than 50 miles from the monitor and the elevation difference between the monitor and the station is less than 500 vertical feet
    2. Among these, identify the closest station for which 50 percent of the pollution monitor's daily readings can be matched to the station's weather data. We then match the four climate variables for this station to the time series of ozone measurements

Empirical Strategy

- In the RD approach, identification of the regulations' effects comes from the change in ozone concentration **within a narrow window around the phase-in of each regulation**. Why?
  1. Because imposition of a gasoline standard affects all cars simultaneously, implying that the standard will cause a step change in emissions almost immediately after implementation; and
  2. Because ozone decomposes overnight, meaning that daily maximum ozone concentrations will respond quickly to changes in emissions
- RD permits an identification assumption that is less restrictive than that of the DD model (which we will study later)
  - While identification of the DD estimator requires that unobserved variables affecting ozone concentrations do so only through a *linear* time trend, the RD model permits unobserved factors to act *nonlinearly over time*, as long as they are not discontinuous when gasoline regulations phase in

Empirical Specification

- RD specification takes the form

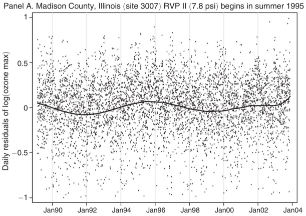
$$ln(y_{it}) = \alpha_i \times Treat_{ct} + \beta_i \boldsymbol{W_{it}} + f_i(Date_t) + \mu_i + \varepsilon_{it}$$

where

- y_{it} is a measure of air pollution at monitor i on date t
- $Treat_{ct}$ is a vector of four variables indicating whether the county c in which monitor i is located is subject to one of four possible regulatory treatments at time t (excluding the baseline RVP standard)
- the treatment effects α_i and the coefficients β_i on the weather variables are monitor-specific
- $f_i(Date_t)$ is an eighth-order Chebychev polynomial in time that is also monitor-specific
- monitor FEs (μ_i) are also included

Descriptives

- Panels B and C exhibit a change in the smooth pollution pattern – and a local jump – after the policy change



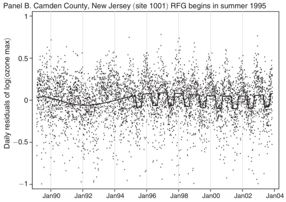


FIGURE 6. DAILY MAXIMUM OZONE CONCENTRATIONS (C

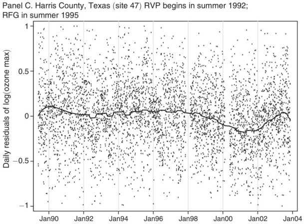


FIGURE 6. DAILY MAXIMUM OZONE CONCENTRATIONS (*Concluded*)

Notes: Fitted lines are the predicted values of the treatment dummies and eighth-order polynomial time trend from estimating equation (3), centered at zero. Points plotted are the residuals from (3) plus the fitted line.

Main Results

- Tables are displayed in the Appendix and emphasize results at the station-level
- In what follows, focus on graphical part

Main Results

- **Aim:** Show continuity of running variable

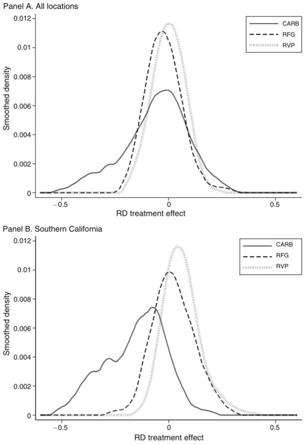


Figure 7. Distribution of Estimated RD Treatment Effects by Policy Type

Notes: The figures display the smoothed cross-monitor distribution of estimated treatment effects from equation (3). The smoother uses an Epanechnikov kernel with a bandwidth of 0.05. The RVP and RFG results shown in panel A do not include California. Panel B includes only Los Angeles, Orange, Riverside, San Bernardino, San Diego, and Ventura counties. These are the only counties in California in which RFG was ever implemented.

Main Results

- **Aim:** Show jump at dicontinuity/cutt-off point

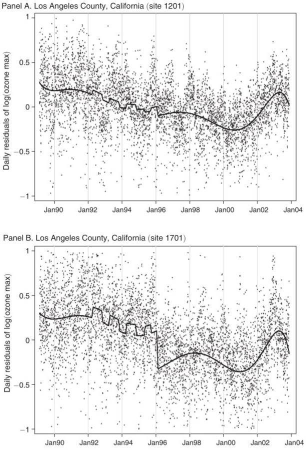


FIGURE 8. DAILY MAXIMUM OZONE CONCENTRATIONS

Notes: Fitted lines are the predicted values of the treatment dummies and eighth-order polynomial time trend from estimating equation (3), centered at zero. Points plotted are the residuals from (3) plus the fitted line. Regulations include RVP (summer 1992–1994), RFG (summer 1995), and CARB (March 1996–).

