

# 大气污染的大数据治理效应测度

王莉, 薛飞

(西北大学经济管理学院, 陕西西安710127)

**摘要:** 大数据治理创新了我国大气污染防治模式。现有文献尚未聚焦于该模式总体效应评价研究。通过提出大数据治理理论分析逻辑, 并基于地级市面板数据给出其总体效应的测度。研究表明, “生态环境大数据建设试点”政策实施, 显著降低试点城市AQI和PM<sub>2.5</sub>、PM<sub>10</sub>、O<sub>3</sub>等污染物排放浓度, 稳健性检验支持该结论。另外通过异质性分析发现, 东部城市的大数据治理效应更为显著, 中部的AQI和PM<sub>2.5</sub>抑制效应大于东部, 中西部对PM<sub>10</sub>浓度和O<sub>3</sub>浓度的抑制效应不明显; 同时非资源型城市的治理效应强于资源型城市, 治理需加大差异化应对。机制分析反映, 大数据治理试点政策显著提高当地政府的科学决策和监管能力, 但公众参与效应相对有限。研究结论具有推动我国环境大数据治理的参考价值。

**关键词:** 大气污染; 大数据治理模式; 治理效应; 双重差分

**中图分类号:** F062.2 **文献标志码:** A **文章编号:** 1000-2154(2023)08-0071-12

**DOI:** 10.14134/j.cnki.cn33-1336/f.2023.08.006

## Measurement of the Effect of Big Data Governance on Air Pollution

WANG Li, XUE Fei

(School of Economics and Management, Northwest University, Xi'an 710127, China)

**Abstract:** Big data governance has innovated China's air pollution control model. The existing literature has not yet focused on the evaluation study of the overall effect of this model. This paper proposes theoretical analysis logic of big data governance and measurement of its overall effect based on prefecture-level city panel data. The study shows that the implementation of the “ecological environment big data construction pilot” policy significantly reduces the AQI and the concentration of PM<sub>2.5</sub>, PM<sub>10</sub>, O<sub>3</sub> and other pollutants in the pilot cities, which is supported by the robustness test. In addition, through the heterogeneity analysis, it is found that the governance effect of big data in eastern cities is more significant, the inhibition effect of AQI and PM<sub>2.5</sub> in central China is greater than that in the east, and the inhibition effect on PM<sub>10</sub> and O<sub>3</sub> concentrations in central and western China is not obvious; at the same time, the governance effect of non-resource cities is stronger than that of resource cities, hence the greater governance necessity to cope with the differentiation. Mechanism analysis reflects that the big data governance pilot policy significantly improves the local government's scientific decision-making and regulatory capacity, but the public participation effect is relatively limited. The relevant data have reference value for promoting environmental big data governance in China.

**Key words:** air pollution; big data governance model; governance effect; difference-in-differences

**收稿日期:** 2023-05-16

**基金项目:** 国家社会科学基金重大项目“工业大数据统计测度理论及应用研究”(21&ZD153); 全国统计科学研究项目“数字经济驱动我国农业高质量发展的统计研究”(2021LY068)

**作者简介:** 王莉, 女, 副教授, 经济学博士, 主要从事经济统计测度与政策有效性评价研究; 薛飞(通讯作者), 男, 讲师, 经济学博士, 主要从事环境经济学及区域政策评估研究。

## 一、引言

作为关系中华民族永续发展的根本大计,我国新时代生态文明建设从理论到实践都发生了历史性、转折性、全局性变化,成就举世瞩目,成为新时代党和国家事业取得历史性成就、发生历史性变革的显著标志。这一历史成就的创建,起步于群众反映强烈的大气雾霾污染突出问题的治理解决。2013年国家出台《大气污染防治行动计划》。习近平总书记则在当年中央经济工作会议上进一步强调,要加大环境治理和生态保护工作力度、投资力度、政策力度,加强污染物减排特别是大气污染防治,推进重点行业、重点区域大气污染治理,加强区域联防联控,把已经出台的大气污染防治十条措施真正落到实处。由此拉开国家环境大治理的大幕,开启生态文明建设的历史进程。进而基于不断深化对生态文明建设规律的认识,形成新时代中国特色社会主义思想,实现由实践探索到科学理论指导的重大转变。<sup>①</sup>

大气污染危害公众健康,影响范围极大。致污因素及机制复杂,至少可归纳为三方面。一是高速工业化的影响。传统工业生产技术是以能量物质转换为目的的,典型的如钢铁、柴油运输等生产。其能量物质转换过程必然产生废弃物即为污染。解决污染唯一途径就是生产的新技术转型。但由于市场体系不具备相应技术转型补偿机制,造成市场失灵。其污染的外部性,只能通过政府干预解决。二是自然环境因素,包括大气流动和地形地势造成的随机性频繁影响。三是社会自然因素的交叉影响。包括城市化与特殊地形造成的雾霾小区域聚集,以及冬季北方大范围燃煤取暖等。上述成因与动态变化带来的巨大挑战,唯有国家干预方能应对。需要解决的问题至少包括大气污染的动态水平监测,治理决策科学性与执行落实程度的识别,基于地理区位、资源禀赋、经济发展水平巨大差异现实,评价地方政府如何落实公众监督与治理主体责任的,等等。凡此,均需要统一规范的数据研究与表达,做到心中有数。国家基于大数据发展战略迅速出台《生态环境大数据建设总体方案》,<sup>②</sup>布局覆盖各地区的雾霾监测体系,集中重点整合2013年以来雾霾监测历史数据,同步集成气象、遥感、排放清单、城市源解析和环境执法数据,形成雾霾案例知识库;构建起应对大气污染的大数据治理推动模式。这里的大数据治理具体指,统筹全国大数据平台建设,基于大数据环保云及管理和应用平台,推进大气环境相关数据资源整合共享,加强科学决策,创新监管模式,完善公共服务的大气污染治理模式。

大气污染的大数据治理模式是政府治理体系一次成功的数字化转型创新实践。其一经推出,即以大数据技术如何发挥有效驱动雾霾污染治理的统筹作用,引发各方面关注。一些研究考察了各类数字信息技术对大气污染物减排和空气质量改善的影响<sup>[1-3]</sup>及其作用机制<sup>[2]</sup>。还有一些将其扩展到生态环境大数据的内涵、框架设计、趋势与挑战、数据应用等方面的理论研究<sup>[4-7]</sup>,以及基于智慧城市建设<sup>[8-9]</sup>、国家级大数据综合试验区<sup>[10]</sup>等大数据应用实践在大气污染治理和减排的积极作用方面的研究。这些侧重某一方面的相关研究对大数据治理实践发挥了积极指导作用,同时也驱动研究向大数据治理建设政策总体效果评价方面聚焦,进一步推动新时代中国特色社会主义思想实践探索。

