Does city-county consolidation policy promote energy efficiency? Evidence from a quasi-natural experiment in China

Weihua Yu[1], Wenjing Liang[1], Xin Yao[2,3*]

[1] Jinhe Center for Economic Research, Xi'an Jiaotong University, Xi'an, China, [2] China Center for Energy Economics Research, School of Economics, Xiamen University, Xiamen, China and [3] Fujian Province Key Laboratory of Philosophy and Social Science for Energy Economics and Green Smart Development, Xiamen University, Xiamen, China

[*] Corresponding author: Xin Yao; Email: lyhneyao@sina.com

Online Appendix

Appendix A. Measurement of energy efficiency

This study employs the super-efficiency slack-based model (SBM) method factoring in the undesirable outputs to quantify the energy efficiency of counties in China under the assumption of variable returns to scale.

First, we will introduce the standard SBM model:

Suppose that the production system consists of j decision-making units (DMUs), where each DMU is composed of three input-output vectors: inputs, desirable outputs, and undesirable outputs. The production process utilizes n units of inputs to generate S_1 units of desirable outputs and S_2 units of undesirable outputs.

The three input-output vectors can be represented as follows: $x \in R^n, y^g \in R^{S_1}, y^b \in R^{S_2}$. We define vector X, Y^g, Y^b：

$$X = \left[x_1, x_2, \cdots, x_j \right] \in R^{n \times j},$$

$$Y^g = \left[y_1^g, y_2^g, \cdots, y_j^g \right] \in R^{S_1 \times j},$$

$$Y^b = \left[y_1^b, y_2^b, \cdots, y_j^b \right] \in R^{S_2 \times j},$$

where $X > 0, Y^g > 0, Y^b > 0,$ the production possibility set can be delineated by the following definition：

$$T = \{(x, y^g, y^b) | x \geq X\theta, y^g \geq Y^g\theta, y^b \leq Y^b\theta, \theta \geq 0\},$$

where Y^g denotes desirable outputs； Y^b denotes undesirable outputs; θ denotes weight vector.

$$\rho = min \frac{1 - \frac{1}{n}\sum_{i=1}^{n} \frac{S_i^-}{x_{i0}}}{1 + \frac{1}{S_1 + S_2}(\sum_{a=1}^{S_1} \frac{S_a^g}{y_{a0}^g} + \sum_{a=1}^{S_2} \frac{S_a^b}{y_{a0}^b})}$$

$$s.t.\begin{cases} x_0 = X\theta + S^- \\ y_0^g = Y^g\theta - S^g \\ y_0^b = Y^b\theta - S^b \\ S^- \geq 0, S^g \geq 0, S^b \geq 0, \theta \geq 0 \\ \sum_{a=1}^{j} \theta_a = 1 \end{cases}$$

where $S = (S^-, S^g, S^b)$ denotes slack vectors of input, desirable outputs, and undesirable outputs; ρ is the efficiency value of the DMU, $0< \rho < 1$. For DMU (x_0, y_0^g, y_0^b), if and only if $\rho = 1$, that is $S^- = S^g = S^b = 0$, it is efficient; if $0 \leq \rho < 1$, it is inefficient.

Super-efficiency SBM model modifies the above model to further compare the efficiency values of DMU with an efficiency of 1:

$$\rho^N = min \frac{\frac{1}{n}\sum_{i=1}^{n} \frac{\bar{x_i}}{x_{i0}}}{\frac{1}{S_1 + S_2}\left(\sum_{a=1}^{S_1} \frac{\overline{y_a^g}}{y_{a0}^g} + \sum_{a=1}^{S_2} \frac{\overline{y_a^b}}{y_{a0}^b}\right)}$$

$$s.t.\begin{cases} \bar{x} \geq \sum_{c=1,\neq k}^{j} \theta_c\, x_c \\ \overline{y^g} \leq \sum_{c=1,\neq k}^{j} \theta_c\, y_c^g \\ \overline{y^b} \geq \sum_{c=1,\neq k}^{j} \theta_c\, y_c^b \\ \bar{x} \geq x_0, \overline{y^g} \leq y_0^g, \overline{y^b} \geq y_0^b, \overline{y^g} \geq 0, \theta \geq 0 \\ \sum_{c=1,\neq k}^{n} \theta_c = 1 \end{cases}$$

where ρ^N represents the efficiency score of the decision-making unit, which can exceed 1, all other variable definitions remain consistent with the above model.