Wrapping-up

- The bar in terms of execution (and robustness) is much higher nowadays than when the paper was published
  - e.g., $p - th$ order polynomials (choice of p always raising suspicions) are replaced with nonparametric local linear estimator
  - e.g., methods to determine optimal window around the discontinuity (although researchers often impose/justify a choice of window)
  - Libraries available in R and Stata
- Nevertheless, an interesting an important paper!
- For those interested in environmental/energy topics: Have a look at the data work...

Overview

- Paper: Lechler and McNamee (2017). "Decentralized despotism? Indirect colonial rule undermines contemporary democratic attitudes".
  - Paper identifies indirect and direct colonial rule as causal factors in shaping support for democracy by exploiting a within-country natural experiment in Namibia
  - Throughout the colonial era, northern Namibia was indirectly ruled through a system of appointed indigenous traditional elites whereas colonial authorities directly ruled southern Namibia.
    - This variation originally stems from where the progressive extension of direct German control was stopped after a rinderpest epidemic in the 1890s, and thus constitutes plausibly exogenous within-country variation in the form of colonial rule
  - Using spatial RDD, find that individuals in indirectly ruled areas are less likely to support democracy and turnout at elections.
  - Findings suggest that that the greater influence of traditional leaders in indirectly ruled areas has socialized individuals to accept non-electoral bases of political authority

Historical Background

- Namibia became a German protectorate in 1884
  - Different ethnic groups with similar political structures
  - Initial focus on S and CE coastal regions which they reached first and where land could be more easily acquired
  - Expansion from the coast by playing off warring local factions and remunerating a number of indigenous elites in CE Namibia for lost landholdings
  - 1897: critical event – a rinderpest epidemic killed 95 percent of the cattle herds in CE and S Namibia
  - Strongest impact on cattle-dependent CE and S Namibia – unlike in N Namibia, the arid land prevented the development of agriculture
  - Result: opportunity to acquire large tracts of land in CE and S Namibia relatively cheaply with lessened collective resistance from weakened indigenous communities
  - Moreover: Government set up a veterinary cordon fence at the boundaries of where at the time its direct control extended in order to protect S and CE cattle herds from rinderpest-infected animals from the N

Hypotheses

- H1: Individuals in indirectly ruled areas are less likely to support democracy as a system of government
- H2: Individuals in indirectly ruled areas are less likely to turnout at elections
- "Our theoretical framework moreover predicts that this relationship is likely being driven by greater contact to traditional leaders and greater respect for authorities in indirectly ruled areas. Thus, whilst we primarily focus on support for democracy as our outcome of interest, we will also test the following secondary hypotheses:"
- H3: Individuals in indirectly ruled areas are more likely to contact traditional authorities
- H4: Individuals in indirectly ruled areas are less likely to support questioning authority

Motivation for RDD: Space

- Cut-off: veterinary cordon fence, which became a Police Zone boundary – divides area directly settled/ruled by Germans authorities vs. area indirectly ruled by system of indigenous elites
- Trade and the permanent movement of people was restricted by the German authorities, and indigenous political structures within the Police Zone were destroyed.

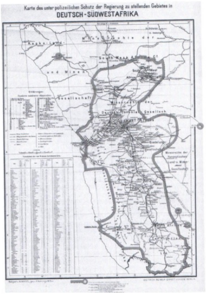


Figure 1: Map of 1907

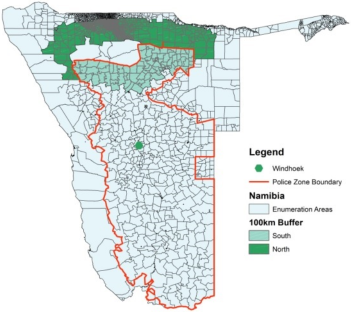


Figure 2: Enumeration areas and buffer

Empirical Strategy

- Estimate

$$Y_{idres} = \beta_0 + \beta_1 Indirectrule_d + X'_{ides}\Gamma + \eta_e + \mu_s + \psi_r + \varepsilon_{idres}$$

