Lecture 15: Regression discontinuity II

PPHA 34600

Prof. Fiona Burlig

Harris School of Public Policy
University of Chicago

From last time: regression discontinuity

As usual, we'd like to run:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

The regression discontinuity:

- Suppose D_i is determined by whether or not X_i lies above a cutoff, c
- We call X_i the "running variable" here
- **Idea**: Having X_i just above or just below c is as good as random...
- ... And there is a discontinuous change in D_i as a result of crossing c

→ We can compare Y_i for units with X_i just above c to Y_i for units with X_i just below c

From last time: regression discontinuity

As usual, we'd like to run:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

The regression discontinuity:

- Suppose D_i is determined by whether or not X_i lies above a cutoff, c
- We call X_i the "running variable" here
- **Idea**: Having X_i just above or just below c is as good as random...
- ... And there is a discontinuous change in D_i as a result of crossing c

→ We can compare Y_i for units with X_i just above c to Y_i for units with X_i just below c

To estimate, run:

$$Y_i = \alpha + \tau D_i + f(X_i) + \varepsilon_i \text{ for } c - h < X_i < c + h$$

where $D_i = \mathbf{1}[X_i \geq c]$

Sharp regression discontinuity

In the most straightforward, or "sharp" RD design:

- $Pr(D_i = 1|X_i \geq c) = 1$ and $Pr(D_i = 1|X_i < c) = 0$
- $Pr(D_i = 1|X_i \geq c) - Pr(D_i = 1|X_i < c) = 1$
- Nobody with $X_i < c$ gets treated
- Everybody with $X_i \geq c$ gets treated
- The probability of treatment jumps from 0 to 100% as X_i crosses c
- $D_i = \mathbf{1}(X_i \geq c)$

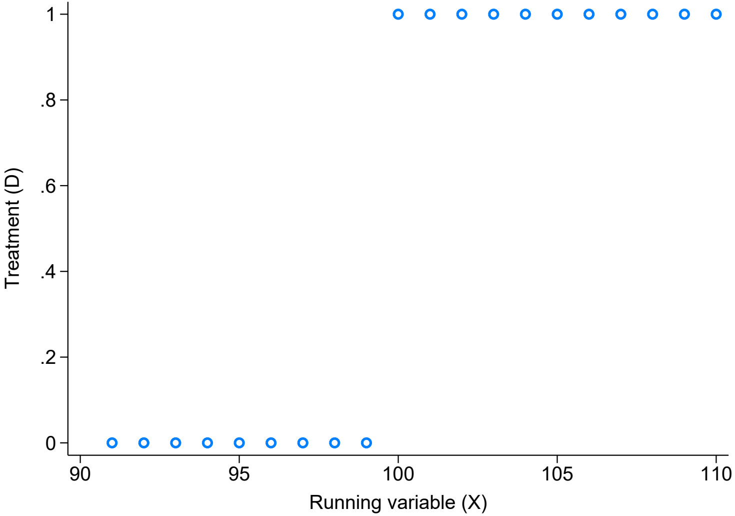
Sharp regression discontinuity

In the most straightforward, or “sharp” RD design:

- $Pr(D_i = 1|X_i \geq c) = 1$ and $Pr(D_i = 1|X_i < c) = 0$
- $Pr(D_i = 1|X_i \geq c) - Pr(D_i = 1|X_i < c) = 1$
- Nobody with $X_i < c$ gets treated
- Everybody with $X_i \geq c$ gets treated
- The probability of treatment jumps from 0 to 100% as X_i crosses c
- $D_i = \mathbf{1}(X_i \geq c)$

→ This is equivalent to **perfect compliance** in the RCT

Sharp regression discontinuity: Treatment assignment



Fuzzy regression discontinuity

Sometimes, treatment status doesn't change by 100% at the cutoff:

- In this case, we implement the **fuzzy regression discontinuity**

Fuzzy regression discontinuity

Sometimes, treatment status doesn't change by 100% at the cutoff:

- In this case, we implement the **fuzzy regression discontinuity**
- Formally:

$$0 < \lim_{x \downarrow c} Pr(D_i = 1 | X_i = x) - \lim_{x \uparrow c} Pr(D_i = 1 | X_i = x) < 1$$

Fuzzy regression discontinuity

Sometimes, treatment status doesn't change by 100% at the cutoff:

- In this case, we implement the **fuzzy regression discontinuity**
- Formally:

$$0 < \lim_{x \downarrow c} Pr(D_i = 1 | X_i = x) - \lim_{x \uparrow c} Pr(D_i = 1 | X_i = x) < 1$$

This implies:

$$Pr(D_i = 1 | X_i \geq c) - Pr(D_i = 1 | X_i < c) = k, \text{ where } 0 < k < 1$$

Fuzzy regression discontinuity

Sometimes, treatment status doesn't change by 100% at the cutoff:

- In this case, we implement the **fuzzy regression discontinuity**
- Formally:

$$0 < \lim_{x \downarrow c} Pr(D_i = 1 | X_i = x) - \lim_{x \uparrow c} Pr(D_i = 1 | X_i = x) < 1$$

  This implies:

$$Pr(D_i = 1 | X_i \geq c) - Pr(D_i = 1 | X_i < c) = k, \text{ where } 0 < k < 1$$

- Some units with $X_i < c$ (may) get treated; more with $X_i \geq c$ do
- Crossing c leads to a change in the *probability* of treatment
- Treatment probability jumps from $0 \leq p \leq 1$ to $0 \leq q \leq 1$ $(q > p)$
- D_i is no longer a deterministic function of X_i

Fuzzy regression discontinuity

Sometimes, treatment status doesn't change by 100% at the cutoff:

- In this case, we implement the **fuzzy regression discontinuity**
- Formally:

$$0 < \lim_{x \downarrow c} Pr(D_i = 1 | X_i = x) - \lim_{x \uparrow c} Pr(D_i = 1 | X_i = x) < 1$$

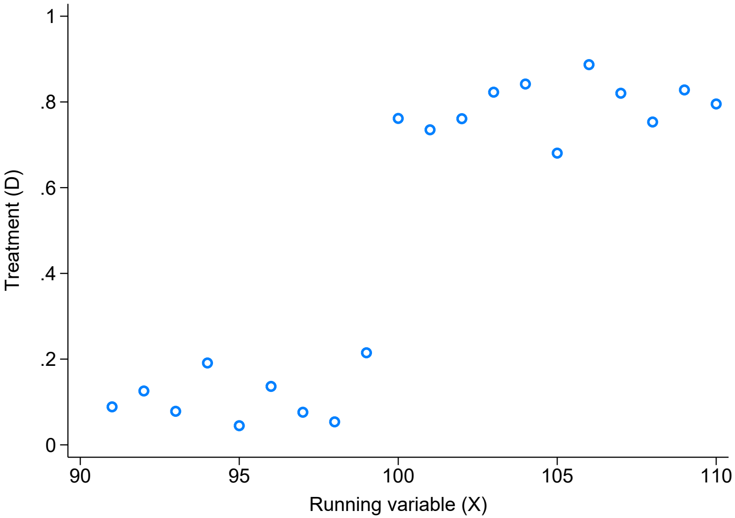
  This implies:

$$Pr(D_i = 1 | X_i \geq c) - Pr(D_i = 1 | X_i < c) = k, \text{ where } 0 < k < 1$$

- Some units with $X_i < c$ (may) get treated; more with $X_i \geq c$ do
- Crossing c leads to a change in the *probability* of treatment
- Treatment probability jumps from $0 \leq p \leq 1$ to $0 \leq q \leq 1$ $(q > p)$
- D_i is no longer a deterministic function of X_i

→ This is equivalent to **imperfect compliance** in the RCT

Fuzzy regression discontinuity: Treatment assignment



Estimating τ with a fuzzy RD

We need to account for the incomplete change in D_i:

- To do this, we estimate two objects:

1. Effect of going from $X_i < c$ to $X_i \geq c$ on our outcome Y_i
   - This will be too close to zero, because there is some noncompliance
   - We call this the **reduced form**

Estimating τ with a fuzzy RD

We need to account for the incomplete change in D_i:

- To do this, we estimate two objects:

1. Effect of going from $X_i < c$ to $X_i \geq c$ on our outcome Y_i
   - This will be too close to zero, because there is some noncompliance
   - We call this the **reduced form**
2. Effect of going from $X_i < c$ to $X_i \geq c$ on treatment status D_i
   - This allows us to quantify the non-compliance
   - We call this the **first stage**

Estimating τ with a fuzzy RD

We need to account for the incomplete change in D_i:

- To do this, we estimate two objects:

1. Effect of going from $X_i < c$ to $X_i \geq c$ on our outcome Y_i
   - This will be too close to zero, because there is some noncompliance
   - We call this the **reduced form**
2. Effect of going from $X_i < c$ to $X_i \geq c$ on treatment status D_i
   - This allows us to quantify the non-compliance
   - We call this the **first stage**

→ This should be feeling familiar...

