Aspirations in the Air: Effect of Development Schemes on AQI

Evidence from a Spatial RD in India

Aayushi Agarwal, Bhaskar Agarwal, Gaurav Banerjee, Nandish Patel

December 7, 2025

Contents

- Background: Why study development reforms in the context of air quality?, Common problems, Exogenous variation...
- Methodology: Sources of data, Wrangling, Merging, Cleaning
- Empirical strategy: Identification, Model, Controls
- Results: State-wise, Aggregated, Placebo Test
- Conclusion: Shortcomings, Way ahead, Q/A

Why look at development schemes?

Mechanism I

AQI $= \downarrow$ f(Development)

- Clean Cooking Adoption
- Waste Management
- Agricultural Residue Management

On the other hand...

Mechanism II

AQI $= \uparrow$ f(Development)

- Infrastructure Development
- Increased Industrial Output
- Rise in Vehicular Emission

Finding the direction of the effect is then an empirical problem

Exogenous variation

- Development policies are implemented non-randomly
- Cannot isolate causal impact at the District level

Government chooses districts to be treated

↓

District boundaries become treatment cutoff

↓

Compare subdistricts along this cutoff

↓

LATE using spatial regression discontinuity

Aspirational Districts Scheme

- Launched in January 2018 by NITI Aayog
- Aimed at rapid transformation of underdeveloped yet aspirational districts
- Focus areas: Basic Infrastructure, Health & Nutrition, Education, Agriculture & Financial Inclusion
- Selected districts using a composite deprivation index

Methodology

Research question

"*How do development policies impact subdistrict level air quality in India?*"

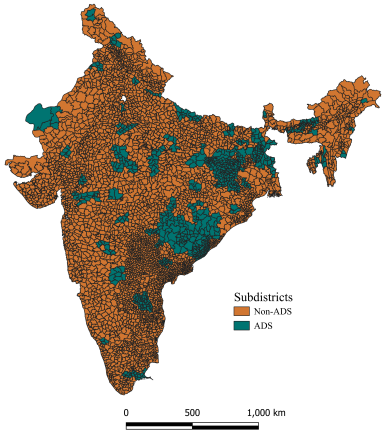
Sources of data:

- Master shapefile- Survey of India (SoI)
- Treatment- NITI Aayog
- AQI- Socioeconomic High-resolution Rural-Urban Geographic Platform (SHRUG)
- Controls- ibid

Subdistricts shapefile

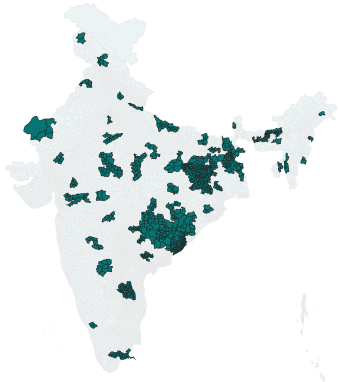
- Source: Survey of India (SoI)
- Tehsil/Taluk level administrative shapefile
- 4723 features, 5 fields, LCC_WGS84

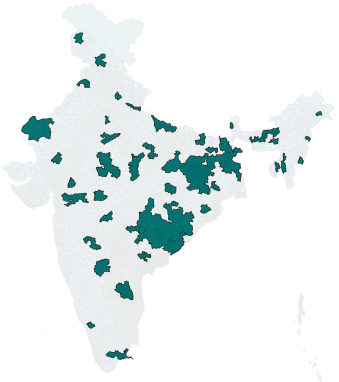
ADS treatment status

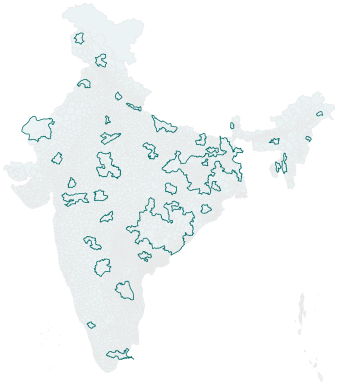


- NITI Aayog only provides names of Districts
- Merging areas with names is a nightmare in India- Differing spellings, Changed names...
- As a result, we fuzzy match district names using Jaro-Winkler Distance
- Filtered for 29 relevant states