本文以我国大气污染的大数据治理总体效应测度为主题,力图给出两方面贡献。一方面,提出大数据治理概念,建立其基于信息驱动而非物质转换驱动的治理总体效应测度分析逻辑,给出大数据技术在大气环境治理中总体作用的理论表达。另一方面,采用基于分散的地区数据测度总体治理效应技术路线,构建准自然实验模型,抓住大数据治理政策地区试点环节,通过试点与非试点城市,政策实施前后的双重差分技术,得到基于建设试点大气污染治理效应的总体评价。进一步,通过异质性分析模型构建,考察大气污染的大数据治理的异质性效应;同时从政府治理能力和公众监督提升视角探析大气环境的大数据治理作用机制。

<sup>①</sup>引自2023年7月18日全国生态环境保护大会上习近平重要讲话, [https://www.mee.gov.cn/ywdt/szyw/202307/t20230718\\_1036581.shtml](https://www.mee.gov.cn/ywdt/szyw/202307/t20230718_1036581.shtml)。

<sup>②</sup>2016年3月7日,环境保护部印发《生态环境大数据建设总体方案》,具体内容见网址 [https://www.mee.gov.cn/gkml/hbb/bgt/201603/t20160311\\_332712.htm](https://www.mee.gov.cn/gkml/hbb/bgt/201603/t20160311_332712.htm)。

## 二、研究设计

### (一) 理论分析逻辑

这里梳理出两方面基本逻辑作为测度研究的分析基础。一是大数据治理如何成为大气环境治理的必然优选模式。二是大数据治理模式怎样有效发挥夯实地方政府主体责任作用。

大数据治理模式是计算机信息系统技术成功应用于政府大气环境治理的表达。基于大气污染等成因及其市场缺失相应治理机制,政府在推动经济社会高质量发展的同时,也必须肩负起治理污染,供给大众环境公共品包括清洁空气的历史责任。基于理论视角考察,如同所有政府行为一样,其相关治理干预均需要基于充分对称信息理性行为,才能实现资源有效配置,取得切实成效。就是说,大气环境的大数据治理是整个治理体系中,力图对所有参与政府干预的行为者提供统一的充分对称的信息环境,使之理性行为的措施。虽然大数据不具备直接改变或消除污染物质的技术功能,但可通过提供相关信息促使去污技术有的放矢、效率提升,最终改善大气环境。显然,基于资源配置的有效性来说,大数据治理是不可或缺的。本文的大气环境大数据治理实践总体效应测度,就是力图对这一模式作用发挥的效应给予检验与评价。当然,实现大数据治理总体效应的测度操作,需要界定诸如何为治理行为者的理性行为,何为统一的充分对称的信息环境,何为治理的总体效应等一系列概念,以及进一步设定测度研究的基本分析逻辑。

首先,设地级城市政府为大数据治理的行为主体,理论上其通过参加国家环境大数据试点建设,获得统一的充分对称的信息,指导其理性治理行为。该设定的现实依据为:其一,我国的大国分级行政管理体制,国家政策与法规落实到地方政府推动执行。特别是中央新的重大决策基本上都是通过试点取得经验再推广全国的。环境的大数据地方试点建设仍然遵从这一传统,这也是夯实地方政府主体责任的制度设计。其二,鉴于大数据治理的技术本质是信息驱动而非物质转换驱动的治理,其相关治理效应表达的是治理行为人在一系列大气致污因素,以及经济社会发展大环境条件下,采用各类具体治理技术行为的综合结果。基于地方政府行为,获取大数据治理总体效应测度信息,逻辑上是合理自洽的。其三,国家环境大数据试点建设包括构建大数据环保云及管理应用平台,推进大气环境相关数据资源整合共享,加强科学决策,创新监管模式,完善公共服务等一系列活动。建设试点的上述活动能够比较充分代表大数据治理行为。

其次,设定大数据治理总体效应的分析逻辑。2013—2022年恰为大气污染防治十年。十年间我国GDP总量增长了69%, $PM_{2.5}$ 浓度下降了57%,重污染天数减少了92%,二氧化硫浓度降至个位数。实现经济快速增长的同时,空气质量明显改善,人民群众生态环境获得感显著提升。面对这一历史性成就,习近平总书记在主持中央政治局第二十九次集体学习时仍然告诫我们,<sup>①</sup>当前我国生态文明建设仍然面临诸多矛盾和挑战,生态环境稳中向好的基础还不稳固,从量变到质变的拐点还没有到来。就是说,大气治理仍在推进中。其要求大数据治理效应的测度聚集在试点建设推进过程的检查。基于总体效应评价分析,发现试点工作存在问题积累经验,提出政策改进空间。相关分析逻辑包括:其一,建立大气质量与地方建设试点的关系,作为治理效应的测度。当然需要给出该测度的有效性检验。其二,通过地区异质性和机制检验,得到夯实地方政府大数据治理主体责任的推进机制。

### (二) 测度模型

地方政府如何基于顶层设计发挥主体责任,推进大数据技术应用于大气环境治理,这里通过治理实践予以考察。原环保部在2016年发布《生态环境大数据建设总体方案》后,重点提出了要开展地方大数据应用试点,各地积极探索应用模式,推动试点成果的推广和实施,开启了大气污染的大数据治理新模式,对于提升环境治理监管效率、提高决策科学性和有效性发挥了重要作用<sup>[2]</sup>。基于理论分析逻辑,那么生态环境

<sup>①</sup>2021年4月30日,中央政治局就新形势下加强我国生态文明建设进行第二十九次集体学习,具体内容见网址 <http://cpc.people.com.cn/n1/2021/0502/c64094-32093920.html>。



大数据建设方案的实施和应用试点的推行,不仅可改善试点城市自身空气质量,还将使空气质量在试点城市和非试点城市之间产生差异,进而识别出生态环境大数据建设对改善城市空气质量的总体效应。本文将“生态环境大数据建设试点”视为一项外生政策冲击,采用双重差分法评估大数据治理对大气污染防治的现实效应。设定基准模型如下:

$$Y_{it} = \alpha + \beta Treat_i \times Post_t + \gamma X_{it} + \vartheta_i + \mu_t + \varepsilon_{it} \quad (1)$$

其中, $Y_{it}$ 表示城市*i*第*t*年的大气污染水平变量; $Treat_i$ 表示生态环境大数据建设试点政策的虚拟变量,若某一地区被获批生态环境大数据建设试点赋值为1,否则赋值为0; $Post_t$ 为政策试点前后的虚拟变量,政策实施当年及之后的年份赋值为1,之前的年份为0。交互项  $Treat_i \times Post_t$  用于评估生态环境大数据建设的大气污染治理效应。 $X_{it}$ 一系列控制变量,包括气象因素、城市规模、产业结构、政府因素等; $\vartheta_i$ 和 $\mu_t$ 分别表示城市固定效应和年份固定效应; $\varepsilon_{it}$ 为随机误差项。在上式中,主要关心的估计系数是 $\beta$ ,若 $\beta < 0$ 且统计显著,则意味着生态环境大数据建设显著提高了试点城市的空气质量,说明大数据治理提升了大气污染治理效果。

