

# The Treatment Effect of the Shanxi Comprehensive Reform Area Policy on PM<sub>2.5</sub> Concentrations: A Study Based on a Quasi-Experiment

Rongxia Zhang (✉ [zhangrongxia@tyut.edu.cn](mailto:zhangrongxia@tyut.edu.cn))

Taiyuan University of Technology <https://orcid.org/0000-0003-2565-0600>

Baichuan Xiang

Taiyuan University of Technology

Wei Li

Taiyuan University of Technology

---

## Research Article

**Keywords:** the Shanxi Comprehensive Reform Area policy, PM 2.5, concentrations, PSM-DID approach, ATT, dynamic effect, mediating effect

**Posted Date:** June 14th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-503016/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

1 **The Treatment Effect of the Shanxi Comprehensive Reform Area Policy on  $PM_{2.5}$**

2 **Concentrations: A Study Based on a Quasi-Experiment**

3 Rongxia Zhang<sup>1,\*</sup>, Baichuan Xiang<sup>1</sup>, Wei Li<sup>1</sup>

4 <sup>1</sup> College of Economics and Management, Taiyuan University of Technology, Taiyuan 030024, China

5 **\*Corresponding author**

6 **Email:** zhangrongxia@tyut.edu.cn; **Tel:**86-13935156902

7

8 **Acknowledgements:** This article is funded by the Shanxi Planning Office of Philosophy and

9 Social Sciences(2019B091). The views of this paper are only from the authors and do not

10 necessarily represent the official views of the sponsors. We would like to appreciate the

11 anonymous reviewers for their valuable comments.

12 **Abstract**

13 Based on panel data on 248 prefecture-level cities in China from 2003 to 2018, this study first estimates  
14 the treatment effect of the Shanxi Comprehensive Reform Area policy on  $PM_{2.5}$  concentrations using a  
15 PSM-DID method. The empirical results show that, on average, the Shanxi Comprehensive Reform Area  
16 policy significantly increased the  $PM_{2.5}$  concentrations of prefecture-level cities in Shanxi Province by  
17 0.211% annually, and the place-based placebo test shows that the treatment effect obtained above is  
18 robust. Second, the dynamic effects show a continuous decrease in incremental effects during 2011-2018,  
19 gradually decreasing from a significant positive increment during 2011-2015 to a zero or even a negative  
20 increment during 2016-2018, indicating that the Shanxi Comprehensive Reform Area policy gradually  
21 increased in environmental friendliness. Third, the influencing mechanisms show that the Shanxi  
22 Comprehensive Reform Area policy influenced  $PM_{2.5}$  concentrations by increasing the intensity of  
23 resource exploitation and decreasing the intensity of environmental regulations, but the capacity of  
24 scientific and technological innovations had no mediating effect on the relationship between the policy  
25 and  $PM_{2.5}$  concentrations. Therefore, the government should further reduce the intensity of resource  
26 exploitation, strengthen the intensity of environmental regulations, and promote environmentally focused  
27 scientific and technological innovations to reduce  $PM_{2.5}$  concentrations in Shanxi Province.

28

29 **Keywords:** the Shanxi Comprehensive Reform Area policy;  $PM_{2.5}$  concentrations; PSM-DID approach;  
30 ATT; dynamic effect; mediating effect

31 **1 Introduction**

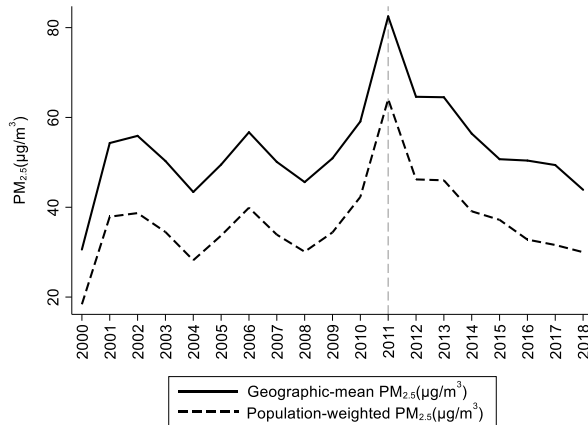
32 Since its reform and opening up, China's economy has maintained strong momentum in its growth and  
33 has had remarkable achievements. However, China's investment-driven growth has also led to serious  
34 environmental pollution (Mendoza 2017). Globally, China's environmental performance is not positive.  
35 According to the *2020 Environmental Performance Index (EPI)*<sup>1</sup> jointly released by Yale University and  
36 Columbia University, China ranks 120<sup>th</sup> out of 180 countries with an EPI score of 37.3, and Denmark  
37 ranks first with a score of 82.5 (Wendling et al. 2020).

38 As a major coal-producing area in China, Shanxi Province experiences severe environmental pollution  
39 (Li et al. 2018). According to the *2019 National Ecological Environment Quality Report* released by the  
40 Ministry of Ecology and Environment, Shanxi Province includes three of the top ten most polluted cities  
41 (Linfen, Taiyuan and Jincheng) among the 168 major cities in terms of the severity of air pollution. The  
42 primary cause of ecological and environmental problems in Shanxi Province is its coal-based economic  
43 development model that is unsustainable in the long term. To explore sustainable development paths for  
44 resource-based regions and boost the optimization and upgrading of their industrial structure, the State  
45 Council of China approved the establishment of the Shanxi National Comprehensive Supporting Reform  
46 Pilot Area for the Transformation of Resource-Based Economy (hereafter referred to as the Shanxi  
47 Comprehensive Reform Area) on December 1<sup>st</sup>, 2010, which is China's only province-wide,  
48 multidimensional and systematic comprehensive reform area for the transformation of a resource-based  
49 economy. Beginning in 2011, the work gradually began to be officially launched. On August 7<sup>th</sup>, 2012,  
50 the State Council officially approved the "Overall Plan for the Shanxi National Comprehensive  
51 Supporting Reform Pilot Area for the Transformation of Resource-Based Economy", which indicated  
52 that the construction of the Shanxi Comprehensive Reform Area had been fully implemented. Since then,  
53 the relevant national ministries and commissions and Shanxi's government departments have  
54 promulgated a series of environmental protection laws and regulations under the framework of the Shanxi  
55 Comprehensive Reform Area, such as the "Special Action Plan for the Environmental Protection of the  
56 Shanxi Comprehensive Reform Area", the "Contingency Plan for Heavily Polluted Weather in the Shanxi  
57 Comprehensive Reform Demonstration Zone" and "Air Pollution Control Measures in Shanxi Province

---

<sup>1</sup> The EPI ranks 180 countries on their environmental health and ecosystem vitality using 32 performance indicators across 11 issue categories. The 2020 rankings are based on the most recently published data, often from 2017 or 2018.

58 from 2013 to 2020”, aiming to reduce pollutant emissions and improve air quality.



59

60 **Fig. 1** Geographic-Mean  $PM_{2.5}$  and Population-Weighted  $PM_{2.5}$  in Shanxi Province<sup>2</sup>

61 As shown in Fig. 1, the geographic-mean and population-weighted  $PM_{2.5}$  concentrations in Shanxi  
62 Province reached a peak in 2011 and have trended downward thereafter. Can we infer from this that the  
63 Shanxi Comprehensive Reform Area policy raised  $PM_{2.5}$  concentrations in 2011 but lowered  $PM_{2.5}$   
64 concentrations from 2012 to 2018? The answer is no. This is because even if Shanxi Province had not  
65 been treated by the Shanxi Comprehensive Reform Area policy,  $PM_{2.5}$  concentrations would still have  
66 been influenced by numerous other factors and would not necessarily have followed their original trend.  
67 Therefore, the increase in  $PM_{2.5}$  concentrations in 2011 and the decline since 2012 were probably not  
68 caused by the Shanxi Comprehensive Reform Area policy alone, but also by other factors, and we need  
69 to disentangle the treatment effect of the Shanxi Comprehensive Reform Area policy from the effects of  
70 those other factors. The treatment effect of the Shanxi Comprehensive Reform Area policy is the  
71 difference in  $PM_{2.5}$  concentrations between the prefecture-level cities in Shanxi Province with the  
72 policy intervention and those without it. The treatment effects of a policy not only answer the questions  
73 of “does it work” or “what is the treatment effect” but also answer the questions of “when does it work,  
74 when not, and why?” (Rodrik and Rosenzweig 2010). This study aims to assess whether the Shanxi  
75 Comprehensive Reform Area policy has improved or deteriorated the air quality in Shanxi Province,  
76 when the policy has worked, and the mechanisms through which it has worked as well, which will allow  
77 for a better understanding of the treatment effects of the policy and provide evidence for future policy

<sup>2</sup> The annual mean  $PM_{2.5}$  data were retrieved from the Atmospheric Composition Analysis Group at Washington University in St. Louis (<https://sites.wustl.edu/acag/datasets/surface-pm2-5/#V4.CH.03>).

78 adjustments.

79 The rest of the paper is structured as follows. Section 2 reviews similar works in the literature. Section  
80 3 describes the methodological framework and data. Section 4 reports the effect of the Shanxi  
81 Comprehensive Reform Area policy on  $PM_{2.5}$  concentrations. Section 5 discusses the mechanisms  
82 through which the policy influences  $PM_{2.5}$  concentrations. The paper ends with conclusions and policy  
83 implications in section 6.

## 84 **2 Literature review**

85 Quasi-experiments, sometimes called nonrandomized or pre-post-intervention experiments, are studies  
86 that aim to identify the cause-and-effect relationships between an intervention and an outcome of interest  
87 but do not use randomization (Eliopoulos et al. 2004). In most cases, “intervention” refers to “treatment”.  
88 Treatment status in a quasi-experiment is determined by nature, politics, an accident, or some other action  
89 rather than by random assignment by the researchers as in randomized trials (Greenstone and Gayer  
90 2009). The data used in a quasi-experiment are observational data, not experimental data. The focus of  
91 the program evaluation in a quasi-experiment is the estimation of the average treatment effect on the  
92 treated (ATT)—computed as the posttreatment average difference between the realized and  
93 counterfactual outcomes of the treated units. However, the fundamental problem with using  
94 observational data to evaluate the ATT is that we cannot observe the counterfactual outcome, which is  
95 the posttreatment outcome of the treated unit had it not been treated. In this case, the major challenge for  
96 researchers is to estimate the missing counterfactual outcome, the outcome in the absence of treatment  
97 for the treated unit. The most widely adopted approaches developed to estimate these counterfactual  
98 values include instrumental variables (IV), regression discontinuity designs (RDD), propensity score  
99 matching (PSM), difference-in-differences (DID), PSM-DID and the synthetic control method (SCM)  
100 (Athey and Imbens 2017). The Shanxi Comprehensive Reform Area policy can be considered a quasi-  
101 experiment because the policy was decided by the government rather than chosen at random. This study  
102 estimates the treatment effects of the Shanxi Comprehensive Reform Area policy on  $PM_{2.5}$   
103 concentrations using a PSM-DID approach within the counterfactual framework.

104 Regarding the evaluation of the treatment effects of the Shanxi Comprehensive Reform Area policy,  
105 there are more theoretical and qualitative studies and fewer empirical and quantitative studies. To date,  
106 only two studies on the China National Knowledge Infrastructure (CNKI) website have studied the

107 treatment effects of the Shanxi Comprehensive Reform Area policy with quantitative methods. Guo and  
108 Guo (2019a) studied the treatment effects of the Shanxi Comprehensive Reform Area policy on the  
109 optimization of the industrial structure using SCM. Guo and Guo (2019b) investigated the treatment  
110 effects of the Shanxi Comprehensive Reform Area policy on the quality of economic development using  
111 DID. Although few studies have focused on the causal relationship between the Shanxi Comprehensive  
112 Reform Area policy and  $PM_{2.5}$  concentrations, we can review those studies that are on the treatment  
113 effects of coordinated regional development strategies since the Shanxi Comprehensive Reform Area  
114 policy is a coordinated regional development strategy.

