Appendix

A. Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Obesity and Air pollution					
BMI (kg/m^2)	6000	22.528	3.611	10.381	66.667
$PM_{2.5}$ (μg/m^3)	5940	49.998	20.483	7.141	93.447
height (cm)	6000	164.511	7.595	120	190
weight (kg)	6000	61.158	11.578	30	160
The geographic data					
latitude(°N)	6000	33.974	6.514	22.528	49.186
longitude (°E)	6000	114.589	7.778	82.051	129.427
area (km^3)	6000	2640.333	4007.451	678	38282
Individual heterogeneity					
yob	6000	1964.792	14.092	1910	1990
edu (year)	5491	9.399	3.677	1	24
gender_dummy	6000	.518	.5	0	1
County characteristic					
population (10,000 person)	6000	70.267	35.438	16	171
employees (person)	6000	44721.887	41708.374	7346	304147
lexp (10,000, ￥)	6000	11.724	.636	10.552	13.842
lrev (10,000, ￥)	6000	10.677	1.149	7.585	13.861
ldep (10,000, ￥)	6000	13.224	1.02	11.006	16.315
lfa (10,000, ￥)	6000	12.841	.99	9.952	14.916

Note: A. Table 1 describes the outcome, assignment variables, and covariates used in this paper. Section 3.4 explains the data sources for each variable in this table. (Missing data: For PM, Xiaohe District of Guizhou Province has rebranded as part of Qianjiang District, and its original county code 520114, which is not used anymore, does not match with the other databases, so there are 60 observations lost. Similarly, Weiyang of Jiangsu Province has changed the name to Ganjiang, Juchao of Anhui Province has changed to Chaohu, Laicheng of Shandong Province has changed to Laiwu, and Liupanshui of Guizhou Province to Shuicheng, but their county code remains the same, so there is no extra data losing; For Edu, 509 individuals did not fill out the year of education in the questionnaire.)

A. Table 2: Definition of Abbreviated Variables

Abbreviation	Variable	Definition
BMI	Obesity	Individuals body mass index
$PM_{2.5}$	Air Pollution	The county-level annual mean particulate matter less than 2.5 micrometers
d	Treatment	Dummy of receiving the central winter heat (latitude>33 degrees north)
v	Assignment	Latitude – 33 Degree North
vd	Interaction	d*v
yob	Age	Year of birth
edu	Education	Number of educated years
lexp	Expenditure	The logarithm of local general budget expenditure
lrev	Revenue	The logarithm of local general budget revenue
ldep	Deposit	The logarithm of average household saving deposit
LFA	Fixed asset	The logarithm of investment in fixed asset

Note: A. Table 2 describes the definitions of the variables and their abbreviations we reported in the regression result tables.

A. Table 3: Covariates are not discontinuous at the threshold

	(1) edu	(2) gender_dummy	(3) yob	(4) longitude	(5) area	(6) lexp	(7) employees	(8) population	(9) ldep	(10) lfa
Conventional	-0.185	-0.0582	-1.140	-4.077	595.9	0.122	-3452.1	-19.18	-0.431	-0.988
	(0.6870)	(0.0516)	(1.6569)	(3.1099)	(564.4598)	(0.6737)	(1.5e+04)	(27.0738)	(0.8612)	(0.6465)
Bias-corrected	0.305	-0.0649	-0.836	-3.218	815.2	0.564	2120.8	-15.87	-0.116	-0.915
	(0.6870)	(0.0516)	(1.6569)	(3.1099)	(564.4598)	(0.6737)	(1.5e+04)	(27.0738)	(0.8612)	(0.6465)
Robust	0.305	-0.0649	-0.836	-3.218	815.2	0.564	2120.8	-15.87	-0.116	-0.915
	(0.7197)	(0.0595)	(2.0611)	(3.5420)	(672.4952)	(0.7352)	(1.8e+04)	(32.0387)	(0.9675)	(0.7588)
N	5491	6000	6000	100	100	100	100	100	100	100

Standard errors in parentheses: * $p < .1$, ** $p < .05$, *** $p < .01$

Note: Columns (1), (2), and (3) report the RD estimates of the coefficient on education, gender, and year of birth (the individual characteristics), using from local linear regression (LLR), following the Equation (2.2) and Equation (3.2) with uniform kernel and bandwidth selected by the method proposed by Imbens and Kalyanaraman (2012). Similarly, Column (4), (5), (6), (7), (8), and (9) reports the RD estimates of the coefficient on longitude, area, expenditure, population, deposit, and fixed asset (the county demographic). The full covariates are listed in A. Table 1. The first row is the conventional estimates, bias-corrected estimates considered the error term, and robust estimates considered both the error term and standard error.

