

Article

The Impact Mechanisms and Paths of the National Big Data Comprehensive Pilot Zones on Urban Pollution Reduction and Carbon Mitigation

Shufen Yang ¹, Luoshi Wu ¹, Weiyong Zou ^{1,*}, Yumeng Li ², Chenyong Shi ³ and Shuoxiang Wang ⁴

¹ School of economics, Guangdong Ocean University, Zhanjiang 524091, China

² School of Economics & Management and Law, University of South China, Hengyang 421009, China

³ School of Business, The University of Leicester, Leicester LE1 7RH, UK; cs920@leicester.ac.uk

⁴ School of Management, The University College of London, London WC1E 6BT, UK

* Correspondence: weiyongzou2023@163.com

Abstract

The National Big Data Comprehensive Pilot Zone (NBDCPZ) constitutes a landmark institutional initiative designed to advance and deploy data elements while strengthening governance capacity for environmental sustainability and low-carbon transition. Drawing on a panel dataset covering 285 Chinese cities at or above the prefectural level over the 2005–2023 period, this study capitalizes on the sequential rollout of the NBDCPZ to establish a quasi-natural experimental design. It employs a staggered difference-in-differences (DID) design to examine the policy's causal effects on urban environmental quality—specifically, its effectiveness in curbing air pollution and carbon emissions—and further identifies the underlying mechanisms driving these effects and sources of contextual heterogeneity. The results demonstrate that the policy yields substantial benefits in curbing urban pollution and advancing carbon reduction; this conclusion remains robust across multiple sensitivity checks. Mechanism analysis reveals that the policy achieves synergistic gains in pollution reduction and carbon mitigation by promoting green innovation and reducing energy consumption, as well as by fostering digital industry agglomeration and accelerating digital innovation. Heterogeneity analysis shows significant structural variation in policy effects: non-resource-based cities exhibit more pronounced carbon mitigation effects, whereas resource-based cities, large cities, and old industrial bases demonstrate stronger pollution control effects.

Keywords: NBDCPZ; pollution reduction and carbon mitigation; quasi-natural experiment



Academic Editor: Bruce Morley

Received: 11 February 2026

Revised: 15 June 2026

Accepted: 30 July 2026

Published: 5 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

As China advances into the era of high-quality development, its economic growth model has fundamentally shifted from a speed-driven paradigm to one centered on quality enhancement. However, due to rapid industrialization and urbanization in earlier stages—coupled with continuous population and economic expansion—environmental pollution and carbon emissions have emerged as critical challenges. These challenges not only hinder sustainable economic growth but also constrain the development of a modernized economic system. Based on the “2024 Report on the State of China’s Ecological Environment (2025)” (https://www.gov.cn/lianbo/bumen/202506/content_7026674.htm accessed on 1 July 2026), the annual average PM_{2.5} concentration in prefecture-level and

above cities reached $29.3 \mu\text{g}/\text{m}^3$ in 2024—approximately six times the World Health Organization’s guideline value. Relevant studies indicate that environmental pollution and carbon emissions are closely associated with cardiovascular diseases, respiratory diseases, neurological disorders, lung cancer, diabetes, and other health conditions [1]. As the world’s largest energy consumer and the largest emitter of carbon dioxide (CO_2), China faces environmental pollution and carbon emissions as major environmental challenges that threaten both economically sustainable growth and public health.

Against this backdrop, the establishment of the NBDCPZ policy offers a new pathway to reconcile the tensions among pollution reduction, carbon mitigation, and economic growth. The “Overall Plan for the Construction of Digital China (2023)” (https://www.gov.cn/xinwen/2023%E2%88%9202/27/content_5743484.htm accessed on 1 July 2026) explicitly calls for “accelerating the synergistic transition of digitalization and greening.” Cutting-edge digital tools—including big data analytics and AI—enable real-time environmental monitoring and intelligent data interpretation. This empowers policymakers and businesses with actionable insights and evidence-based guidance for targeted pollution mitigation and efficient emissions reduction, thereby harmonizing ecological sustainability with economic development. Hence, as a testing ground for the deep integration of digital technology with China’s core endeavor of pollution control and carbon reduction, the NBDCPZ policy holds significant importance for improving the environmental governance system and accelerating a green socio-economic transition.

2. Literature Review

Existing studies have developed three research streams centered on big data applications in pollution reduction and carbon mitigation governance. First, research on the economic and environmental effects of big data technology. Early scholarly attention to big data technology centered predominantly on its implications for corporate growth and evolution [2–4]. As research deepened, scholars discovered that big data contributes to sustainable urban construction [5], optimizes urban energy structures, and facilitates energy conservation and emission reduction—key priorities for China’s long-term sustainability and ecological preservation [6,7]. Nevertheless, much of this scholarship remains confined to descriptive analyses of technological applications and fails to rigorously identify the causal effects of institutional policies. Second, research on the synergistic governance of pollution reduction and carbon mitigation. As a central theme aligned with China’s “dual carbon” goals, pollution abatement and greenhouse gas emission reduction have drawn significant scholarly attention—particularly regarding the co-benefits and policy complementarities between these two environmental strategies [8–10]. Several researchers have also extensively examined key factors influencing pollution and carbon emissions through traditional pathways—such as energy mix composition and technological advancement [11–13], and regional spatial heterogeneity [14]—thereby elucidating the institutional, structural, and technological drivers of the green transition [15]. Nevertheless, the transformative potential of digital technology—as a novel factor of production—has been largely overlooked. The two major research streams have long remained relatively independent, and research on cross-domain integration has only recently gained momentum. Few empirical studies have thus far investigated how big data influences urban carbon emissions [16], and no systematic research has been conducted on the synergistic relationship between big data, pollution reduction, and carbon mitigation. Third, research on the policy effects of the NBDCPZs. The NBDCPZ, designed as a big-data-driven, region-specific policy, seeks to foster the convergence of data elements, thereby positioning itself as a pioneering area for the market-based allocation of data resources [17]. At the macro level, the implementation of the Pilot Zone effectively achieves emission reduction by promoting

industrial advancement [18], elevating GTFP [19], and enhancing the efficiency of energy use [20,21]. At the micro level, it offers extensive data resources and robust policy backing, fostering an enabling environment for corporate green innovation. This not only exerts beneficial ripple effects across the entire industrial value chain and regional environmental management systems but also actively advances urban efforts to curb pollution [22–24]. However, a systematic evaluation of the co-benefits of pollution reduction and carbon mitigation has yet to be established, and the analysis of mechanisms and heterogeneity remains relatively superficial.

Overall, the three research streams remain siloed and lack critical dialogue—especially on how digital governance affects environmental performance. While some studies find that big data and AI directly improve environmental outcomes and cut emissions by boosting resource efficiency, tightening monitoring, and spurring green innovation [7,11], others stress that digital tools yield no automatic environmental gains; their impact hinges on institutional quality, market development, and government capacity [5,18,25]. Consequently, scholars disagree: is digital governance a standalone technical instrument—or a system that only works with aligned institutions and policies?

Current research still faces several limitations. First, causal identification remains weak: most studies rely solely on standard DID, ignoring biases from staggered policy rollout, time-varying treatment effects, and effect heterogeneity. Pilot city selection is also non-random—shaped by local economic strength, digital readiness, and governance capacity—raising endogeneity concerns. Second, existing studies on the NBDPZ's environmental benefits focus narrowly on single indicators—like air pollutant concentrations—while overlooking its broader pollution-carbon co-benefits and underlying mechanisms. Furthermore, existing heterogeneity analyses focus on macro-level disparities—e.g., between China's eastern and western regions—while neglecting granular dimensions like city size, resource endowments, and industrial composition, limiting the theoretical basis for targeted, evidence-based policy design.

This research introduces multiple original contributions across the following dimensions: first, we strengthen causal identification by integrating the standard multi-period DID model with multiple robustness checks—PSM-DID, double machine learning, Goodman-Bacon decomposition, CSDID, and Imputation Estimation. This directly tackles biases from staggered policy rollout, treatment effect heterogeneity, and non-random pilot selection, yielding more reliable estimates and sharper policy insights. Second, we identify two key mechanisms—green technological innovation and digital industrial clusters—through a chain mediation framework, revealing how policy drives pollution reduction and carbon mitigation. Third, we uncover heterogeneity in policy effects on carbon emissions and environmental pollution across three dimensions—resource endowments, city size, and industrial foundation—delivering targeted empirical support for differentiated policymaking.

3. Theoretical Analysis and Research Hypotheses

3.1. Urban Pollution Reduction Effect of NBDPZ Policy

The NBDPZ policy advances digital-driven governance and economic development by integrating data resources, fostering technological innovation and application, and building a digital industrial ecosystem. The primary aim is to foster the market-driven distribution of data resources, improve governmental administrative effectiveness, support industrial structural optimization, and drive an all-encompassing green transition in economic and social development.