### (三) 变量选择与数据处理

为了合理测度与评价基于建设试点的大气污染防治的总体效应,同时考虑到遗漏变量可能造成的内生性问题,还纳入了相关控制变量。各变量设置如下:

**1. 大气污染水平。**区别于大多数研究采用哥伦比亚大学社会经济数据和应用中心或圣路易斯华盛顿大学大气成分分析小组所公布的 $PM_{2.5}$ 浓度数据集,本文锚定2035年“美丽中国建设目标基本实现”以及“十四五”时期“生态文明建设实现新进步”对大气污染指标设定的目标值,并结合《中共中央国务院关于深入打好污染防治攻坚战的意见》提出的既要继续大幅降低 $PM_{2.5}$ 浓度,又要有效遏制臭氧浓度上升趋势,还要努力消除重污染天气。最终选取生态环境部披露的空气质量指数(AQI)及其分项污染物 $PM_{2.5}$ 、 $PM_{10}$ 、 $O_3$ 等指标,以便更为全面地考察中国大数据治理的效应。考虑到绝大多数城市自2014年起开始发布空气质量数据,本文搜集整理生态环境部公布2014年1月1日—2021年12月31日的各地级市空气质量状况原始数据,并统一将其进行平均化处理,并转换为年度数据。

**2. 政策冲击。**根据原环保部公布的生态环境大数据建设试点名单,吉林省、贵州省、江苏省、内蒙古自治区环境保护厅及武汉市、绍兴市等列为全国首批生态环境大数据建设试点单位,设置生态环境大数据建设试点的政策虚拟变量。对各样本单位进行统一赋值,若某一城市获批生态环境大数据建设试点的当年及之后各年取值为1,否则为0。

**3. 控制变量。**导致大气污染的成因复杂,不仅有气象变化等自然因素的随机性影响,又有与经济社会环境因素的交叉影响。本文还纳入了气象和城市经济社会特征两类控制变量。

(1)气象变量。污染物在大气中的运动,与大气的气象条件密切相关。直接影响大气污染的大气层是大气边界层,在边界层内,风向风速、温度、湿度随高度的变化很大<sup>[11]</sup>。大气污染严重的情况往往发生在逆温和静风条件下,气温升高、降水减少、平均风速降低等均不利于污染物的扩散。在利用原始气象数据与大气污染数据进行非参数拟合回归探究相关关系后,同时也借鉴赵仁杰等<sup>[12]</sup>的研究,本文筛选了各地区平均温度、平均风速、年降雨量、相对湿度等气象指标进行控制。

(2)城市经济社会特征。较多研究已关注了经济集聚、城镇化、财政金融制度等因素对大气污染的影响,作为城市经济社会发展的基本要素包括经济和人口规模、财政投入和金融发展水平以及政府支撑科技的力度等宏观变量,对相应城市大气污染水平及其环境治理进程也会产生诸多影响。与孙传旺等<sup>[13]</sup>、宋弘等<sup>[14]</sup>的研究相似,分别以实际GDP的对数、城市年末总人口、公共财政支出占地区生产总值的比重、财政科技支出占财政支出的比重、金融贷款规模占GDP的比值等指标来衡量城市经济社会基本特征,将这些指标作为影响的其他控制变量。

本文采用的是2014—2021年288个地级市的面板数据。大气污染数据来自生态环境部;生态环境大数据建设试点名单来自生态环境部网站,由作者搜集整理而得;气象数据来自美国国家气候数据中心(NC-DC);城市经济特征数据均来源于《中国城市统计年鉴》《中国区域统计年鉴》和各省《统计年鉴》。表1是主要变量描述性统计结果。

表1 主要变量描述性统计

	变量	观测值	均值	标准误	最小值	最大值
被解释变量	AQI	2141	70.497	21.606	23.880	178.053
	PM <sub>2.5</sub> 浓度	2141	42.550	16.372	10.459	129.784
	PM <sub>10</sub> 浓度	2141	75.389	28.892	21.760	233.958
	O <sub>3</sub> 浓度	2141	59.390	9.998	26.576	89.416
解释变量	生态环境大数据建设试点	2141	0.100	0.301	0.000	1.000
机制变量	环境行政处罚案件数	1824	13.717	2.019	4.981	19.374
	单位 GDP 二氧化硫排放量	2131	-4.951	1.205	-9.706	-1.534
	每万人环境污染百度指数	2099	-3.313	0.872	-6.472	-0.697
气象变量	平均温度	2140	2.680	0.445	-0.930	3.255
	年降雨量	2141	6.830	0.610	3.901	7.923
	相对湿度	2141	0.920	0.259	0.153	1.800
	平均风速	2141	70.156	10.026	36.970	84.459
经济社会特征	实际 GDP 的对数	2131	7.558	0.937	5.033	10.674
	年末总人口	2113	5.910	0.707	3.008	8.136
	公共财政支出占 GDP 比重	2111	2.146	2.279	0.058	20.621
	财政科技支出占财政支出比重	2117	19.405	10.802	2.686	87.174
	金融贷款规模占 GDP 比值	2130	1.168	0.637	0.151	9.623

三、实证结果与检验

(一) 基准回归结果

基于建设试点政策效应估计结果,来检验大数据治理对大气污染所发挥的作用。根据表2信息,列(1)至列(4)分别是以 AQI、PM<sub>2.5</sub> 浓度、PM<sub>10</sub> 浓度、O<sub>3</sub> 浓度为被解释变量的回归结果。可以发现,所有回归结果的估计系数均为负值,分别为 -4.451、-3.061、-4.919 和 -2.185,且均通过了至少 5% 的显著性检验,这表明生态环境大数据治理显著降低试点城市大气污染程度。即与非试点城市相比,生态环境大数据建设政策推进过程中,使得试点城市年均 AQI 下降了 4.451,PM<sub>2.5</sub> 浓度下降 3.061ug/m<sup>3</sup>、PM<sub>10</sub> 浓度下降 4.919ug/m<sup>3</sup>,O<sub>3</sub> 浓度下降 2.185ug/m<sup>3</sup>。这一结论与既有文献<sup>[10]</sup>的结论较为一致,肯定了大数据治理在改善空气质量中的积极作用。

