land

Article

Effects of Big Data on $PM_{2.5}$: A Study Based on Double Machine Learning

Xinyu Wei [1], Mingwang Cheng [1], Kaifeng Duan [2,*] and Xiangxing Kong [1]

[1] School of Economics and Management, Tongji University, Shanghai 200092, China; weixinyu@tongji.edu.cn (X.W.)
[2] School of Economics and Management, Fuzhou University, Fuzhou 350108, China
* Correspondence: kefee920729@fzu.edu.cn

Abstract: The critical role of high-quality urban development and scientific land use in leveraging big data for air quality enhancement is paramount. The application of machine learning for causal inferences in research related to big data development and air pollution presents considerable potential. This study employs a double machine learning model to explore the impact of big data development on the $PM_{2.5}$ concentration in 277 prefecture-level cities across China. This analysis is grounded in the quasi-natural experiment named the National Big Data Comprehensive Pilot Zone. The findings reveal a significant inverse relationship between big data development and $PM_{2.5}$ levels, with a correlation coefficient of −0.0149, a result consistently supported by various robustness checks. Further mechanism analyses elucidate that big data development markedly diminishes $PM_{2.5}$ levels through the avenues of enhanced urban development and land use planning. The examination of heterogeneity underscores big data's suppressive effect on $PM_{2.5}$ levels across central, eastern, and western regions, as well as in both resource-dependent and non-resource-dependent cities, albeit with varying degrees of significance. This study offers policy recommendations for the formulation and execution of big data policies, emphasizing the importance of acknowledging local variances and the structural nuances of urban economies.

Keywords: big data development; $PM_{2.5}$; double machine learning; land use; high-quality urban development

Citation: Wei, X.; Cheng, M.; Duan, K.; Kong, X. Effects of Big Data on $PM_{2.5}$: A Study Based on Double Machine Learning. *Land* **2024**, *13*, 327. https://doi.org/10.3390/land13030327

Academic Editor: Chuanrong Zhang

Received: 21 January 2024
Revised: 15 February 2024
Accepted: 1 March 2024
Published: 4 March 2024

1. Introduction

The development of big data plays a crucial role in enhancing high-quality urban development and land use, improving urban air quality, and promoting sustainable urban growth. In the past decade, the global landscape has undergone a significant digital transformation, with big data development at its core. The digital economy's impact on China is comprehensive, evident across all sectors. It harbors the potential to drive sustainable urban development through industrial agglomeration, advancements in green technology, and the acceleration of industrial structural upgrades [1,2]. Moreover, it can facilitate high-quality economic growth via industrial digitalization and the enhancement of human capital [3,4]. The digital economy also plays a critical role in promoting the green transformation of agriculture, advancing high-quality agricultural development, and accelerating rural revitalization [5,6]. Research indicates that the digital economy significantly contributes to enhancing the commonwealth, boosting the competitiveness of imports and exports, and improving environmental quality [7–9]. Additionally, studies have found that the long-term impact of the digital economy on employment is overwhelmingly positive [10]. The digital economy is transforming every aspect of our lives, underscoring the transformative power of big data and digital technologies in fostering a more sustainable, efficient, and dynamic urban development paradigm.

At the same time, air pollution, particularly $PM_{2.5}$ pollution, persists as a formidable global challenge. $PM_{2.5}$ represents a significant environmental and public health hazard.

As a fine particulate matter capable of penetrating deep into the lungs and entering the bloodstream, $PM_{2.5}$ poses a grave threat to human health [11]. Recent years have seen China achieve notable improvements in air quality [12]. Significant strides have been achieved in reducing major air pollutants and enhancing air quality, attributed to initiatives such as the Action Plan for Prevention and Control of Air Pollution and the Three-Year Action Plan for Winning the Battle for the Blue Sky. The average $PM_{2.5}$ concentration in China witnessed a 57% decrease from 2013 to 2022, alongside rapid economic growth. However, despite these advancements, the average annual $PM_{2.5}$ concentrations in 339 prefectural-level cities and above in 2022 still surpassed the latest air quality guideline issued by the World Health Organization, highlighting a distinct disparity in the concentration levels between Europe and the United States. Studies have indicated that $PM_{2.5}$ pollution levels exceed international health standards, underscoring the ongoing need for enhanced pollution prevention and control efforts [13]. The state council's recent promulgation of the Action Plan for Continuous Improvement of Air Quality reflects the Chinese government's steadfast commitment to pollution prevention and treatment. The challenge of particulate matter pollution in China is exacerbated by rapid economic development, urbanization, significant industrial production growth, and increased automobile usage. The 2022 China Ecological Environment Status Bulletin reported that national emissions of particulate matter from exhaust gas amounted to 4,934,000 tons, with industrial sources contributing 62.0% and residential sources 37.0%. The prolonged suspension of fine particles in the atmosphere facilitates the formation of $PM_{2.5}$ pollution. Research indicates that industrial emissions, along with coal and biomass combustion for daily living, are principal sources of $PM_{2.5}$ [14–16]. Additionally, socio-economic and meteorological factors, as well as the generation of secondary particulate matter, further contribute to $PM_{2.5}$ pollution [17–19]. Meteorological conditions exacerbate $PM_{2.5}$ pollution through atmospheric effects and secondary aerosol formation [20]. Population density and economic growth are recognized as primary factors leading to more severe $PM_{2.5}$ pollution [21]. A U-shaped relationship exists between regional economic disparities and $PM_{2.5}$ pollution, suggesting that both extremely high and low levels of economic development can aggravate $PM_{2.5}$ pollution. Furthermore, the interplay between economic development and $PM_{2.5}$ levels can influence population migration, thereby impacting $PM_{2.5}$ concentrations. Consequently, socio-economic factors are intricately linked with $PM_{2.5}$ pollution [22,23].

The establishment of public big data comprehensive pilot zones is recognized as a strategic initiative to address environmental challenges and transform economic growth models. For instance, since Guizhou was designated as the inaugural public big data comprehensive pilot zone, the digital economy's value added has accounted for approximately 37% of the province's GDP, highlighting the pivotal role of digitalization in economic expansion. Furthermore, the implementation of the pilot zone has exerted a favorable influence on environmental protection, notably in reducing carbon emissions and mitigating environmental pollution [24,25]. In the context of China's contemporary economic development, the digital economy assumes a critical role. Its advancement facilitates the fusion of industrialization and informatization, thereby elevating the level of industrialization. This process could potentially reduce $PM_{2.5}$ emissions from industrial sources to a certain degree [26]. Recent studies have indicated that the development of the digital economy significantly impacts the reduction in $PM_{2.5}$ pollution [27]. The influence of the digital economy on $PM_{2.5}$ is primarily mediated through key factors, such as urban quality development and land use planning [28–32]. The exact manner in which the digital economy affects $PM_{2.5}$ pollution is not yet fully understood, necessitating further research to elucidate its mechanisms comprehensively. Such investigations will furnish a theoretical foundation for formulating national economic development and environmental protection policies and strategies.

This paper introduces a double machine learning model to assess the impact of establishing comprehensive big data pilot zones on urban $PM_{2.5}$ concentrations, utilizing panel data from 277 prefecture-level cities across China. It delves into the creation of national big data comprehensive pilot zones in various regions of China and investigates the influence of big data development on the $PM_{2.5}$ levels in these areas. The study highlights regional disparities and examines how big data development affects $PM_{2.5}$ concentrations through mechanisms such as capital and labor reallocation, technological innovation, population migration, and the upgrading of industrial structures. To verify the robustness of these findings, the research conducts several robustness tests.

The paper offers significant contributions in two main areas: firstly, it introduces a novel research angle by analyzing the connection between big data development and $PM_{2.5}$ concentrations within the framework of national policies, underlining the importance of big data policies in environmental management. Secondly, it pioneers a methodological approach by employing a double machine learning model to directly assess the impact of big data development on $PM_{2.5}$ levels. This approach provides a comprehensive theoretical framework that intricately links big data development with environmental governance, offering insights into the multifaceted ways big data initiatives influence $PM_{2.5}$ levels and enhancing the precision of causal inference.

The paper is methodically structured as follows: the second section discusses the rationale and significance of the National Big Data Comprehensive Pilot Zone, the relationship between big data development and environmental pollution, and outlines the research hypotheses. The third section details the design of the double machine learning model, the selection of the variables, and the data sources employed. The fourth section presents an analysis and discussion of the empirical findings. The concluding section summarizes the study's findings and offers policy recommendations based on the research outcomes. This comprehensive study provides a foundation for leveraging big data development to support the formulation of effective environmental policies, offering a fresh perspective on the role of digital technologies in environmental sustainability.