Estimating the reduced form

The effect of crossing the threshold on outcomes is just:

$$\theta = \lim_{x \downarrow c} E[Y_i | X_i = x] - \lim_{x \uparrow c} E[Y_i | X_i = x]$$

Estimating the reduced form

The effect of crossing the threshold on outcomes is just:

$$\theta = \lim_{x \downarrow c} E[Y_i | X_i = x] - \lim_{x \uparrow c} E[Y_i | X_i = x]$$

We can just estimate this as:

$$\hat{\theta} = \bar{Y}(c \leq X_i \leq c + h) - \bar{Y}(c - h \leq X_i < c)$$

Estimating the reduced form

The effect of crossing the threshold on outcomes is just:

$$\theta = \lim_{x \downarrow c} E[Y_i | X_i = x] - \lim_{x \uparrow c} E[Y_i | X_i = x]$$

We can just estimate this as:

$$\hat{\theta} = \bar{Y}(c \leq X_i \leq c + h) - \bar{Y}(c - h \leq X_i < c)$$

Via regression:

$$Y_i = \alpha + \theta \mathbf{1}[X_i \geq c] + \nu_i \text{ for } c - h \leq X_i \leq c + h$$

Note that as before, $\hat{\theta} = \theta$ at the threshold only

Estimating the first stage

The effect of crossing the threshold on treatment status is just:

$$\gamma = \lim_{x \downarrow c} E[D_i | X_i = x] - \lim_{x \uparrow c} E[D_i | X_i = x]$$

Estimating the first stage

The effect of crossing the threshold on treatment status is just:

$$\gamma = \lim_{x \downarrow c} E[D_i | X_i = x] - \lim_{x \uparrow c} E[D_i | X_i = x]$$

We can just estimate this as:

$$\hat{\gamma} = \bar{D}(c \leq X_i \leq c + h) - \bar{D}(c - h \leq X_i < c)$$

Estimating the first stage

The effect of crossing the threshold on treatment status is just:

$$\gamma = \lim_{x \downarrow c} E[D_i | X_i = x] - \lim_{x \uparrow c} E[D_i | X_i = x]$$

We can just estimate this as:

$$\hat{\gamma} = \bar{D}(c \leq X_i \leq c + h) - \bar{D}(c - h \leq X_i < c)$$

Via regression:

$$D_i = \alpha + \gamma \mathbf{1}[X_i \geq c] + \eta_i \text{ for } c - h \leq X_i \leq c + h$$

Estimating the first stage

The effect of crossing the threshold on treatment status is just:

$$\gamma = \lim_{x \downarrow c} E[D_i | X_i = x] - \lim_{x \uparrow c} E[D_i | X_i = x]$$

We can just estimate this as:

$$\hat{\gamma} = \bar{D}(c \leq X_i \leq c + h) - \bar{D}(c - h \leq X_i < c)$$

Via regression:

$$D_i = \alpha + \gamma \mathbf{1}[X_i \geq c] + \eta_i \text{ for } c - h \leq X_i \leq c + h$$

→ In the sharp RD, $\gamma = 1$

→ $\hat{\gamma}$ estimates the change in probability of treatment from crossing c

Putting the pieces together

The fuzzy RD estimator gets you:

$$\tau^{FRD} = \frac{\lim_{x \downarrow c} E[Y_i | X_i = x] - \lim_{x \uparrow c} E[Y_i | X_i = x]}{\lim_{x \downarrow c} E[D_i | X_i = x] - \lim_{x \uparrow c} E[D_i | X_i = x]}$$

Putting the pieces together

The fuzzy RD estimator gets you:

$$\tau^{FRD} = \frac{\lim_{x \downarrow c} E[Y_i|X_i = x] - \lim_{x \uparrow c} E[Y_i|X_i = x]}{\lim_{x \downarrow c} E[D_i|X_i = x] - \lim_{x \uparrow c} E[D_i|X_i = x]}$$

We can estimate this as:

$$\hat{\tau}^{FRD} = \frac{\bar{Y}(c \leq X_i \leq c + h) - \bar{Y}(c - h \leq X_i < c)}{\bar{D}(c \leq X_i \leq c + h) - \bar{D}(c - h \leq X_i < c)} = \frac{\hat{\theta}}{\hat{\gamma}}$$

Putting the pieces together

The fuzzy RD estimator gets you:

$$\tau^{FRD} = \frac{\lim_{x \downarrow c} E[Y_i | X_i = x] - \lim_{x \uparrow c} E[Y_i | X_i = x]}{\lim_{x \downarrow c} E[D_i | X_i = x] - \lim_{x \uparrow c} E[D_i | X_i = x]}$$

We can estimate this as:

$$\hat{\tau}^{FRD} = \frac{\bar{Y}(c \leq X_i \leq c + h) - \bar{Y}(c - h \leq X_i < c)}{\bar{D}(c \leq X_i \leq c + h) - \bar{D}(c - h \leq X_i < c)} = \frac{\hat{\theta}}{\hat{\gamma}}$$

This is just an IV estimator, where $Z_i = \mathbf{1}[X_i \geq c]$!