表2 基准回归结果

被解释变量	AQI	PM <sub>2.5</sub>	PM <sub>10</sub>	O <sub>3</sub>
	(1)	(2)	(3)	(4)
$Treat_i \times Post_t$	-4.451 *** (1.416)	-3.061 ** (1.337)	-4.919 ** (2.194)	-2.185 ** (1.006)
城市控制变量	控制	控制	控制	控制
气象控制变量	控制	控制	控制	控制
城市固定效应	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制
观测值	2079	2079	2079	2079
R <sup>2</sup>	0.932	0.904	0.917	0.789

注：\*、\*\*、\*\*\* 分别表示在 10%、5%、1% 的统计水平下显著；括号内的数值是以城市聚类的稳健标准误；城市经济特征控制变量包括实际 GDP 对数、年末总人口、财政支出占 GDP 比重、财政科技支出占财政支出比重以及金融贷款与 GDP 比值；气象特征控制变量包括城市平均温度、平均风速、年降雨量、相对湿度。下同。因篇幅所限,这里只列出关键的 AQI、PM<sub>2.5</sub>、PM<sub>10</sub>、O<sub>3</sub> 四项指标的估计结果,其他数据信息可向作者备案。

## (二) 平行趋势检验和动态效应

采用双重差分法估算得到有效的大数据治理效应结果的关键前提在于平行趋势假设是否满足,即政策实施之前的处理组与控制组在大气污染水平变动上是否呈现出相同的趋势。同时,大数据治理对大气污染水平的影响作用并不是一蹴而就,随着时间的推移政策可能存在动态效应。借鉴相关研究的处理方式,采用动态双重差分法来检验平行趋势以及政策的动态效应。

$$Y_{it} = \alpha + \beta_k \sum_{k=2014}^{2021} Treat_i \times Post_{tk} + \gamma X_{it} + \vartheta_i + \mu_t + \varepsilon_{it} \quad (2)$$

上式中, $k$ 表示生态环境大数据建设试点实施后的年份。采用政策实施前一年作为基期,即上式中 $k \neq 2015$ 。本文所关注的是估计系数 $\beta_k$ ,反映生态环境大数据建设试点建设第 $k$ 年,试点城市与非试点城市之间的大气污染水平差异。平行趋势检验和动态效应检验的结果如图1所示。可以发现,无论以AQI还是 $PM_{2.5}$ 、 $PM_{10}$ 、 $O_3$ 其余三个分项污染物作为被解释变量, $\beta_{2014}$ 在10%显著性水平上并不显著,表明试点城市与非试点城市在政策实施前并无显著差异,即研究样本满足平行趋势假设。当 $k \geq 2016$ 时, $\beta_k$ 开始出现大幅下降,而这一抑制效应在2020年开始消失。上述结果再次验证了生态环境大数据建设显著改善了试点城市的空气质量状况,且这一抑制效应具有一定持续性。

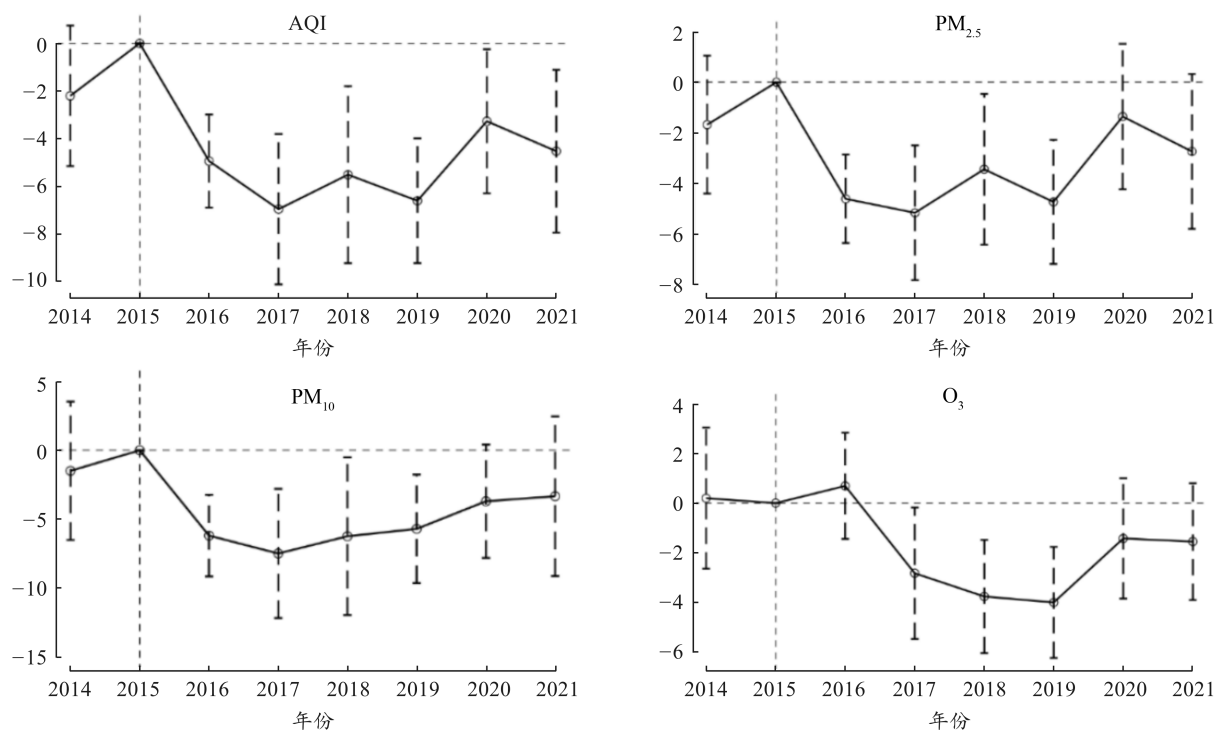


图1 平行趋势检验和动态效应

注:图中纵轴为估计系数;空心圆圈表示 $\beta_k$ 的估计值;虚线表示95%置信区间。

## (三) 安慰剂检验

采用双重差分法进行政策评估的另一个可能的担忧是,由于分析过程中难以控制所有其他政策因素,估计结果的统计显著有可能是来自某些其他随机性因素。为了减少随机因素的影响,这里通过两方面的安慰剂检验来增强估计的稳健性。先借鉴Chetty等<sup>[15]</sup>的做法,按照真实的生态环境大数据建设试点的实施情况,同比例随机生成一个生态环境大数据建设试点城市名单,从而产生一个虚拟的估计系数 $\hat{\beta}^{random}$ ,将上述过程重复进行500次,通过比较得到的500个虚拟估计系数 $\hat{\beta}^{random}$ 与真实估计系数是否存在显著差异,可以判断是否存在其他随机因素对基准估计结果产生影响。图2绘制了 $\hat{\beta}^{random}$ 的分布,可以发现, $\hat{\beta}^{random}$ 分布在零附近,且距离基准回归结果较远,说明通过生态环境大数据建设试点的实施,大数据治理对大气污染水