2. Background, Literature Review, and Hypotheses

2.1. Background

The National Comprehensive Pilot Zone for Big Data stands as a crucial policy in China, aimed at navigating the digitization trend, enhancing the strategic importance of data resources, and propelling the development of the digital economy. This initiative was initially highlighted in the 2014 Government Work Report, signifying the onset of the "big data" era. Concurrently, numerous policy documents, such as the "Action Plan to Promote Big Data Development", were issued. In August 2015, the State Council promulgated the "Action Outline to Promote Big Data Development", recommending pilot projects in specific areas and the advancement of comprehensive big data pilot zones in Guizhou and other regions. By 2016, the pilot zone initiative had expanded to encompass Beijing-Tianjin-Hebei, the Pearl River Delta, Shanghai, Henan Province, Chongqing City, Shenyang City, and the Inner Mongolia Autonomous Region. The establishment of the National Big Data Comprehensive Pilot Zone reflects China's proactive stance in adapting to global digitization trends and fostering the domestic digital economy, supported by a series of policy documents like the Action Plan for Promoting Big Data Development. The National Big Data Comprehensive Pilot Zone has recorded remarkable achievements. According to the National Big Data (Guizhou) Comprehensive Pilot Zone Development Report 2022, the digital economy accounted for approximately 37% of the province's GDP in 2022, maintaining the highest growth rate in the nation for seven consecutive years. The report also details the promotion of the National Integrated Big Data Network National (Guizhou) Hub Node and an investment of 3.913 billion yuan in addition to the construction of 31,000 new 5 G base stations, totaling 84,300 stations.

These pilot zones have effectively spurred digital transformation and development, which are believed to significantly lower regional carbon emissions and meet emission reduction goals. Research has demonstrated that the digital economy has an inverted U-shaped correlation with environmental pollution, where pollutants may initially increase during the early stages of economic development but tend to decline after reaching a certain level of development due to enhanced environmental protection efforts and the advent of new technologies [33]. Moreover, the establishment of big data pilot zones has been linked to increased environmental productivity in urban areas, particularly in cities with high levels of marketization and resource-based economies [34]. Guo et al. (2022) [35] found that the expansion of the digital economy could significantly improve urban air quality, not only fostering the enhancement of local air quality but also motivating reductions in air pollution in neighboring cities. These findings highlight the significant role that big data pilot zones play in urban environmental governance.

2.2. Literature Review and Hypotheses

2.2.1. Literature Review

The establishment of big data pilot zones has been instrumental in driving enterprise digital transformation, leading to a decrease in regional carbon emissions and the fulfillment of emission reduction objectives [36]. Chuai et al. (2019) [37] devised a novel approach to analyze carbon emissions in Nanjing, concentrating on land use and socio-economic factors at a 300 m resolution. Hien et al. (2020) [38] investigated the effects of urban expansion in Hanoi on air quality, assessing ambient NO and SO concentrations in urban and peri-urban districts, thus highlighting the environmental implications of transforming agricultural land into urban areas. Huang et al. (2020) [39] explored big spatial data (BSD) across four case studies, which included detecting polycentric urban structures, evaluating urban vibrancy, estimating $PM_{2.5}$ exposure, and classifying urban land use with deep learning techniques. Wang et al. (2020) [40] developed an air quality prediction model utilizing big data and neural networks, demonstrating its proficiency in learning and forecasting air quality trends. Wu et al. (2021) [41] applied big data and machine learning to uncover the various mechanisms driving urban land use expansion in downtown Huizhou, noting the differential impacts of distinct factors according to land use types and development stages. Guo et al. (2021) [42] examined the correlation between street dynamic vitality and the distribution and combination of land use, identifying a positive relationship between vitality intensity and land use density and emphasizing the significance of transportation facilities and land use diversity in enhancing vitality stability.

Recent research suggests an inverted U-shaped relationship between the digital economy and environmental pollution. As the economy progresses, sulfur dioxide emissions increase but eventually start to decline after reaching a certain development threshold, possibly due to enhanced environmental protection measures and technological advancements [33]. Conversely, another study discovered that the growth of the digital economy significantly mitigates environmental pollution, especially industrial sulfur dioxide pollution, likely a result of technological innovation and the upgrading of industrial structures [4]. Additional studies have also noted improvements in environmental quality and a reduction in pollution attributable to the digital economy, which has facilitated the development of numerous green technologies and industries [43]. Big data is pivotal in the economic transformation towards green technological innovation, enabling enterprises to leverage big data for sustainable practices [33,44]. The advancement of the digital economy directly promotes high-quality green development, with the adjustment of industrial structures and innovation in green technology playing essential roles in this process [45]. Furthermore, the creation of comprehensive big data pilot zones is widely recognized for boosting the green total factor productivity of cities, particularly in highly marketized and resource-based urban areas, where the promotion of green total factor productivity is more pronounced [34]. The role of the digital economy in environmental improvement is further exemplified in green finance, which can serve as a negative regulator in addressing environmental pollution issues [46]. Communication ICT capital, in particular, can enhance carbon emission efficiency and contribute to emission reductions [47], demonstrating the digital economy's potential to support green financial activities and foster the growth of the "green economy", thereby aiding in the alleviation of environmental pollution through the adoption of clean technology and the optimization of industrial structures.

The impact of the digital economy on urban carbon emissions is influenced by spatial dynamics, demonstrating that the effects vary across different regions. In the eastern region, the digital economy does not have a noticeable impact on carbon emissions. Conversely, in the central region, the digital economy exacerbates carbon emissions, while in the western region, it serves to reduce them [48]. Contrarily, another investigation highlighted that the growth of the digital economy resulted in a significant decrease in pollution levels in the eastern region and a notable reduction in CO_2 emissions in the western region [49]. Bai et al. (2022) [50] corroborated these findings, revealing that larger cities, with their advanced digital infrastructure and systems, were more effective in reducing pollution. Nevertheless, the influence of the digital economy and big data is not universally positive. As identified by Ma (2023) [51], an intriguing paradox exists where, in cities with low innovation levels and underdeveloped digital economies, the digital transformation of commercial banks contributes to an increase in $PM_{2.5}$ pollution. This can be attributed to factors like inadequate infrastructure, a lack of expertise, or insufficient regulatory measures, which can impede the effective execution of digitalization initiatives.

Moreover, the development of the digital economy can also mediate carbon emissions through its effects on land use. Rapid urbanization in China has led to an extensive increase in construction land, resulting in the diminution of arable land. Land use changes indirectly influence anthropogenic carbon emissions and, consequently, environmental quality [52,53]. The digital economy could potentially lower carbon emissions from cropland by fostering innovation in green technologies and enhancing the efficiency of cropland's green transformation [54,55]. Furthermore, it can significantly boost the efficiency of urban land use and the ecological integrity of land use practices, promoting a low-carbon development trajectory [56–58]. The digital economy is poised to improve the quality of new urban development, address the imbalance of land resources in urban growth, and elevate the efficacy of urban green development [59,60]. In the realm of digital economic advancement, big data plays a crucial role in optimizing land use and reshaping conventional perspectives on land management. It enhances land planning and management capabilities and supports the development of modern urban areas [61–63]. Additionally, big data facilitates participatory approaches in land use planning, encouraging community involvement [64]. The interaction between land use changes and the digital economy also influences the relationship between digital economic growth and air pollution.

The impact of digital economy development on air pollution remains a contentious topic, with scholars yet to arrive at a consensus. While numerous studies have examined the relationship between big data development and environmental outcomes, the majority have concentrated on carbon emissions, leaving a gap in the literature regarding air pollution. Moreover, there is a scarcity of research employing machine learning for causal inferences in investigations of big data development and its effects on air pollution. This indicates a significant opportunity for further exploration into how digital transformations, driven by big data, influence various forms of environmental pollution, including air quality. The application of machine learning techniques could provide nuanced insights into the complex dynamics between digital economy growth and air pollution, offering a more comprehensive understanding of the potential benefits and challenges associated with digitalization in the context of environmental sustainability.

2.2.2. Hypotheses

Drawing on the Green Solow model [65], this study employs the super-efficiency SBM model to assess the eco-efficiency of 152 Chinese prefecture-level cities from 2003 to 2016. Subsequently, the DID model is utilized to investigate the impact of smart city pilot policies on eco-efficiency. Wu et al. (2022) [66] explored the implications of urbanization on energy conservation and emission reduction across 196 Chinese cities from 2011 to 2018, using a slacks-based approach. They advocated for policies that promote high-quality urban development and ECER, highlighting the critical roles of economic growth, resource allocation, internet technology, and employment structure, with a particular emphasis on innovation, coordination, and green development. This paper delves into high-quality urban development, concentrating on the resource mismatch index, urban green patent technology, and industrial structure upgrading.