Fuzzy RD as IV

Since the fuzzy RD is just an IV, we need the standard IV assumptions:

1. First stage: $E[D_i|X_i \geq c] \neq E[D_i|X_i < c]$ for some i
2. Independence: $Y_i(D_i, \mathbf{1}[X_i \geq c]), D_i(X_i \geq c), D_i(X_i < c) \perp \mathbf{1}[X_i \geq c]$
3. Exclusion restriction: $Y_i(X_i \geq c, D_i) = Y_i(X_i < c, D_i)$ for $D_i \in \{0, 1\}$
4. Monotonicity: $|D_i(X_i \geq c) - D_i(X_i < c)| \geq 0$ for all i

Fuzzy RD as IV

Since the fuzzy RD is just an IV, we need the standard IV assumptions:

1. First stage: $E[D_i|X_i \geq c] \neq E[D_i|X_i < c]$ for some i
2. Independence: $Y_i(D_i, \mathbf{1}[X_i \geq c]), D_i(X_i \geq c), D_i(X_i < c) \perp \mathbf{1}[X_i \geq c]$
3. Exclusion restriction: $Y_i(X_i \geq c, D_i) = Y_i(X_i < c, D_i)$ for $D_i \in \{0, 1\}$
4. Monotonicity: $|D_i(X_i \geq c) - D_i(X_i < c)| \geq 0$ for all i

Assumptions (2) and (3) together buy us:

1. Covariate smoothness:
   $E[Y_i(1)|X_i = x]$ and $E[Y_i(0)|X_i = x]$ are continuous in x

Fuzzy RD as IV

Since the fuzzy RD is just an IV, we need the standard IV assumptions:

1. First stage: $E[D_i|X_i \geq c] \neq E[D_i|X_i < c]$ for some i
2. Independence: $Y_i(D_i, \mathbf{1}[X_i \geq c]), D_i(X_i \geq c), D_i(X_i < c) \perp \mathbf{1}[X_i \geq c]$
3. Exclusion restriction: $Y_i(X_i \geq c, D_i) = Y_i(X_i < c, D_i)$ for $D_i \in \{0, 1\}$
4. Monotonicity: $|D_i(X_i \geq c) - D_i(X_i < c)| \geq 0$ for all i

Assumptions (2) and (3) together buy us:

1. Covariate smoothness:
   $E[Y_i(1)|X_i = x]$ and $E[Y_i(0)|X_i = x]$ are continuous in x

With these assumptions, we get $\tau^{FRD} = \tau^{LATE}$

Fuzzy RD: estimation methods

As with any other IV estimator, we can estimate fuzzy RD via 2SLS:

1. **First stage:**

$$D_i = \alpha + \gamma \mathbf{1}[X_i \geq c] + \nu_i$$

2. **Second stage:**

$$Y_i = \alpha + \tau \hat{D}_i + \varepsilon_i$$

Fuzzy RD: estimation methods

As with any other IV estimator, we can estimate fuzzy RD via 2SLS:

1. **First stage:**

$$D_i = \alpha + \gamma \mathbf{1}[X_i \geq c] + \nu_i$$

2. **Second stage:**

$$Y_i = \alpha + \tau \hat{D}_i + \varepsilon_i$$

We can also estimate fuzzy RD using the first stage and reduced form:

1. **First stage:**

$$D_i = \alpha + \gamma \mathbf{1}[X_i \geq c] + \nu_i$$

2. **Reduced form:**

$$Y_i = \alpha + \theta \mathbf{1}[X_i \geq c] + \varepsilon_i$$

$$\hat{\tau}^{FRD} = \frac{\hat{\theta}}{\hat{\gamma}}$$

Fuzzy RD as IV

What LATE are we estimating with fuzzy RD?

- Just like with the sharp RD, we get the LATE at c
  - Different cutoffs c could generate different LATEs

Fuzzy RD as IV

What LATE are we estimating with fuzzy RD?

- Just like with the sharp RD, we get the LATE at c
  - Different cutoffs c could generate different LATEs
- In addition, we have to worry about **compliers**
  - Not all units go from $D_i = 0$ to $D_i = 1$ as X_i crosses c (this generates the fuzz)
  - We recover treatment effects for only those units who do move

Fuzzy RD as IV

What LATE are we estimating with fuzzy RD?

- Just like with the sharp RD, we get the LATE at c
    - Different cutoffs c could generate different LATEs
- In addition, we have to worry about **compliers**
    - Not all units go from $D_i = 0$ to $D_i = 1$ as X_i crosses c (this generates the fuzz)
    - We recover treatment effects for only those units who do move

$\rightarrow$ We get the LATE for compliers at the threshold

$\rightarrow$ Changing the set of compliers or the threshold (or both) could generate different LATEs

An example: Air pollution in China

Policy issue:

- Local air pollution (PM, SO_x, NO_x) is likely bad for human health
- But how bad?

Approach:

- Look at the Huai River heating policy in China
- Households north of the river got free coal
- This allowed them to heat their houses...
- ...but also led to substantial air pollution:

→ Use a RD model to compare northern to southern cities

What do these Chinese cities look like?