A. Table 4: Different Polynomial Functions for Parametric RD Estimates of Central Heating Policy

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PM_{2.5}$	2.964***	3.163***	-16.79***	-7.738***	-8.515***	-5.389***	-6.811***	-4.261**
	(0.8867)	(0.8155)	(0.9577)	(0.9458)	(1.0779)	(1.2991)	(1.6524)	(1.9159)
N	5940	5940	5940	5940	5940	5940	5940	5940
adj. R^2	0.253	0.420	0.319	0.459	0.364	0.500	0.368	0.506
BMI	0.321	0.461*	-0.709*	-0.497	-0.631	-0.393	-0.315	-0.254
	(0.2603)	(0.2667)	(0.3797)	(0.4183)	(0.5279)	(0.5286)	(0.6296)	(0.6561)
N	6000	5491	6000	5491	6000	5491	6000	5491
adj. R^2	0.021	0.066	0.025	0.068	0.025	0.068	0.025	0.068
Function	linear	linear	2nd	2nd	3rd	3rd	4th	4th
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Standard errors in parentheses: * $p < .1$, ** $p < .05$, *** $p < .01$

Note: In A. Table 4, we report RD estimates of the coefficient on $PM_{2.5}$ and BMI, using different degree latitude from the Qinling Mountains and Huai River and different polynomial function forms(1-4) in degrees latitude from the border interacted with a policy dummy variable, following the Equation (2.1) and Equation (3.1) with the full sample. All standard errors in columns are clustered at the county level. Results are based on the full sample and (2) and (4) include the covariates listed in A. Table 1.

A. Table 5: Different Kernels Non-parametric RD Estimates of Central Heating Policy

<A. Table 5.1>

$PM_{2.5}$	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	-24.98^{***}	-18.94^{*}	-22.32^{***}	-15.82^{**}	-17.86^{**}	-15.18^{*}
	(9.0195)	(10.3495)	(7.0773)	(7.7027)	(7.5090)	(7.9190)
Bias-corrected	-23.48^{***}	-30.22^{***}	-25.94^{***}	-14.38^{*}	-19.80^{***}	-12.87
	(9.0195)	(10.3495)	(7.0773)	(7.7027)	(7.5090)	(7.9190)
Robust	-23.48^{**}	-30.22^{**}	-25.94^{***}	-14.38	-19.80^{**}	-12.87
	(10.6093)	(13.4492)	(8.4729)	(9.2910)	(9.1488)	(9.5226)
Kernel	Uniform	Uniform	Triangle	Triangle	Epanechnikov	Epanechnikov
Controls	No	Yes	No	Yes	No	Yes
N	99	99	99	99	99	99

<A. Table 5.2>

BMI	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	-0.543^{*}	0.297	0.565	0.370	-0.189	0.370
	(0.3176)	(0.3669)	(0.4344)	(0.4414)	(0.3556)	(0.4316)
Bias-corrected	-0.519	0.503	0.692	0.331	-0.0920	0.333
	(0.3176)	(0.3669)	(0.4344)	(0.4414)	(0.3556)	(0.4316)
Robust	-0.519	0.503	0.692	0.331	-0.0920	0.333
	(0.3685)	(0.4451)	(0.5076)	(0.5435)	(0.4298)	(0.5304)
Kernel	Uniform	Uniform	Triangle	Triangle	Epanechnikov	Epanechnikov
Controls	No	Yes	No	Yes	No	Yes
N	6000	5491	6000	5491	6000	5491

Note: In A. Table 5.1 and A. Table 5.2, we separately report RD estimates of the coefficient on $PM_{2.5}$ and BMI, from local linear regression (LLR), following the Equation (2.2) and Equation (3.2) with 3 different kernels, which are Uniform, Triangle and Epanechnikov, and bandwidth selected by the method proposed by Imbens and Kalyanaraman (2012). Results are based on the full sample and (2), (4), and (6) include the covariates listed in A. Table 1. Note that, for $PM_{2.5}$, A. Table 5.1 are based on 99 counties observations, and A. Table 5.2 are based on 5940 individual observations, so estimates of (3) and (4) are higher than (1) and (2). The first row is the conventional estimates, bias-corrected estimates considered the error term, and robust estimates considered both the error term and standard error.