The mismatch between labor and capital resources is identified as a significant contributor to environmental pollution [67]. Studies have investigated city cluster construction and economic agglomeration as viable strategies for diminishing environmental $PM_{2.5}$ levels. The formation of city clusters enhances resource distribution, thereby reducing environmental $PM_{2.5}$ concentrations. Economic agglomeration, on the other hand, lessens $PM_{2.5}$ pollution through the acceleration of industrial clustering and the minimization of resource mismatches [68–70]. The digital economy, as a nascent economic paradigm, has garnered interest for its potential to rectify resource mismatches. Various studies have posited that the digital economy can effectively mitigate resource mismatches, offering a theoretical foundation for the reduction in environmental $PM_{2.5}$ levels [71]. Consequently, this paper introduces hypothesis one:

H1. *The development of big data optimizes resource allocation, mitigates resource mismatches in cities, and thereby reduces $PM_{2.5}$ concentrations.*

The exponential growth of big data significantly impacts numerous sectors, notably in technological innovation. This influence is prominently observed in the digital economy, which is progressively shaping the urban innovation landscape. These advancements are closely associated with green, low-carbon development, playing a crucial role in pollution mitigation. Numerous studies [30–32] have highlighted the digital economy's role in fostering the development of cleaner technologies and greener products. This is facilitated through the promotion of scientific and technological innovations, especially patents in green technology. Innovations driven by data enhance the efficacy of pollution reduction technologies, culminating in cleaner and more sustainable urban ecosystems. Thus, this paper posits hypothesis two:

H2. *The advancement of big data fosters urban innovation, particularly by encouraging patents in green and low-carbon technologies, which in turn reduces $PM_{2.5}$ concentrations.*

The evolution of big data has significantly propelled digital development and transformation across various industries, facilitating the acceleration of enterprises' industrial structure upgrading and transformation [72,73]. Research has demonstrated that the digital economy enhances industrial structure upgrading, exerting a more substantial influence on the rationalization of industrial structures than on the advancement of industrial structures [74,75]. Conversely, other studies have revealed that the development of the digital economy notably encourages the advancement of the manufacturing industrial structure, with the advanced industrial structure playing a more pronounced mediating role in fostering industrial quality development through the digital economy [76–78]. The negative correlation between haze pollution and industrial structure upgrading was established [79], with a particular study in the Yangtze River Economic Zone showing that the optimization and upgrade of the industrial structure could effectively mitigate $PM_{2.5}$ pollution [80]. This research indicates that optimizing industrial structure reduces emissions and that combining industrial structure upgrading with technological innovation significantly diminishes $PM_{2.5}$ pollution [81,82]. Through both separate and combined effects, the digital economy advances industrial structure upgrades and technological innovation, subsequently reducing $PM_{2.5}$ pollution. Hence, this paper introduces hypothesis three:

H3. *Big data development facilitates the upgrading of urban industrial structures and contributes to the reduction in $PM_{2.5}$ concentrations.*

Research has reported the practicality of using population density in urban planning [83]. There is a correlation between population density and land use planning. In western Germany, a correlation between land use and population growth is evident in most areas [84]. Salvati (2012) [85] found that low population density could significantly influence specific land uses, suggesting that population density acts as an indirect marker for shifts in urban and rural land use patterns. Variations in population density are pivotal in elucidating land use allocation [86,87]. Shen (2009) [88] utilized an urban population density development model to simulate and validate sustainable land use and urban development strategies in their research. The interplay between land use planning and population density is significant; urbanization and land expansion directly affect urban population density, with some studies incorporating land use structure as a variable in population density models [89,90]. Consequently, this paper employs population density as a proxy variable to signify land use. Domestic pollution, exacerbated by population agglomeration, accounts for a considerable share of air pollution. Economic growth significantly influences both the total population and population mobility, with population agglomeration, urbanization, and economic development correlating with $PM_{2.5}$ emissions [91,92]. Rapid population agglomeration, especially over short durations, can precipitate a marked increase in pollution emissions from domestic sources, surpassing the atmospheric environment's assimilative capacity and exacerbating environmental pollution [93]. Further research indicates a nonlinear relationship between population density and domestic source pollution emissions; an exceedingly high population density exacerbates air pollution [94]. Hypothesis three posits that big data development can diminish urban population density, thereby reducing $PM_{2.5}$ concentrations [69,95]. Based on these considerations, this paper proposes hypothesis four:

H4. *Big data development streamlines land use planning, thereby mitigating $PM_{2.5}$ concentrations.*

3. Methodology

3.1. Model Construction

This study delves into the influence of big data development on $PM_{2.5}$ concentrations, noting that existing research has predominantly relied on traditional econometric models for causal inferences. These models, including difference-in-difference (DID) and complete control approaches, present considerable limitations. DID models require extensive data sets, while complete control models struggle to accommodate control groups with extreme values. A significant shortcoming of traditional econometric models lies in their need for a predefined model form, rendering them inadequate for capturing nonlinear relationships and misaligned with the complexity of real-world scenarios. Additionally, these models are susceptible to issues, such as endogeneity, multicollinearity, and limited covariate adjustment, leading to biased estimates.

The integration of machine learning techniques with econometric analysis has emerged as a vital research avenue. The double machine learning (DML) framework, introduced by Chernozhukov et al. (2018) [96], addresses the limitations of traditional econometric models by not presupposing the model's structure a priori. It effectively tackles challenges like nonlinear pattern recognition and endogeneity. The current scholarship on double machine learning bifurcates into theoretical explorations and practical applications. For example, Farbmacher et al. (2022) [97] combined double machine learning with a mediation effects analysis, broadening its application from straightforward causal inferences to exploring mediation mechanisms. Yang et al. (2020) [98] utilized double machine learning for a causal analysis, examining the average treatment effect of audit firms. While domestic (i.e., within-country) research on this topic is nascent, primarily concentrating on causal inference applications via double machine learning, studies such as Zhang et al. (2023) [99] have assessed the impact of network infrastructure on inclusive green growth in cities using a double machine learning model, further probing the mechanisms at play through mediation analysis. Wang et al. (2022) [100] developed a double machine learning model to assess the effects of default risk on the bond market, indicating ample scope for theoretical advancements in causal inference within double machine learning frameworks.

Accordingly, this paper adopts double machine learning models to assess the impact of big data development on $PM_{2.5}$, utilizing the partial nonlinear model framework suggested by Chernozhukov et al. (2018) [96], and constructs the model as follows:

$$lnPM_{2.5} = \theta_0 Event + g_0(X) + \xi, E(\xi|Event, X) = 0 \quad (1)$$

$$Event = m_0(X) + \psi, E(\psi|X) = 0 \quad (2)$$

$LnPM_{2.5}$ represents the logarithmic value of $PM_{2.5}$. The variable "*Event*" is an explanatory variable, specifically a policy dummy variable that indicates whether the region is a national big data comprehensive experimental area. Additionally, 0 indicates that the region is not a comprehensive national big data experimental area and 1 indicates that the region is a comprehensive national big data experimental area. $X = (x_1, x_2, \ldots, x_p)$ is a vector of high-dimensional control variables, including per capita GDP, infrastructure development, etc. The form of $\hat{g}_0(X)$ is unknown, but the machine learning model is utilized to estimate the specific form of $\hat{g}_0(X)$. The perturbation terms ξ and ψ have a conditional mean value of 0. This paper focuses on θ_0, which represents the effect of big data development on $PM_{2.5}$. The degree of influence can be obtained through Equations (1) and (2), corresponding to the estimator $\hat{\theta}_0$. The specific derivation process is not described in detail here, but can be found in Chernozhukov et al. (2018) [96]. The final unbiased estimator can be obtained by regressing Equations (1) and (2) twice using machine learning models, such as random forests. This approach ensures objectivity and precision in the estimation process.

$$\hat{\theta}_0 = (\frac{1}{n}\sum_{i\in I}\hat{\psi}_i \text{Event}_i)\frac{1}{n}\sum_{i\in I}\hat{\psi}_i[Y_i - g_0(x)_i] \quad (3)$$

In the formula, i represents the ith observation, I represents the overall observation, and n represents the sample size.

Chernozhukov et al. (2018) [96] provided a theoretical analysis of the unbiased estimator $\hat{\theta}_0$, which reveals that its estimation error can be divided into three parts via the following equations:

$$\sqrt{n}\left(\theta_0 - \hat{\theta}_0\right) = a^* + b^* + c^* \quad (4)$$

$$a^* = \left(E\left[\psi^2\right]\right)^{-1}\frac{1}{\sqrt{n}}\sum_i \psi_i\xi_i \sim N(0, \Sigma) \quad (5)$$

$$b^* = \left(E\left[\psi^2\right]\right)^{-1}\frac{1}{\sqrt{n}}\sum_i \left(\hat{m}_0(X_i) - m_0(X_i)\right)\left(\hat{g}_0(X_i) - g_0(X_i)\right) \quad (6)$$

It can be concluded that the estimator $\hat{\theta}_0$ has a slower rate of convergence than $\frac{1}{\sqrt{n}}$.

$$\left|\sqrt{n}\left(\hat{\theta}_0 - \theta_0\right)\right| \xrightarrow{p} 0 \quad (7)$$

The effect of big data development on $PM_{2.5}$ varies across regions and economic environments. Based on this heterogeneity, we refer to the more general interaction model developed by Chernozhukov et al. (2018) [96] as follows:

$$lnPM_{2.5} = g_0(Event, X) + \xi, E(\xi|Event, X) = 0 \quad (8)$$

$$Event = m_0(X) + \psi, E(\psi|X) = 0 \quad (9)$$

The model yields treatment effects: $\theta_0 = E[g(Event = 1, X) - g(Event = 0, X)]$.