Table 1. Summary statistics

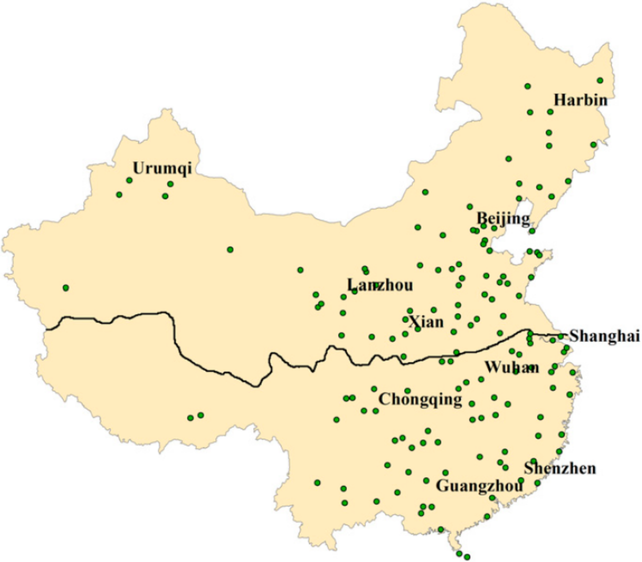
Variable	South (1)	North (2)	Difference in means (3)	Adjusted difference in means (4)	*P* value (5)
Panel 1: Air pollution exposure at China's Disease Surveillance Points					
TSPs, μg/m^3	354.7	551.6	196.8***	199.5***	<0.001/0.002
SO_2, μg/m^3	91.2	94.5	3.4	−3.1	0.812/0.903
NO_x, μg/m^3	37.9	50.2	12.3***	−4.3	<0.001/0.468
Panel 2: Climate at the Disease Surveillance Points					
Heating degree days	2,876	6,220	3,344***	482	<0.001/0.262
Cooling degree days	2,050	1,141	−910***	−183	<0.001/0.371
Panel 3: Demographic features of China's Disease Surveillance Points					
Years of education	7.23	7.57	0.34	−0.65	0.187/0.171
Share in manufacturing	0.14	0.11	−0.03	−0.15***	0.202/0.002
Share minority	0.11	0.05	−0.05	0.04	0.132/0.443
Share urban	0.42	0.42	0.00	−0.20*	0.999/0.088
Share tap water	0.50	0.51	0.02	−0.32**	0.821/0.035
Rural, poor	0.21	0.23	0.01	−0.33*	0.879/0.09
Rural, average income	0.34	0.33	0.00	0.24	0.979/0.308
Rural, high income	0.21	0.19	−0.02	0.27	0.772/0.141
Urban site	0.24	0.25	0.01	−0.19	0.859/0.241
Predicted life expectancy	74.0	75.5	1.54***	−0.24	<0.001/0.811
Actual life expectancy	74.0	75.5	1.55	−5.04**	0.158/0.044

What do we get from the naive estimator?

Table 2. Impact of TSPs (100 μg/m^3) on health outcomes using conventional strategy (ordinary least squares)

Dependent variable	(1)	(2)
ln(All cause mortality rate)	0.03* (0.01)	0.03** (0.01)
ln(Cardiorespiratory mortality rate)	0.04** (0.02)	0.04** (0.02)
ln(Noncardiorespiratory mortality rate)	0.01 (0.02)	0.01 (0.02)
Life expectancy, y	−0.54** (0.26)	−0.52** (0.23)
Climate controls	No	Yes
Census and DSP controls	No	Yes

Enter the regression discontinuity



Estimating the effects of air pollution on mortality

The authors use 2SLS to estimate a fuzzy regression discontinuity model:

$$Pollution_i = \alpha + \gamma \mathbf{1}[Latitude_i \leq river] + f(Latitude_i) + \nu_i$$

$$Y_i = \alpha + \tau \widehat{Pollution}_i + f(Latitude_i) + \varepsilon_i$$

Where:

$Pollution_i$ is a measure of the total suspended particulates in city i

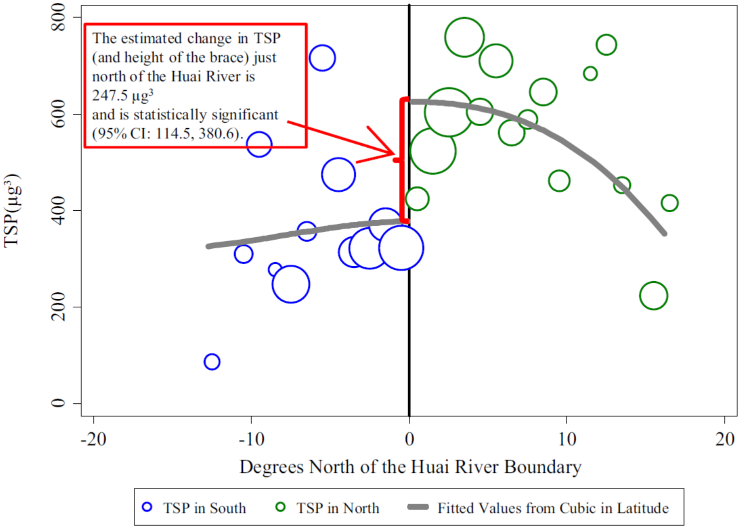
Y_i is life expectancy in city i

$\mathbf{1}[Latitude_i \leq river]$ is equal to one if city i is north of the river