Regarding digital technology innovation, the new growth theory posits that knowledge accumulation and technological advancement are the key determinants of economic

growth. Centered on the development of a data-driven digital economy, the Pilot Zone effectively accelerates the flow of information and dissemination [26]. By optimizing resource allocation and promoting decision-making innovation, it enhances the productivity of green innovation output [27,28]. Furthermore, digital technologies and intelligent platforms are now embedded across key domains—including environmental monitoring, industrial manufacturing, and public administration—driving digital transformation, advancing smart manufacturing, and enabling ecological governance innovation [29]. Regarding resource allocation, data possesses replicable and shareable characteristics. By enabling precision production and intelligent management, it drives traditional production factors to be allocated more efficiently in the production process. Concurrently, it allows for creating more material wealth and corresponding services with fewer physical resources, thereby effectively substituting for traditional material resources. This substitution reduces the intensity of resource use and pollutant generation per unit of output at the source and helps overcome the constraint of diminishing marginal returns of traditional production factors. Moreover, policy synergy can strengthen systematic emission reduction. The big data pilot policy fosters synergistic policy impacts through cross-departmental data sharing and institutional innovations. In summary, the NBDCPZ improves pollutant and carbon monitoring accuracy and efficiency through digital innovation; strengthens front-end control of emissions via optimized resource allocation; and enhances systemic emission reduction through policy synergy—delivering a multidimensional impact on urban environmental governance. Accordingly, this study formulates the following hypothesis:

H1. *The NBDCPZ policy can promote urban pollution reduction and carbon mitigation.*

3.2. Mechanism Analysis of NBDCPZ Policy on Urban Pollution and Carbon Reduction

3.2.1. The Emission Reduction Effect of Green Innovation

Green innovation serves as the core technical underpinning for tackling the intertwined challenges of energy use and environmental degradation, while simultaneously enabling co-benefits between pollution control and greenhouse gas mitigation [30,31]. Inadequate green innovation capacity leads to sluggish industrial upgrading in cities, hindering the green transformation of severely polluted industries. This exacerbates carbon emissions and conventional pollutant discharges, emerging as a critical bottleneck constraining urban green development [32]. From the lens of technological change, green innovation drives technological advancements in energy production and consumption and facilitates the widespread application of clean energy sources, thereby reducing total carbon emissions at the source [8,33]. From the perspective of environmental governance, green innovation not only improves resource utilization and curtails resource consumption but also boosts product recovery rates and mitigates the spillover of environmental pollutants [34].

Nevertheless, traditional urban green innovation is plagued by challenges such as high information barriers and prohibitively high research and development expenditures. The low efficiency of green technology commercialization further impedes its ability to simultaneously address energy demand and environmental emissions. The NBDCPZ policy can systematically stimulate urban green innovation by synergistically leveraging data-driven empowerment and policy incentives [35,36]. On one hand, leveraging big data, the policy breaks down green R&D information silos, integrates multi-source data—global green patents, market demand, and policy signals—and lowers information costs and decision risks for enterprise green R&D, accelerating sustainable technology advancement [3,37]. On the other hand, the Pilot Zone guides the agglomeration of innovation factors into the green technology field through supporting measures such as fiscal subsidies, talent attraction and cultivation, and platform construction, thereby promoting the large-scale

application of green innovation outcomes. The improvement of green innovation levels can directly drive innovations in production processes and optimization of energy structures, significantly reducing urban energy consumption intensity. The reduction in energy consumption intensity not only cuts carbon dioxide emissions from fossil fuel combustion but also simultaneously reduces various pollutant emissions associated with energy consumption, ultimately achieving synergistic improvements in pollution abatement and carbon mitigation. Accordingly, this study formulates the following hypothesis:

H2. *The NBDCPZ policy promotes urban pollution reduction and facilitates low-carbon transition through green innovation.*

3.2.2. The Innovation Effect of Digital Industry Agglomeration

Digital industry clustering provides critical technical support for integrating urban pollution abatement with carbon mitigation, while enabling the value conversion of data factors through accelerated iterative innovation of digital technologies [38]. Insufficient digital industry agglomeration leads to weak urban digital innovation momentum, delayed digital transformation of traditional industries, and low intelligence levels of environmental governance, which hinders the achievement of long-term breakthroughs in pollution abatement and carbon mitigation. From the dimension of industrial transformation, digital industry agglomeration facilitates real-time data collection from the production end, which can be converted into manufacturing knowledge to optimize green processes and enhance production efficiency [39]. From the dimension of environmental governance, digital innovation enables a smart eco-protection system—comprehensive monitoring, precise traceability, and dynamic control—that overcomes regulatory lag and inefficiency, boosting the accuracy and coherence of pollution reduction and carbon mitigation.

The NBDCPZ policy accelerates spatial agglomeration of digital enterprises, high-end talent, and technological capital—through upgraded digital infrastructure, targeted industrial policies, and an improved digital ecosystem—boosting urban digital industry concentration [40,41]. High-level digital industry agglomeration fosters deeper convergence among digital technologies, transforms the operational logic, and drives collaborative upgrades in production, management, and marketing—thereby enhancing digital innovation capacity [42,43]. On one hand, the strengthening of digital innovation capabilities enables the digital transformation of traditional high-energy-consuming industries, optimizing production processes and energy dispatching to reduce energy consumption and emissions at the production end; on the other hand, it facilitates the digital transformation of environmental governance, enabling real-time pollution source monitoring and cross-departmental collaborative supervision, thereby improving emission reduction efficiency at the governance level. Ultimately, these two pathways jointly drive reductions in urban carbon emissions and environmental pollution levels. Accordingly, this study formulates the following hypothesis:

H3. *The NBDCPZ policy promotes urban pollution reduction and facilitates low-carbon transition through the innovation effect of digital industry agglomeration.*

4. Data and Methodology

4.1. Model Specification

From 2015 to 2016, the National Development and Reform Commission and other relevant ministries and commissions successively approved the construction of NBDCPZ. The coverage includes provincial capitals (or provincial administrative centers) such as Guizhou, Guangdong, Hebei, Henan, and Inner Mongolia, as well as Beijing, Shanghai,

and Shenyang, covering a total of 80 cities. This strategic move is dedicated to promoting a balanced and coordinated configuration of the big data sector across China's four major regions and facilitating the accelerated socioeconomic transition toward environmentally sustainable development. This study regards the implementation of the NBDCPZ as a quasi-natural experiment, constructs a multi-period DID model, and explores its impact on mitigating pollutant and carbon emissions at the city level. The empirical model is specified as below:

$$y_{it} = \alpha_1 + \beta_1 treat_i \times post_t + \sum_j \gamma_j Control_{it} + \lambda_i + \delta_t + \varepsilon_{it} \quad (1)$$

The dependent variable y is operationalized using two complementary indicators—per capita carbon emissions ($lnpc_{CO_2}$) and environmental pollution index ($pollution$)—to jointly capture the dual dimensions of pollution abatement and carbon mitigation in city i during year t . The interaction term $treat_i \times post_t$ —a multi-period DID policy dummy—is the primary explanatory variable. Specifically, $treat_i$ is a city-level treatment indicator equal to 1 if city i has been designated as an NBDCPZ, and 0 otherwise; $post_t$ is a time-varying policy exposure indicator equal to 1 for the designation year and all subsequent years, and 0 otherwise. $Control_{it}$ denotes a vector consisting of control variables. The terms λ_i and δ_t are designated as the fixed effects for city and year, respectively. ε_{it} denotes the stochastic error term. The coefficient, β_1 , constructed by interacting the policy dummy with the time dummy, captures the net policy effect.

4.2. Variable Definitions

4.2.1. Dependent Variable

$lnpc_{CO_2}$ and $pollution$ serve as the dependent variables. To avoid information masking and reveal the true performance of environmental governance, this study measures urban pollution reduction and carbon emission reduction indicators from two dimensions: carbon emissions and environmental pollution. Among them, the carbon emission index, adapted from [44], is derived from a 1 km $\times$ 1 km resolution raster dataset provided by the CGER (Center for Global Environmental Research) website, which includes carbon emissions from various activities, including the combustion of fossil fuels. To generate $lnpc_{CO_2}$, this study spatially aggregates the grid-level emissions to the city level and normalizes them by the corresponding year-end population. The $pollution$ is constructed by measuring three sub-indicators—total wastewater discharge, SO_2 emissions, and smoke and dust emissions—using an improved entropy weight method to derive a composite environmental pollution index.

4.2.2. Core Explanatory Variable

$Treat_i \times post_t$ —the NBDCPZ policy—is the core explanatory variable. Specifically, $treat_i$ is a treatment-group dummy variable that equals 1 if city i is designated as a pilot city and 0 otherwise. $post_t$ is a time dummy variable that equals 1 for years t in which the policy has been formally implemented and 0 otherwise.