A. Table 6: The OLS Approach

Linear regression

BMI	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
$PM_{2.5}$	.014	.004	3.79	0	.007	.021	***
edu	.001	.018	0.08	.938	-.035	.037	
gender_dummy	-.603	.113	-5.36	0	-.826	-.38	***
yob	-.051	.005	-10.50	0	-.061	-.041	***
latitude	.075	.016	4.62	0	.043	.107	***
longitude	.011	.012	0.88	.382	-.014	.035	
area	0	0	-1.59	.115	0	0	
lexp	.156	.218	0.72	.475	-.276	.589	
employees	0	0	2.18	.031	0	0	**
population	-.001	.003	-0.33	.744	-.007	.005	
ldep	-.13	.142	-0.91	.362	-.413	.152	
lfa	.01	.115	0.09	.932	-.218	.237	
Constant	118.247	9.924	11.91	0	98.552	137.941	***

Mean dependent var	22.574	SD dependent var	3.632
R-squared	0.071	Number of obs.	5431.000
F-test	26.478	Prob > F	0.000
Akaike crit. (AIC)	29045.842	Bayesian crit. (BIC)	29131.641

*** *p<.01*, ** *p<.05*, * *p<.1*

Note: In A. Table 6, we report the detailed OLS regression result of the association between $PM_{2.5}$ and BMI, following the equation (1). Results are based on the full sample and include the covariates listed in A. Table 1. All standard errors are clustered at the county level.

A. Table 7: The 2SLS Approach

Instrumental variables (2SLS) regression

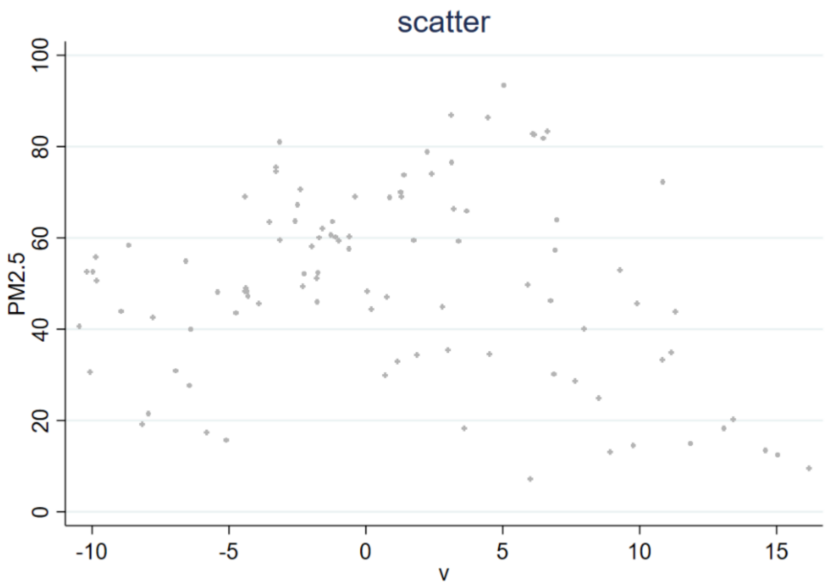
BMI	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
$PM_{2.5}$	.121	.052	2.34	.019	.02	.222	**
v	-.209	.131	-1.60	.11	-.466	.047	
vd	.537	.248	2.17	.03	.052	1.023	**
edu	-.059	.033	-1.78	.076	-.125	.006	*
gender_dummy	-.682	.113	-6.02	0	-.903	-.46	***
yob	-.047	.005	-10.18	0	-.056	-.038	***
longitude	.042	.018	2.37	.018	.007	.077	**
area	0	0	-2.54	.011	0	0	**
lexp	.867	.38	2.28	.023	.122	1.613	**
employees	0	0	2.62	.009	0	0	***
population	-.011	.005	-2.12	.034	-.021	-.001	**
ldep	-.939	.41	-2.29	.022	-1.743	-.135	**
lfa	-.39	.215	-1.81	.07	-.812	.032	*
Constant	111.378	9.554	11.66	0	92.653	130.104	***