$f(Latitude_i)$ is a flexible function of latitude

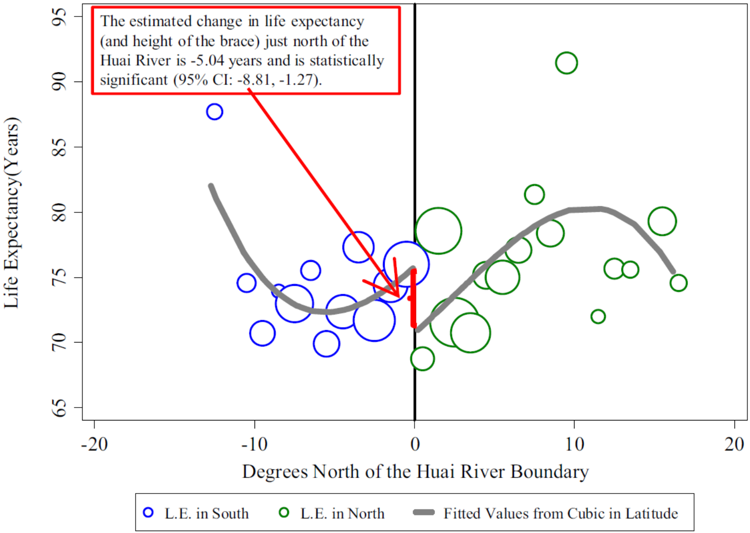
ν_i, ε_i are error terms

First stage



The estimated change in TSP (and height of the brace) just north of the Huai River is 247.5 µg³ and is statistically significant (95% CI: 114.5, 380.6).
TSP(µg³)
800
600
400
200
0
-20
-10
0
10
20
Degrees North of the Huai River Boundary
TSP in South
TSP in North
Fitted Values from Cubic in Latitude

Reduced form



2SLS

Table 3. Using the Huai River policy to estimate the impact of TSPs (100 μg/m³) on health outcomes

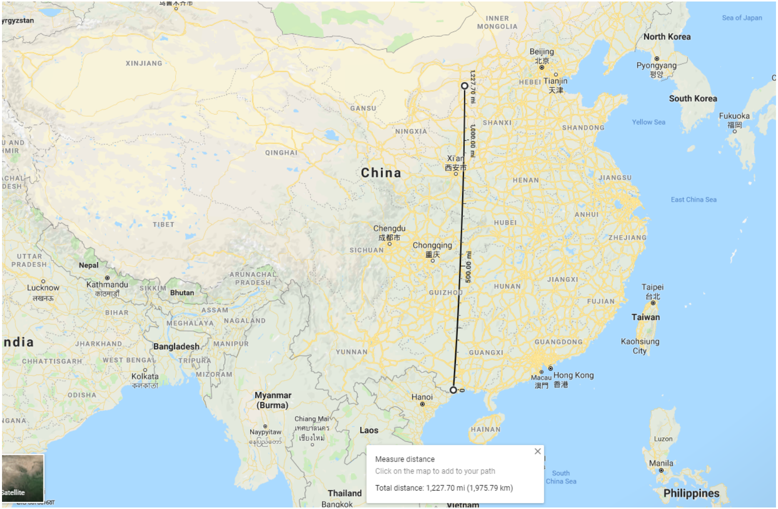
Dependent variable	(1)	(2)	(3)
Panel 1: Impact of "North" on the listed variable, ordinary least squares			
TSPs, 100 μg/m³	2.48*** (0.65)	1.84*** (0.63)	2.17*** (0.66)
ln(All cause mortality rate)	0.22* (0.13)	0.26* (0.13)	0.30* (0.15)
ln(Cardiorespiratory mortality rate)	0.37** (0.16)	0.38** (0.16)	0.50*** (0.19)
ln(Noncardiorespiratory mortality rate)	0.00 (0.13)	0.08 (0.13)	0.00 (0.13)
Life expectancy, y	−5.04** (2.47)	−5.52** (2.39)	−5.30* (2.85)
Panel 2: Impact of TSPs on the listed variable, two-stage least squares			
ln(All cause mortality rate)	0.09* (0.05)	0.14** (0.07)	0.14* (0.08)
ln(Cardiorespiratory mortality rate)	0.15** (0.06)	0.21** (0.09)	0.23** (0.10)
ln(Noncardiorespiratory mortality rate)	0.00 (0.05)	0.04 (0.07)	0.00 (0.06)
Life expectancy, y	−2.04** (0.92)	−3.00** (1.33)	−2.44 (1.50)
Climate controls	No	Yes	Yes
Census and DSP controls	No	Yes	Yes
Polynomial in latitude	Cubic	Cubic	Linear
Only DSP locations within 5° latitude	No	No	Yes