4.2.3. Control Variables

Urbanization Rate ($city$) is obtained by dividing the number of permanent urban residents by the total permanent resident population. For cities lacking statistics on permanent urban residents, the ratio of non-agricultural permanent residents to the total permanent resident population is used. Foreign direct investment (fdi) is measured by the proportion of actual foreign direct investment to GDP. Infrastructure level ($instra$) is defined as the per capita length of the highway network. Financial development level (fin) is quantified as the combined value of financial institutions' total deposits and loans relative to GDP.

Market integration level (*mar*) is quantified using the method of [45], which models market integration as a function of economic scale (GDP) and spatial distance.

4.3. Data Sources

This study excludes cities with substantial missing data, including Zhongwei and Nyingchi, and constructs a panel dataset comprising 285 Chinese cities ranked at the prefecture level or higher over the 2005–2023 period. This paper integrates data from the China City Statistical Yearbook, the CNRDS database, local government work reports, and satellite remote sensing observations from the U.S. NOAA. Individual missing data points are imputed using interpolation and the simple moving average method. Table 1 presents the basic statistical properties of the main variables.

Table 1. Descriptive statistics.

Variable	Symbol	Obs	Mean	Std. Dev.	Min	Max
Per capita carbon emissions	<i>lnpc_{o2}</i>	5415	9.6188	1.0648	0.0000	12.3658
Environmental pollution index	<i>pollution</i>	5415	0.0367	0.0414	0.0001	0.5264
NBDPZ Policy	<i>Treat × Post</i>	5415	0.1189	0.3237	0.0000	1.0000
Urbanization Rate	<i>city</i>	5415	0.5279	0.1714	0.1499	0.9522
Foreign Direct Investment	<i>fdi</i>	5415	1.7197	1.7999	0.0125	8.5935
Human Capital	<i>edu</i>	5415	1.8437	2.3769	0.0047	14.6927
Infrastructure level	<i>infra</i>	5415	32.7587	20.1357	0.4053	180.9980
Financial development level	<i>fin</i>	5415	2.4029	1.1727	0.8959	6.7602
Market integration level	<i>mar</i>	5415	44.1196	24.6901	1.1211	195.9180

5. Analysis of Empirical Results

5.1. Benchmark Regression Results

The estimated coefficients and associated statistical results are presented in Table 2. Column (1) controls only for year- and city-level fixed effects. Empirical results show that *lnpc_{o2}* is significantly negative at the 1% level. Column (2) incorporates all control variables; its coefficient remains -0.0455 —slightly smaller but still significant at the 1% level. Given the sample of 285 prefecture-level and above cities, this modest effect translates into substantial national reductions in total emissions and meaningful environmental improvement. Columns (3) and (4) report results obtained by substituting the dependent variable with *pollution*, yielding consistent results: *pollution* is consistently negative and statistically significant, regardless of whether control variables are included. In comparison with existing literature, the magnitude of the estimated policy effect in the present research largely aligns with those of Pan et al. (2024) [18] and Wang et al. (2025) [46]. These findings suggest that the implementation of the NBDPZ significantly contributes to urban pollution control and carbon abatement, thereby validating Hypothesis H1.

Table 2. Benchmark regression results.

	(1)	(2)	(3)	(4)
Variables	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Pollution</i>	<i>Pollution</i>
<i>Treat × Post</i>	-0.0523^{***} (0.006)	-0.0455^{***} (0.005)	-0.0150^{***} (0.001)	-0.0148^{***} (0.002)
<i>city</i>		0.1236^{***} (0.033)		-0.0103^{*} (0.005)
<i>fdi</i>		0.0039^{***} (0.001)		-0.0001 (0.000)

Table 2. Cont.

	(1)	(2)	(3)	(4)
Variables	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Pollution</i>	<i>Pollution</i>
<i>edu</i>		0.0111 *** (0.003)		−0.0023 *** (0.000)
<i>instra</i>		0.0030 *** (0.000)		0.0002 *** (0.000)
<i>fin</i>		0.0207 *** (0.003)		0.0014 ** (0.001)
<i>mar</i>		−0.0104 *** (0.001)		−0.0009 *** (0.000)
Constant	9.6590 *** (0.001)	9.8432 *** (0.060)	0.0385 *** (0.000)	0.0782 *** (0.008)
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	5415	5415	5415	5415
R-squared	0.993	0.994	0.754	0.758

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors.

5.2. Pre—Trend Test

The pre-trend assumption constitutes the core identification condition of the DID model, ensuring that estimates of policy impacts are not biased by pre-existing systematic disparities. Figures 1a and 1b present pre-trend test results for the dependent variables of *lnpc_{o2}* and *pollution*, respectively. Results reveal that, before the NBDCPZ policy took effect, the estimated coefficients for both dependent variables fluctuated around zero; the 95% confidence intervals of all coefficients included zero and did not reach the conventional statistical significance level, that is, there was no statistical difference from zero. This suggests that, prior to policy implementation, the trajectories of pollution control and carbon reduction evolved similarly in pilot and non-pilot cities, providing empirical support for the pre-trend assumption. From implementation through year 7, the NBDCPZ policy exerted a significantly negative effect on *lnpc_{o2}* and *pollution*, and its effect strengthened progressively over time. This reflects two realities: digital infrastructure, data platforms, and technology diffusion require time to mature; and enterprise digital transformation and green innovation follow multi-stage cycles of R&D, adoption, and scaling. Consequently, the policy delivers persistent, cumulative pollution abatement and carbon mitigation, driving sustained green and low-carbon transformation in pilot cities.

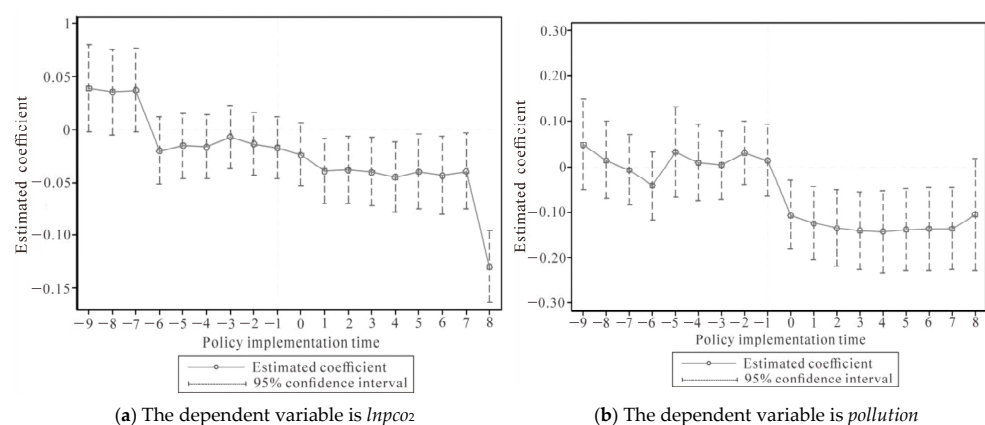


Figure 1. Pre—trend test.

5.3. Robustness Checks

5.3.1. Placebo Test

This study validates the reliability of its findings through placebo tests that randomly assign pilot-city status. Specifically, 80 cities are randomly selected from the sample of 285 cities to form the placebo treatment group, with the remaining cities assigned to the control group. This sampling process is repeated 500 times, and regressions are performed for the two dependent variables—*lnpc_{o2}* and *pollution*. Figure 2a,b present the placebo test results for these two dependent variables, respectively. The results indicate a significant discrepancy between the actual policy effect and the effects of random policy interventions. The kernel density curves corresponding to both dependent variables are symmetrically distributed around zero, approximating a normal distribution, and the vast majority of the estimated coefficients fail to achieve statistical significance. The findings confirm that the NBDPZ drives pollution reduction and carbon mitigation through targeted policies. Digital governance thus delivers sustained improvements in urban environmental performance, enabling local governments to advance digital-enabled green transformation.

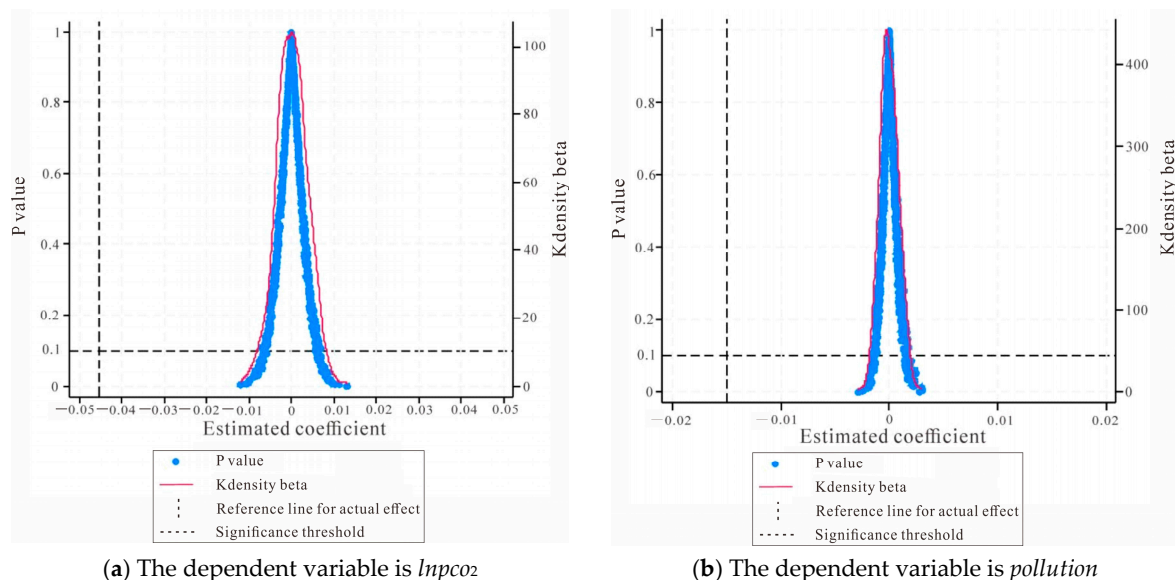


Figure 2. Placebo test.