Appendix B. Tables and figures

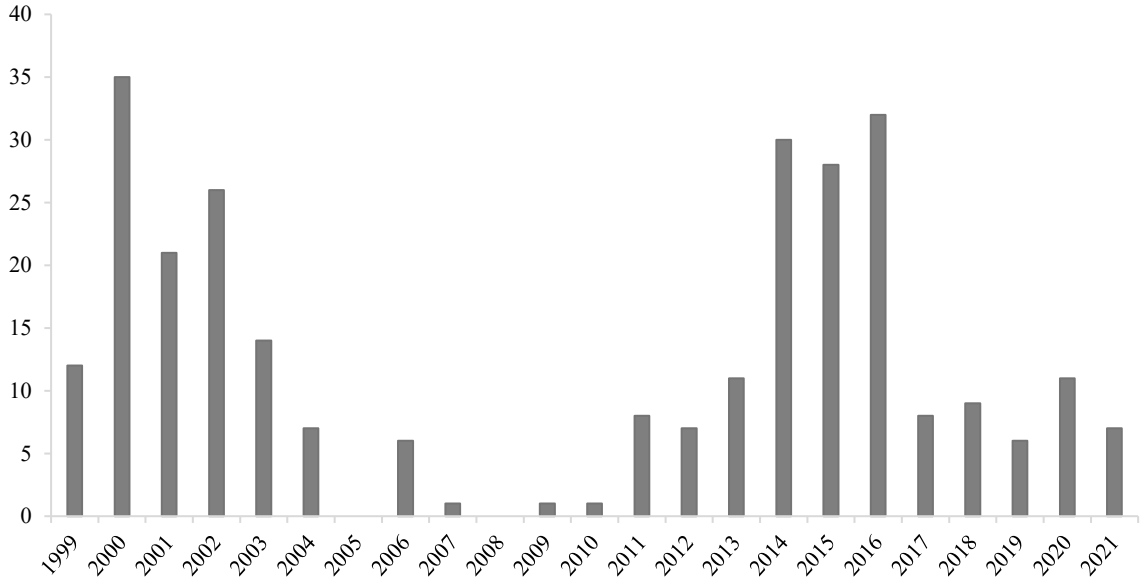

Figure A1. Number of merged counties.