115 In terms of coordinated regional development strategies, such as the Western Development Strategy,  
116 Jing-Jin-Ji Collaborative Development Strategy, and Northeast Revitalization Strategy, scholars have  
117 conducted various studies from the following three perspectives. First, researchers have evaluated the  
118 macro and micro effects of coordinated regional development strategies on the regions treated by the  
119 strategies. At the macro level, existing studies have used provincial-, municipal- or industrial-level panel  
120 data to estimate the average treatment effects of coordinated regional development strategies on the  
121 economy, on technological innovation, and on air pollutants, which include  $CO_2$ ,  $SO_2$  and  $PM_{2.5}$ . (Cai  
122 et al. 2017; Lu 2019; Zhang et al. 2019a; Zhang et al. 2020). At the micro level, published papers have  
123 employed micro data on industrial enterprises to estimate the average treatment effects of coordinated  
124 regional development strategies on indicators such as enterprise investment, productivity, innovation and  
125 pollutants, which include industrial wastewater, industrial  $SO_2$  and industrial smoke (dust) (Ji et al.  
126 2018; Wu et al. 2017). Second, some articles have evaluated the effects of coordinated regional  
127 development strategies in terms of the coordinated growth of East, Central and Western China (Gan et  
128 al. 2011). Third, other articles have studied the effects of coordinated regional development strategies  
129 from the perspective of intraregional differences (Peng and Chen, 2016).

130 No studies based on quantitative methods have been found that study the mechanisms that influence  
131 the relationship between the Shanxi Comprehensive Reform Area policy and environmental pollutants.  
132 Therefore, we can only review a limited number of studies on these influencing mechanisms based on  
133 coordinated regional development strategies. Ji (2020) argued that the Central Rising Strategy aggravated  
134 environmental pollution, most likely by relaxing resource exploitation restrictions, undertaking more FDI  
135 projects and increasing the share of the secondary industry in total industry. Yin and Li (2019) suggested  
136 that the Jing-Jin-Ji Collaborative Development Strategy induced the ecological development of the

137 economy through environmental regulations. Sun et al. (2019) claimed that the Western Development  
138 Strategy provided a strong impetus for economic development by improving the quality of scientific and  
139 technological innovations. In addition, other studies have suggested that the possible influencing  
140 mechanisms of such strategies included fixed asset investments, financial development, and urbanization.  
141 Therefore, because these influencing mechanisms are complex and not directly observable and it is  
142 difficult to identify influencing mechanisms through randomized controlled trials, the papers published  
143 to date have not drawn consistent conclusions. We can only test the influencing mechanisms that are  
144 likely to have a great impact on air quality according to the available literature and to economic theory  
145 on a case-by-case basis.

146 In summary, there are still shortcomings in the literature that call for further improvement. First,  
147 the literature evaluating the Shanxi Comprehensive Reform Area policy is particularly scarce, and  
148 no studies have been found on the environmental effects of this policy. However, as a province with  
149 a typical resource-based economy, Shanxi's ecological and environmental issues are of great  
150 concern to researchers and policy makers. Moreover, ecological restoration is one of the four major  
151 tasks of the Shanxi Comprehensive Reform Area policy, which has been in effect for 10 years, so it  
152 is essential to evaluate the environmental effects of this policy. Second, in terms of methodology,  
153 the published literature initially focused on qualitative as well as descriptive data analyses, and in  
154 recent years, it has focused more on counterfactual analyses based on the DID and PSM-DID  
155 methods. However, there are problems with the practical application of these methods, such as the  
156 failure to satisfy the common trends assumption in DID-based studies and the problem of treating  
157 panel data as pooled data in PSM-DID-based studies. Therefore, the application of these methods  
158 needs to be further standardized.

### 159 **3 Methodological framework and data**

#### 160 **3.1 Methodological framework**

161 This study uses the PSM-DID method proposed by Heckman et al. (1998) to estimate the counterfactual  
162 outcome values that would obtain in the absence of treatment for the treated units. Combining the  
163 advantages of PSM and DID, the PSM-DID method can mitigate selection bias and eliminate the  
164 influence of time-invariant unobservable factors on the outcome variable, thus yielding more robust  
165 estimates. The idea behind the PSM-DID method originates from a matching estimator, and the basic



166 idea is to find a real or synthetic unit in the control group whose observable covariates are as similar as  
167 possible to those of a unit in the treatment group (Zhang and Duan 2020). If the unit's selection into the  
168 treatment group is caused by observable characteristics and time-invariant unobservable characteristics  
169 and if the treated unit  $i$  and untreated unit  $j$  have similar probabilities of being treated, then the  
170 outcomes of unit  $i$  and unit  $j$  can be compared to estimate the treatment effects.

171 The steps for conducting a PSM-DID analysis are as follows (see also Fig. 2):

172 (1) Assign the units into the treatment group and the control group. In our case, 11 prefecture-level  
173 cities in Shanxi Province are assigned to the treatment group, and 237 prefecture-level cities in other  
174 provinces are assigned to the control group.

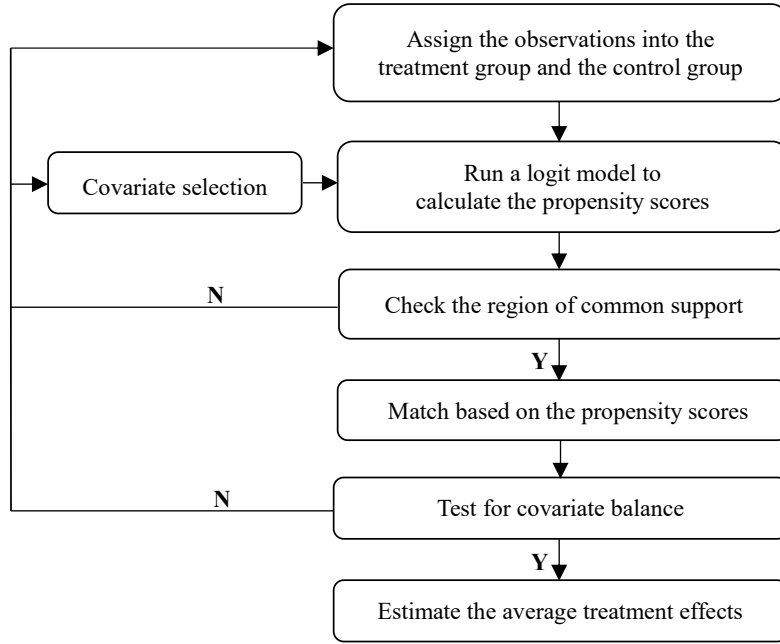
175 (2) Run a logit model to calculate the propensity score for each unit assigned to the treatment group  
176 conditional on all observable covariates that jointly affect assignment to treatment and the outcome of  
177 interest.

178 (3) Check the region of common support, which is the interval over which the propensity scores of  
179 the treatment and the control groups overlap. A large region of common support is generally associated  
180 with a higher probability of achieving better quality matches (Li et al. 2021). If this condition is not  
181 satisfied, the covariates included in the logit model and the units in the control group should be reselected.

182 (4) Match the units from the treatment group with those from the control group based on their  
183 propensity scores. The most commonly used matching methods in the literature are kernel matching and  
184 nearest neighbor matching within a caliper.

185 (5) A balance test is performed to check whether the treated units and the untreated units are  
186 statistically similar in their covariates after matching. If significant differences in covariate values are  
187 found, then the logit model should be respecified and the process should be restarted from step (1).

188 (6) Estimate the average treatment effects based on the differences between each pair of matched units.



**Fig. 2** Steps for conducting a PSM-DID analysis

189

190 In this study, PSM is done first to exclude systematic differences between the treatment group and the  
 191 control group in order to mitigate sample selection bias due to observables; then, DID is performed using  
 192 the treated and matched control units to eliminate the influence of time-invariant unobservable factors  
 193 on  $PM_{2.5}$  concentrations. We address the issue of time-variant confounders by selecting control units  
 194 that are more similar to the units in the treatment group. The benchmark model in this paper is built  
 195 according to the standard DID model as follows:

$$196 \quad Y_{it} = \alpha_0 + \alpha_1 Treated_{it} * Period_{it} + \alpha_2 Treated_{it} + \alpha_3 Period_{it} + \alpha_4 X_{it} + \varepsilon_{it} \quad (1)$$

197 where  $Y_{it}$  is the logarithm of the geographic-mean  $PM_{2.5}$  concentrations for city  $i$  in year  $t$ .  
 198  $Treated_{it}$  is the treatment dummy: if unit  $i \in$  treatment group,  $Treated_{it} = 1$ ; otherwise,  
 199  $Treated_{it} = 0$ . Since the Shanxi Comprehensive Reform Area policy was approved on December 1<sup>st</sup>,  
 200 2010, officially launched in 2011 and fully implemented in 2012, we consider the policy implementation  
 201 year to be 2011.  $Period_{it}$  is the time dummy: if  $t \in (2011, 2018)$ ,  $Period_{it} = 1$ ; otherwise,  
 202  $Period_{it} = 0$ .  $Treated_{it} * Period_{it}$  is the interaction term, and its coefficient is the DID estimator we  
 203 are interested in, which is the ATT, indicating the average degree to which the treated cities have reduced  
 204 or increased their  $PM_{2.5}$  concentrations relative to untreated cities due to the Shanxi Comprehensive  
 205 Reform Area policy (Zhang et al. 2019b).  $X_{it}$  represents the observable time-variant covariates that can  
 206 provide additional explanatory information.  $\varepsilon_{it}$  is an idiosyncratic error term.

207 The DID estimator is illustrated in Table 1. Equation (1) shows that for treated cities in Shanxi  
 208 Province, the  $PM_{2.5}$  concentrations before and after the policy change are  $\alpha_0 + \alpha_2 + \alpha_4 X_{it_0}$  and  $\alpha_0 +$   
 209  $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 X_{it_1}$ ,<sup>3</sup> respectively, and the difference between them,  $\Delta Y_t = \alpha_1 + \alpha_3 + \alpha_4 (X_{it_1} -$   
 210  $X_{it_0})$ , is the change in  $PM_{2.5}$  concentrations, which represents the combined treatment effects of the  
 211 Shanxi Comprehensive Reform Area policy and of other external factors. Similarly, for matched  
 212 untreated cities in the control group, the change in  $PM_{2.5}$  concentrations before and after the policy  
 213 change is  $\Delta Y_c = \alpha_3 + \alpha_4 (X_{it_1} - X_{it_0})$ , which represents the treatment effect of other external factors.  
 214 Therefore, the difference between  $\Delta Y_t$  and  $\Delta Y_c$ ,  $\Delta \Delta Y = \alpha_1$ , is the net treatment effect of the Shanxi  
 215 Comprehensive Reform Area policy. If  $\alpha_1 > 0$ , then the Shanxi Comprehensive Reform Area policy has  
 216 caused the air quality of the treated cities to deteriorate; otherwise, it has improved their air quality.

217 **Table 1** Illustration of the DID Estimator

	Before ( $Period_{it}=0$ )	After ( $Period_{it}=1$ )	After–Before
Treatment ( $Treated_{it} = 1$ )	$\alpha_0 + \alpha_2$ $+ \alpha_4 X_{it_0}$	$\alpha_0 + \alpha_1 + \alpha_2$ $+ \alpha_3 + \alpha_4 X_{it_1}$	$\Delta Y_t = \alpha_1 + \alpha_3$ $+ \alpha_4 (X_{it_1} - X_{it_0})$
Control ( $Treated_{it} = 0$ )	$\alpha_0 + \alpha_4 X_{it_0}$	$\alpha_0 + \alpha_3$ $+ \alpha_4 X_{it_1}$	$\Delta Y_c = \alpha_3 + \alpha_4 (X_{it_1} - X_{it_0})$
Treatment–Control	$\alpha_2$	$\alpha_1 + \alpha_2$	$\Delta \Delta Y = \alpha_1$

218

### 219 3.2 Covariate selection

220 The observable covariates were selected on the basis of the published literature following the stepwise  
 221 procedure proposed by Imbens and Rubin (2015)<sup>4</sup> and are as follows. (1) The economic development  
 222 level: Scholars have conducted research on the impact of economic development on haze pollution and  
 223 have verified that the relationship between these features is an inverted U-shape (Hishan et al. 2019;  
 224 Dong et al. 2018). This study uses GDP per capita to represent the economic development level. (2) The  
 225 scale of fiscal expenditure: Recent studies have found that fiscal expenditure is an important determinant  
 226 of environmental quality, and the net impact of fiscal expenditure on the environment is ambiguous and  
 227 may depend on the composition of the fiscal expenditure, the specific pollutant and the country or region

<sup>3</sup>  $X_{it_0}$  and  $X_{it_1}$  represent the values of the covariates before and after treatment, respectively.

<sup>4</sup> Imbens and Rubin (2015) proposed a stepwise procedure with three stages to select covariates for a model that estimates propensity scores. First, a set of basic covariates are included linearly based on economic theory; second, other covariates are also included linearly and the likelihood ratio test is used to justify the newly included covariates. The third step is to decide which of the interactions and quadratic terms for the selected variates should be included in the model specification.