Mean dependent var	22.574	SD dependent var	3.632
R-squared	.	Number of obs.	5431.000
Chi-square	327.594	Prob > chi2	0.000

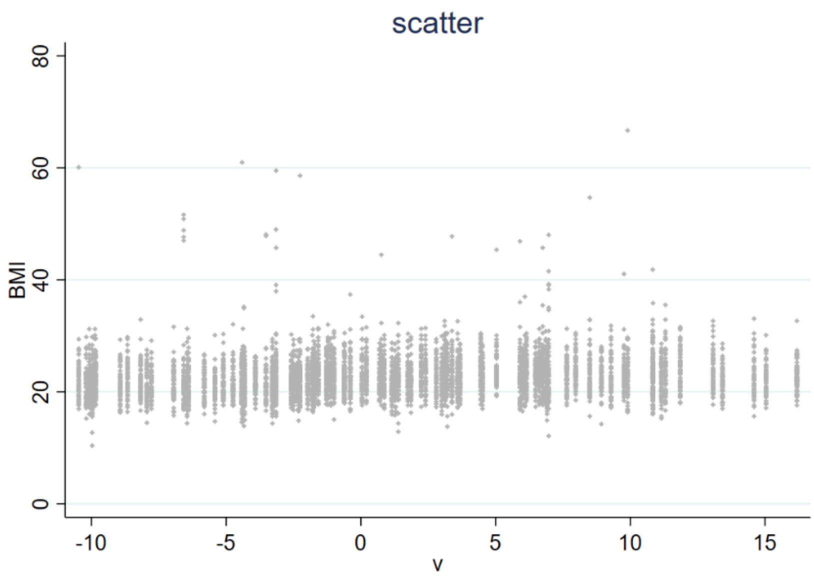
*** *p*<.01, ** *p*<.05, * *p*<.1

Note: In A. Table 6, we report the detailed 2SLS estimates using the difference of degree latitude from the Qinling Mountains and Huai River as the IV and a linear polynomial in degrees latitude from the border interacted with a policy dummy variable, following the equation (4). Results are based on the full sample and include the covariates listed in A. Table 1.

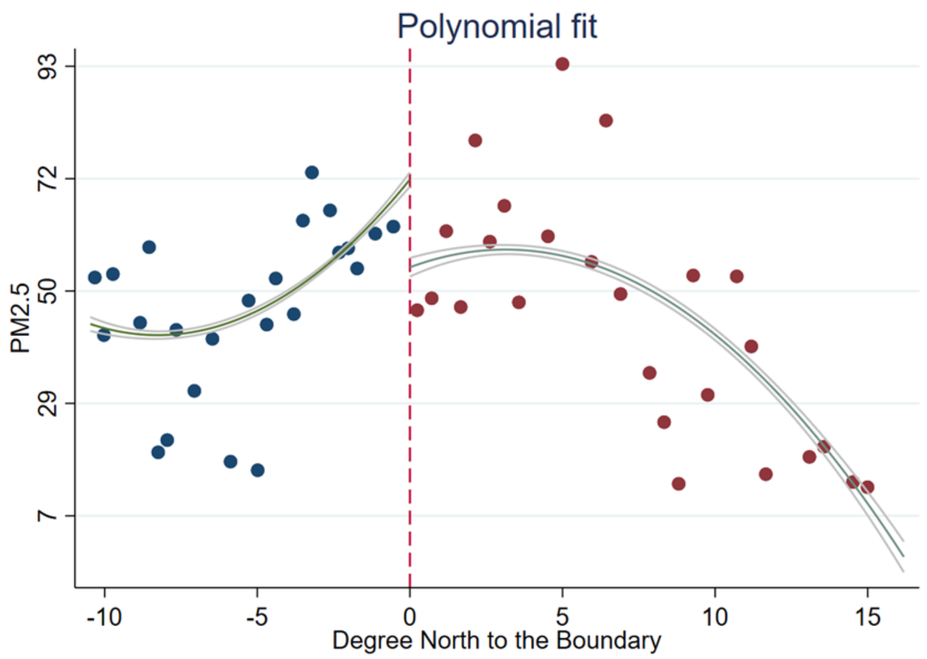
A. Figure 1: Scatters of $PM_{2.5}$ on Degrees North from the North-South line