3.2. Variable Selection and Data Sources

3.2.1. Outcome Variables

The study's primary explanatory variable is the average annual urban concentration of $PM_{2.5}$, which was log-transformed to address the skewed distribution of the data. The data were obtained from the Washington University in St. Louis website (https://sites.wustl.edu/acag/datasets/surface-pm2-5/#V5.GL.02. accessed on 1 November 2023).

3.2.2. Explanatory Variable

The principal explanatory variable in this study is the designation of cities as National Big Data Comprehensive Pilot Zone. This classification serves to examine the impact of big data development policies on $PM_{2.5}$ concentrations.

3.2.3. Control Variables

To ensure the precision and reliability of this analysis, the study incorporates a range of control variables. These include the logarithmic form of GDP per capita, serving as an indicator of the potential influence of economic development on environmental quality. The degree of urban infrastructure development is gauged through the proportion of fixed asset investment in GDP, directly correlating with environmental quality. The financial robustness of local governments, indicated by the ratio of local fiscal expenditure in the general budget to GDP, potentially impacts their capacity and efficiency in environmental conservation efforts. Financial development is assessed by examining the ratio of loans from financial institutions to GDP, offering insights into a city's level of financial development and its association with environmental governance capabilities.

Human capital measurement is based on the ratio of college students to the population at the end of a given period, reflecting the city's educational standing. This factor could influence residents' environmental awareness and the effective implementation of policies. The openness of a city to the global market, measured using the ratio of foreign direct investment to GDP, serves as an indicator of economic openness. The study also accounts for the impact of climatic factors on environmental quality using the logarithmic form of the average annual precipitation and the logarithmic form of the average annual temperature, acknowledging their direct effects on environmental quality.

These control variables are critical for mitigating the influence of the omitted variables and other confounding factors. The economic data for this research were primarily derived from the China urban statistical yearbook. The temperature data were sourced from the global summary of the day station reports, and precipitation data were obtained from the National Meteorological Science Data Sharing Service Platform-China Surface Climate Data Daily Value Dataset (V3.0), ensuring a comprehensive and robust dataset for analysis.

4. Results

4.1. Baseline Results

Table 1 presents the benchmark regression results of this study. Model (1) employs the lasso algorithm, which is particularly adept at handling high-dimensional data, characteristic of this research. This algorithm mitigates the risk of overfitting by imposing penalties on regression coefficients. The analysis reveals that the regression coefficient for the National Big Data Comprehensive Pilot Zone policy on $PM_{2.5}$ concentrations is -0.0149, significant at the 1% level, indicating a substantial negative correlation between big data initiatives and urban $PM_{2.5}$ levels. To validate the robustness of these findings, the study further explores the use of random forests, regression trees, and xgboost algorithms, resulting in Models (2–4), respectively. These algorithms excel in managing nonlinear relationships and complex interactions, offering a nuanced perspective for data analysis. The results consistently demonstrate that, across the different methodologies, the impact of the National Big Data Comprehensive Pilot Zone policy on $PM_{2.5}$ concentrations remains negative and statistically significant at the 1% level. This consistency reinforces the conclusion that big data development effectively reduces urban $PM_{2.5}$ concentrations. Our comparative analysis across diverse algorithms confirms that the National Big Data Comprehensive Pilot Zone policy exerts a consistent and significant negative impact on $PM_{2.5}$ levels across all models, substantiating hypothesis H1 that big data development beneficially influences urban air quality improvement. These findings suggest that further investment and application of big data technology in environmental monitoring and management could significantly enhance environmental quality.

Table 1. Regression to the baseline.

	Partial Linear Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0149 ***	−0.1018 ***	−0.1267 ***	−0.0340 ***
SE	0.0049	0.0185	0.0145	0.0127
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
	Interactive Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.1791 ***	−0.1140 ***	−0.1919 ***	−0.2516 ***
SE	0.0060	0.0181	0.0346	0.0147
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *** significant at 1%.

4.2. Mechanisms

The findings indicate that big data development can significantly decrease $PM_{2.5}$ concentrations in urban areas. However, there is a need for additional research to explore the mechanisms through which big data development influences $PM_{2.5}$ reduction and overall air pollution mitigation. This section delves into the mechanism by which the National Big Data Comprehensive Pilot Zone policy contributes to the reduction in $PM_{2.5}$ levels, employing a double machine learning model. The analysis reveals that the total effects across various mediating paths are statistically significant at the 1% level.

Big data development plays a critical role in fundamentally enhancing resource allocation (as illustrated in Tables 2 and 3). By leveraging big data, the efficiency of resource flow and utilization within the urban economy can be improved, directing capital and labor towards industries and sectors where they are most effectively utilized. This optimization minimizes environmental costs associated with resource mismatches, leading to more efficient and environmentally sustainable production processes. Thus, hypothesis one is confirmed.

Thus, big data development aids in diminishing $PM_{2.5}$ concentrations through fostering green technology innovation, as illustrated in Tables 4 and 5. The advancement of big data has significantly propelled the innovation of green and low-carbon technologies. An increase in the number of patents, encompassing both utility model and invention patents, serves as a robust indicator of innovation activities, highlighting big data's pivotal role in advancing environmental technologies. Patents in green low-carbon technologies, including utility model patents and inventions, play a crucial role in reducing $PM_{2.5}$ levels. These innovations enhance energy efficiency and encourage the adoption of novel environmental technologies in both production and everyday life, thereby lowering $PM_{2.5}$ emissions.

Table 2. High-quality urban development: capital mismatch.

Partial Linear Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0149 ***	−0.1020 ***	−0.1267 ***	−0.0332 ***
SE	0.0049	0.0185	0.0145	0.0129
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
Interactive Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.1787 ***	−0.1152 ***	−0.1919 ***	−0.2520 ***
SE	0.0060	0.0181	0.0346	0.0148
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *** significant at 1%.

Table 3. High-quality urban development: labor mismatch.

Partial Linear Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0149 ***	−0.1040 ***	−0.1267 ***	−0.0310 **
SE	0.0049	0.0185	0.0145	0.0128
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
Interactive Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.1780 ***	−0.1158 ***	−0.1919 ***	−0.2486 ***
SE	0.0060	0.0181	0.0346	0.0147
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; **, ***significant at 5%, and 1%, respectively.

From the perspective of industrial structure upgrading, as detailed in Table 6, big data development expedites the rationalization and advancement of industrial structures towards a more sustainable and environmentally friendly orientation. The incorporation of big data enables traditional industries to refine production processes, heighten resource efficiency, and undertake industrial modernization, leading to decreased energy use and emissions. Furthermore, the emergence of big data fosters the growth of new green industries and supports the shift towards sustainable economic models, thereby improving industrial competitiveness and environmental efficiency. Such industrial restructuring alleviates environmental stress and contributes to the reduction in $PM_{2.5}$ levels.

Table 4. High-quality urban development: number of utility model patents in the city.

Partial Linear Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0180 ***	−0.1053 ***	−0.1267 ***	−0.0392 ***
SE	0.0045	0.0186	0.0145	0.0134
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
Interactive Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.2133 ***	−0.1175 ***	−0.2332 ***	−0.2504 ***
SE	0.0059	0.0179	0.0234	0.0149
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *** significant at 1%.

Table 5. High-quality urban development: number of patents for municipal inventions.

Partial Linear Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0170 ***	−0.1053 ***	−0.1267 ***	−0.0386 ***
SE	0.0050	0.0186	0.0145	0.0131
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
Interactive Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.1994 ***	−0.1171 ***	−0.1919 ***	−0.2481 ***
SE	0.0060	0.0180	0.0346	0.0147
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *** significant at 1%.

Table 6. High-quality urban development: Mediating factor: industrial structure.

Partial Linear Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0273 **	−0.0934 ***	−0.1254 ***	−0.0480 ***
SE	0.0112	0.0186	0.0145	0.0127
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
Interactive Model				
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0583 *	−0.1140 ***	−0.2195 ***	−0.2502 ***
SE	0.0313	0.0178	0.0226	0.0149
Control Variables	Control	Control	Control	Control

Table 6. *Cont.*

	Partial Linear Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *, **, *** significant at 10%, 5%, and 1%, respectively.

Regarding land use, as presented in Table 7, big data development encourages rational population distribution and optimal land use planning. Industries related to big data generate a pull effect on population movements, influencing shifts in population density and leading to more sustainable land use patterns. This adjustment mitigates environmental pressures associated with overpopulation.

Table 7. Land use planning.

	Partial Linear Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0445 ***	−0.1065 ***	−0.1526 ***	−0.0526 ***
SE	0.0105	0.0183	0.0134	0.0104
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
	Interactive Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.1043 ***	−0.1193 ***	−0.2381 ***	−0.2464 ***
SE	0.0230	0.0177	0.0218	0.0139
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *** significant at 1%.