5.3.2. Double Machine Learning (DML)

To mitigate endogeneity concerns stemming from the omission of key variables or measurement biases in the model, this study incorporates the DML methodology to perform robust estimation. First, the study predicts *lnpc_{o2}* and *pollution* using the control variable set, respectively, and derives the corresponding prediction residuals. Second, it predicts the NBDPZ policy variable using the same control variable set to obtain its residual. Finally, it estimates the causal effect of the policy by regressing one set of residuals on the other, thereby mitigating estimation bias induced by model misspecification.

Table 3, columns (1) to (6), reports estimates from three machine learning models: ordinary least squares (OLS), Lasso regression with cross-validation (LassoCV), and Ridge regression with cross-validation (RidgeCV). The latter two models leverage regularization techniques to mitigate overfitting and the “curse of dimensionality.” Columns (1)–(3) present results for the dependent variable per capita carbon emissions, while columns (4)–(6) correspond to the environmental pollution index.

Table 3. Results of double machine learning estimation.

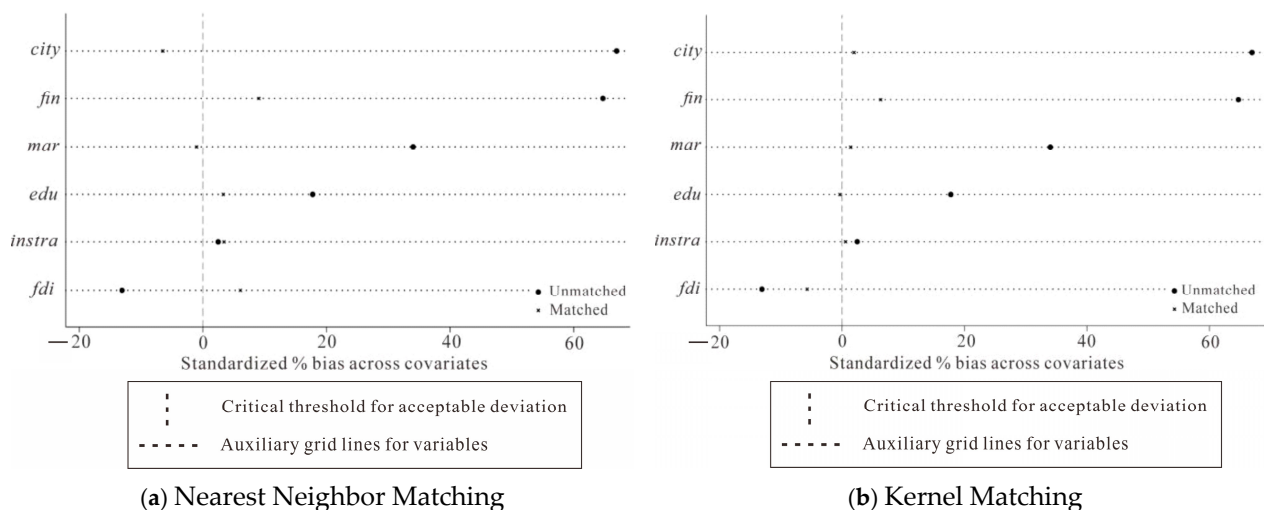
Variable	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Ols</i>	<i>Lassocv</i>	<i>Ridgecv</i>	<i>Ols</i>	<i>Lassocv</i>	<i>Ridgecv</i>
	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Pollution</i>	<i>Pollution</i>	<i>Pollution</i>
<i>Treat</i> × <i>Post</i>	−0.0431 *** (0.005)	−0.0442 *** (0.000)	−0.0431 *** (0.060)	−0.0165 *** (0.002)	−0.0167 *** (0.002)	−0.0165 *** (0.002)
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5415	5415	5415	5415	5415	5415

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors.

Across all models, the NBDCPZ policy significantly reduces *lnpc_{o2}* and *pollution* ($p < 0.01$), confirming robust policy effectiveness. Digital governance thus consistently improves urban environmental performance—delivering both broad applicability and lasting impact.

5.3.3. Propensity Score Matching-Difference-in-Differences (PSM-DID)

To address selection bias arising from both observed and unobserved confounders, this study implements a PSM-DID strategy. Three matching strategies—nearest-neighbor matching and kernel matching—are applied to the sample. Figures 3 and 4 present the post-matching balance test results for the two dependent variables: *lnpc_{o2}* and *pollution*, respectively.

**Figure 3.** Balance test for the dependent variable *lnpc_{o2}*.

The standardized mean differences of all control variables are substantially smaller post-matching than pre-matching, indicating favorable matching performance and effective balancing of sample characteristics between the two groups. The study then conducts DID regressions using the matched sample, with results reported in Table 4. Across all matching methods, the NBDCPZ policy significantly lowers environmental pollution ($p < 0.01$), confirming robust causal effects. Pilot cities' environmental gains stem not from pre-existing advantages—such as development level, resources, or governance capacity—but from

the policy itself. Digital infrastructure and data allocation reforms thus directly yield measurable environmental governance outcomes.

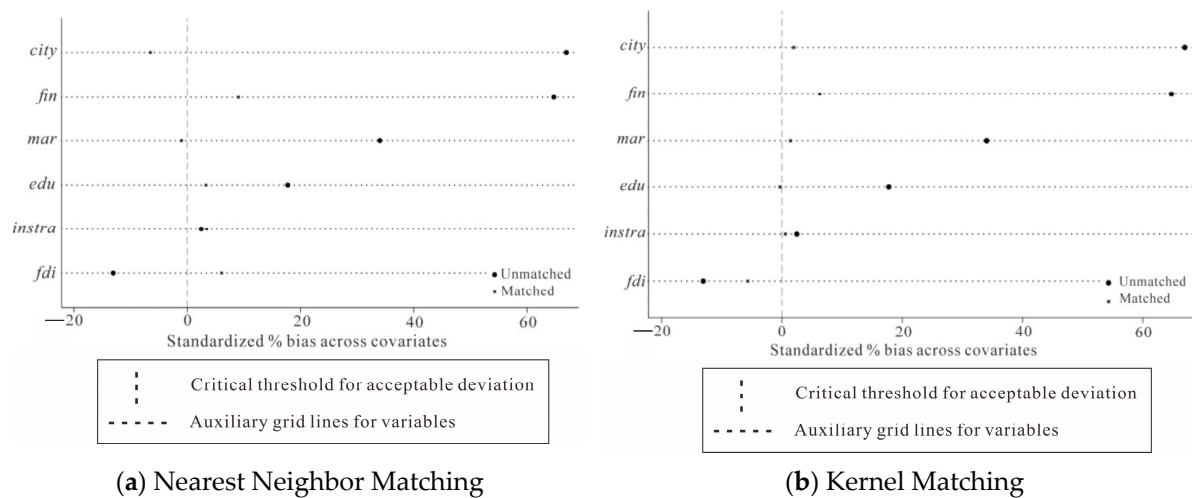


Figure 4. Balance test for the dependent variable *pollution*.

Table 4. PSM-DID Estimation Results.

Variable	(1)	(2)	(3)	(4)
	Nearest Neighbor Matching	Kernel Matching	Nearest Neighbor Matching	Kernel Matching
	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Pollution</i>	<i>Pollution</i>
<i>Treat × Post</i>	−0.0426 *** (0.005)	−0.0455 *** (0.005)	−0.0146 *** (0.002)	−0.0148 *** (0.002)
Constant	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	5415	5415	5415	5415
R-squared	0.994	0.994	0.760	0.759

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors.

5.3.4. Additional Robustness Tests

To rigorously examine the stability and reliability of the baseline estimates, this study conducts a series of robustness checks for the two dependent variables—*lnpc_{o2}* and *pollution*—to rule out potential confounding influences.