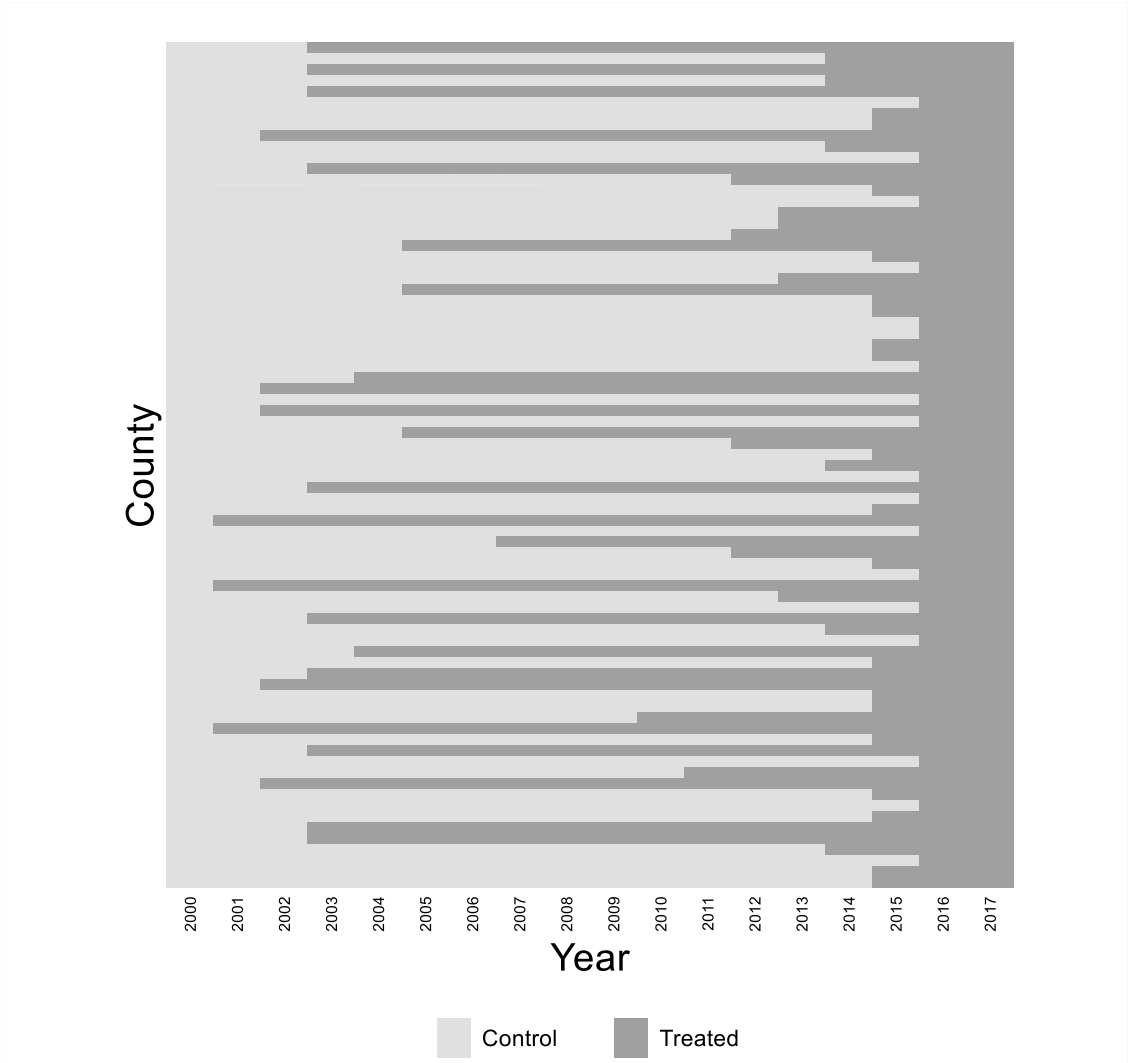

Figure A2. Processing time point diagram.