4.3. Heterogeneous Analysis

To facilitate the formulation of precise and effective regional environmental policies, our study undertakes a heterogeneity analysis. This approach enables us to accurately comprehend and quantify the distinct impact of big data development on environmental governance.

4.3.1. Central, East, and West Regions

Building on the aforementioned study, we conducted a heterogeneity analysis to examine the differential impacts of big data development on $PM_{2.5}$ concentrations across China's central, eastern, and western regions. Given the considerable regional disparities in environmental quality, industrial structure, and technological application levels among these areas, the efficacy of big data policies in environmental governance is likely to vary. Acknowledging these regional differences is crucial for the effective implementation of such policies. Thus, this research seeks to elucidate the heterogeneity in the effects of big data development on $PM_{2.5}$ concentrations across different regions through a causal inference methodology employing double machine learning. According to the findings presented in Table 8, big data development exerts a significant suppressive influence on $PM_{2.5}$ concentrations in the central, eastern, and western regions in certain linear models. In the interaction models, significant results were observed in the central and eastern regions, whereas the western region's results were not statistically significant, albeit indicating a negative impact. This discrepancy may be attributed to the more advanced implementation of big data technologies and environmental regulatory mechanisms in the eastern and central regions, which effectively contributes to air quality improvement. Conversely, despite the western region's lower economic development level and relatively delayed infrastructural development, the findings suggest a potential negative but non-significant impact on $PM_{2.5}$ concentrations. Overall, big data development positively influences air quality by enhancing environmental monitoring and management capabilities. Furthermore, this heterogeneity analysis underscores the importance of considering regional variations when developing and implementing future big data initiatives. For instance, infrastructural development in the central and western regions may require additional focus, whereas the eastern region could benefit from further leveraging big data in environmental monitoring and management.

Table 8. Heterogeneous analysis: central, east, and west regions.

	Partial Linear Model			Interactive Model		
	Central	**Eastern**	**Western**	**Central**	**Eastern**	**Western**
Digital	−0.0982 **	−0.0841 **	−0.1803 ***	−0.1418 ***	−0.0735 ***	−0.0259
SE	0.0403	0.0373	0.0336	0.0121	0.0114	0.0222
Control Variables	Control	Control	Control	Control	Control	Control
Year FE	Control	Control	Control	Control	Control	Control
City FE	Control	Control	Control	Control	Control	Control
Observations	1089	1067	891	1089	1067	891

Note: 'SE' indicates standard error; **, *** significant at 5%, and 1%, respectively.

4.3.2. Resource-Based and Non-Resource-Based Cities

This study posits that the structural characteristics of urban economies, notably the variance in resource dependence, can markedly influence the effectiveness of big data initiatives in environmental governance. It delves into the disparities between "resource cities" and "non-resource cities" regarding the impact of big data development on $PM_{2.5}$ concentrations. Cities reliant on specific natural resources encounter considerable environmental challenges and necessitate structural transformation. Employing the analytical framework of double machine learning for causal inferences, this section investigates and contrasts the influence of big data development on $PM_{2.5}$ levels in resource-based cities. The regression outcomes, depicted in the Table 9, reveal that big data development substantially reduces $PM_{2.5}$ concentrations in both resource-based and non-resource-based cities, indicating big data technology's versatile utility in enhancing air quality. However, in resource-based cities, given their unique challenges, the deployment of big data may assume a more pivotal role in driving environmental improvements and economic transition. In the formulation and implementation of environmental policies related to big data, the structural dynamics of urban economies ought to be thoroughly accounted for. For resource-based cities, strategies should focus on leveraging big data technologies to refine industrial structures, augment resource efficiency, and mitigate environmental pollutants. Conversely, in cities not characterized by resource dependence, greater emphasis could be placed on advancing the sophistication and accuracy of urban governance through big data utilization.

Table 9. Heterogeneous analysis: resource-based and non-resource-based cities.

	Partial Linear Model		Interactive Model	
	Resource-Based City	**Non-Resource-Based City**	**Resource-Based City**	**Non-Resource-Based City**
Digital	−0.1803 ***	−0.0879 ***	−0.1251 ***	−0.1117 ***
SE	0.0336	0.0232	0.0100	0.0076
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	1210	1837	1210	1837

Note: 'SE' indicates standard error; *** significant at 1%.

5. Robustness Check

5.1. Adding the Variable "Broadband China Pilot" to the Baseline Regression

In the robustness analysis, we incorporated additional policy influences that might alter the study's outcomes, notably the "Broadband China" policy. This initiative is designed to enhance the national internet infrastructure, potentially influencing big data's application and evolution indirectly. To mitigate any conceivable effects of this policy on our results, the "Broadband China" policy was integrated into our model to assess the steadfastness of big data development's impact on $PM_{2.5}$ levels. The outcomes, displayed in Table 10, affirm that the suppressive influence of big data development on $PM_{2.5}$ concentrations persists as robust and statistically significant at the 1% level, even when accounting for the "Broadband China" policy's effect. These findings reinforce our core assertion that big data implementation plays a crucial role in diminishing urban $PM_{2.5}$ levels. The impact remains not only statistically significant but also steadfast upon the inclusion of other pivotal policy considerations. The robustness analysis indicates that although initiatives like "Broadband China" may affect environmental quality and technological advancements in urban settings, the contribution of big data policies to environmental preservation stands as independent and substantial.

Table 10. Robustness check A.

	Partial Linear Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0149 ***	−0.1016 ***	−0.1267 ***	−0.0335 ***
SE	0.0049	0.0185	0.0145	0.0127
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
	Interactive Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.1824 ***	−0.1155 ***	−0.1919 ***	−0.2516 ***
SE	0.0060	0.0181	0.0346	0.0147
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *** significant at 1%.

5.2. Inclusion of the Variable "Province-Time Cross-Multiplier" in the Baseline Regression

As part of the robustness analysis, this study further examines potential similarities and temporal dynamics in policy implementation, economic progression, and environmental governance across different provinces. To precisely assess the influence of big data policies on $PM_{2.5}$ levels and mitigate potential biases stemming from provincial characteristics and time variations, a "province-time cross-multiplier" variable is integrated into the analytical model. This methodological approach is designed to account for varying provincial effects over time, thereby enabling a more accurate delineation of big data policy's impact on environmental quality. The findings, detailed in Table 11, reveal that the inclusion of the "province-time cross-multiplier" term does not alter the significant suppressive influence of big data policies on $PM_{2.5}$ concentrations, which remains robust at the 1% significance level. This analysis underscores that the beneficial effect of big data development on air quality enhancement is both significant and stable, even when controlling for potential province-specific and temporal-specific influences. Consequently, this reinforces the preliminary conclusion that big data policies exert a distinct and meaningful reduction on $PM_{2.5}$ levels, unaffected by provincial and temporal variations.

Table 11. Robustness check B.

	Partial Linear Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.0150 ***	−0.1198 ***	−0.1267 ***	−0.0369 ***
SE	0.0046	0.0192	0.0145	0.0128
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047
	Interactive Model			
	Lasso	**Random Forest**	**Regression Tree**	**Xgboost**
Digital	−0.1401 ***	−0.1189 ***	−0.1919 ***	−0.2473 ***
SE	0.0057	0.0192	0.0346	0.0146
Control Variables	Control	Control	Control	Control
Year FE	Control	Control	Control	Control
City FE	Control	Control	Control	Control
Observations	3047	3047	3047	3047

Note: 'SE' indicates standard error; *** significant at 1%.

5.3. Endogeneity

To address the issue of endogeneity, this study employs an instrumental variable, specifically the interaction term between the number of internet users in the country over the past year and the number of landline telephones per 10,000 people in each city in 1984 [101,102]. The outcomes, displayed in Table 12, affirm that the development of big data maintains a significant suppressive influence on $PM_{2.5}$ concentrations, even after the application of instrumental variables. This effect is statistically significant, indicating that the study's findings withstand the robustness test.

Table 12. Robustness check C.

	Lasso	**Random Forest**	**Regression Tree**
Digital	−0.4231 **	−0.5345 ***	−0.4968 ***
SE	0.1720	0.1054	0.1356
Control Variables	Control	Control	Control
Year FE	Control	Control	Control
City FE	Control	Control	Control
Observations	2387	2387	2387

Note: 'SE' indicates standard error; **, *** significant at 5%, and 1%, respectively.

6. Conclusions and Policy Implications

The advancement of big data is pivotal in enhancing air quality, facilitating high-quality urban development, and optimizing land use planning. This study introduces a double machine learning model to ascertain the impact of establishing National Big Data Comprehensive Pilot Zone on urban $PM_{2.5}$ concentrations, utilizing panel data from 277 prefecture-level cities across China. The investigation into the establishment of these pilot zones across various Chinese regions indicates a marked negative correlation between big data development and urban $PM_{2.5}$ levels, a finding that withstands rigorous robustness checks. This evidence suggests that the deployment of big data policies contributes positively to air quality improvement. The research delineates multiple channels through which big data development influences $PM_{2.5}$ concentrations, including the enhancement of urban development quality and the refinement of land use planning. These mechanisms collectively contribute to the reduction in $PM_{2.5}$ levels. Notably, the effect of big data development on $PM_{2.5}$ concentrations exhibits significant regional variation. For example, in the economically advanced eastern region, the link between big data policies and reduced $PM_{2.5}$ concentrations is particularly pronounced. The heterogeneity analysis further demonstrates that big data development exerts a substantial suppressive effect on $PM_{2.5}$ levels across the central, eastern, and western regions. Moreover, big data initiatives significantly curb $PM_{2.5}$ concentrations in both resource-intensive and non-resource-intensive cities, albeit with varying degrees of impact across different urban contexts.

