The Treatment Effect of the Shanxi Comprehensive Reform Area Policy on PM2.5 Concentrations: A Study Based on a Quasi-Experiment

Rongxia Zhang (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

Version of Record: A version of this preprint was published at Environmental Science and Pollution Research on September 8th, 2021. See the published version at https://doi.org/10.1007/s11356-021-16225-y.

The Treatment Effect of the Shanxi Comprehensive Reform Area Policy on $PM_{2.5}$ Concentrations: A Study Based on a Quasi-Experiment

Rongxia Zhang[1,*], Baichuan Xiang[1], Wei Li[1]

[1] College of Economics and Management, Taiyuan University of Technology, Taiyuan 030024, China

***Corresponding author**

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

Acknowledgements: This article is funded by the Shanxi Planning Office of Philosophy and Social Sciences(2019B091). The views of this paper are only from the authors and do not necessarily represent the official views of the sponsors. We would like to appreciate the anonymous reviewers for their valuable comments.

Abstract

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

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

1 Introduction

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

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

[1] 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.

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

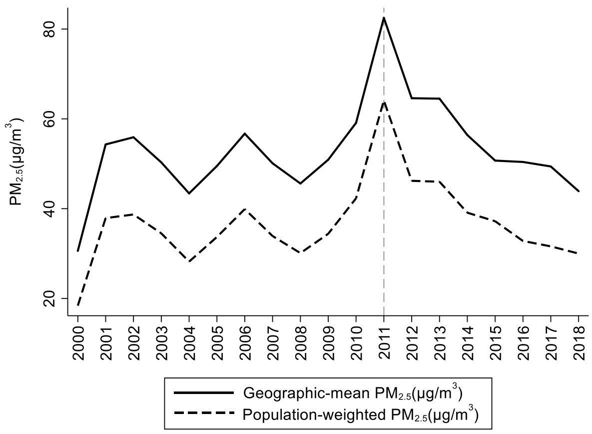

Fig. 1 Geographic-Mean $PM_{2.5}$ and Population-Weighted $PM_{2.5}$ in Shanxi Province[2]

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

[2] 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).

adjustments.

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

2 Literature review

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

Regarding the evaluation of the treatment effects of the Shanxi Comprehensive Reform Area policy, there are more theoretical and qualitative studies and fewer empirical and quantitative studies. To date, only two studies on the China National Knowledge Infrastructure (CNKI) website have studied the treatment effects of the Shanxi Comprehensive Reform Area policy with quantitative methods. Guo and Guo (2019a) studied the treatment effects of the Shanxi Comprehensive Reform Area policy on the optimization of the industrial structure using SCM. Guo and Guo (2019b) investigated the treatment effects of the Shanxi Comprehensive Reform Area policy on the quality of economic development using DID. Although few studies have focused on the causal relationship between the Shanxi Comprehensive Reform Area policy and $PM_{2.5}$ concentrations, we can review those studies that are on the treatment effects of coordinated regional development strategies since the Shanxi Comprehensive Reform Area policy is a coordinated regional development strategy.

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

No studies based on quantitative methods have been found that study the mechanisms that influence the relationship between the Shanxi Comprehensive Reform Area policy and environmental pollutants. Therefore, we can only review a limited number of studies on these influencing mechanisms based on coordinated regional development strategies. Ji (2020) argued that the Central Rising Strategy aggravated environmental pollution, most likely by relaxing resource exploitation restrictions, undertaking more FDI projects and increasing the share of the secondary industry in total industry. Yin and Li (2019) suggested that the Jing-Jin-Ji Collaborative Development Strategy induced the ecological development of the economy through environmental regulations. Sun et al. (2019) claimed that the Western Development Strategy provided a strong impetus for economic development by improving the quality of scientific and technological innovations. In addition, other studies have suggested that the possible influencing mechanisms of such strategies included fixed asset investments, financial development, and urbanization. Therefore, because these influencing mechanisms are complex and not directly observable and it is difficult to identify influencing mechanisms through randomized controlled trials, the papers published to date have not drawn consistent conclusions. We can only test the influencing mechanisms that are likely to have a great impact on air quality according to the available literature and to economic theory on a case-by-case basis.

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

3 Methodological framework and data

3.1 Methodological framework

This study uses the PSM-DID method proposed by Heckman et al. (1998) to estimate the counterfactual outcome values that would obtain in the absence of treatment for the treated units. Combining the advantages of PSM and DID, the PSM-DID method can mitigate selection bias and eliminate the influence of time-invariant unobservable factors on the outcome variable, thus yielding more robust estimates. The idea behind the PSM-DID method originates from a matching estimator, and the basic idea is to find a real or synthetic unit in the control group whose observable covariates are as similar as possible to those of a unit in the treatment group (Zhang and Duan 2020). If the unit's selection into the treatment group is caused by observable characteristics and time-invariant unobservable characteristics and if the treated unit i and untreated unit j have similar probabilities of being treated, then the outcomes of unit i and unit j can be compared to estimate the treatment effects.