228 being investigated (Zhang et al. 2017; Halkos and Paizanos 2013). The ratio of the local governments'  
229 general public expenditure to GDP is used to measure the scale of fiscal expenditure in our study. (3) The  
230 industrial structure: The industrial structure, as an important force driving economic development, is  
231 undoubtedly also an important factor influencing environmental pollution. This study uses the share of  
232 the secondary industry in GDP to characterize the industrial structure. (4) The informatization level:  
233 Information technology helps optimize and upgrade the industrial structure by embedding more efficient  
234 production technologies into all types of production activities (You and Chi 2012), thus reducing  
235 emissions through technology. This study uses the ratio of total postal and telecommunication services  
236 per capita to GDP per capita to represent the informatization level, following Yuan and Zhu (2018). (5)  
237 Human capital: Environmental issues are caused by humans, and the literature agrees that human capital  
238 can represent individual awareness of pollution and its harmful effects. However, it is hard to define and  
239 measure human capital precisely, and the literature often quantifies human capital with educational  
240 attainment (Lan et al. 2012). This study uses the percentage of college students in the total local  
241 population to represent human capital levels. (6) Population agglomeration: Researchers generally agree  
242 that in developed countries, population agglomeration can reduce air pollution. However, in the case of  
243 China, the results are mixed due to issues with data or with the econometric methods used, or to other  
244 issues. Therefore, further research needs to be conducted (Chen et al. 2020). Population density is  
245 employed to represent population agglomeration in this study. (7) The foreign direct investment (FDI)  
246 level: The pollution haven hypothesis tells us that developed countries take advantage of the lower labor  
247 costs and lower environmental standards of developing countries to transfer pollution-intensive  
248 industries to those developing countries through FDI (Lan et al. 2012). The literature suggests that FDI  
249 influences environmental quality through two main channels: economic growth and technological  
250 progress. There is no unanimous conclusion on the effects of FDI on the environment, and its effects are  
251 related to the specific context of the analysis (Pazienza 2014). The FDI level is represented by the ratio  
252 of the product of USD-denominated FDI and the exchange rate to GDP. The covariates included in this  
253 study and the corresponding formulas are shown in Table 2.

254

255 **Table 2** Covariates and their computational formulas

Covariates	Symbols	Computational formulas
------------	---------	------------------------

The economic development level	gdpp	GDP per capita
The scale of fiscal expenditure	locbudper	Local governments' general public expenditure/GDP
The industrial structure	secindper	The value-added of the secondary industry/GDP
The informatization level	inform	Postal and telecommunication services per capita/GDP per capita
Human capital	hiedustuper	Number of college students/total population
Population agglomeration	popdensity	population density
The foreign direct investment level	fdiper	USD-denominated FDI*exchange rate/GDP

### 256 3.3 Data

257 This study uses panel data on 248 prefecture-level cities in China (excluding Beijing, Tianjin,  
258 Shanghai, Chongqing, nonresource-based cities in southern China and cities with missing data) from  
259 2003 to 2018 to evaluate the treatment effect of the Shanxi Comprehensive Reform Area policy on  
260  $PM_{2.5}$  concentrations. The  $PM_{2.5}$  concentration data are from the Atmosphere Composition  
261 Analysis Group of Dalhousie University in Canada, and other data are from the 2004-2019 *China*  
262 *City Statistical Yearbooks* and the Statistical Communiques on National Economic and Social  
263 Development of the prefecture-level cities. This study uses a double logarithmic model; that is, the  
264  $PM_{2.5}$  concentrations and all covariates are in logarithmic form (except for  $Treated_{it} * Period_{it}$ ,  
265  $Treated_{it}$  and  $Period_{it}$ ), because logarithmic transformations can reduce heteroskedasticity and  
266 the resulting coefficients represent elasticities.

## 267 4 The treatment effect of the Shanxi Comprehensive Reform Area policy on $PM_{2.5}$ concentrations

### 268 4.1 Propensity score matching

269 In nonrandomized observational studies, units are not randomly assigned to the treatment or control  
270 groups. In other words, the characteristics of the units in the treatment group differ systematically  
271 from those in the control group, and those characteristics also have a bearing on the incidence of  
272 selection bias (Li et al. 2013). Therefore, significant differences between the two groups may exist,  
273 and the differences observed after implementation of the policy cannot be attributed to the policy  
274 only. Hence, we construct a matched control group that is similar to the treatment group, and the  
275 PSM approach provides a method for balancing the two groups by matching the treatment and  
276 control units according to a set of covariates, the purpose of which is to eliminate selection bias and

277 ensure that the two groups are comparable.

#### 278 4.1.1 Estimating the propensity score

279 PSM reduces the matching criteria from multiple dimensions to a single dimension, the propensity score.

280 Therefore, treated and untreated units with similar propensity scores can be compared to obtain treatment

281 effects (Li et al. 2013). Before matching, a binary logit model needs to be constructed to estimate the

282 conditional probability of assignment to the treatment group given a vector of observable covariates:

$$283 \quad Treated_{it} = \gamma_0 + \gamma_1 X_{it} + \mu_{it} \quad (2)$$

284 where  $Treated_{it} = 1$  if city  $i$  is treated by the Shanxi Comprehensive Reform Area policy and

285  $Treated_{it} = 0$  otherwise.  $\gamma_0$  is the intercept term.  $\gamma_1$  is the vector of regression coefficients.  $X_{it}$

286 is a vector of the observable covariates selected above.  $\mu_{it}$  is the stochastic error term.

287 The estimated results of the logit propensity score model can be found in Table 3, which shows

288 that almost all covariates play a significant role in the city's selection into the treatment group.

289 **Table 3** Estimation of propensity scores (logit)

Variables	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
lngdpp	-0.420***	0.155	-2.71	0.007	-0.723	-0.116
lnlocbudper	0.794***	0.148	5.37	0.000	0.504	1.085
lnsecindper	5.444***	0.526	10.34	0.000	4.413	6.476
lninform	1.429***	0.132	10.81	0.000	1.170	1.689
lnhiedustuper	0.104	0.085	1.23	0.220	-0.062	0.271
lnpopdensity	-0.479***	0.097	-4.93	0.000	-0.669	-0.289
lnfdiper	-0.183***	0.061	-2.99	0.003	-0.304	-0.063
Constant	-11.175***	1.962	-5.70	0.000	-15.020	-7.330
Pseudo $R^2$	0.1791	Number of obs	3968			

290 Note: The dependent variable, Treated, is treatment status, which equals one if a city belongs to the

291 treatment group and zero if it belongs to the control group. Standard errors are clustered at the city level.

292 \*, \*\* and \*\*\* indicate significance at the 10%, 5% and 1% levels, respectively.

#### 293 4.1.2 Verifying the common support assumption

294 After estimating the propensity score, the next step is to ensure that the region of common support or

295 overlap in propensity scores across the treated and untreated cities is sizable because the matching results

296 are considered to be accurate only when a larger share of the observations satisfies the common support

297 assumption (Heckman and Vytlačil, 2001). The region of common support may be different for different

298 matching methods. Tables 4 and 5 show that for kernel matching and nearest neighbor matching within  
 299 a caliper, 102 and 53 observations do not satisfy the common support assumption, and 3866 and 3915  
 300 observations do satisfy the common support assumption, respectively. Both matching algorithms have a  
 301 small amount of sample loss, and the matching results are satisfactory.

302 **Table 4** Common support test (kernel matching)

Treatment assignment	Off support	On support	Total
Untreated	100	3,692	3,792
Treated	2	174	176
Total	102	3,866	3,968

303 **Table 5** Common support test (nearest neighbor matching within a caliper)

Treatment assignment	Off support	On support	Total
Untreated	37	3755	3792
Treated	16	160	176
Total	53	3915	3968

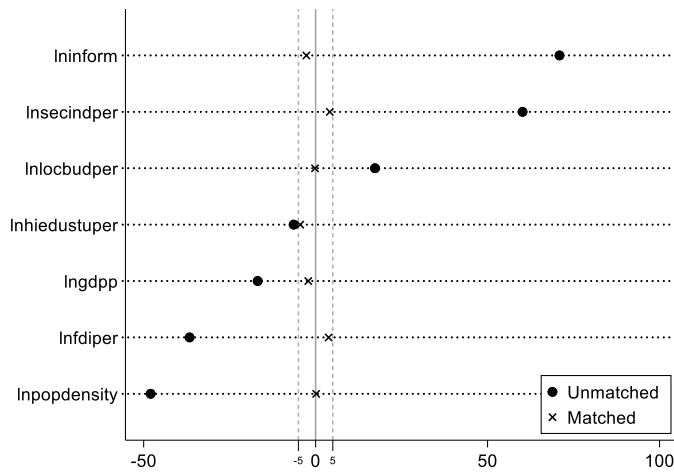
#### 304 4.1.3 Selecting the matching algorithm

305 After verifying the common support assumption, we should select a matching algorithm to choose control  
 306 cities to act as the counterfactuals for the treated cities. Theoretically, there are multiple algorithms that  
 307 perform PSM, each with asymptotically equivalent matching results. This study first matches the cities  
 308 using the kernel matching algorithm, which uses the weighted average of each dimensional feature of  
 309 several untreated cities to construct a counterfactual city, with weights inversely proportional to the  
 310 distance between the treated and untreated cities. Meanwhile, to ensure the robustness of the empirical  
 311 results, this paper also employs 1:1 nearest neighbor matching without replacement within a caliper. With  
 312 this algorithm, treated city  $i$  is matched to the one untreated city that is closest to city  $i$  in terms of  
 313 propensity score within a predefined caliper width of 0.01.

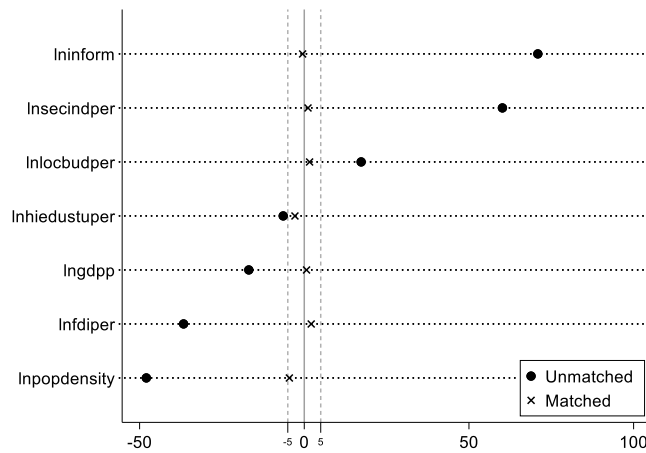
#### 314 4.1.4 Testing the balance condition

315 After the matching procedure is completed, a balance test should be conducted to test whether the  
 316 covariates of the two groups are statistically similar. If significant differences exist, the logit model  
 317 should be respecified, and the process should be repeated from the beginning. Figs. 3 and 4 show the  
 318 results of the balance tests for kernel matching and nearest neighbor matching within a caliper,  
 319 respectively. The results indicate that the absolute values of the standardized percentage biases across the  
 320 matched covariates are less than 5%, and the p-values for all covariates after matching are greater than

321 0.1, indicating that no significant differences exist between the covariates of the treatment and control  
 322 groups after matching. Furthermore, the pseudo R-squared value decreases from 0.178 to 0.001,  
 323 indicating that the after-matching logit model has no explanatory power. Hence, the characteristics of the  
 324 two groups are already very similar after matching, satisfying the balance condition for PSM  
 325 (Rosenbaum and Rubin 1985).