平的影响并未受到其他随机因素的干扰。

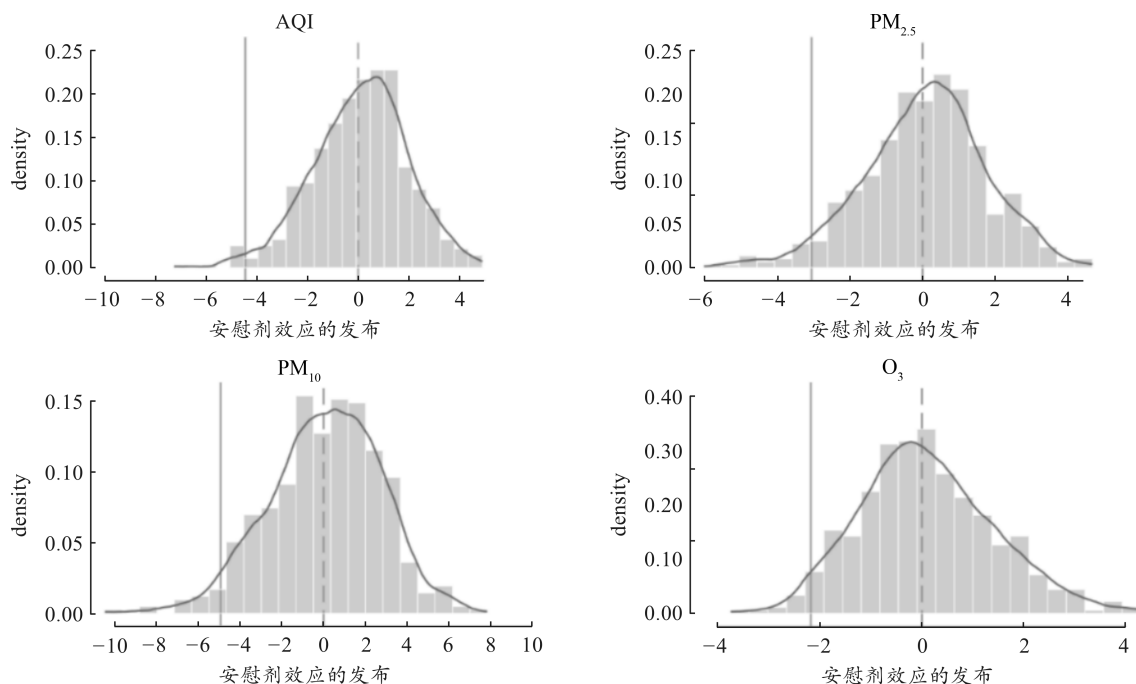


图2 安慰剂检验

此外,借鉴刘瑞明和赵仁杰<sup>[16]</sup>、王锋和葛星<sup>[17]</sup>的做法,通过改变政策实施时间来进行安慰剂检验。具体地,分别将生态环境大数据建设试点的实施年份提前一年,如果此时生态环境大数据建设变量并不显著为负,则说明大气污染水平下降来源于生态环境大数据建设;反之,若生态环境大数据建设变量仍旧显著为负,则说明大气污染水平下降受到其他随机性因素的干扰。表3列(1)是政策实施时间提前一年的结果,可以发现,将试点时间提前一年并不会对大气污染水平产生显著的影响,这间接说明大气污染水平的下降不是由其他因素导致的,而是来自生态环境大数据建设试点的大数据治理成效。

#### (四) 稳健性检验

上文回归结果表明通过大数据治理能够降低试点城市大气污染水平。为了进一步排除混淆因素的干扰,以式(1)的双重差分模型为基准,进行以下四个方面的稳健性检验。

**1. 样本缩尾处理。**对被解释变量的研究样本进行上下1%的缩尾处理,以此排除异常值对结果的可能影响。表3列(2)的估计结果表明,在排除异常值干扰后,生态环境大数据建设的估计系数仍旧在5%水平上显著为负,且估计系数数值与基准回归结果相似,这进一步增强了基准结果的稳健性。

**2. 样本重新筛选。**本文是基于全国所有城市的样本开展分析,但不同城市的行政级别以及污染治理强度可能影响估计结果,因此对样本进一步筛选。一方面,考虑到直辖市和副省级城市在行政级别与初始禀赋上高于其他一般地级市,故剔除了这两类城市,估计结果见表3列(3)。另一方面,考虑到京津冀地区作为全国大气污染重灾区,实施了一揽子的大气污染治理政策,纳入这部分样本也可能造成估计结果偏误。所以,剔除京津冀地区的所属城市后重新回归,估计结果见表3列(4)。可以发现,在样本重新筛选后主要估计系数依旧显著为负,证明估计的结果并未受到样本选择的影响。

**3. 排除其他政策的干扰。**为了避免样本期间其他环境政策或大数据政策影响大气污染治理效应,从而造成基准回归结果出现偏误,因此搜集了自2014年起基于城市层面的大型环境政策或大数据政策,主要包括自2017年起实施的冬季清洁取暖试点城市政策、2016年起实施的国家级大数据综合试验区政策等。将这些政策虚拟变量的交互项加入模型,回归结果见表3列(5)。不难发现,除 O<sub>3</sub> 外,以 AQI 或其他污染物为被



解释变量时,列(5)中的估计系数与基准回归结果较为相似,表明上述两项政策并未造成估计结果的偏误。

4. 倾向匹配得分双重差分法。采用双重差分法虽然可以控制试点城市与非试点城市在政策实施前后的相对差异,但控制样本选择偏差所引起的内生性仍未解决。为缓解样本选择偏差,采用倾向匹配得分双重差分模型,先利用 Logit 模型估计得到的倾向得分来进行近邻匹配,基于匹配后的样本重新进行回归,以缓解样本选择偏差。结果见表3列(6),与基准回归结果一致,且4个空气污染指标都至少通过了5%的显著性检验,这无疑再次强化了本文基准回归结果的可靠性。

表3 稳健性检验

	时间安慰剂检验	上下1%缩尾	剔除特殊城市	剔除京津冀地区	考虑其他政策的影响	倾向匹配得分 双重差分法
	(1)	(2)	(3)	(4)	(5)	(6)
$AQI$	-2.404 (1.892)	-5.199*** (1.349)	-3.892*** (1.430)	-5.789*** (1.349)	-4.929*** (1.456)	-3.477*** (1.055)
$PM_{2.5}$	-1.509 (1.814)	-3.689*** (1.284)	-2.453* (1.345)	-4.168*** (1.281)	-3.801*** (1.356)	-2.172** (0.951)
$PM_{10}$	-3.241 (3.126)	-5.973*** (2.109)	-4.300* (2.307)	-6.770*** (2.109)	-5.852*** (2.194)	-3.163** (1.568)
$O_3$	-2.089 (1.413)	-2.087** (1.002)	-2.059* (1.079)	-2.069** (1.019)	-1.509 (0.972)	-2.837*** (0.821)