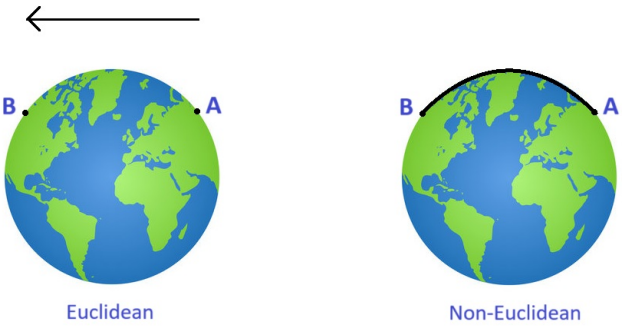
Cutoff





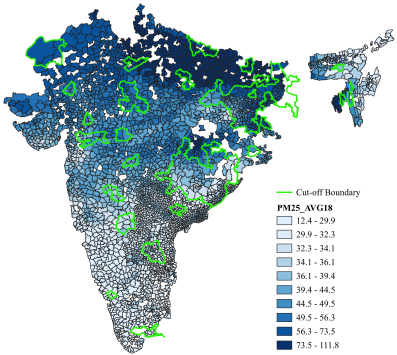


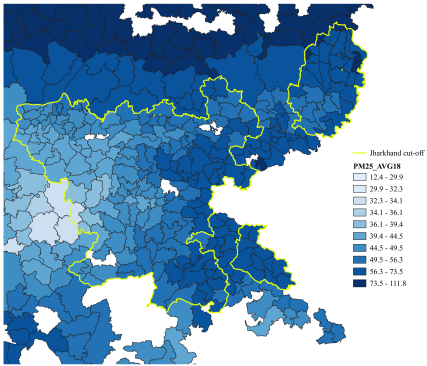
Distance to cutoff



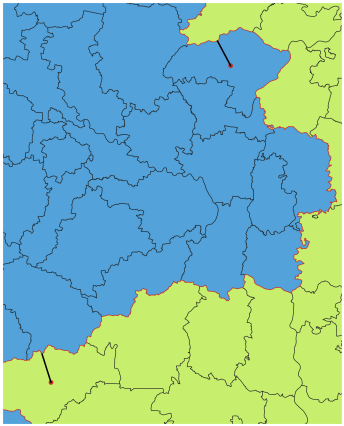
- 'Distance to cutoff' is the perpendicular distance from centroids of subdistricts to the cutoff boundary. +ve for treated and -ve for control.

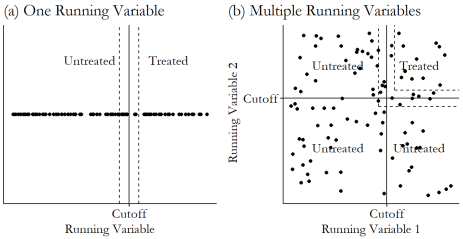
AQI


Cut-off Boundary
PM25_AVG18
12.4 - 29.9
29.9 - 32.3
32.3 - 34.1
34.1 - 36.1
36.1 - 39.4
39.4 - 44.5
44.5 - 49.5
49.5 - 56.3
56.3 - 73.5
73.5 - 111.8


Jharkhand cut-off
PM25_AVG18
12.4 - 29.9
29.9 - 32.3
32.3 - 34.1
34.1 - 36.1
36.1 - 39.4
39.4 - 44.5
44.5 - 49.5
49.5 - 56.3
56.3 - 73.5
73.5 - 111.8

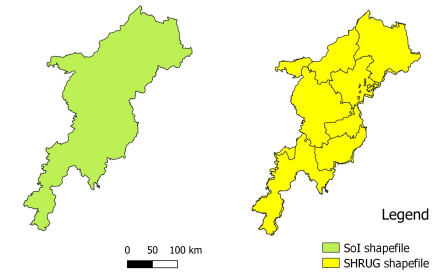
Why controls?





Source: The Effect, Nick Huntington-Klein

Controls



S. No	Controls	Description
1.	pc18_sc_share	Scheduled castes population share
2.	pc18_st_share	Scheduled tribes population share
3.	pc18_lit_share	Literate population share
4.	pc18_rural_share	Rural population share
5.	pc18_work_share	Working population share
6.	pc18_forest_share	Forest cover share

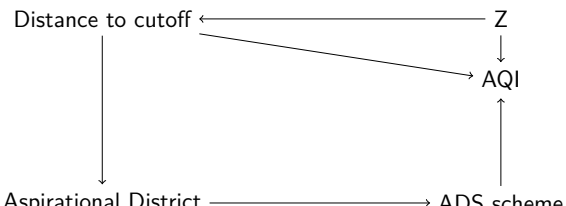
- Source: SHRUG, PC01 and PC11
- Extrapolation:

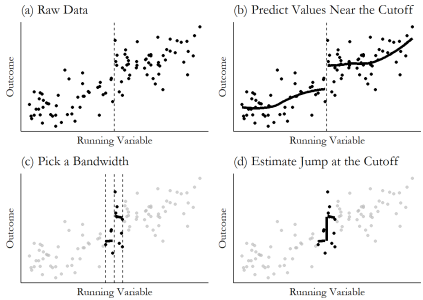
$$\text{Growthrate}_{ij} = \frac{PC11_{ij} - PC01_{ij}}{PC01_{ij}} * 100$$

$$\text{AGR}_{ij} = \frac{Growthrate_{ij}}{10}$$

$$\text{PC18}_{ij} = PC11_{ij}\left(1 + \frac{AGR_{ij}}{100}\right)^{7}$$

Identification





Source: The Effect, Nick Huntington-Klein

Model

Linear

$$pm25\_avg18_i = \beta_0 + \beta_1 A_i + \beta_2 D_i + \beta_3 A_i D_i + \mu_i$$

β_1 is the coefficient of interest

2nd order polynomial

$$pm25\_avg18_i = \beta_0 + \beta_1 A_i + \beta_2 D_i + \beta_3 D_i^2 + \beta_4 A_i D_i + \beta_5 A_i (D_i)^2 + \mu_i$$

Following the recommendations of [Gelman and Imbens, 2019], we do not check for higher order polynomials greater than two

rdrobust

- Bias correction
- MSE optimised bandwidth selection
- Triangularly weighted kernel
- Heteroskedasticity robust standard errors
- Controls
- Restricting geographical area under study

Estimating model with controls

$$pm25\_avg18_{is} = \beta_0 + \beta_1 A_i + \beta_2 D_i + \beta_3 A_i D_i + X_i\gamma + \lambda_s + \mu_i$$

Main RD estimates

Table: State-wise Robust RD Estimates

	Estimate	95% CI	Std. Error	Robust P-Value	Obs	Eff. Obs	Bandwidth	Covs
ANDHRA PRADESH	2.125	[−0.772, 5.021]	1.478	0.151	635	116	13.515	Yes
ANDHRA PRADESH	1.382	[−1.813, 4.577]	1.630	0.397	635	128	15.345	No
BIHAR	14.087	[−128.410, 156.585]	72.704	0.846	79	11	7.035	Yes
BIHAR	−39.645	[−110.808, 31.518]	36.308	0.275	79	5	5.643	No
GUJARAT	3.702	[−14.711, 22.115]	9.394	0.694	201	47	36.846	Yes
GUJARAT	−0.989	[−24.325, 22.346]	11.906	0.934	201	47	34.969	No
JHARKHAND	**-18.790**	**[-35.527,-2.054]**	**8.539**	**0.028**	**256**	**17**	**4.398**	**Yes**
JHARKHAND	−22.381	[−40.013, −4.749]	8.996	0.013	256	43	6.602	No
MADHYA PRADESH	3.868	[−4.513, 12.250]	4.277	0.366	259	77	24.041	Yes
MADHYA PRADESH	6.142	[−2.170, 14.454]	4.241	0.148	259	72	21.989	No
MAHARASHTRA	0.746	[−17.722, 19.214]	9.423	0.937	329	53	15.824	Yes
MAHARASHTRA	−2.357	[−28.844, 24.129]	13.514	0.862	329	54	16.327	No
MIZORAM	8.251	[1.925, 14.578]	3.228	0.011	16	15	124.220	Yes
MIZORAM	23.682	[14.480, 32.884]	4.695	0.000	16	15	124.220	No
RAJASTHAN	29.610	[−25.449, 84.670]	28.092	0.292	230	45	17.098	Yes
RAJASTHAN	31.702	[−26.159, 89.562]	29.521	0.283	230	63	25.969	No
TELANGANA	2.924	[−17.674, 23.522]	10.509	0.781	429	23	5.165	Yes
TELANGANA	−13.303	[−50.241, 23.636]	18.847	0.480	429	21	4.833	No

Aggregated RD estimate

Table: Aggregated Bias-corrected Robust RD Estimates

	Estimate	95% CI	Std. Error	Robust P-Value	Obs	Eff. Obs	Bandwidth	Covs
ALL STATES	1.829	[−1.561, 5.220]	1.730	0.290	3456	889	20.572	Yes
ALL STATES	−1.916	[−17.619, 13.787]	8.012	0.811	3456	1194	30.659	No

Notes. Standard errors are clustered by state.

Placebo test

Table: Placebo Test

Outcome variable: PM2.5 in 2017 (before policy implementation)

	Estimate	95% CI	Std. Error	Robust P-Value	Obs	Eff. Obs	Bandwidth	Covs
JHARKHAND	−17.297	[−36.163, 1.568]	9.626	0.072	256	29	5.261	Yes

Notes. Standard errors are heteroskedasticity robust.

Conclusion

- We find that AQI in 2018 is lower by approximately 18.79 units in Jharkhand for the treated subdistricts
- Evidence for dual mechanism
- Shortcomings: Spatial spillovers, Non-euclidean distance
- Way forward: Village/Town level, Two-running variable approach

Thank You :)