**Fig. 3** Standardized percentage bias across covariates (kernel matching, kernel type=epanechnikov)



**Fig. 4** Standardized percentage bias across covariates (nearest neighbor matching within a caliper)

326 **4.2 The results of the PSM-DID model**

327 The common support test and the balance test show that the PSM results are valid; thus, the DID  
 328 regression from equation (1) can be performed with the matched samples. For comparative purposes, we  
 329 also list the results of the OLS and DID regressions (see Table 6). Column (1) of Table 6 shows the results  
 330 of the OLS regression, for which the coefficient on the variable  $Treated_{it}$  is 0.077, indicating that the  
 331 Shanxi Comprehensive Reform Area policy has a significant positive effect on  $PM_{2.5}$  concentrations in



332 Shanxi Province. However, this estimate is biased by unobservable characteristics associated with  
333 treatment assignment and  $PM_{2.5}$  concentrations. Column (2) shows the results of the DID method,  
334 where the coefficient on the interaction term  $Treated_{it} * Period_{it}$  is significantly positive. However,  
335 the DID model is unlikely to satisfy the common trend assumption, so the estimate is still biased. The  
336 PSM-DID method alleviates the shortcomings of the previous two methods and generates more accurate  
337 results. The coefficients on the interaction term for the kernel matching and the nearest neighbor  
338 matching within a caliper in columns (3) and (4) are 0.211 and 0.223, respectively, both of which are  
339 significant at the 1% level, indicating that the implementation of Shanxi Comprehensive Reform Area  
340 policy resulted in an average annual increase of 0.211% and 0.223% in  $PM_{2.5}$  concentrations during  
341 2011-2018. Therefore, on average, instead of reducing  $PM_{2.5}$  concentrations, the Shanxi  
342 Comprehensive Reform Area policy significantly increased  $PM_{2.5}$  concentrations and deteriorated the  
343 air quality in Shanxi Province. The results also show that the OLS and DID regressions underestimate  
344 the treatment effect.

345

346 **Table 6** Regression results for the benchmark model

Variables	OLS	DID	PSM-DID	
	(1)	(2)	(3) Kernel	(4) NN
$Treated_{it} * Period_{it}$		0.206*** (9.87)	0.211*** (9.45)	0.223*** (11.17)
$Treated_{it}$	0.077*** (3.24)	-0.089 (-1.37)	-0.095 (-1.47)	-0.100 (-1.49)
$Period_{it}$		0.029** (2.45)	0.033*** (3.01)	0.031*** (2.59)
lnGDP	-0.188*** (-18.51)	-0.154*** (-10.47)	-0.166*** (-10.65)	-0.157*** (-10.49)
lnlocbudper	0.047*** (4.71)	0.099*** (10.57)	0.106*** (10.85)	0.099*** (10.71)
lnsecindper	0.423*** (15.72)	0.333*** (7.16)	0.357*** (7.72)	0.335*** (7.15)
lninform	-0.158*** (-14.04)	-0.027*** (-3.07)	-0.023** (-2.54)	-0.028*** (-2.92)
lnhiedustuper	0.033*** (5.31)	-0.025* (-1.81)	-0.024* (-1.66)	-0.024* (-1.75)
lnpopdensity	0.379*** (31.90)	0.344*** (9.09)	0.351*** (9.05)	0.344*** (9.09)

Infdi per	0.031*** (6.70)	0.025*** (5.42)	0.023*** (4.81)	0.024*** (5.06)
Constant	1.12*** (9.57)	2.167*** (9.00)	2.168*** (8.49)	2.177*** (8.96)
R-squared	0.562	0.522	0.508	0.523
Number of obs	3968	3968	3866	3915

347 Note: Cluster-robust standard errors are calculated for all regressions. Columns (3) and (4) represent the  
348 regression results obtained by DID after kernel matching and nearest neighbor matching within a caliper,  
349 respectively.

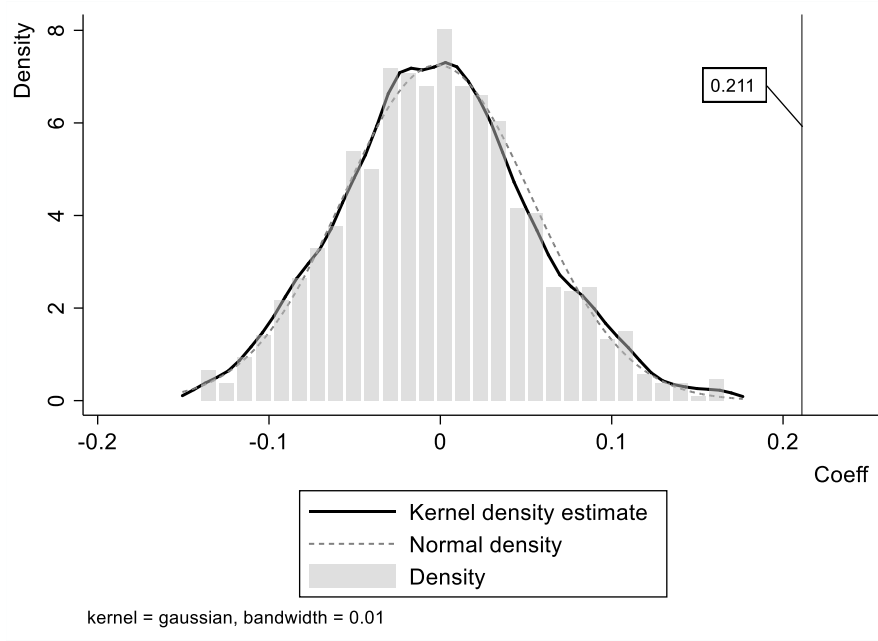
350 The results of the PSM-DID model also suggest the following (see Table 6): (1) Economic  
351 development reduced  $PM_{2.5}$  concentrations significantly, which is mainly because the pattern of  
352 economic development was transformed from extensive economic growth with high energy consumption  
353 and high pollution to intensive economic growth with low energy consumption and low pollution. (2)  
354 Fiscal expenditures significantly increased  $PM_{2.5}$  concentrations, mainly because the effect of  
355 environmentally unfriendly expenditures, such as those on fixed assets, exceeded that of environmentally  
356 friendly expenditures, such as those on environmental protection and education. (3) The share of the  
357 secondary industry in GDP is positively and significantly correlated with  $PM_{2.5}$  concentrations. This  
358 result is consistent with those of much of the literature, since the secondary industry, including industry  
359 and construction, consumes large amounts of energy and generates large amounts of pollution. (4)  
360 Informatization significantly improved air quality mainly because improved informatization indicates  
361 more efficient production technology, which generates less pollution in the production of the same  
362 amount of output. (5) The percentage of college students in the local population reduced  $PM_{2.5}$   
363 concentrations at a significance level of 10% because education levels represent environmental  
364 awareness. (6) An increase in population density significantly increased  $PM_{2.5}$  concentrations and  
365 worsened air quality, primarily because an increase in population density represents an increase in the  
366 demand for transportation, housing and energy consumption. (7) FDI significantly increased  $PM_{2.5}$   
367 concentrations, mainly because FDI in Shanxi Province is concentrated in the traditional manufacturing  
368 industry, which consumes a considerable amount of energy and causes serious pollution.

### 369 **4.3 Placebo test**

370 To ensure the robustness of the above results, this study conducts a place-based placebo test following  
371 Li et al. (2016) and Cantoni et al. (2017). First, 11 cities are randomly selected as the false treatment  
372 group from the total sample of 248 cities, and the other cities are selected as the control group. Second,

373 the value of the variable  $Treated_{it}$  is adjusted accordingly, and the PSM-DID regression is estimated.  
 374 Third, the above steps are repeated 1000 times, and then the robustness of the regression is evaluated  
 375 based on the significance and distribution of the 1000 coefficients on the interaction term  $Treated_{it} *$   
 376  $Period_{it}$ .

377 The density distribution of the estimated coefficients on the interaction term  $Treated_{it} * Period_{it}$   
 378 is approximately normally distributed (see Fig. 5), and the estimated coefficients are distributed around  
 379 0, with a mean value of -0.0005. In the benchmark model, the coefficient on the interaction term for the  
 380 sample matched with the kernel matching method is 0.211 (as shown by the vertical line in Fig. 5), which  
 381 is far from the plotted coefficient distribution. Therefore, the  $PM_{2.5}$  concentrations in the cities included  
 382 in the false treatment group were not affected by the Shanxi Comprehensive Reform Area policy. Hence,  
 383 the significant positive effect of the Shanxi Comprehensive Reform Area policy on the  $PM_{2.5}$   
 384 concentrations of the 11 prefecture-level cities in Shanxi Province was not driven by other unobservable  
 385 factors but by the Shanxi Comprehensive Reform Area policy itself; thus, the results of the benchmark  
 386 model are very robust.



387  
 388 **Fig. 5** The distribution of 1000 coefficient estimates

389 **4.4 Dynamic effects tests**

390 The findings above show that the implementation of the Shanxi Comprehensive Reform Area policy did  
 391 deteriorate the air quality in Shanxi Province, but the findings only present the average effect and not the  
 392 effects in different years. However, along with the implementation of the Shanxi Comprehensive Reform

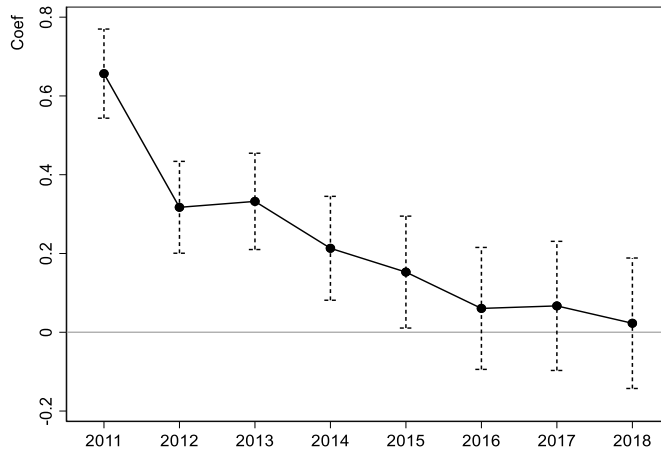
393 Area policy, supporting policies have gradually improved and have been more efficiently implemented,  
 394 so the effects of the Shanxi Comprehensive Reform Area policy on  $PM_{2.5}$  concentrations are expected  
 395 to have changed over time. Therefore, we examine the dynamic effects of the Shanxi Comprehensive  
 396 Reform Area policy on  $PM_{2.5}$  concentrations, and the results are shown in Figs. 5 and 6, where the black  
 397 dots represent the coefficients on the interaction term  $Treated_{it} * Period_{it}$  for each year, and the short  
 398 dashed lines above and below the black dots represent the upper and lower bounds of the 95% confidence  
 399 interval for each coefficient.

400 Fig. 6 presents the coefficients on the interaction term  $Treated_{it} * Period_{it}$  from 2011 to 2018  
 401 without controlling for covariates. The coefficients basically show a decreasing trend and are  
 402 significantly positive from 2011 to 2015 and significantly negative from 2016 to 2018. Fig. 7 shows the  
 403 coefficients on the interaction term  $Treated_{it} * Period_{it}$  from 2011 to 2018 after controlling for  
 404 covariates, which show the same decreasing trend as the coefficients without controlling for covariates  
 405 with the difference that the coefficients are significantly positive from 2011 to 2015 and positive but not  
 406 significant from 2016 to 2018. The results of Figs. 6 and 7 both indicate that the Shanxi Comprehensive  
 407 Reform Area policy reduced the rate at which  $PM_{2.5}$  concentrations increased from 2011 to 2015 and  
 408 significantly reduced  $PM_{2.5}$  concentrations or had an insignificant effect on  $PM_{2.5}$  concentrations  
 409 from 2016 to 2018. Therefore, although the PSM-DID results show that the average treatment effect is  
 410 significantly positive, the dynamic effects test results show that the treatment effect gradually  
 411 transformed from a decreasing increment to an insignificant or even negative effect, and we can conclude  
 412 that the Shanxi Comprehensive Reform Area policy gradually increased in environmental friendliness.



413  
 414

**Fig. 6** Trend in coefficients without covariates



415  
416 **Fig. 7** Trend in coefficients with covariates

417 **5 The mechanisms influencing the effect of the Shanxi Comprehensive Reform Area policy on**  
418  **$PM_{2.5}$  concentrations**

419 **5.1 Mediator variables**

420 By referring to the literature on the mechanisms that influence the effect of coordinated regional  
421 development strategies on the environment in resource-based regions, we summarize the main  
422 influencing mechanisms as follows: (1) First, coordinated regional development strategies may relax  
423 resource exploitation restrictions and increase the intensity of resource exploitation, thereby deteriorating  
424 environmental quality; (2) second, within the framework of a coordinated regional development strategy,  
425 a series of environmental regulations are formulated and implemented to reduce environmental pollution  
426 and resource depletion during the process of economic development; and (3) third, coordinated regional  
427 development mainly relies on scientific and technological innovation, which reduces environmental  
428 pollution and promotes the construction of an ecological civilization. Therefore, we focus on three  
429 mediator variables: the intensity of resource exploitation, the intensity of environmental regulations, and  
430 the capacity of scientific and technological innovation.