注: \*、\*\*、\*\*\*分别表示在10%、5%、1%的统计水平下显著;上述结果中的数值为交互项的估计系数,括号内的数值是以城市聚类的稳健标准误;回归结果均控制城市和年份固定效应。

四、拓展性分析

(一) 异质性分析

上文基于样本整体分析了生态环境大数据建设试点政策对大气污染的总体治理效应,其是代表性的是平均的。但对于夯实地方政府主体责任而言,需要差异性信息支持。这里从地理区位和资源禀赋两个方面进行治理的异质性分析。

1. 地理区位异质性。考虑到各城市在经济发展和对外开放水平等方面的差异,按照地理位置分为东、中、西部城市,并在公式(1)的基础上分别加入中部地区(Middle)和西部地区(West)的虚拟变量与  $Treat_i \times Post_t$  的交互项,具体模型如下所示:

$$Y_{it} = \alpha + \beta_1 Treat_i \times Post_t + \beta_2 Treat_i \times Post_t \times Middle + \beta_3 Treat_i \times Post_t \times West + \gamma X_{it} + \vartheta_i + \mu_t + \varepsilon_{it} \quad (3)$$

上式中,估计系数  $\beta_1$  表示生态环境大数据建设对东部城市大气污染的影响; $\beta_1 + \beta_2$  表示生态环境大数据建设对中部城市大气污染的影响; $\beta_1 + \beta_3$  分别表示生态环境大数据建设对西部城市大气污染的影响。回归结果如表4所示。可以发现,  $Treat_i \times Post_t$  的估计系数均显著为负,说明生态环境大数据建设对东部地区大气污染具有显著影响。以东部为比较基础,观察  $AQI$  和  $PM_{2.5}$  浓度两项指标的估计结果,发现  $Treat_i \times Post_t \times Middle$  的估计系数显著为负,说明大数据治理对中部地区  $AQI$  和  $PM_{2.5}$  浓度的抑制效应大于东部地区。同时发现  $Treat_i \times Post_t \times West$  的估计系数显著为正,且绝对值大于  $Treat_i \times Post_t$  的估计系数,表明生态环境大数据建设尚未在西部地区发挥明显的治理效应;而以  $PM_{10}$  浓度和  $O_3$  浓度为被解释变量时,  $Treat_i \times Post_t \times Middle$  和  $Treat_i \times Post_t \times West$  的估计系数并不显著,表明生态环境大数据建设对中部和西部地区上述两类污染物的治理效应并不明显。这不难理解,东部地区的大数据发展基础较好,数据产量、存量以及其大数据应用水平均高于西部地区,因此东部地区大气污染的大数据治理效应较为明显;中部地区关于雾霾的大数据治理效应较强,但由于中部地区特定的气象条件(例如高温、强烈日照)有利于臭氧的生成和累



积,且臭氧污染机制较为复杂,这给中部地区臭氧治理造成挑战。西部地区生态环境状况较为复杂,生态环境大数据建设的目标呈现出多样化特点,因而其大气污染治理效应并不明显。

表4 地理区位异质性

解释变量	AQI	PM <sub>2.5</sub>	PM <sub>10</sub>	O <sub>3</sub>
	(1)	(2)	(3)	(4)
$Treat_i \times Post_i$	- 5.133 *** (1.568)	- 4.114 *** (1.402)	- 5.613 ** (2.701)	- 3.111 ** (1.256)
$Treat_i \times Post_i \times Middle$	- 4.531 ** (2.285)	- 4.527 ** (1.928)	- 4.143 (3.266)	0.681 (2.180)
$Treat_i \times Post_i \times West$	5.352 * (2.838)	6.592 ** (2.576)	5.128 (5.232)	2.648 (2.184)
城市控制变量	控制	控制	控制	控制
气象控制变量	控制	控制	控制	控制
城市固定效应	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制
观测值	2079	2079	2079	2079
R <sup>2</sup>	0.933	0.905	0.918	0.789

2. 资源禀赋异质性。资源禀赋是影响城市大气污染治理的重要因素。为了进一步分析大数据治理对不同资源禀赋城市大气污染防治的效应,根据《全国资源型城市可持续发展规划(2013—2020年)》中的分类方式,本文将样本城市划分为资源型城市和非资源型城市,并在公式(1)的基础上分别加入资源型城市(Resource)的虚拟变量与  $Treat_i \times Post_i$  的交互项,具体模型如下所示:

$$Y_{it} = \alpha + \beta_1 Treat_i \times Post_i + \beta_2 Treat_i \times Post_i \times Resource + \gamma X_{it} + \vartheta_i + \mu_t + \varepsilon_{it}$$

(4)

上式中,估计系数  $\beta_1$  表示生态环境大数据建设对非资源型城市大气污染的影响; $\beta_1 + \beta_2$  表示生态环境大数据建设对资源型城市大气污染的影响。回归结果如表5所示。可以发现,  $Treat_i \times Post_i$  的估计系数显著为负,而  $Treat_i \times Post_i \times Resource$  的估计系数为正,这说明生态环境大数据建设对非资源型城市的大气污染治理效应强于资源型城市。这可能是因为资源型城市产业结构相对单一,面临严重产业路径依赖和锁定效应,因而大数据治理发挥的作用有限;相较于资源型城市,非资源型城市的产业结构呈现出多样化,这有利于产业结构转型和政策效应的发挥。

表5 资源禀赋异质性

解释变量	AQI	PM <sub>2.5</sub>	PM <sub>10</sub>	O <sub>3</sub>
	(1)	(2)	(3)	(4)
$Treat_i \times Post_i$	- 6.587 *** (1.493)	- 5.108 *** (1.418)	- 6.223 *** (2.388)	- 3.780 *** (1.066)
$Treat_i \times Post_i \times Resource$	6.442 ** (2.496)	6.172 ** (2.405)	3.932 (4.545)	4.810 *** (1.831)
城市控制变量	控制	控制	控制	控制
气象控制变量	控制	控制	控制	控制
城市固定效应	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制
观测值	2079	2079	2079	2079
R <sup>2</sup>	0.933	0.904	0.917	0.790