  - Y: demand for democracy of individual i, living in enumeration area d in region r, belonging to the ethnic group e, being surveyed in round s
  - $Indirectrule$: dummy variable of whether individual lives in an enumeration area which belonged to the indirectly or the directly ruled part of Namibia
  - X : control variables, which includes individual-level characteristics such as age and dummies for income and education as well as distance to Windhoek (capital)
  - η_e, μ_s, ψ_r : ethnicity, survey-round, region FEs

- Paper details the use of a 100km buffer zone around the plausibly exogenous boundary between the two zones (see Figure 2) due to changes in border over time

Checking Balance

- The **Balance Table** (Table 1) is a key validity check for the continuity assumption. It shows that other covariates (like Education, Age) are smooth across the boundary (no significant difference). If "Age" also jumped at the cutoff, we couldn't be sure if our outcome jump was due to the "Indirect rule" or the difference in "Age".
- Authors nevertheless add individual-level controls to all specifications as they are also important determinants of political attitudes (Bratton et al., 2005) and help to identify the effects more precisely
- Also use survey questions as proxies for income and education and age (interesting – see paper)

Table 1: Balancing table for the buffer zone

	(1)	(2)	(3)	(4)
VARIABLES	Without food	Education	Gender	Age
Indirect colonial rule	0.0398	-0.274	0.0232	1.606
	(0.163)	(0.193)	(0.0306)	(1.043)
Observations	1,417	1,406	1,060	1,413
Ethnicity FE	yes	yes	yes	yes
Survey round FE	yes	yes	yes	yes
Mean of DV	1.140	3.814	0.490	35.82

Results from OLS regressions. The sample consists of observations from the 100km buffer zone. Standard errors (clustered by Enumeration Area) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Results: Indirect Rule and Support for Democracy

- Living in the formerly indirectly ruled part of Namibia decreases support for democracy and probability of voting
  - Columns (2), (5): preferred RDD specification including region FEs, individual level controls and distance to Windhoek
  - Columns (3), (6): results of (ordered) probit model are robust

Table 2: Effect of indirect rule on support for democracy and voting

VARIABLES	(1) Support democracy OLS	(2) Support democracy OLS	(3) Support democracy O Probit	(4) Voting OLS	(5) Voting OLS	(6) Voting Probit
Indirect colonial rule	-0.178**	-0.223*	-0.357*	-0.122***	-0.166**	-0.590*
	(0.0746)	(0.133)	(0.198)	(0.0409)	(0.0824)	(0.305)
Distance to Windhoek		-0.0228	-0.0142		-0.0389	-0.0919
		(0.0790)	(0.118)		(0.0574)	(0.223)
Observations	1,347	1,329	1,329	734	723	721
R^2	0.019	0.043		0.049	0.287	
Ethnicity FE	yes	yes	yes	yes	yes	yes
Survey round FE	yes	yes	yes	yes	yes	yes
Region FE	no	yes	yes	no	yes	yes
Individual-level controls	no	yes	yes	no	yes	yes
# clusters	165	165	165	91	91	91
Mean of DV	2.399	2.401	2.401	0.722	0.719	0.718

Results from OLS regressions. Individual-level control variables are age, education dummies and income dummies. The sample consists of observations from the 100km buffer zone. Standard errors (clustered by Enumeration Area) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Robustness: RD Polynomial

- Results are robust to quadratic polynomials

Table 4: Different specifications of RD polynomial

VARIABLES	(1) Support democracy	(2) Support democracy	(3) Support democracy	(4) Voting	(5) Voting	(6) Voting
Indirect colonial rule	-0.208*	-0.246*	-0.259***	-0.192**	-0.127	-0.216***
	(0.111)	(0.127)	(0.0956)	(0.0952)	(0.0815)	(0.0496)
Observations	1,347	1,347	1,347	734	734	734
R^2	0.022	0.029	0.029	0.052	0.063	0.060
Lat/Lon	yes	no	no	yes	no	no
Dist. Windhoek quadr	no	yes	no	no	yes	no
Dist. Boundary quadr	no	no	yes	no	no	yes
Ethnicity FE	yes	yes	yes	yes	yes	yes
Survey round FE	yes	yes	yes	yes	yes	yes
# clusters	165	165	165	91	91	91
Mean of DV	2.399	2.399	2.399	0.722	0.722	0.722

Results from OLS regressions. Columns (1), and (4) include a local linear polynomials in Longitude and Latitude. Columns (2), and (5) include a quadratic polynomial in distance to Windhoek. Columns (3), and (6) include a quadratic polynomial in distance to the boundary. The sample consists of observations from the 100km buffer zone. Standard errors (clustered by Enumeration Area) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Robustness: Additional Controls