431 **5.2 The causal steps procedure for testing mediation**

432 The causal steps procedure for testing mediation popularized by Baron and Kenny (1986) is used to test  
433 whether the intensity of resource exploitation, the intensity of environmental regulations, or the capacity  
434 of scientific and technological innovation play a mediating role in the relationship between the Shanxi  
435 Comprehensive Reform Area policy and  $PM_{2.5}$  concentrations. Mediation is tested through three  
436 regressions as follows.

437 Step 1: The logarithm of geographic-mean  $PM_{2.5}$  concentrations is regressed on  $Treated_{it} *$   
438  $Period_{it}$ , which signifies that the Shanxi Comprehensive Reform Area policy is in effect, and the  
439 regression equation is given by Eq. (1) above.

440 Step 2: The mediator is regressed on the Shanxi Comprehensive Reform Area policy:

$$441 \quad Mediator_{it} = \beta_0 + \beta_1 Treated_{it} * Period_{it} + \beta_2 Treated_{it} + \beta_3 Period_{it} + \beta_4 X_{it} + \varepsilon_{it} \quad (3)$$

442 Step 3: The logarithm of geographic-mean  $PM_{2.5}$  concentrations is regressed on both the Shanxi  
443 Comprehensive Reform Area policy and the mediator:

$$444 \quad Y_{it} = \delta_0 + \delta_1 Treated_{it} * Period_{it} + \theta Mediator_{it} + \delta_2 Treated_{it} + \delta_3 Period_{it} + \delta_4 X_{it} + \varepsilon_{it} \quad (4)$$

445 The results from Eqs. (1)-(3) are used to evaluate the following conditions for the intensity of resource  
446 exploitation, the intensity of environmental regulations, and the capacity of scientific and technological  
447 innovation as mediators of the relationship between the Shanxi Comprehensive Reform Area policy and  
448  $PM_{2.5}$  concentrations: (a) the Shanxi Comprehensive Reform Area policy should be related to the  
449 logarithm of  $PM_{2.5}$  concentrations in equation (1) such that  $\alpha_1$  is significant; (b) the Shanxi  
450 Comprehensive Reform Area policy should be related to the mediator in Equation (3) such that  $\beta_1$  is  
451 significant; (c) the mediator should be related to the logarithm of  $PM_{2.5}$  concentrations in Equation (3)  
452 such that  $\theta$  is significant; and (d) the relation between the Shanxi Comprehensive Reform Area policy  
453 and the logarithm of  $PM_{2.5}$  concentrations in Equation (4) (i.e.,  $\delta_1$ ) should be insignificant or  
454 significant but smaller than the relationship between the Shanxi Comprehensive Reform Area policy and  
455 the logarithm of  $PM_{2.5}$  concentrations in Equation (1) (i.e.,  $\alpha_1$ ). Assuming that the first 3 conditions  
456 are satisfied, complete mediation is inferred if  $\delta_1$  is insignificant, while partial mediation is inferred if  
457  $\delta_1$  remains significant but is significantly smaller than  $\alpha_1$  (Edwards and Lambert, 2007). If at least one  
458 of the coefficients  $\beta_1$  or  $\theta$  is insignificant, a Sobel or bootstrap test is required to determine whether  
459 the mediator variable has a mediating effect.

### 460 **5.3 The mediating effects**

#### 461 5.3.1 The intensity of resource exploitation

462 We use the share of mining employment in total urban employment to represent the intensity of resource  
463 exploitation, which increased from 18% in 2011 to 23% in 2018 in Shanxi Province. The results of the  
464 mediating effects test are shown in column (3) of Table 6 and columns (1) and (2) of Table 7. The results  
465 show that  $\alpha_1$  in equation (1) is 0.211 and significant at the 1% level for the kernel matching algorithm,

466 indicating that the Shanxi Comprehensive Reform Area policy significantly increased  $PM_{2.5}$   
 467 concentrations.  $\beta_1$  in equation (3) is 0.677 and significant at the 1% level, which suggests that the  
 468 Shanxi Comprehensive Reform Area policy significantly increased the intensity of resource exploitation,  
 469 which is extremely detrimental to air quality. In equation (4),  $\theta$  is 0.024 and significant at the 1% level,  
 470 and  $\delta_1$  is 0.165 and significant at the 5% level. Therefore, both the estimated value and the significance  
 471 of  $\delta_1$  are smaller than those of  $\alpha_1$ , suggesting that the influence of the Shanxi Comprehensive Reform  
 472 Area policy on  $PM_{2.5}$  concentrations is reduced after the mediator is accounted for. Therefore, the  
 473 intensity of resource exploitation is a partial mediator of the relationship between the Shanxi  
 474 Comprehensive Reform Area policy and  $PM_{2.5}$  concentrations.

### 475 5.3.2 The intensity of environmental regulations

476 Pollutant emissions per 10,000 yuan of GDP are used to represent the intensity of environmental  
 477 regulations, and the pollutants include industrial wastewater, industrial sulfur dioxide and industrial  
 478 smoke (dust). Theoretically, the more intense environmental regulation is, the lower the pollutant  
 479 emissions per 10,000 yuan of GDP. Column (3) of Table 7 shows that  $\beta_1$  is 0.235 and significant at the  
 480 10% level, indicating that the Shanxi Comprehensive Reform Area policy significantly contributed to the  
 481 increase in pollutant emissions per 10,000 yuan of GDP at the 10% level. Therefore, the policy lowered  
 482 the intensity of environmental regulation even though a series of environmental protection laws and  
 483 regulations had been promulgated. Column (4) of Table 7 shows that  $\theta$  is 0.037 and significant at the  
 484 1% level, and  $\delta_1$  is 0.181 and significant at the 1% level. The estimated value of  $\delta_1$  is smaller than that  
 485 of  $\alpha_1$ , and the significance of  $\delta_1$  ( $p=0.006$ ) is lower than that of  $\alpha_1$  ( $p=0.000$ ). Therefore, the intensity  
 486 of environmental regulation is also a partial mediator of the relationship between the Shanxi  
 487 Comprehensive Reform Area policy and  $PM_{2.5}$  concentrations.

488 **Table 7** Mediating effects tests for the three mediator variables

Variables	Intensity of resource exploitation		Intensity of environmental regulation		Capacity of scientific and technological innovation	
	(1) $\ln$ resexp	(2) $\ln PM_{2.5}$	(3) $\ln$ envreg	(4) $\ln PM_{2.5}$	(5) $\ln$ scitecinn	(6) $\ln PM_{2.5}$
$Treated_{it}$	0.677***	0.165**	0.235*	0.181***	-0.291**	0.195***
* $Period_{it}$	(9.42)	(2.74)	(1.78)	(3.22)	(-2.37)	(6.44)
$Treated_{it}$	1.693***	-0.054	0.028	-0.038	0.143	-0.026
	(4.74)	(-1.50)	(0.29)	(-1.04)	(1.45)	(-0.36)

<i>Period<sub>it</sub></i>	-0.11** (-2.01)	0.032 (0.61)	-0.989*** (-3.69)	0.091 (1.66)	0.097*** (2.68)	0.035 (1.49)
lngdpp	-0.322*** (-4.00)	-0.181*** (-6.18)	-1.071*** (-8.73)	-0.174*** (-6.47)	0.676*** (13.14)	-0.21*** (-5.90)
lnlocbudper	0.103** (2.37)	0.031 (1.65)	-0.466*** (-4.03)	0.047** (2.48)	0.591*** (18.04)	0.03 (1.32)
lnsecindper	1.285*** (7.56)	0.356*** (6.10)	2.025*** (5.98)	0.379*** (7.04)	-0.205 (-1.45)	0.449*** (5.84)
lninform	-0.032 (-0.71)	-0.145*** (-9.92)	0.284*** (6.76)	-0.166*** (-8.61)	-0.042 (-1.18)	-0.159*** (-6.00)
lnhiedustuper	-0.072 (-1.34)	-0.022 (-1.57)	0.05 (1.63)	-0.007 (-0.93)	0.036 (1.13)	-0.0245* (-1.70)
lnpopdensity	-0.509*** (-4.39)	0.39*** (24.87)	-0.325*** (-6.90)	0.383*** (26.60)	0.089** (2.47)	0.373*** (8.39)
lnfdiper	0.018 (0.94)	0.035*** (4.54)	0.131*** (4.99)	0.031*** (3.99)	0.034* (1.82)	0.032*** (3.13)
lnresexp		0.024*** (10.15)				
lnenvreg				0.037*** (5.48)		
lnscitecinn						0.008 (0.58)
Constant	2.037** (1.99)	1.222*** (7.71)	1.572*** (3.08)	1.158*** (7.35)	-7.971*** (-13.38)	1.225*** (2.97)
R-squared	0.2318	0.572	0.675	0.576	0.556	0.556
Number of obs	3866	3866	3866	3866	3866	3866

489

### 490 5.3.3 The capacity of scientific and technological innovation

491 This paper uses the share of science and technology expenditure in total public financial expenditure to  
492 represent the capacity of scientific and technological innovation. Column (5) of Table 7 shows that  $\beta_1$   
493 is -0.291 and significant at the 5% level, which shows that the Shanxi Comprehensive Reform Area  
494 policy significantly reduced the share of science and technology expenditure in total public financial  
495 expenditure. Column (6) of Table 7 shows that  $\delta_1$  is 0.195 and significant at the 1% level, whereas  $\theta$   
496 is 0.008 and not significant even at the 10% level. Therefore, we need to verify the mediating effect of  
497 the capacity of scientific and technological innovation using the Sobel test or bootstrapping. The Sobel  
498 test statistic is -0.0024, and the corresponding p-value is 0.353, so there is no mediating effect. However,  
499 the prerequisites for the Sobel test to be valid are that the mediating effects are normally distributed and  
500 that the sample is large, and the prerequisite of a normal distribution is difficult to meet. Therefore, a



501 bootstrap test is conducted, which shows that the mediating effect is -0.00241, with a corresponding p-  
502 value of 0.347 and a confidence interval of (-0.0074, 0.0026), suggesting that the capacity of scientific  
503 and technological innovation is not a mediator of the relationship between the Shanxi Comprehensive  
504 Reform Area policy and  $PM_{2.5}$  concentrations.

## 505 **6 Conclusions and policy implications**

506 The Shanxi Comprehensive Reform Area policy has attracted considerable attention since its  
507 implementation in 2011. This paper examines the causal effects of this policy on  $PM_{2.5}$   
508 concentrations using the PSM-DID method. The main findings are that the policy significantly  
509 increased  $PM_{2.5}$  concentrations by 0.211% per year on average; however, the dynamic effects  
510 show that the incremental effects of the policy include a decrease in the increase in  $PM_{2.5}$   
511 concentrations from significantly positive increments to zero or even negative increments from 2011  
512 to 2018, indicating that the policy gradually increased in environmental friendliness. The  
513 influencing mechanisms suggest that the Shanxi Comprehensive Reform Area policy influenced air  
514 quality by increasing the intensity of resource exploitation and decreasing the intensity of  
515 environmental regulation, while the capacity of scientific and technological innovation had no  
516 mediating effect.

517 This paper's marginal contribution to the existing knowledge on the treatment effects of the  
518 Shanxi Comprehensive Reform Area policy consists of two parts. First, the average and dynamic  
519 treatment effects of the Shanxi Comprehensive Reform Area policy on  $PM_{2.5}$  concentrations and  
520 the influencing mechanisms between them are analyzed. Second, a naive comparison of  $PM_{2.5}$   
521 concentrations before and after program implementation is shown to be uninformative since  $PM_{2.5}$   
522 concentrations in Shanxi would have changed even if Shanxi Province had not implemented the  
523 policy; therefore, counterfactual analysis is very useful for revealing the treatment effects of major  
524 policies issued by the state.

525 Our main findings have significant implications for policy makers. First, the government should  
526 further reduce the intensity of resource exploitation and strengthen the intensity of environmental  
527 regulation in order to reduce the ecological destruction and environmental pollution caused by  
528 market entities and gradually improve environmental quality and resource utilization. Second, the  
529 Shanxi Comprehensive Reform Area policy did not promote environmental quality through

530 scientific and technological innovation, which is an important driving force for the sustainable  
531 development of the ecological environment. Therefore, Shanxi Province should continue to promote  
532 environmentally focused scientific and technological innovations to solve prominent environmental  
533 problems and promote sustained improvements in environmental quality. Third, the treatment  
534 effects of the Shanxi Comprehensive Reform Area policy were not static but varied over time, so a  
535 dynamic and accurate assessment of its treatment effects is necessary to ensure the accuracy of  
536 policy adjustments and smooth policy operations.