Arising from the research discussed herein, this paper delineates several policy recommendations: Firstly, the endorsement and support for the adoption of big data technology are advocated. Such measures are poised to augment the efficiency and efficacy of environmental governance, particularly within the realms of environmental monitoring and management. Additionally, this document advocates for an escalation in the deployment of digital economy policies, reflecting the beneficial influence of big data advancements on mitigating urban $PM_{2.5}$ levels. Secondly, the crafting of big data policies must account for regional disparities. This study highlights the variable impact of big data advancements across different regions on $PM_{2.5}$ concentrations, with a notable emphasis on the pronounced effects within the economically advanced eastern region. Thus, the formulation and execution of big data strategies should thoroughly incorporate considerations of these regional distinctions. For example, enhanced infrastructural development may be necessitated in the central and western regions, whereas the eastern regions could benefit from a deeper integration of big data into environmental monitoring and management practices. Thirdly, it is crucial to acknowledge the nuances of urban economic structures and resource allocations. The development of big data exerts a notable suppressive influence on $PM_{2.5}$ levels across both resource-centric and non-resource-centric cities, albeit with varying degrees of significance contingent upon the city's specific resource characteristics. Consequently, when developing and applying environmental policies associated with big data, an in-depth understanding of the city's economic framework and resource assets is imperative. In cities dependent on natural resources, the focus should be on leveraging big data technologies to refine industrial configurations and enhance resource utilization efficiency, thereby curtailing environmental pollutants. Conversely, in cities not predicated on natural resources, the emphasis should be placed on advancing the sophistication and accuracy of urban governance to bolster environmental management efforts.

Author Contributions: Formal analysis, X.K.; Writing—original draft, X.W.; Writing—review & editing, K.D.; Supervision, M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Key Program of the National Social Science Foundation of China under Grant number 22AZD048.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: We would like to thank the editor and the anonymous reviewers for their helpful suggestions and comments.

Conflicts of Interest: The authors declare that they have no conflicts of interests.

References

1. Guo, Q.; Ma, X. How Does the Digital Economy Affect Sustainable Urban Development? Empirical Evidence from Chinese Cities. *Sustainability* **2023**, *15*, 4098. [CrossRef]
2. Ma, X.; Feng, X.; Fu, D.; Tong, J.; Ji, M. How does the digital economy impact sustainable development?—An empirical study from China. *J. Clean. Prod.* **2024**, *434*, 140079. [CrossRef]
3. Guo, R.; Zhang, M. Research on the Impact of Industrial Digitalization on High-Quality Economic Development: Evidence from China. *Res. Econ. Manag.* **2023**, *8*, 110. [CrossRef]
4. Guo, B.; Wang, Y.; Zhang, H.; Liang, C.; Feng, Y.; Hu, F. Impact of the digital economy on high-quality urban economic development: Evidence from Chinese cities. *Econ. Model.* **2023**, *120*, 106194. [CrossRef]
5. Hong, M.; Tian, M.; Wang, J. The impact of digital economy on green development of agriculture and its spatial spillover effect. *China Agric. Econ. Rev.* **2023**, *15*, 708–726. [CrossRef]
6. Yao, W.; Sun, Z. The Impact of the Digital Economy on High-Quality Development of Agriculture: A China Case Study. *Sustainability* **2023**, *15*, 5745. [CrossRef]

7. Xu, W. Research on The Impact of Digital Economy on Common Prosperity. Master's Thesis, Chongqing Technology and Business University, Chongqing, China, 2023.
8. Xu, C. Study on the Impact of Digital Economy Development on Export Trade Competitiveness. Master's Thesis, Jiangxi University of Finance and Economics, Nanchang, China, 2023.
9. Cao, Z. Research on the Impact of Digital Economy on Environmental Quality. Ph.D. Thesis, Anhui University, Hefei, China, 2022.
10. Ma, S. A Study on the Impact of the Digital Economy on Labor Employment. Master's Thesis, Sichuan Normal University, Chengdu, China, 2023.
11. Feng, S.; Gao, D.; Liao, F.; Zhou, F.; Wang, X. The health effects of ambient PM 2.5 and potential mechanisms. *Ecotoxicol. Environ. Saf.* **2016**, *128*, 67–74. [CrossRef]
12. Shi, S.; Wang, W.; Li, X.; Xu, C.; Lei, J.; Jiang, Y.; Zhang, L.; He, C.; Xue, T.; Chen, R.; et al. Evolution in disparity of $PM_{2.5}$ pollution in China. *Eco-Environ. Health* **2023**, *2*, 257–263. [CrossRef]
13. Ali, M.A.; Huang, Z.; Bilal, M.; Assiri, M.E.; Mhawish, A.; Nichol, J.E.; de Leeuw, G.; Almazroui, M.; Wang, Y.; Alsubhi, Y. Long-term $PM_{2.5}$ pollution over China: Identification of $PM_{2.5}$ pollution hotspots and source contributions. *Sci. Total Environ.* **2023**, *893*, 164871. [CrossRef] [PubMed]
14. Chen, X.; Yuan, F.; Wang, B.; Wang, Y.; Li, K.; Li, X. Analysis on the Causes of Atmospheric $PM_{2.5}$ Pollution in Anyang City in 2017. *J. Green Sci. Technol.* **2021**, *23*, 122–125. [CrossRef]
15. Cheng, M.; Tang, G.; Lv, B.; Li, X.; Wu, X.; Wang, Y.; Wang, Y. Source apportionment of $PM_{2.5}$ and visibility in Jinan, China. *J. Environ. Sci.* **2021**, *102*, 207–215. [CrossRef]
16. Wu, W.; Chang, X.; Xing, J.; Wang, S.; Hao, J. Assessment of $PM_{2.5}$ Pollution Mitigation due to Emission Reduction from Main Emission Sources in the Beijing-Tianjin-Hebei Region. *Environ. Sci.* **2017**, *38*, 867–875. [CrossRef]
17. Li, W.; Yang, G.; Qian, X. The socioeconomic factors influencing the $PM_{2.5}$ levels of 160 cities in China. *Sustain. Cities Soc.* **2022**, *84*, 104023. [CrossRef]
18. Chen, X. Research into Impact of $PM_{2.5}$ Air Pollution on Household Decisions on Migrations in China. *J. Nanjing Inst. Technol. (Soc. Sci. Ed.)* **2022**, *22*, 53–63. [CrossRef]
19. Zheng, H.; Wang, Z. Analysis of Air Pollution Influencing Factors of PM Secondary Particles by Random Forest. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *804*, 042065. [CrossRef]
20. Sun, W.; Huo, J.; Fu, Q.; Zhang, Y.; Lin, X. Influence of Meteorological Factors and Chemical Processes on the Explosive Growth of $PM_{2.5}$ in Shanghai, China. *Atmosphere* **2022**, *13*, 1068. [CrossRef]
21. Yan, D.; Kong, Y.; Jiang, P.; Huang, R.; Ye, B. How do socioeconomic factors influence urban $PM_{2.5}$ pollution in China? Empirical analysis from the perspective of spatiotemporal disequilibrium. *Sci. Total Environ.* **2020**, *761*, 143266. [CrossRef]
22. Li, Y. Research on the Impact of Regional Economic Differences on Air Pollution. *J. Green Sci. Technol.* **2023**, *25*, 77–82. [CrossRef]
23. Siyu, P. Health, air pollution, and location choice. *J. Environ. Econ. Manag.* **2023**, *119*, 102794.
24. Bu, H.; Li, G.; Yu, X.; Xun, Z. Is smart carbon emission reduction justified in China? Evidence from national big data comprehensive pilot zones. *Heliyon* **2023**, *9*, e17744. [CrossRef]
25. Lyu, Y.; Xiao, X.; Zhang, J. Does the digital economy enhance green total factor productivity in China? The evidence from a national big data comprehensive pilot zone. *Struct. Chang. Econ. Dyn.* **2024**, *69*, 183–196. [CrossRef]
26. Jin, J. The Current State of China's Digital Information Industry and Technology Development. *Digit. Technol. Appl.* **2023**, *41*, 48–50. [CrossRef]
27. Fang, X.; Gao, B.; Cui, S.; Ding, L.; Wang, L.; Shen, Y. Regional Differences in PM 2.5 Environmental Efficiency and Its Driving Mechanism in Zhejiang Province, China. *Atmosphere* **2023**, *14*, 672. [CrossRef]
28. Li, Z.; Li, N.; Wen, H. Digital Economy and Environmental Quality: Evidence from 217 Cities in China. *Sustainability* **2021**, *13*, 8058. [CrossRef]
29. Qi, G.; Wang, Z.; Wang, Z.; Wei, L. Has Industrial Upgrading Improved Air Pollution?—Evidence from China's Digital Economy. *Sustainability* **2022**, *14*, 8967. [CrossRef]
30. Sun, X.; Chen, Z.; Loh, L. Exploring the Effect of Digital Economy on $PM_{2.5}$ Pollution: The Role of Technological Innovation in China. *Front. Environ. Sci.* **2022**, *10*, 904254. [CrossRef]
31. Wu, X.; Qin, Y.; Xie, Q.; Zhang, Y. The Mediating and Moderating Effects of the Digital Economy on $PM_{2.5}$: Evidence from China. *Sustainability* **2022**, *14*, 16032. [CrossRef]
32. Zhao, H.; Cheng, Y.; Zheng, R. Impact of the Digital Economy on $PM_{2.5}$: Experience from the Middle and Lower Reaches of the Yellow River Basin. *Int. J. Environ. Res. Public Health* **2022**, *19*, 17094. [CrossRef] [PubMed]
33. Zhang, X.; Zhong, J.; Wang, H. Does the Development of Digital Economy Affect Environmental Pollution? *Sustainability* **2023**, *15*, 9162. [CrossRef]
34. Li, Y. Research on the Impact of Big Data Development on Urban Green Total Factor Productivity: Quasi Natural Experiment Based on National Big Data Comprehensive Pilot Zone. Master's Thesis, Dongbei University of Finance and Economics, Dalian, China, 2022.
35. Guo, B.; Wang, Y.; Zhang, H. Dose Digital Economy Improve Urban Air Quality: Quasi Natural Experiment Based on National Big Data Comprehensive Pilot Zone. *J. Guangdong Univ. Financ. Econ.* **2022**, *37*, 58–74.
36. Sun, W.; Mao, N.; Lan, F.; Wang, L. Policy Empowerment, Digital Ecosystem and Enterprise Digital Transformation: A Quasi Natural Experiment Based on The National Big Data Comprehensive Zone. *China Ind. Econ.* **2023**, *9*, 117–135. [CrossRef]