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

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

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

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

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

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

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

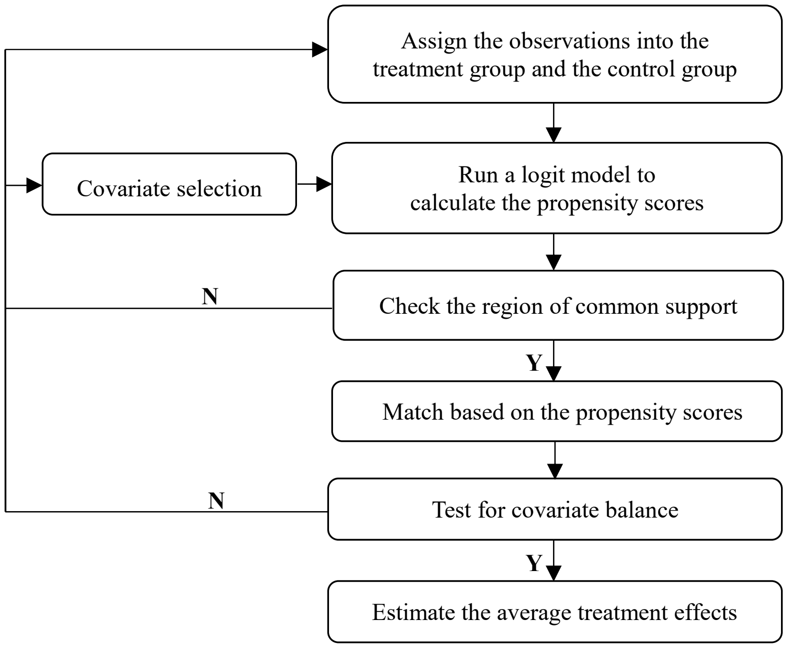

Fig. 2 Steps for conducting a PSM-DID analysis

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

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

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

The DID estimator is illustrated in Table 1. Equation (1) shows that for treated cities in Shanxi Province, the $PM_{2.5}$ concentrations before and after the policy change are $\alpha_0 + \alpha_2 + \alpha_4 X_{it_0}$ and $\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 X_{it_1}$,[3] respectively, and the difference between them, $\Delta Y_t = \alpha_1 + \alpha_3 + \alpha_4(X_{it1} - X_{it0})$, is the change in $PM_{2.5}$ concentrations, which represents the combined treatment effects of the Shanxi Comprehensive Reform Area policy and of other external factors. Similarly, for matched untreated cities in the control group, the change in $PM_{2.5}$ concentrations before and after the policy change is $\Delta Y_c = \alpha_3 + \alpha_4(X_{it1} - X_{it0})$, which represents the treatment effect of other external factors. Therefore, the difference between ΔY_t and ΔY_c, $\Delta\Delta Y = \alpha_1$, is the net treatment effect of the Shanxi Comprehensive Reform Area policy. If $\alpha_1 > 0$, then the Shanxi Comprehensive Reform Area policy has caused the air quality of the treated cities to deteriorate; otherwise, it has improved their air quality.

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	α_2	$\alpha_1 + \alpha_2$	$\Delta\Delta Y = \alpha_1$

3.2 Covariate selection

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

[3] X_{it_0} and X_{it_1} represent the values of the covariates before and after treatment, respectively.

[4] 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.

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

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

3.3 Data

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

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

4.1 Propensity score matching

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

4.1.1 Estimating the propensity score

PSM reduces the matching criteria from multiple dimensions to a single dimension, the propensity score. Therefore, treated and untreated units with similar propensity scores can be compared to obtain treatment effects (Li et al. 2013). Before matching, a binary logit model needs to be constructed to estimate the conditional probability of assignment to the treatment group given a vector of observable covariates:

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

where $Treated_{it} = 1$ if city i is treated by the Shanxi Comprehensive Reform Area policy and $Treated_{it} = 0$ otherwise. γ_0 is the intercept term. γ_1 is the vector of regression coefficients. X_{it} is a vector of the observable covariates selected above. μ_{it} is the stochastic error term.

The estimated results of the logit propensity score model can be found in Table 3, which shows that almost all covariates play a significant role in the city's selection into the treatment group.

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		

Note: The dependent variable, Treated, is treatment status, which equals one if a city belongs to the treatment group and zero if it belongs to the control group. Standard errors are clustered at the city level. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

4.1.2 Verifying the common support assumption

After estimating the propensity score, the next step is to ensure that the region of common support or overlap in propensity scores across the treated and untreated cities is sizable because the matching results are considered to be accurate only when a larger share of the observations satisfies the common support assumption (Heckman and Vytlacil, 2001). The region of common support may be different for different matching methods. Tables 4 and 5 show that for kernel matching and nearest neighbor matching within a caliper, 102 and 53 observations do not satisfy the common support assumption, and 3866 and 3915 observations do satisfy the common support assumption, respectively. Both matching algorithms have a small amount of sample loss, and the matching results are satisfactory.

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

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

4.1.3 Selecting the matching algorithm

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

4.1.4 Testing the balance condition

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

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

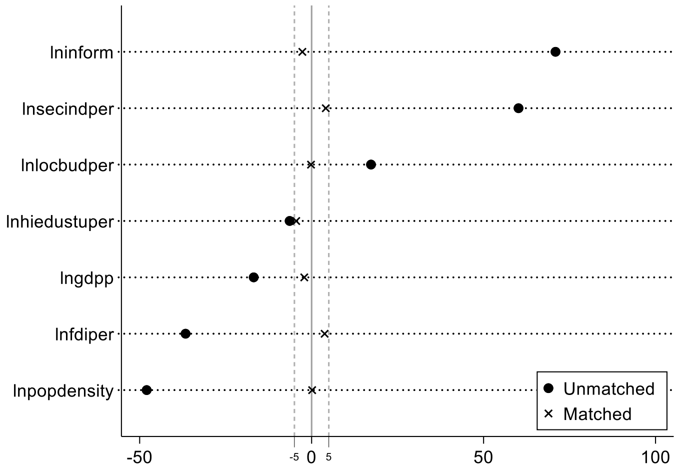

Fig. 3 Standardized percentage bias across covariates (kernel matching, kernel type= epanechnikov)

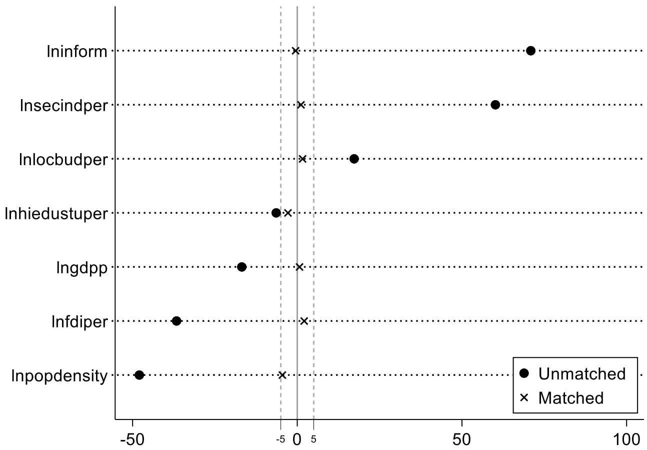

Fig. 4 Standardized percentage bias across covariates (nearest neighbor matching within a caliper)

4.2 The results of the PSM-DID model

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

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

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

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

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

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

4.3 Placebo test

To ensure the robustness of the above results, this study conducts a place-based placebo test following Li et al. (2016) and Cantoni et al. (2017). First, 11 cities are randomly selected as the false treatment group from the total sample of 248 cities, and the other cities are selected as the control group. Second, the value of the variable $Treated_{it}$ is adjusted accordingly, and the PSM-DID regression is estimated. Third, the above steps are repeated 1000 times, and then the robustness of the regression is evaluated based on the significance and distribution of the 1000 coefficients on the interaction term $Treated_{it} * Period_{it}$.