537 This study has some limitations. First, PSM-DID cannot eliminate bias resulting from  
538 unobservable time-variant factors between the treatment group and the control group. To address  
539 this issue, we mainly select prefecture-level cities in northern and central China that are similar to  
540 those in Shanxi Province as the cities in the control group, excluding non-resource-based cities in  
541 southern China. In this way, the bias caused by unobservable factors can be minimized. Second, the  
542 paper only estimates the treatment effect of the Shanxi Comprehensive Reform Area policy on  
543  $PM_{2.5}$  concentrations; however, the treatment effects of the policy are comprehensive, including  
544 economic effects, social effects and effects on other aspects of the environment, which are directions  
545 for future research.

546

547 **Ethics approval and consent to participate:** Not applicable.

548 **Consent for publication:** Not applicable.

549 **Availability of data and materials:** The datasets used and/or analysed during the current study are  
550 available from the corresponding author on reasonable request.

551 **Competing interests:** The authors declare that they have no competing interests.

552 **Funding:** This research was supported by the Shanxi Planning Office of Philosophy and Social  
553 Sciences (2019B091).

554 **Authors' contributions:** RZ initiated and designed the study, reviewed the literature, analyzed the  
555 data, and wrote the manuscript. BX collected the data and helped with the data analysis. WL revised  
556 the manuscript. All authors read and approved the final manuscript.

557

558 **References**

- 559 Athey S, Imbens GW (2017) The state of applied econometrics: Causality and policy evaluation. *J Econ*  
560 *Perspect*, 31(2): 3-32. <https://doi.org/10.1257/jep.31.2.3>
- 561 Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological  
562 research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 51(6): 1173-  
563 1182. <http://dx.doi.org/10.1037/0022-3514.51.6.1173>
- 564 Cai S, Wang Y, Zhao B, Wang S, Chang X, Hao J (2017) The impact of the “Air Pollution Prevention  
565 and Control Action Plan” on  $PM_{2.5}$  concentrations in Jing-Jin-Ji region during 2012–2020. *Sci*  
566 *Total Environ* 580:197-209. <https://doi.org/10.1016/j.scitotenv.2016.11.188>
- 567 Cantoni D, Chen Y, Yang DY, Yuchtman N, Zhang YJ (2017) Curriculum and Ideology. *J Polit Econ*  
568 125(2), 338–392. <https://doi.org/10.1086/690951>
- 569 Chen J, Wang B, Huang S, Song M (2020) The influence of increased population density in China on air  
570 pollution. *Sci Total Environ* 735:139456. <https://doi.org/10.1016/j.scitotenv.2020.139456>
- 571 Dong K, Sun R, Dong C, Li H, Zeng X, Ni G (2018) Environmental Kuznets curve for  $PM_{2.5}$   
572 emissions in Beijing, China: What role can natural gas consumption play? *Ecol Indic* 93, 591-  
573 601. <https://doi.org/10.1016/j.ecolind.2018.05.045>
- 574 Edwards JR, Lambert LS (2007) Methods for integrating moderation and mediation: A general analytical  
575 framework using moderated path analysis. *Psychol Methods* 12(1), 1–22.  
576 <https://doi.org/10.1037/1082-989X.12.1.1>
- 577 Eliopoulos GM, Harris AD, Bradham DD, Baumgarten M, Zuckerman IH, Fink JC, Perencevich EN  
578 (2004) The use and interpretation of quasi-experimental studies in infectious diseases. *Clin*  
579 *Infect Dis* 38:1586-1591. <https://doi.org/10.1086/420936>
- 580 Gan W, Xu X, Yi J (2011) Achievement, problems and countermeasures for implementation of Western  
581 Development Strategy. *Economic Geography*, 2011, 31(1):40-46.  
582 <https://doi.org/10.15957/j.cnki.jjdl.2011.01.014> (In Chinese).
- 583 Greenstone M, Gayer T (2009) Quasi-experimental and experimental approaches to environmental  
584 economics. *J Environ Econ Manage* 57:21-44. <https://doi.org/10.1016/j.jeem.2008.02.004>
- 585 Guo J, Guo S (2019a) Has the establishment of National Comprehensive Supporting Reform Pilot Zone  
586 promoted the optimization of industrial structure — Empirical analysis based on Synthetic

587 Control Method. Finance and Economics 8: 69-81. (In Chinese)

588 Guo S, Guo J (2019b) The establishment of “Comprehensive Reform Zone”, industrial diversification  
589 and high-quality development of resource-based regions. Ind Econ Res (1): 87-98.  
590 <https://doi.org/10.13269/j.cnki.ier.2019.01.008> (In Chinese)

591 Halkos GE, Paizanos EA (2013) The effect of government expenditure on the environment: An empirical  
592 investigation. Ecol Econ 91, 48-56. <https://doi.org/10.1016/j.ecolecon.2013.04.002>

593 Heckman JJ, Ichimura H, Todd P (1998) Matching as an econometric evaluation estimator. Rev Econ  
594 Stud 65(2): 261-294. <https://doi.org/10.1111/1467-937X.00044>

595 Heckman JJ, Vytlacil E (2001) Policy-relevant treatment effects. Am Econ Rev 91(2): 107-111.  
596 <https://doi.org/10.1257/aer.91.2.107>

597 Hishan SS, Sasmoko, Khan A, Ahmad J, Hassan ZB, Zaman K, Qureshi MI (2019) Access to clean  
598 technologies, energy, finance, and food: environmental sustainability agenda and its  
599 implications on Sub-Saharan African countries. Environ Sci Pollut Res Int 26(16): 16503-16518.  
600 <https://doi.org/10.1007/s11356-019-05056-7>

601 Imbens GW, Rubin DB (2015) Causal inference in statistics, social, and biomedical sciences. New York,  
602 NY: Cambridge University Press.

603 Ji X (2020) Research on the impact of central China’s Rising Strategy on urban environmental quality –  
604 An analysis based on PSM-DID method. Inquiry Into Economic Issues 8:157-169. (In Chinese)

605 Ji Y, Wu H, Liu X, Dou J. How do fiscal and tax policies affect firms’ productivity under Western  
606 Development Strategy. Review of Industrial Economics, 2018, 17: 71-93. (In Chinese)

607 Lan J, Kakinaka M, Huang X (2012) Foreign direct investment, human capital and environmental  
608 pollution in China. Environ Resource Econ 51:255-275.[https://doi.org/10.1007/s10640-011-](https://doi.org/10.1007/s10640-011-9498-2)  
609 9498-2

610 Li H, Graham DJ, Majumdar A (2013) The impacts of speed cameras on road accidents: An application  
611 of propensity score matching methods. Accid Anal Prev 60:148-157.  
612 <http://dx.doi.org/10.1016/j.aap.2013.08.003>

613 Li L, Lei Y, Wu S, He C, Yan D (2018) Study on the coordinated development of economy, environment  
614 and resource in coal-based areas in Shanxi Province in China: Based on the multi-objective  
615 optimization model. Resour Policy 55:80-86. <https://doi.org/10.1016/j.resourpol.2017.10.017>

616 Li P, Lu Y, Wang J (2016) Does flattening government improve economic performance? Evidence from

617 China. *J Dev Econ* 123:18-37. <https://doi.org/10.1016/j.jdeveco.2016.07.002>

618 Li X, Guo H, Jin S, Ma W, Zeng Y (2021) Do farmers gain internet dividends from E-commerce adoption?  
619 Evidence from China. *Food Policy*: 102024. <https://doi.org/10.1016/j.foodpol.2021.102024>

620 Lu Z, Huang L, Liu J, Zhou Y, Chen M, Hu J (2019) Carbon dioxide mitigation co-benefit analysis of  
621 energy-related measures in the Air Pollution Prevention and Control Action Plan in the Jing-  
622 Jin-Ji region of China. *Resour Conserv Recycl* X1: 100006.  
623 <https://doi.org/10.1016/j.rcrx.2019.100006>

624 Paziienza P (2014) The relationship between FDI and the natural environment: Facts, evidence and  
625 prospects. Springer International Publishing, Switzerland.  
626 <https://doi.org/10.1007/978-3-319-04301-2>

627 Peng X, Chen Z. Impact evaluation on China's Western Development policy. *China Population,*  
628 *Resources and Environment*, 2016, 26(3): 136-144.  
629 <https://doi.org/10.3969/j.issn.1002-2104.2016.03.017>(In Chinese)

630 Rodrik D, Rosenzweig MR (2010) Preface: Development policy and development economics: An  
631 introduction. In: Rodrik D, Rosenzweig MR (Eds.), *Handbook of Development Economics*, vol.  
632 5, Amsterdam, North-Holland.

633 Rosenbaum PR, Rubin DB (1985) Constructing a control group using multivariate matched sampling  
634 methods that incorporate the propensity score. *Am Stat* 39(1): 33-38.  
635 <https://doi.org/10.1080/00031305.1985.10479383>

636 Sun Y, Li G, Shi D (2019) Western development and technological innovation: spillover or trap? Re-  
637 examination based on PSM-DID. *Journal of Yunnan Finance and Trade Institute* (6): 51-62. (In  
638 Chinese)

639 Mendoza OMV (2017) Infrastructure Development, Income Inequality and Urban Sustainability in the  
640 People's Republic of China (ADB Working Paper 713). Asian Development Bank Institute,  
641 Tokyo.

642 Pesaran MH, Smith RP (2016). Counterfactual analysis in Macroeconometrics: An empirical  
643 investigation into the effects of quantitative easing. *Res Econ*, 70(2): 262-280.  
644 <https://doi.org/10.1016/j.rie.2016.01.004>

645 Yuan H, Zhu C (2018) Does the Western Development promote the industrial structure transformation  
646 and upgrading? — A test based on PSM-DID method. *China Soft Science* (6): 67-81. (In

647 Chinese)

648 Wendling ZA, Emerson JW, de Sherbinin A, Esty DC, et al. (2020). 2020 Environmental Performance  
649 Index. New Haven, CT: Yale Center for Environmental Law & Policy.

650 Wu H, Liu X, Ji Y (2017) How do tax cuts affect firm business efficiency? Based on the quasi-natural  
651 experiment of western development policy. *Journal of Finance and Economics* 43(4):55-67.  
652 <https://doi.org/10.16538/j.cnki.jfe.2017.04.005> (In Chinese)

653 Yin A, Li C (2019) Study on the heterogeneous effect of environmental regulation on economic  
654 ecological development in Beijing, Tianjin and Hebei. *Contemporary Economic Management*,  
655 42(5):66-71. <https://doi.org/10.13253/j.cnki.ddjjgl.2020.05.009> (In Chinese)

656 You H, Chi R (2012) Informatization and the industrial upgrading in China. *Serv Sci* 3(1): 7-20.  
657 Retrieved January 20, 2021, from <http://web.csulb.edu/~rchi/ssj/papers/20121/201212.pdf>

658 Zhang C, Zhou B, Wang Q (2019a) Effect of China's Western Development Strategy on carbon intensity.  
659 *J Clean Prod* 215: 1170-1179. <https://doi.org/10.1016/j.jclepro.2019.01.136>

660 Zhang H, Duan M (2020) China's pilot emissions trading schemes and competitiveness: An empirical  
661 analysis of the provincial industrial sub-sectors. *J Environ Manage* 258: 109997.  
662 <https://doi.org/10.1016/j.jenvman.2019.109997>

663 Zhang H, Duan M, Deng Z (2019b) Have China's pilot emissions trading schemes promoted carbon  
664 emission reductions? – the evidence from industrial sub-sectors at the provincial level. *J Clean*  
665 *Prod* 234: 912-924. <https://doi.org/10.1016/j.jclepro.2019.06.247>

666 Zhang Q, Zhang S, Ding Z, Hao Y (2017) Does government expenditure affect environmental quality?  
667 Empirical evidence using Chinese city-level data. *J Clean Prod* 161:143-152.  
668 <https://doi.org/10.1016/j.jclepro.2017.05.096>

669 Zhang R, Dong S, Li Z (2020) The economic and environmental effects of the Beijing-Tianjin-Hebei  
670 Collaborative Development Strategy — taking Hebei Province as an example. *Environ Sci*  
671 *Pollut Res* 27:35692-35702. <https://doi.org/10.1007/s11356-020-09790-1>