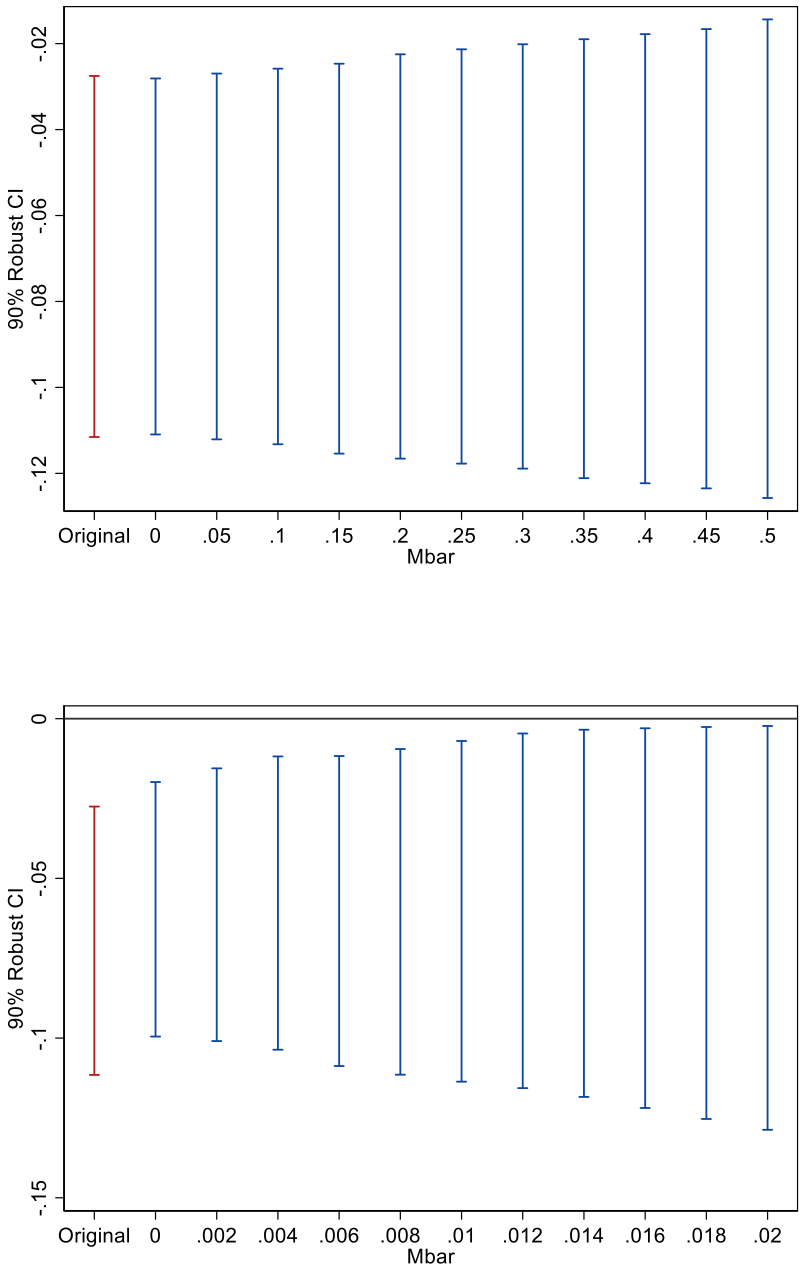

Figure A3. Parallel-trends sensitivity test.

Table A1. PSM-DID analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	lee	lee	lee	lee	lee	lee
Matching method	Nearest Neighbor				radius	kernel
	1:1	1:2	1:3	1:4		
Post	-0.1770	-0.1400	-0.1340	-0.1440	-0.1200	-0.1200
	(0.0605)	(0.0415)	(0.0368)	(0.0334)	(0.0270)	(0.0269)
Control variables	YES	YES	YES	YES	YES	YES
region effect	YES	YES	YES	YES	YES	YES
time effect	YES	YES	YES	YES	YES	YES
city*time effect	YES	YES	YES	YES	YES	YES
N	777	1,528	2,196	2,829	22,263	22,282
R^2	0.896	0.913	0.925	0.930	0.923	0.929

Note: The values enclosed in parentheses indicate county-level clustered standard errors.

Table A2. Bacon decomposition

Type of Comparison	Average treatment effect	Weights
Treatment vs. Never treated	-0.118	0.977
Earlier Treatment vs. Later Comparison	-0.122	0.012
Later Treatment vs. Earlier Comparison	-0.208	0.010

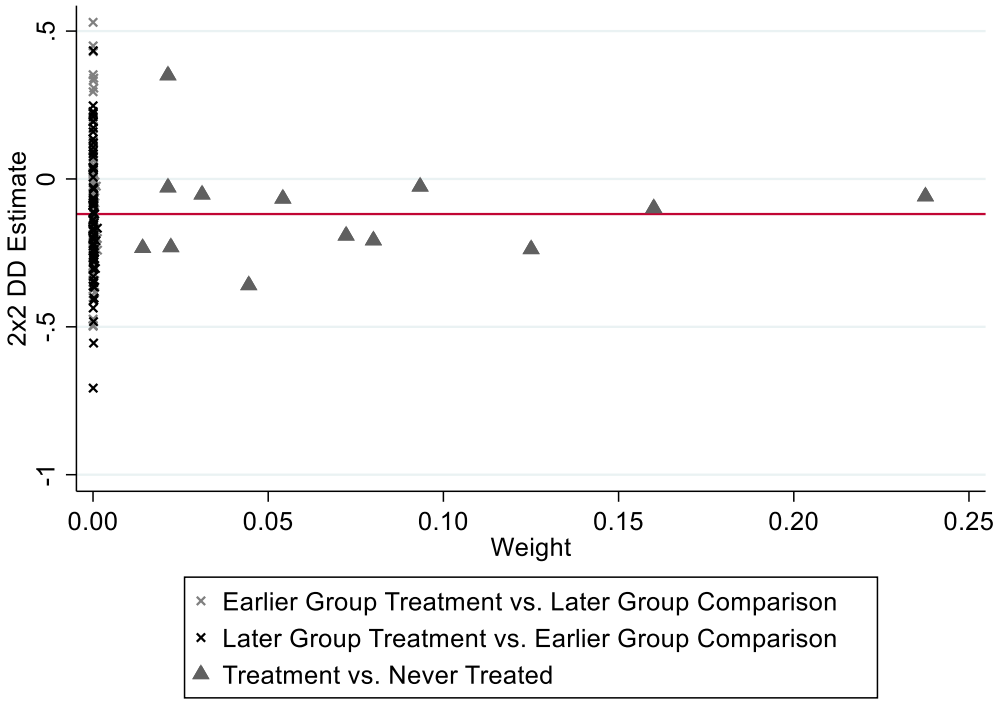

Figure A4. Bacon decomposition.