37. Chuai, X.; Feng, J. High resolution carbon emissions simulation and spatial heterogeneity analysis based on big data in Nanjing City, China. *Sci. Total Environ.* **2019**, *686*, 828–837. [CrossRef] [PubMed]
38. Hien, P.D.; Men, N.T.; Tan, P.M.; Hangartner, M. Impact of urban expansion on the air pollution landscape: A case study of Hanoi, Vietnam. *Sci. Total Environ.* **2020**, *702*, 134635. [CrossRef] [PubMed]
39. Huang, B.; Wang, J. Big spatial data for urban and environmental sustainability. *Geo-Spat. Inf. Sci.* **2020**, *23*, 125–140. [CrossRef]
40. Wang, W.; Yang, S. Research on Air Quality Forecasting Based on Big Data and Neural Network. In Proceedings of the 2020 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, China, 25–27 September 2020; pp. 180–184.
41. Wu, R.; Wang, J.; Zhang, D.; Wang, S. Identifying different types of urban land use dynamics using Point-of-interest (POI) and Random Forest algorithm: The case of Huizhou, China. *Cities* **2021**, *114*, 103202. [CrossRef]
42. Guo, X.; Chen, H.; Yang, X. An Evaluation of Street Dynamic Vitality and Its Influential Factors Based on Multi-Source Big Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 143. [CrossRef]
43. Zhao, Y.; Kong, X.; Ahmad, M.; Ahmed, Z. Digital Economy, Industrial Structure, and Environmental Quality: Assessing the Roles of Educational Investment, Green Innovation, and Economic Globalization. *Sustainability* **2023**, *15*, 2377. [CrossRef]
44. Ren, Y.; Liu, Y.; Hu, Z.; Li, H. Impact of big data development and intellectual property protection on corporate green technology innovation. *China Popul. Resour. Environ.* **2023**, *33*, 157–167.
45. Dan, M.; Qing, Z. Innovation in emerging economies: Research on the digital economy driving high-quality green development. *J. Bus. Res.* **2022**, *145*, 801–813.
46. Zhang, J.; Qian, F. Digital economy enables common prosperity: Analysis of mediating and moderating effects based on green finance and environmental pollution. *Front. Energy Res.* **2023**, *10*, 1080230. [CrossRef]
47. Xu, Q.; Zhong, M.; Cao, M. Does digital investment affect carbon efficiency? Spatial effect and mechanism discussion. *Sci. Total Environ.* **2022**, *827*, 154321. [CrossRef]
48. Wang, X.; Sun, X.; Zhang, H.; Ahmad, M. Digital Economy and Environmental Quality: Insights from the Spatial Durbin Model. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16094. [CrossRef] [PubMed]
49. Hu, J.; Zhao, X.; Wu, D.; Wu, W. Digital economy and environmental governance performance: Empirical evidence from 275 cities in China. *Environ. Sci. Pollut. Res. Int.* **2022**, *30*, 26012–26031. [CrossRef] [PubMed]
50. Bai, F.; Huang, Y.; Shang, M.; Ahmad, M. Modeling the impact of digital economy on urban environmental pollution: Empirical evidence from 277 prefecture-level cities in China. *Front. Environ. Sci.* **2022**, *10*, 991022. [CrossRef]
51. Ma, F.; Fahad, S.; Wang, M.; Nassani, A.A.; Haffar, M. Spatial Effects of Digital Transformation, $PM_{2.5}$ Exposure, Economic Growth and Technological Innovation Nexus: $PM_{2.5}$ Concentrations in China during 2010–2020. *Int. J. Environ. Res. Public Health* **2023**, *20*, 2550. [CrossRef] [PubMed]
52. Liu, T.; Liu, H.; Qi, Y. Construction land expansion and cultivated land protection in urbanizing China: Insights from national land surveys, 1996–2006. *Habitat Int.* **2015**, *46*, 13–22. [CrossRef]
53. Chuai, X.; Huang, X.; Wang, W.; Zhao, R.; Zhang, M.; Wu, C. Land use, total carbon emissions change and low carbon land management in Coastal Jiangsu, China. *J. Clean. Prod.* **2015**, *103*, 77–86. [CrossRef]
54. Tang, Y.; Chen, M. The Impact Mechanism and Spillover Effect of Digital Rural Construction on the Efficiency of Green Transformation for Cultivated Land Use in China. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16159. [CrossRef]
55. Li, J.; Sun, Z.; Zhou, J.; Sow, Y.; Cui, X.; Chen, H.; Shen, Q. The Impact of the Digital Economy on Carbon Emissions from Cultivated Land Use. *Land* **2023**, *12*, 665. [CrossRef]
56. Wu, W.; Feng, Y. Digital Economy, Consumption Upgrading and Urban Land Use Efficiency—Empirical Evidence Based on Wuhan City Circle. *Stat. Decis.* **2022**, *38*, 107–110. [CrossRef]
57. Zhang, C.; Zhang, F.; Wu, W. Digital Economy, Domestic General Circulation and Urban Land Use Efficiency—Empirical Evidence Based on Wuhan City Circle. *Stat. Decis.* **2023**, *39*, 163–167. [CrossRef]
58. Liu, Q.; Jiang, H.; Li, J.; Song, J.; Zhang, X. Antidote or poison? Digital economy and land-use. *Land Use Policy* **2024**, *139*, 107083. [CrossRef]
59. Chen, L.; Zhong, C.; Li, C. Research on the Impact of the Digital Economy on China's New-Type Urbanization: Based on Spatial and Mediation Models. *Sustainability* **2022**, *14*, 14843. [CrossRef]
60. Wang, J.; Wu, H.; Xiao, D.; Luo, X.; Liu, Y.; Wang, W. Can the digital economy address the loss of green development efficiency due to resource mismatch? Evidence from China's land transaction markets. *J. Environ. Plan. Manag.* 2023; *epub ahead of print*. [CrossRef]
61. Huang, H.; Li, Q. Opportunities, Data Sources, and Potential Applications of Land Use Optimization in the Big Data Era. *China Land Sci.* **2017**, *31*, 74–82.
62. Cai, Y. Exploring the innovation of land use planning development model based on big data. *South China Agric.* **2019**, *13*, 171–172. [CrossRef]
63. Pan, L. Analysis of Territorial Spatial Planning and Smart Cities in the Era of Big Data. *Art Design.* **2023**, *2*. [CrossRef]
64. Timo, D.V.W. Potential of big data for pro-active participatory land use planning. *Geoplanning: J. Geomat. Plan.* **2018**, *5*, 205.
65. Yao, T.; Huang, Z.; Zhao, W. Are smart cities more ecologically efficient? Evidence from China. *Sustain. Cities Soc.* **2020**, *60*, 102008. [CrossRef]