672

# Figures

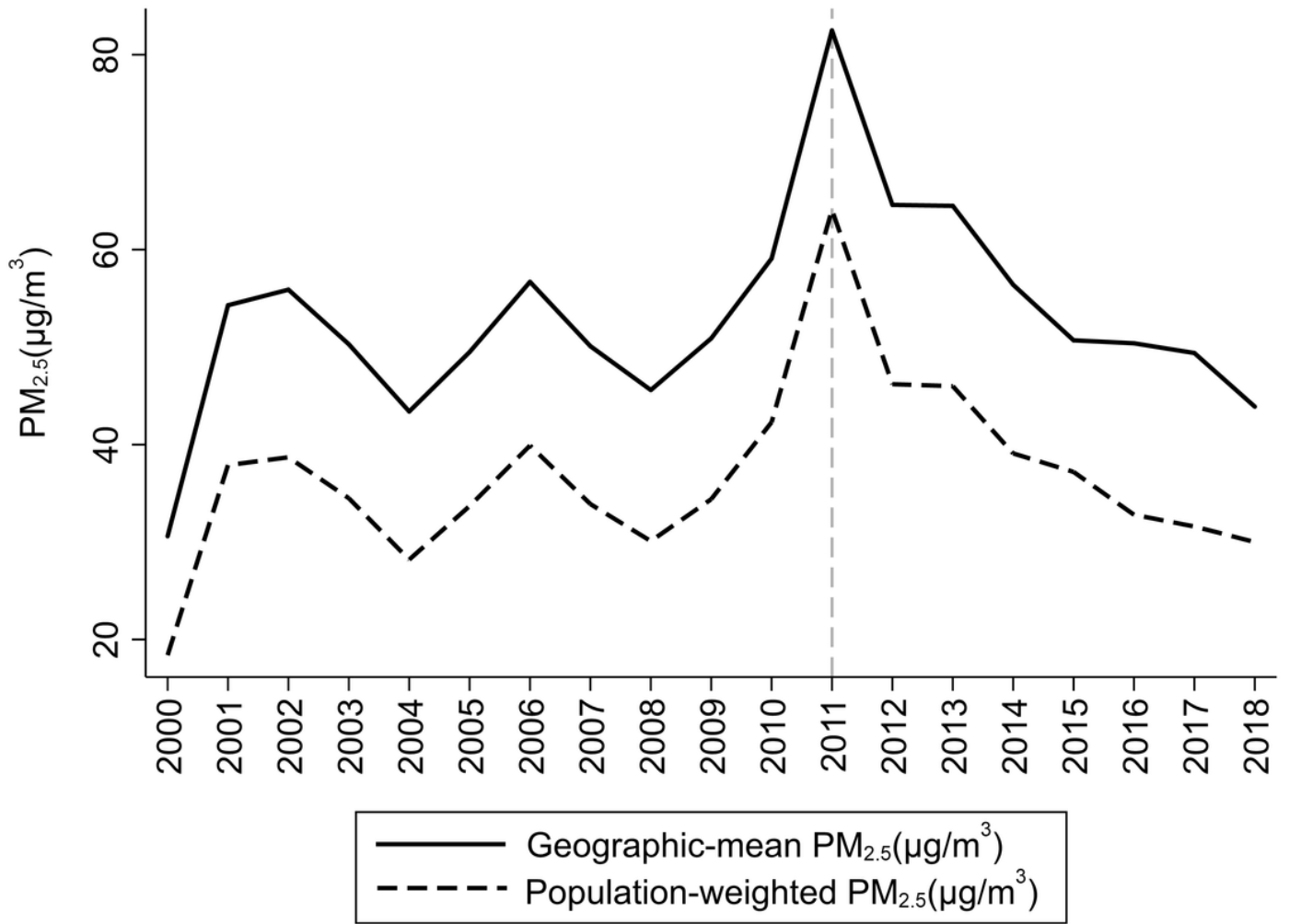
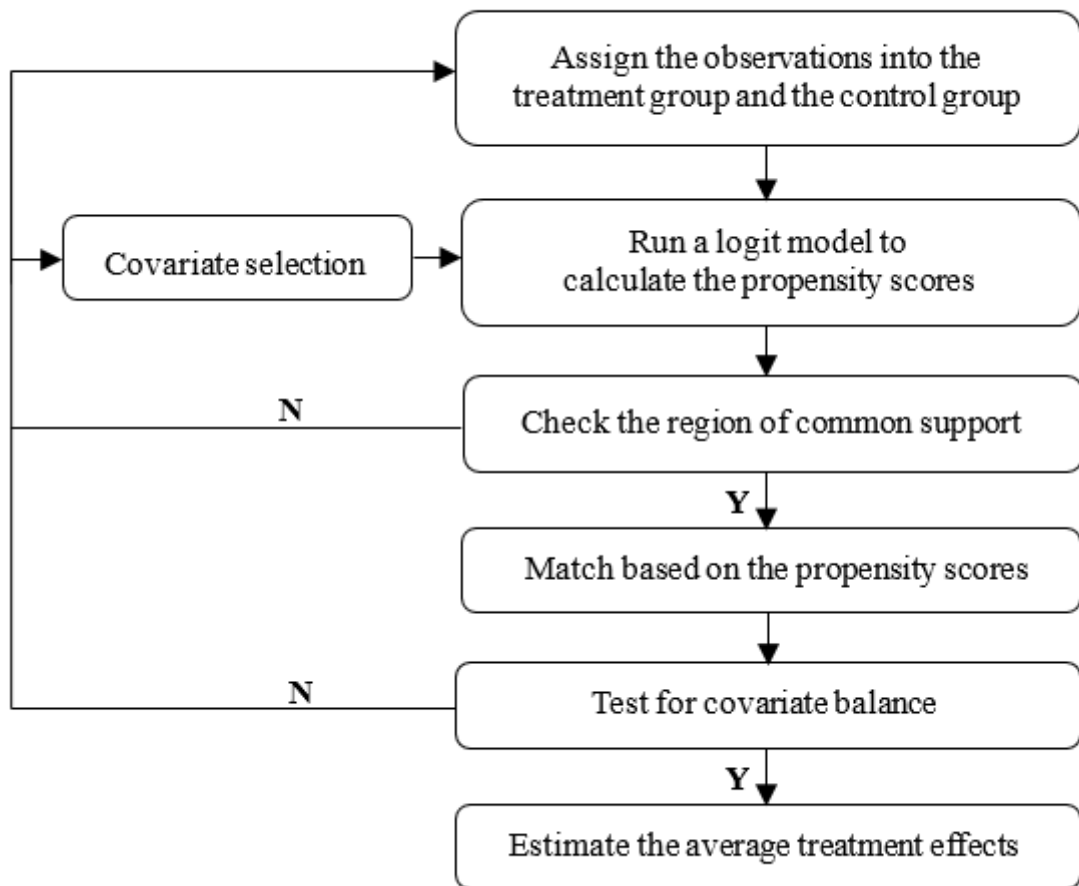


Figure 1

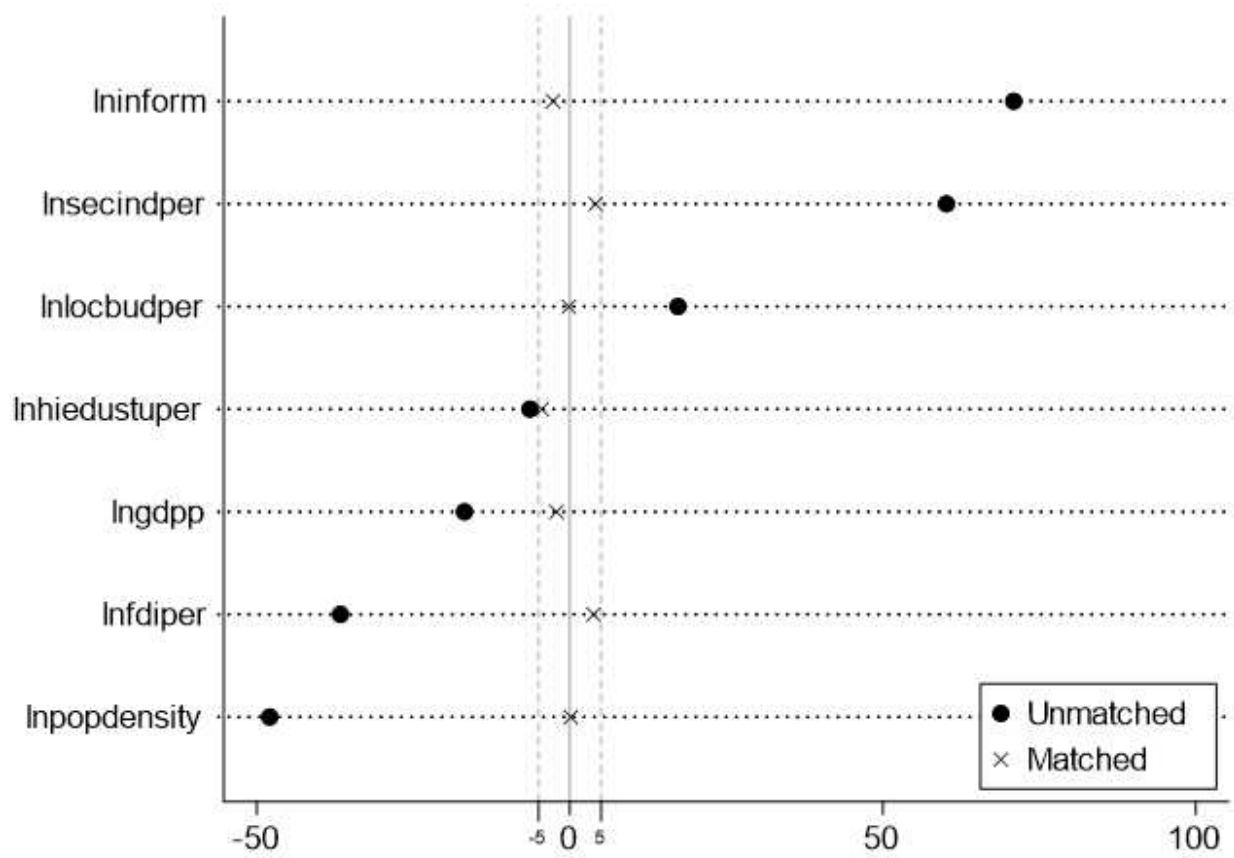
Geographic-Mean PM<sub>2.5</sub> and Population-Weighted PM<sub>2.5</sub> in Shanxi Province