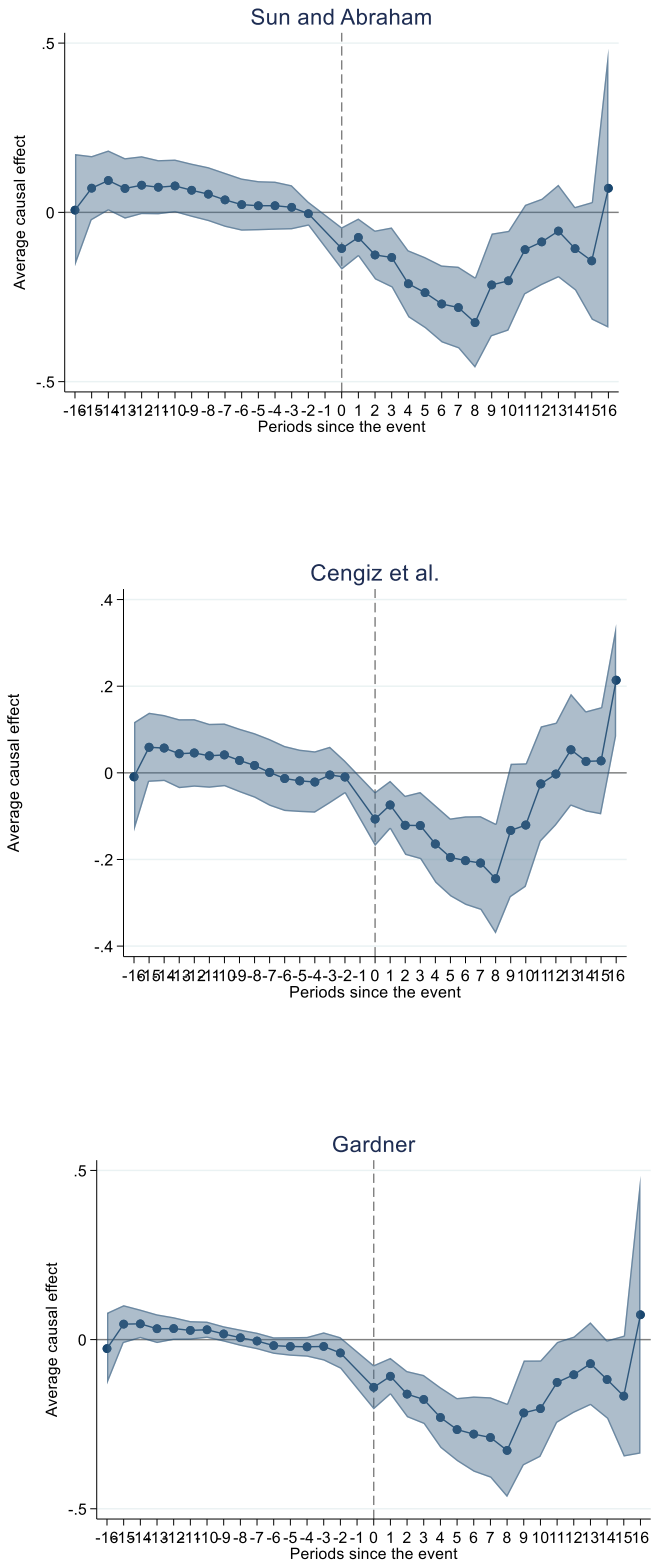

Figure A5. Three new estimation methods of staggered DID model.