66. Wu, H.; Sun, M.; Zhang, W.; Guo, Y.; Irfan, M.; Lu, M.; Hao, Y. Can urbanization move ahead with energy conservation and emission reduction? New evidence from China. *Energy Environ.* **2022**. [CrossRef]
67. Bian, Y.; Song, K.; Bai, J. Market segmentation, resource misallocation and environmental pollution. *J. Clean. Prod.* **2019**, *228*, 376–387. [CrossRef]
68. Wang, S.; Li, M. Impact of spatial structure of urban agglomerations on $PM_{2.5}$ pollution: Based on resource misallocation. *Heliyon* **2023**, *9*, e14099. [CrossRef]
69. Ye, X. A Study of the Impact of the Digital Economy on the Willingness to Stay of the Mobile Population. Master's Thesis, Jiangxi University of Finance and Economics, Nanchang, China, 2023.
70. Zhong, J.; Wei, Y. Spatial effects of industrial agglomeration and open economy on pollution abatement. *China Popul. Resour. Environ.* **2019**, *29*, 98–107.
71. Zhang, Y.; Wang, J. Research on the Impact of Digital Economy on Resource Misallocation: Mechanism and Empirical Evidence. *Contemp. Econ. Manag.* **2023**, *45*, 32–45. [CrossRef]
72. Sun, Z.; Wang, L.; Zhang, X. Digitalization Enables Industrial Transformation and Upgrading: Opportunities, Challenges, and Paths to Realization. *J. Xi'an Jiaotong Univ. (Soc. Sci.)* **2023**, 1–23. [CrossRef]
73. Pang, J.; Zhang, Y.; Jiao, F. The Impact of the Digital Economy on Transformation and Upgrading of Industrial Structure: A Perspective Based on the "Poverty Trap". *Sustainability* **2023**, *15*, 15125. [CrossRef]
74. Xiong, M.; Zhang, F.; Zhang, H.; Wang, H. Digital economy, credit expansion, and modernization of industrial structure in China. *Financ. Res. Lett.* **2023**, *58*, 104500. [CrossRef]
75. Bai, E.; Yang, F. Research on regional heterogeneity of digital economy promoting the transformation and upgrading of China's industrial structure. *China Townsh. Enterp. Account.* **2023**, *12*, 147–149.
76. Hui, N.; Yu, Q.; Gu, Y. Does the Digital Economy Improve the Innovation Efficiency of the Manufacturing Industry? Evidence in Provincial Data from China. *Sustainability* **2023**, *15*, 10615. [CrossRef]
77. Wang, L.; Tian, Y.; Zhou, Z. Research on the Enabling Effect of the Digital Economy on the High-Quality Development of Gansu Province: A Perspective on Industrial Structure Upgrading. *J. Lanzhou Univ. Financ. Econ.* **2023**.
78. Bao, T. Digital Technology Empowers Dual Optimization of Manufacturing Structure: Efficiency Improvement and Green Transformation. *South China J. Econ.* **2023**, *12*, 83–106. [CrossRef]
79. Han, X.; Cao, T. Urbanization level, industrial structure adjustment and spatial effect of urban haze pollution: Evidence from China's Yangtze River Delta urban agglomeration. *Atmos. Pollut. Res.* **2022**, *13*, 101427. [CrossRef]
80. Zeng, H.; Tu, F.; Zhou, M.; Wu, W. A Study on the Influence of Optimiation of Urban Industrial Structure on Haze Pollution in the Yangtze River Economic Belt. *J. East China Univ. Technol. (Soc. Sci.)* **2023**, *42*, 44–52.
81. Li, T.; Xue, L.; Zhao, W. Influencing Mechanism of Industrial Structure Adjustment on Urban $PM_{2.5}$ in China. *Econ. Geogr.* **2023**, *43*, 180–192. [CrossRef]
82. Wang, Y. Spatio-temporal characteristics and optimal path of $PM_{2.5}$ pollution in Yangtze River Delta urban agglomeration. *Environ. Prot. Sci.* **2023**, *49*, 117–125. [CrossRef]
83. Tan, M.; Li, X.; Lu, C.; Luo, W.; Kong, X.; Ma, S. Urban population densities and their policy implications in China. *Habitat Int.* **2008**, *32*, 471–484. [CrossRef]
84. Akimoto, F. The birth of 'land use planning'in American urban planning. *Plan. Perspect.* **2009**, *24*, 457–483. [CrossRef]
85. Salvati, L.; Munafo, M.; Morelli, V.G.; Sabbi, A. Low-density settlements and land use changes in a Mediterranean urban region. *Landsc. Urban Plan.* **2012**, *105*, 43–52. [CrossRef]
86. Stankowski, S.J. Population density as an indirect indicator of urban and suburban land-surface modifications. *US Geol. Surv. Prof. Pap.* **1972**, *800*, 219–224.
87. Pfeffer, M.J.; Schlelhas, J.W.; DeGloria, S.D.; Gomez, J. Population, conservation, and land use change in Honduras. *Agric. Ecosyst. Environ.* **2005**, *110*, 14–28. [CrossRef]
88. Shen, Q.; Chen, Q.; Tang, B.-S.; Yeung, S.; Hu, Y.; Cheung, G. A system dynamics model for the sustainable land use planning and development. *Habitat Int.* **2009**, *33*, 15–25. [CrossRef]
89. Lu, J.; Guldmann, J.-M. Landscape ecology, land-use structure, and population density: Case study of the Columbus Metropolitan Area. *Landsc. Urban Plan.* **2012**, *105*, 74–85. [CrossRef]
90. Song, X.; Feng, Q.; Xia, F.; Li, X.; Scheffran, J. Impacts of changing urban land-use structure on sustainable city growth in China: A population-density dynamics perspective. *Habitat Int.* **2021**, *107*, 102296. [CrossRef]
91. Zhao, X.; Zhou, W.; Han, L. The spatial and seasonal complexity of $PM_{2.5}$ pollution in cities from a social-ecological perspective. *J. Clean. Prod.* **2021**, *309*, 127476. [CrossRef]
92. Mi, Y.; Sun, K.; Li, L.; Lei, Y.; Wu, S.; Tang, W.; Wang, Y.; Yang, J. Spatiotemporal pattern analysis of $PM_{2.5}$ and the driving factors in the middle Yellow River urban agglomerations. *J. Clean. Prod.* **2021**, *299*, 126904. [CrossRef]
93. Wang, X.; Xie, G.; Yue, S. Impact of Growth and Population Aggregation on Urban Environmental Quality and Its Regional Differentiation: A Case Study of 74 Cities Implemented the New Standard for Air Quality during the First Stage. *Econ. Geogr.* **2015**, *35*, 71–76+91. [CrossRef]
94. Liang, D.; Lu, H.; Guan, Y.; Feng, L.; He, L.; Qiu, L.; Lu, J. Population density regulation may mitigate the imbalance between anthropogenic carbon emissions and vegetation carbon sequestration. *Sustain. Cities Soc.* **2023**, *92*, 104502. [CrossRef]

95. Guo, D.; Deng, A.; Shen, Z. Research on Fault-Correction-Mechanism and Spatial Spillover Effect of Digital Economy to Promote Labor Mobility: Panel Data Analysis Based on 11Prefectural-level Cities in Zhejiang Province. *Zhejiang Acad. J.* **2023**, *6*, 153–161. [CrossRef]
96. Chernozhukov, V.; Chetverikov, D.; Demirer, M.; Duflo, E.; Hansen, C.; Newey, W.; Robins, J. Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **2018**, *21*, C1–C68. [CrossRef]
97. Farbmacher, H.; Huber, M.; Lafférs, L.; Langen, H.; Spindler, M. Causal mediation analysis with double machine learning. *Econom. J.* **2022**, *25*, 277–300. [CrossRef]
98. Yang, J.-C.; Chuang, H.-C.; Kuan, C.-M. Double machine learning with gradient boosting and its application to the Big N audit quality effect. *J. Econom.* **2020**, *216*, 268–283. [CrossRef]
99. Zhang, T.; Li, J. Network Infrastructure, Inclusive Green Growth, and Regional Inequality: From Causal Inference Based on Double Machine Learning. *J. Quant. Technol. Econ.* **2023**, *40*, 113–135. [CrossRef]
100. Wang, R.; Peng, F.; Li, W.; Wang, C. Does Terminating Rigid Payment Diminish Financing Cost of Companies? *J. Manag. World* **2022**, *38*, 42–64. [CrossRef]
101. Zhao, T.; Zhang, Z.; Liang, S. Digital Economy, Entrepreneurship, and High-Quality Economic Development: Empirical Evidence from Urban China. *J. Manag. World* **2020**, *36*, 65–76. [CrossRef]
102. Huang, Q.; Yu, Y.; Zhang, S. Internet Development and Productivity Growth in Manufacturing Industry: Internal Mechanism and China Experiences. *China Ind. Econ.* **2019**, *8*, 5–23. [CrossRef]