- Again, results largely robust

Table 5: Effect of indirect rule on support for democracy

VARIABLES	(1) Support democracy	(2) Support democracy	(3) Support democracy	(4) Support democracy	(5) Support democracy
Indirect colonial rule	-0.223*	-0.313**	-0.237*	-0.222*	-0.320**
	(0.133)	(0.141)	(0.135)	(0.126)	(0.147)
Distance to Windhoek	-0.0228	0.00163	-0.0212	-0.0185	0.00273
	(0.0790)	(0.0821)	(0.0792)	(0.0761)	(0.0830)
Performance government		-0.00667			-0.00652
		(0.0288)			(0.0289)
Livestock suitability			-0.0197		-0.0414
			(0.0363)		(0.0421)
Urban				-0.00178	-0.0486
				(0.0713)	(0.0865)
Observations	1,329	1,274	1,329	1,334	1,274
R^2	0.043	0.043	0.043	0.046	0.044
Ethnicity FE	yes	yes	yes	yes	yes
Survey round FE	yes	yes	yes	yes	yes
Region FE	yes	yes	yes	yes	yes
Individual-level controls	yes	yes	yes	yes	yes
# clusters	165	165	165	165	165
Mean of DV	2.401	2.399	2.401	2.397	2.399

Results from OLS regressions. Individual-level control variables are age, education dummies and income dummies. The sample consists of observations from the 100km buffer zone. Standard errors (clustered by Enumeration Area) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Effect of indirect colonial rule on voting

VARIABLES	(1) Voting	(2) Voting	(3) Voting	(4) Voting	(5) Voting
Indirect colonial rule	-0.166**	-0.212**	-0.173*	-0.181*	-0.224**
	(0.0824)	(0.0872)	(0.0888)	(0.0920)	(0.0970)
Distance to Windhoek	-0.0389	-0.00556	-0.0371	-0.0367	-0.00636
	(0.0574)	(0.0564)	(0.0592)	(0.0586)	(0.0572)
Performance government		0.0137			0.0125
		(0.0223)			(0.0226)
Livestock suitability			-0.00430		0.0121
			(0.0194)		(0.0304)
Urban				0.0247	0.0494
				(0.0418)	(0.0598)
Observations	723	687	723	723	687
R^2	0.287	0.285	0.287	0.287	0.286
Ethnicity FE	yes	yes	yes	yes	yes
Survey round FE	yes	yes	yes	yes	yes
Region FE	yes	yes	yes	yes	yes
Individual-level controls	yes	yes	yes	yes	yes
# clusters	91	91	91	91	91
Mean of DV	0.719	0.721	0.719	0.719	0.721

Results from OLS regressions. Individual-level control variables are age, education dummies and income dummies. The sample consists of observations from the 100km buffer zone. Standard errors (clustered by Enumeration Area) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Discussion of Mechanism

- Evidence suggests that the mechanisms are not demographic factors such as education or income but rather are institutional – specifically, the institution of traditional leadership
- Authors argue that the ongoing local role that traditional authorities play in formerly indirectly ruled areas of Namibia acts as a parallel undemocratic hierarchical governance structure and socializes individuals in indirectly ruled areas to accept non-electoral bases of legitimacy
- Paper examines mechanism quantitatively, in addition to other details

Take-aways

- Regression discontinuity design (RDD) is an impact evaluation method that is adequate for programs
  - that use a continuous index to rank potential participants; and
  - that have a cutoff point along the index that determines whether or not potential participants are eligible to receive the program

References

- Gertler et al (2016). Impact Evaluation in Practice, 2nd. Edition. Washington, DC: Inter-American Development Bank and World Bank
  - Chapter 6
- Gertler et al (2016). Impact Evaluation in Practice, 2nd. Edition, Technical Companion (Version 1.0). Washington, DC: Inter-American Development Bank and World Bank.
  - p. 26-28
- Bursztyn and Cantoni (2016). A Tear in the Iron Curtain: The Impact of western Television on Consumption Behavior. *The Review of Economics and Statistics*, 98(1): 25-41.
- Lechler and McNamee (2017). Decentralized despotism? Indirect colonial rule undermines contemporary democratic attitudes. Munich Discussion Paper, No. 2017-7.

References

Surveys: (with many empirical insights, guidance)

- Imbens and Lemieux (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics* 142 (2): 615-635.
- Lee and Lemieux (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature* 48 (2): 281-355.

Density test:

- McCrary (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics* 142 (2): 698-714.