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

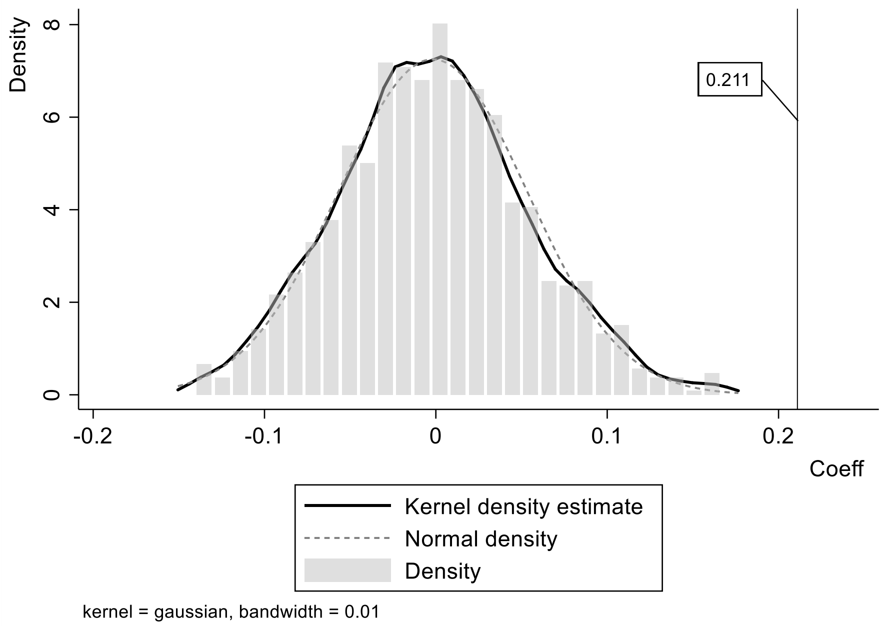

Fig. 5 The distribution of 1000 coefficient estimates

4.4 Dynamic effects tests

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

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

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

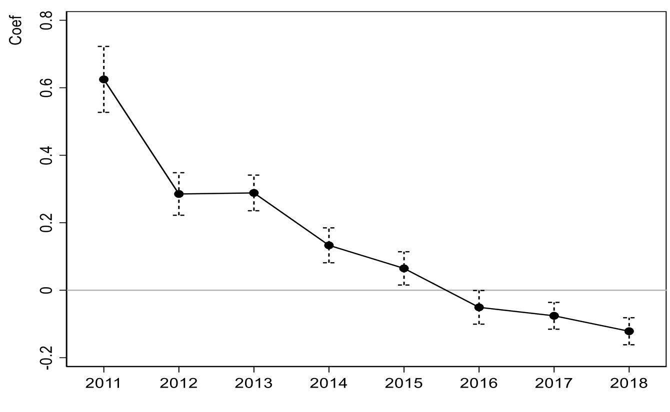

Fig. 6 Trend in coefficients without covariates

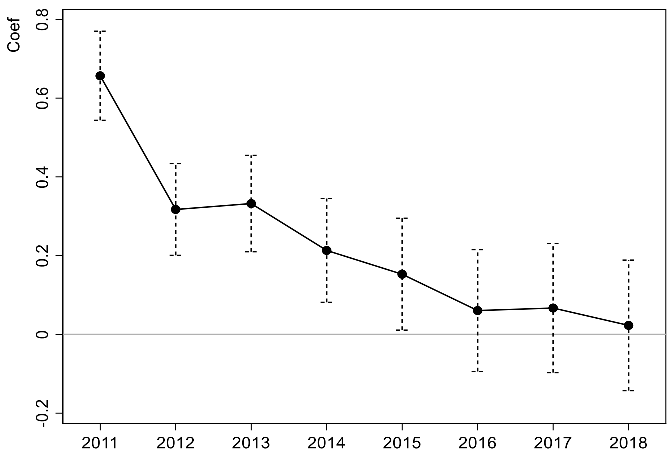

Fig. 7 Trend in coefficients with covariates

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

5.1 Mediator variables

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

5.2 The causal steps procedure for testing mediation

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

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

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

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

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

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

The results from Eqs. (1)-(3) are used to evaluate the following conditions for the intensity of resource exploitation, the intensity of environmental regulations, and the capacity of scientific and technological innovation as mediators of the relationship between the Shanxi Comprehensive Reform Area policy and $PM_{2.5}$ concentrations: (a) the Shanxi Comprehensive Reform Area policy should be related to the logarithm of $PM_{2.5}$ concentrations in equation (1) such that α_1 is significant; (b) the Shanxi Comprehensive Reform Area policy should be related to the mediator in Equation (3) such that β_1 is significant; (c) the mediator should be related to the logarithm of $PM_{2.5}$ concentrations in Equation (3) such that θ is significant; and (d) the relation between the Shanxi Comprehensive Reform Area policy and the logarithm of $PM_{2.5}$ concentrations in Equation (4) (i.e., δ_1) should be insignificant or significant but smaller than the relationship between the Shanxi Comprehensive Reform Area policy and the logarithm of $PM_{2.5}$ concentrations in Equation (1) (i.e., α_1). Assuming that the first 3 conditions are satisfied, complete mediation is inferred if δ_1 is insignificant, while partial mediation is inferred if δ_1 remains significant but is significantly smaller than α_1 (Edwards and Lambert, 2007). If at least one of the coefficients β_1 or θ is insignificant, a Sobel or bootstrap test is required to determine whether the mediator variable has a mediating effect.

5.3 The mediating effects

5.3.1 The intensity of resource exploitation

We use the share of mining employment in total urban employment to represent the intensity of resource exploitation, which increased from 18% in 2011 to 23% in 2018 in Shanxi Province. The results of the mediating effects test are shown in column (3) of Table 6 and columns (1) and (2) of Table 7. The results show that α_1 in equation (1) is 0.211 and significant at the 1% level for the kernel matching algorithm, indicating that the Shanxi Comprehensive Reform Area policy significantly increased $PM_{2.5}$ concentrations. β_1 in equation (3) is 0.677 and significant at the 1% level, which suggests that the Shanxi Comprehensive Reform Area policy significantly increased the intensity of resource exploitation, which is extremely detrimental to air quality. In equation (4), θ is 0.024 and significant at the 1% level, and δ_1 is 0.165 and significant at the 5% level. Therefore, both the estimated value and the significance of δ_1 are smaller than those of α_1, suggesting that the influence of the Shanxi Comprehensive Reform Area policy on $PM_{2.5}$ concentrations is reduced after the mediator is accounted for. Therefore, the intensity of resource exploitation is a partial mediator of the relationship between the Shanxi Comprehensive Reform Area policy and $PM_{2.5}$ concentrations.