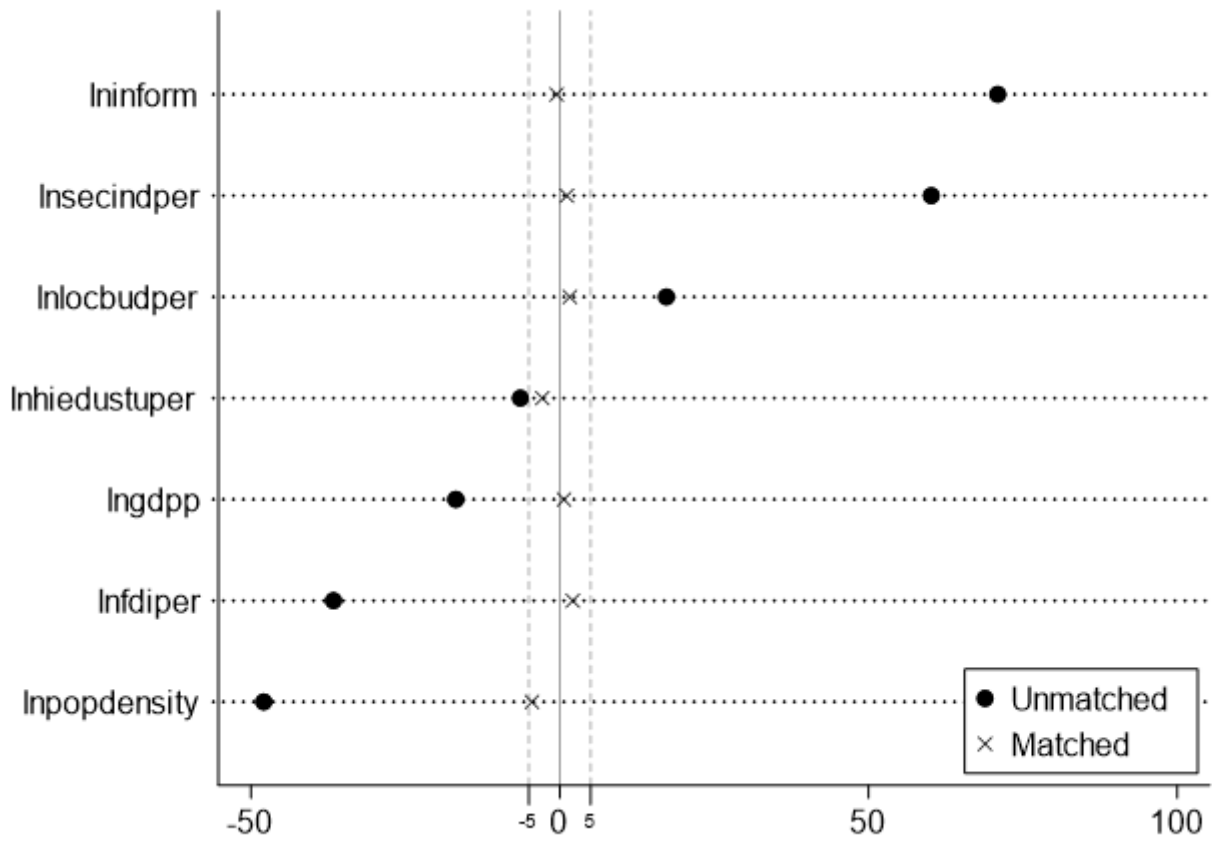
**Figure 2**

Steps for conducting a PSM-DID analysis



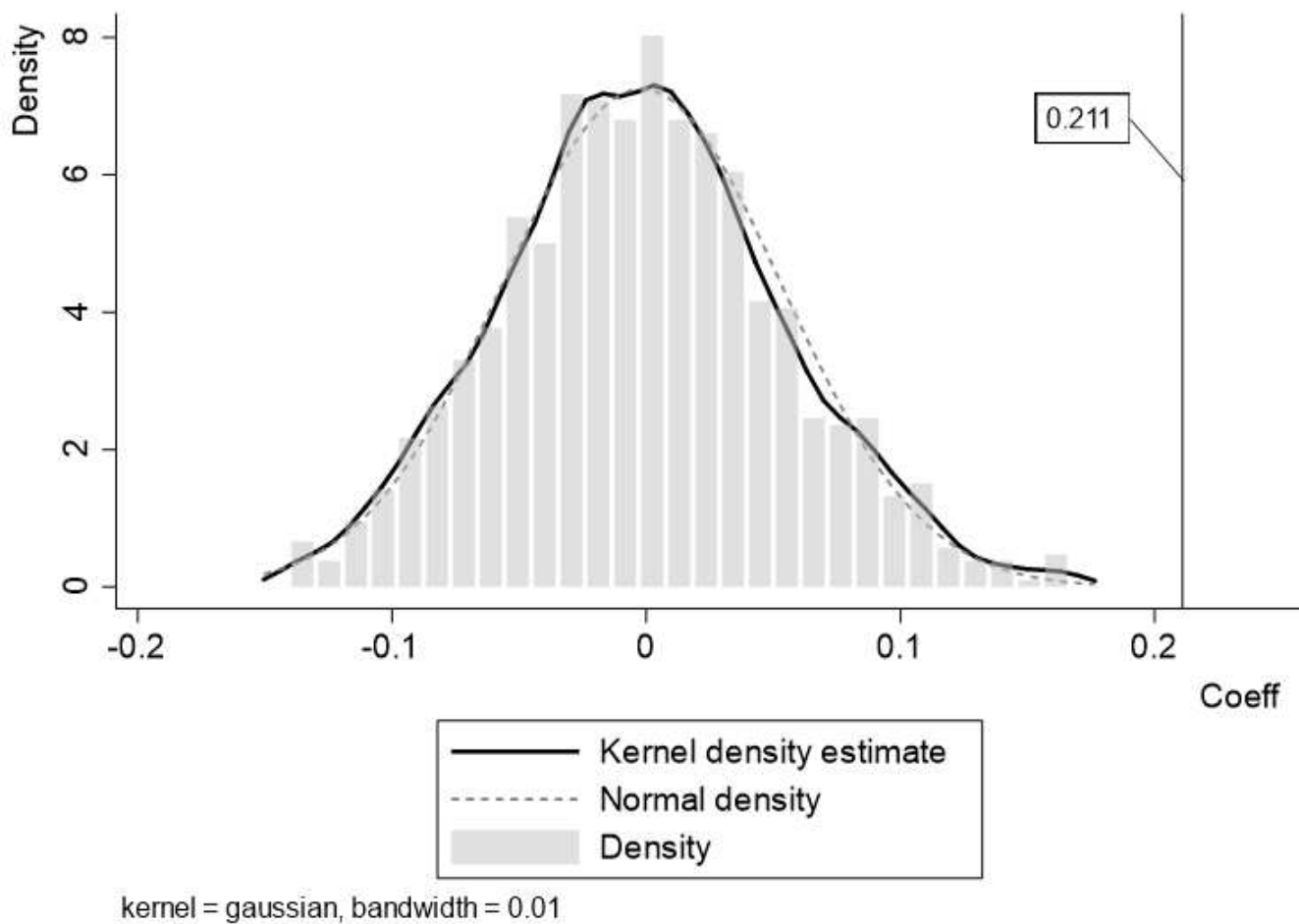
**Figure 3**

Standardized percentage bias across covariates (kernel matching, kernel type= epanechnikov)



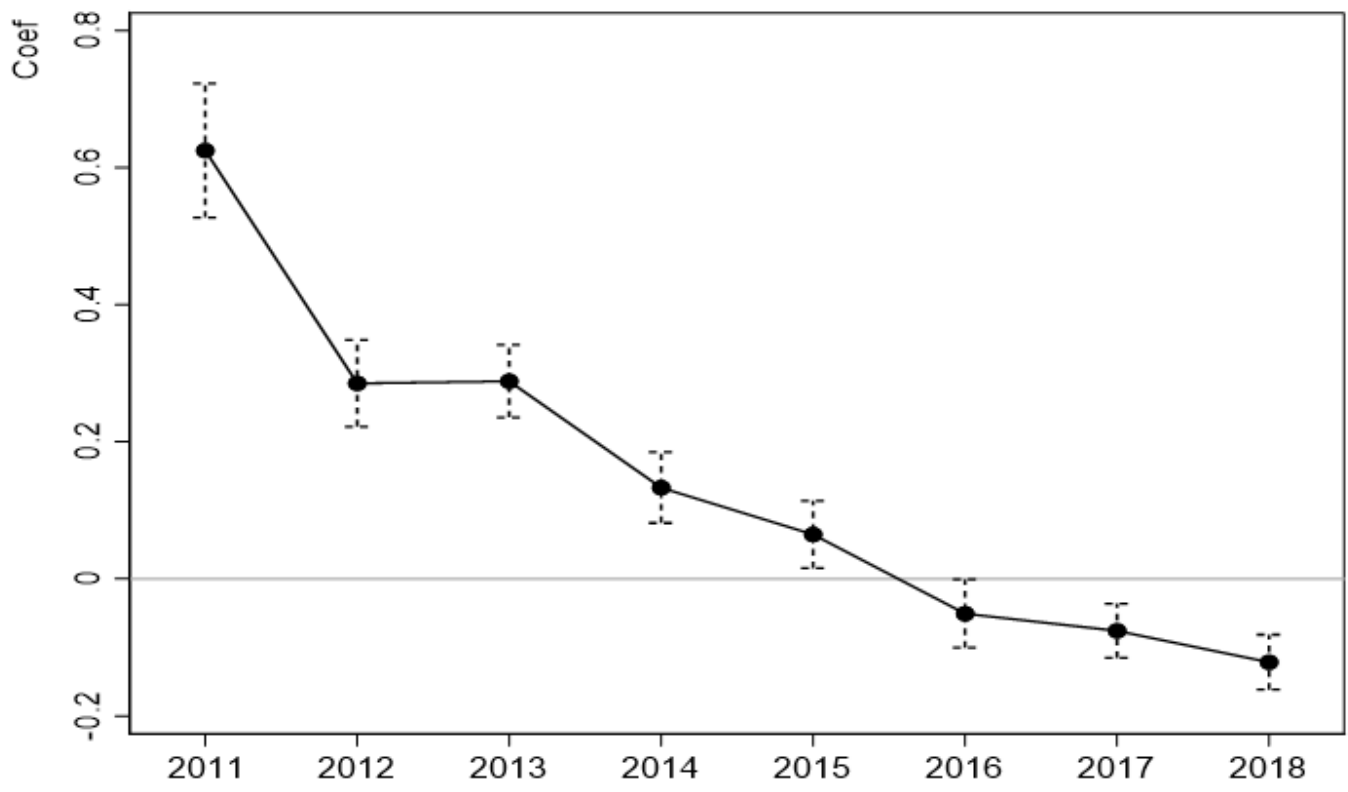
**Figure 4**

Standardized percentage bias across covariates (nearest neighbor matching within a caliper)



**Figure 5**

The distribution of 1000 coefficient estimates



**Figure 6**

Trend in coefficients without covariates

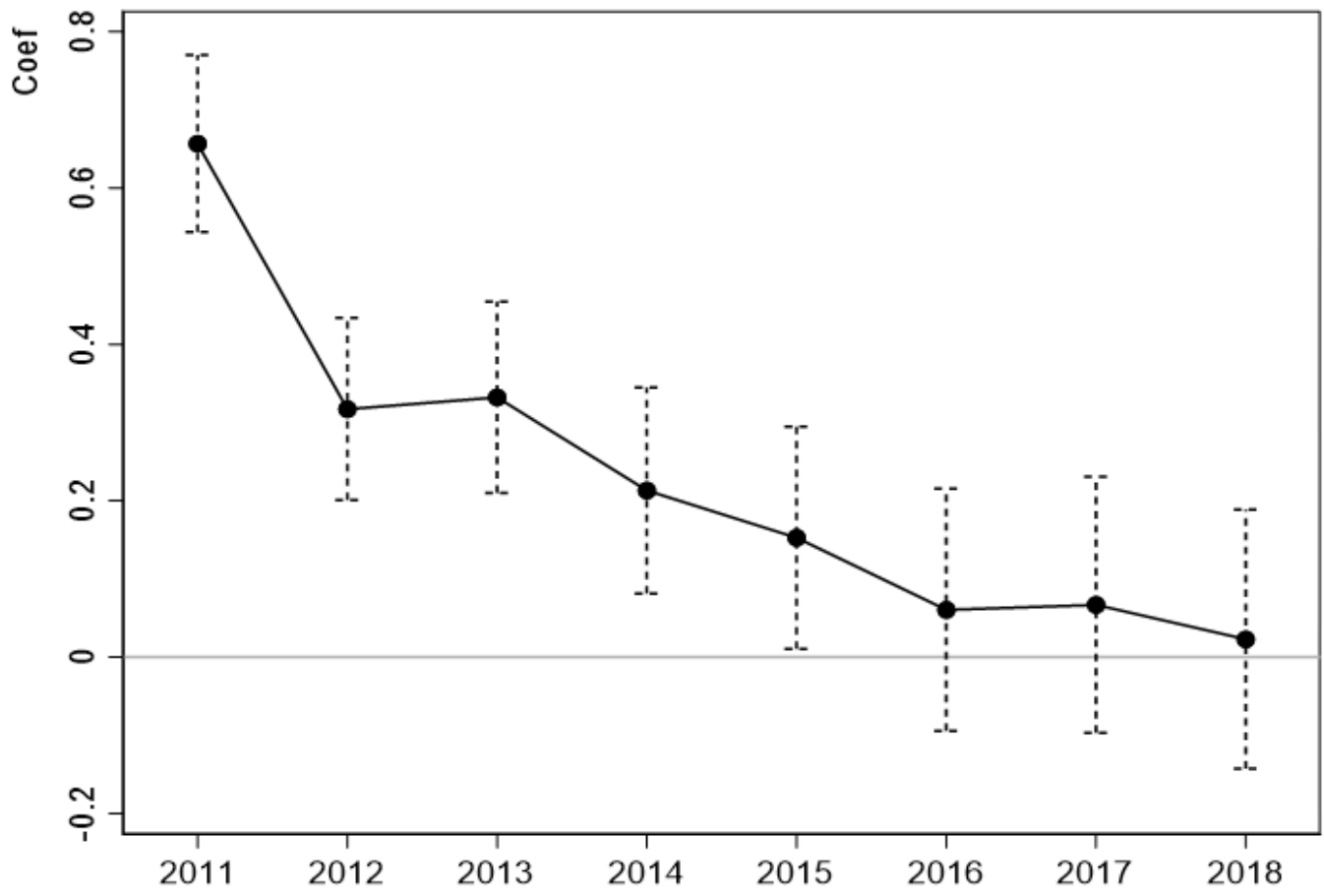


Figure 7

Trend in coefficients with covariates