RD implementation details

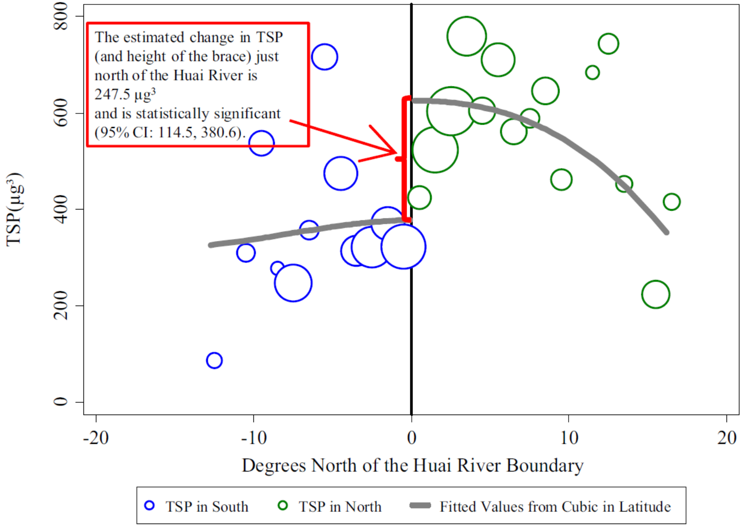
When we estimate RDs, we want to be careful to consider:

- Bandwidth selection
- Functional form

Bandwidth selection



Bandwidth selection



Key conversion: 1 degree ≈ 110 km

Bandwidth selection

Two increasingly popular methods for RD bandwidth selection:

- Imbens and Kalyanaraman (2012)
- Calonico, Cattaneo, and Titiunik (2014a, 2014b, 2015)

→ Both implemented in `R` and STATA with the `rdrobust` package

→ Best practice: do these, but also test sensitivity to alternatives

Bandwidth selection

Robustness checks of choice of functional form for latitude, DSP locations within 5^0 Latitude of Huai River

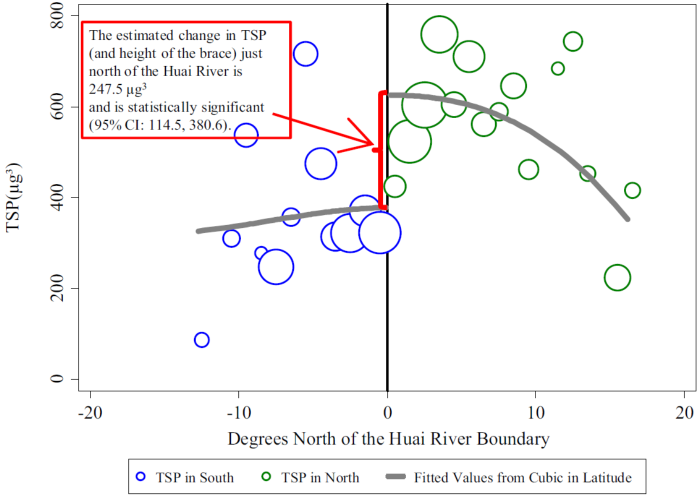
	Linear & Controls	Quadratic & Controls	Cubic & Controls
	(1)	(2)	(3)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares			
TSP (100 μg/m^3)	2.17***	1.18*	0.60
	(0.66)	(0.69)	(0.55)
	[0.576]	[0.0001]	[0.28]
	{232}	{216.4}	{215.8}
ln(All Cause Mortality Rate)	0.30*	0.24	0.33
	(0.15)	(0.17)	(0.20)
	[0.171]	[0.587]	[0.577]
	{21.91}	{22.84}	{23.74}
ln(Cardiorespiratory Mortality Rate)	0.50***	0.40*	0.51**
	(0.19)	(0.22)	(0.24)
	[0.042]	[0.359]	[0.652]
	{56.1}	{56.05}	{57.19}
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.02	0.09
	(0.13)	(0.15)	(0.18)
	[0.999]	[0.931]	[0.66]
	{8.704}	{10.56}	{11.73}