5.3.2 The intensity of environmental regulations

Pollutant emissions per 10,000 yuan of GDP are used to represent the intensity of environmental regulations, and the pollutants include industrial wastewater, industrial sulfur dioxide and industrial smoke (dust). Theoretically, the more intense environmental regulation is, the lower the pollutant emissions per 10,000 yuan of GDP. Column (3) of Table 7 shows that β_1 is 0.235 and significant at the 10% level, indicating that the Shanxi Comprehensive Reform Area policy significantly contributed to the increase in pollutant emissions per 10,000 yuan of GDP at the 10% level. Therefore, the policy lowered the intensity of environmental regulation even though a series of environmental protection laws and regulations had been promulgated. Column (4) of Table 7 shows that θ is 0.037 and significant at the 1% level, and δ_1 is 0.181 and significant at the 1% level. The estimated value of δ_1 is smaller than that of α_1, and the significance of δ_1 (p=0.006) is lower than that of α_1 (p=0.000). Therefore, the intensity of environmental regulation is also a partial mediator of the relationship between the Shanxi Comprehensive Reform Area policy and $PM_{2.5}$ concentrations.

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) lnresexp	(2)ln$PM_{2.5}$	(3) lnenvreg	(4)ln$PM_{2.5}$	(5) lnscitecinn	（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)

$Period_{it}$	-0.11**	0.032	-0.989***	0.091	0.097***	0.035
	(-2.01)	(0.61)	(-3.69)	(1.66)	(2.68)	(1.49)
lngdpp	-0.322***	-0.181***	-1.071***	-0.174***	0.676***	-0.21***
	(-4.00)	(-6.18)	(-8.73)	(-6.47)	(13.14)	(-5.90)
lnlocbudper	0.103**	0.031	-0.466***	0.047**	0.591***	0.03
	(2.37)	(1.65)	(-4.03)	(2.48)	(18.04)	(1.32)
lnsecindper	1.285***	0.356***	2.025***	0.379***	-0.205	0.449***
	(7.56)	(6.10)	(5.98)	(7.04)	(-1.45)	(5.84)
lninform	-0.032	-0.145***	0.284***	-0.166***	-0.042	-0.159***
	(-0.71)	(-9.92)	(6.76)	(-8.61)	(-1.18)	(-6.00)
lnhiedustuper	-0.072	-0.022	0.05	-0.007	0.036	-.0245*
	(-1.34)	(-1.57)	(1.63)	(-0.93)	(1.13)	（-1.70）
lnpopdensity	-0.509***	0.39***	-0.325***	0.383***	0.089**	0.373***
	(-4.39)	(24.87)	(-6.90)	(26.60)	(2.47)	(8.39)
lnfdiper	0.018	0.035***	0.131***	0.031***	0.034*	0.032***
	(0.94)	(4.54)	(4.99)	(3.99)	(1.82)	(3.13)
lnresexp		0.024***				
		(10.15)				
lnenvreg				0.037***		
				(5.48)		
lnscitecinn						0.008
						(0.58)
Constant	2.037**	1.222***	1.572***	1.158***	-7.971***	1.225***
	(1.99)	(7.71)	(3.08)	(7.35)	(-13.38)	(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

5.3.3 The capacity of scientific and technological innovation

This paper uses the share of science and technology expenditure in total public financial expenditure to represent the capacity of scientific and technological innovation. Column (5) of Table 7 shows that β_1 is -0.291 and significant at the 5% level, which shows that the Shanxi Comprehensive Reform Area policy significantly reduced the share of science and technology expenditure in total public financial expenditure. Column (6) of Table 7 shows that δ_1 is 0.195 and significant at the 1% level, whereas θ is 0.008 and not significant even at the 10% level. Therefore, we need to verify the mediating effect of the capacity of scientific and technological innovation using the Sobel test or bootstrapping. The Sobel test statistic is -0.0024, and the corresponding p-value is 0.353, so there is no mediating effect. However, the prerequisites for the Sobel test to be valid are that the mediating effects are normally distributed and that the sample is large, and the prerequisite of a normal distribution is difficult to meet. Therefore, a bootstrap test is conducted, which shows that the mediating effect is -0.00241, with a corresponding p-value of 0.347 and a confidence interval of (-0.0074，0.0026), suggesting that the capacity of scientific and technological innovation is not a mediator of the relationship between the Shanxi Comprehensive Reform Area policy and $PM_{2.5}$ concentrations.

6 Conclusions and policy implications

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

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

Our main findings have significant implications for policy makers. First, the government should further reduce the intensity of resource exploitation and strengthen the intensity of environmental regulation in order to reduce the ecological destruction and environmental pollution caused by market entities and gradually improve environmental quality and resource utilization. Second, the Shanxi Comprehensive Reform Area policy did not promote environmental quality through scientific and technological innovation, which is an important driving force for the sustainable development of the ecological environment. Therefore, Shanxi Province should continue to promote environmentally focused scientific and technological innovations to solve prominent environmental problems and promote sustained improvements in environmental quality. Third, the treatment effects of the Shanxi Comprehensive Reform Area policy were not static but varied over time, so a dynamic and accurate assessment of its treatment effects is necessary to ensure the accuracy of policy adjustments and smooth policy operations.

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

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

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

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

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

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

References

Athey S, Imbens GW (2017) The state of applied econometrics: Causality and policy evaluation. J Econ Perspect, 31(2): 3-32. https://doi.org/10.1257/jep.31.2.3

Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. J Pers Soc Psychol 51(6): 1173-1182. http://dx.doi.org/10.1037/0022-3514.51.6.1173

Cai S, Wang Y, Zhao B, Wang S, Chang X, Hao J (2017) The impact of the "Air Pollution Prevention and Control Action Plan" on $PM_{2.5}$ concentrations in Jing-Jin-Ji region during 2012–2020. Sci Total Environ 580:197-209. https://doi.org/10.1016/j.scitotenv.2016.11.188

Cantoni D, Chen Y, Yang DY, Yuchtman N, Zhang YJ (2017) Curriculum and Ideology. J Polit Econ 125(2), 338–392. https://doi.org/10.1086/690951

Chen J, Wang B, Huang S, Song M (2020) The influence of increased population density in China on air pollution. Sci Total Environ 735:139456. https://doi.org/10.1016/j.scitotenv.2020.139456

Dong K, Sun R, Dong C, Li H, Zeng X, Ni G (2018) Environmental Kuznets curve for $PM_{2.5}$ emissions in Beijing, China: What role can natural gas consumption play? Ecol Indic 93, 591-601. https://doi.org/10.1016/j.ecolind.2018.05.045

Edwards JR, Lambert LS (2007) Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. Psychol Methods 12(1), 1–22. https://doi.org/10.1037/1082-989X.12.1.1

Eliopoulos GM, Harris AD, Bradham DD, Baumgarten M, Zuckerman IH, Fink JC, Perencevich EN (2004) The use and interpretation of quasi-experimental studies in infectious diseases. Clin Infect Dis 38:1586-1591. https://doi.org/10.1086/420936

Gan W, Xu X, Yi J (2011) Achievement, problems and countermeasures for implementation of Western Development Strategy. Economic Geography, 2011, 31(1):40-46. https://doi.org/10.15957/j.cnki.jjdl.2011.01.014 (In Chinese).