Columns (1) and (4) of Table 5 exclude observations from the 2020–2022 pandemic period to mitigate the impact of anomalous years on the findings; the policy coefficient remains statistically significant and negative. Columns (2) and (5) further control for concurrent policies—including the Artificial Intelligence Innovation and Development Pilot and the Cross-Border E-Commerce Pilot—to eliminate interference from policy overlap; the estimated coefficients of the main variable retain both their negative sign and statistical significance. Columns (3) and (6) add two control variables: per capita GDP (measuring economic development level) and fiscal intervention (measuring government intervention

intensity), to alleviate bias arising from omitted variables; the estimation results closely replicate those of the benchmark regression.

Table 5. Results of additional robustness tests.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Excluding Pandemic Years	Controlling for Other Policies	Adding New Control Variables	Excluding Pandemic Years	Controlling for Other Policies	Adding New Control Variables
	<i>Lnpco₂</i>	<i>Lnpco₂</i>	<i>Lnpco₂</i>	<i>Pollution</i>	<i>Pollution</i>	<i>Pollution</i>
<i>Treat × Post</i>	−0.0424 *** (0.006)	−0.0426 *** (0.005)	−0.0376 *** (0.006)	−0.0145 *** (0.002)	−0.0139 *** (0.002)	−0.0134 *** (0.002)
<i>aidid</i>		−0.0782 *** (0.018)			−0.0288 *** (0.006)	
<i>cbedid</i>		−0.0352 *** (0.005)			−0.0094 *** (0.001)	
<i>lnpgdp</i>			0.0585 *** (0.014)			0.0104 *** (0.002)
<i>gov</i>			0.0628 *** (0.021)			0.0099 *** (0.004)
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4560	5415	5415	4560	5415	5415
R-squared	0.994	0.994	0.994	0.780	0.767	0.761

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors.

Robustness checks confirm that the NBDPZ policy consistently and significantly lowers *lnpco₂* and *pollution*. Digital infrastructure and market-oriented data reforms thus deliver sustained, systemic environmental gains in urban governance, offering strong empirical support for digital green governance.

5.3.5. Heterogeneous Treatment Effects

In the multi-period staggered DID framework, estimating the average treatment effect (ATE) via the conventional two-way fixed effects (TWFE) estimator yields biased results when treatment effects vary heterogeneously across time and units. In the context of dynamic policy shocks in particular, the “bad control group” problem is highly likely to arise—specifically, later-treated units use earlier-treated units as their control group, leading to cross-period contamination of the estimated coefficients. In light of this, this paper applies the Bacon decomposition [47] to identify heterogeneous treatment effects in the baseline regression. The results are shown in Table 6.

The results indicate that, for *lnpco₂*, the weights of both potentially biasing comparison groups—“received treatment earlier” and “have not yet received treatment”—are negligible. The same holds for *pollution*. In contrast, the “never received treatment” control group accounts for 99.62% of the weight for both outcomes, and its estimated coefficients closely match the baseline results. Heterogeneous treatment effects therefore exert virtually no influence on the estimates, confirming baseline result robustness.

Table 6. Results of Goodman-Bacon decomposition.

Group	<i>lnpco₂</i>	<i>Pollution</i>	Weight
	Coefficient	Coefficient	
Never_v_timing	−0.0524	−0.0150	0.9962
Early_v_Late	−0.0321	−0.0011	0.0021
Late_v_Early	0.0248	−0.0158	0.0017

We further test robustness using the CSDID method [48] and Imputation Estimation [49]. The results are shown in Table 7. Pre-policy average treatment effects (Pre_avg) for *lnpco₂* (−0.0043) and *pollution* (−0.0001) are insignificant and near zero—confirming no systematic pre-treatment differences and satisfying the parallel trends assumption. Post-policy dynamic effects (Post_avg) are −0.0154 (*lnpco₂*, significant at 10%) and −0.0124 (*pollution*, significant at 1%), demonstrating that the policy persistently reduces both outcomes with cumulative impact. Imputation Estimation yields ATTs of −0.1843 (*lnpco₂*) and −0.0137 (*pollution*), both significant at 1%. The NBDCPZ policy thus significantly reduces urban carbon emissions and pollution—even after fully accounting for cross-city variation in policy timing and treatment effect heterogeneity.

Table 7. Estimation Results for heterogeneous treatment effects.

Variable	CSDID Estimation: Dynamic Average Treatment Effect		Imputation Estimation	
	<i>lnpco₂</i>	<i>pollution</i>	<i>lnpco₂</i>	<i>pollution</i>
Pre_avg	−0.0043 (0.003)	−0.0001 (0.001)		
Post_avg	−0.0154 * (0.009)	−0.0124 *** (0.004)		
ATT			−0.1843 *** (0.060)	−0.0137 *** (0.005)

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors.

Overall, the imputation ATT for *lnpco₂* (−0.1843) exceeds the Post_avg estimate, indicating cumulative amplification of policy effects as implementation progresses and digital infrastructure expands. Thus, the NBDCPZ delivers short-term environmental gains and long-term green transformation momentum—directly supporting the “dual carbon” goals.

6. Extensions and Further Analysis

6.1. Mechanism Identification and Testing

This section uses a chain mediation model to examine how the independent variable affects the dependent variable through sequentially linked mediators—revealing the underlying causal mechanism.

6.1.1. The Emission Reduction Effect of Green Innovation

This study operationalizes Green Innovation (*gi*) as the per capita count of green patent applications. This indicator is constructed through the following steps: first, drawing on the International Patent Classification (IPC) Green List—introduced by the World Intellectual Property Organization (WIPO) in 2010 to facilitate the identification and re-

trieval of environmentally friendly technologies—we match and identify green patents by cross-referencing the IPC codes provided by the China National Intellectual Property Administration. The total count of such identified green patents is then divided by the year-end urban population to yield the per capita measure. Additionally, drawing on the method proposed by [50], urban energy consumption is inferred from nighttime light data, and the urban per capita energy consumption intensity (*lnpny*) is calculated. The basic logic is that the brighter the nighttime lights, the more active the nighttime economic activity, the higher the level of economic development, and the greater the energy consumption.

Results show that the NBDCPZ policy significantly boosts urban green innovation levels, and green innovation further significantly reduces energy consumption intensity. Green innovation—covering energy-saving equipment, clean processes, and energy management technologies—raises energy output per unit input and cuts fossil fuel dependence, reducing total energy use. Lower energy intensity simultaneously cuts CO₂ emissions and co-emitted pollutants like SO₂ and particulate matter. Thus, green innovation reduces pollution and carbon emissions not directly, but by transforming how energy is used and how goods are produced. Table 8 confirms: *gi* significantly reduces *lnpny*, which in turn significantly increases both *lnpco₂* and *pollution*. Hence, improved energy efficiency is the central mediating channel—green innovation drives environmental gains through efficiency gains, not mere correlation. The full mechanism is “NBDCPZ → green innovation → lower energy intensity → reduced pollution and carbon emissions,” supporting Hypothesis H2. This source-level pathway shifts cities from passive governance to proactive emission reduction, improving mitigation efficiency and strengthening endogenous green growth.

Table 8. Mechanism testing for the emission reduction effect of *gi*.

Variable	(1) <i>Gi</i>	(2) <i>Lnpc_{o2}</i>	(3) <i>Pollution</i>	(4) <i>Lnyny</i>	(5) <i>Lnpc_{o2}</i>	(6) <i>Pollution</i>
<i>Treat</i> × <i>Post</i>	0.2656 *** (0.095)					
<i>gi</i>		−0.0144 *** (0.002)	−0.0030 *** (0.000)	−0.0439 *** (0.003)		
<i>lnyny</i>					0.1229 *** (0.009)	0.0090 *** (0.001)
Constant	YES	YES	YES	YES	YES	YES
Control Variables	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	5415	5415	5415	5415	5415	5415
R-squared	0.741	0.994	0.760	0.935	0.994	0.754

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors.

6.1.2. Innovation Effects of Digital Industry Clusters

The digital industry clustering level (*dagglo*) is measured using the location quotient, calculated as the ratio of the share of workers employed in information transmission, software, and information technology services in cities to the corresponding national share in that sector. Following the methodology of [51], urban digital innovation capability (*dinnov*) is quantified by compiling the annual number of digital invention patent applications filed

in each prefecture-level city, as classified under the *Patent Classification System for Key Digital Technologies* (2023) released by the CNIPA.

Column (1) of Table 9 reveals that the policy exerts a statistically significant positive effect on *dagglo*. In turn, *dagglo* strengthens *dinnov*, which significantly lowers per capita carbon emissions and environmental pollution indices. Thus, agglomeration improves environmental outcomes not directly, but via enhanced digital innovation. Theoretically, it fosters key externalities—knowledge spillovers, talent mobility, and collaborative R&D—not merely by increasing firm count, but through spatial concentration of digital firms, universities, research institutes, and tech service platforms. This concentration cuts information search and collaboration costs, accelerates knowledge diffusion and tech iteration, and lifts digital innovation output. Applications of digital innovations—such as big data analytics, AI, industrial internet, and intelligent scheduling—in energy management, pollution monitoring, and production optimization improve resource allocation efficiency and governance precision, cutting both pollutants and carbon emissions. The full pathway—“NBDCPZ → digital agglomeration → digital innovation capacity → reduced emissions and pollution”—is confirmed, supporting Hypothesis H3.