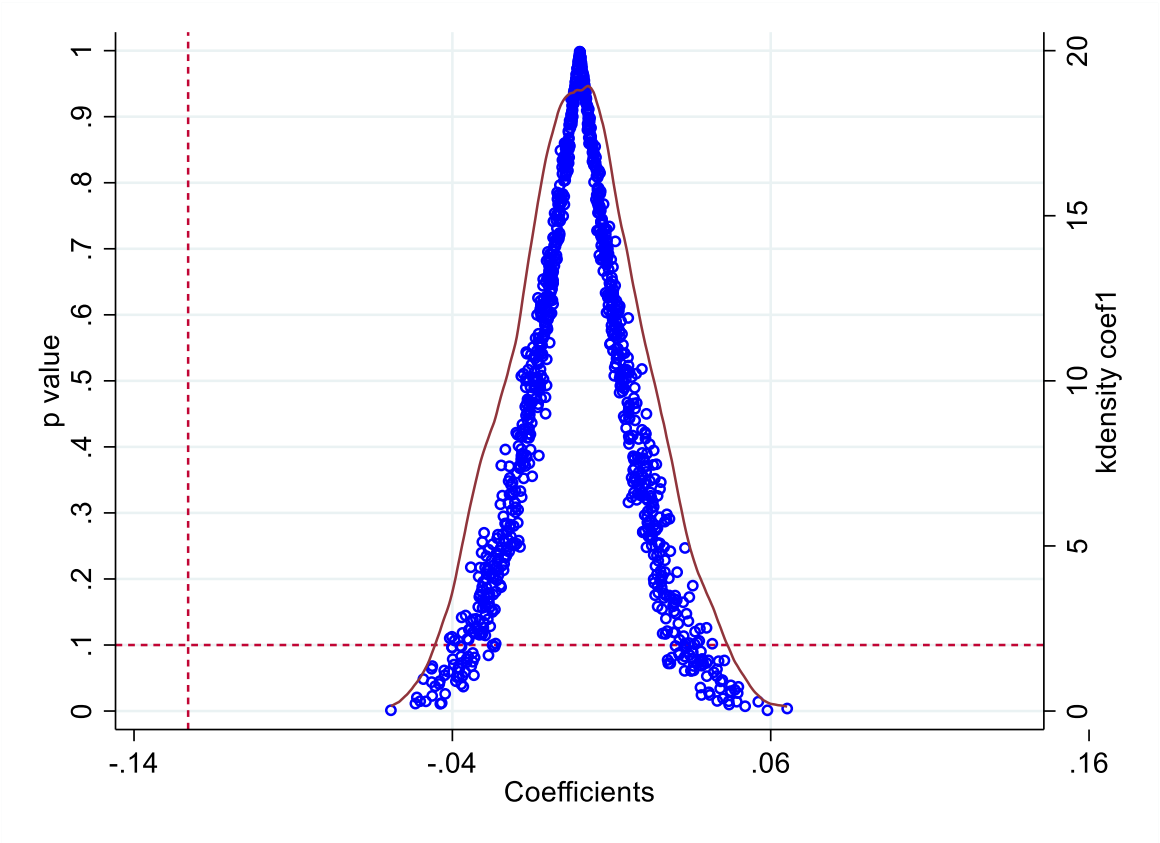

Figure A6. Placebo test repeated 1000 times.

Table A3. Consider two other policies for the same period

	(1)	(2)
	lee	lee
Post	-0.1140	-0.1120
	(0.0296)	(0.0290)
Control variables	YES	YES
region effect	YES	YES
time effect	YES	YES
city*time effect	YES	YES
N	20,925	26,775
R^2	0.913	0.916

Notes: The values enclosed in parentheses indicate county-level clustered standard errors. Column (1) presents the regression results after excluding observations from the low-carbon city pilot program, while Column (2) presents the results after excluding observations from the fiscal energy conservation and emission reduction pilot program.

Table A4. Robustness checks

	(1)	(2)	(3)	(4)
	lee	lee2	lee_w1	lee_w2
Post	-0.1230	-0.0960	-0.1210	-0.0952
	(0.0238)	(0.0187)	(0.0233)	(0.0175)
Control variables	YES	YES	YES	YES
region effect	YES	YES	YES	YES
time effect	YES	YES	YES	YES
city*time effect	YES	YES	YES	YES
CLC or CDPG	YES	NO	NO	NO
N	28,917	28,917	28,917	28,917
R^2	0.917	0.916	0.915	0.910

Notes: The values enclosed in parentheses indicate county-level clustered standard errors. Column (1) controls for a CLC/CDPG dummy to account for the stronger fiscal capacity of county-level cities (CLC) and counties directly under provincial governments (CDPG). Column (2) reports results using an alternative measure of the dependent variable. Columns (3) and (4) report results based on 1% and 5% winsorization of all variables, respectively.

Table A5. Heterogeneity analysis

	(1)	(2)	(3)	(4)
	lee	lee	lee	lee
	SER	WER	SRD	WRD
Post	-0.0633	-0.0918	-0.0904	-0.0395
	（0.0406）	(0.0289)	(0.0322)	(0.0276)
Control variables	YES	YES	YES	YES
region effect	YES	YES	YES	YES
time effect	YES	YES	YES	YES
city*time effect	YES	YES	YES	YES
N	9,433	9,473	9,030	8,776
R^2	0.912	0.949	0.937	0.913

Notes: The values enclosed in parentheses indicate county-level clustered standard errors. SER refers to samples from cities with strong environmental regulation, WER refers to samples from cities with weak environmental regulation, SRD refers to samples from cities with strong resource dependence, and WRD refers to samples from cities with weak resource dependence.