Greenstone M, Gayer T (2009) Quasi-experimental and experimental approaches to environmental economics. J Environ Econ Manage 57:21-44. https://doi.org/10.1016/j.jeem.2008.02.004

Guo J, Guo S (2019a) Has the establishment of National Comprehensive Supporting Reform Pilot Zone promoted the optimization of industrial structure — Empirical analysis based on Synthetic

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Chinese)

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

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

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

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

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

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

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

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

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

Figures

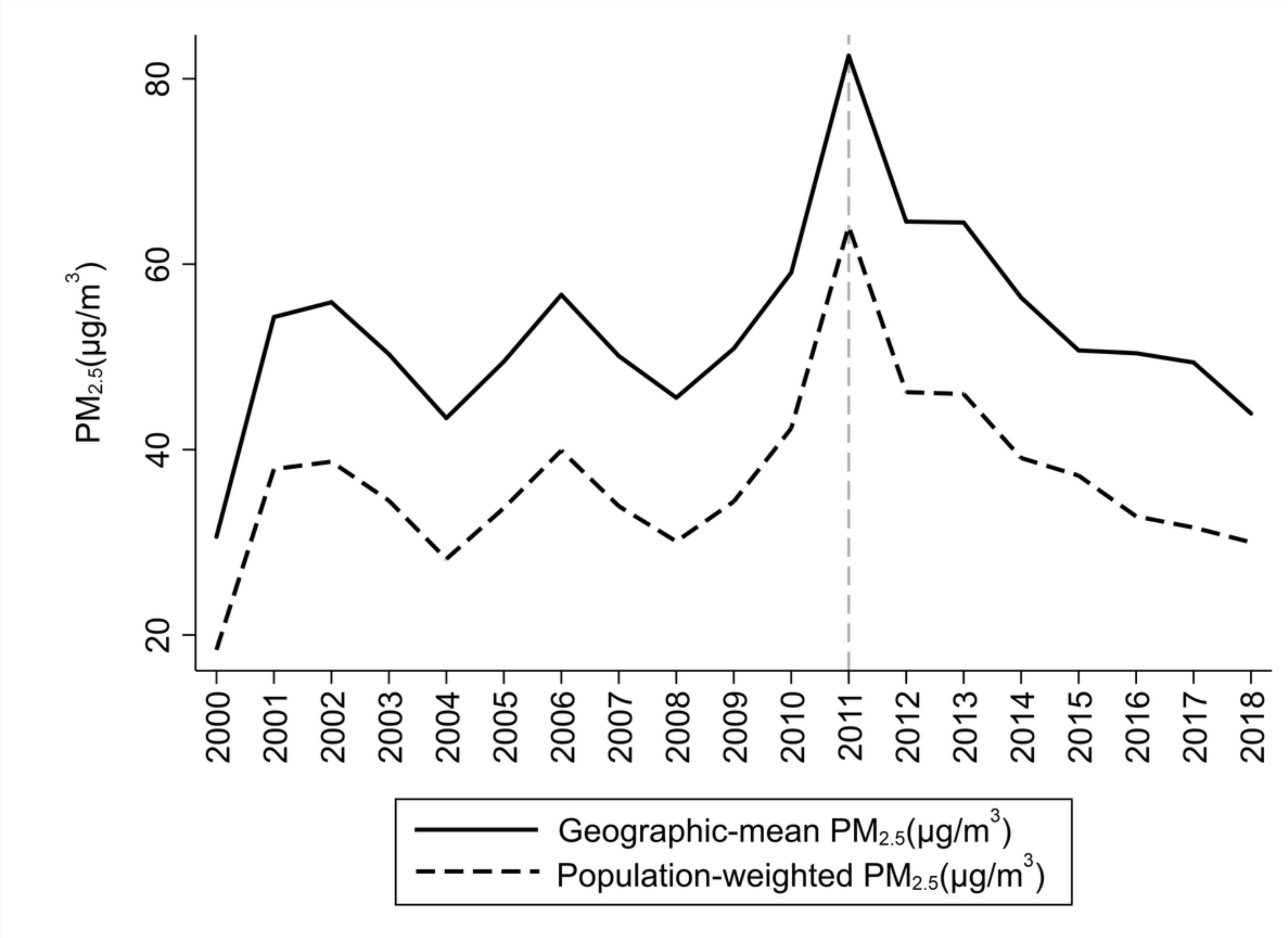

Figure 1

Geographic-Mean PM2.5 and Population-Weighted PM2.5 in Shanxi Province

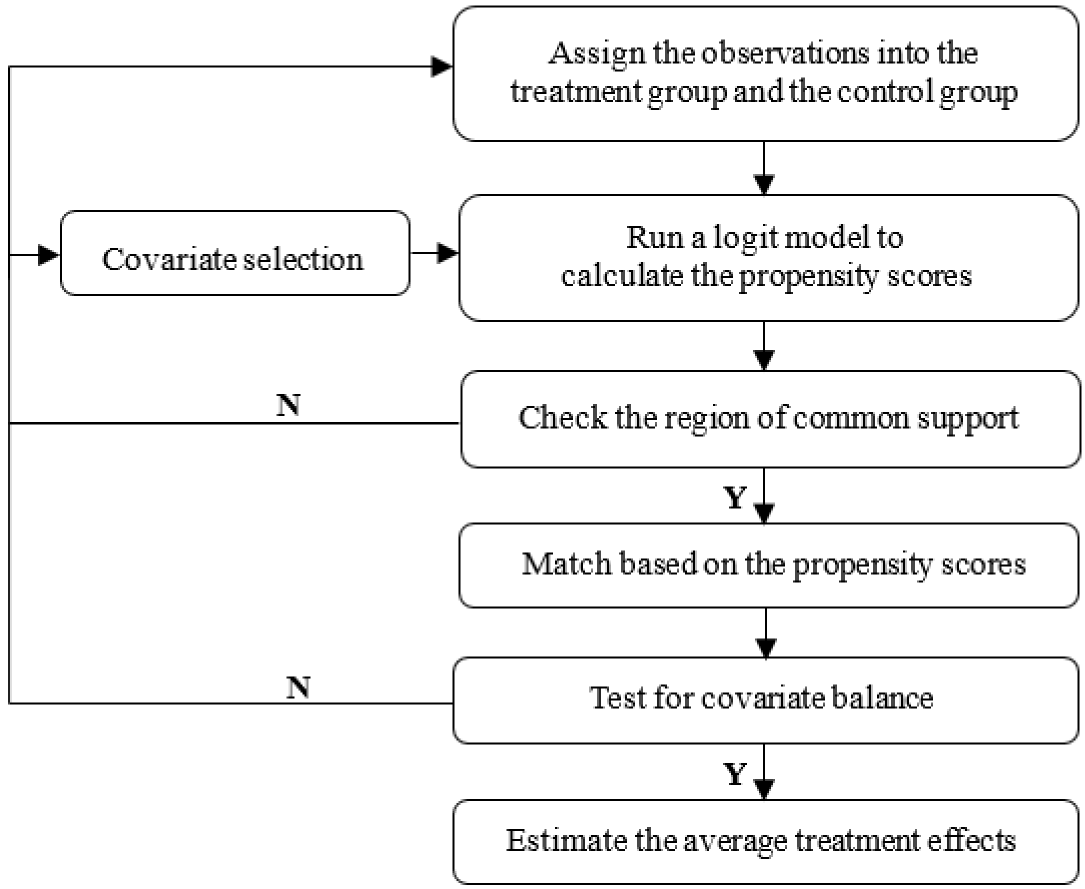

Figure 2

Steps for conducting a PSM-DID analysis

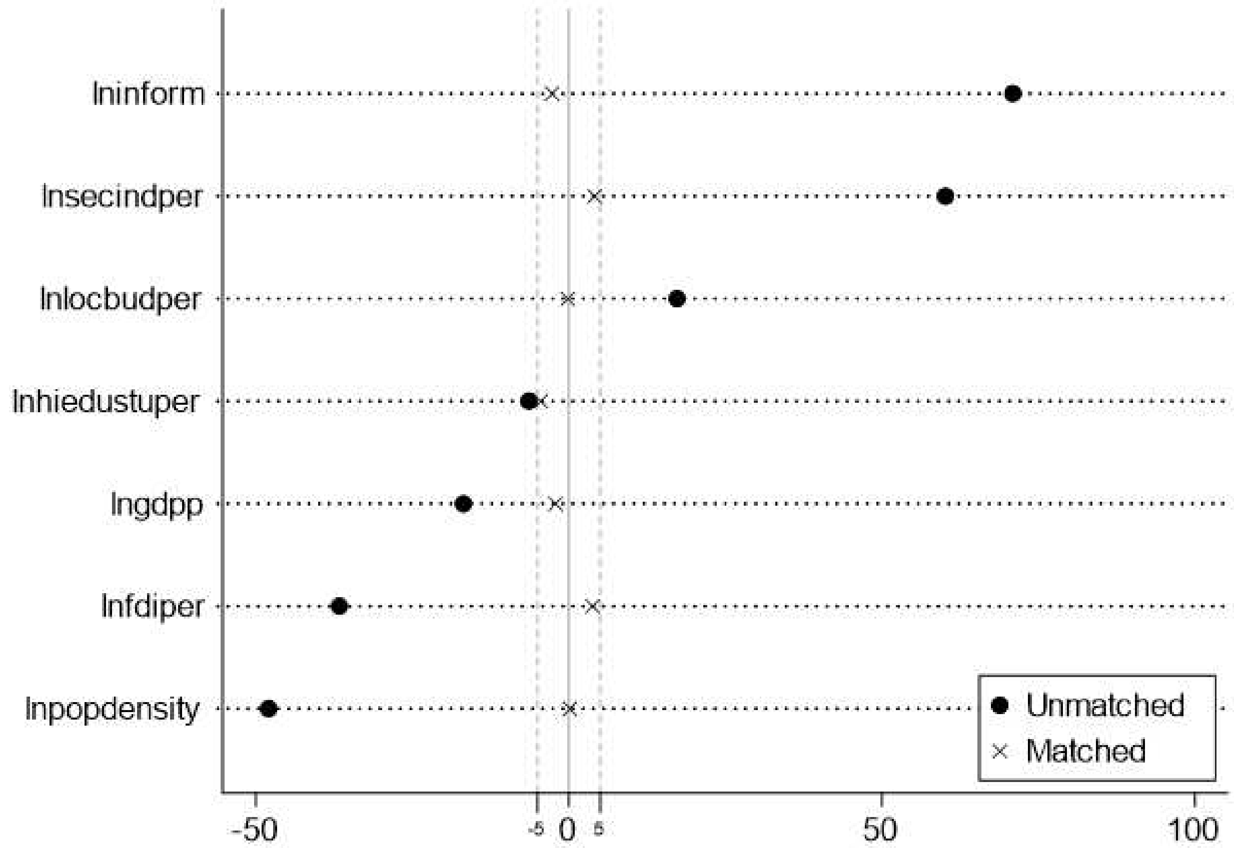

Figure 3

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