Bandwidth selection

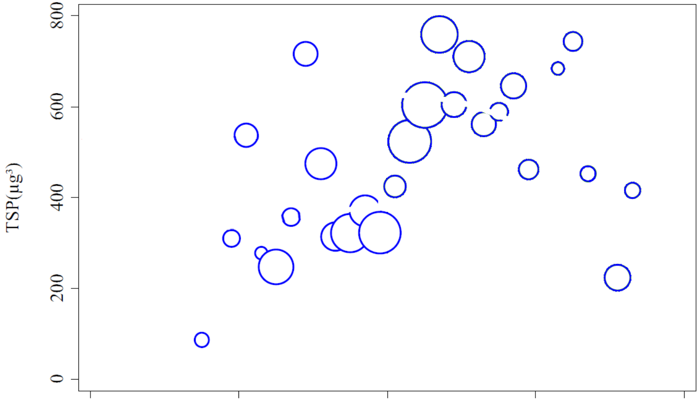
Robustness checks of choice of functional form for latitude, DSP locations within 5^0 Latitude of Huai River

	Linear & Controls	Quadratic & Controls	Cubic & Controls
	(1)	(2)	(3)
Panel 2: Impact of TSP (100 $\mu g/m^3$) on the Listed Variable, Two-stage Least Squares			
ln(All Cause Mortality Rate)	0.14*	0.20	0.56
	(0.08)	(0.17)	(0.56)
ln(Cardiorespiratory Mortality Rate)	0.23**	0.34	0.85
	(0.10)	(0.21)	(0.76)
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.02	0.16
	(0.06)	(0.13)	(0.34)
Life Expectancy (years)	-2.44	-3.44	-9.57
	(1.50)	(3.26)	(10.03)

Functional form



Functional form



Functional form

Robustness checks of choice of functional form for latitude

	Linear & Controls	Quadratic & Controls	Cubic & Controls	Quartic & Controls	Quintic & Controls
	(1)	(2)	(3)	(4)	(5)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares					
TSP (100 μg/m^3)	2.89***	2.63***	1.84***	1.95***	1.52**
	(0.56)	(0.49)	(0.63)	(0.59)	(0.72)
	[0.988]	[0.068]	[0.148]	[0.229]	[0.671]
	{492.4}	{489}	{487.2}	{486.3}	{487.5}
ln(All Cause Mortality Rate)	0.12	0.09	0.26*	0.26**	0.37**
	(0.10)	(0.10)	(0.13)	(0.13)	(0.16)
	[0.276]	[0.215]	[0.035]	[0.908]	[0.409]
	{39.88}	{38.8}	{34.11}	{35.92}	{36.13}
ln(Cardiorespiratory Mortality Rate)	0.13	0.09	0.38**	0.39**	0.47**
	(0.13)	(0.13)	(0.16)	(0.16)	(0.19)
	[0.652]	[0.243]	[0.003]	[0.747]	[0.696]
	{102.3}	{101.5}	{91.92}	{93.34}	{94.62}
ln(Non-Cardiorespiratory Mortality Rate)	0.09	0.05	0.08	0.07	0.19
	(0.10)	(0.09)	(0.13)	(0.12)	(0.14)
	[0.135]	[0.151]	[0.933]	[0.973]	[0.35]
	{43.04}	{41.27}	{43.13}	{45.07}	{44.97}
Life Expectancy (years)	-1.62	-1.29	-5.52**	-5.67**	-5.43*
	(1.66)	(1.68)	(2.39)	(2.36)	(2.94)
	[0.101]	[0.6]	[0.001]	[0.737]	[0.984]
	{757.1}	{758}	{746.8}	{748.2}	{750.2}

Functional form

Robustness checks of choice of functional form for latitude

	Linear & Controls	Quadratic & Controls	Cubic & Controls	Quartic & Controls	Quintic & Controls
	(1)	(2)	(3)	(4)	(5)
Panel 2: Impact of TSP (100 $\mu g/m^3$) on the Listed Variable, Two-stage Least Squares					
ln(All Cause Mortality Rate)	0.04	0.03	0.14**	0.13**	0.24*
	(0.03)	(0.04)	(0.07)	(0.06)	(0.13)
ln(Cardiorespiratory Mortality Rate)	0.05	0.03	0.21**	0.20**	0.31*
	(0.04)	(0.05)	(0.09)	(0.08)	(0.17)
ln(Non-Cardiorespiratory Mortality Rate)	0.03	0.02	0.04	0.04	0.13
	(0.03)	(0.03)	(0.07)	(0.06)	(0.10)
Life Expectancy (years)	-0.56	-0.49	-3.00**	-2.90**	-3.56
	(0.54)	(0.62)	(1.33)	(1.24)	(2.34)

Functional form

Higher-order polynomials should not be used in RD:

- That is, anything with a non-linear term!
- **Intuition:** Endpoints have an outsized impact on polynomials...
- ... but this is really problematic for RD!
- In practice, estimates tend to be very sensitive

→ Instead, use (local) linear regression

Recap

TL;DR:

1. The regression discontinuity is great
2. We can use it with non-binary treatments, and without sharp D_i assignment
3. But getting the details right is key!