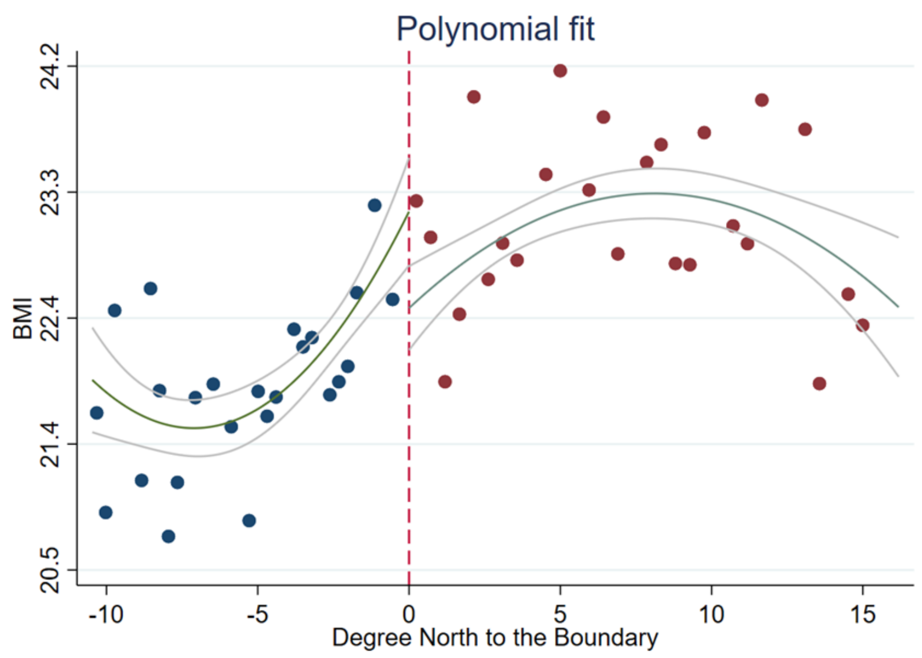
A. Figure 2: Scatters of BMI on Degrees North from the North-South line