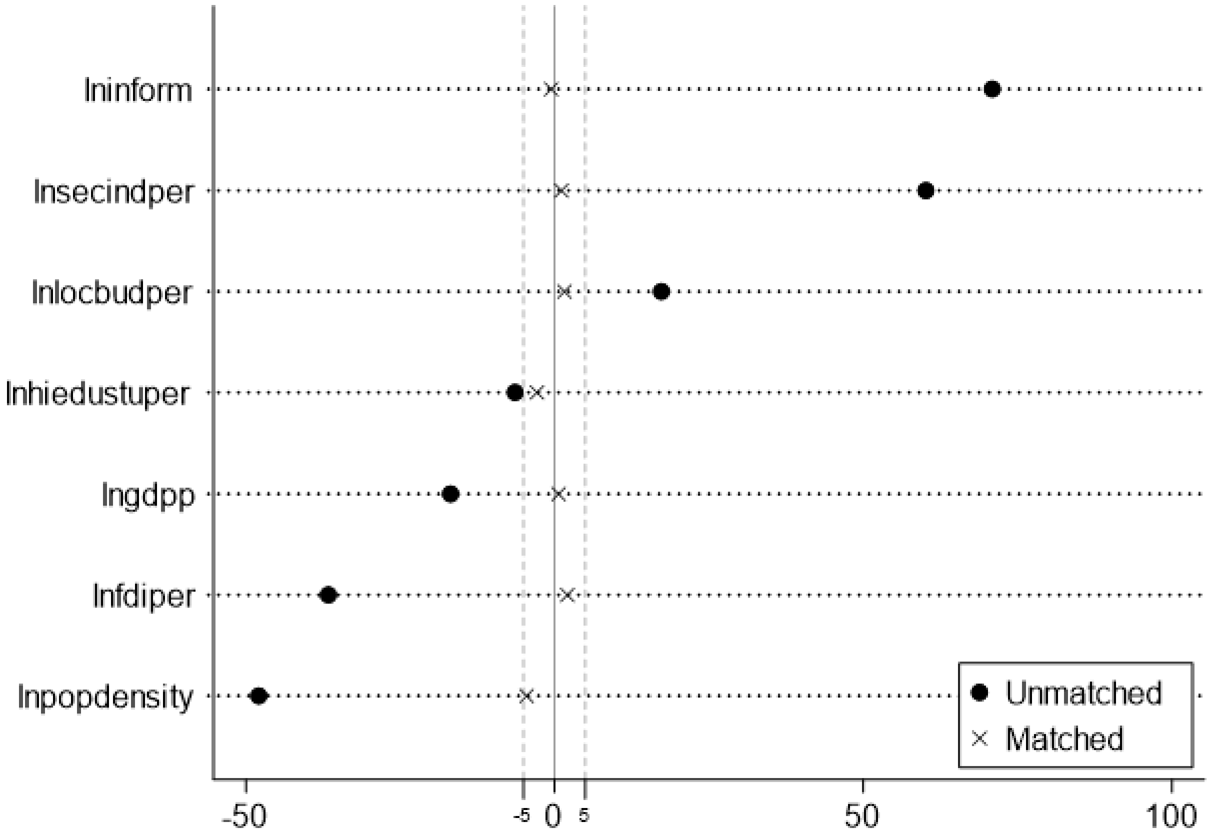

Figure 4

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

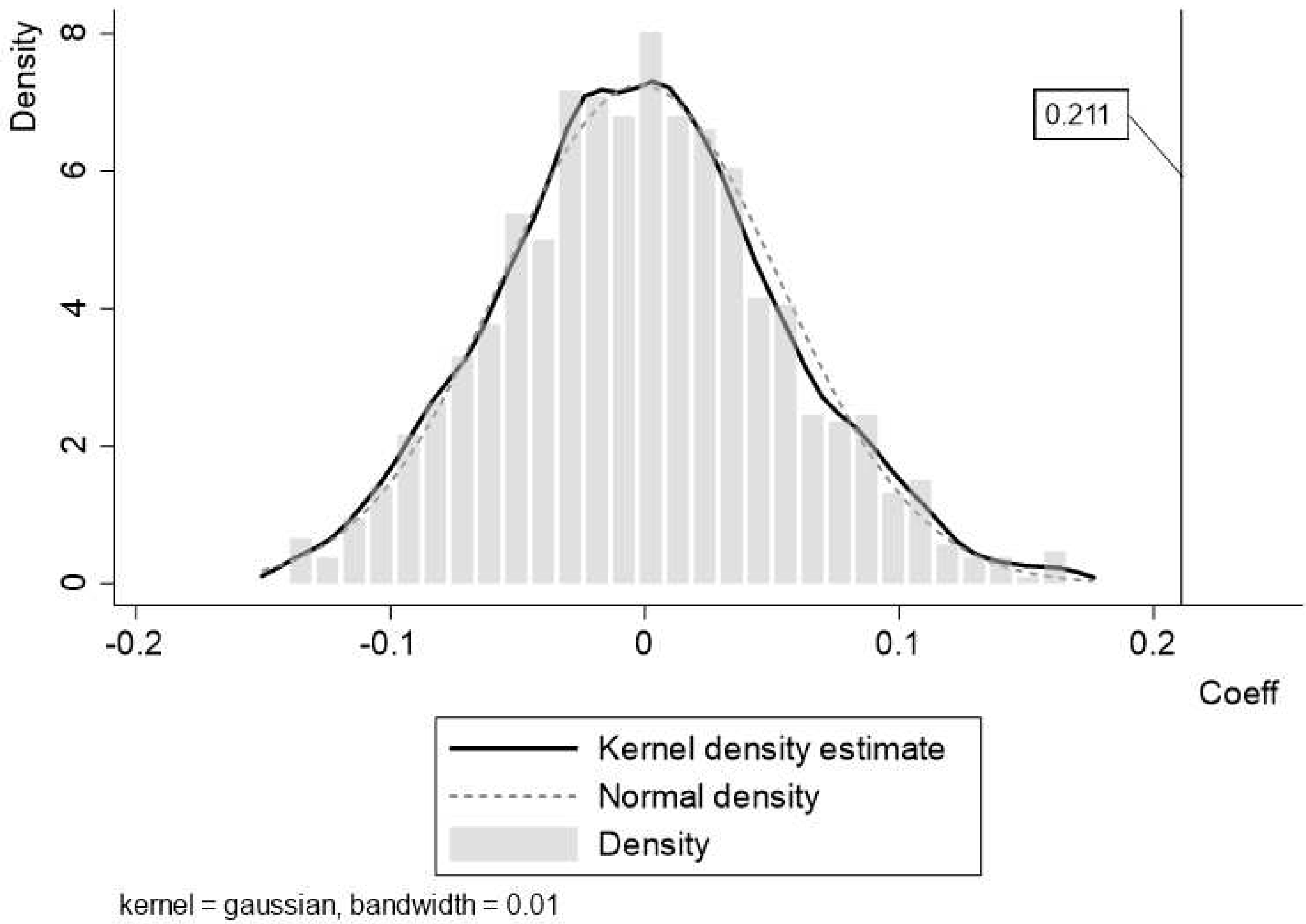

Figure 5

The distribution of 1000 coefficient estimates

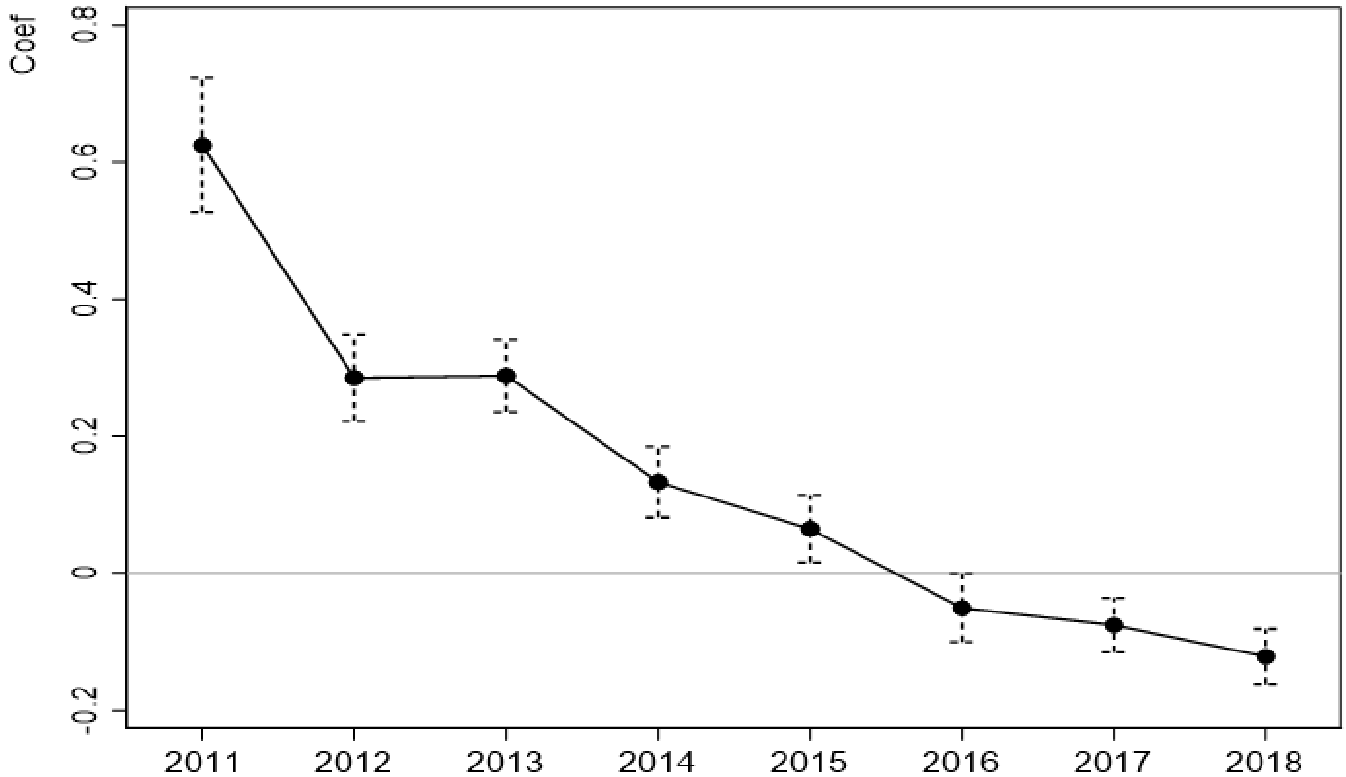

Figure 6

Trend in coefficients without covariates

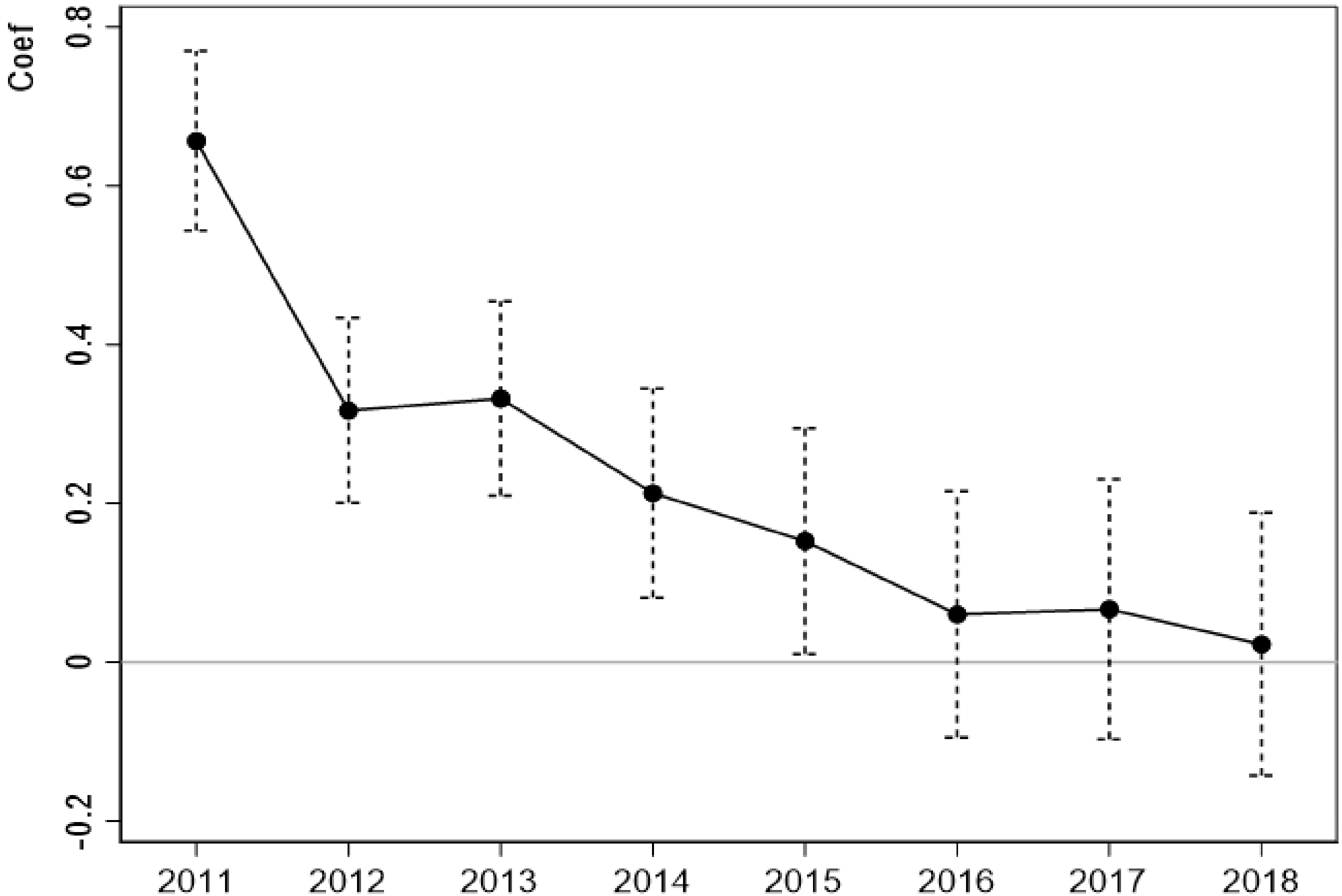

Figure 7

Trend in coefficients with covariates