(二) 机制分析

上文验证了基于大数据技术统筹作用的治理模式确实从总体上有效驱动了雾霾等污染物治理的进

展,展现了政府治理体系的一次成功数字化转型创新实践。其对于提升环境治理监管效率、提高决策科学性和有效性发挥重要作用<sup>[2]</sup>。但大数据治理实践中,鉴于统一的充分对称的信息环境正在建设推进中,导致治理行为者不能实现充分理性行为,可能会对有效发挥夯实地方政府主体责任作用产生影响。这里从是否提升政府环境决策与监管能力、是否提升公众参与度两个机制,进一步考察治理效应。回归结果如表6所示。

**1. 政府环境决策与监管能力。**大数据治理运用大数据、云计算等现代信息技术手段打造的“智慧环保”等平台,能快速搜集和处理涉及环境风险、举报、突发事件等信息,实现各级政府间实时数据传输和共享,从而上级政府能够清晰了解各地区的污染变化情况,从而解决环境监管中的委托代理问题<sup>[18]</sup>,以提高政府环境决策与监管能力。如何衡量政府环境决策与监管能力,从两个指标尝试。(1)环境行政处罚案件数是反映地方政府环境执法监管力度的重要指标<sup>[19]</sup>,进而侧面反映出地方政府环境监管能力,本文将其作为衡量政府环境监管能力的代理变量。由于环境行政处罚案件数仅有省级层面的数据,借鉴范子英和赵仁杰<sup>[20]</sup>的做法,以地级市工业产值占本省工业产值的比重作为权重,通过混频处理后将其转换为城市层面。回归结果见表6第(1)列,估计系数为1.066在1%的显著水平下显著为正,大数据治理使得环境污染的立案数量有所增加。反映出生态环境大数据建设能够使地方政府及时获得各方面数据,对于污染事件迅速反应从而提高处置决策能力,进而遏制污染问题从而提升空气质量。(2)城市大气污染物排放量直接反映了政府环境监管成效,同时也一直是区域环境保护的约束性考核指标。本文采用单位GDP二氧化硫排放量对数值衡量政府环境监管能力。表6中第(2)列汇报了回归结果。可以发现,估计系数在5%的显著水平下显著为负,说明生态环境大数据建设能够提升环境监管的精细化,当精准识别、实时追踪环境数据成为常态时,必然有利于二氧化硫排放量等污染物的减排,进而改善空气质量。

**2. 公众参与程度。**公众作为治理主体之一,是推进环境治理体系和能力现代化的重要组成部分<sup>[18]</sup>。通过生态环境大数据建设的推进,就是要拓展共享共建机制,调动公众积极主动参与,从而使公众参与环境治理的行为及其对政府的监督成为环境治理体系不可缺少的一部分<sup>[21]</sup>。若能落实公众参与的主体责任,对于大气污染治理将发挥积极作用。但对于如何衡量治理过程中公众这一主体参与治理的责任,尚未有恰当指标。本文对每个城市以“环境污染”为关键词的百度搜索指数进行逐个检索并整理,并考虑人口规模,将其与年末总人口之比作为侧面衡量公众参与程度的代理变量。表6第(3)列的结果显示,估计系数虽为正的0.034,但未能通过10%的显著性检验,表明样本期间提高公众参与程度的路径并未发挥明显效应。可能的原因是,目前生态环境大数据建设更多偏向对于空气质量情况的大数据技术监测以及相关数据的搜集,其主动公开的环境信息相对较少,数据资源共享开放水平较低,公众所提出的相关诉求也可能未得到重视和落实<sup>[22]</sup>,从而限制了公众参与在大气污染治理中的积极作用。

综上机制分析结果,在一定程度上反映生态环境大数据试点的建设通过提高政府监管能力确实提升了大气污染治理效果,但通过提高公众参与程度的机制还未显示出明显作用。

表6 机制分析

变量	政府环境决策与监管能力		公众参与程度
	(1)	(2)	(3)
$Treat_i \times Post_i$	1.066 *** (0.107)	-0.099 ** (0.043)	0.034 (0.044)
城市控制变量	控制	控制	控制
城市固定效应	控制	控制	控制
时间固定效应	控制	控制	控制
观测值	1907	2080	2206
$R^2$	0.921	0.970	0.897

## 五、政策空间与启示

在习近平生态文明思想指引下,我国大气环境治理体系与治理能力现代化建设不断推进。中央将空气质量明显改善定位为刚性要求,中央财政大气污染防治专项资金年年增加,累计下达2248亿元,全力以赴坚决打赢蓝天保卫战。本文从生态环境大数据建设试点政策的视角,检验实践中政府干预下推动大气污染的大数据治理模式在这一转型进程中所发挥的总体效应,研究结果对进一步推动我国生态环境大数据治理具有参考价值,为推进数字生态文明建设提供政策依据。

但我国生态环境稳中向好的基础还不稳固,与人民群众对优美生态环境的需要相比仍存差距。城市环境空气质量总体仍未摆脱“气象影响型”,尚有29.8%的城市 $PM_{2.5}$ 平均浓度达不到国家二级标准, $O_3$ 浓度呈波动上升趋势,区域性重污染天气过程时有发生。加之新污染物的出现,特别是日益呈现 $PM_{2.5}$ 和 $O_3$ 双高的复合型污染特征,如何应对这些问题没有先例可循。而“反弹”成为2022年大气污染的高频词,各地区在不同污染物的治理成效上也出现分化,中国大气污染治理这场长跑,尚处于焦灼而关键时刻<sup>[23]</sup>。新一轮污染防治攻坚战从“坚决打好”转变为“深入打好”,意味着触及的矛盾问题层次更深、领域更广,大气污染防治进入深水区。习近平总书记在2023年7月第九次全国环境保护大会上强调,要深化人工智能等数字技术应用,构建美丽中国数字化治理体系,建设绿色智慧的数字生态文明。基于新一代数字技术应用的大数据治理模式,如何实现更加精准治污、科学治污提出更高要求,还有较大持续改进创新空间。

第一,夯实大数据治理模式的技术基础,拓展大气环境智慧监测感知能力建设。大数据治理发挥了相应作用,其得益于大气环境智慧监测体系率先突破,实现了空气质量监测感知的高效化。在试点效应推广中,各地区可进一步加大积极引进和部署先进的卫星遥感、无人机技术及热成像等相关监测技术,基于“天空地”一体化监测网络,综合卫星遥感、地面监测、高精度气象等多源实时监测数据,完善城市的“部、市、县、乡”的多级热点网格联动监管体系,强化大数据管理与分析应用,充分满足各行为主体对数据全面及时和准确的收集和传输。面对区域性的复合型污染新难题,推动微站联网建设,拓展 $PM_{2.5}$ 与 $O_3$ 协同控制监测感知能力建设。开发利用人工智能的自动识别技术和智能化模型等对实时污染源解析与预警,实现工业、扬尘、机动车以及“散乱污”等污染源的精细化监测和督查。以大数据监测感知技术作为基点支撑精准和科学治污,提升政府决策和监管能力。