Table 9. Mechanism test of the innovation effect of *dagglo*.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Dagglo</i>	<i>Lnpc_{o2}</i>	<i>Pollution</i>	<i>Dinnov</i>	<i>Lnpc_{o2}</i>	<i>Pollution</i>
<i>Treat</i> × <i>Post</i>	0.0897 *** (0.021)					
<i>dagglo</i>		−0.0094 ** (0.004)	−0.0088 *** (0.001)	4.5318 *** (0.533)		
<i>dinnov</i>					−0.0036 *** (0.001)	−0.0006 *** (0.000)
Constant	YES	YES	YES	YES	YES	YES
Control Variables	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	5415	5415	5415	5415	5415	5415
R-squared	0.670	0.993	0.757	0.782	0.994	0.758

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors.

6.2. Heterogeneity Analysis

This paper examines the heterogeneous effects from three perspectives: resource endowment, city size, and industrial foundation.

Heterogeneity in resource endowment. Following the categorization framework outlined in the “*National Sustainable Development Plan for Resource-based Cities (2013–2020)*”, the urban sample is stratified into two groups: 115 resource-based cities and 170 non-resource-based cities. Columns (1) and (2) of Tables 10 and 11 show the results. Results show the policy significantly reduces pollution and carbon emissions in both city types, but with clear dimension-specific effects. For *lnpc_{o2}*, the effect is significantly stronger in non-resource-based cities; for *pollution*, it is significantly stronger in resource-based cities. Resource-based cities face carbon path dependence due to long-standing reliance on extraction and processing, limiting short-term carbon reductions. However, their conventional pollutants are highly concentrated, enabling big data tools to deliver rapid

pollution cuts. Non-resource-based cities, with diversified industries, leverage big data-driven upgrades to cut carbon emissions quickly, but with conventional pollutants already low, further reductions offer limited marginal gains.

Table 10. Heterogeneity test 1.

	Resource-Based Cities	Non-Resource-Based Cities	Large Cities	Medium and Small Cities	Old Industrial Bases	Non-Old Industrial Bases
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>	<i>Lnpc_{o2}</i>
<i>Treat × Post</i>	−0.0175 **	−0.0613 ***	−0.0417 ***	−0.0393 ***	−0.0527 ***	−0.0442 ***
	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.007)
Constant	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Empirical <i>p</i> -value	0.029 **		0.009 ***		0.003 ***	
Observations	2185	3230	1957	3458	1805	3610
R ²	0.995	0.992	0.995	0.992	0.991	0.995

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors. The statistical significance of the disparity in adjusted coefficients across groups is assessed using an empirical *p*-value, which is derived from 1000 bootstrap replications.

Table 11. Heterogeneity test 2.

	Resource-Based Cities	Non-Resource-Based Cities	Large Cities	Medium and Small Cities	Old Industrial Bases	Non-Old Industrial Bases
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Pollution</i>	<i>Pollution</i>	<i>Pollution</i>	<i>Pollution</i>	<i>Pollution</i>	<i>Pollution</i>
<i>Treat × Post</i>	−0.0232 ***	−0.0099 ***	−0.0238 ***	−0.0077 ***	−0.0153 ***	−0.0132 ***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)
Constant	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Empirical <i>p</i> -value	0.000 ***		0.000 ***		0.000 ***	
Observations	2185	3230	1957	3458	1805	3610
R ²	0.700	0.782	0.787	0.668	0.681	0.783

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Values in parentheses are robust standard errors. The statistical significance of the disparity in adjusted coefficients across groups is assessed using an empirical *p*-value, which is derived from 1000 bootstrap replications.

Heterogeneity in city size. This study classifies Chinese cities into two categories—large cities and medium- and small-sized cities—according to the “*Notice on Adjusting the Criteria for Classifying City Sizes (2014)*”. Columns (3) and (4) of Tables 10 and 11 present the subsample regression results. The findings indicate that the policy has a significant effect on environmental outcomes in both types of cities, but the effects exhibit dimensional heterogeneity. Regarding *lnpc_{o2}*, the policy impacts observed in both city categories are broadly comparable; regarding *pollution*, the policy effect is stronger in large cities. Large cities have well-developed digital infrastructure and mature smart environmental protection systems, and big data technology can achieve systematic pollution abatement.

However, their industrial structures are relatively optimized, leaving limited marginal space for carbon mitigation. Medium- and small-sized cities have weak digital foundations and face greater difficulties in pollution supervision, limiting the effectiveness of pollution abatement. However, they have greater potential for improving industrial energy efficiency, and big data technology transformation can quickly achieve carbon mitigation.

Heterogeneity in industrial foundation. Following the “*National Adjustment and Transformation Plan for Old Industrial Bases (2013–2022)*”, the sample cities are stratified into two groups: old industrial base cities and non-old industrial base cities. Columns (5) and (6) of Tables 10 and 11 present the results, showing that the NBDCPZ has statistically significant effects on both pollution reduction and carbon emission reduction in the two city categories under examination. The policy has a stronger effect in old industrial base cities than in non-old industrial base cities. The difference in economic effects between the two groups is statistically significant. The former category has a high proportion of traditional heavy industries and a large energy consumption and emissions base. Data empowerment can effectively drive the optimization of manufacturing processes and the upgrade of energy management, unlocking greater emission reduction potential. Non-old industrial base cities feature relatively light industrial structures and more robust environmental governance foundations, yielding comparatively smaller policy-induced marginal improvements.

All the above heterogeneity analyses pass the significance test for differences in adjusted coefficients across groups. These show the NBDCPZ policy has non-uniform effects across cities, constrained by resource endowments, city size, and industrial foundations. Thus, no single optimal model of digital governance applies universally; regions must tailor implementation to their industrial structure, environmental priorities, and digital infrastructure. These findings extend the boundary conditions of research on the digital economy’s environmental effects and offer empirical support for designing targeted digital policies.

7. Conclusions and Recommendations

This study employs a panel dataset covering 285 Chinese cities from 2016 to 2023, leveraging the staggered rollout of the NBDCPZ policy as a quasi-experimental setting to credibly estimate its causal effects. It adopts a multi-period DID framework to comprehensively assess the policy’s impacts on pollutant abatement and greenhouse gas emission reduction, as well as to identify the underlying mechanisms. The key findings emerging from the empirical analysis are summarized below: First, the NBDCPZ policy significantly reduces *lnpco₂* and *pollution*, achieving synergistic gains in pollution abatement and carbon mitigation. The core finding holds consistently across multiple robustness checks and heterogeneity treatment effect tests. Second, the policy exerts its effect through two chain mediation pathways: (1) improved green innovation levels → reduced energy consumption intensity; and (2) increased digital industry agglomeration → enhanced digital innovation capability. Both pathways jointly drive urban pollution abatement and carbon mitigation. Third, the policy effect exhibits significant heterogeneity: non-resource-based cities show more prominent carbon mitigation effects, whereas resource-based cities, large cities, and old industrial base cities demonstrate more significant pollution abatement effects.

Drawing on these findings, we propose the following three targeted policy recommendations: first, accelerate the replication and scaling up of experiences from the NBDCPZ and establish a long-term mechanism for digital-enabled pollution abatement and carbon mitigation. Drawing on experience from existing Pilot Zones, policy expansion should prioritize central and western regions, resource-based cities, and old industrial bases, rather than uniform nationwide coverage. Promotion precision must be grounded in empiri-

cally verified net effects. Crucially, rollout must align with local fiscal capacity and data infrastructure—avoiding waste and implementation resistance. Establish a unified national big data public service platform for pollution abatement and carbon mitigation, integrating multi-source data (e.g., industrial pollution discharge, energy consumption, carbon emissions, and green innovation) to enable real-time pollution source monitoring, precise traceability, and collaborative governance. Platform development must address inconsistent data standards, interdepartmental information silos, and privacy compliance risks, ensuring operability and security. Strengthen market-oriented data allocation mechanisms; clarify legal boundaries for collection, sharing, and use; and balance public interest against corporate trade secrets. Implement tiered access and security reviews to institutionalize digital empowerment of environmental governance.

Second, strengthen the chain mediation mechanism and unblock the core pathways through which digital technology enables pollution abatement and carbon mitigation. On one hand, for the pathway of “green innovation → reduction in energy consumption,” scale up support for green innovation by establishing a dedicated fund for the integrated innovation of big data and green technologies and guiding enterprises to increase R&D investment in green technologies. On the other hand, for the pathway of “digital industry agglomeration → digital innovation upgrading,” accelerate the development of digital industrial clusters by leveraging Pilot Zones to cultivate a cohort of digitally enabled enterprises and industrial parks with core competitiveness; grant digital enterprises preferential policies such as tax reductions and exemptions, venue subsidies, and talent attraction and cultivation support; provide special incentives for digital environmental protection patents and digital tools for carbon accounting. Policy design must account for SMEs’ financing constraints, using phased funding or performance-based incentives to lower upfront investment risk. Governments must also assess local fiscal capacity to prevent subsidy-driven resource misallocation, and directly tackle high retrofit costs, technology mismatch, and talent gaps through targeted training and technical support.