A. Figure 3: Polynomial Fit of $PM_{2.5}$ on Degrees North from the North-South line

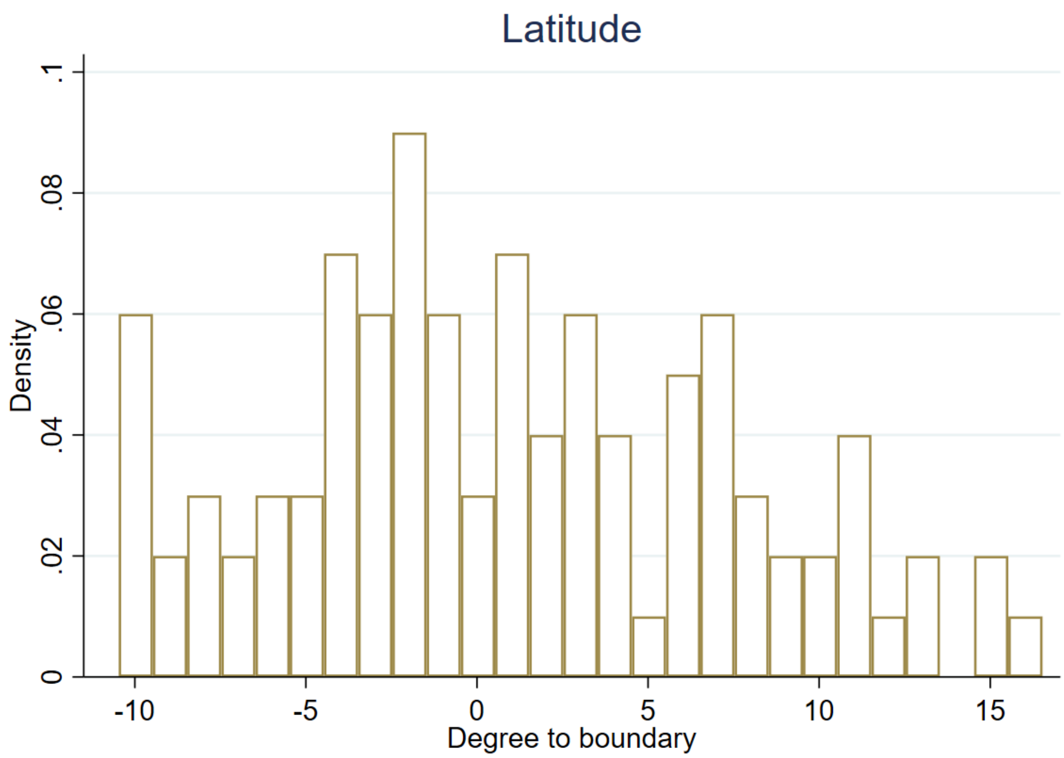


A. Figure 4: Polynomial Fit of BMI on Degrees North from the North-South line

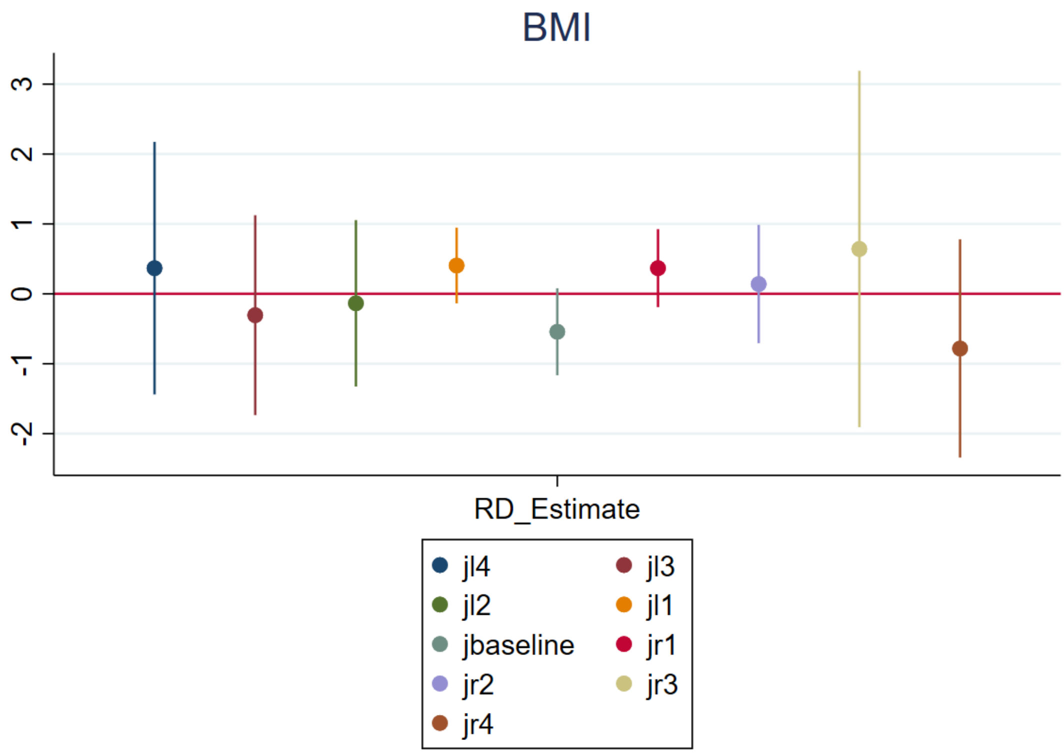


Note: The discontinuity is defined when the distance from the location of observations to the boundary is 0 (when the latitude is 33 degrees north). The quadric fitted value of $PM_{2.5}$ or BMI on the difference of latitude from the boundary is estimated respectively on each side of the division line with 34 bins and confidence interval.

A. Figure 5: Polynomial Fit of BMI on Degrees North from the North-South line



A. Figure 6: Placebo Test for Different Thresholds



Note: The baseline point is the real cut-off of zero, l or r indicates the new cut-off is on the left or right of the threshold, and the following number is the number of the quintile. For instance, 'l4' indicates the new cut-off is the 80 percent point on the left of the minimum difference from county centroids to the border. As shown in the A. Figure 6, RD estimate with real cut-off point is marginally significant from zero(the CI touch a little bit the 0 line). The other placebo cut-offs' RD estimates are not significant, for their confident intervals all go through zero. Though by changing the cut-off point, we don't find any local treatment effect for obesity, the RD design of BMI is still considered as not so valid.