第二,破除数据烟囱和信息壁垒,压实地方政府主体责任。大气污染成因复杂,利益主体诉求不同,治理效应存在明显地区差异。长期以来环境治理过程中存在严重的信息不对称,环境监测虽积累海量数据,但由于各地监测机构各自为政,监测数据无统一规范,准确性和可比性差,数据难以共享,形成了应用“烟囱”和数据“孤岛”林立。大数据治理模式通过建立全国统一的实时在线环境监控系统,实现空气质量、重大污染源、大气状况监测监控全覆盖,通过进一步优化国家、省、市三级监测数据联通共享机制,规范不同时空分辨率的卫星遥感数据,加大投入整合全球多源气象、大气污染、辐射监测数据,融合跨行业数据交叉分析与验证,为合理量化评估污染责任、畅通压力传导机制、动态评估规划提供数据信息,加快政府污染治理能力转型。

第三,创新公众参与的机制设计,激发生态文明建设的内生动力。良好生态环境是最普惠的民生福祉,需激发全社会共享共建的内生动力。但现实中公众参与途径碎片化、监督责任模糊,信息共享开放程度低,限制了参与治理的作用,需创新公众参与的机制设计<sup>[24]</sup>。一方面,不断扩充部长信箱、12369环保热线、微博微信等政民互动的渠道,运用大数据创新政府服务理念和服务方式,完善信息公开督促和审查机制,智能开发实时推送各类便民环境信息产品,高质量满足服务公众的信息需求。另一方面,需搭建有效的反馈响应机制,通过广泛运用大数据、云计算等现代信息技术手段,自动抓取、主题检索、专题聚焦涉及环境风险、环保举报、突发环境事件、社会舆论等互联网海量数据,利用机器学习等算法实现有效筛选和甄别,形成有价值的反馈信息,及时回应公众意见。当然还需大力提升公众的数字素养和数字技能以适应时代需求,并借助智能技术手段推广和宣传“公民生态环境行为规范新十条”,引导公众践行生态文明理念,把建设美丽中国转化为全体人民的自觉行动。

## 参考文献:

- [1] ZHOU X, ZHOU D, WANG Q, et al. How information and communication technology drives carbon emissions: a sector-level analysis for China[J]. *Energy Economics*, 2019, 81(1): 380-392.
- [2] YANG X, WU H, REN S, et al. Does the development of the internet contribute to air pollution control in China? Mechanism discussion and empirical test[J]. *Structural Change and Economic Dynamics*, 2021, 56(1): 207-224.
- [3] 吕民乐, 陈颖瑶. 信息化有利于降低雾霾污染吗? ——基于空间计量模型的实证检验[J]. *南京财经大学学报*, 2021(3): 13-24.
- [4] 赵芬, 张丽云, 赵苗苗, 等. 生态环境大数据平台架构和技术初探[J]. *生态学杂志*, 2017(3): 824-832.
- [5] 刘丽香, 张丽云, 赵芬, 等. 生态环境大数据面临的机遇与挑战[J]. *生态学报*, 2017(14): 4896-4904.
- [6] 陈健. 数字化技术赋能环境治理现代化的路径优化[J]. *哈尔滨工业大学学报(社会科学版)*, 2023(3): 80-90.
- [7] 张毅, 贺桂珍, 吕永龙, 等. 我国生态环境大数据建设方案实施及其公开效果评估[J]. *生态学报*, 2019(4): 1290-1299.
- [8] 石大千, 丁海, 卫平, 等. 智慧城市建设能否降低环境污染[J]. *中国工业经济*, 2018(6): 117-135.
- [9] GAO K, YUAN Y. Is the sky of smart city bluer? Evidence from satellite monitoring data[J]. *Journal of Environmental Management*, 2022, 317(17): 115483.
- [10] 郭炳南, 王宇, 张浩. 数字经济发展改善了城市空气质量吗——基于国家级大数据综合试验区的准自然实验[J]. *广东财经大学学报*, 2022(1): 58-74.
- [11] HUANG Y, GUO B, SUN H, et al. Relative importance of meteorological variables on air quality and role of boundary layer height[J]. *Atmospheric Environment*, 2021, 267(23): 118737.
- [12] 赵仁杰, 钟世虎, 张家凯. 非意图的后果: 政府扁平化改革与空气污染治理[J]. *世界经济*, 2022(2): 162-187.
- [13] 孙传旺, 罗源, 姚昕. 交通基础设施与城市空气污染——来自中国的经验证据[J]. *经济研究*, 2019(8): 136-151.
- [14] 宋弘, 孙雅洁, 陈登科. 政府空气污染治理效应评估——来自中国“低碳城市”建设的经验研究[J]. *管理世界*, 2019(6): 95-108, 195.
- [15] CHETTY R, LOONEY A, KROFT K. Salience and taxation: theory and evidence[J]. *American Economic Review*, 2009, 99(4): 1145-1177.
- [16] 刘瑞明, 赵仁杰. 国家高新区推动了地区经济发展吗? ——基于双重差分方法的验证[J]. *管理世界*, 2015(8): 30-38.
- [17] 王锋, 葛星. 低碳转型冲击就业吗? ——来自低碳城市试点的经验证据[J]. *中国工业经济*, 2022(5): 101-119.
- [18] 潘旭文, 付文林. 环境信息公开与地区空气质量——基于 PM<sub>2.5</sub> 监测的准自然实验分析[J]. *财经研究*, 2022(5): 110-124.
- [19] 杨晓妹, 庞倩男, 王有兴. 环境税对绿色创新的空间溢出效应——基于政府环境监管调节作用的分析[J]. *中国人口·资源与环境*, 2023(1): 50-60.
- [20] 范子英, 赵仁杰. 法治强化能够促进污染治理吗? ——来自环保法庭设立的证据[J]. *经济研究*, 2019(3): 21-37.
- [21] LIU M, ZHANG B, BI J. Appreciating the role of big data in the modernization of environmental governance[J]. *Frontiers of Engineering Management*, 2022, 9(1): 163-169.
- [22] 张国兴, 雷慧敏, 马嘉慧, 等. 公众参与对污染物排放的影响效应[J]. *中国人口·资源与环境*, 2021(6): 29-38.
- [23] 陈松蹊. 空气质量评估报告(十)“3+110”城市 2013—2022 年区域污染状况评估[R]. 北京: 北京大学统计科学中心, 2023.
- [24] 刘华军, 邵明吉, 郭立祥. 新时代的中国大气污染治理之路——历程回顾、成效评估与路径展望[J]. *商业经济与管理*, 2023(2): 66-79.



(责任编辑 郭宝才 王 权)