Third, implement differentiated and precision-targeted policies to enhance policy adaptability and effectiveness. For resource-based cities, prioritize supporting the deployment of big data technologies to establish real-time monitoring systems for industrial pollution sources; strengthen end-of-pipe pollution control; and leverage local green energy advantages to develop green computing power industries, thereby gradually reducing path dependence on carbon-intensive pathways [52]. For non-resource-based cities, expand the application of digital dispatching for clean energy to amplify carbon emission reduction advantages; for large cities, develop benchmark scenarios of smart environmental protection, construct a digital governance system characterized by full-domain perception, intelligent early warning, and collaborative disposal, and consolidate the leading edge in pollution control. For small- and medium-sized cities, increase investment in digital infrastructure; prioritize addressing the shortcomings in digital infrastructure; promote low-cost and lightweight smart environmental protection supervision tools, and rapidly elevate the digitalization level of pollution reduction and carbon emission reduction. For old industrial base cities, establish dedicated digital transformation guidance funds; promote intelligent and green upgrading of heavy chemical industries; and unlock emission reduction potential through industrial restructuring. Policy implementation must account for three key constraints: energy infrastructure upgrade costs, local fiscal capacity, and digital governance capability, avoiding hasty rollout that wastes resources or erodes investment returns. In digitally underdeveloped or fiscally strained regions, adopt a pilot-then-scale approach: launch targeted demonstrations in priority sectors first, then expand only after validating effectiveness and achieving scale. Finally, align short-term spending with long-term environmental gains, and match infrastructure investment to real-world application needs,

advancing digital transformation affordably and cohesively to deliver both decarbonization and high-quality regional growth.

This study contributes to the literature in three ways. First, it leverages national big data Pilot Zones as an exogenous shock and applies a difference-in-differences design to mitigate endogeneity, while also enabling quantitative measurement of “big data” as a policy variable. Second, it identifies two causal pathways via chain mediation: (1) “pilot policy → green innovation → pollution & carbon reduction” and (2) “pilot policy → digital agglomeration → digital innovation → pollution & carbon reduction.” Third, it examines heterogeneity across resource endowments, city size, and industrial base, clarifying where and for whom the findings generalize.

This study has three key limitations and corresponding research opportunities. First, endogeneity: pilot city selection is non-random, and unobserved confounders persist despite robustness checks. Future work should leverage exogenous instruments, regression discontinuity, or natural experiments to strengthen causal identification. Second, mechanism complexity: while we focus on green innovation and digital agglomeration, digital technologies affect environmental performance through multiple channels—including digital finance [53,54], firm-level digital transformation [55], environmental disclosure [56], and green resource allocation efficiency [57]. Next steps include building a unified “digital infrastructure → firm behavior → environmental outcomes” framework using microdata and multi-level governance insights. The theoretical analysis of this study does not explore research mechanisms such as nonlinearity, threshold effects, spatial spillovers, or policy interaction mechanisms, and these aspects could be further strengthened in future research. Third, temporal dynamics: the 2016 big data pilot’s environmental impact likely evolves as infrastructure matures and data accumulates. Longer-term tracking is needed to inform adaptive policy design and exit strategies.

Author Contributions: S.Y.: Data curation, Investigation, Validation, Writing—review & editing. L.W.: Project administration, Writing—review & editing. W.Z.: Conceptualization, Methodology, Investigation, Formal analysis, Writing—original draft, Visualization, Funding acquisition, Resources. Y.L.: Investigation, Data curation, Writing—review & editing. C.S.: Supervision, Investigation, Writing—review & editing. S.W.: Investigation, Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Youth Project of Guangdong Provincial Philosophy and Social Sciences Planning (Grant No. GD24YYJ03), the Scientific Research Start-up Funds of Guangdong Ocean University (Grant No. 060302082319), the Guangdong Philosophy and Social Sciences Planning Special Project for Eastern, Western, and Northern Guangdong (Grant No. GD24YDXZYJ02), and the Youth Project of Natural Science Foundation of Hunan Province (Grant No. 2026JJ70110).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Münzel, T.; Hahad, O.; Sørensen, M.; Lelieveld, J.; Duerr, G.D.; Nieuwenhuijsen, M.; Daiber, A. Environmental risk factors and cardiovascular diseases: A comprehensive expert review. *Cardiovasc. Res.* **2022**, *118*, 2880–2902. [PubMed]
2. Loebbecke, C.; Picot, A. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *J. Strateg. Inf. Syst.* **2015**, *24*, 149–157. [CrossRef]
3. Nambisan, S.; Wright, M.; Feldman, M. The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. *Res. Policy* **2019**, *48*, 103773. [CrossRef]

4. Lanzolla, G.; Pesce, D.; Tucci, C.L. The digital transformation of search and recombination in the innovation function: Tensions and an integrative framework. *J. Prod. Innov. Manag.* **2021**, *38*, 90–113.
5. Liu, R.; Hou, M.; Jing, R.; Bauer, A.; Wu, M. The impact of national big data pilot zones on the persistence of green innovation: A moderating perspective based on green finance. *Sustainability* **2024**, *16*, 9570. [\[CrossRef\]](#)
6. Wang, M.; Wu, Y.; Zhang, X.; Lei, L. How does industrial agglomeration affect internal structures of green economy in China? An analysis based on a three-hierarchy meta-frontier DEA and systematic GMM approach. *Technol. Forecast. Soc. Change* **2024**, *206*, 123560. [\[CrossRef\]](#)
7. Ahmad, N.; Liu, Y.; Žiković, S.; Belyaeva, Z. The effects of technological innovation on sustainable development and environmental degradation: Evidence from China. *Technol. Soc.* **2023**, *72*, 102184. [\[CrossRef\]](#)
8. Xu, L.; Fan, M.; Yang, L.; Shao, S. Heterogeneous green innovations and carbon emission performance: Evidence at China's city level. *Energy Econ.* **2021**, *99*, 105269. [\[CrossRef\]](#)
9. Yi, H.; Zhao, L.; Qian, Y.; Zhou, L.; Yang, P. How to achieve synergy between carbon dioxide mitigation and air pollution control? Evidence from China. *Sustain. Cities Soc.* **2022**, *78*, 103609. [\[CrossRef\]](#)
10. Xu, Y.; Peng, J.; Tian, J.; Fu, S.; Hu, C.; Feng, Y. The impact and mechanism analysis of clean development mechanism on the synergistic effects of pollution mitigation and carbon reduction. *Environ. Res.* **2024**, *260*, 119659. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Lee, K.H.; Min, B. Green R&D for eco-innovation and its impact on carbon emissions and firm performance. *J. Clean. Prod.* **2015**, *108*, 534–542. [\[CrossRef\]](#)
12. Du, K.; Li, P.; Yan, Z. Do green technology innovations contribute to carbon dioxide emission reduction? Empirical evidence from patent data. *Technol. Forecast. Soc. Change* **2019**, *146*, 297–303. [\[CrossRef\]](#)
13. Lin, B.; Ma, R. Green technology innovations, urban innovation environment and CO₂ emission reduction in China: Fresh evidence from a partially linear functional-coefficient panel model. *Technol. Forecast. Soc. Change* **2022**, *176*, 121434. [\[CrossRef\]](#)
14. Wang, L.; Sun, H.; Hu, X.; Leng, Z.; Zha, Z.; Zhang, Z.; Yang, J. Measurement of China's provincial consumption-based PM_{2.5} emissions and its influencing factors from the perspective of spatial heterogeneity. *J. Clean. Prod.* **2021**, *317*, 128367. [\[CrossRef\]](#)
15. Guo, M.; Hamm, N.A.S.; Chen, B. Understanding air pollution reduction from the perspective of synergy with carbon mitigation in China from 2008 to 2017. *J. Environ. Manag.* **2024**, *367*, 122017. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Wei, X.; Jiang, F.; Yang, L. Does digital dividend matter in China's green low-carbon development? Environmental impact assessment of the national big data comprehensive pilot zones policy. *Environ. Impact Assess. Rev.* **2023**, *101*, 107143. [\[CrossRef\]](#)
17. Bu, H.; Li, G.; Yu, X.; Xun, Z. Is smart carbon emission reduction justified in China? Evidence from NBDPCZs. *Heliyon* **2023**, *9*, e17887. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Pan, L.; Zhang, Y.; Li, J.; Lin, F. Digital economy and urban carbon emissions: A quasi-natural experiment of the national big data comprehensive pilot zones in China. *Energy J.* **2024**, *45*, 107–134. [\[CrossRef\]](#)
19. Yan, Y.; Li, Z.; Zhang, Y.; Gong, B. Digital economy and environmental co-benefits: Evidence on air pollution and carbon emissions in China. *Econ. Anal. Policy* **2026**, *90*, 1347–1368. [\[CrossRef\]](#)
20. Hu, S.; Chen, W.; Liu, Y. Policy synergies in urban digital governance for pollution reduction and carbon mitigation: Evidence from Chinese cities. *Sustain. Cities Soc.* **2025**, *130*, 106575. [\[CrossRef\]](#)
21. Luo, Q.; Wang, F.; Wang, C.; Dong, X. Synergistic effects of the dual pilot policy on pollution reduction and carbon mitigation: The role of data empowerment and market mechanisms. *J. Clean. Prod.* **2026**, *554*, 148031. [\[CrossRef\]](#)
22. Wiggberg, M.; Gulliksen, J.; Cajander, Å.; Pears, A. Defining digital excellence: Requisite skills and policy implications for digital transformation. *IEEE Access* **2022**, *10*, 52481–52507. [\[CrossRef\]](#)
23. Wang, Q.; Liu, Y. Unlocking green horizons in China: Digital industry agglomeration and corporate environmental performance. *J. Clean. Prod.* **2025**, *522*, 146260. [\[CrossRef\]](#)
24. Lin, B.; Zhou, Y. Does Internet development affect energy and carbon emission performance? *Sustain. Prod. Consum.* **2021**, *28*, 1–10. [\[CrossRef\]](#)
25. Han, Y.; Qu, S. How can environmental regulation and digital economy synergistically affect regional green innovation in China? Evidence based on a double machine learning model. *Environ. Dev. Sustain.* **2025**. [\[CrossRef\]](#)
26. Wu, C.; Zheng, X.; Zhang, N. Digital governance and environmental protection: Evidence from China's data-sharing policies. *Environ. Impact Assess. Rev.* **2026**, *119*, 108348. [\[CrossRef\]](#)
27. Fu, L. An empirical research on industry structure and economic growth. *Stat. Res.* **2010**, *27*, 79–81.
28. Tang, C.; Xu, Y.; Hao, Y.; Wu, H.; Xue, Y. What is the role of telecommunications infrastructure construction in green technology innovation? A firm-level analysis for China. *Energy Econ.* **2021**, *103*, 105576. [\[CrossRef\]](#)
29. Jin, W. Monitoring and management of green information in water ecological environment based on sensors and big data. *Meas. Sens.* **2024**, *34*, 101255. [\[CrossRef\]](#)
30. Acemoglu, D.; Aghion, P.; Bursztyn, L.; Hémous, D. The environment and directed technical change. *Am. Econ. Rev.* **2012**, *102*, 131–166. [\[CrossRef\]](#) [\[PubMed\]](#)

31. Sun, Y.; Stephen, E. Sustainable development pathways: How fiscal policy, green innovation, and environmental governance shape ecological footprints in G20 economies. *Sustain. Dev.* **2025**, *34*, 170–181. [\[CrossRef\]](#)
32. Zou, W. Can environmental regulation promote urban green innovation? *Econ. Surv.* **2023**, *40*, 24–33.
33. Hao, Y.; Guo, Y.; Wu, H. The role of information and communication technology on green total factor energy efficiency: Does environmental regulation work? *Bus. Strategy Environ.* **2022**, *31*, 403–424. [\[CrossRef\]](#)
34. Barforoush, N.; Etebarian, A.; Naghsh, A.; Shahin, A. Green innovation: A strategic resource to attain competitive advantage. *Int. J. Innov. Sci.* **2021**, *13*, 645–663. [\[CrossRef\]](#)
35. Jing, R.; Xu, Y.; Liu, R. Data element, big data and enterprise innovation: Evidence from data trading platforms and national big data comprehensive pilot zones. *Int. Rev. Financ. Anal.* **2025**, *105*, 104426. [\[CrossRef\]](#)
36. Sun, L.; Huang, C.; Yang, L. Can data factor aggregation enhance enterprise resilience? A quasi-natural experiment employing national big data comprehensive pilot zone. *Int. Rev. Econ. Financ.* **2026**, *106*, 105007. [\[CrossRef\]](#)
37. Mu, W.; Chang, J. Data aggregation and enterprise digital transformation: A quasi-natural experiment based on national big data pilot zones. *Acad. J. Bus. Manag.* **2025**, *7*, 164–171. [\[CrossRef\]](#)
38. Xu, R.Y.; Wang, K.L. How does digital industry agglomeration drive carbon peaking and carbon neutrality in China? The role of digital technology innovation. *J. Clean. Prod.* **2025**, *510*, 145642. [\[CrossRef\]](#)
39. Chen, W.; Zhang, J.; Cai, J. Big data capability, technical knowledge search, and green technology innovation: A mixed methods study. *J. Innov. Knowl.* **2026**, *16*, 101029. [\[CrossRef\]](#)
40. Li, L.; Jin, S.; Zhao, J.; Ma, C.; Diao, Y. Artificial intelligence and digital industry agglomeration development: Theory and empirical analysis from urban economics. *Econ. Anal. Policy* **2026**, *90*, 1071–1088. [\[CrossRef\]](#)
41. Zhao, Y.; Wu, Q.; Yang, W.; Han, X.; Men, L.; Liu, R.; Guo, Y. The impact of the national big data comprehensive pilot zone on digital economy development: Evidence from China. *Inf. Technol. Dev.* **2025**, *31*, 1297–1317. [\[CrossRef\]](#)
42. Wang, K.; Fang, C. A three-step synergy analysis framework for carbon mitigation and pollution reduction in China. *J. Environ. Manag.* **2025**, *386*, 125814. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Lu, D.; Hui, E.C.M.; Shen, J.; Shi, J. Digital industry agglomeration and urban innovation: Evidence from China. *Econ. Anal. Policy* **2024**, *84*, 1998–2025. [\[CrossRef\]](#)
44. Sun, S.; Wang, X.; Gao, C. Can green credit reduce carbon emissions? *China Popul. Resour. Environ.* **2023**, *33*, 37–47.
45. Qian, X.; Zou, W.; Sun, M.; Cheng, Y. Does transportation infrastructure contribute to urban entrepreneurship? Evidence from the high-speed railway in China. *Res. Transp. Bus. Manag.* **2024**, *57*, 101216. [\[CrossRef\]](#)
46. Wang, W.; Liu, Y.; Dong, X. Digital infrastructure's environmental paradox? Evidence from China's national big data comprehensive pilot zones. *Econ. Anal. Policy* **2025**, *88*, 281–307. [\[CrossRef\]](#)
47. Goodman-Bacon, A. Difference-in-differences with variation in treatment timing. *J. Econom.* **2021**, *225*, 254–277. [\[CrossRef\]](#)
48. Callaway, B.; Sant'Anna, P.H.C. Difference-in-differences with multiple time periods. *J. Econom.* **2021**, *225*, 200–230. [\[CrossRef\]](#)
49. Borusyak, K.; Jaravel, X.; Spiess, J. Revisiting event-study designs: Robust and efficient estimation. *Rev. Econ. Stud.* **2024**, *91*, 3253–3285. [\[CrossRef\]](#)
50. Shi, D.; Li, S. Emission trading system and energy efficiency: Measurement and empirical analysis on prefecture-level and above cities. *China Ind. Econ.* **2020**, *9*, 5–23.
51. Li, X.Q.; Shi, D.Q.; Liu, J.J. Digital innovation and urban green total factor productivity: A quasi-natural experiment based on the national big data comprehensive pilot zone. *China Ind. Econ.* **2024**, *5*, 134–152.
52. Wu, L.; Sun, L.; Qi, P.; Ren, X.; Sun, X. Energy endowment, industrial structure upgrading, and CO₂ emissions in China: Revisiting the resource curse in the context of carbon emissions. *Resour. Policy* **2021**, *74*, 102329. [\[CrossRef\]](#)
53. Cao, S.; Nie, L.; Sun, H.; Sun, W. Digital finance, green technological innovation and energy-environmental performance: Evidence from China's regional economies. *J. Clean. Prod.* **2021**, *327*, 129458. [\[CrossRef\]](#)
54. Liu, J.; Song, R.; Fu, Y. Digital finance empowering corporate ESG performance: The dual-engine role of digital transformation and green technological innovation. *Sustainability* **2025**, *17*, 10743. [\[CrossRef\]](#)
55. Guo, D.; Qi, F.; Wang, R.; Li, L. How does digital inclusive finance affect the ecological environment? Evidence from Chinese prefecture-level cities. *J. Environ. Manag.* **2023**, *342*, 118158. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Liu, H.; Liu, W.; Chen, G. Environmental information disclosure, digital transformation, and total factor productivity: Evidence from Chinese heavy-polluting listed companies. *Int. J. Environ. Res. Public Health* **2022**, *19*, 9657. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Xu, Q.; Li, X.; Guo, F. Digital transformation and environmental performance: Evidence from Chinese resource-based enterprises. *Corp. Soc. Responsib. Environ. Manag.* **2023**, *30*, 1816–1840. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.